Symmetry, Integrability and Geometry: Methods and Applications SIGMA **20** (2024), 055, 87 pages

Knots, Perturbative Series and Quantum Modularity

Stavros GAROUFALIDIS [a] *and Don ZAGIER* [bc]

[a)] *International Center for Mathematics, Department of Mathematics, Southern University of Science and Technology, Shenzhen, P.R. China*
E-mail: *stavros@mpim-bonn.mpg.de*
URL: http://people.mpim-bonn.mpg.de/stavros

[b)] *Max Planck Institute for Mathematics, Bonn, Germany*
E-mail: *dbz@mpim-bonn.mpg.de*
URL: http://people.mpim-bonn.mpg.de/zagier

[c)] *International Centre for Theoretical Physics, Trieste, Italy*

Received April 25, 2023, in final form May 26, 2024; Published online June 24, 2024
https://doi.org/10.3842/SIGMA.2024.055

Abstract. We introduce an invariant of a hyperbolic knot which is a map $\alpha \mapsto \mathbf{\Phi}_\alpha(h)$ from $\mathbb{Q}/\mathbb{Z}$ to matrices with entries in $\overline{\mathbb{Q}}[[h]]$ and with rows and columns indexed by the boundary parabolic $\mathrm{SL}_2(\mathbb{C})$ representations of the fundamental group of the knot. These matrix invariants have a rich structure: (a) their (σ_0, σ_1) entry, where σ_0 is the trivial and σ_1 the geometric representation, is the power series expansion of the Kashaev invariant of the knot around the root of unity $\mathrm{e}^{2\pi \mathrm{i}\alpha}$ as an element of the Habiro ring, and the remaining entries belong to generalized Habiro rings of number fields; (b) the first column is given by the perturbative power series of Dimofte–Garoufalidis; (c) the columns of $\mathbf{\Phi}$ are fundamental solutions of a linear q-difference equation; (d) the matrix defines an $\mathrm{SL}_2(\mathbb{Z})$-cocycle W_γ in matrix-valued functions on $\mathbb{Q}$ that conjecturally extends to a smooth function on $\mathbb{R}$ and even to holomorphic functions on suitable complex cut planes, lifting the factorially divergent series $\mathbf{\Phi}(h)$ to actual functions. The two invariants $\mathbf{\Phi}$ and W_γ are related by a refined quantum modularity conjecture which we illustrate in detail for the three simplest hyperbolic knots, the 4_1, 5_2 and $(-2,3,7)$ pretzel knots. This paper has two sequels, one giving a different realization of our invariant as a matrix of convergent q-series with integer coefficients and the other studying its Habiro-like arithmetic properties in more depth.

Key words: quantum topology; knots; 3-manifolds; Jones polynomial; Kashaev invariant; volume conjecture; Chern–Simons theory; asymptotics; quantum modularity conjecture; quantum modular forms; hyperbolic 3-manifolds; dilogarithm; cocycles; $\mathrm{SL}_2(\mathbb{Z})$; denominators; Habiro-like functions; functions near $\mathbb{Q}$; Neumann–Zagier matrices; Nahm sums; q-holonomic modules

2020 Mathematics Subject Classification: 57N10; 57K16; 57K14; 57K10

Part 0: Introduction and overview

In this paper and the companion paper [44], we will define and study three different types of objects that can be associated to a hyperbolic knot:

- periodic functions on $\mathbb{Q}$ with values in $\overline{\mathbb{Q}}$ with striking arithmetic properties and belonging to a generalization of the Habiro ring;
- divergent formal series in an infinitesimal variable h, or more precisely infinite collections of such power series, indexed by a rational number α (here "h" is meant to remind one of Planck's constant and the perturbative expansions of quantum field theory); and

- q-series with integer coefficients, convergent in the unit disk and also thought of via $q = \mathrm{e}^{2\pi \mathrm{i}\tau}$ as holomorphic functions of a variable τ in the upper half-plane.

The first of these generalizes the Kashaev invariant of the knot, while the second and third correspond roughly to the two partition functions $Z(h)$ and $\widehat{Z}(q)$ that are being studied in the ongoing program of Gukov et al. [18, 50] for general 3-manifolds. We will study the first two types of invariants in the present paper, and the functions of q or τ in [44]. In all three cases, we will actually define a whole *matrix* of functions of the type described above, and in all three cases one of the central questions will be the behavior of these functions under the action of the modular group on the rational numbers or on the upper half-plane. Another key aspect is that each of the three types of matrices constructed encodes the same information as the other two and that all three can be interpreted as different realizations of the same abstract object, a square matrix of "functions-near-$\mathbb{Q}$" that we believe is associated to every hyperbolic knot, just as the different types of cohomology groups associated to an algebraic variety over a number field, despite their very different properties, are seen as different realizations of the same underlying "motive".

The starting point for our entire investigation is the Kashaev invariant of a knot and the "quantum modularity" property for its Galois-equivariant extension that was conjectured in [84]. We will review these topics in detail in Section 1, but remind the reader briefly of the basic ingredients here. The Kashaev invariant of a hyperbolic knot K is an element $\langle K\rangle_N$ of $\mathbb{Z}\big[\mathrm{e}^{2\pi\mathrm{i}/N}\big]$ for every $N \in \mathbb{N}$ whose absolute value is conjectured to grow exponentially like e^{cN}, where c is $1/2\pi$ times the hyperbolic volume of the knot complement $\mathbb{S}^3 \smallsetminus K$. This invariant can be extended to a function $J = J^{(K)}$ (we will omit the knot from the notation when it is fixed) from $\mathbb{Q}/\mathbb{Z}$ to $\overline{\mathbb{Q}}$ by Galois equivariance. (This means that we write $\langle K\rangle_N$ as a polynomial in $\mathrm{e}^{2\pi\mathrm{i}/N}$ with rational coefficients and define $J(a/N)$ for all a prime to N as the same polynomial evaluated at $\mathrm{e}^{-2\pi\mathrm{i}a/N}$.) The quantum modularity conjecture gives a formula for the ratio of the values of $J(X)$ and $J(\gamma X)$ as an asymptotic series in $1/X$ as X tends to infinity through integers, or even through rational numbers with bounded denominator, where $\gamma X = \frac{aX+b}{cX+d}$ with $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. The quantitative version of this conjecture is given in equation (1.6) below, via a collection of well-defined formal power series $\{\Phi_\alpha(h)\}_{\alpha\in\mathbb{Q}}$ with algebraic coefficients, but the conjecture can also be visualized in a weaker qualitative form by comparing the graphs of $J(x)$ and of $J(x)/J(\gamma x)$ as functions, as is done in the following figure (taken from [84]), which shows the plots of $\log(J(x))$ and $\log(J(x)/J(-1/x))$ for $K = 4_1$ ("figure 8 knot"), the simplest hyperbolic knot. The former consists of a whole "cloud" of points and has no reasonable extension to the real numbers, whereas the latter does extend to a well-defined function on $\mathbb{R}$, albeit one with infinitely many discontinuities.

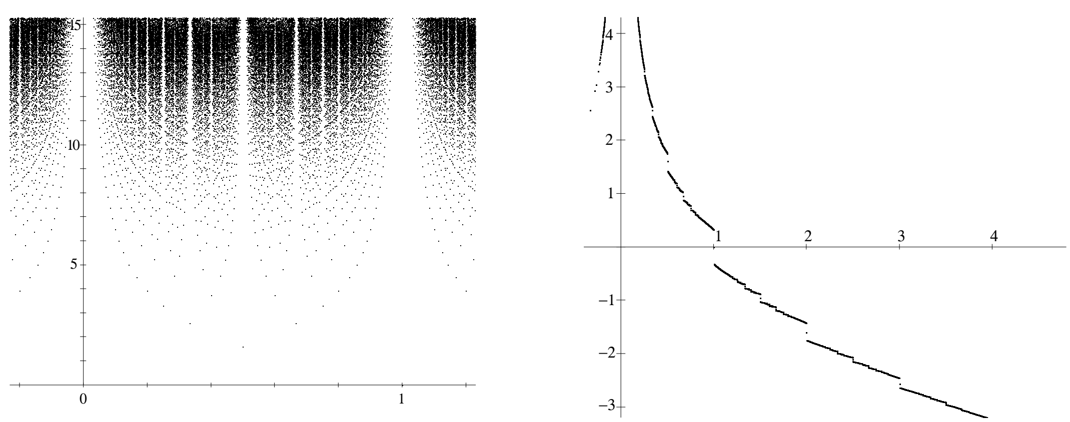

Figure 1. The functions $\log(J(x))$ and $\log(J(x)/J(-1/x))$ for the 4_1 knot.

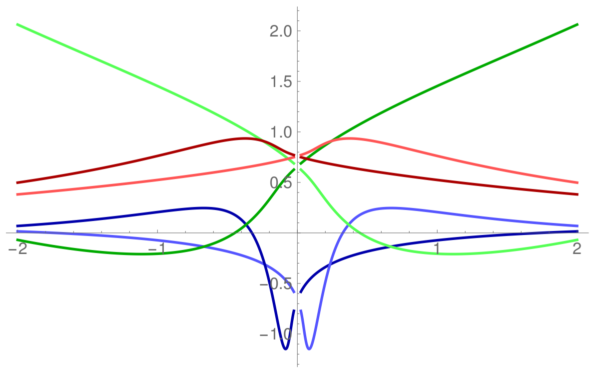

Figure 2. Plots of the six nontrivial entries (rescaled) of the matrix $W_S^{(4_1)}(x)$.

The marked improvement of the graph on the right of Figure 1 as opposed to the one on the left was already very striking and led to the introduction in [84] of a notion of "quantum modular forms" that has proved quite useful and has been exploited and extended by several subsequent authors. But it was also somehow unsatisfactory, because the function $\log J(x) - \log J(-1/x)$ still is far from smooth or even continuous, whereas in all the other examples in [84] the difference $f(x) - f(\gamma x)$ for a quantum modular form $f\colon \mathbb{Q} \to \mathbb{C}$ extended to an analytic function on $\mathbb{R}$ minus a finite set. This problem was "solved" in [84] by defining quantum modular forms by the weak requirement that the difference $f(x) - f(\gamma x)$ was "analytically better behaved" than f itself, rather than demanding that it be analytic on the complement of a finite set. But now it turns out that this cop-out is not needed, since the riddle of the missing smoothness is solved completely by upgrading J to a matrix of which it is only one entry. Specifically, in the new picture, $J(X)$ is replaced by a certain matrix-valued invariant $\mathbf{J}(X) = \mathbf{J}^{(K)}(X)$ (we use boldface letters to indicate matrices) which is defined and studied in the course of this paper (Sections 2.2, 3.1, and 4.1–4.5). For the figure 8 knot this matrix has the form

$$\begin{pmatrix} 1 & J(X) & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix},$$

where each of the six nontrivial components has a "cloudlike" graph like the first plot in Figure 1. But now instead of dividing the *scalar* invariant $J(X)$ by $J(-1/X)$ as before, we look at the *matrix* product $\mathbf{J}(-1/X)^{-1}\widetilde{\mathbf{j}}_S(X)\mathbf{J}(X)$, where $\widetilde{\mathbf{j}}_\gamma(X)$ is the matrix-valued automorphy factor defined in (4.14). Then the graphs of the six nontrivial entries of this product matrix, multiplied by suitable elementary factors to make them real and finite at the origin, look as in Figure 2 and are now smooth functions on the real line![1] The same graph also beautifully illustrates the interrelationship of the different matrix invariants of knots which we spoke of above, because the six curves that are plotted are at the same time canonical lifts to $C^\infty(\mathbb{R})$ of the six components of the matrix of formal power series $\mathbf{\Phi}_\alpha(h)$ that is associated to the knot and to every rational number α (here for $\alpha = 0$). In this way the matrix $\mathbf{J}$ of Habiro-like functions determines the formal power series $\mathbf{\Phi}_\alpha(h)$ (to get values of α other than 0, one would replace $-1/X$ by $\gamma(X)$ for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ with $\gamma(\infty) = \alpha$), and conversely the matrices $\mathbf{\Phi}_\alpha(h)$ determines the matrix $\mathbf{J}$ simply by $\mathbf{J}(\alpha) = \mathbf{\Phi}_\alpha(0)$.

[1] To avoid misconceptions, we note that the other striking property of this picture, its symmetry with respect to the vertical axis, is due to the accidental fact that the 4_1 knot is amphicheiral. But already for the next simplest case of the 5_2 knot, which will be the second of our three standard examples throughout the paper, the matrix $\mathbf{J}$ would be have size 4×4 with 12 non-trivial entries, all complex, and a graph of their 24 real and imaginary parts would be visually unintelligible.

This preliminary discussion and these pictures already give a first impression of the content of this paper. However, giving the exact definitions of the various objects being studied is not at all a straightforward business, because some of these are based only on numerical data and cannot yet be justified by any theoretical considerations, so that they can only reasonably be introduced after the numerical investigations have been presented, while in other cases there are different candidate definitions whose equality is only conjectural. In the body of the paper we will therefore present the material in two stages. Part I introduces the main players: the perturbative power series, their conjectural analytic and number-theoretical properties, and the emergence of an $\mathrm{SL}_2(\mathbb{Z})$-cocycle through their quantum modularity properties. The final (conjectural) statements are given in Section 5, so that a reader who wants to see just the short version of the story right away can skip directly to that section. Part II then contains more detailed information about the definitions and properties of the objects appearing in Part I, including a discussion of the numerical methods used, some of which are quite subtle. The paper ends with an appendix containing tables of some of the functions studied for a few simple hyperbolic knots.

Since the paper contains so many different types of objects, with rather intricate interconnections and taking shape only gradually in the course of the exposition, it seemed useful to end this introduction by giving a detailed overview of the main ingredients. A further reason to include this rather long list here is that it contains a number of items (the unexpected appearance of algebraic units, a description of the Bloch group and extended Bloch group in terms of "half-symplectic matrices", the notions of "asymptotic functions near $\mathbb{Q}$" and of "holomorphic quantum modular forms", a generalization of the Habiro ring to Habiro-like rings associated to number fields other than $\mathbb{Q}$, or a procedure to "evaluate" divergent power series numerically) that are applicable or potentially applicable in domains quite separate from that of quantum knot invariants and that therefore may be of independent interest.

• **Indexing set.** Both the rows and the columns of the matrices associated to a knot K are indexed by a finite set $\mathcal{P}_K$ that can be described either in terms of boundary parabolic representations of the fundamental group of the knot complement $\mathbb{S}^3 \smallsetminus K$ or in terms of flat connections, as explained in detail in Section 2.

• **Lift of the complex volume.** The leading asymptotic exponent of the new matrices is a complex-valued function on the set $\mathcal{P}_K$ which agrees with the complexified volume of the boundary parabolic representation, except that the latter is only well-defined modulo $4\pi^2\mathbb{Z}$. Thus, a consequence of the refined quantum modularity conjecture is that the complexified volume of a hyperbolic knot now has a canonical lift from $\mathbb{C}^2/4\pi^2\mathbb{Z}$ to $\mathbb{C}$.

• **Level.** As already mentioned, the matrices that we study also depend on a rational number α. In all cases they are periodic in α, with the period $N = N_K$ however not always being the same: it is 1 for the 4_1 and 5_2 knots, but 2 for the $(-2,3,7)$-pretzel knot. (These three knots will serve as our standard illustrations throughout the paper.) Similarly, the modular invariance properties are not always under the full modular group $\mathrm{SL}_2(\mathbb{Z})$, but sometimes under the subgroup $\Gamma(N)$. We do not know what this "level" N is in general, although we have a guess (in terms of the quasi-periodicity of the degrees of the colored Jones polynomials), but its appearance in the numerical investigations was striking.

• **Perturbatively and non-perturbatively defined power series.** In the original version of the quantum modularity conjecture, the main statement was the *existence* of a collection of formal power series $\Phi_\alpha(h)$ describing the relationship between $J(X)$ and $J(\gamma X)$ for large X and fixed $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, with no prediction of what this power series was. However, in two papers [14, 15] by Tudor Dimofte and the first author explicit candidates for these power series as perturbative series in h defined by Gaussian integration of a function explicitly given in terms of a triangulation of M^3, rendering the original conjecture much more precise. These series will now form all but the top entry of the second column of our matrix. (The first column is simply $(10\dots0)^t$.) The top entry of the second column is defined in a completely different, non-perturbative way in terms of the expansion near $q = \mathrm{e}^{2\pi\mathrm{i}\alpha}$ of the Kashaev invariant of K seen as an element of the Habiro ring. All of this will be explained in detail in Section 2, while the definitions of the other entries of the matrix $\mathbf{\Phi}_\alpha(h)$, which are again given by perturbatively defined power series in all but the top row and (conjecturally) by elements of the Habiro ring for their top entries, will emerge via the Refined quantum modularity conjecture discussed in Sections 4.

• **Arithmetic aspects.** One of the main themes of this paper is that the topology of a knot involves a large amount of surprisingly complicated algebraic number theory. This is valid both for the values of the Kashaev invariant itself and of its generalizations as given by the matrix $\mathbf{J}(x)$ at rational arguments x and more generally for the coefficients of the entries $\mathbf{\Phi}_\alpha^{(\sigma,\sigma')}(h)$ ($\alpha \in \mathbb{Q}$, $\sigma, \sigma' \in \mathcal{P}_K$), which conjecturally belong to $\overline{\mathbb{Q}}(h)$. Among the most striking things that we found were the occurrence of certain algebraic *units*, which led to the paper [10] with Frank Calegari associating units (modulo nth powers) in the nth cyclotomic extension of an arbitrary number field to elements in the Bloch group of this field, congruence properties of Ohtsuki type (which will be touched on only briefly here but will become a main theme in the planned paper with Peter Scholze and Campbell Wheeler on the construction of Habiro rings associated to any number field), and universal bounds, independent of the knot, for the denominators of the coefficients of the perturbative power series occurring (see Section 9.2).

• **Half-symplectic matrices and the extended Bloch group.** The power series constructed in [14, 15] are given in terms of the so-called Neumann–Zagier data describing the combinatorics of a triangulation of a knot complement. This data takes the form of an $N \times 2N$ integral matrix, where N is the number of simplices of the triangulation, together with a solution (corresponding to the "shape parameters", i.e., the cross-ratios of the vertices of the ideal tetrahedra) of a collection of algebraic equations defined by this matrix. The key property here is that the defining matrix is the upper half of a $2N \times 2N$ symplectic matrix over $\mathbb{Z}$. This leads to a somewhat different description than the standard one of the Bloch group and extended Bloch group. All of this is discussed in Section 6, together with the definition of the perturbative series and a different appearance of the same construction in the context of Nahm sums.

• **Unimodularity and inverse matrix.** Experimentally, we find that the matrices that we construct are always unimodular, and also that there are explicit formulas for their inverses as linear (or, for the top row, quadratic) rather than higher-degree polynomials in the entries of the matrices themselves. Combined with the behavior under complex conjugation, this leads to a kind of generalized unitarity property for our matrices (of course, again only conjecturally, but we will not keep repeating this since most of the properties we are discussing are only conjectural, though based on such extensive data that they are very unlikely not to hold). All of this will be discussed in Section 5. The formula for the inverses of our matrices (apart from the top row) can be interpreted as giving quadratic relations for our power series, a special case of which appears in a recent paper of Gang, Kim and Yoon [22] and which will be described in Section 3.3. These quadratic relations will take on a life of their own in the companion paper [44] in terms of expressions for the "state integrals" defined by Kashaev and others as bilinear expressions in power series in $q = \mathrm{e}^{2\pi\mathrm{i}\tau}$ and $\widetilde{q} = \mathrm{e}^{-2\pi\mathrm{i}/\tau}$.

• **Extension property.** As already stated, the rows and columns of our matrices are indexed by the set $\mathcal{P}_K$ of flat connections. This set has a canonical element, the trivial connection, which we put at the beginning of the list, and the first row and column of each of our matrices then has a completely different nature from the other entries. In particular, as we already saw, the first column always consists of a 1 followed by 0s, so that the entire matrix is in $(1+r)\times(1+r)$ block triangular form, where $1+r = |\mathcal{P}_K|$. This means that these matrices are describing structures which are r-dimensional extensions of 1-dimensional substructures. This can be seen clearly in the q-holonomy discussed below, where the recursions satisfied by elements giving the top row contain a constant term 1 and those of the other rows are the corresponding homogeneous recursions.

• ***q*-holonomy.** As already explained, the second column (the first column being trivial) of our matrices of formal power series has a direct definition in terms of the triangulation of the knot complement and the corresponding Neumann–Zagier data. The other columns are defined in a more complicated way that we still have not understood completely. Roughly speaking, each of the entries of the original column belongs to a "q-holonomic system", meaning that it is part of a sequence of functions of $q = \mathrm{e}^{2\pi\mathrm{i}\alpha}$ that satisfy linear recursions over $\mathbb{Q}[q]$ and hence span a finite-dimensional space, and the other columns belong to, and in fact span, the same space. This applies not only to the matrices of formal power series in h, but also to the Habiro-like matrices $\mathbf{J}$ and to the matrices of q-series studied in [44], and will be discussed in detail in Section 7. The mysterious point here is that the columns of our matrices, which are completely and uniquely defined by the various properties embodied in the refined modularity conjecture, give a canonical basis for these q-holonomic modules, but that even in those situations where we know what the module is or should be, we do not have an a priori description of this basis.

• **Refined quantum modularity.** As stated at the beginning of this introduction, our whole story arises from the quantum modularity conjecture (QMC) made in [84]. In its original form, the QMC says that for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z})$ we have

$$J(\gamma X) \approx (cX+d)^{3/2} J(X) \widehat{\Phi}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right), \qquad X \to \infty$$

as X tends to infinity with a bounded denominator, where $\widehat{\Phi}_\alpha(h)$ for $\alpha \in \mathbb{Q}$ is a "completed" version of $\Phi_\alpha(h)$ obtained by multiplying it by a suitable pure exponential in $1/h$, and where "$\approx$" denotes asymptotic equality to all orders in $1/X$ (or h). This statement already refines the volume conjecture of Kashaev [58] and its arithmetic properties and extension to all orders as described in [16]. In the refined quantum modularity conjecture (RQMC), which will be developed step by step in the course of Sections 3 and 4, it is extended in two different ways. First of all, the above asymptotic statement will be generalized by replacing the function $J\colon \mathbb{Q} \to \mathbb{Q}$ by the other entries, initially of the first column and then by a kind of "bootstrapping" process (see below) of the whole matrix. More importantly, however, the right-hand side will be sharpened by the addition of lower-order terms which are completed versions of other entries of the matrix $\widehat{\mathbf{\Phi}}_\alpha(X)$. Since the addition of an exponentially smaller expression to a divergent power series does not make sense a priori, this requires a process of numerical evaluation by "optimal truncation" and then "smoothed optimal truncation" as listed in the bullet "Numerical aspects" below and discussed in detail in Sections 4.1 and 10.2. The final result gives an asymptotic development to much higher precision of each of the generalized Habiro-functions $\mathbf{J}^{(\sigma,\sigma')}$ evaluated at γX with X tending to infinity as a linear combination of $(r+1)$ of the power series $\widehat{\mathbf{\Phi}}_\alpha(h)$, with $\alpha = a/c$ and $h = 2\pi\mathrm{i}/c(cX+d)$. This can then be written compactly in matrix form as

$$\mathbf{J}^{(K)}(\gamma X) \approx \widetilde{\mathbf{j}}_\gamma(X) \mathbf{J}^{(K)}(X) \widehat{\mathbf{\Phi}}^{(K)}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right)$$

(= equation (4.12)), which is the final version of the RMC. Here the "automorphy factor" $\widetilde{\mathbf{j}}_\gamma(X)$ is a diagonal matrix whose first entry is $(cX+d)^{3/2}$ and whose other entries are pure exponentials in $X + d/c$. Note that this property relates the matrices $\mathbf{J}$ and $\widehat{\Phi}$, and allows in particular to compute the second one from the first one.

• **A matrix-valued cocycle.** The matrix-valued form of RQMC as just stated leads immediately to the definition of an $\mathrm{SL}_2(Z)$-cocycle with coefficients in the space of matrix-valued functions on $\mathbb{Q}$ (or more precisely—and necessary in order to have an $\mathrm{SL}_2(Z)$-module structure—of almost-everywhere-defined functions on $\mathbb{P}^1(\mathbb{Q})$), defined by

$$W_\gamma(x) = \mathbf{J}(\gamma x)^{-1} \widetilde{\mathbf{j}}_\gamma(x) \mathbf{J}(x),$$

where this time the "automorphy factor" $\widetilde{\mathbf{j}}_g(x)$ is a slightly different diagonal matrix, again with elementary entries, as defined explicitly in equations (4.14) and (3.5). The cocycle properties of this automorphy factor imply that W_γ is a multiplicative cocycle, meaning that $W_{\gamma\gamma'}(X)$ is equal to $W_\gamma(\gamma' X)$ times $W_{\gamma'}(X)$. Its remarkable properties are summarized in the next two bullets.

• **Analyticity and holomorphic quantum modular forms.** The most important single discovery of this paper is that the cocycle $W_\gamma(X)$, originally defined on rational numbers by the formula just given, extends to a smooth function on the real numbers. This fact, which might have been found purely experimentally by looking at the graphs of the components of $W_\gamma(X)$, as illustrated by Figure 2 above, and which can also be checked purely experimentally, as explained in Section 5.4, was actually predicted in advance on the basis of the occurrence of the same cocycle $\gamma \mapsto W_\gamma$ with a completely different construction in the companion paper [44] to this one. Specifically, in that paper we construct a matrix $\mathbf{Q}^{\mathrm{hol}}(\tau)$ of holomorphic functions of a complex variable $\tau \in \mathbb{C}\smallsetminus\mathbb{R}$, whose entries are power series with integer coefficients in $q = \mathrm{e}^{2\pi\mathrm{i}\tau}$, and such that the coboundary $\mathbf{Q}^{\mathrm{hol}}(\gamma\tau)^{-1}\mathbf{Q}^{\mathrm{hol}}(\tau)$ extends holomorphically across both the half-lines $\big(\gamma^{-1}(\infty), \infty\big)$ and $\big(-\infty, \gamma^{-1}(\infty)\big)$, with the restrictions to these two half-lines coinciding with the function W_γ there. This extendability of $\mathbf{Q}^{\mathrm{hol}}(\gamma\tau)^{-1}\mathbf{Q}^{\mathrm{hol}}(\tau)$ across subintervals of the real line means that $\mathbf{Q}^{\mathrm{hol}}$ is an example of a "holomorphic quantum modular form", a new type of object that turns out to occur in many other contexts and that will be described briefly in Section 5.4 and in detail in the papers [44, 85].

• **"Functions near $\mathbb{Q}$".** Each component $\Phi_\alpha^{\sigma,\sigma'}(h)$ of the matrix $\mathbf{\Phi}_\alpha$ has a natural completion, as explained in Section 2, defined as its product with a certain exponential in $1/h$ and (in the case of the top row) a half-integral power of h, and it is these completions that appear in the original quantum modularity conjecture and its various extensions. It turns out that the "right way" to think of these collections of series is that they represent one single "asymptotic function near $\mathbb{Q}$" defined by $\mathbf{Q}^{(\sigma,\sigma')}(\alpha - \hbar) = \widehat{\Phi}_\alpha^{\sigma,\sigma'}(2\pi\mathrm{i}\hbar)$, where α varies over $\mathbb{Q}$ and $\hbar$ is infinitesimal. This notion of asymptotic functions near $\mathbb{Q}$ (or simply "functions near $\mathbb{Q}$" for short), which will be defined and explained more carefully in Section 5.3, sheds light on several properties of our knot invariants (and also turns out to occur also in other contexts). In particular, the cocycle W_γ, which was initially defined (almost everywhere) on $\mathbb{Q}$ by the formula $W_\gamma(x) = \mathbf{J}(\gamma x)^{-1}\widetilde{\mathbf{j}}_\gamma(X)\mathbf{J}(x)$, is not a coboundary in the space of functions on $\mathbb{Q}$, but *is* one in the larger space of functions near $\mathbb{Q}$: $W_\gamma(x) = \mathbf{Q}(\gamma x)^{-1}\mathbf{Q}(x)$. The occurrence of the *same* cocycle with two different representations as a coboundary in appropriate matrix-valued $\mathrm{SL}_2(\mathbb{Z})$-modules provides the link between the two papers and the reason for our belief that both the matrix $\mathbf{Q}$ of generalized Habiro functions and the matrix $\mathbf{Q}^{\mathrm{hol}}$ of q-series are realizations of the same underlying motive-like object.

• **Numerical aspects.** Everything in the paper is based on numerical computations, and these have several non-obvious aspects, as discussed in Section 10. In particular, we explain there how Kashaev invariants can be computed rapidly and how one can then use extrapolation techniques to evaluate many coefficients of the power series $\Phi_a^{(\sigma,\sigma')}(h)$ numerically and recognize them as real numbers. The calculations also have a "bootstrapping" aspect in which the successively discovered relations among the series as described by the final refined quantum modularity conjecture permit one to evaluate these series to increasing levels of precision in a recursive way. Finally, in order to identify the correct series in the RQMC, it is crucial to be able to evaluate the divergent series in h occurring, not only up to order h^N for any fixed integer N, but up to exponentially small error terms, where the constant occurring in the exponential can also be successively improved in several steps. This is done by a process of "smoothed optimal truncation" which was originally a second appendix to this paper, but has now been relegated to a separate publication [45] and is also briefly described in Section 10.2.

We end with a few miscellaneous remarks on different aspects of the above constructions.

• **Resurgence aspects.** An important aspect of our paper are the matrices $\mathbf{\Phi}_\alpha(h)$ of factorially divergent series and their distinguished lift $W_\gamma(x)$ to a matrix of analytic functions. In this connection we find several properties that can be classified under the general heading of "resurgence." On the one hand, we find experimentally that the coefficients of each entry of $\mathbf{\Phi}_\alpha(h)$ are given asymptotically as *integer* linear combinations of certain divergent expansions involving the coefficients themselves multiplied by gamma factors. The integrality of the so-called Stokes constants is a phenomenon observed in the current paper and further studied in [27, 28]. A different connection of our results with the usual resurgence properties of the perturbative series $\mathbf{\Phi}_\alpha(h)$ is the method of smoothed optimal truncation mentioned just above, which can be seen as an alternative approach to lifting these power series to actual functions than the standard method via Borel resummation and Padé approximation. Of course, the final emergence of a canonical lift coming from the analyticity properties of the cocycle $W_\gamma(x)$ eventually makes both numerical procedures obsolete in our case, but this cocycle could not be found without having them first.

• **Equivalence of the various invariants.** We observe that all of our invariants, assuming their conjectured properties, determine each other and in particular all are determined by the colored Jones polynomials and perhaps even by the Kashaev invariant alone. The point of the paper is therefore not that our new invariants can distinguish knots, but that they reveal new properties and make connection with $\mathrm{SL}_2(\mathbb{C})$-representations of the fundamental group, hyperbolic geometry and non-trivial number theory.

• **Eighth roots of unity and the Dedekind eta function.** A further minor surprise was the appearance of Dedekind sums at several places in the calculations, which we had not expected. Notably, this occurred in the construction of state sums for the non-Habiro-like elements of our matrix $\mathbf{J}$ (see Section 7.2) and in the ubiquitous but mysterious 8th root of unity that enters all of our asymptotic formulas and that is related to the 8th root of unity occurring in the modular transformation behavior of η^3.

• **3 + 1: a possible alternative interpretation.** Finally, we mention a point that will not be discussed in the paper at all, but may give a different way of looking at the objects studied here. Namely, the (σ, σ') entry of the matrix $\mathbf{J}(X)$ that we have associated to a knot K, and hence also to its complement $M = \mathbb{S}^3 \smallsetminus K$, can be thought of as numbers associated to the "Witten cylinder", which is the 4-manifold $M \times \mathbb{R}$ equipped with a pair of boundary-parabolic $\mathrm{SL}_2(\mathbb{C})$ connections σ and σ' on its two ends, together with an integer k, called the "level" in complex Chern–Simons theory [80, 81] and related to the rational number X by $k+2 = \mathrm{den}(X)$. This suggests a possible interpretation of the entries of $\mathbf{J}(X)$ as expectation values of some yet-to-be-defined $(3+1)$-dimensional theory on $M \times \mathbb{R}$.

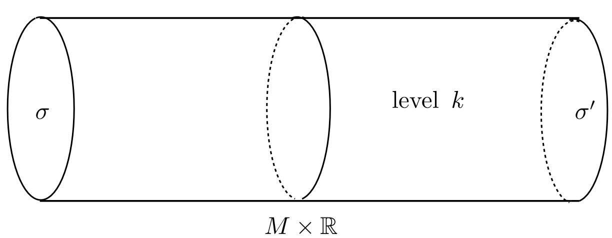

Figure 3. The Witten cylinder.

• **Relation with the Kauffman bracket skein module.** The objects that we have studied have another connection with a 3+1 dimensional theory that goes though the Kauffman bracket skein module and a conjecture of Witten (see, for instance, [79]), now a theorem of Gunningham–Jordan–Safronov [51]. Recently, Lê and the first author defined an explicit map from the Kauffman bracket skein module of an integer homology 3-sphere to the Habiro ring [36]. The image of this map (with can be effectively computed, and is naturally a topological invariant) generates a finite-rank $\mathbb{Z}[q^{\pm}]$-module which is a subset of the Habiro ring, and a basis of this module conjecturally coincides with the top row of the **J**-matrix.

We end this introduction with a disclaimer. In this paper, we are trying only to present interesting new phenomena and are *not* striving for maximum generality. Not only are we restricting our attention to knot complements rather than arbitrary 3-manifolds, but we will usually assume that the knots being considered have whatever properties (e.g., that their parabolic character varieties are 0-dimensional) are convenient for our exposition. In any case most of the material presented is empirical, based on extensive experiments with a few very simple knots having all of these special properties, and we prefer not to speculate on what modifications might be needed in more general situations. Of course, we expect that the whole story is quite general, and ongoing calculations by Campbell Wheeler that appear in his thesis [76] indeed already confirm that very similar types of statements will hold, for instance, for the Witten–Reshitikhin–Turaev invariant of certain closed hyperbolic 3-manifolds.

Part I: The main story

1 The original quantum modularity conjecture

The starting point for this paper is the quantum modularity conjecture of [84], which itself is a refinement of the famous volume conjecture of Kashaev [58]. Let us recall them briefly here. Kashaev defined an invariant $\langle K\rangle_N \in \mathbb{Z}[\mathrm{e}^{2\pi\mathrm{i}/N}]$ for every knot K and positive integer N and conjectured that if K is hyperbolic knot the absolute values of these numbers grow exponentially like $\mathrm{e}^{c_K N}$, where $2\pi c_K$ is the hyperbolic volume of $S^3 \smallsetminus K$. A more precise version of the conjecture [48] says that there is a full asymptotic expansion

$$\langle K\rangle_N \sim N^{3/2}\mathrm{e}^{v(K)N}\Phi^{(K)}\left(\frac{2\pi\mathrm{i}}{N}\right) \tag{1.1}$$

valid to all orders in $1/N$ where $v(K)$ is the suitably normalized complexified volume of K and where $\Phi^{(K)}(h)$ is a (divergent) power series. (Here and from now on we use the abbreviation $\zeta_n = \mathbf{e}(1/n)$ for $n \in \mathbb{N}$, where $\mathbf{e}(x) := \mathrm{e}^{2\pi\mathrm{i}x}$.) It was further conjectured in [16] and in [23] that $\Phi^{(K)}(h)$ has algebraic coefficients, and more precisely that it belongs to $\zeta_8\delta^{-1/2}F_K[[h]]$, where F_K is the trace field of K and δ some nonzero element of F_K.

For instance, for the simplest hyperbolic knot 4_1 (figure eight knot), the Kashaev invariant is given explicitly by

$$\langle 4_1\rangle_N = \sum_{n=0}^{N-1}\bigl|(\zeta_N;\zeta_N)_n\bigr|^2 \tag{1.2}$$

(see [59, equation (2.2)]) with $(q;q)_n := \prod_{j=1}^{n}\bigl(1-q^j\bigr)$, with values for $N = 1,\dots,6$ and 100 given numerically by

N	1	2	3	4	5	6	$\cdots$	100
$\langle 4_1\rangle_N$	1	5	13	27	$46+2\sqrt{5}\approx 50.47$	89	$\cdots$	8.2×10^{16}

Here the trace field F_K is $\mathbb{Q}\bigl(\sqrt{-3}\bigr)$ and the series $\Phi^{(K)}(h)$ begins

$$\Phi^{(4_1)}(h) = \frac{1}{\sqrt[4]{3}}\left(1+\frac{11}{72\sqrt{-3}}h+\frac{697}{2\bigl(72\sqrt{-3}\bigr)^2}h^2+\frac{724351}{30\bigl(72\sqrt{-3}\bigr)^3}h^3+\cdots\right). \tag{1.3}$$

Similarly, for the knot 5_2, where the Kashaev invariant is given by formula (A.2) below, the trace field is $\mathbb{Q}(\xi)$, where ξ is the root of $\xi^3-\xi^2+1=0$ with negative imaginary part, and $\Phi^{(K)}(\hbar)$ has an expansion starting

$$\Phi^{(5_2)}(h)=\frac{\zeta_8}{\sqrt{3\xi-2}}\biggl(1+\frac{117\xi^2-222\xi+203}{24(3\xi-2)^3}h \\ +\frac{117279\xi^2-209229\xi+157228}{2\bigl(24(3\xi-2)^3\bigr)^2}h^2+\cdots\biggr). \tag{1.4}$$

The passage from the volume conjecture to the quantum modularity conjecture (QMC for short) begins with the observation that the Kashaev invariant $\langle K\rangle_N$ is the value at $x=-1/N$ of a 1-periodic function $J(x)=J^{(K)}(x)$ on the rational numbers (i.e., it satisfies $J(x+1)=J(x)$ for all x), defined uniquely by the further requirement that it is a Galois-invariant function of $\mathbf{e}(x)$. (The uniqueness holds because every primitive N-th root of unity is a Galois conjugate of ζ_N.) As shown by Murakami and Murakami [64], this function can also be identified with an evaluation of the colored Jones polynomial $J_{K,N}(q)$ [57, 73] by $J(x)=J_{K,N}(\mathbf{e}(-x))$ for any $N\in\mathbb{Z}$ with $Nx\in\mathbb{Z}$. In [84], it was found that (1.1) is just the special case $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)=\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ of the more general (and of course still conjectural) statement that

$$J^{(K)}\left(\frac{aN+b}{cN+d}\right)\sim(cN+d)^{3/2}\mathrm{e}^{v(K)(N+d/c)}\Phi^{(K)}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cN+d)}\right) \tag{1.5}$$

to all orders in N as $N\to\infty$ for any matrix $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\mathrm{SL}_2(\mathbb{Z})$ with $c>0$, where $\Phi^{(K)}_\alpha(\hbar)$ is a power series with algebraic coefficients depending on $\alpha\in\mathbb{Q}/\mathbb{Z}$, with $\Phi^{(K)}_0=\Phi^{(K)}$. This does not yet look like a modularity statement, but in [84] it was further observed that (1.5) holds also for non-integral values of N (which we then denote by X for clarity) but with one crucial modification, namely that we have

$$J^{(K)}\left(\frac{aX+b}{cX+d}\right)\sim(cX+d)^{3/2}\mathrm{e}^{v(K)(X+d/c)}\Phi^{(K)}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right)J^{(K)}(X) \tag{1.6}$$

to all orders in $1/X$ as $X\to\infty$ in $\mathbb{Q}$ with bounded denominator, with the same series $\Phi^{(K)}_\alpha(\hbar)$ as before but now with the additional factor $J^{(K)}(X)$ depending only on X modulo 1. (Here the condition of X having bounded denominator was included in the original conjecture, and will be retained for its refinements in this paper, because all of our experiments were done for X with simple fractional part. However, it will be a consequence of the smoothness statements cited above and discussed in Section 5.2 that in fact (1.6) remains true for any sequence of rational numbers X tending to infinity, and we have checked this experimentally for many cases.)

Formula (1.6) expresses the QMC in a quantitative form, in terms of specific power series with algebraic coefficients, while the plots of $J^{(K)}(X)$ and $J^{(K)}(X)/J^{(K)}(\gamma X)$ for $K=4_1$ and $\gamma=S$ that were shown (on a logarithmic scale because the functions grow exponentially) in the introduction presents the same conjecture in a more qualitative visual form. Both ways of looking at the conjecture will be refined greatly during the course of this paper.

The QMC (1.6), or even its special case (1.5), give us not just one, but a whole collection of power series $\Phi^{(K)}_\alpha(\hbar)$ associated to any knot. These series $\Phi^{(K)}_\alpha(\hbar)$ have a striking arithmetical structure. For example, in [84] we found that for $K=4_1$ and α with denominator 5 that

$$\Phi^{(4_1)}_\alpha(\hbar)=\pm\sqrt[4]{3}\,\varepsilon_\alpha^{1/5}\bigl(A^{(\alpha)}_0+A^{(\alpha)}_1h+\cdots\bigr)\qquad\text{if}\quad\alpha\in\frac{1}{5}\mathbb{Z}\smallsetminus\mathbb{Z}, \tag{1.7}$$

where ε_α is a unit (whose fifth root must be chosen appropriately) of the cyclotomic field $\mathbb{Q}(\zeta_{15})=F_{4_1}(\zeta_5)$ ($\zeta_m:=\mathbf{e}(1/m)$) (explicitly $\varepsilon_\alpha=\frac{Z_\alpha^4-1}{Z_\alpha(Z_\alpha+1)^2}$, with $Z_\alpha:=\mathbf{e}\bigl(\alpha-\frac{1}{3}\bigr)$) and the coefficients $A^{(\alpha)}_n$

belong to the same field, e.g., $A_0^{(\alpha)} = 1 + Z_\alpha^2 + Z_\alpha^{-2} - Z_\alpha^4 - Z_\alpha^{-4}$, a prime of norm 29. More generally, in the appendix to this paper extensive evidence is given for the conjecture that the power series $\Phi_\alpha^{(K)}(\hbar)$ for any knot K and rational number α always belongs to $\mu\delta^{-1/2}\sqrt[c]{\varepsilon}F_{K,c}[[h]]$ with the same δ as before and some $8c$-th root of unity μ, where c is the denominator of α, $F_{K,c}$ the field obtained by adjoining the c-th root of unity $\mathbf{e}(\alpha)$ to F_K, and ε is an S-unit of $F_{K,c}$ for some finite set of primes S of F_K independent of c. This experimentally discovered property of quantum invariants of knots in turn suggested the purely number-theoretical conjecture, which was then proved in [10], that to an *arbitrary* number field F and element of the Bloch group of F one can canonically associate a sequence of S-units, well defined up to c-th powers, in the c cyclotomic extension $F(\zeta_c)$ for all $c \geq 1$, with S independent of c.

We will give more details about this and other arithmetic properties of the series $\Phi_\alpha^{(K)}$ (such as estimates of the denominators of their coefficients) in Section 9, and will give a complete proof in the case of the 4_1 knot in Section 8. (This case and a few others were proven independently by Bettin and Drappeau [6].) We will also give detailed numerical evidence for several other knots, for several values of $\alpha \in \mathbb{Q}/\mathbb{Z}$, and to a relatively high degree in the power series $\Phi_\alpha^{(K)}(\hbar)$, in the appendix to this paper. The calculations required to obtain these values are not at all trivial, since one has to be able to calculate the Kashaev invariants for (many) arguments with large denominators and then use very precise extrapolation methods to be able to find the coefficients to high enough accuracy to recognize them numerically as algebraic numbers.

Presenting the numerical evidence for the QMC was the initial motivation for this paper, and this already led to interesting numerical observations, such as the appearance of the near unit ε or the denominator estimates mentioned above. But in the course of doing these calculations we discovered that the QMC is only part of a much larger story involving a whole collection of power series $\{\Phi_\alpha^{(K,\sigma)}(\hbar)\}_{\alpha\in\mathbb{Q}/\mathbb{Z},\sigma\in\mathcal{P}_K}$ indexed by a certain finite set $\mathcal{P}_K$ (defined below) as well as by an index in $\mathbb{Q}/\mathbb{Z}$ as before. In the rest of Part I, we explain what these power series are, how they are related to each other, and how they lead to new invariants and to a whole series of successive refinements of the original quantum modularity conjecture.

2 A collection of formal power series

2.1 The indexing set $\mathcal{P}_K$

The power series mentioned above are labeled by a finite set $\mathcal{P}_K$ that coincides with the set of boundary parabolic $\mathrm{SL}_2(\mathbb{C})$-representations of $\Gamma := \pi_1(S^3 \smallsetminus K)$ (or equivalently, of flat connections on $S^3 \smallsetminus K$ whose restriction to the peripheral subgroup of Γ are parabolic) whenever the latter is finite. For a hyperbolic knot, this set has three distinguished elements, denoted σ_0, σ_1 and σ_2, corresponding respectively to the trivial representation, the geometric representation (given by the natural embedding of Γ into the isometry group of $\mathbb{H}^3$) and the antigeometric representation, which is its complex conjugate and corresponds to the geometric representation of the orientation-reversed hyperbolic knot. We denote by $\mathcal{P}_K^{\mathrm{red}} = \mathcal{P}_K \smallsetminus \{\sigma_0\}$ the reduced set of nontrivial representations (or connections), and often number the elements of $\mathcal{P}_K$ as $\sigma_0, \sigma_1, \dots, \sigma_r$, where $r := |\mathcal{P}_K^{\mathrm{red}}|$ will be called the *rank* of the knot. (We hope that the superscript “red” will not confuse the reader into thinking that the representations in $\mathcal{P}_K^{\mathrm{red}}$ are reducible; in fact, quite the opposite is true.) The points of $\mathcal{P}_K^{\mathrm{red}}$ correspond to the solutions (in $\mathbb{C}$) of a set of polynomial equations (the so-called Neumann–Zagier equations coming from a triangulation of $S^3 \smallsetminus K$) as explained in detail in Section 6. In particular, $\mathcal{P}_K^{\mathrm{red}}$ comes equipped with an action of the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, so to every element $\sigma \in \mathcal{P}_K^{\mathrm{red}}$ is associated a number field F_σ (given either as the field generated by the coordinates of the solution of the NZ equations or as the fixed field of the stabilizer of σ in $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$), called its trace field, together with an embedding, also denoted σ, of F_σ into $\overline{\mathbb{Q}} \subset \mathbb{C}$. The field F_{σ_1} coincides with the trace field F_K of K

as introduced above and F_{σ_2} is the same field with the complex conjugate embedding into $\mathbb{C}$. Two more important invariants of $\sigma \in \mathcal{P}_K^{\rm red}$ are an element ξ_σ of the Bloch group (or third K-group) of F_σ and a complexified volume $\mathbf{V}(\sigma) = \mathbf{V}(K,\sigma)$ (= i times the usual volume plus the Chern–Simons invariant), obtained as the image of ξ_σ under the Borel regulator map or via the dilogarithm, or alternatively its renormalized version $v(\sigma) = v(K,\sigma) = \mathbf{V}(\sigma)/2\pi\mathrm{i}$, which for the geometric representation is the same as the number $v(K)$ occurring in (1.1). (We will usually omit the letter K in this and all similar notations when the knot is not varying.) We extend all of these invariants to $\mathcal{P}_K$ by setting $F_{\sigma_0} = \mathbb{Q}$, $\mathbf{V}(s_0) = 0$, $\xi_{\sigma_0} = 0$. In Section 6 of Part II, we will give more details about the set $\mathcal{P}_K$ and its invariants, and also say something about the situation when the variety of parabolic representations contains positive-dimensional components. In Sections 5 and 6, we will also describe a large (matrix- rather than vector-valued) collection of power series associated to K.

Before explaining how to associate a formal power series to each $\sigma \in \mathcal{P}_K$ and $\alpha \in \mathbb{Q}/\mathbb{Z}$, we first would like to make the above definitions more tangible by describing $\mathcal{P}_K$ explicitly for three simple examples, the knots 4_1 and 5_2 already used above and the $(-2,3,7)$ pretzel knot (henceforth simply $(-2,3,7)$), which will be our basic examples throughout the paper. They have ranks 2, 3, and 6, respectively. For $K = 4_1$, the only elements of $\mathcal{P}_K$ are the three universal ones σ_0, σ_1 and σ_2, the first corresponding to the trivial $\mathrm{SL}_2(\mathbb{C})$-representation with trace field $\mathbb{Q}$ and the other two both with trace field $\mathbb{Q}\big(\sqrt{-3}\big)$, but with the complex embedding $\sqrt{-3} \mapsto -\mathrm{i}\sqrt{3}$ in the second case. The corresponding volumes are $\mathbf{V}(\sigma_0) = 0$, $\mathbf{V}(\sigma_1) = \mathrm{i}V$, and $\mathbf{V}(\sigma_2) = -\mathrm{i}V$, where $V = 2.02\ldots$ is the usual volume, and are all real because the knot 4_1 is amphicheiral. $\big($In general, the mirror knot $\overline{K}$ of a knot has trace fields $F_{\overline{K},\sigma} = \overline{F_{K,\sigma}}$ and $\mathbf{V}(\overline{K},\sigma) = \overline{\mathbf{V}(K,\sigma)}$.$\big)$ For $K = 5_2$ we again have only two essentially different fields $\mathbb{Q}$ and $\mathbb{Q}(\xi)$ with $\xi^3 - \xi^2 + 1 = 0$ (the cubic field with discriminant -23), the latter with three embeddings σ_1, σ_2, and σ_3 corresponding to choosing the root $\xi \in \mathbb{C}$ with negative, positive, or zero imaginary part, respectively. But for the third knot $K = (-2,3,7)$, $\mathcal{P}_K$ consists of seven elements, the trivial representation σ_0, the three representations σ_1, σ_2, σ_3 corresponding to the trace field of K (which is the same as that of 5_2, with its embeddings numbered the same way), and three further elements σ_4, σ_5 and σ_6 corresponding to the field $\mathbb{Q}(\eta)$ with $\eta^3 + \eta^2 - 2\eta - 1 = 0$ (the abelian cubic field with discriminant 49) together with the three embeddings into $\mathbb{C}$ given by sending η to $2\cos(2\pi/7)$, $2\cos(4\pi/7)$ and $2\cos(6\pi/7)$, respectively. In general, to each knot K we associate the algebra $\mathcal{A}_K = \mathbb{Q} \times \mathcal{A}_K^{\rm red}$ defined as the product of the abstract fields F_σ with σ ranging over representatives of the Galois orbits of $\mathcal{P}_K$, so that $\mathcal{P}_K$ $\big($resp. $\mathcal{P}_K^{\rm red}\big)$ can be identified with the set of all algebra maps from $\mathcal{A}_K$ $\big($resp. $\mathcal{A}_K^{\rm red}\big)$ to $\mathbb{C}$; then for our three basic examples, we have

$$\mathcal{A}_{4_1}^{\rm red} = \mathbb{Q}\big(\sqrt{-3}\big), \qquad \mathcal{A}_{5_2}^{\rm red} = \mathbb{Q}(\xi), \qquad \mathcal{A}_{(-2,3,7)}^{\rm red} = \mathbb{Q}(\xi) \times \mathbb{Q}(\eta). \tag{2.1}$$

2.2 Four constructions of the power series $\Phi_\alpha^{(K,\sigma)}(h)$

We will now describe several different approaches to obtaining the formal power series $\Phi_\alpha^{(K,\sigma)}(h)$ associated to an element σ of $\mathcal{P}_K$ and a number $\alpha \in \mathbb{Q}/\mathbb{Z}$.

