LINK HOMOLOGY AND FROBENIUS EXTENSIONS II

MIKHAIL KHOVANOV AND LOUIS-HADRIEN ROBERT

ABSTRACT. The first two sections of the paper provide a convenient scheme and additional diagrammatics for working with Frobenius extensions responsible for key flavours of equivariant SL(2) link homology theories. The goal is to clarify some basic structures in the theory and propose a setup to work over sufficiently non-degenerate base rings. The third section works out two related SL(2) evaluations for seamed surfaces.

CONTENTS

1. INTRODUCTION

1.1. **Road map for Frobenius extensions in $U(2)$-equivariant Khovanov homology.** A building block for the Khovanov homology [18, 3] is a specific $(1+1)$-dimensional TQFT associated to a commutative Frobenius algebra of rank two. Different versions of this TQFT exist, including those studied in [4, 22] and in many other papers [10, 33, 13, 5]. These different versions may yield different types of information. The first striking example is the construction

by Rasmussen [30] of a lower bound on the slice genus of knots based on degenerating one of these TQFTs into the Lee $(1+1)$-dimensional TQFT [27]. An attempt to sort through some of these theories, including Bar-Natan theories [4], had been made in [22].

A number of theories can be constructed via the natural action of $U(2)$ and its subgroups G on $\mathbb{S}^2 \cong \mathbb{CP}^1$, see [22]. In those theories the homology R_G of the empty link is the G-equivariant cohomology $H^*(*, \mathbb{Z})$ of a point, while the homology A_G of the unknot is the G-equivariant cohomology $H^*(\mathbb{S}^2, \mathbb{Z})$. When $G = SU(2)$ this theory allows to recover the Rasmussen invariant and for $G = U(1)$ with the diagonal embedding into $U(2)$ it produces variations of the Rasmussen invariant. In the case $G = U(2)$ rings R_G and A_G are the rings of symmetric polynomials in two variables and a particular rank two extension of the latter, respectively.

$GL(N)$ foam evaluation [20, 31, 25] formulas require working with the entire ring of polynomials in N variables rather than its subring of the symmetric polynomials. One proves that foam evaluation is symmetric, but individual terms are not and many computations need to be done in the larger ring.

Passing from the ring of symmetric polynomials to the bigger ring of all polynomials corresponds to downsizing from $U(N)$ to $U(1)^{\times N}$-equivariant cohomology, for the standard embedding $U(1)^{\times N} \subset U(N)$. Reducing to this smaller subgroup enlarges the ground ring and makes it easier to manipulate.

There are further potential benefits to working with all polynomials rather than symmetric ones. A larger ground ring delivers more possibilities for new constructions, due to the option to break the symmetry of the polynomial variables from the symmetric group acting on them to a smaller group.

We study $N = 2$ case of this construction in the paper and $U(1) \times U(1)$-equivariant homology, including carefully writing down various diagrammatic relations in the Frobenius algebras that control the theory.

Since the first version of the paper came out, two developments showed the benefit of using $U(1) \times U(1)$-equivariant homology over $U(2)$-equivariant one:

- R. Akhmechet constructed $U(1) \times U(1)$-equivariant homology for links in the solid torus [1]. Earlier attempts by several researchers to build $U(2)$-equivariant homology for link in the solid torus have not been successful. R. Akhmechet's idea is to break the symmetry between the two generators of the polynomial ground ring when defining maps for generating cobordisms in the solid torus.
- T. Sano showed that the the sign indeterminacy in the extension of Khovanov homology to tangle cobordisms can be eliminated in $U(1) \times U(1)$-equivariant theory [32]. In the non-equivariant and in the $U(2)$-equivariant theory it's unknown whether such elimination is possible. Prior to Sato's work, the indeterminacy was resolved by adding defects to cobordisms, see Clark–Morrison–Walker [13] and Caprau [11], or by introducing $GL(2)$ foams, see Blanchet [7].

Together with the use of $U(1)^{\otimes N}$ in foam evaluation, these examples easily justify taking a close look at $U(1)^{\otimes N}$-equivariant link homology. We expect that the number of uses of such

"symmetry breaking" to find new structures and refinements in link homology will continue to grow, for both $N = 2$ and arbitrary N.

In this note, without trying to be all-inclusive, we go through several key variations on the usual $U(2)$-equivariant Frobenius pair (R, A), with R and A being the equivariant $U(2)$ cohomology groups with $\mathbb{Z}$ coefficients of a point and a 2-sphere, respectively.

2. A CUBE OF FOUR RANK TWO FROBENIUS EXTENSIONS

2.1. **Road map for Frobenius extensions in $U(2)$-equivariant Khovanov homology.** We collect the four Frobenius pairs into Figure 1 below; see also Figure 3 for a more detailed view of this cube.
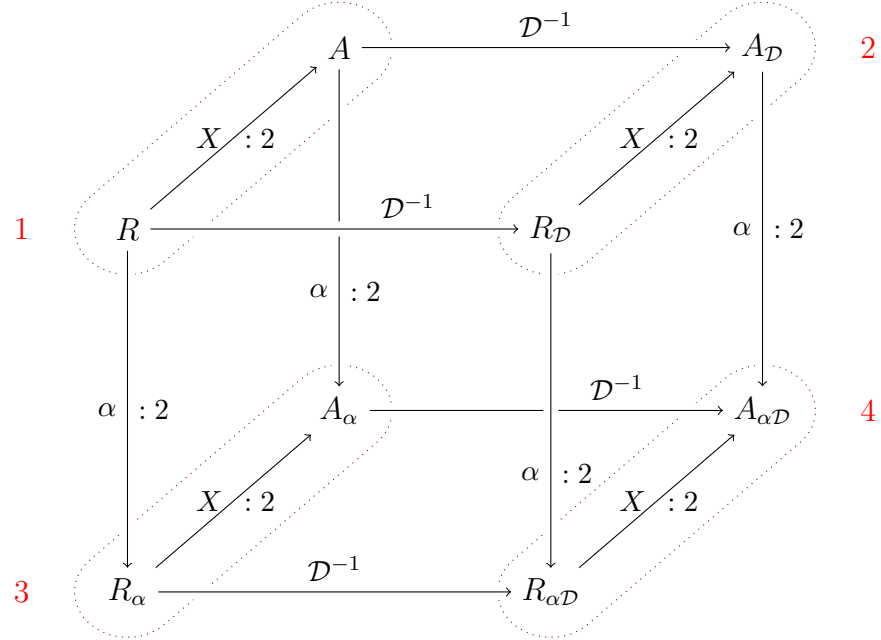

FIGURE 1. Cube of four Frobenius extensions; red numbers $1 - 4$ on the sides provide links to descriptions of these extensions.

The diagram consists of four graded Frobenius extensions of degree two

(1) $$(R, A), \ (R_{\mathcal{D}}, A_{\mathcal{D}}), \ (R_{\alpha}, A_{\alpha}), \ (R_{\alpha\mathcal{D}}, A_{\alpha\mathcal{D}})$$

and maps between them. In each of these extensions (R_*, A_*) Frobenius algebra A_* is a free graded R_* module of rank two with a basis $\{1, X\}$, with the trace map $\epsilon : A_* \longrightarrow R_*$ given by

(2) $$\epsilon(1) = 0, \quad \epsilon(X) = 1,$$

and the comultiplication

(3) $$\begin{aligned} \Delta(1) &= X \otimes 1 + 1 \otimes (X - E_1), \\ \Delta(X) &= X \otimes X - E_2 1 \otimes 1. \end{aligned}$$

Inclusions of rings $R_* \subset A_*$ in the extensions are shown by the four arrows that go diagonally pointing northeast. Horizontal and vertical arrows represent certain inclusions of ground rings R_* and induced inclusions of rings A_*.

1. The first extension (R, A) consists of graded rings
- $R = \mathbb{Z}[E_1, E_2]$, $\deg(E_1) = 2, \deg(E_2) = 4$,
- $A = R[X]/(X^2 - E_1 X + E_2)$, $\deg(X) = 2$.

Algebra A is a commutative Frobenius R-algebra of rank two with a basis $\{1, X\}$. The trace and comultiplication are given by formulas (2), (3).

Algebra A is a free graded R-module with homogeneous basis $\{1, X\}$. Elements X and $E_1 - X$ in A are roots of the polynomial $x^2 - E_1 x + E_2$ with coefficients in R, and there is an R-linear involution σ of the algebra A transposing X and $E_1 - X$. We can somewhat restore this symmetry between X and $E_1 - X$ in the notation and denote

$$(4) \qquad\qquad X_1 \;=\; X, \quad X_2 \;=\; E_1 - X,$$

so that

$$X_1 + X_2 = E_1, \quad X_1 X_2 = E_2$$

and

$$(5) \qquad\qquad \sigma(X_1) = X_2, \quad \sigma(X_2) = X_1, \quad \sigma(1) = 1,$$

with $\sigma(a + b X_1) = a + b X_2$ for $a, b \in R$. We have

$$(6) \qquad \Delta(1) \;=\; X_1 \otimes 1 - 1 \otimes X_2 \;=\; -(X_2 \otimes 1 - 1 \otimes X_1),$$
$$(7) \qquad \Delta(X_1) \;=\; X_1 \otimes X_1 - E_2\, 1 \otimes 1,$$
$$(8) \qquad \Delta(X_2) \;=\; E_2\, 1 \otimes 1 - X_2 \otimes X_2.$$

Comultiplication introduces a sign which breaks the symmetry between the two roots, and comultiplication does not commute with σ. Likewise, the trace map breaks the symmetry between the roots,

$$(9) \qquad\qquad \epsilon(X_1) = 1, \quad \epsilon(X_2) = -1,$$

and $\epsilon\sigma = -\epsilon$.

In Figure 2 we recall the usual diagrammatic conventions for 2-dimensional TQFTs and commutative Frobenius algebras, with our distinguished generator X denoted by a dot.

The definition of ϵ reads diagramatically as follows:

$$(10) \qquad\qquad \begin{array}{c}\text{(sphere)}\end{array} = 0 \quad \text{and} \quad \begin{array}{c}\text{(sphere with dot)}\end{array} = 1.$$

The composition of comultiplication and multiplication

$$(11) \qquad m\Delta \;:\; A \longrightarrow A \otimes A \longrightarrow A, \; m\Delta(1) = m(X_1 \otimes 1 - 1 \otimes X_2) = X_1 - X_2$$

is an A-module map taking 1 to the difference of roots $X_1 - X_2$. Consequently, this map is the multiplication by $X_1 - X_2$. Thus, a genus one cobordism with one boundary component, when
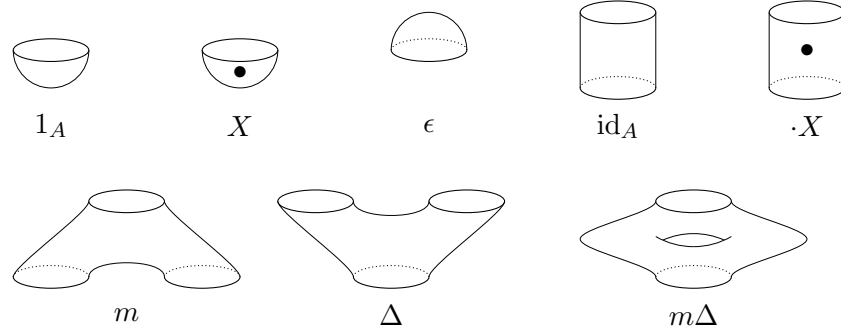
FIGURE 2. Cobordisms and the Frobenius structure maps dictionary.

viewed as a cobordism from the empty one-manifold to $\mathbb{S}^1$, represents the element $X_1 - X_2 = m\Delta(1) \in A$. A two-torus with two boundary components, when viewed as a cobordism between the circles, describes the map $A \longrightarrow A$ which is the multiplication by $X_1 - X_2$.

The trace map ϵ determines an R-linear symmetric bilinear form on A,

$$(,) \ : \ A \otimes A \longrightarrow R, \ (a, b) = \epsilon(ab).$$

This is a unimodular form on A with dual bases $(1, X)$ and $(X - E_1, 1)$. This pair of dual bases can also be written $(1, X_1)$ and $(-X_2, 1)$.

The identity map on A decomposes in these bases as usual:

(12) $\qquad a \ = \ X_1 \otimes \epsilon(a) - 1 \otimes \epsilon(X_2 a) \ = \ -X_2 \otimes \epsilon(a) + 1 \otimes \epsilon(X_1 a), \ a \in A.$

We can write this identity diagrammatically as the *neck-cutting* relation.

(13)

Notice that we use a dot of the second type (hollow dot, a little circle with white area inside) to denote the second root X_2, borrowing the idea and notation from [6]. The following notation may also be useful:

(14) $\qquad\qquad\qquad\qquad X_\bullet := X_1 = X, \ \ X_\circ := X_2 = E_1 - X,$

where we use $\bullet$ and $\circ$ as subscripts for root elements instead of $1, 2$.

A solid dot $\bullet$ placed on a vertically positioned annulus denotes the multiplication by $X_1 = X$ operator on A, see the rightmost diagram in Figure 2. Defining relation in A translates into the solid dot reduction relation

(15)

The dual (or hollowed) dot ($\circ$) denotes the multiplication by $X_2 = E_1 - X$. A torus with a boundary component is the difference of a dot and a dual dot on a disk with the same boundary.

(16)
$$\text{(torus diagram)} = \text{(bowl with } \bullet) - \text{(bowl with } \circ) = \text{(bowl with } \star) \quad \text{with} \quad \boxed{\star} = \boxed{\bullet} - \boxed{\circ}.$$

Relation (15) holds with solid dot replaced by hollow dot. We denote by $\star$ the difference of the solid and hollow dots, and use a similar subscript in X:

(17)
$$X_\star := X_1 - X_2 = X_\bullet - X_\circ = 2X - E_1 = m\Delta(1) \in A.$$

Note that $X_\star$ is in A but not in R. With our notations, a torus with one boundary component equals a disk with a star dot on it, see equation (16). In this sense, having a star dot on a surface is equivalent to adding a handle to it.

Recall that the discriminant of the quadratic polynomial $x^2 - E_1 x + E_2$ is the square of the difference $X_1 - X_2$ of roots and belongs to R,

(18)
$$\mathcal{D} = (X_1 - X_2)^2 = X_*^2 = E_1^2 - 4E_2 \in R.$$

The degree $\deg(\mathcal{D}) = 4$. Diagrammatically,

(19)
$$\boxed{\star^2} = \boxed{\star\star} = \mathcal{D} \boxed{}.$$

A closed surface of genus two evaluates to $\epsilon(\mathcal{D}) = 0$, as in [4, Section 9.2]. A closed surface of genus three evaluates to $2\mathcal{D}$:

(20)

$$= 2\mathcal{D}.$$

Significance of closed genus three surfaces and their evaluation was pointed out by Bar-Natan in [4, Section 9.2], where the Lee theory was tied to specializing the value of the closed genus three surface to 8.