If $\sigma = \sigma_1$ is the geometric representation, then $\Phi_\alpha^{(K,\sigma)}(h)$ is by definition just the power series $\Phi_\alpha^K(h)$ whose existence is asserted by the quantum modularity conjecture, and for σ in the Galois orbit of σ_1 we simply apply Galois conjugation to this series (at the level of its n-th power if α has denominator n), with some special consideration for the roots of unity occurring. For example, for the knot $K = 4_1$ the series $\Phi_\alpha^{(K,\sigma)}(h)$ for $\alpha = 0$ and $\alpha = a/5$ are the ones given by (1.3) and (1.7), respectively, for $\sigma = \sigma_1$. We then get $\Phi_0^{(K,\sigma_2)}$ simply by replacing $\sqrt{-3}$ by $-\sqrt{-3}$ (or, in this case, replacing h by $-h$ and multiplying by i) in (1.3), and $\Phi_{a/5}^{(4_1,\sigma_2)}$ is obtained from (1.7) by performing the same operation on both $\varepsilon_{a/5}$ and the coefficients $A_n^{(a/5)}$. Similarly, if K is the 5_2 knot then the value of $\Phi_\alpha^{(K,\sigma)}(h)$ at $\alpha = 0$ is given by (1.4) if $\sigma = \sigma_1$, and the values for $\sigma = \sigma_2$ or σ_3 are obtained simply by replacing ξ by its Galois conjugates. In general, the coefficients of these power series lie in the product of a certain root of unity, the c-th root (where c is the denominator of α) of a unit in F_σ, and a conjugate of the same factor $\delta^{-1/2}$ as in the original QMC as described in Section 1. More details about the arithmetic of these numbers will be given in Section 9.

The reader may have noticed that the QMC asserts the existence of the power series $\Phi_\alpha^{(K,\sigma)}(h)$ for $\sigma = \sigma_1$ but gives no clue about how to define them (and especially how to define those for σ not a Galois conjugate of σ_1) given a hyperbolic knot K. A definition of the power series $\Phi_\alpha^{(K,\sigma)}(h)$ for all $\sigma \neq \sigma_0$ was given by Tudor Dimofte and the first author in the two papers [14] (for $\alpha = 0$) and [15] (for general α). What is more, the definition of the series uses as input the gluing equation matrices of an ideal triangulation of the knot complement, along with a solution to the Neumann–Zagier equations. Roughly speaking, one associates to an ideal triangulation of the knot complement a collection of polynomial equations (the Neumann–Zagier equations) whose solutions correspond to the elements of $\mathcal{P}_K^{\mathrm{red}}$, the solution for each $\sigma \in \mathcal{P}_K^{\mathrm{red}}$ being a collection of algebraic numbers (the shape parameters) belonging to the field F_σ. One then associates to each solution of these equations and for each $\alpha \in \mathbb{Q}/\mathbb{Z}$ a certain integral that is evaluated perturbatively by the standard method of Gaussian integration and Feynman diagrams (with a possible ambiguity of multiplication by a power of $\mathbf{e}(\alpha)$). This process, whose details will be reviewed in Section 6, is completely effective and gives, for instance, the three power series

$$\Phi_0^{((-2,3,7),\sigma_j)}(h) = \frac{\xi_j}{\sqrt{6\xi_j - 4}} \Bigg(1 - \frac{33\xi_j^2 - 123\xi_j + 128}{24(3\xi_j - 2)^3} h - \frac{104172\xi_j^2 - 183417\xi_j + 130189}{2\big(24(3\xi_j - 2)^3\big)^2} h^2 + \cdots \Bigg) \tag{2.2}$$

for the elements σ_1, σ_2, and σ_3 of $\mathcal{P}_{(-2,3,7)}$, where ξ_1, ξ_2, ξ_3 are the Galois conjugates of ξ as numbered above, and the three totally different power series

$$\Phi_0^{((-2,3,7),\sigma_{j+3})}(h) = \sqrt{\frac{\eta_j - 2}{14}} \left(1 - \frac{43\eta_j^2 - 21}{168} h - \frac{3928\eta_j^2 + 63\eta_j - 1491}{2 \cdot 168^2} h^2 + \cdots \right) \tag{2.3}$$

for the elements σ_4, σ_5, and σ_6, where $\eta_j = 2\cos(2\pi j/7)$ are the Galois conjugates of η in the ordering given above. The coefficients of the power series $\Phi^{(K,\sigma)}(h)$ for all $\sigma \in \mathcal{P}_K$ have similar arithmetic properties to the special case when σ is Galois conjugate to σ_1.

As well as the "straight" power series $\Phi_\alpha^{(K,\sigma)}(h)$, we will also need the *completed functions*

$$\widehat{\Phi}_\alpha^{(K,\sigma)}(h) = \mathrm{e}^{\mathbf{V}(\sigma)/c^2 h} \Phi_\alpha^{(K,\sigma)}(h), \qquad c = \mathrm{den}(\alpha), \quad \sigma \neq \sigma_0, \tag{2.4}$$

which for the moment we think of as a purely formal expression (the exponential of a Laurent series in h with a simple pole) but which will be given a more precise sense later (cf. Section 10.2). It is this combination that appear in all of our asymptotic formulas, e.g., the right-hand side of (1.5) would become

$$(cN + d)^{-3/2} \widehat{\Phi}_{a/c}^{(K)} \left(\frac{2\pi \mathrm{i}}{cN + d}\right)$$

in this notation. We should also mention here that in [15] the series $\Phi_\alpha^{(K,\sigma)}(h)$ is defined only up to an $2n$-th root of unity, where n is the denominator of α. The generalized QMC that we will present in the next section eliminates this ambiguity (at least up to a net sign depending on σ but not on α).

An idea that will be crucial for this paper is that we have to associate power series $\Phi_\alpha^{(\sigma)}(h) \in \overline{\mathbb{Q}}[[h]]$ to the trivial representation $\sigma = \sigma_0$ as well as to the non-trivial ones to get a coherent total picture. Here there is no Neumann–Zagier data and we use instead a completely different construction based on the Habiro ring. Recall that this ring is defined by

$$\mathcal{H} := \varprojlim \mathbb{Z}[q]/((q;q)_n), \tag{2.5}$$

where $(x;q)_n = \prod_{i=0}^{n-1}\left(1-q^i x\right)$ denotes the q-Pochhammer symbol or "shifted quantum factorial". As mentioned in the introduction, Habiro showed in [52] that the Galois-equivariantly extended Kashaev invariant $J^{(K)}(\alpha)$ ($\alpha \in \mathbb{Q}$) is the evaluation at $q = \mathbf{e}(\alpha)$ of a uniquely defined element, which we will denote by $\mathcal{J}^{(K)}(q)$, of this ring. We then *define* the power series $\Phi_\alpha^{(K,\sigma)}(h)$ for $\sigma = \sigma_0$ by

$$\Phi_\alpha^{(K,\sigma_0)}(h) = \mathcal{J}^{(K)}\big(\mathrm{e}(\alpha)\mathrm{e}^{-h}\big) \in \mathbb{Q}[\mathbf{e}(\alpha)][[h]].$$

For example, for $K = 4_1$ we have the explicit representation (equivalent to (1.2) for $q = \zeta_N$)

$$\mathcal{J}^{(4_1)}(q) = \sum_{n=0}^{\infty}\left(q^{-1};q^{-1}\right)_n (q;q)_n = \sum_{n=0}^{\infty}(-1)^n q^{-n(n+1)/2}(q;q)_n^2 \tag{2.6}$$

of $\mathcal{J}^{(K)}(q)$ as an element of the Habiro ring, and setting $q = \mathrm{e}^{-h} \in \mathbb{Q}[[h]]$ we find

$$\Phi_0^{(4_1,\sigma_0)}(h) = 1 - h^2 + \frac{47}{12}h^4 - \frac{12361}{360}h^6 + \frac{10771487}{20160}h^8 - \cdots \tag{2.7}$$

(which happens to be even because the knot 4_1 is amphicheiral), while the Kashaev invariant for our second standard example 5_2 is given by formula (A.2) below and we find

$$\Phi_0^{(5_2,\sigma_0)}(h) = 1 + h + \frac{5}{2}h^2 + \frac{49}{6}h^3 + \frac{797}{24}h^4 + \frac{19921}{120}h^5 + \cdots.$$

For the $(-2,3,7)$ knot, we have no convenient Habiro-like formula for the Kashaev invariant, but there is still a method (explained in Part II) to obtain its expansion to any order in h at any root of unity just from the values at roots of unity, the expansion at $q = 1$ beginning

$$\Phi_0^{((-2,3,7),\sigma_0)}(h) = 1 - 12h + 129h^2 - \frac{7275}{4}h^3 - \frac{384983}{8}h^4 + \cdots.$$

Note that the complexified volume vanishes for the trivial representation, so that (2.4) would suggest that we should define the completion $\widehat{\Phi}_\alpha^{(K,\sigma_0)}(h)$ to be $\Phi_\alpha^{(K,\sigma_0)}(h)$. But in fact, for reasons that will appear clearly in Section 3, it turns out to be better to define $\widehat{\Phi}_\alpha^{(K,\sigma_0)}(h)$ in this case by

$$\widehat{\Phi}_\alpha^{(K,\sigma_0)}(h) = \left(\frac{ch}{2\pi\mathrm{i}}\right)^{3/2}\Phi_\alpha^{(K,\sigma_0)}(h). \tag{2.8}$$

We have now described constructions of the power series $\Phi_\alpha^{(K,\sigma)}(h)$ for every $\sigma \in \mathcal{P}_K$, but based on very disparate ideas: if σ is the geometric representation or is Galois conjugate to it, we use the quantum modularity conjecture and Galois covariance, for other representations σ different from the trivial one we use a perturbative approach (which is given in [14] and [15] and conjectured there to agree with the first definition when $\sigma = \sigma_1$), and for the trivial representation we define $\Phi_\alpha^{(K,\sigma)}(h)$ by a completely different formula based on the Habiro ring. In fact, as already mentioned in the introduction, there is even a fourth approach in which the series $\Phi_\alpha^{(K,\sigma)}$ are obtained from the asymptotics as q tends radially to $\mathbf{e}(\alpha)$ of certain q-series with integral coefficients. (This connection will not be discussed further here but will be the main theme of [44].) It is then natural to ask why we consider these different series as being similar at all and why we denote them in the same way. In the next two sections, we will present a whole series of properties that justify this.

3 Interrelations among the power series $\Phi^{(\sigma)}_\alpha(h)$

In this section, we describe four empirically found properties, of very different natures, that link and motivate the formal power series introduced above.

3.1 The generalized quantum modularity conjecture

The function $J^{(K)}$ from $\mathbb{Q}/\mathbb{Z}$ to $\overline{\mathbb{Q}}$, which was originally defined as the Galois-equivariant extension of the Kashaev invariant $\langle K\rangle_N$, has now re-appeared as the constant term $\Phi^{(K,\sigma_0)}_\alpha(0)$ of one of a collection of formal power series $\Phi^{(K,\sigma)}_\alpha(h)\in\overline{\mathbb{Q}}[[h]]$ indexed by the elements σ of a finite set $\mathcal{P}_K$ associated to the knot. This suggests that we should look also at the constant terms of the other series as well, i.e., that we should study the functions (*generalized Kashaev invariants*)

$$J^{(K,\sigma)}\colon\ \mathbb{Q}/\mathbb{Z}\to\overline{\mathbb{Q}},\qquad J^{(K,\sigma)}(\alpha):=\Phi^{(K,\sigma)}_\alpha(0) \tag{3.1}$$

for all $\sigma\in\mathcal{P}_K$. These functions turn out to have beautiful arithmetic properties generalizing in a non-obvious way the Habiro-ring property of the original functions $J^{(K)}=J^{(K,\sigma_0)}$. These will be the subject of the subsequent paper [38] and, apart from a few numerical examples, will not be discussed further here. Instead, we will concentrate on the asymptotic properties of the new functions (3.1). In particular, we can ask whether these functions satisfy an analogue of the quantum modularity conjecture for $J^{(K)}$. The answer turns out to be positive, but to involve a number of successive refinements arising from the numerical data. We will present the simplest version here and the strongest versions, which require more preparation, in Sections 4 and 5.

We start once again with the simplest knot $K=4_1$. Here the function $J^{(K,\sigma_0)}(\alpha)=J^{(K)}(\alpha)$ is the one given by (1.2) (with ζ_N replaced by $\mathbf{e}(\alpha)$) whereas the new functions $J^{(K,\sigma_1)}(\alpha)$ and $J^{(K,\sigma_2)}(\alpha)$ are given explicitly by

$$J^{(K,\sigma_1)}(\alpha)=\frac{1}{\sqrt{c}\sqrt[4]{3}}\sum_{Z^c=\zeta_6}\prod_{j=1}^{c}\bigl|1-q^jZ\bigr|^{2j/c},\qquad c=\mathrm{den}(\alpha),\quad q=\mathbf{e}(\alpha) \tag{3.2}$$

and $J^{(K,\sigma_2)}(\alpha)=\mathrm{i}J^{(K,\sigma_1)}(-\alpha)=\overline{J^{(K,\sigma_1)}(\alpha)}$. The original QMC says that $J^{(4_1)}\bigl(\frac{aX+b}{cX+d}\bigr)$ is asymptotically equal to $(cX+d)^{3/2}\widehat{\Phi}^{(4_1)}_{a/c}\bigl(\frac{2\pi\mathrm{i}}{c(cX+d)}\bigr)J^{(4_1)}(X)$ for any matrix $\bigl(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\bigr)\in\mathrm{SL}_2(\mathbb{Z})$ as X tends to infinity with bounded denominator, where the "completion" $\widehat{\Phi}$ is defined by (2.4). When we look at the corresponding asymptotics for the two new functions and for the two simple matrices $\bigl(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\bigr)$ and $\bigl(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\bigr)$ of $\mathrm{SL}_2(\mathbb{Z})$, we see a similar behavior, but with two major differences: the "automorphy factor" $(cX+d)^{3/2}$ is no longer there, and there is a new exponential factor involving the complex volume. Explicitly, what we find experimentally is

$$J^{(4_1,\sigma_1)}(-1/X)\sim \mathrm{e}^{v(K)/(\mathrm{num}(X)\cdot\mathrm{den}(X))}J^{(4_1,\sigma_1)}(X)\widehat{\Phi}^{(4_1)}_0\left(\frac{2\pi\mathrm{i}}{X}\right) \tag{3.3}$$

(here "num" and "den" denote the numerator and denominator) and

$$J^{(4_1,\sigma_1)}(X/(2X+1))\sim \mathrm{e}^{v(K)/((X+\frac12)\cdot\mathrm{den}(X)^2)}J^{(4_1,\sigma_1)}(X)\widehat{\Phi}^{(4_1)}_{1/2}\left(\frac{2\pi\mathrm{i}}{2(2X+1)}\right) \tag{3.4}$$

and similarly for $\Phi^{(\sigma_2)}$ but with $v(K)$ replaced by $v(K,\sigma_2)=-v(K)$. The two equations (3.3) and (3.4) can be written uniformly in the form

$$J^{(4_1,\sigma_1)}(\gamma X)\sim \mathrm{e}^{v(K)\lambda_\gamma(X)}J^{(4_1,\sigma_1)}(X)\widehat{\Phi}^{(4_1)}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right)$$

for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, where $\lambda_\gamma(x)$ is defined for $x = r/s \in \mathbb{Q}$ with r and s coprime by

$$\lambda_\gamma(x) := \frac{1}{\mathrm{den}(x)^2\big(x - \gamma^{-1}(\infty)\big)} = \frac{c}{s(cr+ds)} = \pm\frac{c}{\mathrm{den}(x)\mathrm{den}(\gamma x)}. \tag{3.5}$$

The experiments show that the same thing happens for other knots K and all representations σ, i.e., we can formulate the *generalized quantum modularity conjecture* (GQMC)

$$(cX+d)^{-\kappa(\sigma)}\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}J^{(K,\sigma)}(\gamma X) \sim J^{(K,\sigma)}(X)\widehat{\Phi}^{(K)}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right), \tag{3.6}$$

as $X \to \infty$ with bounded denominator (as usual), and where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ with $c > 0$, and where the *weight* $\kappa(\sigma)$ of the representation $\sigma \in \mathcal{P}_K$ is defined by

$$\kappa(\sigma) = \begin{cases} 3/2 & \text{if } \sigma = \sigma_0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

Notice that (3.6) coincides with the original QMC when $\sigma = \sigma_0$ because in this case the factor $\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}$ on the left-hand side of (3.6) is identically 1. We also see that the two different definitions (2.4) and (2.8) of $\widehat{\Phi}^{(K,\sigma)}$ for $\sigma \neq \sigma_0$ and $\sigma = \sigma_0$ can now be written in a uniform way as

$$\widehat{\Phi}^{(K,\sigma)}_\alpha(h) = |c\hbar|^{\kappa(\sigma)}\mathrm{e}^{v(\sigma)/c^2\hbar}\Phi^{(K,\sigma)}_\alpha(h), \qquad c = \mathrm{den}(\alpha), \quad \hbar := h/2\pi\mathrm{i}, \tag{3.8}$$

which will also be convenient at many other points. Notice that the convention $\hbar = h/2\pi\mathrm{i}$ is almost, but not quite, the same as the one used in ordinary quantum mechanics, and also that the factor $2\pi\mathrm{i}$ relating h and $\hbar$ is the same as that used in our two different normalizations $\mathbf{V}(\sigma)$ and $v(\sigma) = \mathbf{V}(\sigma)/2\pi\mathrm{i}$ of the volume, so that $\mathrm{e}^{v(\sigma)/c^2\hbar} = \mathrm{e}^{\mathbf{V}(\sigma)/c^2h}$.

We end this subsection by proving a cocycle property of the arithmetic function $\lambda_\gamma(X)$ that will be needed in Section 5.

Lemma 3.1. *For all $\gamma, \gamma' \in \mathrm{PSL}_2(\mathbb{Z})$ and $x \in \mathbb{Q} \smallsetminus \{\gamma'^{-1}(\infty), (\gamma\gamma')^{-1}(\infty)\}$, we have*

$$\lambda_{\gamma\gamma'}(x) = \lambda_\gamma(\gamma' x) + \lambda_{\gamma'}(x).$$

Proof. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, $\gamma' = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$ and $\gamma\gamma' = \left(\begin{smallmatrix} a'' & b'' \\ c'' & d'' \end{smallmatrix}\right)$. Then $c''d' - c'd'' = c$, and hence

$$\lambda_{\gamma\gamma'}(x) - \lambda_{\gamma'}(x) = \frac{c''}{s(c''r+d''s)} + \frac{c'}{s(c'r+d's)} = \frac{c}{s(c'r+d's)(c''r+d''s)} = \lambda_\gamma(\gamma' x)$$

as required. A more enlightening way to say this is that $\lambda_\gamma(\gamma'(\infty)) = C(\gamma\gamma', \gamma')$, where

$$C(\gamma_1, \gamma_2) := \frac{c\big(\gamma_1\gamma_2^{-1}\big)}{c(\gamma_1)c(\gamma_2)} = \gamma_2^{-1}(\infty) - \gamma_1^{-1}(\infty),$$

which is a coboundary and hence a cocycle. ■

3.2 Lifting the QMC from constant terms to power series

In the previous subsection, we generalized the original QMC by replacing the Kashaev invariant $J^{(K)}$ by the generalized Kashaev invariants $J^{(K,\sigma)}$ for any $\sigma \in \mathcal{P}_K$. This in turn will be further refined in Section 4 by adding terms of exponentially lower order to the right-hand side of the asymptotic formula. Here we discuss instead a different refinement.

Our starting point, just as in Section 3.1, is that the Kashaev invariant $J^{(K)}(\alpha)$ is equal to the constant term $\Phi_\alpha^{(\sigma_0)}(0)$ of the power series $\Phi_\alpha^{(\sigma_0)}(h)$ as defined in Section 2, so that the original QMC (1.6) can be rewritten as

$$\Phi_{\gamma X}^{(\sigma_0)}(0) \sim (cX+d)^{3/2}\Phi_X^{(\sigma_0)}(0)\widehat{\Phi}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right)$$

for X tending to infinity with fixed fractional part or with bounded denominator. (Here we again omit the knot K from the superscripts when it is not varying to avoid cluttering up the notations. Recall also that $c>0$.) It is then natural to ask whether this asymptotic formula can be lifted to a corresponding statement for the full series $\Phi_\alpha^{(\sigma_0)}(h)$ rather than just its constant term. The answer is affirmative, but with a little twist,

$$\Phi_{\gamma X}^{(\sigma_0)}(h^*) \sim (cX+d)^{3/2}\Phi_X^{(\sigma_0)}(h)\widehat{\Phi}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cx+d)}\right), \tag{3.9}$$

where $x = X - \hbar$ with $\hbar$ as in (3.8) and $h^* = \frac{h}{(cx+d)(cX+d)}$.

Let us explain what the asymptotic expansion (3.9) means in the simplest case of the figure 8 knot. Recall that $\Phi_X^{(\sigma_0)}(h) = \mathcal{J}\big(\mathbf{e}(X)\mathrm{e}^{-h}\big)$ where $\mathcal{J}(q) = \mathcal{J}^{(4_1)}(q)$ is the element of the Habiro ring given by (2.6), related to $J(X) = J^{(4_1)}(X)$ by $J(X) = \mathcal{J}(\mathbf{e}(X))$. Since the Habiro ring is closed under the operator $q\mathrm{d}/\mathrm{d}q$, it contains the function $\mathcal{J}'$ defined by

$$\mathcal{J}'(q) := q\frac{\mathrm{d}}{\mathrm{d}q}\mathcal{J}(q) = \sum_{n=1}^{\infty}\big(q^{-1};q^{-1}\big)_n(q,q)_n\sum_{k=1}^{n}k\frac{1+q^k}{1-q^k}. \tag{3.10}$$

We then define the formal derivative $J'\colon \mathbb{Q}/\mathbb{Z} \to 2\pi\mathrm{i}\overline{\mathbb{Q}}$ by $J'(X) = 2\pi\mathrm{i}\mathcal{J}'(\mathbf{e}(X))$. Then the statement of (3.9) in this case is

$$\frac{1}{(cX+d)^2}\frac{J'\big(\frac{aX+b}{cX+d}\big)}{J\big(\frac{aX+b}{cX+d}\big)} - \frac{J'(X)}{J(X)} \approx -\frac{2\pi\mathrm{i}}{(cX+d)^2}\frac{\widehat{\Phi}'_{a/c}\big(\frac{2\pi\mathrm{i}}{c(cX+d)}\big)}{\widehat{\Phi}_{a/c}\big(\frac{2\pi\mathrm{i}}{c(cX+d)}\big)}, \tag{3.11}$$

interpreted in the following sense. The left-hand side of (3.11) is $2\pi\mathrm{i}$ times an algebraic number belonging to some fixed cyclotomic field for each fixed element $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and bound on the denominator of X, while the right-hand side is defined only as a divergent power series in $(cX+d)^{-1}$. The claim is then that when we compute both sides of (3.11) for fixed γ and for X tending to infinity with bounded denominator, using (3.10) to compute the terms $J'(X)$ and $J'(\gamma X)$ as exact algebraic numbers, the two expressions agree numerically to all orders in $1/X$, and this is the statement that we verified numerically for several elements γ and sequences of large rational numbers X. Note that (3.11) is almost what we would get if we differentiated the original QMC formula (1.6) logarithmically (which of course we are not allowed to do since it is only an asymptotic statement valid for large rational numbers X with fixed denominator and hence is rigid), except that then we would have an extra term $\frac{3}{2}\frac{c}{cX+d}$ which is in fact not present because equation (3.9) contains $(cX+d)^{3/2}$ rather than $(cx+d)^{3/2}$.

All of this was for the trivial connection σ_0. If we consider instead an arbitrary element σ of $\mathcal{P}_K$, then what we find is the obvious combination of (3.6) (which was only for the constant terms $\Phi(0)$) and (3.9) (which gave the “twist” needed to include h), namely

$$(cX+d)^{-\kappa(\sigma)}\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}\Phi_{\gamma X}^{(\sigma)}(h^*) \sim \Phi_X^{(\sigma)}(h)\widehat{\Phi}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cx+d)}\right), \tag{3.12}$$

with $x = X-\hbar$ and $h^* = h/(cx+d)(cX+d)$ as in (3.9).

Equation (3.12) differs in two notable ways from the original QMC (1.6): the appearance of the "tweaking factor" $\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}$ and the change of infinitesimal variable from h to h^*. In fact, the first is explained very simply by replacing the two power series $\Phi^{(\sigma)}$ in (3.12) by their completions as defined in (3.8), because a short calculation shows that the number $\lambda_\gamma(X)$ defined in (3.5) is equal to the difference between $1/\mathrm{den}(X)^2\hbar$ and $1/\mathrm{den}(\gamma X)^2\hbar^*$ with $\hbar^* := \hbar^*/2\pi\mathrm{i} = \hbar/(cx+d)(cX+d)$, so that (3.12) becomes simply

$$\widehat{\Phi}^{(K,\sigma)}_{\gamma X}(h^*) \sim (cx+d)^{-\kappa(\sigma)}\widehat{\Phi}^{(K,\sigma)}_X(h)\widehat{\Phi}^{(K,\sigma_1)}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cx+d)}\right), \tag{3.13}$$

where we have now again included the complete labels of the $\widehat{\Phi}$ series for clarity. In this version both the tweaking factor $\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}$ and the automorphy factor $(cX+d)^{3/2}$ have been absorbed into the completed power series, but then producing a new automorphy factor $(cx+d)^{-3/2}$. Finally, the "twisting" from h to h^* is partly motivated by the calculation just given and the simplifications in (3.13), but more conceptually by observing that $x = X - \hbar$ implies $\gamma x = \gamma X - \hbar^*$. Equation (3.13) will then take on an even more natural form in terms of the notion of "functions near $\mathbb{Q}$" that will be introduced in Section 5.

3.3 Quadratic relations

The next interconnection among the power series $\Phi^{(K,\sigma)}_\alpha(h)$ associated to a given knot K that we discover (experimentally, as always) from the examples is that they satisfy an unexpected quadratic relation, namely

$$\sum_{\sigma\in\mathcal{P}^{\mathrm{red}}_K} \Phi^{(K,\sigma)}_\alpha(h)\Phi^{(K,\sigma)}_{-\alpha}(-h) = 0. \tag{3.14}$$

Notice that this relation is non-trivial even at the level of its constant term, where it says, for example, that the value of the generalized Kashaev invariant $J^{(5_2,\sigma)}(\alpha)$ defined in the last subsection belongs to the kernel of the trace map from $\mathbb{Q}(\xi,\zeta_\alpha)$ to the trace field $\mathbb{Q}(\xi)$ of 5_2 for every rational number α. The special case of this when $\alpha = 0$ was observed independently by Gang, Kim and Yoon [22].

The relation (3.14) is practically vacuous for the figure 8 knot, since in that case it follows immediately from the identity $\Phi^{(4_1,\sigma_2)}_\alpha(h) = \mathrm{i}\Phi^{(4_1,\sigma_1)}_{-\alpha}(-h)$ mentioned at the beginning of Section 2.2. (Stated differently, if we multiply the series (1.3) by its value at $-h$, we obtain an element of $\sqrt{-3}\mathbb{Q}\big[\big[h^2\big]\big]$, so that the trace down to $\mathbb{Q}$ vanishes, and similarly for (1.7).) But for the 5_2 knot the identity is non-trivial even at $\alpha = 0$, where (1.4) gives

$$\Phi^{5_2}(h)\Phi^{5_2}(-h) = \frac{1}{3\xi-2} + \frac{102\xi^2-183\xi+135}{(3\xi-2)^7}h^2 - \frac{143543\xi^2-252029\xi+190269}{4(3\xi-2)^{13}}h^4 + \cdots$$

in which one can check that the three coefficients given, and in fact all coefficients up to order h^{108}, lie in the kernel of the trace map from $\mathbb{Q}(\xi)$ to $\mathbb{Q}$. Notice, by the way, that the series here has much simpler coefficients (specifically, much smaller denominators) than the individual factors as given by (1.4). This is a special case of a more general phenomenon that will be discussed in [38]. When we look at (3.14) for this knot but other values of α, the same thing happens: the m-th root of a unit in $\mathbb{Q}(\xi,\zeta_m)$ that is a common factor of each $\Phi^{(5_2)}_\alpha(h)$ when α has denominator m cancels when we multiply the series at α and $-\alpha$, and the series in $\mathbb{Q}(\xi,\zeta_m)[[h]]$ that we find, although it is no longer even when α is different from 0 or 1/2, always has coefficients lying in the kernel of the trace map from $\mathbb{Q}(\xi,\zeta_m)$ to $\mathbb{Q}(\zeta_m)$.

The above illustrates the relation (3.14) for our second simplest knot 5_2. For our third standard example $K = (-2, 3, 7)$, this relation is even more surprising because now $\mathcal{P}_K$ has two Galois orbits, as discussed in Section 2, and the quadratic relation relates them to one another. Specifically, if we consider separately the contributions from σ_i for $1 \le i \le 3$ and for $4 \le i \le 6$, then equation (2.2) gives

$$\begin{aligned}\sum_{j=1}^{3} & \Phi^{(K,\sigma_j)}(h)\Phi^{(K,\sigma_j)}(-h) \\ &= \mathrm{Tr}_{\mathbb{Q}(\xi)/\mathbb{Q}}\left(\frac{\xi^2}{2(3\xi-2)} + \frac{605\xi^2 - 1217\xi + 878}{2^4(3\xi-2)^7}h^2 + \cdots\right) \\ &= \frac{1}{2} + 0h^2 - \frac{13}{2^6}h^4 + \frac{2987}{2^{11}\cdot 3}h^6 + \frac{3517753}{2^{16}\cdot 5}h^8 - \frac{110362454561}{2^{19}\cdot 3^3\cdot 5\cdot 7}h^{10} - \cdots\end{aligned}$$

and equation (2.3) gives

$$\begin{aligned}\sum_{j=4}^{6} & \Phi^{(K,\sigma_j)}(h)\Phi^{(K,\sigma_j)}(-h) \\ &= \mathrm{Tr}_{\mathbb{Q}(\eta)/\mathbb{Q}}\left(\frac{\eta-2}{2\cdot 7} + \frac{18811\eta^2 - 78046\eta + 67485}{2^8\cdot 3\cdot 7^4}h^2 + \cdots\right) \\ &= -\frac{1}{2} + 0h^2 + \frac{13}{2^6}h^4 - \frac{2987}{2^{11}\cdot 3}h^6 - \frac{3517753}{2^{16}\cdot 5}h^8 + \frac{110362454561}{2^{19}\cdot 3^3\cdot 5\cdot 7}h^{10} + \cdots .\end{aligned}$$

Each of these two series belongs to $\mathbb{Q}\big[\big[h^2\big]\big]$. Computing many more terms $\big($we went up to $O\big(h^{38}\big)\big)$, we find that their sum vanishes, confirming the quadratic relation in a very striking way and at the same time showing a subtle interdependence between the two cubic number fields associated to this knot. We note, however, that these are only two of the three number fields making up the algebra $\mathcal{A}_{(-2,3,7)} = \mathbb{Q}\times\mathbb{Q}(\xi)\times\mathbb{Q}(\eta)$ as defined in (2.1). We have not found *any* relation between the power series $\Phi_\alpha^{(K,\sigma_0)}(h)$ or $\Phi_\alpha^{(K,\sigma_0)}(h)\Phi_{-\alpha}^{(K,\sigma_0)}(-h)$ and the power series $\Phi_\alpha^{(K,\sigma)}(h)$ for $\sigma \ne \sigma_0$. This is reflected in the fact that the summation in (3.14) is over $\mathcal{P}_K^{\mathrm{red}}$ and not over all of $\mathcal{P}_K$.

We end this subsection by mentioning that, as well as the quadratic relation (3.14), there are also *bilinear* expressions in the $\Phi^{(K,\sigma)}$ that are not zero, but (experimentally, and in some cases provably) are *convergent* rather than factorially divergent power series. This will be discussed briefly in Section 5.4 and in detail in the companion paper [44]. Here we give only a numerical example. In Proposition 5.2 below, we will give certain explicit bilinear combinations of the Φ-series which we believe are the Taylor expansions of analytic functions and hence have a positive (and known) radius of convergence. In the simplest case (corresponding in the notation of Proposition 5.2 to the (σ_1, σ_1) component of the matrix $W_S^{(4_1)}(1+x)$, where $S = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ as usual, combined with (5.9)), this power series is given by

$$\mathrm{e}^{-v(4_1)}\Phi(2\pi\mathrm{i}x)\Phi\left(-\frac{2\pi\mathrm{i}x}{1+x}\right) - \mathrm{e}^{v(4_1)}\Phi\left(\frac{2\pi\mathrm{i}x}{1+x}\right)\Phi(-2\pi\mathrm{i}x), \tag{3.15}$$

with $\Phi = \Phi_0^{(4_1)}$ as given in (1.3). The power series Φ has coefficients growing like $n!$ times an exponential function (the precise asymptotics will be described in the next subsection) and has 100th coefficient of the order of 10^{94}, but the combination (3.15) has radius of convergence 1 and, for instance, 100th coefficient of order 10^{-3}. Notice that if we replace all Φ's in (3.15) by the corresponding $\widehat{\Phi}$'s, then the prefactors $\mathrm{e}^{\pm v(4_1)}$ disappear.

3.4 Asymptotics of the coefficients

The third interrelationship between the series $\Phi_\alpha^{(\sigma)}$ for different elements σ of $\mathcal{P}_K$ arises via the asymptotics of their coefficients.

For both theoretical and numerical purposes, we need to be able to compute the "values" of the divergent series $\Phi_\alpha^{(\sigma)}(h)$ for very small h, and for this we need to know how their coefficients grow. We will write $A_\alpha^{(\sigma)}(n) = A_\alpha^{(K,\sigma)}(n)$ for the coefficient of h^n in $\Phi_\alpha^{(K,\sigma)}(h)$.

As usual, we start with the simplest example $K = 4_1$, $\sigma = \sigma_1$ (geometric representation), and $\alpha = 0$, where $\Phi_\alpha^{(K,\sigma)}(h)$ is just the series (1.3). Let us write just $A(n)$ for its n-th coefficient $\big($so $A(0) = 3^{-1/4}$, $A(1) = 11A_0/72\sqrt{-3}\big)$. Calculating many coefficients and using a standard numerical extrapolation method that is recalled in Part II, we find that $A(n)$ grows factorially like $(n-1)!\lambda^{-n}\big(c_0 + c_1 n^{-1} + c_2 n^{-2} + \cdots\big)$ for some constants λ and c_i. The numbers λ and c_0 are easily recognized to be $2\mathbf{V}(K) = 2\mathrm{i}\mathrm{Vol}(K)$ and $3A(0)/2\pi$, respectively, but the further coefficients c_i have more and more complicated expressions. It turns out that a much more convenient representation for the asymptotics is as a sum of shifted factorials $(n-1-\ell)!$ rather than of terms $n!/n^\ell$, because in this version we find the expansion

$$A(n) \sim \frac{3}{2\pi}\sum_{\ell\geq 0}(-1)^\ell A(\ell)\frac{(n-\ell-1)!}{(2\mathbf{V}(4_1))^{n-\ell}} \tag{3.16}$$

with easily recognizable coefficients to all orders. If we now recall that $\Phi_0^{(4_1,\sigma_2)}(h)$ equals $\mathrm{i}\Phi_0^{(4_1,\sigma_1)}(-h)$ and hence $A_0^{(4_1,\sigma_2)}(n) = (-1)^n A_n \mathrm{i}$, then we can recognize (3.16) as one of a pair of coupled asymptotic expansions

$$A_0^{(\sigma_1)}(n) \sim \frac{3}{2\pi\mathrm{i}}\sum_{\ell\geq 0}A_0^{(\sigma_2)}(\ell)\frac{(n-1-\ell)!}{(2\mathbf{V}(4_1))^{n-\ell}}, \qquad A_0^{(\sigma_2)}(n) \sim \frac{-3}{2\pi\mathrm{i}}\sum_{\ell\geq 0}A_0^{(\sigma_1)}(\ell)\frac{(n-\ell-1)!}{(-2\mathbf{V}(4_1))^{n-\ell}}.$$

This already looks quite nice, but the picture becomes even clearer when we consider also the coefficients $B(0) = 1$, $B(1) = 0$, $B(2) = -1$, $\ldots$ of the third series $\Phi_0^{\sigma_0}$ as given in (2.7). Since the $B(n)$ vanish for n odd, it would first seem that one has to give separate asymptotic formulas according to the parity of n, but a better way is to write $B(n) = A_0^{(\sigma_0)}(n)$ as a sum of *two* asymptotic expansions labelled by the two other elements σ_1 and σ_2 of $\mathcal{P}_K$:

$$B(n) \sim \sqrt{2\pi}\sum_{\ell\geq 0}A_0^{(\sigma_1)}(\ell)\frac{\Gamma\big(n-\ell+\frac{3}{2}\big)}{(-\mathbf{V}(4_1))^{n-\ell+3/2}} - \sqrt{2\pi}\sum_{\ell\geq 0}A_0^{(\sigma_2)}(\ell)\frac{\Gamma\big(n-\ell+\frac{3}{2}\big)}{\mathbf{V}(4_1)^{n-\ell+3/2}}. \tag{3.17}$$

Here we observe that the expressions $2\mathbf{V}(4_1)$, $-2\mathbf{V}(4_1)$, $-\mathbf{V}(4_1)$ and $\mathbf{V}(4_1)$ occurring in the denominators of the last two formulas can be written in a uniform way as $\mathbf{V}(\sigma_1) - \mathbf{V}(\sigma_2)$, $\mathbf{V}(\sigma_2) - \mathbf{V}(\sigma_1)$, $\mathbf{V}(\sigma_0) - \mathbf{V}(\sigma_1)$ and $\mathbf{V}(\sigma_0) - \mathbf{V}(\sigma_2)$, respectively. Exactly analogous asymptotic statements turn out to hold for the coefficients of the series $\Phi_\alpha^{(4_1,\sigma)}$ also for $\alpha \neq 0$, with the same coefficients, leading for this knot to the uniform conjectural statement

$$A_\alpha^{(K,\sigma)}(n) \sim (2\pi)^{\kappa_\sigma - 1}\sum_{\sigma'\neq\sigma}M_K(\sigma,\sigma')\sum_{\ell\geq 0}A_\alpha^{(\sigma')}(\ell)\frac{\Gamma(n-\ell+\kappa_\sigma)}{\big(\mathbf{V}(\sigma)-\mathbf{V}(\sigma')\big)^{n-\ell+\kappa_\sigma}}, \tag{3.18}$$

for all elements $\sigma \in \mathcal{P}_K$ and all $\alpha \in \mathbb{Q}$, where κ_σ is defined as in (3.7) and where the coefficients $M_K(\sigma,\sigma')$ are integers independent on α, given for $K = 4_1$ by

$$M_{4_1} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & -3 \\ 0 & 3 & 0 \end{pmatrix}.$$

Experiments with our other two standard sample knots 5_2 and $(-2,3,7)$ reveal the same asymptotic behavior (3.18), with the matrix M_K given in these two cases by

$$M_{5_2} = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 0 & 0 & 4 & -3 \\ 0 & -4 & 0 & -3 \\ 0 & 3 & 3 & 0 \end{pmatrix}, \qquad M_{(-2,3,7)} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 2 & 2 \\ 0 & 0 & 0 & 1 & 2 & 2 & 2 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We end this subsection by making a number of remarks about the asymptotic formula (3.18) and the matrices M_K.

1. The coefficients of the matrices M_K are much simpler invariants of K than the coefficients of the power series $\Phi_\alpha^{(\sigma)}$, because they are rational integers rather than algebraic numbers and also do not depend on α. It would be of considerable interest to have an direct topological definition of these numbers rather than just an indirect one in terms of the (still conjectural) asymptotic formula (3.18). One possibility in Section 2 is that they are related to the counting of flow lines in Floer homology. They are also presumably the same as the skew-symmetric matrices of "Stokes indices" as recently introduced by Kontsevich [61].

2. The different forms of the asymptotics of the coefficients of $\Phi_\alpha^{(\sigma)}$ for $\sigma = \sigma_0$ and $\sigma \neq \sigma_0$ are directly related to the different weights and different completions of these series as given in equation (3.8).

3. A different asymmetry between the trivial and non-trivial representations is seen in the fact that $M_K(\sigma, \sigma_0)$ always vanishes but $M_K(\sigma_0, \sigma)$ does not, meaning that the large-index coefficients of the $\Phi^{(\sigma_0)}$ series "see" the small-index coefficients of the Φ^σ series for $\sigma \neq \sigma_0$ but not vice versa. It is interesting to note that similar "one-way phenomenon" regarding the matrices appearing in [78], see also Gukov et al. [49, 50].

4. In all three examples given above, we further observe that apart from their first column, which vanishes, and first row, which does not, the matrices M_K are skew-symmetric, i.e., $M_K(\sigma, \sigma') = -M_K(\sigma', \sigma)$ for $\sigma, \sigma' \neq \sigma_0$. This phenomenon, which we expect to hold for all knots, will be shown below to be a formal consequence of the quadratic relation (3.14).

5. We also observe that the lower 4×4 block of the matrix $M_{(-2,3,7)}$ vanishes identically. In view of the numbering of the indices, this means that $M_K(\sigma, \sigma')$ vanishes whenever σ and σ' are both real and distinct from σ_0. This in fact holds for all knots and is a special case of the more general identity $M_K\big(\overline{\sigma}, \overline{\sigma'}\big) = -M_K(\sigma, \sigma')$ for all $\sigma, \sigma' \neq \sigma_0$, which we can prove easily (assuming that the expansion (3.18) is correct) simply by taking the complex conjugate of (3.18) and noting that $\mathbf{V}(\overline{\sigma})$ and $A_\alpha^{(\overline{\sigma})}(n)$ are the complex conjugates of $\mathbf{V}(\sigma)$ and $A_{-\alpha}^{(\sigma)}(n)$, respectively (and, of course, that the coefficients of M_K are real). The minus sign arises from the pure imaginary prefactor $(2\pi \mathrm{i})^{-1}$ in (3.18).

6. A corollary of (3.18) is the growth estimate

$$A_\sigma^{(\sigma)}(n) = O\big(n^{\kappa_\sigma - 1} n! \Delta(\sigma)^{-n}\big),$$

where

$$\Delta(\sigma) = \Delta(K, \sigma) = \min_{M_K(\sigma,\sigma') \neq 0} \big|\mathbf{V}(\sigma) - \mathbf{V}(\sigma')\big|.$$

This estimate will be important for the optimal truncation that is used in the next section and discussed in more detail in Section 10 and in [45].

7. We should also mention that there is still some sign ambiguity in the definition of the matrix M_K. At the moment, even assuming the validity of the various conjectures presented in the next two sections, we can only normalize the power series $\Phi_\sigma^{(\sigma)}(h)$ up to the ambiguity of a sign $\varepsilon_\sigma \in \{\pm 1\}$ independent of α but depending on σ, and making this change would multiply $M_K(\sigma,\sigma')$ by $\varepsilon_\sigma\varepsilon_{\sigma'}$ (which would not affect either of the properties mentioned in **3** and **4** above). Similarly, when $\sigma=\sigma_0$ the formula defining $M_K(\sigma,\sigma')$ has an inherent ambiguity coming from the choice of sign of square-root of $V_\sigma - V_{\sigma'} = -V_{\sigma'}$ in (3.18) (only in the first term; the choices for the other terms are then determined in the obvious way), so that each of the matrix entries $M_K(\sigma_0,\sigma')$ is actually only well defined up to sign. Of course, it is possible that there is some canonical way to normalize everything to eliminate these ambiguities, but we do not yet know how to do this.

8. Actually, however, there is a problem with all of these statements that we have glossed over so far but that does need to be addressed. This is that the right-hand side of (3.18) does not really make sense as it stands, since the terms on the right-hand side are given by divergent series and hence can be computed only up to some level of precision, but at the same time have exponentially different orders of growth, so that it is not a priori clear what it means to add them. In the case of 4_1, we did not see this problem, because there is only one term in (3.18). This point will be discussed briefly in Section 10.2 and in detail in [45].

4 Refining the quantum modularity conjecture

In this section, we will show how one can go beyond the original QMC as described in Section 1 or its generalization as described in Section 3.1. We will present this via a series of successive refinements, each one found experimentally and building on its predecessors. This will culminate in the complete, though of course still conjectural, definition (in Sections 4.1–4.4) of the matrices $\mathbf{J}$ and $\mathbf{\Phi}$ discussed in the introduction and of the final refinement (in Section 4.5) of the original quantum modularity conjecture.

4.1 Improving the quantum modularity conjecture: optimal truncation

The QMC in its original form says that $J^{(K)}(-1/X)$ agrees with $X^{3/2}J^{(K)}(X)\widehat{\Phi}_0^{(K)}(2\pi\mathrm{i}/X)$ to all orders in $1/X$ as X tends to infinity with fixed denominator, with a similar statement when $-1/X$ is replaced by $\frac{aX+b}{cX+d}$ for any $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. A natural question is whether we can do better than this and obtain an asymptotic estimate, or even a precise asymptotic formula, for the *difference* of these two expressions. At first sight this seems to makes no sense, since $\widehat{\Phi}_0^{(K)}(h)$ $\big($or more generally $\widehat{\Phi}_{a/c}^{(K)}(h)\big)$ is given in terms of a divergent power series that a priori does not have a numerical value but rather gives only an approximation up to any given order in h. But we can remedy this by replacing the series $\Phi(h)=\Phi_0^{(K)}(h)$ or $\Phi_{a/c}^{(K)}(h)$ by its "optimal truncation" $\Phi(h)^{\mathrm{opt}}$ obtained by truncating the divergent infinite series at the value of N (depending on h) where the terms of this series become smallest in absolute value, a little like what is done in physics when for instance the magnetic moment of the electron is computed to high accuracy by truncating a divergent sum of Feynman integrals at a suitably small term. If $\Phi(h)=\sum A_n h^n$ with A_n growing like $n!/B^n$ for some complex number B, then this "naive optimal truncation" is given by $\sum_{n=0}^N A_n h^n$ with N chosen near to $|B/h|$. Then the first term neglected, and hence also the expected error, is of the order of magnitude of e^{-N}, so we have a way to define $\Phi(h)$ up to an exponentially small error rather than only up to fixed powers of h. Of course, to get a completely well-defined function $\Phi(h)^{\mathrm{opt}}$ we would have to fix a prescription for choosing N, say as the floor or ceiling or nearest integer to $|B/h|$ (and perhaps also dividing by 2 the last term retained), but since the terms with $n\approx|B/h|$ are all very small the specific choice is not important and we will do better later anyway.