Likewise, a surface of genus two with one boundary circle equals a disk times $\mathcal{D}$, see below and equations (16), (19):

$$= \text{(genus two diagram)} = \text{(bowl with } \star\star) = \mathcal{D} \text{ (bowl)}.$$

A closed surface of odd genus $2n + 1$ evaluates to $2\mathcal{D}^n$, of even genus $2n$ to zero. A surface of odd genus $2n + 1$ with a dot evaluates to $E_1 \mathcal{D}^n$, of even genus $2n$ with a dot to $\mathcal{D}^n$.

We have

(21) $\qquad \sigma(X_\star) = -X_\star, \quad \epsilon(X_\star) = 2, \quad \Delta(X_\star) = X_1 \otimes X_1 + X_2 \otimes X_2 - 2E_2 1 \otimes 1.$

Exact sequences of A-bimodules

(22) $\qquad 0 \longrightarrow A\{4\} \xrightarrow{\Delta} A \otimes_R A\{2\} \xrightarrow{\ell_X - r_X} A \otimes A \xrightarrow{m} A \longrightarrow 0,$

where $\{k\}$ is the grading shift up by k, glue into an infinite 2-periodic resolution of the identity A-bimodule A by free A-bimodules

(23) $\qquad \ldots \xrightarrow{\ell_{X_1} - r_{X_1}} A \otimes_R A\{4\} \xrightarrow{m\Delta} A \otimes_R A\{2\} \xrightarrow{\ell_{X_1} - r_{X_1}} A \otimes_R A \xrightarrow{m} A \longrightarrow 0.$

Two-dimensional TQFT (R, A), with its applications to link homology, was introduced by Bar-Natan [4]. It's also the theory labelled (R_5, A_5) in [22], with parameters h, t there corresponding to $E_1, -E_2$.

2. Frobenius extension $(R_\mathcal{D}, A_\mathcal{D})$ consists of graded rings

- $R_\mathcal{D} = R[\mathcal{D}^{-1}] = \mathbb{Z}[E_1, E_2, \mathcal{D}^{-1}]$, where $\mathcal{D}$ is the discriminant, see equation (18). Since $\deg(\mathcal{D}^{-1}) = -4$, graded algebra $R_\mathcal{D}$ is nontrivial in all even degrees.
- $A_\mathcal{D} = R_\mathcal{D}[X]/(X^2 - E_1 X + E_2)$.

Thus, algebra $R_\mathcal{D}$ is obtained from R by inverting the discriminant element (localizing along it). Algebra $A_\mathcal{D} = A \otimes_R R_\mathcal{D}$, so that the Frobenius extension $(R_\mathcal{D}, A_\mathcal{D})$ is obtained from (R, A) by inverting the discriminant $\mathcal{D}$. Algebra $A_\mathcal{D}$ is a free $R_\mathcal{D}$-module of rank two with a homogeneous basis $\{1, X\}$. Formulas (2) and (3) for the trace and the comultiplication hold in this Frobenius extension, as well as all other formulas derived above for the extension (R, A).

Since $\mathcal{D}$ is invertible of degree 4, there is an isomorphism of graded free $R_\mathcal{D}$-modules $R_\mathcal{D} \cong R_\mathcal{D}\{4\}$, so that in the Grothendieck group of graded modules one can work at most over $\mathbb{Z}[q]/(q^4 - 1)$ rather than over $\mathbb{Z}[q, q^{-1}]$ for graded R-modules.

Separability: Given a unital homomorphism $\psi : S \longrightarrow T$ of commutative rings, one can form the enveloping algebra $T^e = T \otimes_S T$, acting on T by

$$(t_1 \otimes t_2)t = t_1 t t_2$$

and a homomorphism of T^e-modules

$$m : T \otimes_S T \longrightarrow T, \quad m(a \otimes b) = ab.$$

Homomorphism ψ is called *separable* if there exists a homomorphism $\mu : T \longrightarrow T \otimes_S T$ of T^e-modules such that $m \circ \mu = \mathrm{Id}_T$,

$$m \circ \mu \ : \ T \longrightarrow T \otimes_S T \longrightarrow T.$$

Remark: Separability can be defined for noncommutative S-algebras T as well, changing the notion of the enveloping algebra to $T^e = T \otimes_S T^o$, where T^o is the opposite algebra of T.

Separability is equivalent to the existence of an idempotent $e \in T \otimes_S T$ such that $m(e) = 1$ and $eu = 0$ for any $u \in \ker(m)$. Such e is called a *separability idempotent* and is equal to $\mu(1)$.

Separability is also equivalent to the condition that the exact sequence of T^e-modules

$$0 \longrightarrow \ker(m) \longrightarrow T \otimes_S T \xrightarrow{\ m\ } T \longrightarrow 0$$

splits. T^e-submodule $\ker(m)$ is generated by the complementary idempotent $1 - e$, in the separable case. We refer to [17, 14] for more information on separable extensions.

Proposition 2.1. *The ring extension $R_{\mathcal{D}} \longrightarrow A_{\mathcal{D}}$ is separable.*

Proof. Specializing the above notations to our case, we have the algebra

$$A_{\mathcal{D}}^e = A_{\mathcal{D}} \otimes_{R_{\mathcal{D}}} A_{\mathcal{D}}.$$

Recall that $X_* = X_1 - X_2$ is the difference of roots, denoted by star dot $\star$ on a surface. From equation (18) we see that $\pm X_*$ are the two square roots in A of the discriminant $\mathcal{D} \in R$.

With $\mathcal{D}$ invertible, we can define the inverse of the $\star$ dot as the ratio $X_\star/\mathcal{D}$,

$$(24) \qquad \boxed{\star^{-1}} = \mathcal{D}^{-1} \boxed{\ \star\ }$$

This diagram describes $X_\star^{-1} = \mathcal{D}^{-1} X_\star \in A_{\mathcal{D}}$. Since $m\Delta = (X_1 - X_2)\,\mathrm{Id}_{A_{\mathcal{D}}} = X_\star\,\mathrm{Id}_{A_{\mathcal{D}}}$, see equations (11) and (16), and $X_\star$ is invertible in $A_{\mathcal{D}}$, we can rescale the comultiplication to the map

$$(25) \qquad \Delta_{\mathcal{D}} := \ell_{X_\star} \mathcal{D}^{-1} \Delta = \ell_{X_\star^{-1}} \Delta \ : \ A_{\mathcal{D}} \longrightarrow A_{\mathcal{D}}^e$$

to split the short exact sequence

$$(26) \qquad 0 \longrightarrow \ker(m) \longrightarrow A_{\mathcal{D}} \otimes_{R_{\mathcal{D}}} A_{\mathcal{D}} \xrightarrow{\ m\ } A_{\mathcal{D}} \longrightarrow 0.$$

Here $\ell_{X_\star}$ denotes the left multiplication by $X_\star$ on $A_{\mathcal{D}}^e$, but the right multiplication by the same element works as well, since Δ is a homomorphism of $A_{\mathcal{D}}^e$-modules.

Diagrammatically, $\Delta_{\mathcal{D}}$ is given by dotting the copants with the star inverse,

$$(27) \qquad \Delta_{\mathcal{D}} = \quad\boxed{\star^{-1}}\quad = \mathcal{D}^{-1} \quad\boxed{\ \star\ }\quad ,$$

Note that placing a star dot on a surface is equivalent to adding a handle, see (16). One can think of the inverse $\star^{-1}$ on a surface as an *antihandle* (of genus minus one). Informally, $\Delta_{\mathcal{D}}$ is given by copants of genus (-1). In the theory $(R_{\mathcal{D}}, A_{\mathcal{D}})$ we can turn a connected component of a surface into a "negative genus" surface by placing a suitable negative power of the $\star$ dot on it.

We have $m\Delta_{\mathcal{D}} = \mathrm{id}_{A_{\mathcal{D}}}$. Diagrammatically, $m\Delta$ adds a handle to the annulus representing the identity map of $\mathrm{id}_{A_{\mathcal{D}}}$, the state space of the circle. Presence of the star dot inverse on the surface for $\Delta_{\mathcal{D}}$ cancels out that handle.

The separability idempotent is

$$(28) \qquad \begin{aligned} e &= \Delta_{\mathcal{D}}(1) = \mathcal{D}^{-1}\Delta(X_\star) = \mathcal{D}^{-1}(X_1 \otimes X_1 + X_2 \otimes X_2 - 2E_2 1 \otimes 1) \\ &= \mathcal{D}^{-1}(X_1 \otimes 1 - 1 \otimes X_2)^2 = \mathcal{D}^{-1}\Delta(1)^2 \in A_{\mathcal{D}}^e. \qquad \square \end{aligned}$$

Diagrammatically, e is a genus one cobordism from the empty 1-manifold to the union of two circles times the scalar $\mathcal{D}^{-1}$. The handle in the genus one cobordism can be substituted by $X_\star$, see below, writing e as tube cobordism decorated by $\star$ dot times $\mathcal{D}^{-1}$. Equivalently, we can write e as a tube with $\star^{-1}$ dot on it, thus a 'genus minus one' cobordism with two boundary circles at the top.

The endomorphism ℓ_e of $\frac{e}{\mathcal{D}}$ given by multiplication by e can be depicted as follows:

Idempotent e gives a factorization

$$A_{\mathcal{D}}^e = A_{\mathcal{D}}^e e \times A_{\mathcal{D}}^e (1-e) \cong A_{\mathcal{D}} \times A_{\mathcal{D}}$$

of the algebra $A_{\mathcal{D}}^e$ into the direct product of two copies of $A_{\mathcal{D}}$.

Idempotent e is *strongly separable*, that is, fixed by the involution of $A_{\mathcal{D}}^e$ which permutes the two factors in the tensor product [17, section 5.3]. This is due to $\Delta_{\mathcal{D}}(1)$ being invariant under this involution.

Remark: Separability is closely related to vanishing of higher Hochschild homology groups. For a separable extension $S \subset T$ and any T-algebra A, there is an isomorphism $\mathrm{HH}_n^T(A) \cong \mathrm{HH}_n^S(A)$, see [28, Theorem 1.2.13]. Hochschild had shown [15] that an extension $k \subset T$ of a field k is separable iff the Hochschild homology $\mathrm{HH}_1(T, M) = 0$ for all T-bimodules M. Hochschild homology plays an important role in link homology, starting with the work of Przytycki [29] on the relation between Hochschild homology and the homology of $(2, n)$-torus links.

Our original extension $R \longrightarrow A$ is not separable, since it's easy to check that the multiplication map $m : A \otimes A \longrightarrow A$ does not split.

3. The third extension (R_α, A_α) consists of graded rings
- $R_\alpha = \mathbb{Z}[\alpha_1, \alpha_2]$, $\deg(\alpha_1) = \deg(\alpha_2) = 2$,
- $A_\alpha = R_\alpha[X]/((X - \alpha_1)(X - \alpha_2))$, $\deg(X) = 2$.

The involution $\alpha_1 \leftrightarrow \alpha_2$, denoted σ_α and induced by the permutation of indices, acts on R_α, and we identify the subring of invariants under this involution with the ring $R = \mathbb{Z}[E_1, E_2]$ via identifications $E_1 = \alpha_1 + \alpha_2$, $E_2 = \alpha_1 \alpha_2$. Under the inclusion $R \subset R_\alpha$, the latter is a free graded R-module of rank 2 with a homogeneous basis $\{1, \alpha_1\}$, for example.

Recall that A is a graded R-algebra as well. Graded R-algebras A and R_α are isomorphic, and there are two possible isomorphisms, taking $X \in A$ to either α_1 or α_2 in R_α. Notice that $X = X_1, E_1 - X = X_2$ in A and α_1, α_2 in R_α are roots of the polynomial $y^2 - E_1 y + E_2$ with coefficients in R. Any R-algebra isomorphism $A \cong R_\alpha$ will take roots X_1, X_2 to roots α_1, α_2 in

some order. Despite the existence of these isomorphisms, rings A and R_α have different origins and carry different topological interpretations in our story. This pair of isomorphisms between A and R_α does not immediately carry over to the topological side.

We can rewrite the defining relation in A_α as

$$0 = (X - \alpha_1)(X - \alpha_2) = X^2 - E_1 X + E_2,$$

recognizing the defining relation in A, but over a larger ground ring, to obtain a canonical isomorphism

$$(29) \qquad\qquad A_\alpha \;=\; R_\alpha \otimes_R A,$$

which can be taken as the definition of A_α.

Elements $X_1, X_2 \in A$ and $\alpha_1, \alpha_2 \in R_\alpha$ are roots in A_α of $y^2 - E_1 y + E_2$. Furthermore, we have

$$X_1 + X_2 = \alpha_1 + \alpha_2 = E_1, \;\; X_1 X_2 = \alpha_1 \alpha_2 = E_2.$$

The four elements $X_i - \alpha_j$, $i, j \in \{1, 2\}$, are zero divisors in R_α:

$$(X_1 - \alpha_1)(X_1 - \alpha_2) = (X_2 - \alpha_1)(X_2 - \alpha_2) = 0.$$

Up to a sign, that's only two zero divisors, since $X_2 - \alpha_1 = -(X_1 - \alpha_2)$ and $X_2 - \alpha_2 = -(X_1 - \alpha_1)$.

The comultiplication map can be simply written using $X - \alpha_i$, $i = 1, 2$:

$$
\begin{aligned}
\Delta(1) &= (X - \alpha_1) \otimes 1 + 1 \otimes (X - \alpha_2) = (X - \alpha_2) \otimes 1 + 1 \otimes (X - \alpha_1), \\
\Delta(X - \alpha_1) &= (X - \alpha_1) \otimes (X - \alpha_1), \\
\Delta(X - \alpha_2) &= (X - \alpha_2) \otimes (X - \alpha_2).
\end{aligned}
$$

We can think of $X - \alpha_i$ as shifted dots and denote them on diagrams by a small circle with i in it, ① and ②. We display below some skein relations on shifted dots. In particular, the comultiplication formulas can be interpreted as simple neck-cutting relations.

It may also be convenient to denote $X_① = X - \alpha_1, X_② = X - \alpha_2 \in A_\alpha$. Relations above can be rewritten algebraically, for example $X_① X_② = 0$.

Then $X_\star = X_① + X_②$, that is, $\boxed{\star} = \boxed{①} + \boxed{②}$.

In the ring R_α discriminant factorizes, $\mathcal{D} = (\alpha_1 - \alpha_2)^2$, which is just like its factorization $\mathcal{D} = (X_1 - X_2)^2 = X_\star^2$ in the isomorphic ring A. By analogy, one may also denote $\alpha_\star = \alpha_1 - \alpha_2$.

Involutions on R_α and A_α. Compared to R, the ring R_α has an additional symmetry, an involution σ_α, mentioned above, permuting α_1 and α_2. This R-linear involution extends to A_α transposing the roots α_1, α_2 and fixing X:

$$\sigma_\alpha(\alpha_1) = \alpha_2, \ \sigma_\alpha(\alpha_2) = \alpha_1, \ \sigma_\alpha(X) = X.$$

σ_α fixes the subalgebra R and permutes $X_①$ and $X_②$,

(30) $$\sigma_\alpha(X_①) = X_②, \ \ \sigma_\alpha(X_②) = X_①.$$

We can extend R-algebra involution σ on A to an R_α-algebra involution of A_α, also denoted σ, with

$$\sigma(X - \alpha_1) = \alpha_2 - X, \ \sigma(X - \alpha_2) = \alpha_1 - X, \ \sigma(1) = 1.$$

Algebra involutions σ and σ_α of A_α commute, and their composition $\sigma\sigma_\alpha$ is an R-linear involution which negates shifted dots,

$$\sigma\sigma_\alpha(X_①) = -X_①, \ \sigma\sigma_\alpha(X_②) = -X_②.$$

We write down a summary of these involutions of algebra A_α in Table 1.