Using the description of the asymptotics of the coefficients of the series $\Phi_\alpha^{(K,\sigma)}(h)$ given in Section 3.4 above, we can compute their optimal truncations explicitly. Starting as usual with the simplest example $K = 4_1$ and $\alpha = 0$ (and also $\Phi = \Phi^{(\sigma_1)} = \Phi^{(4_1,\sigma_1)}$, the series occurring in the QMC), we have from (3.16) the estimate $A_n = O(n!/(2V)^n)$ with $V = \mathrm{Vol}(4_1) = 2.02988\ldots$, so the optimal truncation occurs for N near $2V/|h|$. The expected error in $\Phi(h)$ for $h = 2\pi\mathrm{i}/X$ is therefore of the order of $\mathrm{e}^{-2v(4_1)X}$, with $v(4_1) = V/2\pi = 0.32306\ldots$, and since the completed function $\widehat{\Phi}(h) = \mathrm{e}^{V/h}\Phi(h)$ grows like $\mathrm{e}^{v(4_1)X}$, this means that not only the relative but even the absolute expected error in $\widehat{\Phi}(h)^{\mathrm{opt}}$ is exponentially small in this case. As a numerical example, we consider the value $X = 100$. The Kashaev invariant $\langle 4_1\rangle_{100}$, which we can compute to arbitrary precision from (2.6) with $q=\zeta_{100}$, has the approximate value 81985188380512462.9310054954341, while the corresponding value $100^{3/2}\widehat{\Phi}\big(\frac{2\pi\mathrm{i}}{100}\big)^{\mathrm{opt}}$ (obtained in this case by retaining the first 66 coefficients of the divergent series) has the numerical value 81985188380512461.9269943535808 with an expected error of the order of 10^{-12}. We see immediately that these two numbers are not equal within the accuracy of the computation, so that the most obvious first guess for a more precise version of the QMC is not true. But when we look at the difference of these two numbers we find the numerical value

$$\langle 4_1\rangle_{100} - 100^{3/2}\widehat{\Phi}\left(\frac{2\pi\mathrm{i}}{100}\right)^{\mathrm{opt}} \approx 1.00401114185,$$

which is very close to 1. Repeating the experiment for other large integral values of X, we find that this difference has the asymptotic expansion $1 - h^2 + \frac{47}{12}h^4 + \cdots$ (with $h = 2\pi\mathrm{i}/X$ as before), which we recognize easily as the power series $\Phi^{(4_1,\sigma_0)}(h)$ as given in (2.7), and a numerical computation shows that indeed the optimal truncation of that series at $h = \frac{2\pi\mathrm{i}}{100}$ has precisely the same value 1.00401114185, to the same precision. Repeating the calculations with other integral and non-integral values of X and also for $J^{(4_1)}(\gamma X)$ for matrices $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ other than $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$, we find the same behavior in all cases, leading to the conjectural asymptotic formula

$$\begin{aligned}&(cX+d)^{-3/2}J^{(4_1)}\left(\frac{aX+b}{cX+d}\right)\\ &\qquad\overset{?}{\approx} J^{(4_1)}(X)\widehat{\Phi}^{(4_1,\sigma_1)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right) + \widehat{\Phi}^{(4_1,\sigma_0)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right)\end{aligned} \tag{4.1}$$

for any matrix $\gamma = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and for $X \to \infty$ with fixed (or bounded) denominator, with the coefficient of the second series $\widehat{\Phi}^{(4_1,\sigma_0)}$ being 1 for all X and γ.

Here the natural question arises whether one can improve the precision of (4.1) even further by adding to the right-hand side a *third* term involving $\widehat{\Phi}^{(4_1,\sigma_2)}$, the last of the three completed series for the 4_1 knot. But for the moment we can't even make sense of this since the intrinsic error in the optimal-truncation values of both $\widehat{\Phi}^{(4_1,\sigma_1)}(h)$ and $\widehat{\Phi}^{(4_1,\sigma_0)}(h)$ has exponential decay of the order of $\mathrm{e}^{-v(K)X}$ (for the first function because it grows like $\mathrm{e}^{+v(K)X}$ and has a relative error $\mathrm{e}^{-2v(K)X}$, as we have already seen, and for the second because it grows only like a power of X but has a larger relative error $\mathrm{e}^{-v(K)X}$ by virtue of (3.17)). This is the same as the order of growth of the third function $\widehat{\Phi}^{(4_1,\sigma_2)}(h)$, so that dividing the difference of the left- and right-hand sides of (4.1) by $\widehat{\Phi}^{(4_1,\sigma_2)}(h)$, with all $\widehat{\Phi}$-series defined by optimal truncation, would give meaningless values. We will return to this problem in Section 4.3 below. Before doing that, however, we first look at two other knots for which a new phenomenon appears that is not visible for 4_1.

4.2 New elements of the Habiro ring

For the knot $K = 5_2$ the set $\mathcal{P}_K$ has four elements: the Habiro one, the geometric and antigeometric ones, and the one corresponding to the real embedding of the cubic field $F_K = \mathbb{Q}(\xi)$. However, it has only three distinct real volumes: the geometric volume $\mathrm{Im}\,\mathbf{V}(\sigma_1) = \mathrm{Vol}(K)$ (with the numerical value $2.82812\ldots$), the anti-geometric volume $\mathrm{Im}\,\mathbf{V}(\sigma_2) = -\mathrm{Vol}(K)$, and 0 for both $\sigma = \sigma_0$ and $\sigma = \sigma_3$, and consequently only three distinct orders of growth (one exponentially large, one exponentially small, and two of polynomial growth) of the corresponding $\widehat{\Phi}$-functions $\widehat{\Phi}^{(K,\sigma)}(h)$. (For simplicity we concentrate for the moment only on $\alpha = 0$ and omit it from the notations.) This means that in the analogue of (4.1) there is only one term that is too small to be seen numerically when we replace the Φ-series by their optimal truncation, so that here one can hope to see *three* distinct terms on the right. To test this, we take the same values $N = 100$, $h = 2\pi\mathrm{i}/N$ as before. Then $J^{(K)}\big(-\frac{1}{N}\big)$ is of the order of magnitude of 10^{22} and the difference $N^{-3/2}J^{(K)}\big(-\frac{1}{N}\big) - \widehat{\Phi}^{(K,\sigma_1)}\big(h\big)^{\mathrm{opt}} - \widehat{\Phi}^{(K,\sigma_0)}\big(h\big)^{\mathrm{opt}}$ is of the order of 1 just as before, but now when we divide this difference by $\widehat{\Phi}^{(K,\sigma_3)}\big(h\big)^{\mathrm{opt}}$ we obtain $2+(1.22-5.23\mathrm{i})\cdot 10^{-9}$, strongly suggesting that the limiting value of this difference as X tends to infinity through integers exists and is equal to 2. Further experiments for non-integral values of X and for other matrices $\gamma = \big(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\big)$ then lead to the new conjectural asymptotic statement

$$(cX+d)^{-3/2} J^{(5_2)}\left(\frac{aX+b}{cX+d}\right) \overset{?}{\approx} J^{(5_2)}(X)\widehat{\Phi}_{a/c}^{(5_2,\sigma_1)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right) + \widehat{\Phi}_{a/c}^{(5_2,\sigma_0)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right) + Q^{(5_2)}(X)\widehat{\Phi}_{a/c}^{(5_2,\sigma_3)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right) \tag{4.2}$$

for all γ and all X tending to infinity with bounded denominator, where $Q^{(5_2)}(x)$ (which is a temporary notation, only for this knot) is a function that is independent of γ but that, unlike the constant coefficient 1 of the Habiro term $\widehat{\Phi}^{(K,\sigma_0)}(h)$, is not independent of x. Instead, $Q^{(5_2)}(x)$ is numerically found to be a periodic function of period 1 taking on simple algebraic values, the first few being

$$Q^{(5_2)}(0) = 2, \qquad Q^{(5_2)}\big(\tfrac{1}{2}\big) = 8, \qquad Q^{(5_2)}\big(\pm\tfrac{1}{3}\big) = \frac{37 \pm 3\sqrt{-3}}{2}, \qquad Q^{(5_2)}\big(\pm\tfrac{1}{4}\big) = 29 \pm 13\mathrm{i}.$$

(These values were found experimentally, using a Chinese-remainder-type interpolation, and the existence of such functions for all knots is not known.) Looking at more values (specifically, for all x with denominator up to 200), we find that $Q^{(5_2)}(x)$ belongs to $\mathbb{Z}[\mathbf{e}(x)]$ and is Galois-invariant, so we can write it as $\mathcal{Q}^{(5_2)}(\mathbf{e}(x))$ where $\mathcal{Q}^{(5_2)}(q)$ is an element of $\mathbb{Z}[q]$ for every root of unity q, the first values being given by

$\mathrm{Ord}(q)$	1	2	3	4	5	6
$\mathcal{Q}^{(5_2)}(q)$	2	8	$20+3q$	$29+13q$	$69+27q+37q^2+2q^3$	$-46+69q$

This suggests that $q \mapsto \mathcal{Q}^{(5_2)}(q)$ might be an element of the Habiro ring $\mathcal{H}$ defined in (2.5), just as we know is the case for the coefficient $J^{(5_2)}(X)$ of the first $\widehat{\Phi}$-term in (4.2). This hypothesis can be tested numerically, because a well-known property of any element $\mathcal{Q} \in \mathcal{H}$ (originally observed by Ohtsuki [68] in the context of the WRT-invariants of integer homology spheres even before the Habiro ring had been formalized) is that it satisfies an infinite number of congruences, the simplest of which is that $\mathcal{Q}(\zeta_p)$ for every prime number p should be congruent modulo p to $c_0 + c_1\pi_p + c_2\pi_p^2 + \cdots + c_{p-2}\pi_p^{p-2}$, where $\pi_p = \zeta_p - 1$ is the prime dividing p in $\mathbb{Q}(\zeta_p)$ and where the c_i are rational integers independent of p. This means in our case that the coefficient of x^i in the polynomial $Q_p(1+x) \in \mathbb{Z}[x]$ should be congruent modulo p to a *fixed* integer $c_i \in \mathbb{Z}$ for all primes $p > i+1$, and testing this for the numerically obtained polynomials Q_n, we find that it is indeed true, with $Q_p(\zeta_p) \equiv \mathrm{Oh}(\pi_p) \pmod{p}$ for a power series $\mathrm{Oh}(x) \in \mathbb{Z}[[x]]$ beginning

$$\mathrm{Oh}(x) = 2 - 3x + 8x^2 - 28x^3 + 120x^4 - 614x^5 + 3669x^6 - 25125x^7 + O\big(x^8\big).$$

In fact, later we were able to guess an explicit formula, given below in Section 7.1, that is manifestly in the Habiro ring and that reproduces the values of the polynomials $Q_n(q)$ and power series $\mathrm{Oh}(x)$ as given above. But in many other cases, including the $(-2,3,7)$ knot discussed below, we cannot give even conjectural explicit formulas of the required kind, and in such cases it is important to be able to have a numerical test of the Habiro-ness of a periodic function.

Notice that the right-hand side of (4.2) contains only three of the four completed power series $\widehat{\Phi}^{5_2,\sigma}_{a/c}$. Just as for the 4_1 knot, this is not because the last one isn't really there, but because our approximate evaluations are not accurate enough at this point to detect the remaining term, which is exponentially small. We will correct this in Section 4.3.

We were able to carry out similar calculations for the $(-2,3,7)$ pretzel knot, though the numerical analysis required here was much more arduous due to the larger number of series involved. Recall that this knot has rank 6, so that $\mathcal{P}_K$ contains seven elements. What makes the calculation feasible at all is that five of these seven elements are real (the Habiro one and the ones corresponding to the real embedding of $\mathbb{Q}(\xi)$ and to all three embeddings of $\mathbb{Q}(\eta)$), so that only one of the seven $\widehat{\Phi}$-functions is exponentially small and hence invisible with optimal truncation. (Actually, the fact that the other terms apart from the geometric one are of the order of 1 is not quite enough: one also has to verify by using the formulas of Section 3.4 and the numerical values of the complex volumes $v(K,\sigma_i)$ that the absolute error made in calculating the exponentially large dominant term $\widehat{\Phi}^{(K,\sigma_1)}(h)$ using optimal truncation is exponentially small.) We find a formula exactly analogous to (4.2), but now with six terms on the right, namely

$$(cX+d)^{-3/2}J^{(-2,3,7)}\left(\frac{aX+b}{cX+d}\right) \overset{?}{\approx} \sum_{\substack{0\le j\le 6\\ j\ne 2}} Q_j^{(-2,3,7)}(X)\widehat{\Phi}_{a/c}^{((-2,3,7),\sigma_j)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right), \tag{4.3}$$

where $j=2$ is omitted for the same reason as in (4.2) (viz., that the corresponding term is too small to see at this stage) and where $Q_1^{(-2,3,7)}(x) = J^{(-2,3,7)}(x)$, and $Q_0^{(-2,3,7)}(x)\equiv 1$, and the four new periodic functions $Q_j^{(-2,3,7)}(x) = \mathcal{Q}_j^{(-2,3,7)}(\mathbf{e}(x))$ take values in $\mathbb{Z}[\mathbf{e}(x)]$ just as before, the first values (for $j\ne 0,2$) being

$\mathrm{Ord}(q)$	1	2	3	4	5	6
$\mathcal{Q}_1^{(-2,3,7)}(q)$	1	1	$-5+6q$	$17-8q$	$-21-27q-5q^2+4q^3$	$-107+108q$
$\mathcal{Q}_3^{(-2,3,7)}(q)$	-4	-12	$-15-10q$	$-16-2q$	$-36-20q-29q^2-24q^3$	$23+14q$
$\mathcal{Q}_4^{(-2,3,7)}(q)$	2	-10	$-16-12q$	$-46q$	$-8-44q-38q^2-48q^3$	$116-24q$
$\mathcal{Q}_5^{(-2,3,7)}(q)$	-2	-6	$-14-6q$	$8-10q$	$32q-4q^2-10q^3$	$-82+122q$
$\mathcal{Q}_6^{(-2,3,7)}(q)$	2	2	$4-8q$	$10-12q$	$-4-36q-44q^2-34q^3$	$136-16q$

Just as with the 5_2 knot, we can verify the Ohtsuki property for these functions to a large number of terms and thus convince ourselves that each one belongs to the Habiro ring, even though in this case we do not know an explicit formula that makes this property manifest.

4.3 Smoothed optimal truncation

We already mentioned at the end of Section 4.1 that it would be natural to expect a third term in (4.1) involving the missing $\widehat{\Phi}$-function $\widehat{\Phi}^{(K,\sigma_2)}(h)$, and the same applies even more strikingly to the two knots discussed in Section 4.2, where we were obliged to omit the σ_2-term in both (4.2) and (4.3) because it would have been absorbed in the error terms of the other $\widehat{\Phi}$'s and hence could not be seen numerically if these values were defined by naive optimal truncation. However, there is a more precise way of turning the divergent series $\Phi(h) = \Phi_\alpha^\sigma(h)$ into functions that are defined up to exponentially rather than merely polynomially small errors, but with a much better exponent than before, by replacing the naive optimal truncation $\Phi(h)^{\mathrm{opt}}$ by a *smoothed* version $\Phi(h)^{\mathrm{sm}}$. The details of the procedure to do this are somewhat complicated and play no role for the story we are telling here, so will be given in detail in a separate publication [45] and described briefly in Section 10.2, the only important point here being that the improvement is sufficiently good, at least for our three standard knots, that we can unambiguously identify the periodic coefficients of the missing $\widehat{\Phi}$-terms.

We start as usual with the knot $K=4_1$ and the series $\Phi(h)=\Phi_0^{(\sigma_1)}(h)$ whose initial terms are given in (1.3). In Section 4.1, we considered $X=100$, $h=\frac{2\pi\mathrm{i}}{X}$ and saw that the number $100^{3/2}\widehat{\Phi}(h)^{\mathrm{opt}}\approx 8.195\times 10^{16}$ had an error of the order of 10^{-12}, which was more than sufficient to identify its difference with $\langle 4_1\rangle_{100}$ unambiguously as $\widehat{\Phi}^{(\sigma_0)}(h)$ but not enough to see a possible contribution from the much smaller $\widehat{\Phi}^{(\sigma_2)}(h)$. If we replace optimal by smooth truncation, then the error in $\widehat{\Phi}(h)$ decreases from (approximately) 10^{-15} to 10^{-44} and the error in $\widehat{\Phi}^{(\sigma_0)}(h)$ from 10^{-15} to 10^{-42}. We can therefore compute the difference of the left- and right-hand sides of (4.1) (for $X=100$, $\gamma=S$) to 42 digits, finding that it vanishes, and since the remaining $\widehat{\Phi}$-value $\widehat{\Phi}^{(\sigma_2)}(h)$ has the much larger order of 10^{-14}, we see that this quantity, if it occurs at all, must have coefficient 0. But when we replace $X=100$ by $100\frac{1}{3}$, we find that the difference $X^{-3/2}J(-1/X)-J\bigl(\frac{1}{3}\bigr)\widehat{\Phi}^{(\sigma_1)}(h)^{\mathrm{sm}}-\widehat{\Phi}^{(\sigma_0)}\bigl(h\bigr)^{\mathrm{sm}}$ no longer vanishes but instead is equal to $\widehat{\Phi}^{(\sigma_2)}\bigl(h\bigr)^{\mathrm{sm}}$ times $-1.732050807568877293527446341\mathrm{i}$, which coincides to this accuracy with $-\sqrt{-3}$. Doing the same for other large values of X with small denominators and other γ, we find that (4.1) with all Φ-values interpreted by smooth rather than optimal truncation can be improved to

$$(cX+d)^{-3/2}J^{(4_1)}\left(\frac{aX+b}{cX+d}\right)\overset{?}{\approx}\sum_{j=0}^{2}Q_j^{(4_1)}(X)\widehat{\Phi}^{(4_1,\sigma_j)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right), \tag{4.4}$$

where, just as for the $(-2,3,7)$ pretzel knot, $Q_0^{(4_1)}(x)=1$, $Q_1^{(4_1)}(X)=J^{(4_1)}(x)$ and $Q_2^{(4_1)}$ is a 1-periodic functions, the notations in each case being a shorthand for $Q_{\sigma_j}^{(4_1)}$. The first few values of the periodic functions $Q_i^{(4_1)}(x)=\mathcal{Q}_i^{(4_1)}(\mathbf{e}(x))$ for $i=1$ and 2 are given by

$\mathrm{Ord}(q)$	1	2	3	4	5	6
$\mathcal{Q}_1^{(4_1)}(q)$	1	5	13	27	$44-4q^2-4q^3$	89
$2\mathcal{Q}_2^{(4_1)}(q)$	0	0	$-2-4q$	$-14q$	$-15-30q-22q^2-8q^3$	$46-92q$

Just as with the functions $Q^{(5_2)}(x)$ and $Q_i^{(-2,3,7)}(x)$ $(i=3,4,5,6)$ found for the 5_2 and $(-2,3,7)$ knots in the previous subsection, the function $Q_2^{(4_1)}$ (whose values we found by the numerical procedure just outlined for all x with denominators up to 200), multiplied by 2, turned out to always belong to $\mathbb{Z}[\mathbf{e}(x)]$ and to satisfy all of the necessary Ohtsuki-type congruences near 0 and 1/2 required for it to be an element of the Habiro ring. In this case, following a tip by Campbell Wheeler, we were actually able to guess a formula that reproduced all of the numerically found values and (after multiplication by 2) was visibly in the Habiro ring, namely the following simple variant of equation (2.6):

$$\mathcal{Q}_2^{(4_1)}(q)=\frac{1}{2}\sum_{n=0}^{\infty}\bigl(q^{n+1}-q^{-n-1}\bigr)\bigl(q^{-1};q^{-1}\bigr)_n(q;q)_n. \tag{4.5}$$

When we recompute the examples of Section 4.2 with smooth rather than optimal truncation, the situation is exactly similar and we are able to add a $\widehat{\Phi}_\alpha^{(K,\sigma_2)}(h)$ term to the right-hand sides of both (4.2) and (4.3), obtaining for both knots a conjectural approximate formula of the form

$$(cx+d)^{-3/2}J^{(K)}\left(\frac{aX+b}{cX+d}\right)\overset{?}{\approx}\sum_{\sigma\in\mathcal{P}_K}Q_\sigma^{(K)}(X)\widehat{\Phi}_{a/c}^{(K,\sigma)}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right), \tag{4.6}$$

where $Q_{\sigma_0}^{(K)}(x)=1$ and $Q_{\sigma_1}^{(K)}(x)=J^{(K)}(x)$. In Sections 7.1 and 9, we give more information about these numbers including a formula for $Q_{\sigma_2}^{(5_2)}(x)$ as an element of the Habiro ring.

4.4 Strengthening the generalized quantum modularity conjecture

So far we have generalized the original quantum modularity conjecture (1.6) in two very different ways: in Section 3.1, we extended it from the Kashaev invariant $J = J^{(\sigma_0)}$ to the functions defined in (3.1), and in the last three subsections we refined it by adding additional terms of lower order to the right-hand side. Not surprisingly, these two can be combined, but with some new aspects.

If we repeat the calculations described in the previous subsection (using smooth truncation for all the Φ-series occurring) but with the function $J^{(K)} = J^{(K,\sigma_0)}$ replaced by the function defined in (3.1) with $\sigma \neq \sigma_0$, then instead of (4.6) we find

$$\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)} J^{(K,\sigma)}\left(\frac{aX+b}{cX+d}\right) \overset{?}{\approx} \sum_{\sigma'\in\mathcal{P}_K^{\mathrm{red}}} J^{(K,\sigma,\sigma')}(X)\widehat{\Phi}^{(K,\sigma')}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right), \tag{4.7}$$

where $J^{(K,\sigma,\sigma')}$ are 1-periodic functions on $\mathbb{Q}$ with $J^{(K,\sigma,\sigma_1)}(x) = J^{(K,\sigma)}(x)$ (cf. (3.6)). There are, however, three main differences with (4.6). The first is that the automorphy factor $(cX+d)^{3/2}$ is replaced for $\sigma \neq \sigma_0$ by the factor $\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}$ involving the σ-th volume $v(\sigma)$ (which is zero for $\sigma = \sigma_0$). The second is that the Habiro power series $\widehat{\Phi}^{(K,\sigma_0)}_{a/c}$, which in (4.6) had the coefficient 1, now does not occur at all. The third is that the new functions $J^{(K,\sigma,\sigma')}(x)$ are now no longer elements of the Habiro ring when considered as functions of $q = \mathbf{e}(x)$, as was the case for the functions $J^{(K,\sigma_0,\sigma')}(x) = Q^{(K)}_{\sigma'}(x)$. But they are still $\overline{\mathbb{Q}}$-valued and have various “Habiro-like” properties, including the following:

- $J^{(K,\sigma,\sigma')}(x)$ for $x \in \mathbb{Q}/\mathbb{Z}$ is the constant term of a power series $\Phi^{(K,\sigma,\sigma')}_x(h)$ lying in the same space as the power series $\Phi^{(K,\sigma)}_x(h)$, as discussed briefly after (1.7) and in more detail in Section 9, i.e., it belongs to $\mu\delta^{-1/2}\sqrt[m]{\varepsilon}F_\sigma(\zeta_m)[[h]]$ with the same root of unity μ, the same element δ_σ of $F_\sigma^\times$, the same set S of primes of F_σ (independent of x) and the same S-unit $\varepsilon = \varepsilon_x$ of $F_\sigma(\zeta_m)$ as for $\Phi^{(K,\sigma)}_x(h)$.
- one can interpret $\Phi^{(K,\sigma,\sigma')}_x(h)$ as $\mathcal{Q}\big(\mathbf{e}(x)\mathrm{e}^{-h}\big)$ where $\mathcal{Q}$ is an element of a Habiro ring $\mathcal{H}_{F_\sigma}$ generalizing the ordinary Habiro ring $\mathcal{H} = \mathcal{H}_{\mathbb{Q}}$ whose definition and arithmetic properties will be discussed in a planned joint paper with Peter Scholze and Campbell Wheeler [38]. In particular, for primes p that split completely in F_σ, there are congruence properties modulo p relating, for instance, the first p coefficients of $\mathcal{Q}\big(\mathrm{e}^{-h}\big)$ to the value of $\mathcal{Q}(\zeta_p)$.

We can write equation (4.7) more uniformly by allowing the case of $\sigma = \sigma_0$, but remembering that there is then an automorphy factor $(cX+d)^{-3/2}$ that is not present for $\sigma \in \mathcal{P}_K$. Then all of the formulas found so far can be collected into a single conjectural formula

$$\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}(cX+d)^{-\kappa(\sigma)} J^{(K,\sigma)}\left(\frac{aX+b}{cX+d}\right) \overset{?}{\approx} \sum_{\sigma'\in\mathcal{P}_K} J^{(K,\sigma,\sigma')}(X)\widehat{\Phi}^{(K,\sigma')}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right) \tag{4.8}$$

valid for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and all $\sigma \in \mathcal{P}_K$ for $X \to \infty$ with bounded denominator, where $J^{(K,\sigma,\sigma')}$ are periodic functions belonging to a generalized Habiro ring and satisfying $J^{(K,\sigma,\sigma_0)} = \delta_{\sigma,\sigma_0}$ and $J^{(K,\sigma_0,\sigma)} = Q^{(K)}_\sigma$ (as introduced in the previous section), with the weight κ_σ and the multiplier $\lambda_\gamma(X)$ defined as in (3.5) and (3.7).

As a concrete illustration of the refined QMC (4.8) we take once again the figure 8 knot with $\sigma = \sigma_1$. Here, (4.8) involves three terms $\sigma' = \sigma_0, \sigma_1, \sigma_2$, with two of the three coefficients already known (the first vanishes and the second is $J^{(4_1,\sigma_1)}(X)$) but with the third one being a new periodic function on $\mathbb{Q}$ given explicitly by

$$J^{(4_1,\sigma_1,\sigma_2)}(x) = \frac{\mathrm{i}}{2\sqrt[4]{3}\sqrt{c}} \sum_{Z^c=\zeta_6} \big(Zq - Z^{-1}q^{-1}\big)\prod_{j=1}^{c}\big|1-q^jZ\big|^{2j/c},$$
$$c = \mathrm{den}(x), \quad q = \mathbf{e}(x), \tag{4.9}$$

which is related to the function given in (3.2) in exactly the same way as $Q^{(4_1)}_{\sigma_2}(x)$ and $J^{(4_1)}(x)$ are related by (4.5) and (2.6). Likewise, the refined QMC (4.8) for 4_1 and for $\sigma = \sigma_2$ leads to the periodic function $J^{(4_1,\sigma_2,\sigma_2)}(x) = -\mathrm{i}J^{(4_1,\sigma_1,\sigma_2)}(-x)$. We also find the *bilinear identity*

$$J^{(4_1,\sigma_1,\sigma_1)}(x)J^{(4_1,\sigma_2,\sigma_2)}(x) - J^{(4_1,\sigma_1,\sigma_2)}(x)J^{(4_1,\sigma_2,\sigma_1)}(x) = 1 \tag{4.10}$$

for all $x \in \mathbb{Q}/\mathbb{Z}$. $\big($Note also that $J^{(\sigma_1,\sigma_1)}(x) = J^{(\sigma_1)}(x)$ and $J^{(\sigma_2,\sigma_1)}(x) = J^{(\sigma_2)}(x)$.$\big)$ This identity will be generalized to all knots in Section 5.

4.5 The refined quantum modularity conjecture

The refinement of the quantum modularity conjecture that we have obtained so far, equation (4.8), has two noteworthy aspects. One is that, although we find new collections of "Habiro-like" functions $J^{(\sigma,\sigma')}$ for the asymptotic expansion as $X \to \infty$ of the functions $J^{(\sigma_0,\sigma)}(\gamma X)$ for different $\sigma \in \mathcal{P}_K$ (here we continue the practice of omitting the knot from all notations when it is not varying), these arise as coefficients of the *same* completed formal power series $\widehat{\Phi}^{(\sigma')}(h)$ as we found for the initial Galois-extended Kashaev invariant $J^{(\sigma_0)}$. The other is that among the new coefficients $J^{(K,\sigma,\sigma')}$, the subset corresponding to $\sigma' = \sigma_1$ coincides precisely with the set of functions $J^{(\sigma)}$ whose asymptotic behavior near fixed rational points is being studied. It is therefore natural to ask whether the functions $J^{(K,\sigma,\sigma')}$ for σ' different from σ_1 also have a quantum modularity property, and if so, what new power series are involved. In this final subsection, we will study both questions and give our (nearly) final version of the QMC.

As usual, we look first at the 4_1 knot. Here, as well as the three periodic functions $J^{(\sigma_0,\sigma_1)}(x) := J(x)$ and $J^{(\sigma_0,\sigma_1)}(x) := J^{(\sigma_1)}$ and $J^{(\sigma_0,\sigma_2)} := J^{(\sigma_2)}$ we had studied earlier, we found two new periodic functions $J^{(\sigma_1,\sigma_2)}(x)$ and $J^{(\sigma_2,\sigma_2)}(x)$, given explicitly by formulas (4.9) and related to the others by (4.10). If we look numerically at the asymptotics of both functions with $x = -1/X$ for $X \to \infty$ with bounded denominator, we find

$$X^{-3/2}J^{(\sigma_0,\sigma_2)}\left(-\frac{1}{X}\right) \sim J^{(\sigma_0,\sigma_1)}(X)\widehat{\Psi}^{(1)}\left(\frac{2\pi\mathrm{i}}{X}\right),$$
$$\mathrm{e}^{-v(\sigma_1)\lambda_S(X)}J^{(\sigma_1,\sigma_2)}\left(-\frac{1}{X}\right) \sim J^{(\sigma_1,\sigma_1)}(X)\widehat{\Psi}^{(1)}\left(\frac{2\pi\mathrm{i}}{X}\right)$$

with the *same* completed power series

$$\widehat{\Psi}^{(1)}(h) = \mathrm{e}^{\mathbf{V}(4_1)/h}\Psi^{(1)}(h), \qquad \Psi^{(1)}(h) = \frac{\mathrm{i}\sqrt[4]{3}}{2}\left(1 - \frac{37}{72\sqrt{-3}}h - \frac{1511}{2(72\sqrt{-3})^2}h^2 + \cdots\right)$$

in both cases. Based on the analogy with the asymptotics of the functions $J^{(\sigma_1,\sigma_1)}(-1/X)$ as given in (4.4), we would now expect the more accurate approximations

$$X^{-3/2}J^{(\sigma_0,\sigma_2)}\left(-\frac{1}{X}\right) \approx \widehat{\Psi}^{(0)}\left(\frac{2\pi\mathrm{i}}{X}\right) + J^{(\sigma_0,\sigma_1)}(X)\widehat{\Psi}^{(1)}\left(\frac{2\pi\mathrm{i}}{X}\right) + J^{(\sigma_0,\sigma_2)}(X)\widehat{\Psi}^{(2)}\left(\frac{2\pi\mathrm{i}}{X}\right),$$
$$\mathrm{e}^{-v(\sigma_1)\lambda_S(X)}J^{(\sigma_1,\sigma_2)}\left(-\frac{1}{X}\right) \approx J^{(\sigma_1,\sigma_1)}(X)\widehat{\Psi}^{(1)}\left(\frac{2\pi\mathrm{i}}{X}\right) + J^{(\sigma_1,\sigma_2)}(X)\widehat{\Psi}^{(2)}\left(\frac{2\pi\mathrm{i}}{X}\right),$$
$$\mathrm{e}^{-v(\sigma_2)\lambda_S(X)}J^{(\sigma_2,\sigma_2)}\left(-\frac{1}{X}\right) \approx J^{(\sigma_2,\sigma_1)}(X)\widehat{\Psi}^{(1)}\left(\frac{2\pi\mathrm{i}}{X}\right) + J^{(\sigma_2,\sigma_2)}(X)\widehat{\Psi}^{(2)}\left(\frac{2\pi\mathrm{i}}{X}\right), \tag{4.11}$$

where $\widehat{\Psi}^{(2)}(h)$ is the completed series $\mathrm{e}^{-\mathbf{V}(4_1)/h}\Psi^{(2)}(h)$ with $\Psi^{(2)}(h) = -\mathrm{i}\Psi^{(1)}(-h)$ and where $\widehat{\Psi}^{(0)}(h)$ is the completed series $(h/2\pi\mathrm{i})^{3/2}\Phi^{(0)}(h)$ (cf. (2.8)) with

$$\Phi^{(0)}(h) = -h + \frac{11}{6}h^3 - \frac{1261}{120}h^5 + \frac{611771}{5040}h^7 - \cdots$$

the power series in $h\mathbb{Q}\big[\big[h^2\big]\big]$ obtained by replacing $q = \mathbf{e}(X)$ by $q = \mathrm{e}^{-h}$ in formula (4.5), as well of course as similar formulas for $J^{(\sigma,\sigma')}(\gamma X)$ for other matrices $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}_2(\mathbb{Z})$ with the completed power series $\widehat{\Psi}^{(j)}(h)$ replaced by suitable new completed power series

$$\widehat{\Psi}^{(j)}_{a/c}(h) = \mathrm{e}^{\mathbf{V}(\sigma_j)/ch}(h/2\pi\mathrm{i})^{\kappa(\sigma_j)}\Psi^{(j)}_{a/c}(h)$$

but with the same periodic coefficients.

To test (4.11) or its generalizations to other $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ directly we would need to find many terms of the power series $\Psi^{(j)}_\alpha(h)$ and carry out the smoothed optimal truncation as described earlier in this section, because the different exponential growths of their completions would mean that the contributions with $j \neq 1$ would not be numerically visible at the level of mere formal power series. This could be done, but an easier test of the prediction is to take linear combinations of the first two or last two lines in (4.11) to eliminate the dominant $\widehat{\Psi}^{(1)}$-term. This (together with (4.10)) produces the two new asymptotic predictions

$$\begin{aligned}
&X^{-3/2}J^{(\sigma_1,\sigma_1)}(X)J^{(\sigma_0,\sigma_2)}\left(-\frac{1}{X}\right) - \mathrm{e}^{-v(\sigma_1)\lambda_S(X)}J^{(\sigma_0,\sigma_1)}(X)J^{(\sigma_1,\sigma_2)}\left(-\frac{1}{X}\right)\\
&\qquad \approx J^{(\sigma_1,\sigma_1)}(X)\widehat{\Psi}^{(0)}\left(\frac{2\pi\mathrm{i}}{X}\right),\\
&\mathrm{e}^{-v(\sigma_1)\lambda_S(X)}J^{(\sigma_2,\sigma_1)}(X)J^{(\sigma_1,\sigma_2)}\left(-\frac{1}{X}\right) - \mathrm{e}^{-v(\sigma_2)\lambda_S(X)}J^{(\sigma_1,\sigma_1)}(X)J^{(\sigma_2,\sigma_2)}\left(-\frac{1}{X}\right)\\
&\qquad \approx \widehat{\Psi}^{(2)}\left(\frac{2\pi\mathrm{i}}{X}\right),
\end{aligned}$$

both of which can be tested directly since they do not involve functions of different orders of growth on the right, and both of which we confirmed numerically to very high precision. We omit the details, having given more than enough descriptions of analogous numerical calculations in this section already.

Generalizing the above discussion to other knots, we find as our nearly-final version of the QMC the asymptotic statement

$$\begin{aligned}
&(cX+d)^{-\kappa(\sigma)}\mathrm{e}^{-v(\sigma)\lambda_\gamma(X)}J^{(K,\sigma,\sigma')}\left(\frac{aX+b}{cX+d}\right)\\
&\qquad \overset{?}{\approx} \sum_{\sigma''\in\mathcal{P}_K} J^{(K,\sigma,\sigma'')}(X)\widehat{\Phi}^{(K,\sigma'',\sigma')}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right)
\end{aligned}$$

for $X \in \mathbb{Q}$ tending to infinity with bounded denominator and for every $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ with $c > 0$, where the functions $J^{(K,\sigma,\sigma')}$ are the "Habiro-like" functions that we found in Section 4.4, given as the constant terms of certain power series $\Phi^{(K,\sigma,\sigma')}_\alpha(h) \in \overline{\mathbb{Q}}[[h]]$, and where $\widehat{\Phi}^{(K,\sigma,\sigma')}_\alpha(h)$ are the completions defined by

$$\widehat{\Phi}^{(K,\sigma,\sigma')}_\alpha(h) = (ch/2\pi\mathrm{i})^{\kappa(\sigma)}\mathrm{e}^{\mathbf{V}(\sigma)/c^2h}\Phi^{(K,\sigma,\sigma')}_\alpha(h), \qquad \sigma,\sigma' \in \mathcal{P}_K.$$

To get the final version, we upgrade this statement about constant terms to a statement about the full (completed) power series in the same way as we did in 3.2, obtaining:

Refined quantum modularity conjecture (RQMC): *For fixed* $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ *with* $c > 0$*, we have*

$$\widehat{\Phi}^{(K,\sigma,\sigma')}_{\gamma X}(h^*) \overset{?}{\approx} (cx+d)^{\kappa(\sigma)} \sum_{\sigma''\in\mathcal{P}_K} \widehat{\Phi}^{(K,\sigma,\sigma'')}_X(h)\widehat{\Phi}^{(K,\sigma'',\sigma')}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cx+d)}\right)$$

to all orders in $1/X$ *as* $X \in \mathbb{Q}$ *tending to* $+\infty$ *with bounded denominator, where*

$$x = X - \hbar \text{ and } h^* = h/(cx+d)(cX+d).$$

We end this section by observing that the two versions of the refined quantum modularity conjecture that we just stated can both be written more succinctly in matrix form as

$$\mathbf{J}^{(K)}(\gamma X) \approx \widetilde{\mathbf{j}}_\gamma(X)\mathbf{J}^{(K)}(X)\widehat{\mathbf{\Phi}}^{(K)}_{a/c}\left(\frac{2\pi \mathrm{i}}{c(cX+d)}\right) \tag{4.12}$$

and

$$\widehat{\mathbf{\Phi}}^{(K)}_{\gamma X}(h^*) \overset{?}{\approx} \mathbf{j}_\gamma(x)\widehat{\mathbf{\Phi}}^{(K)}_{X}(h)\widehat{\mathbf{\Phi}}^{(K)}_{a/c}\left(\frac{2\pi \mathrm{i}}{c(cx+d)}\right) \tag{4.13}$$

as $X \to \infty$ with bounded denominator for a fixed knot K and element $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, where $\mathbf{J}^{(K)}$ and $\widehat{\mathbf{\Phi}}^{(K)}$ denote the matrices of Habiro-like functions and completed formal power series, with columns and rows indexed by $\mathcal{P}_K$, with entries $J^{(K,\sigma,\sigma')}(x)$ and $\Phi^{(K,\sigma,\sigma')}_{a/c}(h)$, respectively, and where $\mathbf{j}$ and $\widetilde{\mathbf{j}}$ are the matrix-valued automorphy factors defined by

$$\mathbf{j}_\gamma(x) = \mathbf{diag}\big(|cx+d|^{\kappa(\sigma)}\big), \qquad \widetilde{\mathbf{j}}_\gamma(x) = \mathbf{diag}\big(\mathrm{e}^{v(\sigma)\lambda_\gamma(x)}|cx+d|^{\kappa(\sigma)}\big), \tag{4.14}$$

the second of which is the "tweaked" version of the first. Note that both of these factors are unchanged if we replace γ by $-\gamma$, and hence are actually automorphy factors on $\mathrm{PSL}_2(\mathbb{Z})$. Also, from the fact that λ is an additive cocycle (see Lemma 3.1) we deduce that both $\mathbf{j}$ and $\widetilde{\mathbf{j}}$ are matrix-valued cocycles on $\mathrm{PSL}_2(\mathbb{Z})$, meaning that they satisfy

$$\mathbf{j}_{\gamma\gamma'}(x) = \mathbf{j}_{\gamma'}(x)\mathbf{j}_\gamma(\gamma' x), \qquad \widetilde{\mathbf{j}}_{\gamma\gamma'}(x) = \widetilde{\mathbf{j}}_{\gamma'}(x)\widetilde{\mathbf{j}}_\gamma(\gamma' x), \tag{4.15}$$

for all $\gamma, \gamma' \in \mathrm{PSL}_2(\mathbb{Z})$. This will be important in the next section.

5 The matrix-valued cocycle associated to a knot

Let us define, for a fixed knot K (suppressed from the notation as usual), matrix $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and number $x \in \mathbb{Q} \smallsetminus \{\gamma^{-1}(\infty)\}$, an $(r+1)\times(r+1)$ matrix $W_\gamma(x)$ by

$$W_\gamma(x) = \mathbf{J}(\gamma x)^{-1}\widetilde{\mathbf{j}}_\gamma(x)\mathbf{J}(x), \tag{5.1}$$

where $\widetilde{\mathbf{j}}$ is the automorphy factor defined in (4.14). (This formula makes sense because the matrix $\mathbf{J}$ is conjecturally invertible, and even unimodular, as discussed in (5.8) below.) This function has remarkable properties. On the one hand, the refined quantum modularity conjecture as presented above can now be rewritten as the asymptotic statement

$$W_\gamma(X) \approx \widehat{\mathbf{\Phi}}_{a/c}\left(\frac{2\pi \mathrm{i}}{c(cX+d)}\right)^{-1} \tag{5.2}$$

for $X \in \mathbb{Q}$ tending to infinity with bounded denominator. In particular, unlike the completely discontinuous function $\mathbf{J}(x)$ in terms of which it is defined, $W_\gamma(X)$ has an asymptotic behavior at infinity that depends only on X as a real number and not on its numerator and denominator separately, and in Section 5.2 we will present very strong evidence that this is true not only asymptotically at infinity, but also for finite values of the argument, so that $W_\gamma(x)$ becomes a smooth (and in fact real analytic) function of its argument away from the singularity at $x = \gamma^{-1}(\infty)$. On the other hand, the cocycle property (4.15) of $\widetilde{\mathbf{j}}$ immediately implies that the function $\gamma \mapsto W_\gamma(\cdot)$ is a cocycle for the group $\mathrm{PSL}_2(\mathbb{Z})$ acting on the multiplicative group of almost-everywhere-defined invertible matrix-valued functions on $\mathbb{P}^1(\mathbb{Q})$, meaning that it satisfies

$$W_{\gamma\gamma'}(x) = W_\gamma(\gamma' x) W_{\gamma'}(x) \tag{5.3}$$

for all γ and γ' in $\mathrm{PSL}_2(\mathbb{Z})$. But this cocycle property then immediately extends by continuity to imply that W_γ on $\mathbb{R}$ is also a $\mathrm{PSL}_2(\mathbb{Z})$-cocycle, but now in the space of piecewise smooth matrix-valued functions on $\mathbb{P}^1(\mathbb{R})$. We can then use the smoothness to define a canonical lift of each of the formal power series $\Phi_\alpha^{(\sigma,\sigma')}(h)$ to an actual function of h, simply by requiring (5.2) to be an exact rather than merely an asymptotic equality.

These various properties will be described in detail in this section. The first subsection treats the elementary properties (behavior under complex conjugation, determinant, and inverse) of the matrices $\mathbf{J}(x)$ and $W_\gamma(x)$. The discussion of the smoothness properties and the lifting of the perturbative series $\Phi_\alpha^{(\sigma,\sigma')}(h)$ to well-defined functions of h will be given in Section 5.2, while the brief final subsection treats the expected equality between the cocycle $W_\gamma(x)$ and the cocycle constructed in the companion paper [44] using state integrals, which gives the real explanation for its smoothness and even analyticity.

5.1 The Habiro-like matrix and the perturbative matrix

In Section 4, we saw how successive refinements of the original quantum modularity conjecture (1.6) led to a whole matrix $\mathbf{J}^{(K)}(x)$ of generalized Kashaev invariants and to a collection of matrices $\mathbf{\Phi}_\alpha^{(K)}(h)$ of formal power series having $\mathbf{J}^{(K)}(\alpha)$ as their constant term. The existence of these new matrices and the description of their properties is the main content of this paper. We emphasize that, although the refined QMC which led to the definition of these matrices and to the means of finding them numerically is still conjectural, the matrices themselves are well-defined, at least in terms of a chosen triangulation: Their first columns are trivial (a one followed by r zeros). Their second columns were defined in Section 2 in terms of the original Kashaev invariant and of the perturbative series defined in [14, 15]. The further columns of the $\mathbf{\Phi}$-matrix can also be given by Gauss-type integrals like those in [14, 15], and in principle one could also always find explicit formulas for the $\mathbf{J}$ matrix, as has been written out for the 4_1 knot in detail in Section 4.3 (equations (4.5) and (4.9)) and will be discussed more generally in Sections 7.1–7.3 in the context of q-holonomic systems, with full details for the knot 5_2. In general, it is not known that these quantities are topological invariants, since their definitions depend a priori on the choice of an ideal triangulation and are believed but not proven to be invariant under Pachner moves. But we expect this invariance to be true, and in any case the new matrices are completely computable in practice, as we seen, and have extremely interesting properties. In this subsection, we look at the properties that are directly visible, and in the following one at the deeper properties of the associated cocycle W.

Extension property. From their very definitions, the matrices $\mathbf{J}$ and $\mathbf{\Phi}$ (now omitting the knot from the notation) both have a $(1+r)\times(1+r)$ block triangular form

$$\mathbf{J}(x) = \begin{pmatrix} 1 & Q(x) \\ 0 & \mathbf{J}^{\mathrm{red}}(x) \end{pmatrix}, \qquad \mathbf{\Phi}_\alpha(h) = \begin{pmatrix} 1 & \mathcal{Q}\big(\mathbf{e}(\alpha)\mathrm{e}^{-h}\big) \\ 0 & \mathbf{\Phi}_\alpha^{\mathrm{red}}(h) \end{pmatrix}, \tag{5.4}$$

where $Q(x) = \big(Q^{(\sigma)}(x)\big)_{\sigma\in\mathcal{P}_K^{\mathrm{red}}}$ is the vector of length r whose entries are given by the periodic functions found in Section 4, $\mathcal{Q}(q) = \big(\mathcal{Q}^{(\sigma)}(q)\big)_{\sigma\in\mathcal{P}_K^{\mathrm{red}}}$ is the corresponding function in terms of $q = \mathbf{e}(x)$ (which we believe to be elements of the Habiro ring $\mathcal{H}\otimes\mathbb{Q}$ and therefore to be defined not just at roots of unity, but also at infinitesimal deformations of roots of unity), and $\mathbf{J}^{\mathrm{red}}$ and $\mathbf{\Phi}_\alpha^{\mathrm{red}}$ are $r\times r$ matrices with rows and columns indexed by the elements of $\mathcal{P}_K^{\mathrm{red}}$. We are mainly interested in the larger matrices, but we will sometimes want to consider the "reduced" matrices separately because they sometimes occur separately (notably in Section 3.3, where only $\mathcal{P}_K^{\mathrm{red}}$ occurs in (3.14), and in the statements below about the inverse matrices of $\mathbf{J}^{\mathrm{red}}$ and $\mathbf{\Phi}_\alpha^{\mathrm{red}}$). This block triangular property, trivial though it is, should have a deeper meaning as the statement that the $(r+1)$-dimensional objects associated to a knot K (specifically, the q-holonomic modules that will be the subject of Section 7), with a basis parametrized by the set of representations $\mathcal{P}_K$, are in fact extensions of r-dimensional objects with a basis parametrized by $\mathcal{P}_K^{\mathrm{red}}$ by something one-dimensional.

Behavior under complex conjugation. The next point is the following compatibility with complex conjugation, namely

$$\overline{\mathbf{J}^{\sigma,\sigma'}(-x)} = \begin{cases} \mathbf{J}^{\sigma,\sigma'}(x) & \text{if } \sigma \text{ is real,} \\ -\mathrm{i}\mathbf{J}^{\overline{\sigma},\sigma'}(x) & \text{if } \sigma \text{ is not real,} \end{cases} \tag{5.5}$$

where "σ real" means $\sigma = \overline{\sigma}$. In matrix form this becomes

$$\overline{\mathbf{J}(-x)} = B\mathbf{J}(x), \tag{5.6}$$

where B is the unimodular symmetric unitary matrix with $B^{(\sigma,\sigma')}$ equal to 1 if $\sigma' = \sigma = \overline{\sigma}$, to $-\mathrm{i}$ if $\sigma' = \overline{\sigma} \neq \sigma$, and to 0 if $\sigma' \neq \overline{\sigma}$. The symmetry (5.5) has as the special case $(\sigma,\sigma') = (\sigma_0,\sigma_1)$ the behavior $\overline{J(-x)} = J(x)$ of the Kashaev invariant for rational numbers x under complex conjugation, which holds because the colored Jones polynomials have real (even integer) coefficients or alternatively because J is an element of the Habiro ring. The same argument applies conjecturally to all $J^{(\sigma_0,\sigma)}$, since they also belong to the Habiro ring, and if we use the full RQMC it also suffices to establish the general case. Actually, equation (5.6) can be strengthened to

$$\overline{\mathbf{\Phi}_{-\alpha}(\overline{h})} = B\mathbf{\Phi}_\alpha(h), \tag{5.7}$$

which specializes at $h = 0$ to (5.6). We remark that equation (5.7) remains true if we replace both $\mathbf{\Phi}$'s by their completions $\widehat{\mathbf{\Phi}}$ (except for the top rows, which differ by a factor of i), because $\overline{V(\sigma)} = V(\overline{\sigma})$ for all $\sigma \in \mathcal{P}_K^{\mathrm{red}}$.

Unimodularity. The next statement, generalizing equation (4.10), is that the matrices $\mathbf{J}$ and even $\mathbf{\Phi}^{\mathrm{red}}$, are experimentally found to be unimodular. More precisely, this is definitely true for the 4_1 and 5_2 knots, for which we have closed formulas for all of the entries of the Habiro-like matrices and can compute numerically; for other knots, we are convinced, and willing to conjecture, that the determinant is ± 1, but we have no really convincing reason except aesthetics that it should always be $+1$. The unimodularity implies in particular that the $\mathbf{J}$-matrices are always invertible, a fact that is of course crucial even to define the cocycle W in (5.1). Notice that it is compatible with (5.6) and (5.7), since the matrix B is unimodular.