Involution	linearity	α_1	α_2	X_1	X_2	$X_①$	$X_②$
σ	R_α-linear	α_1	α_2	X_2	X_1	$-X_②$	$X_①$
σ_α	A-linear	α_2	α_1	X_1	X_2	$X_②$	$X_①$
$\sigma\sigma_\alpha$	R-linear	α_2	α_1	X_2	X_1	$-X_①$	$-X_②$

TABLE 1. Summary of involutions σ, σ_α and $\sigma\sigma_\alpha$.

It may also be convenient to relabel σ into σ_X, by analogy with σ_α.

Equivariant cohomology: Passing from (R, A) to (R_α, A_α) corresponds to passing from $U(2)$-equivariant cohomology of a point and a 2-sphere to $U(1) \times U(1)$-equivariant cohomology of these topological spaces, where $U(1) \times U(1)$ is a maximal torus in $U(2)$. Group $U(2)$ acts on $\mathbb{S}^2$ via the identification of the latter with $\mathbb{CP}^1$ and descending from the standard action of $U(2)$ on $\mathbb{C}^2$.

4. The fourth extension $(R_{\alpha\mathcal{D}}, A_{\alpha\mathcal{D}})$ consists of graded rings
- $R_{\alpha\mathcal{D}} = \mathbb{Z}[\alpha_1, \alpha_2, (\alpha_1 - \alpha_2)^{-1}]$, $\deg(\alpha_1) = \deg(\alpha_2) = 2$,
- $A_{\alpha\mathcal{D}} = R_{\alpha\mathcal{D}}[X]/((X - \alpha_1)(X - \alpha_2))$, $\deg(X) = 2$.

This extension is obtained from the third extension (R_α, A_α) by inverting the discriminant $\mathcal{D} = (\alpha_1 - \alpha_2)^2$. Equivalently, one can invert $\alpha_1 - \alpha_2$. An element of $A_{\alpha \mathcal{D}}$ can be written uniquely as $f_1 \cdot 1 + f_2 \cdot X$ where $f_1, f_2 \in R_{\alpha \mathcal{D}}$. Let

$$(31) \qquad e_1 \ = \ \frac{X - \alpha_1}{\alpha_2 - \alpha_1}, \quad e_2 \ = \ \frac{X - \alpha_2}{\alpha_1 - \alpha_2}.$$

These two elements of $A_{\alpha \mathcal{D}}$ are mutually-orthogonal complementary idempotents,

$$(32) \qquad 1 = e_1 + e_2, \ e_1^2 = e_1, \ e_2^2 = e_2, \ e_1 e_2 = e_2 e_1 = 0.$$

Consequently, the ring $A_{\alpha \mathcal{D}}$ decomposes as the direct product

$$(33) \qquad A_{\alpha \mathcal{D}} \ = \ R_{\alpha \mathcal{D}} e_1 \times R_{\alpha \mathcal{D}} e_2.$$

The entire Frobenius algebra structure decouples as well:

$$(34) \qquad \Delta(e_1) \ = \ (\alpha_2 - \alpha_1) \cdot e_1 \otimes e_1, \quad \Delta(e_2) \ = \ (\alpha_1 - \alpha_2) \cdot e_2 \otimes e_2,$$

$$(35) \qquad \epsilon(e_1) = (\alpha_2 - \alpha_1)^{-1}, \quad \epsilon(e_2) = (\alpha_1 - \alpha_2)^{-1}.$$

Both $X_\star = X_1 - X_2$ and $\alpha_\star$ are invertible in $A_{\alpha \mathcal{D}}$, since both square to $\mathcal{D}$. The ratio

$$\frac{X_\star}{\alpha_\star} = X_\star \alpha_\star^{-1} = \frac{2X - E_1}{\alpha_1 - \alpha_2} = e_2 - e_1$$

is a degree zero invertible element other than -1 that squares to 1. Elements

$$\{1, X_\star \alpha_\star^{-1}, -1, -X_\star \alpha_\star^{-1}\} = \{1, e_2 - e_1, -1, e_1 - e_2\}$$

constitute a subgroup of degree zero invertible elements in $A_{\alpha \mathcal{D}}$ isomorphic to $\mathbb{Z}/2 \times \mathbb{Z}/2$. Since the ring $A_{\alpha \mathcal{D}}$ is not an integral domain, it may contain a finite non-cyclic subgroup of invertible elements. Since $X_\star$ is invertible, it can be put in the denominator instead of $\alpha_\star$ to recover the same idempotents:

$$(36) \qquad e_1 \ = \ \frac{\alpha_1 - X_1}{X_2 - X_1}, \quad e_2 \ = \ \frac{\alpha_1 - X_2}{X_1 - X_2}.$$

2.2. **Summary of basic properties.** In Figure 3 we redraw the cube of Frobenius extensions, adding some information about the rings involved.

Let us summarize this diagram. In it, three of the Frobenius extensions listed in (1) are obtained by base changes from the original extension (R, A):

$$(37) \quad (R_\alpha, A_\alpha) \cong R_\alpha \otimes_R (R, A), \ (R_\mathcal{D}, A_\mathcal{D}) \cong R_\mathcal{D} \otimes_R (R, A), \ (R_{\alpha \mathcal{D}}, A_{\alpha \mathcal{D}}) \cong R_{\alpha \mathcal{D}} \otimes_R (R, A).$$

The original extension (R, A) is given by the upper left diagonal arrow.

Diagonal arrows: Diagonal arrows in the diagram, pointing northeast, denote inclusions $R_* \subset A_*$ for each of the four Frobenius extensions (1). In each case A_* is a free R_*-module of rank two with a basis $\{1, X\}$. The rank and the additional basis (and generating) element X are indicated by writing ": 2" and "X" next to the arrows.

Horizontal arrows: The four horizonal arrows correspond to localizing each of the four commutative rings R, A, R_α, A_α in the vertices of the left square facet of the cube by inverting the
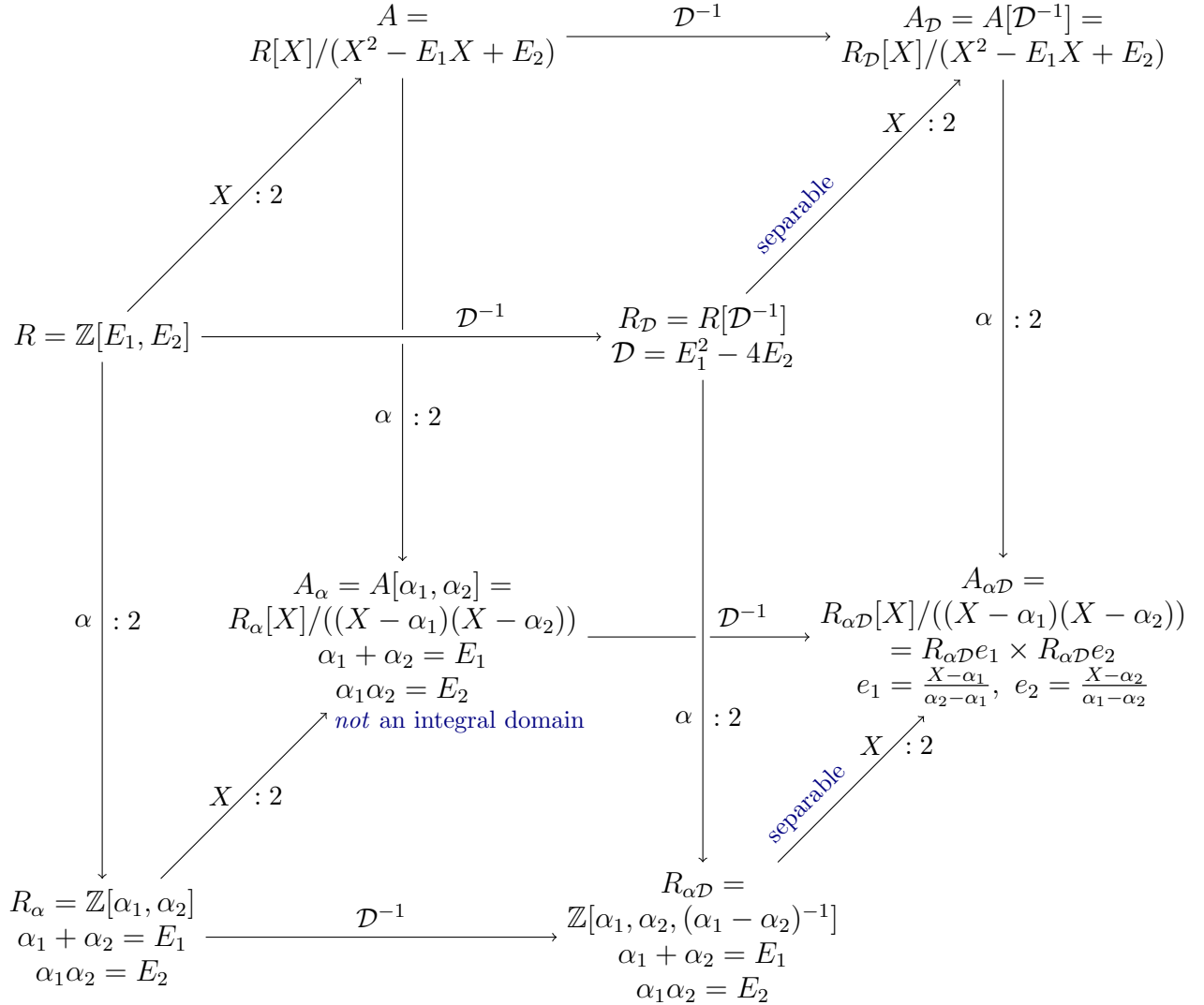
FIGURE 3. Cube of the four Frobenius extensions, with the rings defined

discriminant $\mathcal{D} \in R$. Each arrow is the inclusion of one of these four rings into the localized ring. We put $\mathcal{D}^{-1}$ by these arrows to indicate this operation. The two Frobenius extensions on the right side of the diagram are separable, but not the two on the left.

Downward arrows: The four vertical downward arrows correspond to enlarging each of the four rings in the vertices of the top square by adding roots α_1, α_2 of polynomial $y^2 - E_1 y + E_2$. This extension of rings is a rank two extension, with the larger ring a free rank two module over the smaller ring with a basis $\{1, \alpha_i\}$ for $i \in \{1, 2\}$. We indicated this transformation by writing "α" and "$: 2$" next to each vertical arrow. In particular, the leftmost vertical arrow denotes the

inclusion

$$R = \mathbb{Z}[E_1, E_2] \ \subset R_\alpha = \mathbb{Z}[\alpha_1, \alpha_2], \ \alpha_1 + \alpha_2 = E_1, \ \alpha_1\alpha_2 = E_2.$$

Thus, downward vertical arrows denote an extension where we tensor each of the four rings $R, A, R_\mathcal{D}, A_\mathcal{D}$ with R_α over R. In particular,

$$A_\alpha = A \otimes_R R_\alpha, \ \ R_{\alpha\mathcal{D}} = R_\mathcal{D} \otimes_R R_\alpha, \ \ A_{\alpha\mathcal{D}} = A_\mathcal{D} \otimes_R R_\alpha.$$

As we've mentioned, R_α is a free rank two graded R-module with a basis $\{1, \alpha_1\}$ (or $\{1, \alpha_2\}$). In fact, (R, R_α) is a rank two Frobenius extension isomorphic to the extension (R, A). There are two R-algebra isomorphisms $A \cong R_\alpha$: one of them takes X to α_1 and, necessarily, $E_1 - X$ to α_2, while the other takes the roots $(X, E_1 - X)$ of $x^2 - E_1 x + E_2$ to (α_2, α_1), correspondingly. We will not be using these isomorphisms in the paper.

Zero divisors and idempotents: Ring A_α contains zero divisors $\pm(X - \alpha_1)$, $\pm(X - \alpha_2)$, with $(X - \alpha_1)(X - \alpha_2) = 0$. In particular, it's not an integral domain. Additionally, with the determinant inverted, these zero divisors rescale into idempotents e_1, e_2 in $A_{\alpha\mathcal{D}}$, where they decompose the ring into the direct product of two copies of the ground ring $R_{\alpha\mathcal{D}}$.

Pushouts: Rings $R_{\alpha\mathcal{D}}$ and $A_{\alpha\mathcal{D}}$, together with the maps into them in the front and back squares of the diagram in Figures 1 and 3, are pushouts (and colimits) of commutative ring diagrams

(38) $$R_\alpha \longleftarrow R \longrightarrow R_\mathcal{D}, \quad A_\alpha \longleftarrow A \longrightarrow A_\mathcal{D},$$

correspondingly. In the extension $R \subset R_\alpha$, the ring R_α is free of rank two as an R-module. In particular, this is a flat and a faithfully flat extension, ditto for $A \subset A_\alpha$. The inclusion $R \subset R_\mathcal{D}$ is a localization of commutative rings and thus a flat extension. The back square of the four A_*'s in the diagram is obtained via the base change $R \longrightarrow A$ from the front square pushout of commutative rings R_*'s.

Diagonal plane symmetries of the cube: As mentioned earlier, rings A and R_α are isomorphic as R-algebras, via two possible algebra automorphisms that take X to α_1 or α_2 and $E_1 - X$ to α_2 or α_1, correspondingly:

$$\tau_1 : A \xrightarrow{\cong} R_\alpha, \ \ \tau_1(X) = \alpha_1, \ \tau_1(E_1 - X) = \alpha_2,$$
$$\tau_2 : A \xrightarrow{\cong} R_\alpha, \ \ \tau_2(X) = \alpha_2, \ \tau_2(E_1 - X) = \alpha_1.$$

Composition $\tau_1\tau_2^{-1} = \tau_2\tau_1^{-1}$ is an R-linear automorphism of R_α that transposes α_1, α_2. Likewise, $\tau_1^{-1}\tau_2 = \tau_2^{-1}\tau_1$ is an R-linear automorphism of A that transposes $X_1 = X$ and $X_2 = E_1 - X$.

Ring A_α is a free R-module of rank four with a basis, for instance, $\{1, X, \alpha_1, \alpha_1 X\}$. Isomorphisms τ_1, τ_1^{-1} extend to a algebra involution of A_α, also denoted τ_1, given by

(39) $$\tau_1(\alpha_1) = X, \ \tau_1(X) = \alpha_1, \ \tau_1(\alpha_2) = E_1 - X, \ \tau_1(E_1 - X) = \alpha_2,$$

or, in the other notation,

(40) $$\tau_1 : \alpha_1 \leftrightarrow X_1, \ \alpha_2 \leftrightarrow X_2.$$

Thus, τ_1 adds sign to $X - \alpha_1$ but fixes $X - \alpha_2$ in A_α. It acts with eigenvalue -1 on the subspace $R(X - \alpha_1)$ and with eigenvalue 1 on $R(X - \alpha_2)$. Likewise, $\tau_2(X - \alpha_1) = X - \alpha_1$ and $\tau_2(X - \alpha_2) = -(X - \alpha_2)$.