Inverse/Unitarity. The final property that we want to mention, again only conjectural, is more surprising. This is that the inverse of $\mathbf{J}^{\mathrm{red}}$ (but not of the full matrix $\mathbf{J}$, for which we have no corresponding guess) can be given explicitly by the formula

$$\mathbf{J}^{\mathrm{red}}(x)^t \mathbf{J}^{\mathrm{red}}(-x) = \overline{B^{\mathrm{red}}}, \tag{5.8}$$

where we have set $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & B^{\mathrm{red}} \end{smallmatrix}\right)$. In fact, even this statement can be strengthened, namely to

$$\mathbf{\Phi}_\alpha^{\mathrm{red}}(h)^t \mathbf{\Phi}_{-\alpha}^{\mathrm{red}}(-h) = \overline{B^{\mathrm{red}}}, \tag{5.9}$$

which specializes to (5.8) when we set $h = 0$. In view of (5.6), the first of these equations can be rewritten as

$$\overline{\mathbf{J}^{\mathrm{red}}(x)^t} B^{\mathrm{red}} \mathbf{J}^{\mathrm{red}}(x) = B^{\mathrm{red}},$$

which we see as a kind of unitarity or rather sesqui-unitarity, since if B^{red} were the identity matrix they would simply say that the matrices $\mathbf{J}^{\mathrm{red}}(x)$ and $\mathbf{\Phi}^{\mathrm{red}}_{\alpha}(h)$ are unitary. Note that equation (5.9) remains true also if we replace Φ by $\widehat{\Phi}$, since the volume factors cancel, and also that the very special case $\sigma = \sigma' = \sigma_1$, for which the right-hand side of (5.9) vanishes, is just the quadratic relation (3.14) that was discussed in Section 3.3. It is also worth remarking explicitly that the expression on the left of (5.9) is a priori an element of $\mathbb{Q}(\zeta_c)[[h]]$ ($c = \mathrm{den}(\alpha)$), at least if the predicted algebraic properties of the power series $\Phi^{(\sigma,\sigma')}_{\alpha}$ as discussed in Section 9.1 are true, because the extra factors (root of unity and c-th root of an S-unit in $F_{\sigma''}(\zeta_c)$) cancel in the products $\Phi^{(\sigma'',\sigma)}(h)\Phi^{(\sigma'',\sigma')}(-h)$ and because the sum over σ'' implicit in the matrix multiplication gets us from $F_{\sigma''}(\zeta_c)$ down to $\mathbb{Q}(\zeta_c)$. Finally, we should mention that equation (5.8) also gives us a formula for the inverse of the full matrix $\mathbf{J}(x)$, because of the block triangular form of the latter as given in (5.4), namely

$$\mathbf{J}(x)^{-1} = \begin{pmatrix} 1 & -Q(x)B^{\mathrm{red}}\mathbf{J}^{\mathrm{red}}(-x)^t \\ 0 & B^{\mathrm{red}}\mathbf{J}^{\mathrm{red}}(-x)^t \end{pmatrix},$$

in which the elements of the top row are bilinear in the entries of $\mathbf{J}(x)$ and $\mathbf{J}(-x)$ rather than merely linear as is the case for the other rows.

However, the real interest to us of the final point above is not just that there are explicit formulas for the inverses of the matrices $\mathbf{J}^{\mathrm{red}}$ (or even $\mathbf{J}$) and $\mathbf{\Phi}$, but above all that the inverse of $\mathbf{J}^{\mathrm{red}}$ is expressed *linearly* (more correctly, sesquilinearly) in terms of the entries of the matrix itself. This means in particular that the entries of the cocycle W^{red} (the bottom right $r \times r$ block of W) are expressed *bilinearly* in terms of those of $\mathbf{J}^{\mathrm{red}}$. This remark will come into its own in the sequel [44], where this reduced cocycle will arise in a completely different way as a bilinear combination in functions of $q = \mathrm{e}^{2\pi\tau}$ and $\widetilde{q} = \mathrm{e}^{-2\pi/\tau}$ as a consequence of the factorization of state integrals.

We end this subsection by listing the properties of the function W_γ that it inherits by virtue of its definition (5.1) from the corresponding properties of $\mathbf{J}$ listed above. These will become important in the next subsection, when we extend W_γ from $\mathbb{Q}$ to $\mathbb{R}$.

The "extension property" is immediate: the matrix $W_\gamma(x)$ has the block triangular form $\left(\begin{smallmatrix} 1 & 0 \\ 0 & W^{\mathrm{red}}_\gamma(x) \end{smallmatrix}\right)$ for an $r \times r$ "reduced" matrix $W^{\mathrm{red}}_\gamma(x)$ which is again a cocycle. The complex conjugation property for W takes the form

$$\overline{W_\gamma(x)} = W_{\varepsilon\gamma\varepsilon}(-x) \tag{5.10}$$

for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $x \in \mathbb{Q}$, where $\varepsilon = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. (Note that $\varepsilon\gamma\varepsilon \in \mathrm{SL}_2(\mathbb{Z})$.) This is a consequence of the following short calculation using (5.6) and the easy conjugation behavior of the automorphy factor $\widetilde{\mathbf{j}}_\gamma(x)$:

$$\overline{W_\gamma(x)} = \overline{\mathbf{J}(x)^{-1}\widetilde{\mathbf{j}}_\gamma(-x)\mathbf{J}(\gamma(-x))} = \mathbf{J}(x)^{-1}B^{-1}\overline{\widetilde{\mathbf{j}}_\gamma(-x)}B\mathbf{J}(-\gamma(x)) = W_{\varepsilon\gamma\varepsilon}(-x).$$

A nice consequence of (5.10) is that we can now extend equation (5.2), which described the asymptotics of $W_\gamma(X)$ as X tends to infinity with bounded denominator on the assumption of the RQMC, to give the corresponding asymptotic behavior of $W_\gamma(X)$ also as $X \to -\infty$:

$$W_\gamma(X) \approx B\widehat{\mathbf{\Phi}}_{a/c}\left(\frac{2\pi\mathrm{i}}{c(cX+d)}\right)^{-1}, \qquad X \to -\infty. \tag{5.11}$$

The third property is that the determinant of $W_\gamma(x)$ is given by

$$\det W_\gamma(x) = |j(\gamma,x)|^{-3/2},$$

where $j(\gamma,x)$ is defined as $cx+d$ for $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$. This follows from the (conjectural) unimodularity of $\mathbf{J}$, the definition of $\widetilde{\mathbf{j}}_\gamma(x)$, and the fact that $\sum_{\sigma\in\mathcal{P}_K} v(\sigma)$ vanishes ("Galois descent"). Finally, from (5.8) we immediately deduce the corresponding formula for the inverse matrix of $W^{\mathrm{red}}_\gamma(x)$:

$$W^{\mathrm{red}}_\gamma(x)^{-1} = \overline{B^{\mathrm{red}}}\,W^{\mathrm{red}}_{\varepsilon\gamma\varepsilon}(-x)^t.$$

In this connection, we note that (5.1) and (5.8) also imply that

$$W^{\mathrm{red}}_\gamma(x) = \left(\overline{B^{\mathrm{red}}}\right)^{-1}\mathbf{J}^{\mathrm{red}}(-\gamma x)^t\widetilde{\mathbf{j}}^{\mathrm{red}}_\gamma(x)\mathbf{J}^{\mathrm{red}}(x).$$

In other words, W^{red} is bilinear in the entries of the matrices $\mathbf{J}$, an important property that is also shared by the functions defined by state integrals.

5.2 Smoothness

We now come to the really exciting point. The cocycle $W_\gamma(x)$ is defined in terms of the "Habiro-like" matrix $\mathbf{J}$ by (5.1). The entries of $\mathbf{J}^{(4_1)}$, one of which was shown in Figure 1 of the introduction, would all have a "cloudlike" structure like the one seen there. But when one graphs the entries of the matrix $W_\gamma(x)$, they are all smooth! For instance, Figure 4 shows the graphs of the six nontrivial components of the 3×3 matrix $W_S(x)$ for the figure 8 knot (with three of them divided by i to make them real), where $S=\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ as usual, plotted in each case for all rational numbers in $(0,2]$ with denominator at most 40 (so for roughly 1000 data points).

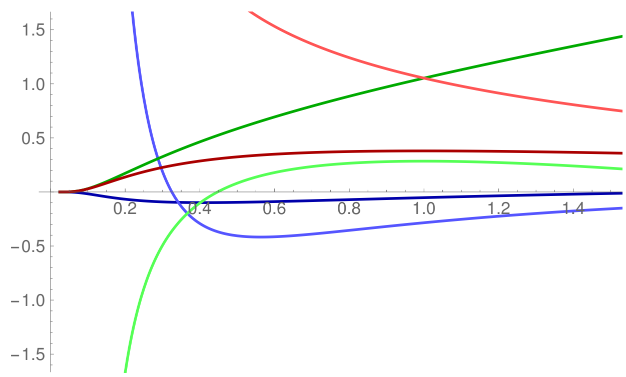

Figure 4. Plots of the six nontrivial entries of the matrix $W_S(x)$ for the 4_1 knot.

We formalize this by stating the following conjecture:

Conjecture 5.1. *The function W_γ defined on $\mathbb{Q}\smallsetminus\{\gamma^{-1}(\infty)\}$ extends to a C^∞ function on $\mathbb{R}\smallsetminus\{\gamma^{-1}(\infty)\}$.*

A first consequence of this is that the cocycle property (5.3), which held for the restriction of W to $\mathbb{P}^1(\mathbb{Q})$ by equation (5.1) and the cocycle property of $\widetilde{\mathbf{j}}$, is then automatically true for the extended function on $\mathbb{P}^1(\mathbb{R})$, even though there is no longer any "coboundary-like formula" of type (5.1). This new cocycle now takes values in the much smaller space of almost-everywhere-defined matrix-valued functions on $\mathbb{P}^1(\mathbb{R})$.

The conjectural smoothness of the function W_γ has another important consequence that was already mentioned in the introduction to this section, namely that we can invert the asymptotic statement (5.2) to get a definition of *exact* matrix-valued functions for all $\alpha\in\mathbb{Q}$ which are smooth on all of $\mathbb{R}$ and whose Taylor expansions (after the "Wick rotation" $h\mapsto\hbar=h/2\pi\mathrm{i}$) agree with the divergent power series $\mathbf{\Phi}_\alpha(h)$. To do this, we simply define a new function $\mathbf{\Phi}^{\mathrm{exact}}_\alpha$ by requiring (5.2) to be an exact rather than just an asymptotic equality, i.e., by defining

$$\mathbf{\Phi}_\alpha(2\pi\mathrm{i}x)^{\mathrm{exact}} := \mathbf{diag}\left(|cx|^{-\kappa(\sigma)}\mathrm{e}^{-v(\sigma)/c^2x}\right)W_\gamma\left(\frac{1}{c^2x}-\frac{d}{c}\right)^{-1} \qquad (5.12)$$

for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ with $\alpha = \gamma(\infty) = a/c$. This should then be an everywhere smooth and almost everywhere analytic function on $\mathbb{R}$ whose Taylor expansion at 0 agrees with the original divergent series. The six functions obtained in this way from the non-trivial elements of $W_S(x)$ for the figure 8 knot as plotted above (and multiplied by suitable powers of i to make them all real) are the ones shown in Figure 2 in the introduction.

Apart from the numerical data, there are at least four reasons why we should expect this smoothness property of the function W_γ to hold:

1. At the simplest level, equation (5.2) tells us that the matrix $W_\gamma(X)$ is at least asymptotically smooth in the limit as $X \to \infty$ through rational numbers with bounded denominator, since it agrees to all orders in $1/X$ with a power series in $1/X$ with coefficients that do not depend on the denominator or other arithmetic properties of X. Stated more visually, if we were to display the components of $W_\gamma(X)$ by plotting their values, for instance, for rational values of X between 1000 and 1001 and with denominators less than 100, then these data points would have to lie on a very smooth curve to very high precision.

2. In fact this same argument can be pushed much further, since by using the cocycle property (5.3) for $x = X$ tending to infinity with bounded denominator we get a description of the asymptotic behavior of $W_g(x)$ in the neighborhood of any rational point, not merely at infinity, and hence an explicit formula for its Taylor expansion at any rational point near which it has a smooth expansion. This will be carried out in Proposition 5.2 below.

3. But the real reason that we expected the smoothness property is much deeper and also predicts (and in some cases leads to a proof of) much more: the entries of the matrix-valued function $W_\gamma(\cdot)$ for a fixed γ extend to functions that are not merely smooth, but actually *analytic*, on $\mathbb{R} \smallsetminus \{-d/c\}$. This comes from the study of q-series associated to a knot and their relation to state integrals, as carried out in the companion paper [44] to this one, and will be discussed in more detail in the final subsection of this section.

4. Finally, once one expects the real-analyticity, one can check it numerically using only the matrices studied in this paper, without any reference to either q-series or state integrals, by computing many Taylor coefficients of W_γ at any rational point using Proposition 5.2 and seeing that they now grow only polynomially rather than factorially. This point too will be discussed in more detail in Section 5.4 below.

In the context just of this paper, where we are considering only functions on $\mathbb{Q}$ and formal power series in h, but not holomorphic functions in the upper or lower half-planes or on cut planes, we cannot justify the statement about analyticity or even continuity of the entries of the matrix $W_\gamma(\cdot)$, i.e., we cannot show that the function $W_\gamma(\cdot)\colon \mathbb{Q} \smallsetminus \{-d/c\} \to \mathbb{C}$ has any natural extension to a matrix-valued function on $\mathbb{R} \smallsetminus \{-d/c\}$. However, as indicated in point **2** above, we can deduce a weaker statement if we assume the RQMC. To explain what this means, we must first discuss the various possible senses in which a function $f\colon \mathbb{Q} \to \mathbb{C}$ can be continuous or differentiable. There are at least three different natural notions. Usually one considers the set of rational numbers with either the discrete topology or else the topology inherited from their embedding into the reals. In the first sense, of course every function from $\mathbb{Q}$ to $\mathbb{C}$ is continuous (i.e., $f(\alpha + \varepsilon_i) \to f(\alpha)$ for any sequence of rational numbers $\alpha + \varepsilon_i$ converging to $\alpha \in \mathbb{Q}$, since any such sequence is eventually constant) and in fact even C^∞ (with the "Taylor expansion" of f at an arbitrary rational point α being just the constant power series $f(\alpha)$). In the second sense, one means that $f(\alpha + \varepsilon_i) \to f(\alpha)$ or $f(\alpha + \varepsilon_i) = P_{\alpha,d}(\varepsilon_i) + o\big(\varepsilon_i^d\big)$ as $i \to \infty$ for every α and every $d \in \mathbb{N}$, where $P_{\alpha,d}$ is a polynomial of degree d and the ε_i are a sequence whose absolute values tend to 0 as i tends to ∞. Such a function of course need not extend as a C^∞ or even continuous function to $\mathbb{R}$ $\big($an obvious counterexample being $f(x) = 1/\big(x - \sqrt{2}\big)\big)$, but if it does then this extension is unique, so that the space $C^\infty_{\mathrm{strong}}(\mathbb{Q})$ of smooth functions in this sense contains $C^\infty(\mathbb{R})$ as a subspace. But there is a third, weaker, sense, in which one requires $f(\alpha+\varepsilon_i)=f(\alpha)+o(1)$ or $f(\alpha+\varepsilon_i)=P_{\alpha,d}(\varepsilon_i)+o\big(\varepsilon_i^d\big)$ only for sequences $\{\varepsilon_i\}$ of rational numbers that have bounded *numerators* but denominators tending to infinity (so that in particular they tend to 0 in the usual sense). We then have the strict inclusions

$$C^\infty(\mathbb{R}) \subsetneqq C^\infty_{\text{strong}}(\mathbb{Q}) \subsetneqq C^\infty_{\text{weak}}(\mathbb{Q}) \subsetneqq C^\infty_{\text{discrete}}(\mathbb{Q}) = \mathbb{C}^{\mathbb{Q}}.$$

An example (courtesy of Peter Scholze) of a function $f\colon \mathbb{Q}\to\mathbb{R}$ that is C^∞ in the weak sense but not in the strong sense is given by choosing a sequence of rational numbers $\{x_n\}$ tending to 0 and disjoint intervals $I_n\ni x_n$ with I_n containing no rational numbers with numerator $\le n$; then define f to be 0 at $x=0$ and to be the restriction of a C^∞ function on $\mathbb{R}^*$ supported on $\bigcup_n I_n$ and with $f(x_n)=1$, in which case f is obviously smooth in the strong sense always from 0 and in the weak sense at 0 (since the values of f on any sequence of rational numbers tending to 0 with bounded numerators stabilizes to 0), but is not even continuous at 0.

After this lengthy preliminary discussion, we can state the result on the smoothness properties of the cocycle W_γ, with an explicit formula for the power series of $W_\gamma(\alpha+\varepsilon)$ near any $\alpha\in\mathbb{Q}$. We will write ε as $-\hbar$ to match our previous conventions.

Proposition 5.2 (assuming RQMC)**.** *The function W_γ belongs to $C^\infty_{\text{weak}}\big(\mathbb{Q}\smallsetminus\big\{\gamma^{-1}(\infty)\big\}\big)$ for every $\gamma\in\mathrm{PSL}_2(\mathbb{Z})$. Explicitly, $W_\gamma(\alpha-\hbar)$ for $\alpha\ne\gamma^{-1}(\infty)$ and $\hbar$ tending to 0 with bounded numerator is given to all orders in $\hbar$ by the power series*

$$W_\gamma(\alpha-\hbar)\approx \mathbf{\Phi}_{\gamma\alpha}(2\pi\mathrm{i}\hbar^*)^{-1}\,\mathbf{diag}\left(\left|\frac{\operatorname{den}(\gamma\alpha)\hbar^*}{\operatorname{den}(\alpha)\hbar}\right|^{\kappa(\sigma)}\mathrm{e}^{v(\sigma)\lambda_\gamma(\alpha)}\right)\mathbf{\Phi}_\alpha(2\pi\mathrm{i}\hbar), \tag{5.13}$$

with $\hbar^*=\hbar/((c\alpha+d)(c\alpha-c\hbar+d))$ if $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$.

Proof. By the definition of weak smoothness on $\mathbb{Q}$, we have to show that $W_\gamma(\alpha+\varepsilon)$ for fixed $\gamma\in\mathrm{SL}_2(\mathbb{Z})$ and $\alpha\in\mathbb{Q}$ is given to all orders by a power series in ε depending only on α and γ as ε tends to 0 through rational numbers with bounded numerator. If we write α as a'/c' with a' and c' coprime and extend $\left(\begin{smallmatrix}a'\\c'\end{smallmatrix}\right)$ to a matrix $\gamma'=\left(\begin{smallmatrix}a'&b'\\c'&d'\end{smallmatrix}\right)\in\mathrm{SL}_2(\mathbb{Z})$, then the condition of ε having bounded numerator is easily seen to be equivalent to the condition that $\alpha+\varepsilon=\gamma'X$ with X tending to $\pm\infty$ with bounded denominator. We consider first the case when $X\to+\infty$ (meaning that $\alpha+\varepsilon$ tends to α from the left). By the cocycle property (5.3) and the basic asymptotic property (5.2) of W_γ, we have

$$\begin{aligned}W_\gamma(\alpha+\varepsilon)&=W_\gamma(\gamma'X)=W_{\gamma\gamma'}(X)W_{\gamma'}(X)^{-1}\\&\approx\widehat{\mathbf{\Phi}}_{a''/c''}\left(\frac{2\pi\mathrm{i}}{c''(c''X+d'')}\right)^{-1}\widehat{\mathbf{\Phi}}_{a'/c'}\left(\frac{2\pi\mathrm{i}}{c'(c'X+d')}\right),\end{aligned}$$

where we have written $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ and $\gamma\gamma'=\gamma''=\left(\begin{smallmatrix}a''&b''\\c''&d''\end{smallmatrix}\right)$ and where $\approx$ as usual means that the two expressions being compared are equal to all orders in $1/X$ as X tends to infinity with bounded denominator or ε to zero with bounded numerator. If we now use our previous conventions, writing

$$-\varepsilon=\hbar=\frac{1}{c'(c'X+d')},\qquad \hbar^*=\frac{1}{c''(c''X+d'')}=\frac{c'^2\hbar}{c''(c''-cc'\hbar)},$$

and also use that the "tweaking function" λ_γ satisfies

$$\frac{1}{c'^2\hbar}-\frac{1}{c''^2\hbar^*}=\left(X+\frac{d'}{c'}\right)-\left(X+\frac{d''}{c''}\right)=\frac{c}{c'c''}=\lambda_\gamma(\alpha), \tag{5.14}$$

then we get equation (5.13) above. This equation expresses $W_\gamma(\alpha+\varepsilon)$ as a product of three matrices of power series in $\hbar=-\varepsilon$ and hence shows that it is itself such a matrix. This proves the assertion in the first case $X\to+\infty$. To treat the case $X\to-\infty$ we use equation (5.11) instead of (5.2) and find that $W_\gamma(\alpha-\varepsilon)$ is given by the *same* formula as a product of three matrices of power series for $\varepsilon>0$ as it was for $\varepsilon<0$, because the prefactors B in (5.11) cancel. This completes the proof that W_γ is a two-sided smooth function on the rational numbers in the weak sense. ■

Corollary 5.3. *The function $x\mapsto \mathbf{\Phi}_\alpha(2\pi i x)^{\mathrm{exact}}$ on $\mathbb{Q}$ defined by (5.12) is differentiable in the weak sense for every $\alpha\in\mathbb{Q}$.*

Proof. This follows directly from Proposition 5.2 away from $x=0$, since the diagonal prefactor in (5.12) is smooth away from 0, and $\mathbf{\Phi}_\alpha(2\pi i x)^{\mathrm{exact}}$ simply agrees with $\mathbf{\Phi}_\alpha(2\pi i x)$ to all orders in x as $x\to0$, so it is smooth there too. ■

As a final remark in this subsection, we recall that in order to specify the cocycle $\gamma\mapsto W_\gamma$ completely, it suffices to give its values for the two special matrices $T=\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $S=\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$, since these generate the whole modular group. The function $W_T(x)$ is elementary (constant and conjugate to a diagonal matrix of Nth roots of unity, where N is the level of the knot). In the case where the level is 1, such as the 4_1 or 5_2 knots, W_T is simply the identity matrix and the whole cocycle is determined by the single matrix-valued function $W_S(x)$. The fact that $T\mapsto1$, $S\mapsto W_S$ extends to a cocycle on the whole group is then equivalent to the requirement that $W_S(x)$ satisfies the symmetry property $W_S(x)=W_S(-1/x)^{-1}$, together with the three-term Lewis functional equation

$$W_S(x)=W_S(x+1)W_S(x/(x+1)),$$

familiar from the theory of period polynomials of holomorphic modular forms on $\mathrm{SL}_2(\mathbb{Z})$ or of period functions in the sense of [63] of Maass forms on $\mathrm{SL}_2(\mathbb{Z})$. Our cocycles thus belong in some sense to the same family as periods of modular forms.

5.3 "Functions near $\mathbb{Q}$"

We now come to an important and somewhat subtle point. In the calculation that we gave to prove Proposition 5.2, we used only the refined quantum modularity conjecture in its "rational version" (4.12), since the statement of Proposition 5.2 involves only the values of W_γ at rational arguments. If we had used instead the full version (4.13) of the RQMC, we would have obtained a stronger version of the "weakly smooth" condition that applies to approximating a rational number not just by rational numbers that differ from it by a small rational number with bounded numerator, but also by infinitesimal variations of such numbers. To make sense of a statement of this type, we now introduce a notion that will shed more light on the two cocycles $\gamma\mapsto\lambda_\gamma$ and of $\gamma\mapsto W_\gamma$ and that is also relevant in connection with the notion of "holomorphic quantum modular forms" that will be touched on briefly in Section 5.4 and developed in more detail in [44] and in the survey paper [85]. This is the notion of *asymptotic functions near* $\mathbb{Q}$. The basic idea here is to specify a particular type of asymptotic behavior (such as a formal power series) in an infinitesimal neighborhood of every rational point, where "neighborhood" can mean that we approach the rational number from the right and the left on the real line, or in other contexts from above and below in the upper and lower complex half-planes. Since there are many types of behavior that may be of interest, and since it is hard to give a general definition that includes all of the examples and all of the properties that one wishes to include, we will restrict here to the particular classes that arise in the context of knot invariants.

The simplest version of this notion is just given by a collection $\{f_\alpha(\varepsilon)\}_{\alpha\in\mathbb{Q}}$ of formal power series with complex coefficients indexed by the rational numbers. Here we want to think of the infinitesimal power series variable ε as the difference between the rational number α and an infinitesimally nearby real "number" $\alpha+\varepsilon$, i.e., we want to think of the whole collection of power series $\{f_\alpha\}$ as a single "asymptotic function near $\mathbb{Q}$", i.e., as a "function" f defined in infinitesimal neighborhoods of all rational points α by $f(\alpha+\varepsilon)=f_\alpha(\varepsilon)$. Of course f is not a function at all in the traditional sense, since one cannot evaluate it at numerical values of its argument, but as we will see in a moment, this point of view is nevertheless very fruitful. It originally showed up in the paper [84], where "quantum modular forms" were first defined simply as functions on $\mathbb{Q}$ (more precisely, as almost-everywhere-defined functions on $\mathbb{Q}$) but then upgraded to a notion of "strong quantum modular forms" where the original values at rational numbers became the constant terms of a collection of formal power series.

The set of asymptotic functions near $\mathbb{Q}$ (from now on we omit the quotation marks, trusting the reader to remember that these are not actually functions) of this special type forms a ring via pointwise addition and multiplication if we think of its elements as collections of formal power series, and by straight addition and multiplication if we think of them as functions defined in infinitesimal neighbourhoods of all rational points. To understand its elements, it is helpful to think of the following two extreme cases.

(i) Each $f_\alpha(\varepsilon)$ is the Taylor expansion $\sum f^{(n)}(\alpha)\varepsilon^n/n!$ of a function $f\in\mathbb{C}^\infty(\mathbb{R})$ at the point α. Here the various asymptotic expansions near rational points fit together nicely into a single smooth function on $\mathbb{R}$.

(ii) Each $f_\alpha(\varepsilon)$ is the formal power series expansion at $q=\mathrm{e}^{2\pi\mathrm{i}(\alpha+\varepsilon)}$ of an element $A(q)$ of the Habiro ring $\mathcal{H}=\varprojlim\mathbb{Z}[q]/((q;q)_n)$. Here the different power series do not in general fit together smoothly at all, and even their constant terms jump around wildly, as illustrated in Figure 1 in the introduction.

At first sight this definition seems to be pointless because we are not requiring any compatibility at all between the different power series f_α and therefore the ring we have just introduced is canonically isomorphic to the direct product $\mathbb{C}[[\varepsilon]]^{\mathbb{Q}}=\prod_\alpha\mathbb{C}[[\varepsilon]]$ of one copy of the power series ring $\mathbb{C}[[\varepsilon]]$ for every rational number α. The point, however, is that if we pass to the quotient ring $\mathfrak{F}_{0,0}\approx\prod_\alpha\mathbb{C}[[\varepsilon]]/\bigoplus_\alpha\mathbb{C}[[\varepsilon]]$ of asymptotic functions in the neighborhood of all but a finite set of rational points, then the modular group $\Gamma_1=\mathrm{SL}_2(\mathbb{Z})$ acts by sending f to $f\circ\gamma$ for $\gamma\in\Gamma_1$, and this action does not simply permute the different power series f_α but twists them as well. Specifically, if $f(x)$ is represented near α by $f(\alpha+\varepsilon)=f_\alpha(\varepsilon)$ then $f(\gamma(x))$ is represented near $\alpha^*=\gamma(\alpha)$ by $f_{\alpha^*}(\varepsilon^*)$ rather than simply by $f_{\alpha^*}(\varepsilon)$, where $\varepsilon^*=\gamma(\alpha+\varepsilon)-\gamma(\alpha)$, or more explicitly $\varepsilon^*=\varepsilon/(c\alpha+d)(c\alpha+d+c\varepsilon)$ if $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$. This is precisely the twist that we already encountered in Section 3.2 (equations (3.9) and (3.12)) and in Proposition 5.2 above, except that we have changed the previous variable h to $-2\pi\mathrm{i}\varepsilon$ here. (The rescaling of h by a factor $2\pi\mathrm{i}$ was introduced for our knot invariants only to make the power series coefficients algebraic and there is no reason to make this change of variable in the general situation.)

We now generalize the above notion by introducing two complex parameters v and κ and considering the vector space of asymptotic functions near $\mathbb{Q}$ whose local form $f_\alpha(\varepsilon)=f(\alpha+\varepsilon)$ in a real infinitesimal neighborhood of any $\alpha\in\mathbb{Q}$ is given by

$$f_\alpha(\varepsilon)=|\mathrm{den}(\alpha)\varepsilon|^\kappa\mathrm{e}^{-v/\mathrm{den}(\alpha)^2\varepsilon}\phi_\alpha(\varepsilon)\qquad(5.15)$$

for some power series $\phi_\alpha(\varepsilon)\in\mathbb{C}[[\varepsilon]]$. Again we pass to the quotient $\mathfrak{F}_{\kappa,v}$ of almost-everywhere-defined asymptotic functions on $\mathbb{Q}$, i.e., we identify two collections of completed power series if they differ for only finitely many α. The space $\mathfrak{F}_{\kappa,v}$ is a free module of rank 1 over the ring $\mathfrak{F}_{0,0}$ introduced above, and is again isomorphic to $\prod_\alpha\mathbb{C}[[\varepsilon]]/\bigoplus_\alpha\mathbb{C}[[\varepsilon]]$ via $f\mapsto\{\phi_\alpha\}$, but with a different action of $\mathrm{SL}_2(\mathbb{Z})$ than before. Specifically, $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ sends f to the function near $\mathbb{Q}$ that corresponds via (5.15) to the collection of power series $\{\phi^*_\alpha(\varepsilon) = \mathrm{e}^{v\lambda_\gamma(\alpha)}\phi_{\alpha^*}(\varepsilon^*)\}$ with α^* and ε^* as above and with the "tweaking cocycle" $\lambda_\gamma(\alpha)$ introduced in (3.5). Alternatively, in terms of the variable $x = \alpha + \varepsilon$ infinitesimally near $\alpha \in \mathbb{Q}$ we can write this action as the "slash action" (familiar from the theory of modular forms if κ is an even integer, and from the theory of the principal series representation of $\mathrm{SL}_2(\mathbb{R})$ if not) given by $(f|_\kappa\gamma)(x) = |cx+d|^{-\kappa} f(\gamma x)$ for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, where $|cx+d|^{-\kappa} := |c\alpha + d|^{-\kappa}(1 + c\varepsilon/(c\alpha+d))^{-\kappa}$. This also explains the reason for including the perhaps strange-looking factors $\mathrm{den}(\alpha)$ and $\mathrm{den}(\alpha)^{-2}$ in (5.15), since without them there would be no action of the modular group. We also point out that, because of the "tweaking" factor $\mathrm{e}^{v\lambda_\gamma(\alpha)}$ in the definition of the action, $\mathfrak{F}_{\kappa,v}$ is a free module of rank 1 over the ring $\mathfrak{F}_{0,0}$ as a vector space, but not as an $\mathrm{SL}_2(\mathbb{Z})$-module: one cannot choose a generator in an $\mathrm{SL}_2(\mathbb{Z})$-invariant way.

Of course from the point of view of this paper the main reason for introducing the parameters κ and v and the definition (5.15) is that this is exactly the behavior that we found experimentally from the refined quantum modularity conjecture, with $v = v(\sigma)$ and $\kappa = \kappa(\sigma)$ being the normalized volume and weight associated to a parabolic flat connection σ and with $\phi_a(\varepsilon)$ and $f_\alpha(\varepsilon)$ being the power series $\Phi^{(K,\sigma)}_\alpha(h)$ $\big($or more generally $\Phi^{(K,\sigma,\sigma')}_\alpha(h)\big)$ and its completion $\widehat{\Phi}^{(K,\sigma)}_\alpha(h)$ $\big($or $\widehat{\Phi}^{(K,\sigma,\sigma')}_\alpha(h)\big)$ as in (3.8), with $h = -2\pi\varepsilon$. But it is worth noting that the space $\mathfrak{F}_{\kappa,v}$ also contains classical modular forms on the full modular group, since a holomorphic modular form $f(\tau)$ of (necessarily even) weight k on $\mathrm{SL}_2(\mathbb{Z})$ canonically defines a function near $\mathbb{Q}$ of type (5.15) with κ equal to $-k$, with v equal to $-2\pi\mathrm{i}$ times the valuation of f at infinity (the smallest exponent of $q = \mathbf{e}(\tau)$ in the Fourier expansion of $f(\tau)$), and with each power series $\phi_\alpha(\varepsilon)$ reducing to its constant term $\phi_\alpha(0)$, as one sees easily by using the modular transformation property of f to compute the asymptotic development of $f(\alpha+\varepsilon)$ for $\alpha \in \mathbb{Q}$ and ε tending to 0 with positive imaginary part. More generally, mock modular forms (whose definition we omit) also define elements of $\mathfrak{F}_{\kappa,v}$, where κ is again the negative of the weight, but in that case the power series $\phi_\alpha(\varepsilon)$ are in general factorially divergent rather than constant. We do not elaborate on any of this since it is far from the theme of this paper, but it is nice to observe that classical modular and mock modular forms have properties in common with the asymptotic functions occurring here.

There are two further points that we should mention in connection with the definition (5.15). One is that the absolute value appearing there is only appropriate for ε real, which is our original situation when we think of $\alpha+\varepsilon$ as being a deformation of the rational number α on the real line or when we take $\varepsilon = -1/c(cX+d)$ with X a rational number tending to infinity as in the RQMC. But when we consider functions near $\mathbb{Q}$ in the complex as well as in the real domain, the absolute value sign would destroy holomorphy. If κ is an even integer, the problem does not arise, since we can simply replace $|\varepsilon|^\kappa$ by ε^κ, which is holomorphic. If this is not the case then if we consider only functions near rational points in the upper or lower half-plane, we can still replace $|\varepsilon|^\kappa$ in the definition by ε^κ, which makes sense because ε has a well-defined logarithm in either half-plane. (We will never encounter functions in $\mathfrak{F}_{\kappa,v}$ for $\kappa \neq 0$ that are defined in a 360^o complete complex neighborhood of α; our functions will either be defined for nearby real points or for nearby non-real points, or sometimes in a cut plane $\mathbb{C} \smallsetminus (-\infty, 0]$ or $\mathbb{C} \smallsetminus [0,\infty)$, in which case we can extend $|\varepsilon|^\kappa$ holomorphically as ε^κ or $(-\varepsilon)^\kappa$, respectively.) However, when κ is not an integer and we want to discuss the $\mathrm{SL}_2(\mathbb{Z})$ action on $\mathfrak{F}_{\kappa,v}$, then we have to include some kind of multiplier system, as familiar from the theory of modular forms of arbitrary weight. Again, we omit details.

The second minor comment is that one can further generalize $\mathfrak{F}_{0,0}$ by introducing a level N as well as the parameters κ and v. This generalization is necessary if we want to include modular or mock modular forms of level N (say on $\Gamma = \Gamma_0(N)$ or $\Gamma(N)$) into our definition, but also for our knot invariants if the knot has a level > 1, as we found to be the case for the $(-2,3,7)$-pretzel knot. Here the power series $\phi_\alpha(\varepsilon)$ and their completions will have period N rather than 1 with respect to α, and more importantly, the number v in (5.15) is no longer constant but must be replaced by a number v_α that depends on the Γ-equivalence class ("cusp") of α. Again we omit details, since this is not our main subject.

We now return to the functions studied in this paper and to the reason why we introduced asymptotic functions near $\mathbb{Q}$ in the first place. Consider first the tweaking function defined by equation (3.5). We showed in Lemma 3.1 that the map $\gamma \mapsto \lambda_\gamma$ is a cocycle in the space of almost-everywhere-defined functions on $\mathbb{P}^1(\mathbb{Q})$. It is easily checked that it is not a coboundary in that space. But if we extend λ_γ to a function near $\mathbb{Q}$ by setting $\lambda_\gamma(\alpha+\varepsilon) = \lambda_\gamma(\alpha)$ (constant power series), then equation (5.14) says that it now is a coboundary: $\lambda_\gamma(x) = \mu(x) - \mu(\gamma x)$ where μ is the function near $\mathbb{Q}$ defined by $\mu(\alpha+\varepsilon) = 1/\mathrm{den}(\alpha)^2\varepsilon$. More interestingly, if for each σ and σ' in $\mathcal{P}_K$ we define a function near $\mathbb{Q}$ by $Q^{(K,\sigma,\sigma')}(\alpha-\hbar) = \widehat{\Phi}_\alpha^{(K,\sigma,\sigma')}(h)$ and then put them together as a matrix-valued function $\mathbf{Q}$ near $\mathbb{Q}$ given by $\mathbf{Q}(\alpha-\hbar) = \widehat{\mathbf{\Phi}}_\alpha^{(K)}(h)$, then using equation (5.14) again we see that the complicated equation (5.13) can be replaced by the much simpler equation

$$W_\gamma(x) = \mathbf{Q}(\gamma x)^{-1}\mathbf{Q}(x). \tag{5.16}$$

Notice that in this equation the (σ,σ')-entry on the left-hand side is the sum over $\sigma'' \in \mathcal{P}_K$ of the product of the (σ,σ'')-entry of $\mathbf{Q}(\gamma x)^{-1}$ and the (σ'',σ')-entry of $\mathbf{Q}(x)$, which belong to $\mathfrak{F}_{-\kappa(\sigma''),-v(\sigma'')}$ and $\mathfrak{F}_{\kappa(\sigma''),v(\sigma'')}$, respectively. Thus each of the terms of the sum belongs to $\mathfrak{F}_{0,0}$ and we never encounter the problem of having to make sense of sums of asymptotic functions of different orders of growth. The fact that the entries of W_γ all belong to $\mathfrak{F}_{0,0}$ is, of course, a necessary prerequisite for the final statement that they actually belong to its subring $C^\infty(\mathbb{R})$.

Equation (5.16) tells us the cocycle $\gamma \mapsto W_\gamma$, which was not a coboundary in the space of almost-everywhere-defined matrix-valued functions on $\mathbb{P}^1(\mathbb{Q})$ or of piecewise smooth functions on $\mathbb{P}^1(\mathbb{R})$, becomes one when we pass to the space of matrix-valued functions near $\mathbb{Q}$. Both of these can be seen as manifestations of a general phenomenon that one finds in almost all mathematical contexts where notions of homology or cohomology play a role: even though one is really only interested in cocycles that are not coboundaries, the cocycles that one studies are almost always constructed as coboundaries in some bigger space.

5.4 Analyticity

In Section 5.2, we discussed the surprising smoothness properties of the function W_γ on $\mathbb{R} \smallsetminus \{\gamma^{-1}(\infty)\}$. In this subsection, we come to a point much deeper than the smoothness, namely analyticity properties of functions defined in a cut plane. These functions are closely related to *state integrals*. Such integrals appeared originally in the work of Hikami, Andersen, Kashaev and others (see, for example, [1, 16, 53]) in relation to the partition function of complex Chern–Simons theory and to quantum Teichmüller theory, and reappear in our context in [44], the companion paper to this one. We refer to these papers for details and describe the main points here in qualitative form only.

State integrals are analytic functions with several key features:

- They are holomorphic for all $\tau \in \mathbb{C}' = \mathbb{C} \smallsetminus (-\infty, 0]$.
- Their restrictions to $\mathbb{C} \smallsetminus \mathbb{R}$ factorize bilinearly as finite sums of products of a q-series and a $\tilde{q}$-series, where $q = \mathbf{e}(\tau)$ and $\tilde{q} = \mathbf{e}(-1/\tau)$; see [4, 30].
- Their evaluation at positive rational numbers also factorizes bilinearly as a finite sum of a product of a periodic function of τ and a periodic function of $-1/\tau$; see [29].

They are defined as multidimensional integrals of a product of quantum dilogarithms times the exponential of a quadratic form. The quantum dilogarithm, invented by Faddeev [19, 20], is a remarkable meromorphic function of two variables. The structure of its poles implies that the state integrals are holomorphic functions of τ in the cut plane $\mathbb{C}'$. The quantum dilogarithm is also a quasi-periodic function with two quasi-periods, and this has two consequences, one of which is directly related to the third "feature" above, and the other to the second "feature" and to the paper [44].

The first consequence is the fact that one can apply the residue theorem to give an exact formula for the values of the state-integrals at positive rational numbers. Such a formula was given explicitly for the one-dimensional state integrals considered in [29, equation (1), Theorem 1.1], and those one-dimensional state integrals cover the case of the three knots that we consider here, namely the 4_1, 5_2 and $(-2,3,7)$ pretzel knot. It turns out that equation (15) of [29] applied to the case of $(A,B)=(1,2)$ gives a function on $\mathbb{Q}^+$ which is none other than one of our four entries of $W_\gamma^{(4_1)}$ when $\gamma=S$. To get the other three entries of $W_S^{(4_1)}$ one can apply the proof of [29] to a 2×2 matrix of state-integrals of the 4_1 knot introduced in [27, Theorem 3]. And finally, to get the full matrix $W_\gamma^{(4_1)}$ for all γ, one can apply the proof of [29] to a 2×2 matrix of state-integrals of the 4_1 knot that depend on a modular version of Faddeev's quantum dilogarithm [32].

The second consequence is perhaps even more interesting. Not only is each component of the state integral matrix a finite sum of products of a q-series and a $\tilde{q}$-series, but this sum precisely corresponds to matrix multiplication and says that the whole state integral matrix $W_S(\tau)$, whose restriction to a real half-line is our function W_S, factors as the product of a matrix of $\tilde{q}$-series multiplied by a matrix of q-series. More explicitly, $W_S(\tau)$ factors in the upper and lower complex half-planes as $\mathbf{Q}^{\rm hol}(-1/\tau)\mathbf{j}_S^{\rm hol}(\tau)^{-1}\mathbf{Q}^{\rm hol}(\tau)$, where $\mathbf{Q}^{\rm hol}(\tau)$ is an $(r+1)\times(r+1)$ matrix with holomorphic and periodic entries and $\mathbf{j}_S^{\rm hol}(\tau)$ is a diagonal matrix of automorphy factors. Furthermore, the equivariant extension W_γ of the state integrals mentioned above is again a holomorphic function in the cut plane whose restriction to the real half-line is our cocycle W_γ from Section 5.3 and whose restriction to $\mathbb{C}\smallsetminus\mathbb{R}$ factors for every γ as $\mathbf{Q}^{\rm hol}(\gamma(\tau))\mathbf{j}_\gamma^{\rm hol}(\tau)^{-1}\mathbf{Q}^{\rm hol}(\tau)$ with the same periodic function $\mathbf{Q}^{\rm hol}(\tau)$. The fact that this quotient extends analytically across a half-line, even though the matrix-valued holomorphic function $\mathbf{Q}^{\rm hol}(\tau)$ does not, is an example of a (matrix-valued) *holomorphic quantum modular form*, a new and quite general context that is discussed in much more detail in [44, 85], and of which the mock modular forms mentioned in the previous subsection give another nice example. The fact that $W_\gamma(\tau)$ factors as $\mathbf{Q}^{\rm hol}(\gamma(\tau))\mathbf{j}_\gamma^{\rm hol}(\tau)^{-1}\mathbf{Q}^{\rm hol}(\tau)$ is another instance of the general principle ("cocycles are constructed as coboundaries in some larger space") mentioned at the end of Section 5.3. So we have now represented the original cocycle $\gamma\mapsto W_\gamma$ on the real line as a coboundary in two different worlds: functions defined in a small open neighborhood of $\mathbb{P}^1(\mathbb{R})\smallsetminus X$ in $\mathbb{P}^1(\mathbb{C})$ for some finite set X, and asymptotic functions near $\mathbb{Q}$. But in fact these two representations $W_\gamma(\tau)=\mathbf{Q}^{\rm hol}(\gamma(\tau))\mathbf{j}_\gamma^{\rm hol}(\tau)^{-1}\mathbf{Q}^{\rm hol}(\tau)$ and $W_\gamma(x)=\mathbf{Q}(\gamma(x))\widetilde{\mathbf{j}}_\gamma(x)^{-1}\mathbf{Q}(x)$ are not independent: as we will see in [44], the periodic holomoprhic function $\mathbf{Q}^{\rm hol}$ has an asymptotic development as one approaches any rational number from above or below in $\mathbb{C}\smallsetminus\mathbb{R}$, and this is a representative of the *same* asymptotic functions near $\mathbb{Q}$ that we obtained from the Habiro-like functions on $\mathbb{P}^1(\mathbb{Q})$. It is this manifestation of the same abstract object in two completely different realizations that we referred to in the opening paragraph of this paper as an analogue in our context of the notion of motives.

The above discussion explains why one can expect, and in a few cases even prove, the analyticity of the cocycle function $W_\gamma(x)$. But it also seems worth observing that, once one has predicted this analyticity, one can check it numerically using only the matrices studied in this paper, without any reference to either q-series or state integrals. Specifically, Proposition 5.2 gives the Taylor expansion of W_γ at any rational point, and since the coefficients of this series are effectively computable, we can calculate a large number of them and see experimentally that the series has a non-zero radius of convergence, as was already done in equation (3.15) for the special case of the expansion of W_S for the 4_1 knot around $x=1$. In fact, the coefficients can be computed in two different ways, either by using the refined quantum modularity conjecture numerically with the help of optimal and smooth truncation of divergent series, as was done in Section 4, or else by using the *exact* formulas (when they are available, e.g., for the 4_1 and 5_2 knots) for W_γ on $\mathbb{Q}$ to compute the values of this function at many rational points near a given point and then interpolating numerically by the method recalled in "Step 3" of Section 10.1. In this way, we can verify the predicted real-analyticity to high precision and in a very convincing way using only the data coming from the Kashaev invariant and its associated functions. The simplest example is equation (3.15) given in Section 2 for the 4_1 knot and $\gamma = S$. The improvement of convergence in this case is very dramatic: the 150th coefficient of $\Phi(2\pi\mathrm{i}x)$ (the last one that we computed) is about 10^{284}, but the 150th coefficient of the bilinear combination of power series occurring on the right-hand side of (3.15) is only 0.002!

But here we can actually do even more; by changing the variables one gets a new series that not only again (conjecturally and experimentally) has radius of convergence 1, but that now also gives numerical confirmation of the prediction that $W_S(x)$ extends holomorphically to the whole cut plane. Specifically, if we make the change of variables $1+x = \left(\frac{1+w}{1-w}\right)^2$, under which $x=0$ corresponds to $w=0$ and the condition $1+x \in \mathbb{C}'$ is equivalent to $|w|<1$, then we get a power series $B(w) \in \mathbb{R}\left[\left[w^2\right]\right]$ defined by

$$B(w) = \mathrm{e}^{-v(4_1)}\Phi\left(\frac{8\pi\mathrm{i}w}{(1-w)^2}\right)\Phi\left(-\frac{8\pi\mathrm{i}w}{(1+w)^2}\right) - \mathrm{e}^{v(4_1)}\Phi\left(-\frac{8\pi\mathrm{i}w}{(1-w)^2}\right)\Phi\left(\frac{8\pi\mathrm{i}w}{(1+w)^2}\right) \quad (5.17)$$

(with $\Phi(x) \in \mathbb{R}[[x]]$ again given by (1.3)) which should have radius of convergence 1. In fact, the numerical calculation, described in [44], show that the 150th coefficient of B is about $-7.5\cdot 10^{10}$, again far smaller than the original 10^{284}. The fact that this number is much bigger than the corresponding number 0.002 for the bilinear combination (3.15) is not because the series $B(w)$ is worse than the one in (3.15), but precisely because it is better: in order to get the full domain $\mathbb{C}'$ of holomorphy of $W_S(x)$ we have had to produce a power series that has singularities on the entire unit circle rather than at just one point, and the coefficients correspondingly have much larger growth (namely exponential in the square-root of the index, just as in the Hardy–Ramanujan partition formula, rather than being only of polynomial growth, or in this case even of polynomial decay). But in any case, whether we use (5.17) or just (3.15), we see that the single divergent power series $\Phi(h)$, which describes the asymptotic behavior of $W_S^{(4_1,\sigma_1,\sigma_1)}(x)$ near either ∞ or 0, suffices in an explicit manner to determine this function everywhere on all of $\mathbb{R}^*$. For general knots, the corresponding statement would only hold if we consider the entire matrix $\mathbf{\Phi}$ rather than just one entry. In fact, as the whole discussion of Sections 4 and 5 shows, if we assume the whole RQMC, then at least in favorable cases it is probably true that the single power series $\Phi_0^{(K)}(h)$ coming from the modularity of the original Kashaev invariant actually determines everything.

We can summarize this whole subsection as the following conjecture for the cocycle W_γ.

Conjecture 5.4. *The function W_γ on $\mathbb{Q}\smallsetminus\left\{\gamma^{-1}(\infty)\right\}$ extends to a real-analytic function on $\mathbb{R}\smallsetminus\left\{\gamma^{-1}(\infty)\right\}$, and its restriction to each component of $\mathbb{R}\smallsetminus\left\{\gamma^{-1}(\infty)\right\}$ extends to a holomorphic function on the cut plane consisting of this half-line and $\mathbb{C}\smallsetminus\mathbb{R}$.*

5.5 The non-hyperbolic case

The main thrust of this paper, and all of the examples which we have treated in detail, concern the case of hyperbolic knots, for which the volume is positive. We expect that matrix-valued cocycles exist for nonhyperbolic 3-manifolds, with or without boundary, and know that this is so for the example of the complement of the trefoil [82] (where the corresponding invariant is sometimes known by the name Kontsevich–Zagier series) as well as for WRT invariant of the Poincaré homology sphere (a spherical 3-manifold), which was studied by Lawrence and Zagier [62]. In these examples and many others that have been treated since, the series that occur are Taylor series of mock modular forms, and we think that this will always happen for manifolds for which all of the volumes vanish modulo $\pi^2\mathbb{Q}$ (e.g., torus knots, Seifert-fibered manifolds or, in the closed case, spherical manifolds). When it happens, the entries in the $\mathcal{P}_K^{\mathrm{red}}$-part of the matrix are the product of an elementary exponential term $\big($a rational power of $\mathrm{e}^{\pi^2/h}\big)$ and a rational power of q, so that the corresponding Φ-series is purely exponential in h, while the entries in the top row of the matrix (which are again elements of the Habiro ring) still have factorially divergent h-series as in the hyperbolic case, are now elementary functions, with coefficients that are special values of Dirichlet L-series and a Borel transform which is simply a trigonometric function. However, we should emphasize that this simple behavior is not expected for all non-hyperbolic knots or manifolds, but only for those for which *all* solutions of the Neumann–Zagier equations are torsion in the Bloch group, so that all volumes $v(\sigma)$ are rational multiples of $2\pi\mathrm{i}$. Some knots, the like $(2,1)$-cabling of the 4_1 knot, are non-hyperbolic, so have vanishing volume $\mathbf{V}(\sigma_1)$ modulo $4\pi^2$ but have some $\mathbf{V}(\sigma)$ with non-zero imaginary part, and then one expects to find non-trivial h-series.