Isomorphisms τ_1, τ_2 of A_α extend to isomorphisms from $A_\mathcal{D}$ to $R_{\alpha\mathcal{D}}$ and to automorphisms of $A_{\alpha\mathcal{D}}$. They also restrict to identity automorphisms of both R and $R_\mathcal{D}$.

In this way, they act on the entire cube in Figure 3, as automorphisms of algebras $R, A_\alpha, R_\mathcal{D}$, and $A_{\alpha\mathcal{D}}$ (in the four vertices of the cube that lie on a plane through two horizontal edges), isomorphisms between A and R_α and isomorphisms between $A_\mathcal{D}$ and $R_{\alpha\mathcal{D}}$.

Properties of link homology for these four extensions.

Given an oriented link L and its diagram D with n crossings, we can construct 2^n resolutions of D into diagrams of planar circles and then use the Frobenius pair (R, A) and associated 2-dimensional TQFT to build a commutative cube with tensor powers of A (with grading shifts) placed in the vertices. Collapsing into a complex, we obtain a complex $C(D)$ of free graded R-modules. Choosing a base point on D which is not a crossing turns $C(D)$ into a complex of graded A-modules. Homology $\mathrm{H}(D)$ of $C(D)$ is a bigraded R-module, which is almost never free. With a choice of base point, $\mathrm{H}(D)$ is a graded A-module.

To the Reidemeister moves $D_1 \sim D_2$ one assigns chain homotopy equivalences $C(D_1) \cong C(D_2)$ between complexes of free graded R-modules, with induced isomorphisms on bigraded homology group $\mathrm{H}(D_1) \cong \mathrm{H}(D_2)$. With a base point and a Reidemeister move away from the base point, homotopy equivalences and homology isomorphisms become that of A-modules. For a k-component unlink U_k, the homology $\mathrm{H}(U_k) \cong A^{\otimes k}$, a free A-module of rank 2^{k-1}.

For two basepoints p_1, p_2 located on the same component and separated by a single crossing, multiplication maps on $C(D)$ by X_1 at p_1 and by X_2 at p_2 are chain homotopic.

Base changes from R to R_α, $R_\mathcal{D}$, $R_{\alpha\mathcal{D}}$ allow to define chain complexes and the corresponding homology groups

$$
\begin{align}
(41) \qquad C_\alpha(D) &= R_\alpha \otimes_R C(D), \quad \mathrm{H}_\alpha(D) = \mathrm{H}(C_\alpha(D)), \\
(42) \qquad C_\mathcal{D}(D) &= R_\mathcal{D} \otimes_R C(D), \quad \mathrm{H}_\mathcal{D}(D) = \mathrm{H}(C_\mathcal{D}(D)), \\
(43) \qquad C_{\alpha\mathcal{D}}(D) &= R_{\alpha\mathcal{D}} \otimes_R C(D), \quad \mathrm{H}_{\alpha\mathcal{D}}(D) = \mathrm{H}(C_\alpha(D)).
\end{align}
$$

We can arrange these four types of complexes and groups into a three-dimensional cube, where diagonal wiggly arrows denote passage to homology.

$$
\begin{array}{ccc}
\mathrm{H}(D) & \xrightarrow{\ \mathcal{D}^{-1}\ } & \mathrm{H}_{\mathcal{D}}(D) \\
& & \\
C(D) & \xrightarrow{\ \mathcal{D}^{-1}\ } & C_{\mathcal{D}}(D) \qquad \alpha : 2 \\
\alpha : 2 & & \alpha : 2 \\
\alpha : 2 \qquad \mathrm{H}_{\alpha}(D) & \xrightarrow{\ \mathcal{D}^{-1}\ } & \mathrm{H}_{\alpha\mathcal{D}}(D) \\
& \alpha : 2 & \\
C_{\alpha}(D) & \xrightarrow{\ \mathcal{D}^{-1}\ } & C_{\alpha\mathcal{D}}(D)
\end{array}
$$

(diagonal wiggly arrow labeled *homology*)

The four terms in the vertices of the front square are complexes $C_*(D)$ obtained from $C(D)$ by suitable base changes $R \longrightarrow R_*$. They are complexes of free graded R_*-modules. Define the four homology groups $\mathrm{H}_*(D)$ in the back square as the homology of these complexes. They are naturally modules over R_* and, if a base point is picked, modules over A_*.

Flatness of the extensions $R \longrightarrow R_*$ implies that the natural maps

$$
(44) \qquad\qquad\qquad \mathrm{H}(D) \otimes_R R_* \longrightarrow \mathrm{H}_*(D)
$$

are isomorphisms of R_*-modules, also giving isomorphisms passing from diagrams to links

$$
(45) \qquad\qquad\qquad \mathrm{H}(L) \otimes_R R_* \xrightarrow{\ \cong\ } \mathrm{H}_*(L)
$$

These homology theories extend to tangles as usual [19], by first extending $SL(2)$-equivariant rings H^n over R (with $H^1 \cong A$) via base change $R \longrightarrow R_*$ to rings H^n_*. Bimodules and bimodule homomorphisms for flat tangles and their cobordisms extend as well, resulting in suitable 2-functors from the 2-category of flat tangle cobordisms to the 2-category of homogeneous bimodules over H^n_*, over all n.

Extension to tangle cobordisms also works as usual. Due to flatness of our base changes, invariance of maps induced by tangle cobordisms up to an overall sign follows from the corresponding invariance for the theory built out of (R, A). For the invariance of link and tangle cobordism maps we refer to the original papers [4, 16, 21]. This invariance up to a sign can be stated, for tangle cobordisms, in the language of maps between complexes of graded bimodules in the homotopy category.

Isomorphisms (45) are functorial in the link cobordism category, up to an overall sign.

Taking care of the sign requires modifying the theory and working with seamed circles or defect lines as in Caprau [10, 11], Clark–Morrison–Walker [13] and Vogel [33], or $GL(2)$ foams

as in Blanchet [7]. Defect lines in these theories are discussed below, in Section 4.1, mostly via the evaluation approach simplified from $GL(2)$ foams to surfaces with seamed circles.

α-homology: As a complex of graded R-modules,

$$C_\alpha(D) \cong C(D) \cdot 1 \oplus C(D) \cdot \alpha_1,$$

thus isomorphic to two copies of $C(D)$, with a shift. Multiplication by α_1 commutes with taking cohomology. As graded R-modules,

$$\mathrm{H}_\alpha(D) \cong \mathrm{H}(D) \oplus \mathrm{H}(D) \cdot \alpha_1,$$

and, with a choice of a base point in D, that's an isomorphism of graded A-modules.

We see that the α-homology $\mathrm{H}_\alpha(L)$, as a graded A-module, is isomorphic to the sum of two copies of $\mathrm{H}(L)$, one with a grading shift by $\{2\}$, the degree of α_1, but has an additional symmetry induced by the transposition of α_1 and α_2 and boasts zero divisors in the homology of the unknot, viewed as a commutative Frobenius algebra via the pants, cup and cap cobordisms.