Part II. Complements

6 Half-symplectic matrices and their perturbative series

In Section 2, we introduced a finite set $\mathcal{P}_K$ associated to a knot K and the formal power series $\Phi_\alpha^{(K,\sigma)}(h)$ for each $\alpha \in \mathbb{Q}$ and $\sigma \in \mathcal{P}_K^{\mathrm{red}} = \mathcal{P}_K \smallsetminus \{\sigma_0\}$, as defined by Dimofte and the first author in [14, 15] in terms of the Neumann–Zagier data of a triangulation of the knot complement. In this section, we provide details and also a somewhat more general construction, depending on more general data consisting of a "half-symplectic matrix" (defined below), an integral vector, and a solution of the associated Neumann–Zagier equations. This more general class has a q-holonomic structure that will be studied in Section 7 and will also include the formal power series $\Phi_\alpha^{(K,\sigma,\sigma')}(h)$ $(\sigma,\sigma' \in \mathcal{P}_K)$ that we found in Sections 4 and 5, as well as the asymptotic series of Nahm sums near roots of unity. These half-symplectic matrices give a new perspective on the classical Bloch group and the extended Bloch group.

6.1 Half-symplectic matrices and the Bloch group

To each knot K and each element $\sigma \in \mathcal{P}_K$ there is an associated element of the Bloch group (or third algebraic K-group) of $\overline{\mathbb{Q}}$ that plays a central role for many of the constructions and that can be described in terms of the Neumann–Zagier data of a triangulation of the knot complement. In fact, this construction produces an invariant lying in a set defined by "half-symplectic matrices" (= upper halves of symplectic matrices over $\mathbb{Z}$) which is a refinement of the usual Bloch group that has several nice aspects and seems not to have been considered in the literature. In this subsection and the following one, we will describe this set and how one obtains elements in it from the data of a triangulation. In the final subsection, we will explain how to associate a formal power series in h to any such element, the two cases of primary interest being the matrix of power series $\Phi_\alpha^{(K,\sigma,\sigma')}(h)$ associated to a knot and the power series describing the asymptotics of Nahm sums near rational points.

For each positive integer N, we denote by $\mathbf{H}_N$ the set of $N \times 2N$ *half-symplectic matrices*, by which we mean matrices $H = (AB) \in M_{N\times 2N}(\mathbb{Z})$ satisfying the two conditions

(i) the $2N$ columns of H span $\mathbb{Z}^N$ as a $\mathbb{Z}$-module, and

(ii) the matrix AB^t is symmetric.

The name refers to the fact that such matrices arise as the upper half of symplectic matrices, i.e., of matrices $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \mathrm{GL}_{2N}(\mathbb{Z})$ satisfying $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} D^t & -B^t \\ -C^t & A^t \end{smallmatrix}\right)$. To each $H \in \mathbf{H}_N$, we associate the generically zero-dimensional variety V_H defined as the set of N-tuples $z = (z_1, \ldots, z_N)$ in $\left(\mathbb{A}^1 \smallsetminus \{0,1\}\right)^N$ $\left(\mathbb{A}^1 = \text{affine line}\right)$ satisfying the equations

$$V_H\colon \quad \prod_{j=1}^N z_j^{A_{ij}} = (-1)^{(AB^t)_{ii}} \prod_{j=1}^N (1-z_j)^{B_{ij}}, \qquad i = 1, \ldots, N. \tag{6.1}$$

To define the associated power series, we will need both an element of $V_H(\mathbb{C})$ and a slightly stronger discrete datum than H, namely a pair (or triple)

$$\Xi = (H, \nu) = ((AB), \nu) \qquad \text{with} \quad \nu \in \operatorname{diag}\bigl(AB^t\bigr) + 2\mathbb{Z}^N. \tag{6.2}$$

Equation (6.1) can then be written in abbreviated form as $z^A = (-1)^\nu (1-z)^B$.

We observe that there is a second description of the variety V_H as the set of N-tuples $x = (x_1, \ldots, x_N)$ satisfying the trinomial equations

$$1 = (-1)^{\sum_j C_{ji}\nu_j} \prod_{j=1}^N x_j^{A_{ji}} + (-1)^{\sum_j D_{ji}\nu_j} \prod_{j=1}^N x_j^{B_{ji}}, \qquad i = 1, \ldots, N \tag{6.3}$$

or in abbreviated form $1 = (-1)^{C^t\nu} x^{A^t} + (-1)^{D^t\nu} x^{B^t}$, which is isomorphic to V_H via the bijections $x \mapsto z = (-1)^{D^t\nu} x^{B^t} = 1 - (-1)^{C^t\nu} x^{A^t}$ and $z \mapsto x = z^{-C^t}(1-z)^{D^t}$. The x are the Ptolemy coordinates as discussed in Section 6.2 below in the case of knots and their logarithms are the vectors w used below. Note that the signs $(-1)^{C^t\nu}$ and $(-1)^{D^t\nu}$ in (6.3) formulas do not depend on the choice of ν, since its value modulo 2 is fixed by (6.2). They do depend on the choice of symplectic completion (CD) of the half-symplectic matrix (AB), but only in a trivial way: any other choice (C^*D^*) of (CD) has the form $(CD) + S(AB)$ for some symmetric integral $N \times N$ matrix S (this corresponds to multiplying $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ on the left by the symplectic matrix $\left(\begin{smallmatrix} 1 & 0 \\ S & 1 \end{smallmatrix}\right)$), and this simply replaces x by $(-1)^{S^t\nu}x$, i.e., it changes the signs of some of the x_i.

To any complex solution z of the system of equations (6.1) one can associate a complex volume $\mathbf{V}(z)$ in $\mathbb{C}/4\pi^2\mathbb{Z}$ that is defined roughly as the sum of the dilogarithms of the z_i plus a suitable logarithmic correction. More concretely, the imaginary part of $\mathbf{V}(z)$ is a well-defined real number given as $\sum_j D(z_j)$, where $D(z) = \operatorname{Im}\bigl(\mathrm{Li}_2(z) + \log|z|\log(1-z)\bigr)$ is the Bloch–Wigner dilogarithm, which is single-valued. To define the full value of $\mathbf{V}(z)$ modulo $4\pi^2$ requires more work, because the function $\mathrm{Li}_2(z)$ itself, defined by analytic continuation from its value $\sum_{n\ge1} z^n/n^2$ for $|z|<1$, is multivalued on $\mathbb{C} \smallsetminus \{0,1\}$. However, the function $F(v) = \mathrm{Li}_2(1-\mathrm{e}^v)$ has the derivative $v/(\mathrm{e}^{-v}-1)$, which is meromorphic with residues in $2\pi\mathrm{i}\mathbb{Z}$. Hence, F is a well-defined function from $\mathbb{C} \smallsetminus 2\pi\mathrm{i}\mathbb{Z}$ to $\mathbb{C}/4\pi^2\mathbb{Z}$, satisfying the easily checked functional equation $F(v + 2\pi\mathrm{i}n) = F(v) - 2\pi\mathrm{i}n\log(1-\mathrm{e}^v)$ for $n \in \mathbb{Z}$. (See [86].) We can then define

$$\mathbf{V}(z) = \mathbf{V}_\Xi(z) \in \mathbb{C}/4\pi^2\mathbb{Z}$$

by the formula

$$\mathbf{V}(z) = \sum_{j=1}^N \left(F(v_j) + \frac{u_j v_j}{2} + \frac{\pi\mathrm{i}\nu_j}{2}(Cu - Dv)_j - \frac{\pi^2}{6} \right), \tag{6.4}$$

where u_j and v_j are any choice of logarithms of z_j and $1-z_j$ satisfying $Au - Bv = \pi\mathrm{i}\nu$ (which automatically exist as a consequence of the conditions on ν in (6.2) and the condition (i) on H) and where (CD) is the bottom half of a completion of H to a full symplectic matrix. To see that this number, which is only well-defined modulo $4\pi^2$, is independent of the choice of u and v, we observe that any other choice (u^*v^*) of logarithms of z and $1-z$ satisfying $Au^*-Bv^*=\pi\mathrm{i}\nu$ has the form $(u^*v^*)=(uv)+2\pi\mathrm{i}\big(B^tA^t\big)n$ for some $n\in\mathbb{Z}^N$ (this follows easily from the conditions (i) and (ii)), and then using the functional equation of F, we find

$$\mathbf{V}^*-\mathbf{V}=\pi\mathrm{i}n^t\big(-2Au+Au+Bv+2\pi\mathrm{i}AB^tn-\pi\mathrm{i}\nu\big)=2\pi^2\big(n^t\nu-n^tAB^tn\big),$$

which is 0 modulo $4\pi^2$ because AB^t is symmetric and integral with diagonal congruent to ν modulo 2. On the other hand, the expression (6.4) *does* depend on the choice of the $2N\times N$ integral matrix (CD), but only very mildly, by a multiple of $\pi^2/2$, since changing (CD) to $(CD)+S(AB)$ for some symmetric integral $N\times N$ matrix S changes the right-hand side of (6.4) by $-\frac{\pi^2}{2}\nu^tS\nu$. We believe, but have not checked, that it should be possible to lift the formula (6.4) to a formula giving $\mathbf{V}_\Xi(z)$ modulo $4\pi^2$ rather than just modulo $\pi^2/2$ in terms of H alone by adding to the right-hand side a term $r_M\pi^2/2$ where $\mathbf{e}(r_M/8)$ is the 8th root of unity occurring in the transformation law of Siegel theta series with characteristics under the action of the symplectic matrix $M=\left(\begin{smallmatrix}A&B\\C&D\end{smallmatrix}\right)$ as given by Igusa [56, Theorem 3, p. 182].

We make two small remarks on the above formulas before proceeding. The first is that the term $\frac{\pi}{2}\nu^t(Cu-Dv)$ in (6.4) is needed, not only to make the expression on the right independent of the choice of logarithms u and v modulo $4\pi^2$ (it is already independent of this choice modulo π^2 even if this term is omitted), but in order to get the right imaginary part: the imaginary part of $F(v)$ is $D(z)+\mathrm{Im}(u\bar{v})/2$ for $\mathrm{e}^u=1-\mathrm{e}^v=z$, where $D(z)$ is the Bloch–Wigner dilogarithm as above, and it is only if we include the term with $Cu-Dv$ in (6.4) that its imaginary part has the correct value $\sum_j D(z_j)$. The other is that the vector $w:=Dv-Cu$ whose scalar product with ν gave the correction term in (6.4) also gives a parametrization of the $N\times 2$ matrix (uv) as $\big(B^tA^t\big)w+\mathrm{i}\pi\big(D^t,C^t\big)\nu$. $\big($To see this, just write the relationship of (uv) to ν and w as $M\left(\begin{smallmatrix}u\\-v\end{smallmatrix}\right)=\left(\begin{smallmatrix}\pi\mathrm{i}\nu\\-w\end{smallmatrix}\right)$ and use the formula for M^{-1}.$\big)$ This is simply the logarithmic version of the alternative characterization of the variety V_H given in (6.3), with $w=\log x$. Changing (uv) by $2\pi\mathrm{i}\big(B^tA^t\big)n$ with $n\in\mathbb{Z}^N$ corresponds to taking a different logarithm w of the same x.

We next turn to the relation between half-symplectic matrices and the Bloch group. The latter is an abelian group $\mathcal{B}(F)$ which is defined for any field F of characteristic zero as the quotient of the kernel of the map $d\colon\mathbb{Z}[F]\to\Lambda^2(F^\times)$ sending $[x]$ to $x\wedge(1-x)$ for $x\neq 0,1$ by the subgroup generated by the 5-term relation of the dilogarithm. But the precise definition varies slightly in the literature because of delicate 2- and 3-torsion issues arising from the particular definition of the exterior square (for instance, does one require $x\wedge x=0$ for all x or just $x\wedge y=-y\wedge x$?) and the particular choice of the 5-term relation, which potentially comes in 5^6 versions obtained from one another by replacing each of the 5 arguments by its images under the group generated by $x\mapsto 1/x$ and $x\mapsto 1-x$. In fact, we will need the extended Bloch group as introduced by Neumann [66] and studied further by Zickert and others in [46, 88], but here also there are several versions. We recall the definition from [88] here, and then describe a small refinement and the relation to half-symplectic matrices.

Denote by $\widehat{\mathbb{C}}$ the set of pairs of complex numbers (u,v) with $\mathrm{e}^u+\mathrm{e}^v=1$. This is an abelian cover of $\mathbb{C}^\times\smallsetminus\{0,1\}$ via $z=\mathrm{e}^u=1-\mathrm{e}^v$, with Galois group isomorphic to $\mathbb{Z}^2$. The extended Bloch group $\widehat{\mathcal{B}}(\mathbb{C})$ as defined in [46, 88] is the kernel of the map $\widehat{d}\colon\mathbb{Z}\big[\widehat{\mathbb{C}}\big]\to\Lambda^2(\mathbb{C})$, where $\Lambda^2(\mathbb{C})$ is defined by requiring only $x\wedge y+y\wedge x=0$ (rather than $x\wedge x=0$, which is stronger by 2-torsion) and where $\widehat{d}$ maps $[u,v]:=[(u,v)]\in\mathbb{Z}\big[\widehat{\mathbb{C}}\big]$ to $u\wedge v$, divided by the lifted version of the 5-term relation, namely, the $\mathbb{Z}$-span of the set of elements $\sum_{j=1}^5(-1)^j[u_j,v_j]$ of $\mathbb{Z}\big(\widehat{\mathbb{C}}\big)$ satisfying $(u_2,u_4)=(u_1+u_3,u_3+u_5)$ and $(v_1,v_3,v_5)=(u_5+v_2,v_2+v_4,u_1+v_4)$. There is an extended regulator map from $\widehat{\mathcal{B}}(\mathbb{C})$ to $\mathbb{C}/4\pi^2\mathbb{Z}$ given by mapping $\sum[u_j,v_j]$ to $\sum\mathcal{L}(u_j,v_j)$, where $\mathcal{L}(u,v)=F(v)+\frac{1}{2}uv-\frac{\pi^2}{6}$, which one can check vanishes modulo $4\pi^2$ on the lifted 5-term relation. One can also define $\widehat{\mathcal{B}}(F)$ for any subfield F of $\mathbb{C}$, such as an embedded number field, by replacing $\widehat{\mathbb{C}}$ by the subset $\widehat{F}$ consisting of pairs (u,v) with $\mathrm{e}^u=1-\mathrm{e}^v\in F$. The relation of the Bloch group and the extended Bloch group to algebraic K-theory is that $B(F)$ for any field F is isomorphic up to torsion to the algebraic K-group $K_3(F)$ [71], with the Borel regulator map from $K_3(\mathbb{C})$ to $\mathbb{C}/\pi^2\mathbb{Q}$ being given at the level of the Bloch group by dilogarithms, while the extended Bloch group of a number field $F \subset \mathbb{C}$ is isomorphic to $K_3^{\mathrm{ind}}(F)$ [88], for which the Borel regulator lifts to $\mathbb{C}/4\pi^2\mathbb{Z}$.

We now extend this group slightly by replacing $\mathbb{Z}\big[\widehat{\mathbb{C}}\big]$ by the larger group $\mathbb{Z}\big[\widehat{\mathbb{C}}\big] \oplus \mathbb{C}$ and $\widehat{d}$ by a map from this group to $\Lambda^2(\mathbb{C})/(\mathrm{i}\pi \wedge \mathrm{i}\pi)$, still given on $\mathbb{Z}\big[\widehat{\mathbb{C}}\big]$ by $[u,v] \mapsto u \wedge v$ but now also defined on $\mathbb{C}$ by $\widehat{d}(x) = x \wedge (x + \pi\mathrm{i})$, which despite appearances is a linear map because of the antisymmetry of $\wedge$. We then divide the kernel of this new $\widehat{d}$ by a larger set of relations, namely the same lifted 5-term relation as before (with $\mathbb{C}$ component equal to 0) together with the relations $([u,v]+[v,u]-[u',v']-[v',u'],0)$ and $([u,v]+[-u,v-u+\pi\mathrm{i}],u)$ for all (u,v) and (u',v') in $\widehat{\mathbb{C}}$, corresponding to the elements $[z]+[1-z]$ and $[z]+[1/z]$. The extended regulator map to $\mathbb{C}/4\pi^2\mathbb{Z}$ is now defined by mapping $\big(\sum[u_j,v_j],x\big)$ to $\sum \mathcal{L}(u_j,v_j) - x\pi\mathrm{i}/2$, which agrees with the previous definition when x is 0 and which can be checked to vanish also on the new relations. The advantage of this further extension of the Bloch group is that the solutions (u,v) of the logarithmic Neumann–Zagier equations (i.e., the set of $(u,v) \in \mathbb{C}^{2N}$ with $(u_j,v_j) \in \widehat{\mathbb{C}}$ for each j and $Au - Bv = \pi\mathrm{i}\nu$ with ν as in (6.2)) now give an element of $\widehat{\mathcal{B}}(\mathbb{C})$, namely the class ξ of the pair $\big(\sum_{j=1}^N [u_j,v_j], w\nu^t\big)$, where $w = Cu - Dv$ as before. Using the parametrization $(u,v) = \big(B^t A^t\big)w + \big(C^t D^t\big)\nu\mathrm{i}\pi$ discussed above, we check easily that the image of this in $\Lambda^2(\mathbb{C})$ under $\widehat{d}$ is $\big(\nu^t C D^t \nu\big)(\mathrm{i}\pi)\wedge(\mathrm{i}\pi)$, and its image under the regulator map is precisely the number $\mathbf{V}(z)$ defined in (6.4). When (u,v) comes from a triangulation of a 3-manifold, then the effect of the extended 5-term relation is precisely that of a $(2,3)$-Pachner move (changing one triangulation to another by replacing two tetrahedra with a common face by three tetrahedra with the same set of vertices), so that the element $\xi \in \widehat{\mathcal{B}}(\mathbb{C})$ is a topological invariant of the manifold.

We end this subsection by explaining briefly how half-symplectic matrices actually give a new description of the extended Bloch group as a quotient of Sp_∞ by suitable relations. Here for convenience we are writing Sp_N rather than Sp_{2N} for the group of symplectic matrices of size $2N \times 2N$ over $\mathbb{Z}$, and Sp_∞ for the direct limit of these groups with respect to the natural inclusions $\mathrm{Sp}_N \hookrightarrow \mathrm{Sp}_{N+1}$. It also turns out to be more convenient to define Sp_N as the space of matrices M satisfying $MJ_N^* M^t = J_N^*$ instead of $MJ_N M^t = J_N$ used above, where $J_N = \big(\begin{smallmatrix} 0 & -1_N \\ 1_N & 0 \end{smallmatrix}\big)$ and J_N^* is the block diagonal matrix with N copies of J_1 on the diagonal, in which case the inclusion just sends M to $M^+ = \big(\begin{smallmatrix} M & 0 \\ 0 & 1_2 \end{smallmatrix}\big)$, and similarly the lifted 5-term relations become much simpler with this convention. The relations that we divide by are roughly as follows. The first is stability (identify $[M]$ and $[M^+]$). A second is that we identify $[M]$ and $\big[\big(\begin{smallmatrix} 1 & 0 \\ S & 1 \end{smallmatrix}\big)M\big]$ with S integral and symmetric are equivalent. (This corresponds to working with half-symplectic rather than full symplectic matrices.) A third is that we identify $M \in \mathrm{Sp}_N$ with $\big(\begin{smallmatrix} g & 0 \\ 0 & g^{t-1} \end{smallmatrix}\big)M$ for any $g \in \mathrm{GL}_N(\mathbb{Z})$. (This corresponds to permuting the N relations (6.1) or multiplying one of them by a monomial in the others.) A fourth is to identify M with $M\big(\begin{smallmatrix} P & 0 \\ 0 & P \end{smallmatrix}\big)$ for any $N \times N$ permutation matrix P, corresponding in the geometric case to changing the numbering of the N simplices, and yet another (which maybe can be omitted) corresponds to relabelling the edges so that the shape parameter z goes to z' or z''. The main one, of course, is a symplectic-matrix version of the 5-term relation. This was first discovered in the special case corresponding to the Nahm sums (6.5) by Sander Zwegers in an unpublished 2011 conference talk and then given in various versions for arbitrary symplectic matrices by Dimofte and the first author in [14] and in unpublished work by Campbell Wheeler and Michael Ontiveros (MPIM). The set of equivalence classes becomes an abelian group by setting $[M]+[M']$ equal to the class of $\big(\begin{smallmatrix} M & 0 \\ 0 & M' \end{smallmatrix}\big)$ and $-[M]$ to the class of M^{-1}. To get a map from this group to the extended Bloch group of $\mathbb{C}$, we have to first enlarge it by looking at equivalence classes, not just of half-symplectic matrices H (which is enough by the second of the equivalence relations listed above), but of pairs consisting of a half-symplectic matrix $H = (AB)$ together with a solution $(u,v) \in \mathbb{C}^{2N}$ of the logarithmic NZ equations $Au - Bv = \pi\nu$ with ν as in (6.2), with corresponding lifts of the 5-term and of the various other relations. The map from this larger group to the extended Bloch group is then the one described in the previous paragraph. It is injective because the 5-terms relations defining the extended Bloch group all lift to corresponding relations at the (half-) symplectic level. It is also surjective, as one can show using elements of the set $\mathrm{Sp}_{N,N'} = \{M \in M_{N\times N'}(\mathbb{Z}) \mid MJ_{2N}M^t = J_{2N'}\}$ of "non-square symplectic matrices" (note that this set is just Sp_N if $N = N'$ and reduces to 0 if $N' > N$) together with the obvious composition maps $Sp_{N,N'} \times \mathrm{Sp}_{N',N''} \to \mathrm{Sp}_{N,N''}$, in order to eliminate superfluous relations. (Roughly speaking, if $\sum_{j=1}^{N}[u_j, v_j]$ is the $\mathbb{Z}[\widehat{\mathbb{C}}]$-component of an element of $\widehat{\mathcal{B}}(\mathbb{C})$ as defined above, then we define $N' \le N$ as the rank of the group generated by all u_j and v_j and obtain an element of $Sp_{N,N'}$ by writing the u's and v's in terms of these generators, which then always satisfy a collection of NZ equations.) A more detailed discussion of this and of the whole relationship between half-symplectic matrices and Bloch groups, including our versions of the 5-term relation lifted to symplectic and half-symplectic matrices, is also given in [43].

This concludes our discussion of half-symplectic matrices and the equations (6.1). These objects arise in (at least) two different contexts, in 3-dimensional topology and in the study of special q-hypergeometric series (Nahm sums). The former is of course the one that is of most relevance for this paper, and will be discussed in more detail in the next subsection, but after that we will also say something about Nahm sums because they will play a role in the sequel [44] to this paper and also because they give the most elementary approach to defining the associated formal power series that are our main subject of interest.

6.2 Ideal triangulations and the Neumann–Zagier equations

In 3-dimensional geometry, the shape of an ideal tetrahedron in $\mathbb{H}^3$ is encoded by a complex number ("shape parameter") $z \in \mathbb{C} \smallsetminus \{0, 1\}$, the tetrahedron being isometric to the convex hull of the four points $0, 1, \infty, z \in \mathbb{P}^1(\mathbb{C}) = \partial(\mathbb{H}^3)$. The shape z has three forms z, $z' = 1/(1-z)$ and $z'' = 1 - 1/z$, each corresponding to the choice of a pair of opposite edges of the tetrahedron as shown in Figure 5.

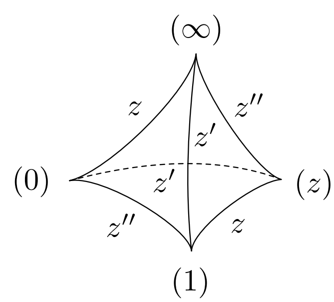

Figure 5. A tetrahedron with shape parameters.

An ideal triangulation of a 3-manifold with torus boundary components give rise to an equation (6.1), where the variables z_i solving the equations (6.1) are the shape parameters of the tetrahedra and the equations are the "gluing conditions" relating the shape parameters of the tetrahedra incident on the various edges of the triangulation and on the cusp. These gluing equations originated in the work of Thurston [72] and further studied in [67], where the key symplectic property of the matrices (AB) was found. We explain very briefly how this works for 3-manifolds whose boundary component is a torus, equipped with an ideal triangulation with N tetrahedra. Each edge of the triangulation gives rise to a gluing equation asserting that the product of the shape parameters of all tetrahedra incident to that edge equals to 1. Every peripheral curve (i.e., a curve in the boundary torus of the 3-manifold) also has an equation of this form (often called the holonomy equation, following Thurston), obtained by setting the product of the shape parameters as the curve intersects the triangulated boundary, equal to 1. The product of the gluing equations corresponding to all edges is identically 1, so one gluing equation is redundant and can be removed and replaced by the holonomy equation of a nontrivial peripheral curve. Since there are N edges, this gives a collection of N gluing equations. If one is interested in the geometric solution that describes the complete hyperbolic structure, where all the shape parameters have positive imaginary part, the above gluing equations are replaced by their stronger logarithmic form, where the right-hand side is now $2\pi\mathrm{i}$ for each edge and 0 for the peripheral curve. Using the fact that the three shape parameters z, $z'=1/(1-z)$ and $z''=1-1/z$ satisfy the relation $zz'z''=-1$, and in logarithmic form $\log z+\log z'+\log z''=\pi\mathrm{i}$, it follows that we can eliminate one of the three variables at each tetrahedron (after choosing a pair of opposite edges for each tetrahedron). Doing so, the logarithmic form of the gluing equations now become linear equations for $\log z_i$ and $\log(1-z_i)$, whose coefficients give rise to the Neumann–Zagier matrices A and B, and where right-hand side is a distinguished flattening ν that should satisfy the mod 2 congruence given in (6.2). (This congruence can presumably be deduced from the "parity condition" for ideal triangulations proved by Neumann [65], but we have not checked this.) Neumann–Zagier's theorem is that the above matrix $(A|B)$ is the upper half of a symplectic matrix with integer entries. Note that the corresponding pairs (H,z) and (Ξ,z) are called "NZ datum" and "enhanced NZ datum" in [14]. Note also that a different choice of opposite edges in each tetrahedron cyclically permutes the triple (z_j,z_j',z_j'') and changes the corresponding Neumann–Zagier matrices, but does not change the corresponding element of $\widehat{\mathcal{B}}(\mathbb{C})$.

The connection between gluing equations and symplectic matrices involves not only the shapes of ideal tetrahedra, but also their Ptolemy variables. The latter is an assignment of nonzero complex numbers x_i at each edge of an ideal triangulation that satisfy the Ptolemy equations, namely at each tetrahedron we have a quadratic equation $x_1x_2\pm x_3x_4\pm x_5x_6=0$ (with suitable signs). The signs require either ordered triangulations or a choice of a Ptolemy cocycle and a detailed description is given in [40, equation 12.2] and also in [26, Section 3]. (The equivalence between the shape and the Ptolemy description of character varieties of surfaces is discussed in detail by Fock–Goncharov [21].) In the 3-dimensional case of a knot complement, these x_j are exactly the ones introduced in (6.3) (and related to $w=Du-Cv$ by $x_j=\mathrm{e}^{w_j}$), which here becomes a system of quadratic trinomial relations after rescaling because in each of the column of the gluing equation matrices there are at most six non-zero entries, corresponding to the six edges of the tetrahedron corresponding to that column.

We mention in passing that the variety defined by just the first $N-1$ edge gluing equations is 1-dimensional (for a suitably chosen triangulation) and that this curve maps to the $\mathrm{PSL}_2(\mathbb{C})$-character variety (via the developing map which assigns a solution to the gluing equations a $\mathrm{PSL}_2(\mathbb{C})$-representation of the fundamental group of the manifold, well-defined up to conjugation). The $\mathrm{PSL}_2(\mathbb{C})$-character variety maps to $\mathbb{C}^*\times\mathbb{C}^*$ (modulo a $\mathbb{Z}/2^2$ quotient) and its image is described by the vanishing of the A-polynomial $A(\ell,m)$ (where ℓ is the longitude) as introduced subsequently in [11]. The variety obtained by adding taking the first $N-1$ relations together with the relation $m^p\ell^q=1$ for coprime integers p and q corresponds to the compact 3-manifold obtained by doing a (p,q) Dehn surgery on the knot complement. A detailed discussion of the choices involved to write down these matrices can be found in [14, Section 2, Appendix A]. All of this data is standard in knot theory, and is computed explicitly for any given knot complement (or more generally, an ideal triangulation of a cusped hyperbolic 3-manifold) by the computer implementation of `SnapPy` [12].

Once an ideal triangulation Δ of a 3-manifold M as above has been fixed, a solution z of its gluing equations gives rise via a developing map to a representation ρ_z (i.e., a group homomorphism) of $\pi_1(M)$ in $\mathrm{PSL}_2(\mathbb{C})$, well-defined up to conjugation. If we choose the Neumann–Zagier equations as above, the representation ρ_z is boundary-parabolic and gives rise to an element of the extended Bloch group [87] and has a well-defined complex volume; see [66] and also [40]. Thus, if Δ is an ideal triangulation of a the complement of a knot K, we have a map $z \mapsto \rho_z$ from $V_H(\mathbb{C})$ to $\mathcal{P}_K$, and the complex volume of $\rho_z \in \mathcal{P}_K$ coincides with the complex volume of z, as follows from the work of Neumann [66] and Zickert [88] on the extended Bloch group.

There are, however, several subtleties of the above construction which we should point out. For instance, there exist triangulations of hyperbolic knots for which the map $V_H(\mathbb{C}) \to \mathcal{P}_K^{\mathrm{red}}$ is not onto or even for which the complex solutions set $V_H(\mathbb{C})$ is empty (this can happen even for triangulations of the complement of the 4_1 knot). In this paper, we will ignore these issues and assume that we are dealing with ideal triangulations for which the map is onto. We will further restrict our attentions to knots for which the set $\mathcal{P}_K$ is finite. (There are known to be knots for which the variety $\mathcal{P}_K$ has strictly positive dimension, but they are too complicated for the calculations in this paper to be carried out. We believe that in such cases the right indexing set of our formal power series would be the set of components of $\mathcal{P}_K$ or of the variety $V_H(\mathbb{C})$.)

An alternative approach to the definition of the set $\mathcal{P}_K$ comes from the branches of the A-polynomial curve above the point $m = 1$, where m is the eigenvalue of the meridian. Even if the $\mathrm{SL}_2(\mathbb{C})$ character variety of decorated representations of a knot complement has positive-dimensional components, its image in $\mathbb{C}^* \times \mathbb{C}^*$, as given by the eigenvalue of the meridian and the longitude, is one-dimensional, and (ignoring any zero-dimensional components) is defined by the zeros of the A-polynomial of the knot. The A-polynomial is discussed in detail in the appendix of [8]. We will focus on knots that satisfy the property that the number of parabolic representations σ coincides with the degree of the A-polynomial of a knot with respect to the longitude. Note that the Galois group $\mathrm{Gal}\big(\overline{\mathbb{Q}}/\mathbb{Q}\big)$ acts on the set $\mathcal{P}_K$ of boundary parabolic representations. What's more, in a boundary parabolic representation, the longitude has eigenvalue ± 1 and this partitions the set $\mathcal{P}_K$ into two subsets $\mathcal{P}_K^{\pm}$, each of which is stable under the action of $\mathrm{Gal}\big(\overline{\mathbb{Q}}/\mathbb{Q}\big)$. The geometric representation lies in $\mathcal{P}_K^-$; see [9, Lemma 2.2].

6.3 Nahm sums and the perturbative definition of the Φ-series

In this final subsection, we describe how to attach to $\Xi = ((AB), \nu)$ as in (6.2), a solution z of the equation (6.1) and a number $\alpha \in \mathbb{Q}$ a completed formal power series belonging to $\mathrm{e}^{\mathbf{V}(z)/\mathrm{den}(\alpha)^2 h}\mathbb{C}[[h]]$. As already stated in Section 2, this was done in [14] (for $\alpha = 0$) and [15] (for general α) in the context of knot complements and Neumann–Zagier data. However, there is a completely different situation where the same formal power series are attached to the same data (Ξ, z, α), namely the asymptotics near roots of unity of special q-hypergeometric series called Nahm sums. Since these are a little more elementary we will use them to explain the derivation of the formal power series.

We begin by recalling what Nahm sums are. The simplest one is defined by

$$F_{A,b}(q) = \sum_{n_1,\dots,n_N \geq 0} \frac{q^{\frac{1}{2}n^tAn + b^tn}}{(q;q)_{n_1}\cdots(q;q)_{n_N}} \in \mathbb{Z}[[q]], \tag{6.5}$$

where A is an even positive definite symmetric matrix in $M_N(\mathbb{Z})$ and b an element of $\mathbb{Z}^N$. Changing each n_j by 1, we see that the stationary points of the summand (i.e., the places where nearby terms are asymptotically equal, giving the expected main contributions to the whole sum) are given in the limit $q \to 1$ by $q^{n_j} = z_j + o(1)$, where $z = (z_1, \dots, z_N)$ is a solution of Nahm's equation $1 - z = z^A$ (which is the special case $B = \mathbf{1}_N$ of (6.1), with $(A\mathbf{1}_N)$ being half-symplectic). Nahm observed that for any solution z of this equation the element $\sum_i [z_i]$ belongs to the Bloch group $\mathcal{B}(\mathbb{C})$ and conjectured that $F_{A,b}(q)$ (up to a rational power of q, and considered as a function of τ with $q = \mathbf{e}(\tau)$) can only be a modular function if at least one solution of the Nahm equation has a trivial class in the Bloch group, and conversely that $F_{A,b}$

(again up to a power of q and as a function of τ) is a modular function of τ for *some* b if all solutions of the Nahm equation have trivial class in the Bloch group. The first assertion was proved in [10]; the second is still open.

If we now generalize the Nahm sum to

$$F_\Xi(q) = \sum_{n\in\mathbb{Z}^N,B^tn\geq 0} \frac{(-1)^{\nu^t n} q^{\frac{1}{2}(n^tAB^tn+\nu^tn)}}{(q;q)_{(B^tn)_1}\cdots(q;q)_{(B^tn)_N}} = \sum_{\substack{m,n\in\mathbb{Z}^N_{\geq 0}\times\mathbb{Z}^N\\ m=B^tn}} \frac{(-1)^{\nu^t n} q^{\frac{1}{2}(n^tAm+\nu^tn)}}{(q;q)_{m_1}\cdots(q;q)_{m_N}}, \tag{6.6}$$

with $\Xi = ((AB),\nu)$ as in (6.2), which is still a power series in q because of the congruence condition on ν, then the same consideration as before shows that the stationary points of the sum correspond via $z = q^{B^tn}$ to the solutions of the equation (6.1). A formal computation of the contribution of the summands near these stationary point will lead to the perturbative series in h that we are looking for, where $q = \mathrm{e}^{-h}$ with $h \to 0$, and in some cases one can show that these h-series actually do describe the radial asymptotics of the Nahm sum (see [42], where this is shown to be the case for the original Nahm sum (6.5) and the real solution of the Nahm equation), but in general the calculation is purely formal because unless all z_i are between 0 and 1 the series corresponding to h will not correspond to any subsum of (6.5) or (6.6). We also mention that there are even more general Nahm sums whose n-th summand $\left(n \in \mathbb{Z}^N\right)$ is the product of a root of unity, a power of q given by a quadratic function of n, and a product of Pochhammer symbols (possibly to integer powers) with linear forms in n as arguments, which occur in several places in quantum topology, e.g., the 3D-index [13] and many of the q-series in [44].

We now explain how to associate to the datum $\Xi = (H,\nu)$ and point z on V_H a formal power series $\Phi_\alpha^{(\Xi,z)}(h)$ for each $\alpha \in \mathbb{Q}/\mathbb{Z}$. We will do this first for the easier case $\alpha = 0$, and then discuss how the formula changes in the general case. The calculations for $\alpha = 0$ were done first for the simplest Nahm sum (6.5) in [83] and [75] and for general half-symplectic matrices (AB) (though under the assumption that B is invertible over $\mathbb{Q}$) in the context of knots in [14]. The power series that were obtained in these two different contexts were syntactically identical, and this coincidence persisted for general α, with the perturbative series of [15] being syntactically equal to the asymptotics of Nahm sums at roots of unity [42, Section 5], with the formulas in all cases being given in terms of (sums of) formal Gaussian integrals. It is for this reason that we can use the easier Nahm sums to motivate the precise form of the integral to be studied. We only sketch the argument, referring to the papers above for more details.

We begin by rewriting the first definition in (6.6) in the form

$$F_{(AB),\nu}(q) = \frac{1}{(q;q)_\infty^N} \sum_{n\in\mathbb{Z}^N} (-1)^{\nu^tn} q^{\frac{1}{2}(n^tAB^tn+\nu^tn)} \prod_{j=1}^N \left(q^{(B^tn)_j+1};q\right)_\infty, \tag{6.7}$$

where we no longer have to restrict to n with $B^tn \geq 0$ because $\left(q^{m+1};q\right)_\infty$ vanishes for $m \in \mathbb{Z}_{<0}$. We must assume for now that the symmetric matrix AB^t is positive definite to ensure the convergence of the series (6.6) or (6.7), but this is not important at the end since the final formulas will be purely algebraic and make sense without this assumption. The key point is that if $q = \mathrm{e}^{-h}$ with h small then the sum will be approximated to all orders by the corresponding integral, with $\mathbb{Z}^N$ replaced by $\mathbb{R}^N$ and the summation sign by an integral sign. (This is a consequence of the Poisson summation formula, which represents the sum over $\mathbb{Z}^N$ of a sufficiently smooth function of sufficiently rapid decay as the sum over $\mathbb{Z}^N$ of its Fourier coefficient, whose constant term is the integral corresponding to the original sum and whose other terms are of smaller order.) We then look at the expansion of the integrand around its stationary points and approximate each by a Gaussian times a power series in a small local variable, as is always done in perturbation theory. The stationary points are indexed by the complex points z of V_H, as already indicated, the correspondence being given by $q^{(B^t n)_j} \sim z_j$. On the other hand, for $z \in \mathbb{C}^*$, $q = \mathrm{e}^{-h}$ with h tending to 0, and t either fixed or growing more slowly than any power of $1/h$, we have the asymptotic formula

$$\begin{aligned}\frac{1}{\left(z\mathrm{e}^{t\sqrt{h}};q\right)_\infty} &\sim \exp\left(\sum_{m=0}^{\infty} \frac{B_m\left(t/\sqrt{h}\right)}{m!}\mathrm{Li}_{2-m}(z)h^{m-1}\right)\\ &= \exp\left(\frac{\mathrm{Li}_2(z)}{h} + \left(\frac{t}{\sqrt{h}} - \frac{1}{2}\right)\log\left(\frac{1}{1-z}\right) + \frac{t^2}{2}\frac{z}{1-z} + (\mathrm{small})\right), \qquad (6.8)\end{aligned}$$

where $B_m(t)$ denotes the m-th Bernoulli polynomial and "(small)" is an explicit power series in t and $\sqrt{h}$ with no constant term in $\sqrt{h}$. (The first statement is [83, Lemma, p. 53] and the second follows because all contributions from $B_m\left(t/\sqrt{h}\right)$ with $m \geq 2$ except for the quadratic part of the B_2-term are small.) Inserting this into the parts near the stationary points of the integral corresponding to the sum (6.7), we find after some calculation that the total contribution of the stationary part corresponding to a given solution z of (6.1) is $\mathrm{e}^{\mathbf{V}(z)/h}$ times an explicit power series in h (initially in $\sqrt{h}$, but then in h because of the parity properties of Bernoulli polynomials) which is written out in [14]. We only mention here that the power series obtained has coefficients in $\mathbb{Q}(z)$ (and hence in F_σ in our application to knots) except for a prefactor $\det(A + B\,\mathbf{diag}(z_j/(1-z_j)))^{-1/2}$ coming from the determinant of the quadratic part of the Gaussian.

When $\alpha = a/c$ is not integral, the calculations, done in [15] in the general case (still with B invertible) and in [42] for the special Nahm sums (6.5), are much more complicated and we refer to those papers for the explicit formulas. A key point is that the stationary points of the integral are now indexed by the c-th roots of the solutions z of (6.1), but with the quadratic form appearing in the Gaussian depending only on z and not on the choice of c-th root. This means that each of the formal power series $\Phi_\alpha^{(\Xi,z)}(h)$ has the form of a sum over $(\mathbb{Z}/c\mathbb{Z})^N$ (after choosing some fixed $\sqrt[c]{z}$) of expressions similar to those occurring for the simpler case $c = 1$. The reader can get a feeling for the nature of the formulas appearing by looking at Section 8 of this paper, where they are carried out in detail for the Kashaev invariant of the 4_1 knot, this case however being deceptively simple because of the positivity of all of the terms occurring.

We make one final remark. The specific formulas given in [14, 15] gave only the series $\Phi_\alpha^{(K,\sigma)}$ as discussed in Section 2, i.e., only the first column of our matrix $\mathbf{\Phi}^{(K)}$, because the vector ν was always assumed to be the one coming from the geometric "flattening". By varying ν, one can get the other columns of $\mathbf{\Phi}$. This variation produces a q-holonomic system that turns out to be closely related to the ones for the generalized Kashaev invariants that will be discussed in Section 7.1. This will be the theme of the next section.

7 Two q-holonomic modules

In Part I we were led by the refined quantum modularity conjecture to find an entire matrix $\mathbf{J}^{(K)}$ of Habiro-like functions generalizing the Kashaev invariant, the first column being the vector of formal power series found in [14, 15]. In this section, we will study the structure of the other columns of this matrix and will see that they have a natural "q-holonomic structure" in terms of an infinite collection of functions that satisfy a recursion of finite length and hence all lie in a finite-dimensional module. In Section 7.1, we explain this in detail for the case of the 4_1 knot, where explicit formulas for the entries of $\mathbf{J}^{(K)}$ were already given in Part I. In the next subsection, we give the corresponding formulas for the 5_2 knot, where the matrix in question has size 4×4 instead of 3×3. These are considerably more complicated than for the 4_1 knot and have the interesting new feature that Dedekind sums appear. In Section 7.3, we explain how these formulas could be guessed. This ansatz involves studying the q-hypergeometric series defining the original Kashaev invariant via stationary points and formal Gaussian summation, analogous to what was done in Section 6.3 for Nahm sums and what will be done in Section 8 for the Kashaev invariant of the 4_1 knot. In the final subsection, we discuss the point already alluded to at the end of Section 6 that the power series in h studied there have a q-holonomic structure with the *same* coefficients as the one associated to the matrix $\mathbf{J}$. This is for the momenta purely experimental and is one of the many mysteries associated with the subject. In Section 7.4, we also briefly mention two further conjectural objects associated to knots (or more generally to half-symplectic matrices) that we believe share the same q-holonomic structure.

7.1 Descendant Habiro-like functions

It turns out that the first row of $\mathbf{J}$ and the first row of $\widehat{\mathbf{\Phi}}$ are basis elements of the span of an inhomogeneous recursion, and the same holds (but now with the corresponding homogeneous recursion) for each of the remaining rows of $\mathbf{J}$ and of $\widehat{\mathbf{\Phi}}$ as well as for the matrix of q-series of [44]. To illustrate how this works, we give the complete formulas for the matrix for the 4_1 knot. The corresponding formulas for the 5_2 knot, which are considerably more complicated and illustrate several further refinements (like the appearance of Dedekind sums), will be given in Section 7.3.

Collecting together our previous results for the 4_1 knot for the reader's convenience, we obtain that the matrix $\mathbf{J} = \mathbf{J}^{(4_1)}$ of periodic functions on $\mathbb{Q}$ has the form

$$\mathbf{J}(x) = \begin{pmatrix} 1 & J^{(0,1)}(x) & J^{(0,2)}(x) \\ 0 & J^{(1,1)}(x) & J^{(1,2)}(x) \\ 0 & J^{(2,1)}(x) & J^{(2,2)}(x) \end{pmatrix} = \mathcal{J}(q) = \begin{pmatrix} 1 & \mathcal{J}^{(0,1)}(q) & \mathcal{J}^{(0,2)}(q) \\ 0 & \mathcal{J}^{(1,1)}(q) & \mathcal{J}^{(1,2)}(q) \\ 0 & \mathcal{J}^{(2,1)}(q) & \mathcal{J}^{(2,2)}(q) \end{pmatrix}$$

(with $q = \mathbf{e}(x)$ and omitting K as usual), where the elements of the first row are given by

$$\begin{aligned} \mathcal{J}^{(0,1)}(q) &= \mathcal{Q}_1^{(4_1)}(q) = \mathcal{J}^{(4_1)}(q) = \sum_{n=0}^{\infty} (q;q)_n \big(q^{-1};q^{-1}\big)_n, \\ \mathcal{J}^{(0,2)}(q) &= \mathcal{Q}_2^{(4_1)}(q) = \frac{1}{2}\sum_{n=0}^{\infty} \big(q^{n+1} - q^{-n-1}\big)(q;q)_n \big(q^{-1};q^{-1}\big)_n \end{aligned} \tag{7.1}$$

(equations (2.6) and (4.5)), with $\mathcal{Q}_i^{(4_1)}(q)$ being the elements of the Habiro ring defined and tabulated in Section 4.3, and that the elements of the other two rows are given by

$$\begin{aligned} \mathcal{J}^{(1,1)}(q) &= \frac{1}{\sqrt{c}\sqrt[4]{3}} \sum_{Z^c=\zeta_6} \prod_{j=1}^{c} \big|1 - q^j Z\big|^{2j/c}, \\ \mathcal{J}^{(2,1)}(q) &= \frac{\mathrm{i}}{\sqrt{c}\sqrt[4]{3}} \sum_{Z^c=\zeta_6^{-1}} \prod_{j=1}^{c} \big|1 - q^j Z\big|^{2j/c}, \\ \mathcal{J}^{(1,2)}(q) &= \frac{1}{2\sqrt{c}\sqrt[4]{3}} \sum_{Z^c=\zeta_6} \big(Zq - Z^{-1}q^{-1}\big) \prod_{j=1}^{c} \big|1 - q^j Z\big|^{2j/c}, \\ \mathcal{J}^{(2,2)}(q) &= \frac{\mathrm{i}}{2\sqrt{c}\sqrt[4]{3}} \sum_{Z^c=\zeta_6^{-1}} \big(Zq - Z^{-1}q^{-1}\big) \prod_{j=1}^{c} \big|1 - q^j Z\big|^{2j/c} \end{aligned} \tag{7.2}$$

(equations (3.2), (4.9) and the accompanying text). The syntactical similarity between equations (7.1) and equations (7.2) is striking, and leads directly to the q-holonomy.

To see this, we rewrite the two formulas in (3.2) as

$$\mathcal{J}^{(0,1)}(q) = \mathcal{H}_0^{(0)}(q), \qquad \mathcal{J}^{(0,2)}(q) = \frac{1}{2}\big(q\mathcal{H}_1^{(0)}(q) - q^{-1}\mathcal{H}_{-1}^{(0)}(q)\big), \tag{7.3}$$

where $\{\mathcal{H}_m^{(0)}(q)\}_{m\in\mathbb{Z}}$ is the sequence of elements of the Habiro ring defined by

$$\mathcal{H}_m^{(0)}(q) = \sum_{n=0}^{\infty}(q;q)_n\big(q^{-1};q^{-1}\big)_n q^{mn}, \qquad m\in\mathbb{Z}. \tag{7.4}$$

It is easy to see that this sequence satisfies the recursion relation

$$q^{m+1}\mathcal{H}_{m+1}^{(0)}(q) + (1-2q^m)\mathcal{H}_m^{(0)}(q) + q^{m-1}\mathcal{H}_{m-1}^{(0)}(q) = 1, \qquad m\in\mathbb{Z} \tag{7.5}$$

(a similar, but homogeneous, recursion relation for the descendants of certain q-series associated to the 4_1 knot was given in [27, equation (14)] and used in [44]) and also that the $\mathbb{Q}[q^{\pm}]$-module they span is free of rank 3 with the top row of the matrix $\mathbf{J}$ as a basis. If we now introduce two further sequences of functions of q (or periodic functions of x, where $q=\mathbf{e}(x)$) by

$$\begin{aligned}\mathcal{H}_m^{(1)}(q) &= \frac{1}{\sqrt{c}\sqrt[4]{3}}\sum_{Z^c=\zeta_6} Z^m\prod_{j=1}^{c}\big|1-q^jZ\big|^{2j/c},\\ \mathcal{H}_m^{(2)}(q) &= \frac{\mathrm{i}}{\sqrt{c}\sqrt[4]{3}}\sum_{Z^c=\zeta_6^{-1}} Z^m\prod_{j=1}^{c}\big|1-q^jZ\big|^{2j/c},\end{aligned} \tag{7.6}$$

then (7.2) says that the non-trivial elements of the second and third rows of $\mathbf{J}$ are given by

$$\mathcal{J}^{(i,1)}(q) = \mathcal{H}_0^{(i)}(q), \qquad \mathcal{J}^{(i,2)}(q) = \frac{1}{2}\big(q\mathcal{H}_1^{(i)}(q) - q^{-1}\mathcal{H}_{-1}^{(i)}(q)\big), \qquad i=1,2.$$

Furthermore, we see that the first column of the matrix $\mathbf{J}$, trivial though it is, nevertheless belongs to the same q-holonomic module as the other columns, since as well as equations (7.3) and (7.6) we also have the relation

$$\mathcal{J}^{(i,0)}(q) = q\mathcal{H}_1^{(i)}(q) + \mathcal{H}_0^{(i)}(q) + q^{-1}\mathcal{H}_{-1}^{(i)}(q), \qquad i=0,1,2, \tag{7.7}$$

as we see by specializing the recursion (7.5) and its counterparts for $\mathcal{H}_m^{(1)}$ and $\mathcal{H}_m^{(2)}$ to $m=0$. Then the quantitative version of the "syntactical similarity" noted above is that we can write the formulas (7.1) and (7.2) or (7.6) and (7.7) uniformly and more compactly in matrix form as

$$\mathcal{J}^{(4_1)}(q) = \begin{pmatrix}\mathcal{H}_{-1}^{(0)}(q) & \mathcal{H}_0^{(0)}(q) & \mathcal{H}_1^{(0)}(q)\\ \mathcal{H}_{-1}^{(1)}(q) & \mathcal{H}_0^{(1)}(q) & \mathcal{H}_1^{(1)}(q)\\ \mathcal{H}_{-1}^{(2)}(q) & \mathcal{H}_0^{(2)}(q) & \mathcal{H}_1^{(2)}(q)\end{pmatrix}\begin{pmatrix}q^{-1} & 0 & \frac{1}{2}q\\ 1 & 1 & 0\\ q & 0 & -\frac{1}{2}q^{-1}\end{pmatrix}. \tag{7.8}$$

Note that none of these equations are unique, since any one of them could be written in infinitely many other ways by using the q-holonomy property, e.g., we could specialize the recursions to any value of m other than 0 to get formulas for $\mathcal{J}^{(i,0)}(q)$ different from (7.7). Similarly, we could rewrite (7.8) by taking three other columns (or linear combinations of columns) of the $\mathcal{H}$-matrix for the first factor on the right, with the corresponding new matrix of Laurent polynomials in the second factor.

More interesting is that there is also nothing sacred about the particular collection $\mathcal{H}^{(i)}(q)$ of functions of q that we chose to define our q-holonomic system, and that there infinitely many other collections, even with completely different indexing sets (e.g., $\mathbb{Z}^2$ instead of $\mathbb{Z}$) that could be used instead and that might have been found it we had given a different combinatorial description of knot. However, the module over $\mathbb{Q}[q,q^{-1}]$ that they generate is at least conjecturally intrinsic to the knot and is simply the span of the columns of $\mathbf{J}^{(K)}$, which therefore constitute a canonical basis indexed by $\mathcal{P}$. This is one of the most mysterious aspects of our matrix invariants. We will return to it in at the end of this section in connection with other possible representations of the same abstract q-holonomic module.

We end the subsection with a final remark. Despite the apparent similarity in the formulas for the elements of the first row and all other rows of the matrix $\mathbf{J}$, there is a crucial difference between formulas like (7.1) or (7.4) for the top rows of our matrix and formulas like (7.2) or (7.6) for the other rows: the former are sums over the lattice points of a cone and hence satisfy an inhomogeneous linear q-difference equation, whereas the latter are sums over periodic groups $\mathbb{Z}/c\mathbb{Z}$ and hence have no boundary terms and satisfy a homogeneous equation. Another difference, to which we hope to return in [38] in the context of Habiro rings for general number fields, is that (7.1) and (7.4) obviously give algebraic integers when q is a root of unity, whereas (7.2) or (7.6) give algebraic integers in some non-evident way, since it is not obvious (but in fact true) that the sums in these formulas are divisible by $\sqrt{c}$. We will find exactly the same behavior for the elements of the $\mathbf{J}$-matrix for the 5_2 knot in the next subsection.

7.2 The J-matrix for the 5_2 knot

In this subsection, we describe that analogues of the formulas just given for our second standard knot 5_2, because as usual the figure 8 knot has such special properties that some of the interesting features are obscured.

The Kashaev invariant of the 5_2 knot is given by

$$\mathcal{J}^{(5_2)}(q)=\sum_{m=0}^{\infty}\sum_{k=0}^{m}q^{-(m+1)k}\frac{(q;q)_m^2}{\left(q^{-1};q^{-1}\right)_k}. \tag{7.9}$$

(See [59, equation 2.3].) This is manifestly an element of the Habiro ring. We generalize it to the two-parameter family of elements of the Habiro ring given by

$$\mathcal{H}_{a,b}^{(0)}(q)=\sum_{m=0}^{\infty}\sum_{k=0}^{m}q^{-(m+1)k+am+bk}\frac{(q;q)_m^2}{\left(q^{-1};q^{-1}\right)_k},\qquad a,b\in\mathbb{Z}. \tag{7.10}$$

These again form a q-holonomic module in the sense of [77], meaning that they satisfy recursions like (7.5) (though in this case more complicated, and omitted here) and hence generate a $\mathbb{Q}\left[q,q^{-1}\right]$-module of finite rank. Here the rank is 4 and the q-holonomic module is generated (as we expect to hold for every knot) by the first row of the matrix $\mathbf{J}$ of the knot,

$$\mathcal{J}^{(0,0)}(q)=-\mathcal{H}_{0,0}^{(0)}(q)+q^{-1}\mathcal{H}_{-1,0}^{(0)}(q)+\mathcal{H}_{0,-1}^{(0)}(q)=1,\qquad \mathcal{J}^{(0,1)}(q)=\mathcal{H}_{0,0}^{(0)}(q),$$
$$\mathcal{J}^{(0,2)}(q)=\mathcal{H}_{0,0}^{(0)}(q)-q^{-1}\mathcal{H}_{-1,0}^{(0)}(q),\qquad \mathcal{J}^{(0,3)}(q)=2\mathcal{H}_{0,0}^{(0)}(q)-q^{-1}\mathcal{H}_{-1,0}^{(0)}(q)+\mathcal{H}_{-1,1}^{(0)}(q).$$

Just as in the case of the 4_1 knot, we find that the further three rows are given by the *same* linear combinations of three other two-parameter families $\mathcal{H}_{a,b}^{(i)}$ $(1\le i\le 3)$ of functions, i.e., we have

$$\mathcal{J}^{(i,0)}(q)=-\mathcal{H}_{0,0}^{(i)}(q)+q^{-1}\mathcal{H}_{-1,0}^{(i)}(q)-\mathcal{H}_{0,-1}^{(i)}(q),\qquad \mathcal{J}^{(i,1)}(q)=\mathcal{H}_{0,0}^{(i)}(q),$$
$$\mathcal{J}^{(i,2)}(q)=\mathcal{H}_{0,0}^{(i)}(q)-q^{-1}\mathcal{H}_{-1,0}^{(i)}(q),\qquad \mathcal{J}^{(i,3)}(q)=2\mathcal{H}_{0,0}^{(i)}(q)-q^{-1}\mathcal{H}_{-1,0}^{(i)}(q)+\mathcal{H}_{-1,1}^{(i)}(q)$$

for $i=0,1,2,3$. The formulas for the functions for $i\ne 0$, whose origin will be indicated in Section 7.3, are of the same type as the corresponding ones for the 4_1 knot (equation (7.2)), though considerably more complicated, but are completely different from (7.10), namely

$$\mathcal{H}_{a,b}^{(i)}(x)=\frac{1}{c\sqrt{3\xi_i-2}}\frac{\theta_{1,i}^{c-1}\mathcal{D}_\zeta(\zeta\theta_{1,i})^2}{\mathcal{D}_\zeta\left(\zeta^{-1}\theta_{2,i}^{-1}\right)}\sum_{k,m \bmod c}\zeta^{-(k+1)m}\theta_{1,i}^{-m}\theta_{2,i}^{-k}\frac{(\zeta\theta_{1,i};\zeta)_k^2}{\left(\zeta^{-1}\theta_{2,i}^{-1};\zeta^{-1}\right)_m}, \tag{7.11}$$

where $c = \mathrm{den}(x)$, $\zeta = \mathbf{e}(-x)$ and $\theta_{1,i}^c = -\xi_i^{-3}$ and $\theta_{2,i} = \xi_i^{-2}$ are any choice of c-th roots of $-\xi_i^{-3}$ and ξ_i^{-2} and ξ_1 (resp., ξ_2, ξ_3) the complex root of the equation $\xi^3 - \xi^2 + 1 = 0$ (as in Section 2.1) with negative (resp., positive, zero) imaginary part. Here $\mathcal{D}_\zeta(x)$ is the renormalized version of the cyclic quantum dilogarithm $D_\zeta(x)$ defined for $q = \mathbf{e}(a/c)$ by

$$\mathcal{D}_q(x) = \mathrm{e}^{-2\pi \mathrm{i} s(a,c)/2} D_q(x) = \mathrm{e}^{-2\pi \mathrm{i} s(a,c)/2} \exp\left(\sum_{j=1}^{c-1} \frac{j}{c} \log\left(1 - q^j x\right)\right), \tag{7.12}$$

where $s(a,c)$ is the Dedekind sum (cf. [55, 69]) and where the logarithm is the principal one away from the cut at the negative real axis and is defined on the cut as the average of the principal branches just above and just below. The cyclic quantum dilogarithm appears in the expansion of Faddeev's quantum dilogarithm at roots of unity (see, for example, [29, 60]) and plays a key role in the definition of the near units associated to elements of the Bloch group [10].

It is worth mentioning that the formulas (7.2) and (7.6) for the 4_1 knot can also be written in terms of the modified cyclic quantum dilogarithm $\mathcal{D}_q$, because $\prod_{j=1}^c \left|1 - q^j Z\right|^{2j/c}$ can be rewritten as $\mathcal{D}_q(Z)\mathcal{D}_{q^{-1}}\left(Z^{-1}\right)$. In fact, we expect formulas of this type, involving multiplicative combinations of the $\mathcal{D}_q$'s corresponding to the combinations defining the element of the Bloch group of F corresponding to the knot, to exist for all knots.

7.3 State-sums

In this subsection, we explain where the formulas just given come from. More precisely, we discuss a heuristic method to discover a formula for the first column of the matrix $\mathbf{J}^{(5_2)}$ given a formula for its top entry i.e., for the Kashaev invariant of the knot. This method produces periodic functions similar to the constant term of the formal power series $\Phi_\alpha^{(\sigma)}(h)$ discussed in Section 6. It also generalizes to the further columns, by replacing the Kashaev invariant by the other in its top row (i.e., in the row of the matrix that is expected always to have entries belonging to the rational Habiro ring), thus producing predictions for the entire matrix $\mathbf{J}$. This is useful in particular for the numerical confirmation of the generalized quantum modularity conjecture.

Our starting point is the formula (7.9) for the Kashaev invariant of the 5_2 knot. Let

$$b_{k,\ell}(q) = q^{-(\ell+1)k} \frac{(q;q)_\ell^2}{\left(q^{-1};q^{-1}\right)_k}$$

denote the summand of the Kashaev invariant of 5_2 in equation (7.9). The function $b_{k,\ell}(q)$ (which is proper q-hypergeometric in the sense of [77]) satisfies the linear q-difference equations

$$\frac{b_{k+1,\ell}(q)}{b_{k,\ell}(q)} = q^{-\ell}\left(1 - q^{k+1}\right)^2, \qquad \frac{b_{k,\ell+1}(q)}{b_{k,\ell}(q)} = q^{-(k+1)} \frac{1}{1 - q^{-\ell-1}}, \tag{7.13}$$

whose right-hand sides are in $\mathbb{Q}\left(q, q^k, q^\ell\right)$. It follows that for natural numbers r, s we have

$$\frac{b_{k+r,\ell}(q)}{b_{k,\ell}(q)} = q^{-r\ell}\left(q^{k+1};q\right)_r^2, \qquad \frac{b_{k,\ell+s}(q)}{b_{k,\ell}(q)} = q^{-(k+1)s} \frac{1}{\left(q^{-\ell-1};q^{-1}\right)}$$

and hence

$$\frac{b_{k+r,\ell+s}(q)}{b_{k,\ell}(q)} = q^{-ks-r\ell-(r+1)s} \frac{\left(q^{k+1};q\right)_r^2}{\left(q^{-\ell-1};q^{-1}\right)_s}.$$

Setting $q^k = z_1$, $q^\ell = z_2$, $q = 1$ and equating the ratios of equations (7.13) to 1, we get the gluing equations for (z_1, z_2)

$$z_2^{-1}(1 - z_1)^2 = 1, \qquad z_1^{-1}\left(1 - z_2^{-1}\right)^{-1} = 1. \tag{7.14}$$

Although the summation for the Kashaev invariant when q is a primitive N-th root of unity is a subset of $[0, N-1]^2$ and when $\left(q^k, q^\ell\right)$ is near (z_1, z_2) is outside the summation range, we will pretend that we have performed analytic continuation. Choose $\zeta = \mathbf{e}(a/c)$ where $(a, c) = 1$ and $c > 0$ and $(\theta_1, \theta_2) = \left(z_1^{1/c}, z_2^{1/c}\right)$. In other words, we choose θ_i to be arbitrary c-th roots of z_i $(i = 1, 2)$. Then we can define $a_{r,s}(\theta_1, \theta_2; \zeta)$ by

$$a_{r,s}(\theta_1, \theta_2; \zeta) = \frac{b_{k+r,\ell+s}(q)}{b_{k,\ell}(q)}\Big|_{q^k=\theta_1, q^\ell=\theta_2, q=\zeta} = \zeta^{-(r+1)s}\theta_1^{-s}\theta_2^{-r}\frac{(\zeta\theta_1;\zeta)_r^2}{\left(\zeta^{-1}\theta_2^{-1};\zeta^{-1}\right)_s}.$$

The principle of equipeaked Gaussians in the asymptotics of $J(\gamma X)$ with $\gamma = \left(\begin{smallmatrix} a & b \\ c & s \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ (as used in Section 8.2 for the case of the 4_1 knot) suggests the expression

$$S(\theta_1, \theta_2; \zeta) = \sum_{r,s=0}^{c-1} a_{r,s}(\theta_1, \theta_2). \tag{7.15}$$

The first observation is that the sum in equation (7.15) is c-periodic, i.e., that $r, s \in \mathbb{Z}/c\mathbb{Z}$. This follows from the fact that (z_1, z_2) satisfy the gluing equations (7.14). A second observation, which we will not make use of, if the fact that $b_{k,\ell}(q)$ determines $a_{r,s}(\theta_1, \theta_2)$ according to the above definitions. Conversely, $a_{r,s}(\theta_1, \theta_2; \zeta)$ determines $b_{k,\ell}(q)$ by $a_{r,s}(1, 1; \zeta) = b_{r,s}(\zeta)$. A curious consequence of this is that $S(1, 1; \zeta) = J(\zeta)$ recovers the Kashaev invariant. The gluing equations (7.14) can be solved as follows: $z_1 = -\xi^{-3}$, $z_2 = \xi^{-2}$, where $\xi^3 - \xi^2 + 1 = 0$. The three solutions give rise to the three embeddings of the trace field of 5_2 into the complex numbers. For $\zeta = \mathbf{e}(a/c)$, let $F_c = F(\zeta)$ and $F_{G,c} = F_c(\theta_1, \theta_2)$, giving extensions $F \subset F_c \subset F_{G,c}$, where $F_{G,c}/F_c$ is an abelian Galois (Kummer) extension with group $(\mathbb{Z}/c\mathbb{Z})^2$ and $S(\theta_1, \theta_2) \in F_{G,c}$. To find how $S(\theta_1, \theta_2; \zeta)$ transform under the Galois group, we compute

$$\frac{a_{r,s}(\zeta\theta_1, \theta_2; \zeta)}{a_{r+1,s}(\theta_1, \theta_2; \zeta)} = \theta_2(1 - \zeta\theta_1)^{-2} = a_{1,0}(\theta_1, \theta_2; \zeta)^{-1},$$
$$\frac{a_{r,s}(\theta_1, \zeta\theta_2; \zeta)}{a_{r,s+1}(\theta_1, \theta_2; \zeta)} = \zeta\theta_1\left(1 - \zeta^{-1}\theta_2^{-1}\right) = a_{0,1}(\theta_1, \theta_2; \zeta)^{-1}$$

(where the left-hand side of the above equations is independent of r and s hence it must equal to the right-hand side). Since the sum in equation (7.15) is c-periodic, it follows that

$$S(\zeta\theta_1, \theta_2; \zeta) = S(\theta_1, \theta_2; \zeta)\theta_2(1 - \zeta\theta_1)^{-2} = S(\theta_1, \theta_2; \zeta)a_{1,0}(\theta_1, \theta_2; \zeta)^{-1}, \tag{7.16a}$$
$$S(\theta_1, \zeta\theta_2; \zeta) = S(\theta_1, \theta_2; \zeta)\zeta\theta_1\left(1 - \zeta^{-1}\theta_2^{-1}\right) = S(\theta_1, \theta_2)a_{0,1}(\theta_1, \theta_2; \zeta)^{-1}. \tag{7.16b}$$

To fix the Galois invariance of $S(\theta_1, \theta_2; \zeta)$, we consider the product

$$P(\theta_1, \theta_2; \zeta) = \prod_{r=0}^{c-1}\left(1 - \zeta^{r+1}\theta_1\right)^{2r}\prod_{s=0}^{c-1}\left(1 - \zeta^{-s-1}\theta_2^{-1}\right)^{-s}.$$

We can rewrite the above product using the cyclic quantum dilogarithm function (7.12) as follows

$$P(\theta_1, \theta_2; \zeta) = z_1^{-1}z_2^{-1}\frac{D_\zeta(\theta_1)^2}{D_{\zeta^{-1}}\left(\theta_2^{-1}\right)}. \tag{7.17}$$

From the transformation property for the cyclic quantum dilogarithm

$$\frac{D_\zeta(x)}{D_\zeta\left(\zeta^{-1}x\right)} = \frac{(1-x)^c}{1-x^c}, \qquad D_\zeta(x)D_{\zeta^{-1}}(x) = (1-x^c)^c(1-x)^c$$

and the fact that (z_1, z_2) solve the gluing equations (7.14), we obtain that

$$P(\zeta\theta_1, \theta_2; \zeta) = P(\theta_1, \theta_2; \zeta)\big(\theta_2^{-1}(1-\zeta\theta_1)^2\big)^c = P(\theta_1, \theta_2; \zeta) a_{1,0}(\theta_1, \theta_2; \zeta)^c, \tag{7.18a}$$

$$P(\theta_1, \zeta\theta_2; \zeta) = P(\theta_1, \theta_2; \zeta)\big(\zeta^{-1}\theta_1^{-1}\big(1-\zeta^{-1}\theta_2^{-1}\big)^{-1}\big)^c = P(\theta_1, \theta_2; \zeta) a_{0,1}(\theta_1, \theta_2; \zeta)^c. \tag{7.18b}$$

Equations (7.16) and (7.18) imply that

$$P^{1/c}(\theta_1, \theta_2; \zeta) S(\theta_1, \theta_2; \zeta) \in \varepsilon^{1/c} F_c,$$

where ε is a unit, which in fact coincides with the one constructed in [10].

The expression given in the above equation, after multiplication by a prefactor, coincides with $\mathcal{H}^{(i)}_{0,0}(x)$ of equation (7.11) if we choose θ_1 and θ_2 corresponding to the root ξ_i of $\xi^3 - \xi^2 + 1 = 0$.

In this way, we have succeeded in guessing the entries of the first column of the matrix $\mathbf{J}^{(5_2)}$ of the 5_2 knot starting from the formula (7.9) for its Kashaev invariant. All of this seems to reek a little of "black magic". But the same method applied to the case of the 4_1 knot (whose Kashaev invariant is given in (7.1)) reproduces the formulas given in (7.2). In fact, we believe that this will work for any knot, giving each entry of the first column of the $\mathbf{J}$-matrix as a sum of products of cyclic quantum dilogarithms with summands modelled on the solution of the Neumann–Zagier gluing equations of the knot triangulation in the same way that the expression (7.17) is modelled on the gluing equations (7.14).

7.4 The q-holonomic module of formal power series

We now explain one of the most mysterious aspects of our story, the appearance of two very different realizations of the same q-holonomic system in the contexts of state sums and of perturbative formal power series. In fact, as we will indicate briefly at the end, we believe that there are actually four q-holonomic systems, of totally different origins, given by recursions with the same Laurent polynomials as coefficients.

We begin by recalling the derivation of the perturbative series in h from Nahm sums, as described in Section 6.3. The Nahm sums $F_\Xi(q)$ as defined in (6.6), with H fixed and ν varying over $\mathrm{diag}\big(AB^t\big) + 2\mathbb{Z}^N$, form a module of finite rank over the ring $R = \mathbb{Z}\big[q^{\pm 1}\big]$ of Laurent polynomials in q. For instance, for the original Nahm sums as defined in (6.5), we have the recursions

$$F_{A,b}(q) - F_{A,b+e_k}(q) = q^{\frac{1}{2}e_k^t A e_k + b^t e_k} F_{A,b+Ae_k}(q), \qquad k = 1, \dots, N$$

(here e_k denotes the k-th basis vector of $\mathbb{Z}^N$), as one sees by noting that $(1-q^{n_k})/(q;q)_{n_k}$ vanishes if $n_k = 0$ and equals $1/(q;q)_{n_k-1}$ if $n_k \geq 1$, so that the difference on the left corresponds simply to shifting the multi-index n by e_k. A similar but more complicated calculation (again corresponding to the shift $n \mapsto n + e_k$ in the definition of the sum and using the relationship between Pochhammer symbols with nearby indices) gives a collection of N recursion relations among the various $F_{((AB),\nu)}$ with fixed (AB). This system is always q-holonomic [77], meaning in particular that the solution space is finite-dimensional.

When we discussed the asymptotic behavior of the Nahm sum (6.6) in Section 6.3, we first rewrote the sum as in (6.7) and then replace the sum over $n \in \mathbb{Z}^N$ by an integral over $x \in \mathbb{R}^N$. It is then clear that exactly the same argument (replacing x by $x + e_k$) shows that the formal power series arising from Gaussian integrals near the various stationary points of the integral satisfy the same system of recurrences, and hence also form a q-holonomic module. In favorable cases, including all the ones we have looked at, the rank of this system will be equal to the cardinality of $\mathcal{P}_K$, because the characteristic variety of the system coincides with the variety V_H as defined in Section 6.1.

The surprising discovery is that the abstract q-holonomic module that we obtain this way is the *same* as the one that we found in the first three subsections of this section from the Habiro-like functions, i.e., although the functions of q are completely different and are even defined in different places, the modules in question are spanned by sequences of elements indexed in the same way and satisfying the same recursions over $\mathbb{Q}\left[q^{\pm 1}\right]$, and moreover that the special basis indexed by $\mathcal{P}_K$ is given in both systems by the *same* linear combination of these elements. (Compare the discussion at the end of Section 7.1.) We believe that this coincidence of two q-holonomic structures will hold, not only for the matrix invariants of knot complements, but more generally for corresponding objects associated to any half-symplectic matrix in the sense of Section 6. This will be further studied in [38].

However, a big surprise of the sequel [44] to this paper is that the very same q-holonomic structure actually occurs a third time in terms of the q-series coming from state integrals that are studied there. We believe that this coincidence holds because these three objects are simply different realizations of the same "function-near-$\mathbb{Q}$" belonging to a generalized Habiro ring, with the "nearness" being realized for the Habiro-like functions by approaching a given rational number through nearby rational numbers of slowly growing height, and in the case of the q-series by approaching a rational number from above in the upper half-plane (or equivalently, approaching a root of unity radially in the unit disk). We have checked the agreement of the relations over $\mathbb{Q}\left[q, q^{-1}\right]$ among the columns of the $\mathbf{J}$- and $\mathbf{\Phi}$-matrices for both the 4_1 and 5_2 knots (although we do not describe that verification in this paper because the specific formulas that were used for $\mathbf{J}$ for these two knots in Sections 7.1 and 7.2–7.3 are not the same as the ones coming from ideal triangulations and Neumann–Zagier data and are rather complicated), while the corresponding verification for the q-series for the same two knots is given in [27] and discussed in [44]. We conjecture that these recursive relations coincide with the ones defined in current work of Rinat Kashaev and the first author [31] from the colored Jones polynomials. But we should emphasize that we still have no idea why any of these q-holonomic modules has a canonical basis indexed by $\mathcal{P}$.

8 Proof of the modularity conjecture for the 4_1 knot

In this section, we give our proof of the quantum modularity conjecture for the figure 8 knot, announced several years ago. Another proof was given by [6], as well as proofs of the quantum modularity conjecture for a few other knots, but we give our proof for completeness and because the point of view here is somewhat different from the one there.

We denote by $J(x)$ the J-function for the 4_1 knot, as given explicitly by equation (2.6) with $q = \mathbf{e}(x)$. We have to show that

$$J\left(\frac{aX+b}{cX+d}\right) \sim (cX+d)^{3/2} J(X) \widehat{\Phi}_{a/c}\left(\frac{2\pi \mathrm{i}}{c(cX+d)}\right) \tag{8.1}$$

to all orders in $1/X$ as X tends to infinity with bounded denominator with $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ fixed and $c > 0$, where $\widehat{\Phi}_{a/c}(h) = \mathrm{e}^{V/c^2 h} \Phi_{a/c}(h)$ is the completed version of a formal power series $\Phi_{a/c}(h)$ with algebraic coefficients and where

$$V = \mathrm{Vol}\left(S^3 \smallsetminus 4_1\right) = 2.02988\ldots$$

is the volume of the complement of this knot. We give the proofs separately for the special case $\gamma = S$, $\alpha := a/c = 0$ and for the general case, since all the main ideas are already visible for the former and the details are much simpler.

8.1 The case of $\alpha = 0$

We begin with the case $\alpha = 0$, which makes it clear where the factor $J(X)$ in equation (8.1) comes from. We use the notation

$$P_n(x) = |(q;q)_n|^2,$$

for $q = \mathbf{e}(x)$ with x rational, so that $P_n(x)$ is the n-th summand in the definition of $J(x)$, and denote by c_r $(r \geq 0)$ the numbers defined by the Taylor expansion

$$\cot\left(\frac{\pi}{6} - \frac{x}{2}\right) = \sum_{r=0}^{\infty} c_r \frac{x^r}{r!},$$

the first values being given by the table

r	0	1	2	3	4	5	6	7
c_r	$\sqrt{3}$	2	$2\sqrt{3}$	10	$22\sqrt{3}$	182	$602\sqrt{3}$	6970

Note that these numbers can also be written $c_r = 2\,\mathrm{Im}(\mathrm{i}^{-r}\mathrm{Li}_{-r}(\mathbf{e}(1/6)))$ and hence have a natural extrapolation backwards by $c_{-1} = 0$, $c_{-2} = -V$. We want to study $J(-1/X) = J(1/X)$ as X tends to infinity in the fixed residue class $\beta \pmod 1$, with β rational. The summands in (1.2) are all positive (that is what makes this case much easier to treat than the general one), and it is easy to find their local peaks, which occur near $n = \big(m + \frac{5}{6}\big)X$ for $0 \leq m < \mathrm{den}(\beta)$, where $\mathrm{den}(b)$ is the denominator of b. (Notice that the terms for larger values of m are 0 anyway, since $P_n(x)$ vanishes for $n \geq \mathrm{den}(x)$.) The following proposition, which is valid at a fixed peak even for X real, gives the asymptotic value of the summand $P_n(x)$ for n in each of these peaks. As usual, $B_r(x)$ denotes the r-th Bernoulli polynomial.

Proposition 8.1. *Fix an integer $m \geq 0$, and set $M = m + \frac{5}{6}$. Then for X tending to infinity and n an integer of the form $MX - \nu$ with $|\nu| \ll X$ we have the asymptotic expansion*

$$\log P_n\left(\frac{1}{X}\right) \;\sim\; \frac{V}{2\pi}X + \log X + \log P_m(X) + \sum_{k=1}^{\infty} c_{k-1}\frac{B_{k+1}(\nu)}{(k+1)!}\left(-\frac{2\pi}{X}\right)^k. \tag{8.2}$$

Note: By the above remark we can omit the first term and sum over $k \geq -1$ instead.

Proof. We first note that for $q = \mathbf{e}(1/X)$ we have

$$\log\left(\frac{P_n(1/X)}{P_{n-1}(1/X)}\right) = \log\big|1-q^n\big|^2 = \log\left|1 - \mathbf{e}\left(\frac{1}{6} + \frac{\nu}{X}\right)\right|^2 = -\sum_{k=1}^{\infty}\frac{c_{k-1}}{k!}\left(\frac{2\pi\nu}{X}\right)^k$$

(here we have used that $\frac{\mathrm{d}}{\mathrm{d}x}\log|1-\mathbf{e}(x)|^2 = 2\pi\cot(\pi x)$), in agreement with equation (8.2) to all orders in $1/X$ since $B_{r+1}(\nu+1) - B_{r+1}(\nu) = (r+1)\nu^r$. This proves (8.2) up to a power series independent of n (but depending a priori on α and m). The full assertion uses the shifted Euler–Maclaurin summation formula; we omit the details. ■

Note that equation (8.2) does not make sense if X is rational and $m \geq \mathrm{den}(X)$, since then $P_n(1/X)$ and $P_m(X)$ vanish, but we will use it only in the exponentiated form

$$P_n\left(\frac{1}{X}\right) \;\sim\; P_m(X)X\mathrm{e}^{V/h}\exp\left(\sum_{r\geq 1}(-1)^r c_r \frac{B_{r+1}(\nu)}{(r+1)!}h^r\right), \qquad h = \frac{2\pi}{X}, \tag{8.3}$$

which holds also in this case. The key point here is that the only dependence on m of the expression on the right-hand side is the factor $P_m(X)$, which equals $P_m(\beta)$ if X goes to infinity in the fixed class $\mathbb{Z}+\beta$ modulo 1. Moreover, since $B_2(\nu)=\nu^2+O(\nu)$ and $B_{r+1}(\nu)=O(\nu^{r+1})$ for $r>2$, the exponential factor in equation (8.3) has the form $\mathrm{e}^{-\sqrt{3}\nu^2h/2}\phi(\nu\sqrt{h},\sqrt{h})$, where

$$\phi(\varepsilon\sqrt{h},\sqrt{h}) = \exp\left(\frac{\sqrt{3}}{2}h\left(\varepsilon-\frac{1}{6}\right)+\sum_{r\ge 2}(-1)^r c_r\frac{B_{r+1}(\varepsilon)}{(r+1)!}h^r\right).$$

The contribution to $J(1/X)$ from the m-th peak is equal to $P_m(\beta)X\mathrm{e}^{V/h}$ times the sum of this exponential factor over all ν with $|\nu|\ll X$ in a fixed residue class $\nu_0 \pmod 1$, where $\nu_0\equiv -MX \pmod 1$. But by the Poisson summation formula and the fact that the Fourier transform of a Gaussian decays exponentially, we have that

$$\sum_{\substack{\nu\equiv\nu_0 (\mathrm{mod}\ 1)\\ |\nu|\ll X}} \mathrm{e}^{-\sqrt{3}\nu^2h/2}\phi(\nu\sqrt{h},\sqrt{h}) = \int_{-\infty}^{\infty}\mathrm{e}^{-\sqrt{3}\nu^2h/2}\phi(\nu\sqrt{h},\sqrt{h})\mathrm{d}\nu$$

$$= \sqrt{X}I_{\sqrt{3}}\big[\phi(t,\sqrt{h})\big], \tag{8.4}$$

to all orders in h, where I_λ for $\lambda>0$ denotes the linear map from $\mathbb{C}[[t]]$ to $\mathbb{C}$ defined by

$$I_\lambda\big[\phi(t)\big] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\lambda t^2/2}\phi(t)\mathrm{d}t, \qquad I_\lambda\big[t^n\big] = \begin{cases}(n-1)!!\lambda^{-(n+1)/2} & \text{if } n \text{ is even,}\\ 0 & \text{if } n \text{ is odd.}\end{cases}$$

Here $(n-1)!!$ as usual denotes the "double factorial" $(n-1)\times(n-3)\times\cdots\times 3\times 1$. A discussion of the estimates that prove equation (8.4) is given in [75, 83] and [42]. It follows that the contribution to $J(1/X)$ from the m-th peak is equal to $P_m(\beta)X^{3/2}\mathrm{e}^{V/h}\Phi_0(h)$ to all orders, where $\Phi_0(h)$ is the power series given by equation (8.5). Hence, $J(1/X)$ itself equals $J(\beta)X^{3/2}\mathrm{e}^{V/h}\Phi_0(h)$ to all orders, as claimed. Note that $\Phi_0(h)$ equals $3^{-1/4}$ times a power series in h with coefficients in $\mathbb{Q}(\sqrt{3})$ and leading coefficient 1, since $I_\lambda[\phi(t)]$ has coefficients in $\lambda^{-1/2}\mathbb{Q}(\lambda)$ for any power series $\phi(t)$ with rational coefficients. (That it is a power series in h rather than merely $\sqrt{h}$ follows from the fact that I_λ annihilates odd functions.)

This concludes the proof of equation (8.1) when $\alpha=0$, with $\Phi_0(h)$ given by

$$\Phi_0(h) = I_{\sqrt{3}}\big(\phi(t\sqrt{h},\sqrt{h})\big). \tag{8.5}$$

8.2 The general case

We now apply the same analysis to the expansion of $J(x)$ around an arbitrary rational number α. The second part of the argument, replacing sums by integrals and computing them by using the functional I_λ, is unchanged, but the analogue of Proposition 8.1 is now slightly more complicated, since the asymptotic formula for $P_n(x)$ near the m-th peak depends on both m and the residue class of n modulo the denominator of α. We use the notations given above, i.e., $x=\frac{aX+b}{cX+d}$ with X tending to infinity in the class $\beta \pmod 1$ and $M=m+\frac{5}{6}$ with $0\le m<\mathrm{den}(\beta)$, but now we also fix a residue class $r \pmod c$ and consider n satisfying

$$n\equiv r+md \pmod c, \qquad n=\frac{M}{\hbar}+c\nu$$

with $|\nu|\ll X$. (Notice that ν has the opposite sign to the one used for $c=1$.)

Proposition 8.2. *For fixed $m<\mathrm{den}(\beta)$ and $r\in\mathbb{Z}/c\mathbb{Z}$ and for n and X tending to infinity as in (5), we have the asymptotic formula*

$$\log P_n(x) \sim \log\left(\frac{1}{\hbar}\right)+\log P_m(\beta)+\sum_{k=-1}^{\infty}C_k^{(r)}(\nu)h^k, \tag{8.6}$$

valid to all orders in h, where $C_k^{(r)}$ *is the polynomial of degree* $k-1$ *defined by*

$$C_k^{(r)}(\nu) = -2\,\mathrm{Re}\left[\frac{\mathrm{i}^{-k}}{(k+1)!}\sum_{j=1}^{c}\mathrm{Li}_{1-k}\big(\zeta_\alpha^{r+j}Z\big)B_{k+1}\left(\nu+\frac{j}{c}\right)\right], \qquad r\in\mathbb{Z}/c\mathbb{Z}, \quad k\geq -1 \quad (8.7)$$

with $\zeta_\alpha = \mathbf{e}(\alpha)$ *and* $Z = \mathbf{e}(-5/(6c))$.

The proof of this proposition, which we omit, is similar to that of Proposition 8.1, the main point again being that the difference of the right-hand sides of (8.6) for n and $n-1$ is given by

$$\begin{aligned}\sum_{k=-1}^{\infty}\big[C_k^{(r)}(\nu)-C_k^{(r-1)}(\nu-1)\big]h^k &= -2\,\mathrm{Re}\left[\sum_{k=0}^{\infty}\mathrm{Li}_{1-k}(\zeta_\alpha^r Z)\frac{(-\mathrm{i}h\nu)^k}{k!}\right]\\ &= \log\big|1-\zeta_\alpha^r Z\mathrm{e}^{-\mathrm{i}\nu h}\big|^2 = \log|1-q^n|^2\\ &= \log P_n(x) - \log P_{n-1}(x),\end{aligned}$$

because

$$q^n = \mathbf{e}\left(n\left(\alpha-\frac{1}{c(cX+d)}\right)\right) = \mathbf{e}\left((r+md)\alpha - \frac{m+5/6}{c} - \nu\hbar\right) = \zeta_\alpha^r Z\mathrm{e}^{-\mathrm{i}\nu h}$$

and

$$\begin{aligned}C_k^{(r)}(\nu)-C_k^{(r-1)}(\nu-1) &= -2\,\mathrm{Re}\left[\frac{\mathrm{i}^{-k}}{(k+1)!}\left(\sum_{j=1}^{c}-\sum_{j=0}^{c-1}\right)\mathrm{Li}_{1-k}\big(\zeta_\alpha^{r+j}Z\big)B_{k+1}\left(\nu+\frac{j}{c}\right)\right]\\ &= -2\,\mathrm{Re}\left[\mathrm{i}^{-k}\mathrm{Li}_{1-k}(\zeta_\alpha^r Z)\frac{B_{k+1}(\nu+1)-B_{k+1}(\nu)}{(k+1)!}\right]\\ &= -2\,\mathrm{Re}\left[\mathrm{Li}_{1-k}(\zeta_\alpha^r Z)\frac{(\nu/\mathrm{i})^k}{k!}\right].\end{aligned}$$

Note that in the above calculation we used that $C_{-1}^{(r)}(\nu)$ is independent of both ν and r. In fact, it is given by

$$C_{-1}^{(r)}(\nu) = 2\,\mathrm{Im}\left[\sum_{j=1}^{c}\mathrm{Li}_2\big(\zeta_\alpha^{r+j}Z\big)\right] = \frac{2}{c}\,\mathrm{Im}[\mathrm{Li}_2(Z^c)] = \frac{V}{c},$$

where we have used the well-known "distribution" property of the dilogarithm. The corresponding distribution property of the 1-logarithm $\mathrm{Li}_1(z) = -\log(1-z)$ shows that $C_0^{(r)}(\nu)$ is also independent of ν and is given by

$$C_0^{(r)}(\nu) = \sum_{j=1}^{c}\left(\nu+\frac{j}{c}-\frac{1}{2}\right)\log\big|1-\zeta_\alpha^{r+j}Z\big|^2 = \log E_r(\alpha),$$

where $E_r(\alpha)$ is the real algebraic number defined by

$$E_r(\alpha) = \prod_{j=1}^{c}\big|1-\zeta_\alpha^{r+j}Z\big|^{2j/c}, \qquad r\in\mathbb{Z}/c\mathbb{Z}. \quad (8.8)$$

Hence, the exponentiated version of (6) can be written

$$P_n(x) \;\sim\; P_m(\beta)\frac{\mathrm{e}^{V/c\hbar}}{\hbar}E_r(\alpha)\exp\left(\sum_{k=1}^{\infty}C_k^{(r)}(\nu)h^k\right),$$

where again the exponential factor at the end has the form $\mathrm{e}^{-c\sqrt{3}\nu^2 h/2}\phi_{c,r}\big(\nu\sqrt{h},\sqrt{h}\big)$ with

$$\phi_{c,r}\big(\varepsilon\sqrt{h},\sqrt{h}\big) = \exp\bigg(\widetilde{C}_1^{(r)}(\nu) + \sum_{k\geq 2} C_k^{(r)}(\nu)h^k\bigg),$$

where $\widetilde{C}_1^{(r)}(\nu)$ is given by the same formula as the right-hand side of (8.7) (with $k=1$) and with $B_2(x) = x^2 - x + 1/6$ replaced by $B_2(x) - x^2$. Note that $\phi_{c,r}(t,\varepsilon)$ is a power series in ε with coefficients in $\mathbb{Q}(\zeta_\alpha, Z)[t]$ and leading coefficient 1. The same reasoning as for $c=1$ now shows that the sum of the values of $P_n(x)$ for n running over the m-th peak and in the residue class $r + md \pmod{c}$ is equal to $\hbar^{-3/2}\mathrm{e}^{V/c\hbar}P_m(\beta)\Phi_\alpha^{(r)}(h)$ for some power series $\Phi_a^{(r)}(h)$ with leading coefficient $E_r(\alpha)$, and summing this over all r gives equation (1.5) for the 4_1 knot with

$$\Phi_\alpha^{(4_1,\sigma_1)}(h) = 3^{-1/4}c^{-1/2}\sum_{r \pmod{c}} \Phi_\alpha^{(r)}(h).$$

Note that the formal Gaussian integration formula for the power series $\Phi_\alpha(h)$ requires to expand the integrand up to order $O\big(h^{3k+1}\big)$ in order to obtain the coefficient of h^k in the series $\Phi_\alpha(h)$.

It remains to look at the units $E_r(\alpha)$. Write F for $\mathbb{Q}(\mathbf{e}(1/6))$, the trace field of the figure 8 knot, and F_c for its cyclotomic extension $F(\zeta_\alpha) = \mathbb{Q}(Z)$. We claim that both $E_r(\alpha)/E_0(\alpha)$ and $\prod_{r \pmod{c}} E_r(\alpha)$ belong to F_c. The second claim follows from the first, since it is clear that $E_0(\alpha)^c$ belongs to F_c, and the first claim follows from the calculation

$$\frac{E_r(\alpha)}{E_{r-1}(\alpha)} = \frac{\prod_{j=0}^{c}\big|1-\zeta_\alpha^{r+j}Z\big|^{2j/c}}{\prod_{j=0}^{c-1}\big|1-\zeta_\alpha^{r+j}Z\big|^{2(j+1)/c}} = \frac{\big|1-\zeta_\alpha^{r}Z\big|^{2}}{\prod_{n \pmod{c}}\big|1-\zeta_\alpha^{n}Z\big|^{2/c}} = \big|1-\zeta_\alpha^{r}Z\big|^{2},$$

from which we get by induction the formula

$$E_r(\alpha) = E_0(\alpha)\big|(\zeta_\alpha Z, \zeta_\alpha)_r\big|^2$$

for all r. In particular, we can write our asymptotic formula to leading order as

$$J\left(\frac{aX+b}{cX+d}\right)\Big/ J(X) \;\sim\; \frac{cE_0(\alpha)S(\alpha)}{3^{1/4}}X^{3/2}\exp\left(\frac{V}{2\pi}\left(X+\frac{d}{c}\right)\right)$$

as $X\to\infty$ with bounded denominator, where

$$S(\alpha) = \sum_{r \pmod{c}}\frac{E_r(\alpha)}{E_0(\alpha)} = \sum_{n=0}^{c-1}\big|(\zeta_\alpha Z,\zeta_\alpha)_n\big|^2 \in F_c. \tag{8.9}$$

It is the factor $S(\alpha)$ which for $c=5$ contains the funny prime π_{29} occurring in [84, p. 14], while $E_0(a)$ is the unit analyzed in [10]. Note that the special properties (9.3) of this unit are clear from the definition (8.8) since if c is prime to 6 and we denote by σ_k the Galois automorphism of F_c over F sending a primitive c-th root of unity to its k-th power, then it is easy to see from (8) that $\sigma_k(E_r(a/c)^c) = E_r(ka/c)^c$ and that the quotient $E_r(ka/c)^k/E_r(a/c)$ belongs to F_c.

For other knots K, there is a similar story, but we can no longer rigorously prove anything, since the terms in the sum defining $J^K(x)$ are no longer positive (or even real) and there is cancellation. However, this sum still has the form of an N-dimensional sum of products or quotients of Pochhammer symbols, where N is the dimension of some terminating q-hypergeometric series (related to the number of simplices in a triangulation of $S^3\smallsetminus K$), and we can formally look at the parts of this sum where the summands are locally constant ("stationary phase"), even if those "parts" now lie outside of the original domain of summation. This leads to a conjectural, but completely explicit, formula of the same general form as (2), and in particular to an asymptotic formula like (9), but with $E_r(\alpha) = E_r^K(\alpha)$ now depending on an element r of $(\mathbb{Z}/c\mathbb{Z})^N$ rather than just $\mathbb{Z}/c\mathbb{Z}$ and with the sum in (10) replaced by one over $(\mathbb{Z}/c\mathbb{Z})^N$. For the 5_2 knot and its sister, the $(-2,3,7)$ pretzel knot, we worked out the corresponding expressions and for small values of c obtained both the units and the pre-factors $S^K(\alpha) \in F_c$ (F = trace field of K) that we had previously found numerically. These are, however, much more complicated than in the 4_1 case; for instance, the factor $\mathfrak{p}_{29} = 2 - \varepsilon_1^{(a)} + \varepsilon_2^{(a)} + 2\varepsilon_3^{(a)}$, a prime of norm 29 that occurred for the 4_1 knot and $c = 5$ (see equation (1.7) and the discussion in the next section) is replaced for the 5_2 knot by $\mathfrak{p}_7^2\mathfrak{p}_{43}$ if $c = 3$ and by $\mathfrak{p}_{9491}\mathfrak{p}_{1227271}$ if $c = 5$, where each $\mathfrak{p}_p$ denotes a prime of norm p in $\mathbb{Q}(\xi)$.

9 Arithmetic aspects

In this section, we discuss the arithmetic properties of the power series $\Phi_\alpha^{(K,\sigma)}(h)$, in particular the identification of the number fields in which their coefficients lie and the integrality properties of these coefficients.

9.1 Algebraic number theory aspects

A detailed study of the power series $\Phi_\alpha^{(K,\sigma)}(h)$ (or more generally $\Phi_\alpha^{(K,\sigma,\sigma')}(h)$) reveals several interesting algebraic number theoretical aspects, especially concerning the field of definition, transformation under the action of the Galois group, and above all the appearance of non-trivial algebraic units.

We begin by looking in more detail at the series $\Phi_{a/5}^{(4_1)}$ because this example is quite illuminating. The first few terms of the series were given in [84, p. 670], as

$$\Phi_{a/5}^{(4_1,\sigma_1)}(h) = \sqrt[4]{3}\,\sqrt[10]{E^{(a)}}\Big(\big(2 - E_1^{(a)} + E_2^{(a)} + 2E_3^{(a)}\big) + \frac{2678 - 943E_1^{(a)} + 1831E_2^{(a)} + 2990E_3^{(a)}}{2^3 3^2 5^2\sqrt{-3}}h + \cdots\Big), \tag{9.1}$$

where $E_k^{(a)} = 2\cos\big(\frac{2\pi(6a-5)k}{15}\big)$ and $E^{(a)} = E_2^{(a)}/\big(E_1^{(a)}\big)^3 E_3^{(a)}$, except that the formula was given there in terms of $\log\Phi$, which introduced spurious denominators in all terms of the expansion. Actually, this is one of the first insights from the numerical calculations: earlier papers had always worked with the logarithm, which is what one sees if one does a Feynman diagram expansion and looks at the contribution of connected graphs only, but (as in many other combinatorial problems) one gets much simpler numbers by looking at the exponentiated sum, corresponding to summing over all graphs rather than just the connected ones. In the case at hand, this meant that the coefficients of h and h^2 in [84] contained mysterious powers of the prime

$$\pi_{29}^{(a)} = 2 - E_1^{(a)} + E_2^{(a)} + 2E_3^{(a)}$$

of $\mathbb{Q}\big(\cos\big(\frac{2\pi}{15}\big)\big)$, which simply disappear as soon as one goes from the logarithm of the series to the series itself. But the few terms of $\Phi_{a/5}^{(4_1,\sigma_1)}(h)$ given in (9.1) also suffice to illustrate several other key arithmetic points:

(a) The most striking feature of (9.1) is the appearance of the 10th root of the algebraic unit $E^{(a)}$ as a prefactor. From this and the corresponding numbers found for other values of α and for other knots we were led to conjecture the existence of algebraic units in cyclotomic extensions of any number field determined by elements of the Bloch group of this field, a prediction that was then confirmed in the joint paper [10] with Frank Calegari.

(b) The case of the 4_1 knot has the somewhat misleading special property that the Kashaev invariant (1.2) is always positive, so that we seem to be seeing elements in the real part $\mathbb{Q}\big(\cos\big(\frac{2\pi}{15}\big)\big)$ of the cyclotomic extension $F(\zeta_5)=\mathbb{Q}(\zeta_{15})$ of the trace field $F=F_{4_1}=\mathbb{Q}\big(\sqrt{-3}\big)$ rather than in this cyclotomic extension itself. In particular, as was not observed in [84], the unit $E^{(a)}$ is, up to sign, the *square* of an element of $F(\zeta_5)$, so that its 10 root is, up to a root of unity, in fact a fifth root of a unit in this larger field. Specifically, we have

$$\sqrt{-E^{(a)}}=\big(\zeta_{15}^{(a)}-\big(\zeta_{15}^{(a)}\big)^{-1}\big)E_2^{(a)}/E_1^{(a)},$$

permitting us to rewrite (9.1) in the form (1.7) given in Section 1. This, too, turned out to be true for the general case studied in [10], where one associates to a number field F and an element of its Bloch group the c-th root of a unit (or at least S-unit for a finite set of primes S independent of c) in $F(\zeta)$ for every primitive c-th root of unity ζ, and not a $(2c)$-th root. This unit, for $F=F_\sigma$ and $\zeta=\mathbf{e}(\alpha)$, is expected to appear as a prefactor in $\Phi_\alpha^{(K,\sigma)}(h)$ for every K, σ, and α.

(c) Apart from the unit prefactor $\sqrt{-E^{(a)}}$ (which equals $\varepsilon_{a/5}$ in the notation of (1.7)), there is a further prefactor $3^{1/4}$ that coincides with the torsion $\delta(4_1)^{-1/2}=\big(\sqrt{-3}\big)^{-1/2}$ up to a root of unity and an element of F_{4_1}.

(d) After we remove these factors, the remaining power series has coefficients in the cyclotomic extension $F(\zeta_5)$ of the trace field.

(e) The denominators of this remaining power series, when we calculate it to many more terms using the methods described in Section 10.1, have powers of 3 (the ramified prime already occurring in (c)) and D_n, where

$$D_n=2^{3n+v_2(n!)}\prod_{\substack{p\ \text{prime}\\ p>2}}p^{\sum_{i\ge0}[n/p^i(p-2)]}. \tag{9.2}$$

(Note that the exponent $v_p(D_n)$ of $p>2$ in D_n can be written as $r+v_p(r!)=v_p((pr)!)$ where $r=[n/(p-2)]$.) We will return to this point in the next subsection.

(f) The unit $\varepsilon_{a/5}$ occurring in (a) and (b), the term under the square-root sign in (c), and the coefficients of the "remaining power series" as defined in (d) are not only in $F(\zeta_5)$, but transform under the Galois group $\{\zeta\mapsto\zeta^r\}_{5\nmid r}$ of $F(\zeta_5)/F$ in the "obvious" way, i.e., each of these numbers is a polynomial in $\zeta=\mathrm{e}^{2\pi\mathrm{i}a/5}$ whose coefficients lie in F_{4_1} and are independent of a.

(g) Finally, the unit $\varepsilon_{a/5}$ of $F(\zeta_5)$, considered in the quotient $F(\zeta_5)^\times/F(\zeta_5)^{\times5}$, transforms under the action of the Galois group $\mathrm{Gal}(F(\zeta_5)/F)=(\mathbb{Z}/5\mathbb{Z})^\times$ in two different ways:

$$\underline{\sigma}_r(\varepsilon_{a/5})=\varepsilon_{ar/5}=(\varepsilon_{a/5})^{1/r},\qquad r\in(\mathbb{Z}/5\mathbb{Z})^\times, \tag{9.3}$$

where $\underline{\sigma}_r$ is the Galois automorphism defined by $\underline{\sigma}_r(\zeta_5)=\zeta_5^r$.

We conjecture that these properties (a)–(g) hold for all hyperbolic knots K, all representations σ in $\mathcal{P}_K$ and all roots of unity ζ_α, with F replaced by the trace field F_σ and 5 by the denominator of α, as well as a few other small modifications (in particular, that instead of a unit one may get an S-unit for small finite set S of primes, essentially the ones occurring in the shape parameters of a triangulation of $S^3\smallsetminus K$, which was empty for 4_1). In other words, the power series $\Phi_\alpha^{(K,\sigma)}(h)$ can be written in the form

$$\Phi_\alpha^{(K,\sigma)}(h)=\mu_{\sigma,\alpha}\cdot(\varepsilon_{\sigma,\alpha})^{1/c}\cdot\delta_\sigma^{-1/2}\sum_{n=0}^{\infty}\widetilde{A}_{\alpha,n}^{(K,\sigma)}h^n,\qquad \widetilde{A}_{\alpha,n}^{(K,\sigma)}\in F_\sigma(\zeta_\alpha) \tag{9.4}$$

(so that $\widetilde{A}^{(K,\sigma)}_{\alpha,n}$ is the product of an algebraic number independent of n and the coefficient denoted $A^{(K,\sigma)}_{\alpha}(n)$ in Section 3.4), where F_σ is defined as in Section 2, $\mu_{\sigma,\alpha}$ is an $(8c)$-th root of unity, and $\varepsilon_{\sigma,\alpha} \in F_\sigma(\zeta_\alpha)^\times$ is a near-unit, canonically defined only up to c-th powers, that transforms up to c-th powers as in (9.3) (with 5 replaced everywhere by c) and that conjecturally depends only on the element of the Bloch group $B(F_\sigma)$ determined by σ and in fact coincides with the near-unit that was constructed in [10], and with the same denominator bound D_n as in (9.2), independent of K, σ and α.

9.2 Denominators and integrality properties

The universal denominator statement given in formula (9.2) above was found empirically on the basis of the extensive numerical data for the 4_1, 5_2 and the $(-2,3,7)$ pretzel knots presented in the appendix to this paper. In this section, we prove it for the denominators of the power series defined in terms of Gaussian-type integrals in [15]. This proof only applies to $\sigma \in \mathcal{P}_K^{\mathrm{red}}$, since there is no such integral representation for the trivial representation, but the corresponding denominator statement is true here also and can in fact be strengthened because the power series in that case come from the Habiro ring, as explained at the end of this section.

Theorem 9.1. *For each knot K, representation $\sigma \in \mathcal{P}_K$, and number $\alpha \in \mathbb{Q}/\mathbb{Z}$, we have*

$$D_n \widetilde{A}^{(K,\sigma)}_{\alpha,n} \in \mathcal{O}_S\big[\zeta_\alpha, c^{-1}\big],$$

where $\Phi^{(K,\sigma)}_{\alpha}(h)$ is as in [15], D_n is as in (9.2), $\widetilde{A}^{(K,\sigma)}_{\alpha,n}$ is as in (9.4), c is the denominator of α, $\mathcal{O}$ is the ring of integers of F_σ and S is a finite set of primes of F_σ that depends on K but not on n or on α.

The first few values of D_n are given by

$$1,\ 24,\ 1152,\ 414720,\ 39813120,\ 6688604160,\ 4815794995200,\ 115579079884800,$$
$$22191183337881600,\ 263631258054033408000,\ 88580102706155225088000,$$
$$27636992044320430227456000,\ 39797268543821419527536640000,\ \ldots.$$

Campbell Wheeler pointed out to us that the above sequence appears (with no proof) to equal to the sequence `A144618` of the online-encyclopedia of integer sequences [70], the latter related to Stirling's formula with half-shift $D_n = \mathrm{den}(a_n)$, where

$$z! \sim \sqrt{2\pi}(z+1/2)^{z+1/2}\mathrm{e}^{-z-1/2}\sum_{n=0}^{\infty}\frac{a_n}{(z+1/2)^n}, \qquad z \to \infty.$$

The numbers D_n grow rapidly, for example

$$D_{50} = 2^{197}3^{72}5^{19}7^{11}11^{5}13^{4}17^{3}19^{2}23^{2}29^{1}31^{1}37^{1}41^{1}43^{1}47^{1}$$

or $D_{50}/24^{50}50! = 5^7 7^3 11^1 13^1 17^1$. $\big($In general, $D_n/24^n n!$ is an integer whose nth root tends to $\prod_{p\geq 5} p^{2/(p-1)(p-2)} = 1.8592481285\ldots.\big)$ We give the first 50 values of D_n by tabulating the ratio $\delta_n = D_n/3D_{n-1}$ (after removing the power of 2, and omitting the values equal to 1):

n	δ_n	n	δ_n	n	δ_n	n	δ_n	n	δ_n	n	δ_n
3	$3{\cdot}5$	11	13	20	7	27	$3^3{\cdot}5{\cdot}11{\cdot}29$	35	$7^2{\cdot}37$	42	$3{\cdot}5{\cdot}23$
5	7	12	$3{\cdot}5$	21	$3{\cdot}5{\cdot}23$	29	31	36	$3^2{\cdot}5{\cdot}11$	44	13
6	$3{\cdot}5$	15	$3{\cdot}5^2{\cdot}7{\cdot}17$	22	13	30	$3{\cdot}5^2{\cdot}7{\cdot}17$	39	$3{\cdot}5{\cdot}41$	45	$3^2{\cdot}5^2{\cdot}7{\cdot}11{\cdot}17{\cdot}47$
9	$3^2{\cdot}5{\cdot}11$	17	19	24	$3{\cdot}5$	33	$3{\cdot}5{\cdot}13$	40	7	48	$3{\cdot}5$
10	7	18	$3^2{\cdot}5{\cdot}11$	25	7	34	19	41	43	50	7

We also remark that $D_{n_1}D_{n_2}|D_{n_1+n_2}$ for all $n_1, n_2 \geq 0$ and hence that the subgroup

$$R_D[[h]] = \left\{ \sum_{n=0}^{\infty} \frac{a_n}{D_n} h^n \;\Big|\; a_n \in R \text{ for all } n \right\}$$

of $K[[h]]$ is a subring for every subring R of a field K of characteristic zero.

Proof. We give the proof only for the case $\alpha = 0$, $c = 1$, using the formulas in [14]. The general case can be proved along the same lines using the more complicated formulas in [15], in which the Bernoulli numbers are replaced by Bernoulli polynomials, but we do not give the details here. We will also ignore the prime 2 in our proof since it behaves somewhat differently and in any case can be added to the finite set of excluded primes S in the statement of the theorem.

The power series $\Phi_0^{(K,\sigma)}(h)$ attached to a triangulation $\mathcal{T}$ of $\mathbb{S}^3 \smallsetminus K$ were defined in [14] as formal Gaussian integrals $\langle f_{\mathcal{T}} \rangle$ of the formal power series

$$f_{\mathcal{T}}(x; z) = \exp\left(\sum_{j=1}^{N} \sum_{\substack{r,k \geq 0 \\ 2r+k-2>0}} \frac{B_r}{r!} \frac{(-x_j)^k}{k!} \mathrm{Li}_{2-r-k}(z_j) h^{r+\frac{k}{2}-1} \right) \tag{9.5}$$

in a multi-variable $x = (x_1, \ldots, x_N)$, where $z_1, \ldots, z_N$ are the shape parameters of $\mathcal{T}$ and where $\langle f \rangle = \langle f(x) \rangle_Q$ is the mean value defined by Gaussian integration with respect to a certain quadratic form Q with coefficients in the field F_σ. This form is essentially the one given by the symmetric matrix $A + B\,\mathbf{diag}(z_j/(1-z_j))$ that occurred in the discussion of (6.7) in Section 6.3, and the function (9.5) is essentially the product of the functions occurring in (6.8), except that the terms with $r + k = m$ fixed were combined there into a single Bernoulli polynomial $B_m(x)$ for $m \geq 3$ or $B_2(x) - x^2$ for $m = 2$, and that the normalizations used in [14] were slightly different from the ones used in Section 6.3.

We now expand the exponential in (9.5) as the product of the exponentials of the monomials in the sum, and recall that $\mathrm{Li}_{2-m}(z) \in \mathbb{Z}[1/(1-z)]$ for every $m \geq 2$, to deduce that $\Phi_\alpha^{(K,\sigma)}(h) = \langle f_{\mathcal{T}} \rangle$ is an R-linear combination of Gaussian averages $\langle T \rangle$ of terms T of the form

$$T = \prod_{j=1}^{N} \prod_{\substack{r,k \geq 0 \\ 2r+k \geq 3}} \frac{1}{\lambda_j(r,k)!} \left(\frac{B_r}{r!} \frac{x_j^k}{k!} h^{r+\frac{k}{2}-1} \right)^{\lambda_j(r,k)} \tag{9.6}$$

with non-negative multiplicities $\lambda_j(r,k)$, where $R = R_{\mathcal{T},\sigma}$ denotes the ring generated over $\mathbb{Z}$ by the numbers $(1-z_j)^{-1}$. We write the monomial T as $c(T) h^n \mathbf{x}^{\mathbf{K}}/\mathbf{K}!$ with

$$n = \sum_{j,r,k} \lambda_j(r,k) \left(r + \frac{k}{2} - 1 \right), \qquad K_j = \sum_{r,k} \lambda_j(r,k) k \tag{9.7}$$

and where for notational convenience have written $\mathbf{x}$ and $\mathbf{K}$ for the N-tuples $(x_1, \ldots, x_N)$ and $(K_1, \ldots, K_N)$ (thus deviating from the convention in the rest of the paper where boldface denotes matrices) and $\mathbf{x}^{\mathbf{K}}/\mathbf{K}!$ for the divided power $\prod x_j^{K_j}/K_j!$. To prove the theorem, we have to bound both the denominators of the numerical coefficient $c(T)$ and the further denominators coming from the Gaussian averaging $\mathbf{x}^{\mathbf{K}}/\mathbf{K}! \mapsto \langle \mathbf{x}^{\mathbf{K}}/\mathbf{K}! \rangle$.

We begin with the latter question. For this, we recall first that the Gaussian average $\langle f \rangle_Q$ is given by $\mathrm{e}^{\Delta_Q}(f)\big|_{x=0}$ for any power series f, where Δ_Q is the Laplacian associated to Q, and hence is equal to $\Delta_Q^{\ell}(f)/\ell!$ if f is a homogeneous polynomial of degree 2ℓ. (For polynomials of odd degree it of course vanishes trivially.) The Laplacian Δ_Q is a quadratic polynomial in the derivatives $\partial_i = \partial/\partial x_i$, and we can enlarge the ring R by adjoining to it the coefficients of this polynomial, so since the image of any divided factorial under any product $\partial_1^{\ell_1}\cdots\partial_N^{\ell_N}$ is an integer, we then certainly have that the Gaussian integral $\langle \mathbf{x}^{\mathbf{K}}/\mathbf{K}!\rangle$ is $1/\ell!$ times an element of R. Unfortunately, it turns out that this estimate is not good enough for our purposes, and we have to work a little harder.

Writing Δ_Q as an R-linear combination of binomials $\partial_i\partial_j$ with $1 \le i \le j \le N$, and applying the multinomial theorem, we see that $\Delta_Q^\ell/\ell!$ is an R-linear combination of terms $\prod_{i\le j}(\partial_i\partial_j)^{m_{ij}}/m_{ij}!$ with $m_{ij}\in\mathbb{Z}_{\ge 0}$. Define an even symmetric $N\times N$ matrix M by setting $M_{ij}=M_{ji}=m_{ij}$ for $i<j$ and $M_{ii}=2m_{ii}$. Then $\prod_{i\le j}(\partial_i\partial_j)^{m_{ij}}=\prod_j \partial_j^{K_j}$ with $\mathbf{K}=M\mathbf{1}$, where $\mathbf{1}$ is the vector consisting of N 1's, and this sends $\mathbf{x}^{\mathbf{K}}/\mathbf{K}!$ to 1 and all other monomials to 0. It follows that a universal denominator of $\langle \mathbf{x}^{\mathbf{K}}/\mathbf{K}!\rangle$ is the number

$$\Delta(\mathbf{K}) := \text{l.c.m.}\biggl\{\prod_{1\le i\le N}(M_{ii}/2)!\prod_{1\le i<j\le N}M_{ij}!\,\biggm|\, M=M^t\in M_{N,N}(\mathbb{Z}_{\ge 0})\text{ even},\ M\mathbf{1}=\mathbf{K}\biggr\}.$$

Notice that this does divide $\ell!$, because ℓ is the sum of the diagonal $M_{ii}/2$ and of the M_{ij} with $i<j$, so that this statement refines the bound given above, but $\Delta(\mathbf{K})$ is in general much smaller than $\ell!$ and this improvement will be needed for the proof. A usually sharper multiplicative upper bound for $\Delta(\mathbf{K})$ is the largest integer S whose square divides the product of the $K_j!$ (the proof of this also uses only the integrality of multinomial coefficients), and then of course $\Delta(\mathbf{K})$ also divides the g.c.d. of $\ell!$ and S, which in general is smaller than either one. (For instance, for $\mathbf{K}=(6,9,9,10)$ we have $\ell! = 355687428096000$, $S=3135283200$, and $(\ell!,S)=S/3$.) Either of these two latter upper bounds would be sufficient for our proof, but in fact there is an easy upper bound that is stronger than either one of them and is extremely sharp (in particular, it is equal to $\Delta(\mathbf{K})$ for all $\mathbf{K}$ with $N\le 4$ and $\max(K_j)\le 30$), namely

$$\Delta^*(\mathbf{K}) := \prod_{p\text{ prime}} p^{\delta_p(\mathbf{K})} \quad\text{with}\quad \delta_p(K_1,\dots,K_N) := \sum_{s\ge 1}\left[\frac{1}{2}\sum_{j=1}^N\left[\frac{K_j}{p^s}\right]\right].$$

To show that $\Delta(\mathbf{K})$ divides $\Delta^*(\mathbf{K})$, we need $\sum_i V_p(M_{ii}/2)+\sum_{i<j}V_p(M_{ij}) \le \delta_p(\mathbf{K})$ for every prime p and every even symmetric matrix M with non-negative entries and row sums $\mathbf{K}$, where $V_p(m) := v_p(m!)$ for any $m\ge 0$ denotes the largest power of p dividing $m!$. In view of the standard formula $V_p(m)=\sum_{s\ge1}[m/p^s]$, it suffices for this to show that

$$\sum_i [M_{ii}/2q] + \sum_{i<j}[M_{ij}/q] \le \frac{1}{2}\sum_j [K_j/q]$$

for every prime power q, and this follows immediately from the obvious facts $[x/2q]=[[x/2]/q]$ and $[x]+[y]+\cdots \le [x+y+\cdots]$ valid for arbitrary real numbers $x,y,\dots$.

Now going back to our main problem, we now see that it suffices to show that the product $\Delta^*(\mathbf{K})c(T)$ has denominator dividing D_n for all terms T as above, with $\mathbf{K}=(K_1,\dots,K_N)$ and n defined by (9.7). We will prove this one prime at a time (ignoring the prime 2), which is natural in view of the fact that the upper bound $\Delta^*(\mathbf{K})$ is defined by its prime power decomposition. To do this, we will split both the term T and the corresponding numerical coefficient $c(T)$, and also each of the N factors T_j and $c(T_j)$ of which they are comprised, as the product of four factors in a way depending on the prime p being studied, labelled "s" (terms with $r=0$ and k smaller than p), "b" (terms with $r=0$ and k bigger than or equal to p), "1" (terms with $r=1$), and "≥ 2" (terms with $r\ge 2$), with a similar splitting of the individual weights K_j into the sum of four pieces $K_j^{(s)}=\sum_{3\le k<p}\lambda_j(0,k)k$, $K_j^{(b)}=\sum_{k\ge p}\lambda_j(0,k)k$, $K_j^{(1)}=\sum_{k\ge1}\lambda_j(1,k)k$, and $K_j^{(\ge2)}=\sum_{r\ge2,k\ge0}\lambda_j(r,k)k$. We also decompose the number n (the

exponent of h in T) in (9.7) as $\ell + n' - t$ with

$$\ell := \frac{1}{2}\sum_{j=1}^{N} K_j, \qquad n' := \sum_{\substack{1\le j\le N\\ r\ge 2, k\ge 0}} \lambda_j(r,k)(r-1), \qquad t := \sum_{\substack{1\le j\le N\\ k\ge 3}} \lambda_j(0,k)$$

and also split t as $t^{(s)} + t^{(b)}$ according as $3 \le k \le p-1$ or $k \ge p$ in the summation, with each of $t^{(s)}$ and $t^{(b)}$ splitting into the sum over $1 \le j \le N$ of pieces $t_j^{(s)}$ and $t_j^{(b)}$ in the obvious way.

The numerical coefficient $c(T)$ can be decomposed as

$$c(T) = \prod_{1\le j\le N} \frac{K_j!}{P_j(T)} \cdot \prod_{\substack{1\le j\le N\\ r\ge 2, k\ge 0}} \frac{1}{\lambda_j(r,k)!}\left(\frac{B_r}{r!}\right)^{\lambda_j(r,k)} \tag{9.8}$$

with

$$P_j(T) = \prod_{0\le r\le 1, k\ge 1} \lambda_j(r,k)!\,k!^{\lambda_j(r,k)} \cdot \prod_{r\ge 2, k\ge 0} k!^{\lambda_j(r,k)},$$

which we can split up further as $P_j^{(s)}(T)P_j^{(b)}(T)P_j^{(1)}(T)P_j^{(\ge 2)}(T)$. The reason that we have included the factor $\lambda_j(r,k)$ into the definition of $P_j(T)$ for $0 \le r \le 1$ but not for $r \ge 2$ is that $\lambda!k!^{\lambda}$ divides $(k\lambda)!$ for all $\lambda \ge 0$ if k is strictly positive but not if $k = 0$, and the terms with $r \ge 2$ can have $k = 0$. Then the product $P_j^{(b)}(T)P_j^{(1)}(T)P_j^{(\ge 2)}(T)$ divides $\big(K_j - K_j^{(s)}\big)!$, while the first factor $P_j^{(s)}(T)$ divides $t_j^{(s)}!$ up to a p-adic unit because here k is always less than p and therefore $k!$ is not divisible by p. (Here we have made repeated use of the integrality of multinomial coefficients.) On the other hand, by Lemma 9.2 below and the submultiplicativity of D_n, the second factor in (9.8) has denominator dividing $D_{n'}$. Putting this all together, we deduce that $c(T)$ is $G(T)/D_{n'}$ times a p-adic integer for every p (always different from 2 and not dividing the denominators of the elements of R), where $G(T) = \prod_{j=1}^{N}\big(K_j!/t_j^{(s)}!\big(K_j - K_j^{(s)}\big)!\big)$. Using the submultiplicativity of D_n again, this reduces the problem to showing that $\delta_p(\mathbf{K}) \le v_p(D_{\ell-t}G(T))$ for each p, and in view of the definitions of $\delta_p(\mathbf{K})$ and D_n and of the above-mentioned formula $V_p(m) = \sum_{s\ge 1}[m/p^s]$ for the p-adic valuation of factorials, this in turn will follow if we can show that

$$\left[\frac{1}{2}\sum_{j=1}^{N}\left[\frac{K_j}{q}\right]\right] \le \left[\frac{\ell - t}{q^*}\right] + \sum_{j=1}^{N}\left(\left[\frac{K_j}{q}\right] - \left[\frac{t_j^{(s)}}{q}\right] - \left[\frac{K_j - K_j^{(s)}}{q}\right]\right) \tag{9.9}$$

for each prime power $q = p^s$ with $s \ge 1$, where $q^* := p^{s-1}(p-2)$.

For this final step, we first note that

$$\frac{\ell - t}{q^*} = \sum_{j=1}^{N} \frac{K_j^{(s)} + K_j^{(b)} + K_j^{(1)} + K_j^{(\ge 2)} - 2t_j^{(s)} - 2t_j^{(b)}}{2q^*} \ge \sum_{j=1}^{N} \frac{K_j - 2t_j^{(s)}}{2q}$$

since

$$\frac{\big(K_j^{(b)} - 2t_j^{(b)}\big)}{q^*} \ge \left(1 - \frac{2}{p}\right)\frac{K_j^{(b)}}{q^*} = \frac{K_j^{(b)}}{q}$$

(because $k \ge p$ in the terms defining $K_j^{(b)}$ and $t_j^{(b)}$) and $q^* < q$. Using that $\left[\frac{x}{2q}\right] = \left[\frac{1}{2}\left[\frac{x}{q}\right]\right]$, we deduce that

$$\left[\frac{\ell - t}{q^*}\right] \ge \left[\frac{1}{2}\sum_{j=1}^{N}\left[\frac{K_j - 2t_j^{(s)}}{q}\right]\right]$$

and hence (since $x \le y$ certainly implies $[x/2] \le [y/2]$) the inequality (9.9) will follow if we have the inequality

$$\left[\frac{K_j}{q}\right] \le \left[\frac{K_j - 2t_j^{(s)}}{q}\right] + 2\left(\left[\frac{K_j}{q}\right] - \left[\frac{t_j^{(s)}}{q}\right] - \left[\frac{K_j - K_j^{(s)}}{q}\right]\right)$$

for every $1 \le j \le N$. But this inequality is trivial, since $K_j - K_j^{(s)} \le K_j - 3t_j^{(s)} \le K_j - 2t_j^{(s)}$ (because every k in the definition of $K_j^{(s)}$ is ≥ 3) and

$$\left[\frac{K_j}{q}\right] \ge \left[\frac{K_j - 2t_j^{(s)}}{q}\right] + 2\left[\frac{t_j^{(s)}}{q}\right].$$

This completes the proof of Theorem 9.1 modulo that of the following lemma. ∎

Lemma 9.2. *For any integers $r \ge 2$ and $\lambda \ge 0$, we have*

$$\frac{1}{\lambda!}\left(\frac{B_r}{r!}\right)^{\lambda} \in \frac{1}{D_{\lambda(r-1)}}\mathbb{Z}. \tag{9.10}$$

Proof. We prove this one prime at time. By well-known results of von Staudt and Clausen, the Bernoulli number B_r ($r > 0$) has p-adic valuation -1 if $(p-1)|r$ and B_r/r is p-integral if $p-1$ does not divide r. From this we deduce that the p-adic valuation of the denominator of $B_r/r!$ is bounded above by $\left[\frac{r}{p-1}\right]$, which is $\le \left[\frac{r-1}{p-2}\right]$ since $\frac{r-1}{p-2} \ge \frac{r}{p-1}$ if $r \ge p-1$ and both expressions vanish otherwise. The p-adic valuation of the denominator of $\frac{1}{\lambda!}\left(\frac{B_r}{r!}\right)^{\lambda}$ is therefore bounded above by $V_p(\lambda) + \lambda\left[\frac{r-1}{p-2}\right]$. On the other hand, from the definition of D_n we have $v_p(D_n) = m + V_p(m) = V_p(pm)$, where $m = \left[\frac{n}{p-2}\right]$. We must therefore show that

$$V_p(\lambda) + \lambda\left[\frac{r-1}{p-2}\right] \le V_p\left(p\left[\frac{\lambda(r-1)}{p-2}\right]\right)$$

for all $\lambda \ge 0$ and $r \ge 2$. For this, we make a case distinction: if $2 \le r < p-1$, then

$$\text{l.h.s.} = V_p(\lambda) = V_p\left(p\left[\frac{\lambda}{p}\right]\right) \le V_p\left(p\left[\frac{\lambda}{p-2}\right]\right) \le \text{r.h.s.},$$

while if $r \ge p-1$, then we set $h = \left[\frac{r-1}{p-2}\right] \ge 1$ and have instead

$$\text{l.h.s.} = V_p(\lambda) + \lambda h \le \lambda h + V_p(\lambda h) = V_p(p\lambda h) \le \text{r.h.s.}$$

because $[\lambda x] \ge \lambda[x]$ for any positive real number x. ∎

We end this subsection with several further observations concerning the denominators and integrality properties of the coefficients of our divergent power series. The first is that the bound in Theorem 9.1 is not only sharp in the strong sense that it is best possible for *every* integer $n \ge 0$ and not merely that it is attained for some n, but that this optimality is reached in two very different extreme ways: in the above lemma if $r = p-1$ and $\lambda \ge 0$ is arbitrary (in which case both sides of (9.10) have the same value $V_p(p\lambda)$) and again in the calculation (9.9) in the case when only $K_j^{(b)}$ occurs and all k_i are equal to p, so that $K_j^{(b)}$ is exactly $pt_j^{(b)}$ (in other words whenever the dominating contribution in (9.6) comes from the terms with $(r,k) = (p-1,0)$ or $(0,p)$). The fact that two completely different mechanisms lead to the same function $n \mapsto D_n$ suggests that this function may be a more fundamental one than appears at first sight and may have a broader domain of applicability.

The second observation is that the universal denominator statement given by Theorem 9.1 can be sharpened by considering the series at the logarithmic level, or equivalently, by studying the denominators of the contributions from connected rather than from all Feynman diagram. This was motivated by the observation of Peter Scholze that the logarithm of the series $\Phi_0^{(4_1)}(h)$ in (1.3), which we had calculated up to order $O\big(h^{150}\big)$, had coefficients with smaller denominators than those of the series itself. Specifically, he found experimentally that the first occurrence of p^k for small primes p ($\neq 2,3$) and $k \geq 1$ in $\log \Phi_0^{(4_1)}(h)$ occurred for the coefficient of h^n with $n = k(p-1)-1$ rather than $n = k(p-2)$ as for the power series $\Phi_0^{(4_1)}(h)$ itself. At first sight this statement seems to contradict the intuition mentioned at the beginning of the section that the arithmetic of the series $\Phi_\alpha(h)$ is much simpler if one does not take their logarithms. But in fact both statements are true! The reason is that in general $\Phi_\alpha(h)$ is a linear combination of finitely many power series corresponding to the stationary points of the function being integrated (specifically, there are c^M such series, where c is the denominator of α and M can be taken to be the number of tetrahedra in a triangulation of the knot complement), and it is not reasonable to take logarithms of sums. But for $\alpha = 0$ there is only one summand, so here it is reasonable to take the logarithm, and for general α the logarithm of each of the finitely many summands of $\Phi_a(h)$ coming from the contribution to the state integral near an individual stationary point is indeed simpler than this summand itself, because it corresponds to a sum over only connected rather than over all Feynman diagrams. The final statement is given in the following theorem.

Theorem 9.3. *For each integer $n \geq 1$, define*

$$D_n^{\mathrm{conn}} = \prod_{p \text{ prime}} p^{[(n+1)/(p-1)]}.$$

Then the coefficient of h^n in $\log \Phi_0^{(K,\sigma)}(h)$ for any knot K and any $\sigma \in \mathcal{P}_K$ has denominator dividing D_n^{conn} apart from a finite set of primes depending only on K and σ. More generally, for any $\alpha \in \mathbb{Q}$ we have $\Phi_\alpha^{(K,\sigma)}(h) \in R \otimes \exp\big(\sum_{n\geq 1} Rh^n/D_n^{\mathrm{conn}}\big) \subset R_D[[h]]$.

The first few values of D_n^{conn} are given by

2, 12, 24, 720, 1440, 60480, 120960, 3628800, 7257600, 479001600, 958003200, 2615348736000, 5230697472000, ...

This sequence too appears in the online-encyclopedia of integer sequences [70] under the name `A091137` and with the formula given above, and coincides with the denominator of the Todd polynomials given in Hirzebruch's book [54, Lemmas 1.5.2 and 1.7.3] without proof and quoted from the paper [2].

The proof of the above theorem (which actually implies Theorem 9.1) is similar to the proof of Theorem 9.1 and is omitted.

Note that the "connected denominators" D_n^{conn} are considerably smaller than the "additive denominators" D_n: $D_n/n!$ is an integer growing exponentially like $(44.621\ldots+o(1))^n$, as already mentioned, while $D_n^{\mathrm{conn}}/(n+2)!$ is an integer of subexponential growth.

The final remarks concern the relation of the above results with the known integrality properties of elements of the Habiro ring. The proof of Theorem 9.1 as given above only works for the power series $\Phi_\alpha^{(\sigma)}$ with $\sigma \neq \sigma_0$, because the perturbative expansion does not apply to the case $\sigma = \sigma_0$. However, as we know, this remaining case is actually simpler because it belongs to the Habiro ring and therefore satisfies $\Phi_0^{(\sigma_0)}(h) \in \mathbb{Z}\big[\big[\mathrm{e}^h - 1\big]\big]$, and more generally $\Phi_\alpha^{(\sigma_0)}(h) \in \mathbb{Z}(\mathbf{e}(\alpha))\big[\big[\mathbf{e}(\alpha)\mathrm{e}^{-h} - 1\big]\big]$. The corresponding property no longer holds for $\sigma \neq \sigma_0$, even for $\alpha = 0$ and the figure 8 knot, and in some sense should not even be expected, because the "natural" invariant here is the completed power series $\widehat{\Phi}_0^{(4_1)}$, which contains a transcendental factor $\mathrm{e}^{\mathbf{V}(4_1)/h}$. However, if we consider the product $\sqrt{3}\Phi(h)\Phi(-h) = -\sqrt{-3}\Phi^{(1)}(h)\Phi^{(2)}(h)$, which would be unchanged if we replaced the power series by their completions, then we *do* find experimentally that it belongs to the ring $\mathbb{Z}[1/3]\big[\big[\mathrm{e}^{-h}-1\big]\big]$, i.e., after the change of variables from $q=\mathrm{e}^{-h}$ to $q=1+x$ it becomes a 3-integral power series in x. We expect, and have checked numerically, that the same is also true for 5_2 if one multiplies all three series $\Phi_0^{(5_2,\sigma_i)}$ $(i=1,2,3)$, and for $(-2,3,7)$ for both products of three series corresponding to the two number fields $\mathbb{Q}(\xi)$ and $\mathbb{Q}(\eta)$ corresponding to this knot. These properties are explained by the properties of Habiro rings for general number fields as being developed in [38].

Finally, we found experimentally that we can obtain a power series that is already integral (away from 2 and 3) in $\mathrm{e}^{-h}-1$ from $\Phi(h)=\Phi^{(4_1,\sigma_1)}(h)$ *without* multiplying it by $\Phi(-h)=-\mathrm{i}\Phi^{(4_1,\sigma_2)}(h)$ if we multiply instead by $\mathcal{E}^{(4_1,\sigma_1)}(h):=\exp\big(-\sum_{r\geq 1}^{\infty}\frac{|B_{r+1}|}{(r+1)!}C_r h^r\big)$ with C_r as in Section 8.1. (Notice that this implies the previous statement becomes $\mathcal{E}(h)\mathcal{E}(-h)=1$.) We expect that there will be a similar correction factor $\mathcal{E}^{(K,\sigma)}(h)$ for any $\Phi_\alpha^{(K,\sigma)}(h)$ and that the corrected Φ-series can be seen as the h-deformed versions of the units constructed in [10], and hope to study this too in [38].

10 Numerical aspects

In this section, we describe how the power series whose existence is predicted by the modularity conjecture can be computed effectively via a numerical computation of the Kashaev invariant, extrapolation, and recognition of algebraic numbers in a known number field. We also describe other methods that are applicable to the power series $\Phi_\alpha^{(K,\sigma)}(h)$ for σ different from σ_1, as well as the smooth truncation methods of "evaluating" factorially divergent power series at non-zero arguments that were discussed in Section 4.3. The actual numerical data for several knots will then be presented in the appendix.

10.1 Computing the power series $\Phi_\alpha^{(K,\sigma)}(h)$

In this subsection, we explain the various methods that can be used to compute the coefficients of the power series $\Phi_\alpha^{(K,\sigma)}(h)$ for all $\alpha\in\mathbb{Q}$ and $\sigma\in\mathcal{P}_K$ numerically and then exactly as algebraic numbers.

Step 1: Computing the colored Jones polynomials. To compute the Kashaev invariant $\langle K\rangle_N$ of a knot K, we use the Murakami–Murakami formula $\langle K\rangle_N=J_N\big(\mathrm{e}^{2\pi\mathrm{i}/N}\big)$, where $J_n(q)=J_{K,n}(q)$ is the n-th colored Jones polynomial, together with a theorem of T.T.Q. Lê and the first author [35] that asserts the existence of a recursion relation for the polynomials $J_n(q)$. This reduces the problem to that of computing/guessing this recursion relation concretely for a given knot. This in turn has been solved for several knots in joint work of the first author, X. Sun and C. Koutschan [33, 34, 39]. The solution required a modulo p computation of the N-th colored Jones polynomial (for several primes p and several thousand values of N), together with a careful guess of the supporting coefficients of such a recursion. In particular, the recursion was computed in [39] for the twist knots K_p with $|p|\leq 15$, was guessed in [33] for the $(-2,3,3+2p)$ pretzel knots with $|p|\leq 5$, and was computed (when $p=2$) or guessed (when $p=3,4,5$) in [34] for the double twist knots $K_{p,p}$ with $2\leq p\leq 5$.

Step 2: Computing the Kashaev invariant for large N. In order to get numerical information about the asymptotics of the Kashaev invariant $\langle K\rangle_N$, we need to be able to compute it numerically to high precision for large values of N, say of the order of $N=5000$. Although both the Kashaev invariant and the colored Jones polynomial are given by finite-dimensional terminating q-hypergeometric sums, and the latter have been programmed in `Mathematica` [3], these programs can only give the value of the colored Jones polynomial and of the Kashaev invariant for modest values of N, say, up to $N=20$, which is far less than we need for the numerical extrapolation. By using the recursion, we can compute $J_n(\zeta_N)$ numerically to high precision for N

large and $0 < n < N$. (This is far faster than computing the colored Jones polynomials $J_n(q)$ for these values of n and substituting $q = \zeta_N$ at the end.) However, this does not give the Kashaev invariant $\langle K \rangle_N = J_{K,N}(\zeta_N)$ because the recursion gives $P(q, q^n)J_n(q)$ as a linear combination of a bounded number of previous values $J_{n-i}(q)$, where $P(q, x)$ is a fixed polynomial that is always divisible by $1 - x$. To overcome this, we use the recursion relation and its first r derivatives, where $(1-x)^r \| P(q, x)$, to get a recursion for the length-r vector $\big(J_n(q), J_n'(q), \ldots, J_n^{(r)}(q)\big)$. We can the use the recursion to compute the whole vector numerically for $q = \zeta_N$ and $0 < n < N$, and the single value $J_n(q)$ for $n = N$. It follows that

Proposition 10.1. *The Kashaev invariant $\langle K \rangle_N$ of a knot K can be computed in $O(N)$ steps.*

This linear-time algorithm, which seems to be new even for the Kashaev invariant, can be used equally well to compute $J(\gamma N)$ for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ fixed and N tending to infinity, or even $J(\gamma X)$ with $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ fixed and X tending to infinity with fixed fractional part, since this simply the value of $J_n(\gamma X)$ with n equal to the denominator of γX and can be calculated by the same trick. Note that the computation takes time $O(N)$ numerically and $O\big(N^3\big)$ if we work over $\mathbb{Q}[\zeta_N]$.

Step 3: Computing $\boldsymbol{\Phi_\alpha^{(K,\sigma_1)}(h)}$. Once we know how to compute $J(\gamma N)$ for large integers N (or even $J(\gamma X)$ for large X with bounded denominator), we can we can obtain the first few coefficients of the power series $\Phi_\alpha^{(K,\sigma_1)}(h)$ numerically for a fixed rational number $\alpha = \gamma(\infty)$ by combining the quantum modularity conjecture (1.5) (or (1.6)) together with the extrapolation method of the second author (as described in detail in [47] or the appendix of [41]) or the closely related Richardson transform [5, Chapter 8]. This is quite effective and gives, for instance, the first hundred coefficients of the series (1.3) or forty coefficients of the series (1.4) in only a few minutes of computing time. We should mention, however, that this extrapolation method requires either exact numbers or else very high precision (often several hundred or even thousand decimal digits in the calculations we did.)

Step 4: Recognizing the coefficients exactly. Given that the coefficients of $\Phi_\alpha^{(K,\sigma_1)}(h)$ are conjecturally algebraic numbers in a specific number field, we can then test numerically by using the standard LLL (Lenstra–Lenstra–Lovasz) algorithm to approximate the numerically computed coefficients by rational linear combinations of a basis of this field. If this works to high precision with coefficients that are not too large and have only small primes in the denominator, then we have considerable confidence that the approximate equality is an exact one. The method is self-verifying in the sense that the success at each stage requires the correctness of the answer guessed at the previous stage.

Step 5: Computing $\boldsymbol{\Phi_\alpha^{(K,\sigma_0)}(h)}$. In this step, we explain how to compute the expansion of the element of the Habiro ring at a root of unity, in linear time. More precisely, we have:

Proposition 10.2. *The series $\Phi_\alpha^{(K,\sigma_0)}(h) + O(h)^{N+1}$ is computable in $O(N)$ steps.*

This follows from the fact that the expansion of the Habiro element at $q = \zeta_\alpha \mathrm{e}^h$ up to $O(h)^{N+1}$ requires cN terms of the cyclotomic polynomial of K, which is a linear combination of the first cN colored Jones polynomials of K. An alternative formula, inspired by Mahler's ideas on p-adic interpolation, gives the following expansion of the Habiro element:

$$\Phi_\alpha^{(K,\sigma_0)}(h) = \sum_{k=1}^{cn} \widehat{J}_n^K\big(\zeta_\alpha \mathrm{e}^h\big) + O\big(h^{n+1}\big), \tag{10.1}$$

where c is the denominator of α, $\zeta_\alpha = \mathbf{e}(\alpha)$ and

$$\widehat{J}_n^K(q) = \sum_{k=1}^{n} (-1)^{n-k} \binom{n}{k}_q q^{\frac{k(k-1)}{2}} J_{n-k+1}^K(q) q^{-\frac{n(n+1)}{2}}$$

and $\binom{n}{k}_q = (q;q)_n/((q;q)_k(q;q)_{n-k})$ is the q-binomial symbol and $(q;q)_n = (1-q)(1-q^2)\cdots(1-q^n)$ is the q-Pochhammer symbol. The right-hand side of (10.1) gives a well-defined formal power series since $\widehat{J}^K_{cn}(\zeta_\alpha \mathrm{e}^h)$ lies in $h^n\mathbb{Q}[[h]]$.

Step 6: Computing the remaining power series. Once we have the leading term in the original QMC, we can obtain the numerical terms by successively subtracting the corrections coming from the values of σ different from σ_1 as explained in Sections 4.1 (using "optimal truncation") and 4.3 (using the more precise "smooth truncation" described in Section 10.2 and in more detail in [45]), where numerical examples were also given.

Step 7: Alternative methods. In Steps 2 and 3, we discussed how to obtain the coefficients of $\Phi^{(K,\sigma_1)}_\alpha(h)$ from the original QMC together with the high-speed high-precision computation of Jones polynomials at roots of unity and numerical extrapolation; in Steps 4 and 5 we explained how to get the series $\Phi^{(K,\sigma)}_\alpha(h)$ for σ Galois conjugate to σ_1 and for $\sigma = \sigma_0$, respectively; and in Step 6 we described how to obtain the remaining series by using the *refined* quantum modularity conjecture together with optimal truncation. However, there are at least two other ways to get these other series that are of interest and are sometimes more efficient.

The first way is to use the formal Gaussian integration of Section 6 and [14, 15]. This method uses exact arithmetic and allows the computation of few terms $A^{(K,\sigma)}_\alpha(k)$ (in practice $k \le 5$) as exact algebraic numbers for many knots (such as those with ideal triangulations with up to 15 ideal tetrahedra). See also [37].

The other, which is completely different, is based on the asymptotics near roots of unity of the holomorphic functions in the upper half plane (generalized Nahm sums) that we study in [44]. Since Nahm sums converge quadratically, the values of those functions at $\tau = \alpha + \mathrm{i}/N$ can be computed in $O(\sqrt{N})$ steps and after extrapolation give a numerical computation of the algebraic numbers $A^{(K,\sigma)}_\alpha(k)$. This method, when applicable, is not only much faster (time $O(\sqrt{N})$ rather than $O(N)$), but also has the major advantage of allowing the simultaneous numerical computation of $A^{(K,\sigma)}_\alpha(k)$ for all α of a fixed denominator and for all σ in a Galois orbit which (after multiplication by D_k and by a suitable S-unit) reduces the problem of recognizing the list of coefficients $A^{(K,\sigma)}_\alpha(k)$ as algebraic numbers to the problem of recognizing numerically computed *integers*, albeit of growth $k!^2C^k$. This allowed us to compute, for instance, 100 coefficients of the series $\Phi^{(K,\sigma)}_\alpha(h)$ for the 5_2 knot for $\alpha = 0$ and $a = 1/2$ and for all three representations σ in the Galois orbit of the geometric representation, and it allowed us to compute 37 terms of the series of the $(-2,3,7)$ pretzel knot for $\alpha = 0$ and for both Galois orbits (each of size 3) of nontrivial representations σ. It is remarkable that this method allows the computation of series for representations not Galois conjugate to the geometric one, though not for the trivial representation σ_0.

Orientation conventions. Finally, we have to discuss an annoying technical point, namely, the choice of a consistent set of conventions for the two classical invariants (namely the trace field and the complex volume of a knot) and the two quantum invariants (namely the colored Jones polynomial and the Kashaev invariant of a knot). These conventions are especially important since the tables of knots rarely distinguish a knot from its mirror, and (for instance) the name 5_2 of the unique hyperbolic 5-crossing knot does not convey this distinction.

On the other hand, replacing a knot K by its mirror K^* reverses the orientation of the knot complement $M_K = S^3 \smallsetminus K$, which implies that

- $F_{K^*} = \overline{F_K}$ and $v(K^*) = \overline{v}(K)$.
- $J_{K^*,N}(q) = J_K(q^{-1})$ and $\langle K^*\rangle_N = \overline{\langle K\rangle_N}$ and $J^{K^*}(x) = J^K(-x) = \overline{J^K(x)}$.

Thus, a random orientation convention for K might not match the asymptotics whose coefficients are in a fixed subfield of the complex numbers, and not on its complex conjugate subfield.

The Jones (hence, also the colored Jones) polynomial $J_K(q) \in \mathbb{Z}[q^{\pm 1}]$ of a knot (or a link) K is uniquely determined by the following skein-relation [57]

$$qJ_{(\times)}(q) - q^{-1}J_{(\times)}(q) = \left(q^{1/2} - q^{-1/2}\right)J_{(\asymp)}(q), \qquad J_{\mathrm{unknot}}(q) = 1. \tag{10.2}$$

In particular, for the left-hand trefoil 3_1, we have: $J_{3_1}(q) = -q^4 + q^3 + q^1$. Moreover, the colored Jones polynomial $J_{K,N}(q) \in \mathbb{Z}[q^{\pm 1}]$ is normalized to be 1 at the unknot, and to equal to the Jones polynomial when $N = 2$.

Note that the `SnapPy` program [12] for computing the Jones polynomial agrees with (10.2), whereas the `Mathematica` program `KnotAtlas` [3] polynomial differs by replacing q by $1/q$:

`L = Link(braid_closure=[-1,-1,-1])`
`L.jones_polynomial()`
`-q^4 + q^3 + q^1`
	`Jones[BR[2, {-1, -1, -1}]][q]`
	`-q^-4 + q^-3 + q^-1`

We will be using the consistent orientation convention for M_K of the `SnapPy` program (when K is given as the closure of a braid, or via a planar projection, or via an augmented DTcode or Gauss code), which has the added advantage that it also gives shapes of tetrahedra corresponding to the hyperbolic structure (exactly or numerically), as well as the trace field (exactly) and the complex volume $v(K)$ (exactly or numerically).

10.2 Optimal truncation and smoothed optimal truncation

One of the main numerical aspects concerns smoothed optimal truncation, which was originally an appendix to an earlier draft of this paper but has now been relegated to a planned independent publication [45] because the methods are applicable to many problems outside the realm of quantum topology. This is a method for the numerical summation of factorially divergent series when only a finite number of coefficients are known and we do not have information about the possible analytic continuation of the Borel transform, which is the method usually used.

We already explained in Section 4.1 the naive optimal truncation, $\Phi(h)^{\mathrm{opt}}$ of a factorially divergent series $\Phi(h) = \sum_{n=0}^{\infty} A_n h^n$, defined simply as $\sum_{n=0}^{N} A_n h^n$ where N is the approximate value of n at which the term $A_n h^n$ takes on its minimum absolute value, given explicitly by $N = |B/h|$ if A_n grows like $n!B^{-n}$. The idea of smoothed optimal truncation is very simply to replace $\Phi(h)^{\mathrm{opt}}$ by a "smoothed" version $\Phi(h)^{\mathrm{smooth}}$ which is defined as $\Phi(h)^{\mathrm{opt}} + \varepsilon_N(h)$ where the exponentially small correction term $\varepsilon_N(h)$ depends on the cutoff parameter N in such a way that $\Phi(h)^{\mathrm{smooth}}$ does not jump when one changes N by 1. This means simply that we require $\varepsilon_{N-1}(h) - \varepsilon_N(h) = A_n h^n$. Of course, if we knew how to solve this equation exactly, then the function $\Phi(h)^{\mathrm{smooth}}$ would be completely independent of N and would give us a canonical way to lift the power series $\Phi(h)$ to an actual function. This is not the case, but if A_n has a known asymptotic expansion, which is true for all of the series in this paper (see Section 3.4) then we can define $\varepsilon_N(h)$ asymptotically as the product of e^{-N} and a power series in $1/n$ chosen in such a way that the desired equality $\varepsilon_{N-1}(h) - \varepsilon_N(h) = A_n h^n$ is true asymptotically to all orders in $1/n$. The details of how to do this if A_n has the asymptotic form $B^{-n}\sum_{\ell=0}^{\infty} c_\ell \Gamma(n+\kappa-\ell)$ for some real number κ and some numerical coefficients c_ℓ (as in equations (3.16), (3.17) or (3.18)) are given in [45] and will not be repeated here. The only thing that is of importance to us here is that the result of the smoothing gives an evaluation of $\Phi(h)$ that is independent of all choices (and hence gives a predicted "right" definition of the corresponding function) up to an error that is exponentially small with a better exponent than that given by naive optimal truncation. Specifically, the new error is $\mathrm{e}^{-N}(1+|C|)^{-N}$ rather than simply e^{-N} as before if, as is always the case for us, the coefficients c_ℓ themselves grow factorially like $\ell! C^{-\ell}$. Examples of this dramatic numerical improvement were given in Section 4.3, where in one case the error in evaluating a series $\Phi(h)$ was of the order of 10^{-29} using naive optimal truncation but of the order of 10^{-56} using smoothed optimal truncation.

A Numerical data for five sample knots

In this appendix, we present numerical data that support the quantum modularity conjecture for a choice of knots. Initially, we hoped that pairs of geometrically similar knots (that have identical trace fields and equal elements in the Bloch group, modulo torsion—henceforth called "sisters") might have identical or nearly identical series $\Phi_\alpha^{(K,\sigma_1)}$. With this in mind, and having already performed the computations for the 4_1 knot, we were led to consider its sister, the `m003` census manifold. The latter is not a knot complement (it is the complement of a knot in a lens space), but its 5-fold cyclic cover is the complement of the (twisted) 5-chain link, with a computable Kashaev invariant, to be compared with the 5-th power of the Kashaev invariant of the 4_1 knot. No relation between the series $\Phi_\alpha^{(K,\sigma_1)}$ was found for this pair, but the units $\varepsilon(K)_\alpha$ did match (up to roots of unity). We then tried the 5_2 knot, whose sister is the $(-2,3,7)$ pretzel knot. Here again the series $\Phi_\alpha^{(K,\sigma_1)}$ were different, but the units $\varepsilon(K)_\alpha$ matched. The final example of the 6_1 knot was chosen because its Bloch group has rank 2 and its $\mathrm{SL}_2(\mathbb{C})$-character variety is more complicated, making the verification of the QMC, the Galois invariance property (9.3) and the match with the unit of [10] more subtle, but again all three were verified numerically. For this knot we did not make any "sister" computations.

Recall the coefficients $A_\alpha^{(K,\sigma)}(k)$ of the power series $\Phi_\alpha^{(K,\sigma)}(h)$ are algebraic numbers. In this appendix, we present the numerically obtained data for $A_\alpha^{(K,\sigma)}(k)$ written in the form

$$A_\alpha^{(K,\sigma)}(k) = C_\alpha^{(K,\sigma)} \widetilde{A}_\alpha^{(K,\sigma)}(k), \tag{A.1}$$

where $C_\alpha^{(K,\sigma)} = \mu_{\sigma,\alpha} \cdot (\varepsilon_{\sigma,\alpha})^{\frac{1}{c}} \cdot \delta_\sigma^{-\frac{1}{2}}$ is given in (9.4). Note however that $C_\alpha^{(K,\sigma)}$ and $\widetilde{A}_\alpha^{(K,\sigma)}(k)$ are not canonically defined numbers, only their product is. (Since we are focusing on the geometric representation σ_1, and we are fixing the knot K, we omit the superscript (K,σ) from the notation in the right-hand side of the above equation.) We will further specify a choice of an algebraic number λ_c such that $\lambda_c^k D_k \widetilde{A}_\alpha(k) \in \mathcal{O}_{F(\zeta_\alpha)}$ is an algebraic integer, where c is the denominator of α, F is the trace field of the knot and D_n is the universal denominator (9.2). Using a basis of the free abelian group $\mathcal{O}_{F(\zeta_\alpha)}$, we can represent the above algebraic integers by lists of integer numbers.

A.1 The figure eight knot

In this appendix, we discuss the numerical aspects of the quantum modularity conjecture for the simplest hyperbolic 4_1 knot, for which we currently know how to prove the modularity conjecture. (The proof was presented in Section 8.) Needless to say, the numerically obtained results agree with the exact computation of the expansion coefficients given in Section 8. Some information about the numerical aspects of the Kashaev invariant of the 4_1 knot were already presented in the introduction, but we give some additional data (e.g., for the expansion near seventh roots of unity), since this is the most accessible knot numerically and also to illustrate the formulas occurring in the proof.

Since the knot is fixed in this section, and so is the geometric representation σ_1, we will suppress them from the notation. As mentioned in Section 1, the trace field of the 4_1 knot is $\mathbb{Q}(\sqrt{-3})$ and the torsion is $\delta = \sqrt{-3}$. Some terms of the series $\Phi_\alpha(h)$ when $c = 1$ or $c = 5$ were given in equations (1.3) and (9.1), respectively. The numerical method allows us to compute and identify the power series $\Phi_{a/c}(h)$ to any desired precision. Although the coefficients of the series $\Phi_{a/c}(h)$, divided by the constant term, is an element of the number field $\mathbb{Q}(\zeta_{3c})$ (when c is coprime to 3) or $\mathbb{Q}(\zeta_c)$ when 3 divides p, a judicious choice of the constant $C_{a/c}$, combined with the Galois invariance of the coefficients allows us to list the coefficients $\widetilde{A}_{a/p}(k)$ for $p \neq 3$ prime and for $a = 1, \dots, p-1$ by giving a $p-1$ tuple of elements in the trace field $\mathbb{Q}(\zeta_3) = \mathbb{Q}\big(\sqrt{-3}\big)$ of the knot. Furthermore, since the knot is amphicheiral, it follows that $\widetilde{A}_{a/p}(k)$ is real or purely imaginary (for k odd or even, respectively), and combined with the above discussion, allows to list the vector of coefficients $\big(\widetilde{A}_{1/p}(k), \dots, \widetilde{A}_{(p-1)/p}(k)\big)$ by a $(p-1)$-dimensional vector of rational numbers. Our numerical extrapolation method allows us to compute this tuple efficiently, and what is more, our code is self-correcting in several ways: if a wrong denominator for $\widetilde{A}_\alpha(k)$ is guessed for some k, its factorization in primes involves prime larger than $k+1$, and the precision of the computation drops in the next step by a factor of two. As a result, we were able to compute 100 terms of the series $\Phi_0(h)$ when $c = 1$, and the results agree with the computations given in [16] as well as computations obtained by a different method by the first author.

In addition to this, we computed the constant term $\Phi_\alpha(0)$ for all α with denominator a prime less than 100, and confirmed that its norm agrees with the predictions of [15, Section 4.1] for $c \leq 19$. We also computed 20 terms of the series $\Phi_\alpha(h)$ for all α with denominator a prime less than 100.

To present a sample of our computations, we start with the special case of $c = 1, 2, 3, 6$, where the c-th root of unity is in the trace field C_α and λ_α

α	0	$1/2$	$1/3$	$2/3$	$1/6$	$5/6$
C_α	$3^{-1/4}$	$3^{1/4}$	$2 \cdot 3^{-1/12}$	$3^{7/12}$	$2^2 \cdot 3^{1/12}$	$3^{17/12}$
λ_c	$72\sqrt{-3}$	$72\sqrt{-3}$	$24\sqrt{-3}$		$36\sqrt{-3}$	

it turns out that $\lambda_c^k D_k \widetilde{A}_\alpha(k)$ are integers being given by

k	$\lambda_1^k D_k \widetilde{A}_0(k)$	$\lambda_2^k D_k \widetilde{A}_{1/2}(k)$	$\lambda_3^k D_k \widetilde{A}_{1/3}(k)$	$\lambda_3^k D_k \widetilde{A}_{2/3}(k)$	$\lambda_6^k D_k \widetilde{A}_{1/6}(k)$	$\lambda_6^k D_k \widetilde{A}_{5/6}(k)$
0	1	1	1	1	1	1
1	11	41	37	25	579	201
2	697	12625	7785	6449	1224117	782865
3	724351	48022429	21535937	18981677	39903107571	29648832381
4	278392949	72296210981	24220768661	21569737445	535664049856461	412895509718949
5	244284791741	252636824949503	63245072194611	56749680285647	16693882665527364525	13164162601119392223

When $c = 4$ and $a = \pm 1 \bmod 4$, with the choice $C_{a/4} = \pm\big(3\big(2 \pm \sqrt{3}\big)\big)^{-1/4}$ and $\lambda_4 = 6\sqrt{-3}$, we can write

$$\lambda_4^k D_k \widetilde{A}_{\pm 1/4}(k) = \tilde{B}_{1/4}(k) \pm \tilde{B}_{-1/4}(k)\mathrm{i},$$

where $B_4(k) = \big(\tilde{B}_{1/4}(k), \tilde{B}_{-1/4}(k)\big) \in \mathbb{Z}^2$ with the first six values are given by

$B_4(0)$	$\langle 1, 2\rangle$
$B_4(1)$	$\langle 365, 370\rangle$
$B_4(2)$	$\langle 311785, 420386\rangle$
$B_4(3)$	$\langle 4219048201, 6325027802\rangle$
$B_4(3)$	$\langle 24805519728725, 38098972914250\rangle$
$B_4(4)$	$\langle 340419470401244075, 531593492940700894\rangle$
$B_4(5)$	$\langle 25036998069742932352139, 39557220304220645794918\rangle$

Finally, when $c = p$ is a prime different from 3, we found out for that for the primes less than 100, the constant C_α of (A.1) can be taken to be

$$C_\alpha = 3^{(-2\pm1)/4} p^{1/2} (\varepsilon_\alpha)^{1/p} \qquad \text{for} \quad p = \pm 1 \bmod 6,$$

where

$$\varepsilon_\alpha = \prod_{|k| \le \frac{p-1}{2}} (\varepsilon(p'k\alpha))^k, \qquad p' = \mp 1/4 \bmod p, \qquad \varepsilon(x) = 2\cos 2\pi(x - 1/3).$$

Note that the unit ε_α in $\mathbb{Q}(\zeta_{3p})$ that appears in the choice of C_α agrees, up to p-th powers of units, with the theoretically computed unit from equation (8.8) (for $r = 0$) below. With the above choice of C_α, the numbers $A_\alpha(k)$ lie in the field $\mathbb{Q}(\zeta_{3p})$, satisfy the Galois invariance described in detail in the introduction, and this allows them to be expressed in terms of vectors $B_p(k) = \langle \tilde{B}_{1/p}(k), \dots, \tilde{B}_{(p-1)/p}(k) \rangle \in \mathbb{Z}^{p-1}$ as follows:

$$\lambda_p^k D_k \widetilde{A}_{a/p}(k) = \sum_{b=1}^{p-1} \eta(ab/p) \tilde{B}_{b/p}(k), \qquad \eta(x) = 2\sin(2\pi(x - 1/3)),$$

where $\lambda_p = 3p^2\sqrt{-3}/2$. The vectors $B_p(k)$ for $k \le 20$ and p a prime less than 100 were numerically obtained and for $p = 5$ and $p = 7$ are given by

$B_5(0)$	$\langle -1, -4, -4, -6 \rangle$
$B_5(1)$	$\langle -55, -5140, -7660, -9690 \rangle$
$B_5(2)$	$\langle -7586065, -48629140, -58401700, -81382470 \rangle$
$B_5(3)$	$\langle -1066837647875, -5818148628500, -6620399493500, -9407838821250 \rangle$
$B_5(4)$	$\langle -51952598327049125, -274293246490488500, -309180073069692500, -440171876888046750 \rangle$
$B_5(5)$	$\langle -5814113396376116334625, -29960825153926862627500, -33500926926525556664500, -47835527737950677253750 \rangle$

and

$B_7(0)$	$\langle -20, 7, 2, 5, -14, -8 \rangle$
$B_7(1)$	$\langle -98140, 8267, -19670, 27937, -39214, -16576 \rangle$
$B_7(2)$	$\langle -2199415652, 426208447, -172006030, 524259533, -1237405358, -619260152 \rangle$
$B_7(3)$	$\langle -676432728043100, 166452454682479, -15638648253886, 168799271208365, -406506539584838, -215671594628336 \rangle$
$B_7(4)$	$\langle -86350611733284233860, 22591735955847949331, -702673247614974230, 21808440520527403561, -52829131820839184902, -28340444966866544008 \rangle$
$B_7(5)$	$\langle -25671367091358132079572196, 6928168872402051353797277, 10873595841062215161670, 6492789075493742592974935, -15896921084389159954206466, -8579075179324647599719264 \rangle$

A.2 The sister of the figure eight knot

Its quotient by $\mathbb{Z}/5\mathbb{Z}$, which is a knot in the lens space $L(5,1)$ rather than the 3-sphere, is the sister of the 4_1 knot, with the same trace field and same Bloch group invariant. We therefore expect to find similarities between the asymptotic power series associated to K_1 and to K_2.

Next, we discuss the case of the sister of the 4_1 knot, the manifold `m003` in the hyperbolic knot census [12], which is not the complement of a knot in the 3-sphere, but is the complement of a nullhomologous knot in the Lens space $L(5,1)$. This complicates things since the sister knot has no Jones (hence, no colored Jones) polynomial, and although it has a Kashaev invariant, a formula for it is not available to us. However, the 5-fold cyclic cover of the sister of the 4_1 knot is the 5-chain link L in S^3 (denoted by 10^5_3 and also by `L10n113`). This is a famous link because virtually every census manifold is a Dehn filling on it [17]. The link L has a colored Jones polynomial $J_{L,N}(q)$ (with all components colored by the same N-dimensional representation) with a formula available from [74] and a Kashaev invariant. More precisely, we have

$$J_{L,N}(q) = -\frac{1}{1-q^N} \sum_{n=0}^{N-1} \left(q^{n+1} - q^{-n}\right) c(N,n)(q)^2 c(N,n)\left(q^{-1}\right)^3,$$

where

$$c(N,n)(q) = \frac{q^{-Nn}}{(q;q)_n} \sum_{k=0}^{N-n-1} q^{-Nk} \prod_{j=k+1}^{n+k} \left(1 - q^{N-j}\right)\left(1 - q^j\right).$$

The above formula is $O(N^2)$ can be rewritten in terms of an $O(N)$ formula that has a recursion relation. However the latter has the disadvantage that the middle term of the summand ($k = N/2$) now vanishes when evaluated at $\mathbf{e}(1/N)$. To overcome this, we compute the sum from both sides by differentiation. Having done so, we tested the QMC and no suprises were found. We computed 10 terms of the series $\Phi^L_\alpha(h)$ when $\alpha = 0$ (given below) and 8 terms when $\alpha = 1/2$.

We now give the data for $\alpha = 0$. The trace field of L is $\mathbb{Q}(\sqrt{-3})$, same as for the 4_1 knot. The complex volume of L is given by

$$\mathbf{V}(L) = 5\mathbf{V}(4_1) - 3\pi^2$$

and its torsion is given by

$$\delta(L) = 2^7\sqrt{-3}.$$

Since L is a link, in (1.1), we should replace the exponent $3/2$ by $5/2$. With these changes, and with the notation of equation (A.1) we get algebraic integers $12^k D_k A^L_0(k)$ in the ring $\mathbb{Z}[\sqrt{-3}]$ and the first 10 are given as follows:

k	$12^k D_k \widetilde{A}^{K_2}_0(k)$
0	1
1	$-115\sqrt{-3} + 279$
2	$-49050\sqrt{-3} + 53286$
3	$-112270440\sqrt{-3} + 163969920$
4	$-131463532440\sqrt{-3} + 2948624280$
5	$4388324675760\sqrt{-3} - 163377997734672$
6	$-155232475000358400\sqrt{-3} + 1614884631367642560$
7	$-45605159081520871392 0\sqrt{-3} - 409415976078904226880$
8	$1201424680509251029718400\sqrt{-3} - 2426468490157451971144320$
9	$2808436744203602304238816 89600\sqrt{-3} + 767958533539384912591107225600$

However, we failed to find any relation between the series for the 4_1 knot and for the 5-fold cover of its sister.

A.3 The 5_2 knot

The pair of the 4_1 knot and its sister from the previous section in unsatisfactory in two ways. For one, the quantum modularity conjecture is proven for the 4_1 knot. Moreover, the sister of 4_1 (and its 5-fold cover) is not a knot. The next simplest hyperbolic knot after 4_1 is the 5_2 knot, whose sister is the mirror of the $(-2, 3, 7)$ pretzel knot. Sister (or geometrically similar) knots have a decomposition into a finite number of congruent ideal tetrahedra, hence they have the same trace field and equal elements in the Bloch group, modulo torsion.

A formula for the Kashaev invariant of the 5_2 knot was given in [59, equation (2.3)],

$$J^{5_2}(x) = \sum_{m=0}^{N-1}\sum_{k=0}^{m} q^{-(m+1)k}\frac{(q;q)^2_m}{(q^{-1};q^{-1})_k}, \qquad q = \mathbf{e}(x), \tag{A.2}$$

where N is the denominator of $x \in \mathbb{Q}$. After multiplication of the above by $\mathbf{e}(x)$, it agrees with the evaluation of the colored Jones polynomial $J_{5_2,N}(\mathbf{e}(x))$, where the Jones polynomial of 5_2 is $J_{5_2}(q) = q - q^2 + 2q^3 - q^4 + q^5 - q^6$. The formula (A.2) allows a computation of the Kashaev invariant in $O(N^2)$ steps, and a simplification of it was found by one of the authors [16, Section 4.1]

$$J^{5_2}(x) = \sum_{m=0}^{N-1} (q;q)_m^2 \left((q^{-1};q^{-1})_m \sum_{k=0}^{m} \frac{q^{-k^2}}{(q^{-1};q^{-1})_k^2} \right), \qquad q = \mathbf{e}(x)$$

that allows an $O(N)$-step computation of the Kashaev invariant. An alternative computation of the latter in $O(N)$-steps can be performed using the recursion relation for the colored Jones polynomial of 5_2 [39].

As mentioned in Section 2.1, the trace field of 5_2 is $F = \mathbb{Q}(\xi)$, where

$$\xi^3 - \xi^2 + 1 = 0, \qquad \xi = 0.877438833\ldots - 0.74486176661\ldots\mathrm{i}. \tag{A.3}$$

The trace field has three embeddings labeled by σ_j for $j = 1, 2, 3$ (as discussed in Section 2.1) and their volumes are given by

$$\mathbf{V}(\sigma_1) = -3R(\xi_1) + \frac{2\pi^2}{3} = 3.0241283\ldots + 2.8281220\ldots\mathrm{i},$$
$$\mathbf{V}(\sigma_2) = -3R(\xi_2) + \frac{2\pi^2}{3} = 3.0241283\ldots - 2.8281220\ldots\mathrm{i},$$
$$\mathbf{V}(\sigma_3) = 3R(\xi_3/(1-\xi_3)) - \frac{\pi^2}{3} = -1.1134545\ldots,$$

where $R(x)$ denotes the Rogers dilogarithm defined by

$$R(x) = \mathrm{Li}_2(x) + \frac{1}{2}\log(x)\log(1-x) - \frac{\pi^2}{6} \quad \text{for} \quad x \in \mathbb{C} \smallsetminus \big((-\infty,-1] \cup [1,\infty)\big).$$

The torsion of the 5_2 knot is given by $\delta(5_2) = 3\xi - 2$.

• Modularity at 0: We choose $\varepsilon(5_2)_0 = 1$, and with the notation of equation (A.1) the first eleven terms are given as follows:

k	$\big(2^3\xi^5(3\xi-2)^3\big)^k D_k \widetilde{A}_0^{5_2}(k)$
0	1
1	$-12\xi^2 + 19\xi - 86$
2	$-1343\xi^2 - 12052\xi + 14620$
3	$1381097\xi^2 + 36300408\xi - 10373787$
4	$-939821147\xi^2 - 7647561573\xi - 5587870829$
5	$114451233224986\xi^2 - 51239666382079\xi - 6305751988731$
6	$-226352740098764127\xi^2 - 631762147829071739\xi - 129887540980528920 8$
7	$-757944502306007361580\xi^2 + 1425054483652604079482\xi + 265478262327318024601 1$
8	$1678503382295602455791664 6\xi^2 - 2226340168480665471705515\xi$ $-1493068435487079406709635 8$
9	$-3735848035153601836654158090473\xi^2 - 3510831690088210470322102227368\xi$ $-4492249598242655768929879548 54$
10	$-3434598496412884157487348707287829 1\xi^2 + 2508523188778967552190692107808941 4\xi$ $+5236440463427011037040111108936206 5$

• Modularity at 1/2: We choose $\varepsilon(5_2)_{1/2} = \xi^{-5}$ and with the notation of equation (A.1) the first six terms are given as follows:

k	$\big(2\xi^5(3\xi-2)^3\big)^k D_k \widetilde{A}^{5_2}_{1/2}(k)$
0	$\xi+2$
1	$307\xi^2 - 138\xi - 628$
2	$-573109\xi^2 - 168712\xi + 457975$
3	$2096955561\xi^2 + 5077310601\xi + 1165885531$
4	$6470888990010\xi^2 - 5414463743327\xi - 10380246225743$
5	$289484322041800655\xi^2 - 138373378538474483\xi - 156775910252412286$

• Modularity at 1/3: Here, the constant term $\Phi^{5_2}_{1/3}(0)$ was numerically computed to high precision

$$\Phi^{5_2}_{1/3}(0) = -1.3478490468923913068\ldots - 1.5706460265356353326\ldots \mathrm{i}$$

but was not initially recognized. To identify it, we used the formula (A.2) for the Kashaev invariant and performed a theoretical computation analogous to the constants $S(\alpha)$ and $E_0(\alpha)$ (given in (8.8) and (8.9)) of the 4_1 knot which produced the primes

$$\mathfrak{p}_7 = \big(\xi^2-1\big)\zeta_6 - \xi + 1, \qquad \mathfrak{p}_{43} = 2\xi^2 - \xi - \zeta_6$$

of norm 7 and 43 respectively in the number field $F_3 = \mathbb{Q}(\xi, \zeta_3)$. Note that the same primes appear in [15, Section 6.2]. In addition, the above constant involves $\delta(5_2)^{-1/2}$ and a number whose third power is in F_3. After some experimentation, we concluded that

$$\Phi^{5_2}_{1/3}(0) = \mathbf{e}(1/36)\frac{1}{\sqrt{3\xi-2}}\mathfrak{p}_7^2\mathfrak{p}_{43}.$$

It follows that a representative of the unit at $\alpha = 1/3$ is given by

$$\varepsilon(5_2)_{1/3} = \mathbf{e}(1/12).$$

It was a bit of a surprise to find that the unit is torsion although the Bloch group of F_{6_1} has rank 1. On the other hand 3 (as well as 2 and a few other primes) are exceptional ones in the work [10].

Once the constant term was recognized, it turned out that we needed to separate one factor of $\mathfrak{p}_7$ in the constant term $\Phi^{5_2}_{1/3}(0)$ from the remaining terms, in order to avoid spurious denominators. With the choice of $C_{1/3} = \mathbf{e}(1/36)(3\xi-2)^{-\frac{1}{2}}\mathfrak{p}_7\mathfrak{p}_{43}$ and the notation of (A.1), the first seven terms were found to be as follows:

k	$\big(\xi^5(3\xi-2)^3\big)^k D_k \widetilde{A}^{5_2}_{1/3}(k)$
0	$\big(-\xi^2+2\xi-2\big)\zeta_6 + \big(2\xi^2-4\xi\big)$
1	$\big(717\xi^2-822\xi+947\big)\zeta_6 + \big(-2226\xi^2+1856\xi+106\big)$
2	$\big(-680145\xi^2+1283633\xi-1844797\big)\zeta_6 + \big(4731470\xi^2-1215426\xi+785050\big)$
3	$\big(-4879664798\xi^2-15547118437\xi+26771206405\big)\zeta_6$ $+\big(-20691193336\xi^2-35194065214\xi-73160959238\big)$
4	$\big(237593851209955\xi^2-123624865686699\xi+65688152000880\big)\zeta_6$ $+\big(-455730563794746\xi^2+258640669065738\xi+244974132213716\big)$
5	$\big(-855911925398142865 4\xi^2+9164193255880569642\xi-8506396294603249043\big)\zeta_6$ $+\big(-8914434881967188748\xi^2-7549553228397039176\xi+21232362162256499338\big)$
6	$\big(1206971041591026374138836\xi^2-1471979903142920023426465\xi+1526039068996370402375484\big)\zeta_6$ $+\big(2034143372251380409655636\xi^2+5390411863643322238842526\xi-9353922586016634666664696\big)$

where $\zeta_6 = \mathbf{e}(1/6)$.

A.4 The $(-2,3,7)$ pretzel knot

Next, we discuss the case of a sister the 5_2 knot, namely the (mirror of) the $(-2,3,7)$ pretzel knot. Note that the trace fields of 5_2 and $(-2,3,7)$ coincide, which allow us to use the notation of (A.3).

Unlike the case of the 4_1 and 5_2 knots, the Kashaev invariant of $(-2,3,7)$ can only be computed via the recursion of the colored Jones polynomial which was guessed in [33], with the convention that the Jones polynomial of $(-2,3,7)$ is given by $J^{(-2,3,7)}(q) = q^{-5} + q^{-7} - q^{-11} + q^{-12} - q^{-13}$. The above inhomogeneous recursion has order 6, maximal degree $(6, 58, 233)$ with respect to the shift variable, the q^n and the q variables, and contains a total of 90 terms, which can be found in [25]. In contrast, the A-polynomial of the $(-2,3,7)$ knot has maximal degree $(6, 55)$ with respect to the (L, M) variables and contains 12 terms. In addition, we multiply the Kashaev invariant of $(-2,3,7)$ by q^{-4}.

Since $(-2,3,7)$ is a sister of the 5_2 knot, they have a common trace field $\mathbb{Q}(\xi)$ given in (A.3). The trace field has three embeddings labeled by σ_j for $j = 1, 2, 3$ (as discussed in Section 2.1) and their complex volumes are given by

$$\begin{aligned}
\mathbf{V}(\sigma_1) &= -3R(\xi_1) + \frac{\pi^2}{3} = 4.6690624\ldots + 2.8281220\ldots \mathrm{i},\\
\mathbf{V}(\sigma_2) &= -3R(\xi_2) + \frac{\pi^2}{3} = 4.6690624\ldots - 2.8281220\ldots \mathrm{i},\\
\mathbf{V}(\sigma_3) &= 3R(\xi_3/(\xi_3 - 1)) + \frac{\pi^2}{3} = 0.5314795\ldots.
\end{aligned}$$

The torsion of $(-2,3,7)$ are given by $\delta((-2,3,7)) = -2(3\xi - 2)\xi^{-2}$.

- Modularity at 0: Using the notation of (A.1), we write

$$\left((2\xi^2 - 6)^3/\xi^5\right)^k D_k \widetilde{A}_0^{(-2,3,7),\sigma_1}(k) = \left(1, \xi, \xi^2\right) \cdot B_0^{(-2,3,7),\sigma_1}(k),$$

where $B_0^{(-2,3,7),\sigma_1}(k) \in \mathbb{Z}^3$ is a vector of integers with the first 11 values given by

k	$B_0^{(-2,3,7),\sigma_1}(k)$
0	$\langle 1, 0, 0\rangle$
1	$\langle -33, 128, -90\rangle$
2	$\langle 79245, -104172, 50944\rangle$
3	$\langle 333329999, -597644460, 317584318\rangle$
4	$\langle -12580573862099, 16160668928488, -9152599685200\rangle$
5	$\langle 275061075796915969, -366241217321535656, 209464837107544698\rangle$
6	$\langle -2146405978510041319481 7, 28432876033981872108244, -16179201892533998639888\rangle$
7	$\langle 3955272505750951827643863 1, -5234180126812342136382858 0, 2983803694262051507735620 6\rangle$
8	$\langle 249767901145868199725688538645, -330081248453503483229302323376, 187971265625750854805584690976\rangle$
9	$\langle -370092578601781010983364074225995049 9, 490307503368489853625660494993135832 0, -279420414366630973064161391574723931 0\rangle$
10	$\langle 392518725914904741935043787434245408953117, -519977480066306945985500543478969169892188, 296298336548750157536627179710807871873120\rangle$

- Modularity at 1/2: If we choose $\varepsilon_{1/2}((-2,3,7)) = 2\xi^5$, with the notation of (A.1), the first four terms are given by

k	$\left(4\xi(3\xi - 2)^3\right)^k D_k \widetilde{A}_{1/2}^{((-2,3,7),\sigma_1)}(k)$
0	1
1	$-225\xi^2 + 404\xi - 249$
2	$87535\xi^2 - 158073\xi + 123948$
3	$1981731163\xi^2 - 3465695160\xi + 2508787814$

- Modularity at 1/3: Here the constant term and the next two coefficients of the power series $\Phi_{1/3}^{(-2,3,7)}(h)$, $\Phi_{2/3}^{(-2,3,7)}(h)$ were computed to high precision, and using as a guidance the appearance of primes of norm 373 (conjectured in [14, Section 6.2]), we identified the constant terms

$$\Phi_{1/3}^{(-2,3,7)}(0) = \mathbf{e}(2/9)\sqrt{-\frac{27}{2(3\xi-2)}}\mathfrak{p}_{373}, \qquad \Phi_{2/3}^{(-2,3,7)}(0) = \mathbf{e}(5/9)\sqrt{-\frac{27}{2(3\xi-2)}}\mathfrak{p}'_{373},$$

where $\mathfrak{p}_{373} = \xi^2 + 2\xi\zeta_6 + 1$ and $\mathfrak{p}'_{373} = \xi^2 + 2\xi(1-\zeta_6) + 1$ are primes in $\mathbb{Q}(\xi, \zeta_6)$ of norm 373. It follows that the unit at $\alpha = 1/3$ is given by

$$\varepsilon((-2,3,7))_{1/3} = \mathbf{e}(2/3).$$

The units of 5_2 and $(-2,3,7)$ at $\alpha = 1/3$ match up to a 24-th root of unity.

As mentioned in Section 2.1, the $(-2,3,7)$ pretzel knot has 6 parabolic nonabelian representations that come in two Galois orbits of size 3 each: one is defined over the trace field (the cubic field of discriminant -23 discussed above), and another defined over the real field $\mathbb{Q}(\eta)$, the abelian field of discriminant 49. At first glance, the latter three parabolic representations (which are $\mathrm{SL}_2(\mathbb{R})$ representations of zero volume) are not seen by the Kashaev invariant. Yet, one can detect them using the asymptotics of the coefficients of the former three representations as explained in Section 10.2.

In the subsequent paper [44], we used the 6 pairs of q-series associated the $(-2,3,7)$ pretzel knot and their asymptotics to compute 37 terms of all six series $\Phi_0^{((-2,3,7),\sigma_j)}(h)$ for $j = 1, \dots, 6$. Below, we give the first 11 terms of the series associated to the abelian number field $\mathbb{Q}(\eta)$ given in Section 2.1. Consider the embeddings σ_{3+j} of the above field for $j = 1,2,3$ given in Section 2.1 which send η to $2\cos(2\pi j/7)$ and let $C_0^{((-2,3,7),\sigma_{3+j})} = \sqrt{(\eta_j - 2)/14}$. The complex volumes of σ_{3+j} are given by

$$\mathbf{V}(\sigma_4) = -\frac{1}{21}\pi^2, \qquad \mathbf{V}(\sigma_5) = \frac{1}{14}\pi^2, \qquad \mathbf{V}(\sigma_6) = -\frac{1}{42}\pi^2$$

and the torsion equals to $\delta((-2,3,7),\sigma_{3+j}) = 14/(\eta_j - 2)$. Using the notation of (A.1), we write

$$7^k D_k \widetilde{A}_0^{((-2,3,7),\sigma_{3+j})}(k) = \big(1, \eta_j, \eta_j^2\big) \cdot B_0^{((-2,3,7),\sigma_{3+j})}(k),$$

where $B_0^{((-2,3,7),\sigma_{3+j})}(k) \in \mathbb{Z}^3$ is a vector of integers with the first 11 values given by

k	$B_0^{((-2,3,7),\sigma_{3+j})}(k)$
0	$\langle 1, 0, 0\rangle$
1	$\langle 43, 0, -21\rangle$
2	$\langle 3928, 63, -1491\rangle$
3	$\langle -9658210, -2570400, 8759835\rangle$
4	$\langle -12802855375, 9661452255, 660110430\rangle$
5	$\langle -42833879089694, 5736063757095, 23026249581258\rangle$
6	$\langle -360522404258392495, -58094689166990595, 278695629206010765\rangle$
7	$\langle 1084805198860949781 65, 114336214602228319050, -161431920455740612440\rangle$
8	$\langle 420957357301236147078125, -601694281205047856100870, 211820529501946639071105\rangle$
9	$\langle 2760519033900938317917577959 50, -10532914689553665256032353437 5, -930622983726598964569771715 25\rangle$
10	$\langle 383716984951192990315815672002 1580, 171203475578865055126294086051 2280,$ $-384064713086317258381330638345 6135\rangle$

A.5 The 6_1 knot

In this appendix, we look at one further knot (this time without a sister), the 6_1 knot, for two reasons. Firstly, the trace field is $F_{6_1} = \mathbb{Q}(\xi)$, a number field of discriminant 257 (a prime) where

$$\xi^4 + \xi^2 - \xi + 1 = 0, \qquad \xi = 0.5474\ldots + 0.5856\ldots\mathrm{i}.$$

The trace field has two complex embeddings, so its Bloch group has rank two, giving a nontrivial test for the unit $\varepsilon(6_1)_\alpha$. Secondly, the $\mathrm{SL}_2(\mathbb{C})$ character variety (and the corresponding A-polynomial) is a curve whose quotient modulo the involution $\iota\colon (M,L) \mapsto \big(M^{-1}, L^{-1}\big)$ is not a rational curve. It was observed by Borot that his recent work with Eynard [7] suggested a mechanism (based on the *topological recursion*) that could explain at least a weak part of the modularity conjecture, namely that the asymptotics of $J^K(\epsilon)$ (as ϵ tends to zero through rational numbers with bounded denominators), is always given by the same series $\Phi_0^K(\epsilon)$ up to a constant factor, not predicted by their model. However, Borot could make this argument precise only in the case where the space of holomorphic differentials of the corresponding spectral curve was anti-invariant under the involution $\iota\colon (M,L)\mapsto \big(M^{-1},L^{-1}\big)$. This condition is equivalent to the statement to the rationality of the quotient of the spectral curve by ι. This led us to conduct a final experiment for the 6_1 knot. The question here was whether the modularity conjecture itself might fail, or had to be modified in the context where the argument based on the work of Borot–Eynard no longer applied. Fortunately, however, we found no anomalies.

To fix conventions, the 6_1 knot is the closure of the braid word $[1,2,3,2,-4,-1,-3,2,-3,4,-3,2]$ where j (respectively, $-j$) corresponds to the standard generator s_j $\big($respective, $s_j^{-1}\big)$ of the braid group in 4 stands, and with Jones polynomial $q^{-4}-q^{-3}+q^{-2}-2q^{-1}+2-q+q^2$. The complex volume is given by

$$\mathbf{V}(6_1) = -(2R_1+R_2+R_3)-\frac{4}{3}\pi^2 = -3.0788629\ldots+3.1639632\ldots\mathrm{i}$$

where

$$R_1 = R\big({-\xi^2}\big)-\frac{1}{2}\pi\mathrm{i}\log\big({-\xi^2}\big)+\pi\mathrm{i}\log\big(1+\xi^2\big),$$
$$R_2 = R\big(1-\xi^3\big)-\pi\mathrm{i}\log\big(1-\xi^3\big)+\frac{1}{2}\pi\mathrm{i}\log\big(\xi^3\big), \qquad R_3 = R(1-\xi).$$

The torsion, a prime in F_{6_1} of norm 257 is given by

$$\delta(6_1) = 1+\xi+4\xi^2+\xi^3.$$

Here, the Kashaev invariant of the 6_1 knot was not computed using the formula given in [59, equation (24)], since the latter is an $O(N^3)$ computation, but rather in $O(N)$ steps using the recursion relation of the colored Jones polynomial. The rather complicated inhomogeneous recursion has order 4, has maximal degree $(4,15,31)$ with respect to the shift variable, the q^n and the q variables, and contains a total of 346 terms, which can be found in [24, 39]. In contrast, the A-polynomial of the 6_1 knot has maximal degree $(4,8)$ with respect to the (L,M) variables and contains 21 terms. Due to the complexity of the recursion, we were forced to use precision 3000 in `pari` when $N=1000$. The first three coefficients of $\Phi_0^{6_1}(h)$ were numerically computed at $\alpha=0$, and using the prediction of [14] and the notation of‘(9.4), those algebraic numbers were identified as follows:

$$\Phi_0^{6_1}(h) = \frac{1}{\sqrt{\delta}}\Bigg(1+\frac{194\xi^3-331\xi^2+207\xi-245}{2^3\cdot 3\cdot\delta(6_1)^3}h + \frac{-154734\xi^3-34354\xi^2+127399\xi-119864}{2^7\cdot 3^2\cdot\delta(6_1)^6}h^2+O\big(h^3\big)\Bigg).$$

Acknowledgments

The authors would like to thank the anonymous referees for their extraordinarily careful reading of the manuscript and their detailed suggestions to improve the exposition.

References

[1] Andersen J.E., Kashaev R., A TQFT from quantum Teichmüller theory, *Comm. Math. Phys.* **330** (2014), 887–934, arXiv:1109.6295.

[2] Atiyah M.F., Hirzebruch F., Cohomologie-Operationen und charakteristische Klassen, *Math. Z.* **77** (1961), 149–187.

[3] Bar-Natan D., KnotAtlas, 2005, available at `http://katlas.org`.

[4] Beem C., Dimofte T., Pasquetti S., Holomorphic blocks in three dimensions, *J. High Energy Phys.* **2014** (2014), no. 12, 177, 118 pages, arXiv:1211.1986.

[5] Bender C.M., Orszag S.A., Advanced mathematical methods for scientists and engineers, Int. Ser. Monogr. Pure Appl. Math., McGraw-Hill Book Co., New York, 1978.

[6] Bettin S., Drappeau S., Modularity and value distribution of quantum invariants of hyperbolic knots, *Math. Ann.* **382** (2022), 1631–1679, arXiv:1905.02045.

[7] Borot G., Eynard B., All order asymptotics of hyperbolic knot invariants from non-perturbative topological recursion of A-polynomials, *Quantum Topol.* **6** (2015), 39–138, arXiv:1205.2261.

[8] Boyd D.W., Rodriguez-Villegas F., Mahler's measure and the dilogarithm. (II) (2005),, arXiv:math.NT/0308041.

[9] Calegari D., Real places and torus bundles, *Geom. Dedicata* **118** (2006), 209–227, arXiv:math.GT/0510416.

[10] Calegari F., Garoufalidis S., Zagier D., Bloch groups, algebraic K-theory, units, and Nahm's conjecture, *Ann. Sci. Éc. Norm. Supér. (4)* **56** (2023), 383–426, arXiv:1712.04887.

[11] Cooper D., Culler M., Gillet H., Long D.D., Shalen P.B., Plane curves associated to character varieties of 3-manifolds, *Invent. Math.* **118** (1994), 47–84.

[12] Culler M., Dunfield N., Weeks J., SnapPy, a computer program for studying the topology of 3-manifolds, 2015, available at `http://snappy.computop.org/`.

[13] Dimofte T., Gaiotto D., Gukov S., Gauge theories labelled by three-manifolds, *Comm. Math. Phys.* **325** (2014), 367–419, arXiv:1108.4389.

[14] Dimofte T., Garoufalidis S., The quantum content of the gluing equations, *Geom. Topol.* **17** (2013), 1253–1315, arXiv:1202.6268.

[15] Dimofte T., Garoufalidis S., Quantum modularity and complex Chern–Simons theory, *Commun. Number Theory Phys.* **12** (2018), 1–52, arXiv:1511.05628.

[16] Dimofte T., Gukov S., Lenells J., Zagier D., Exact results for perturbative Chern–Simons theory with complex gauge group, *Commun. Number Theory Phys.* **3** (2009), 363–443, arXiv:0903.2472.

[17] Dunfield N.M., Thurston W.P., The virtual Haken conjecture: experiments and examples, *Geom. Topol.* **7** (2003), 399–441, arXiv:math.GT/0209214.

[18] Ekholm T., Gruen A., Gukov S., Kucharski P., Park S., Sulkowski P., $\widehat{Z}$ at large N: from curve counts to quantum modularity, *Comm. Math. Phys.* **396** (2022), 143–186, arXiv:2005.13349.

[19] Faddeev L.D., Discrete Heisenberg–Weyl group and modular group, *Lett. Math. Phys.* **34** (1995), 249–254, arXiv:hep-th/9504111.

[20] Faddeev L.D., Kashaev R.M., Quantum dilogarithm, *Modern Phys. Lett. A* **9** (1994), 427–434, arXiv:hep-th/9310070.

[21] Fock V., Goncharov A., Moduli spaces of local systems and higher Teichmüller theory, *Publ. Math. Inst. Hautes Études Sci.* **103** (2006), 1–211, arXiv:math.AG/0311149.

[22] Gang D., Kim S., Yoon S., Adjoint Reidemeister torsions from wrapped M5-branes, *Adv. Theor. Math. Phys.* **25** (2021), 1819–1845, arXiv:1911.10718.

[23] Garoufalidis S., Chern–Simons theory, analytic continuation and arithmetic, *Acta Math. Vietnam.* **33** (2008), 335–362, arXiv:0711.1716.

[24] Garoufalidis S., Twist knots data, 2010, available at `http://people.mpim-bonn.mpg.de/stavros/publications/twist.knot.data`.

[25] Garoufalidis S., Pretzel knots data, 2012, available at `http://people.mpim-bonn.mpg.de/stavros/publications/pretzel.data`.

[26] Garoufalidis S., Goerner M., Zickert C.K., The Ptolemy field of 3-manifold representations, *Algebr. Geom. Topol.* **15** (2015), 371–397, arXiv:1401.5542.

[27] Garoufalidis S., Gu J., Mariño M., The resurgent structure of quantum knot invariants, *Comm. Math. Phys.* **386** (2021), 469–493, arXiv:2007.10190.

[28] Garoufalidis S., Gu J., Mariño M., Peacock patterns and resurgence in complex Chern–Simons theory, *Res. Math. Sci.* **10** (2023), 29, 67 pages, arXiv:2012.00062.

[29] Garoufalidis S., Kashaev R., Evaluation of state integrals at rational points, *Commun. Number Theory Phys.* **9** (2015), 549–582, arXiv:1411.6062.

[30] Garoufalidis S., Kashaev R., From state integrals to q-series, *Math. Res. Lett.* **24** (2017), 781–801, arXiv:1304.2705.

[31] Garoufalidis S., Kashaev R., The descendant colored Jones polynomials, *Pure Appl. Math. Q.* **19** (2023), 2307–2334, arXiv:2108.07553.

[32] Garoufalidis S., Kashaev R., Zagier D., A modular quantum dilogarithm and invariants of 3-manifolds, in preparation.

[33] Garoufalidis S., Koutschan C., The noncommutative A-polynomial of $(-2,3,n)$ pretzel knots, *Exp. Math.* **21** (2012), 241–251, arXiv:1101.2844.

[34] Garoufalidis S., Koutschan C., Irreducibility of q-difference operators and the knot 7_4, *Algebr. Geom. Topol.* **13** (2013), 3261–3286, arXiv:1211.6020.

[35] Garoufalidis S., Lê T.T.Q., The colored Jones function is q-holonomic, *Geom. Topol.* **9** (2005), 1253–1293, arXiv:math.GT/0309214.

[36] Garoufalidis S., Lê T.T.Q., From 3-dimensional skein theory to functions near $\mathbb{Q}$, *Ann. Inst. Fourier*, to appear, arXiv:2307.09135.

[37] Garoufalidis S., Sabo E., Scott S., Exact computation of the n-loop invariants of knots, *Exp. Math.* **25** (2016), 125–129, arXiv:6458.2015.

[38] Garoufalidis S., Scholze P., Wheeler C., Zagier D., The Habiro ring of a number field, in preparation.

[39] Garoufalidis S., Sun X., The non-commutative A-polynomial of twist knots, *J. Knot Theory Ramifications* **19** (2010), 1571–1595, arXiv:0802.4074.

[40] Garoufalidis S., Thurston D.P., Zickert C.K., The complex volume of $\mathrm{SL}(n,\mathbb{C})$-representations of 3-manifolds, *Duke Math. J.* **164** (2015), 2099–2160, arXiv:1111.2828.

[41] Garoufalidis S., van der Veen R., Asymptotics of classical spin networks (with an appendix by Don Zagier), *Geom. Topol.* **17** (2013), 1–37, arXiv:0902.3113.

[42] Garoufalidis S., Zagier D., Asymptotics of Nahm sums at roots of unity, *Ramanujan J.* **55** (2021), 219–238, arXiv:1812.07690.

[43] Garoufalidis S., Zagier D., Hyperbolic 3-manifolds, the Bloch group, and the work of Walter Neumann, *Celebratio Math.*, 2023, available at `https://celebratio.org/Neumann_WD/article/1106`.

[44] Garoufalidis S., Zagier D., Knots and their related q-series, *SIGMA* **19** (2023), 082, 39 pages, arXiv:2304.09377.

[45] Garoufalidis S., Zagier D., Resummation of factorially divergent series, in preparation.

[46] Goette S., Zickert C.K., The extended Bloch group and the Cheeger–Chern–Simons class, *Geom. Topol.* **11** (2007), 1623–1635, arXiv:0705.0500.

[47] Grünberg D.B., Moree P., Zagier D., Sequences of enumerative geometry: congruences and asymptotics, *Exp. Math.* **17** (2008), 409–426, arXiv:math.NT/0610286.

[48] Gukov S., Three-dimensional quantum gravity, Chern–Simons theory, and the A-polynomial, *Comm. Math. Phys.* **255** (2005), 577–627, arXiv:hep-th/0306165.

[49] Gukov S., Mariño M., Putrov P., Resurgence in complex Chern–Simons theory, arXiv:1605.07615.

[50] Gukov S., Pei D., Putrov P., Vafa C., BPS spectra and 3-manifold invariants, *J. Knot Theory Ramifications* **29** (2020), 2040003, 85 pages, arXiv:2040003.

[51] Gunningham S., Jordan D., Safronov P., The finiteness conjecture for skein modules, *Invent. Math.* **232** (2023), 301–363, arXiv:1908.05233.

[52] Habiro K., On the quantum sl_2 invariants of knots and integral homology spheres, in Invariants of Knots and 3-manifolds, *Geom. Topol. Monogr.*, Vol. 4, Geometry & Topology Publications, Coventry, 2002, 55–68, arXiv:math.GT/0211044.

[53] Hikami K., 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, arXiv:math.QA/0604094.

[54] Hirzebruch F., Topological methods in algebraic geometry, *Class. Math.*, Springer, Berlin, 1995.

[55] Hirzebruch F., Zagier D., The Atiyah–Singer theorem and elementary number theory, *Math. Lect. Ser.*, Vol. 3, Publish or Perish, Inc., Boston, MA, 1974.

[56] Igusa J., Theta functions, *Grundlehren Math. Wiss.*, Vol. 194, Springer, New York, 1972.

[57] Jones V.F.R., Hecke algebra representations of braid groups and link polynomials, *Ann. of Math.* **126** (1987), 335–388.

[58] Kashaev R., A link invariant from quantum dilogarithm, *Modern Phys. Lett. A* **10** (1995), 1409–1418, arXiv:q-alg/9504020.

[59] Kashaev R., The hyperbolic volume of knots from the quantum dilogarithm, *Lett. Math. Phys.* **39** (1997), 269–275, arXiv:q-alg/9601025.

[60] Kashaev R., Mangazeev V., Stroganov Yu., Star-square and tetrahedron equations in the Baxter–Bazhanov model, *Internat. J. Modern Phys. A* **8** (1993), 1399–1409.

[61] Kontsevich M., Talks on resurgence, July 20, 2020 and August 21, 2020.

[62] Lawrence R., Zagier D., Modular forms and quantum invariants of 3-manifolds, *Asian J. Math.* **3** (1999), 93–107.

[63] Lewis J., Zagier D., Period functions for Maass wave forms. I, *Ann. of Math.* **153** (2001), 191–258, arXiv:math.NT/0101270.

[64] Murakami H., Murakami J., The colored Jones polynomials and the simplicial volume of a knot, *Acta Math.* **186** (2001), 85–104, arXiv:math/9905075.

[65] Neumann W.D., Combinatorics of triangulations and the Chern–Simons invariant for hyperbolic 3-manifolds, in Topology '90 (Columbus, OH, 1990), *Ohio State Univ. Math. Res. Inst. Publ.*, Vol. 1, de Gruyter, Berlin, 1992, 243–271.

[66] Neumann W.D., Extended Bloch group and the Cheeger–Chern–Simons class, *Geom. Topol.* **8** (2004), 413–474, arXiv:math/0307092.

[67] Neumann W.D., Zagier D., Volumes of hyperbolic three-manifolds, *Topology* **24** (1985), 307–332.

[68] Ohtsuki T., A polynomial invariant of rational homology 3-spheres, *Invent. Math.* **123** (1996), 241–257.

[69] Rademacher H., Zur Theorie der Dedekindschen Summen, *Math. Z.* **63** (1956), 445–463.

[70] Sloane N., Online encyclopedia of integer sequences, available at `https://oeis.org`.

[71] Suslin A.A., K_3 of a field, and the Bloch group, *Proc. Steklov Inst. Math.* **1991** (1991), no. 4, 217–239.

[72] Thurston W., The geometry and topology of 3-manifolds, *Universitext*, Springer, Berlin, 1977.

[73] Turaev V.G., The Yang–Baxter equation and invariants of links, *Invent. Math.* **92** (1988), 527–553.

[74] van der Veen R., Proof of the volume conjecture for Whitehead chains, *Acta Math. Vietnam.* **33** (2008), 421–431, arXiv:math.GT/0611181.

[75] Vlasenko M., Zwegers S., Nahm's conjecture: asymptotic computations and counterexamples, *Commun. Number Theory Phys.* **5** (2011), 617–642, arXiv:1104.4008.

[76] Wheeler C., Modular q-difference equations and quantum invariants of hyperbolic three-manifolds, Ph.D. Thesis, University of Bonn, 2023.

[77] Wilf H.S., Zeilberger D., An algorithmic proof theory for hypergeometric (ordinary and "q") multi-sum/integral identities, *Invent. Math.* **108** (1992), 575–633.

[78] Witten E., Analytic continuation of Chern–Simons theory, in Chern–Simons Gauge Theory: 20 Years After, *AMS/IP Stud. Adv. Math.*, Vol. 50, American Mathematical Society, Providence, RI, 2011, 347–446, arXiv:1001.2933.

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

[80] Witten E., Two lectures on the Jones polynomial and Khovanov homology, in Lectures on Geometry, Clay Lect. Notes, Oxford University Press, Oxford, 2017, 1–27, arXiv:1401.6996.

[81] Witten E., Searching for new invariants of 4-manifolds and knots, January 2020, Auckland, New Zealand.

[82] Zagier D., Vassiliev invariants and a strange identity related to the Dedekind eta-function, *Topology* **40** (2001), 945–960.

[83] Zagier D., The dilogarithm function, in Frontiers in Number Theory, Physics, and Geometry. II, Springer, Berlin, 2007, 3–65.

[84] Zagier D., Quantum modular forms, in Quanta of Maths, *Clay Math. Proc.*, Vol. 11, American Mathematical Society, Providence, RI, 2010, 659–675.

[85] Zagier D., Holomorphic quantum modular forms, in preparation.

[86] Zagier D., Gangl H., Classical and elliptic polylogarithms and special values of L-series, in The Arithmetic and Geometry of Algebraic Cycles (Banff, AB, 1998), *NATO Sci. Ser. C Math. Phys. Sci.*, Vol. 548, Kluwer, Dordrecht, 2000, 561–615.

[87] Zickert C.K., The volume and Chern–Simons invariant of a representation, *Duke Math. J.* **150** (2009), 489–532, arXiv:0710.2049.

[88] Zickert C.K., The extended Bloch group and algebraic K-theory, *J. Reine Angew. Math.* **704** (2015), 21–54, arXiv:0910.4005.