$\alpha\mathcal{D}$-homology: Homology theory $\mathrm{H}_{\alpha\mathcal{D}}$ associated to the pair $(R_{\alpha\mathcal{D}}, A_{\alpha\mathcal{D}})$ is essentially the Lee homology. The decomposition of $A_{\alpha\mathcal{D}}$ into the direct product (33) via the idempotents in equation (31) parallels corresponding decomposition in the Lee homology. In particular, arguments in Lee [27] and Rasmussen [30] apply here as well and show that $\mathrm{H}_{\alpha\mathcal{D}}(L)$ is a free $R_{\alpha\mathcal{D}}$-module of rank 2^k, where k is the number of connected components of link L.

Multiplication by $\mathcal{D}$ is a degree four isomorphism of $\mathrm{H}_{\alpha\mathcal{D}}(L)$. Quantum grading, lifting powers of q, is $\mathbb{Z}/4$-periodic in this theory. Invertibility of $\mathcal{D}$ can be used to reduce the homology to one with a $\mathbb{Z}/4$-grading in the q-direction, see the remark below.

Furthermore,

(46) $$\mathrm{H}_{\alpha\mathcal{D}}(L) \cong \mathrm{H}_\mathcal{D}(L) \otimes_{R_\mathcal{D}} R_{\alpha\mathcal{D}} = \mathrm{H}_\mathcal{D}(L) \cdot 1 \oplus \mathrm{H}_\mathcal{D}(L) \cdot \alpha_1.$$

Consequently, $\mathrm{H}_\mathcal{D}(L)$ is "half the size" of $\mathrm{H}_{\alpha\mathcal{D}}(L)$.

Remark: There are versions of $C_\mathcal{D}(D)$ and $\mathrm{H}_\mathcal{D}(D)$ chain complexes and homology groups with the q-grading reduced to $\mathbb{Z}/4$. To define this theory, form the $\mathbb{Z}/4$-graded quotient ring $R_\lambda = R/(\mathcal{D} - \lambda)$ (another convenient notation is $R_{\lambda,\mathcal{D}}$), where $\lambda \in \{1, -1\}$ is an invertible degree zero element of R (or a more general invertible element of a ground ring different from R). We pick invertible λ to make $\mathcal{D}$ invertible in R_λ. Ring $A_\lambda = A \otimes_R R_\lambda$, which can also be denoted $A_{\lambda,\mathcal{D}}$, is defined via the base change. The pair (R_λ, A_λ) is a rank two Frobenius extension. Since $\deg(\mathcal{D})$ is four, rings R_λ, A_λ are $\mathbb{Z}/4$-graded only, ditto for the complexes $C_\lambda(D) = C(D) \otimes_R R_\lambda$ and $\mathrm{H}_\lambda(D) = \mathrm{H}(C_\lambda(D))$. A similar quotient construction works for the Lee homology $\mathrm{H}_{\alpha\mathcal{D}}(D)$.

3. Ring and module involutions and defect lines

3.1. **Involution σ and Galois action.** Algebra A carries an R-linear involution σ given by $\sigma(X) = E_1 - X$. As an algebra involution, it satisfies $\sigma(1) = 1$. It transposes the roots $X_1 = X, X_2 = E_1 - X$ in A of the polynomial $y^2 - E_1 y + E_2$ with coefficients in $R \subset A$, $\sigma(X_1) = X_2, \sigma(X_2) = X_1$. Ground ring R is also the subring of σ-invariants of A. Notice that the (-1)-eigenspace of σ is zero, which is not surprising since we are not working over a field.

Via a base change $R \longrightarrow R_*$ this involution extends to an involution σ of the R_*-algebra A_*, for any of the four Frobenius extensions (R_*, A_*) in (1). In fact, it works for any base change $R \longrightarrow R_*$ of commutative rings, but we restrict to these four cases.

Involution σ acts R_*-linearly on the free R_*-module A_* with a basis $\{1, X\}$. In this basis the action is given by the matrix

$$(47) \qquad\qquad \sigma \longmapsto \begin{pmatrix} 1 & E_1 \\ 0 & -1 \end{pmatrix} .$$

Let $A_*[\sigma]$ be the crossed product of A_* with the group ring of the order two group $\mathbb{Z}_2$ generated by σ. Elements of $A_*[\sigma]$ have the form $a_0 \cdot 1 + a_1 \cdot \sigma$, with the multiplication rule that moving σ to the right of $a \in A_*$ produces $\sigma(a) \cdot \sigma$:

$$(48) \qquad\qquad b \cdot \sigma \cdot a = b\,\sigma(a) \cdot \sigma, \ \ a, b \in A_*, \ \ \sigma^2 = 1.$$

This gives an R_*-linear action of the cross product $A_*[\sigma]$ on A_*, that is, a homomorphism from the cross-product to the endomorphism algebra

$$(49) \qquad\qquad A_*[\sigma] \longrightarrow \mathrm{End}_{R_*}(A_*) \cong \mathrm{Mat}(2, R_*),$$

the latter isomorphic to the algebra of 2×2 matrices with coefficients in R_*, where we picked the basis $\{1, X\}$ of A_* as a free R_*-module.

Algebra $A_*[\sigma]$ is four-dimensional over R_*, with a basis $\{1, \sigma, X, X\sigma\}$. Its basis elements act on $A_* = R_* \cdot 1 \oplus R_* \cdot X$ as the following 2×2 matrices.

$$(50) \qquad 1 \longmapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma \longmapsto \begin{pmatrix} 1 & E_1 \\ 0 & -1 \end{pmatrix}, \ X \longmapsto \begin{pmatrix} 0 & -E_2 \\ 1 & E_1 \end{pmatrix}, \ X\sigma \longmapsto \begin{pmatrix} 0 & E_2 \\ 1 & 0 \end{pmatrix} .$$

In the definition of a Galois extension of a commutative ring below we specialize to the case of a cyclic group of order two.

Definition 3.1 ([2, 12],[14, Theorem 12.2.9]). Suppose given an extension of commutative rings $S \subset T$ with an involution τ acting S-linearly on T with S the fixed subring and T a projective S-module. (S, T) is called a *Galois extension* of commutative rings with the cyclic Galois group $C_2 = \{1, \tau\}$ of order two if the natural S-algebra homomorphism $T[\tau] \longrightarrow \mathrm{End}_S(T)$ is an isomorphism. Here $T[\tau]$ is the cross product of T with the group ring of the order two group $\{1, \tau\}$.

Proposition 3.2. (R_*, A_*) *is a Galois extension with the cyclic group* $\{1, \sigma\}$ *if and only if the discriminant* $\mathcal{D} = E_1^2 - 4E_2$ *is invertible in* R_*.

Proof. The question is whether the four matrices in (50) are a basis of the free R_*-module $\mathrm{Mat}(2, R_*)$. Writing down these matrices in the column form, we get the matrix

$$U = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & E_1 & -E_2 & E_2 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & E_1 & 0 \end{pmatrix} .$$

A square matrix with coefficients in a commutative ring S is invertible if and only if its determinant is invertible in S. The determinant of U is $-\mathcal{D}$, thus invertible iff $\mathcal{D}$ is. $\qquad\square$

Among our four rings R_*, the discriminant is invertible in $R_\mathcal{D}$ and $R_{\alpha\mathcal{D}}$.

Corollary 3.3. *Extensions* $(R_\mathcal{D}, A_\mathcal{D})$ *and* $(R_{\alpha\mathcal{D}}, A_{\alpha\mathcal{D}})$ *are Galois with the involution* σ. *Extensions* (R, A) *and* (R_α, A_α) *are not Galois with this involution.*

Field extensions of degree two: Let F be a field. A homomorphism $\psi : R \longrightarrow F$ is determined by $a_1 = \psi(E_1), a_2 = \psi(E_2)$ in F. To ψ we can associate the base change ring $A_\psi = A \otimes_R F$, which is an F-vector space with basis $\{1, X\}$ and multiplication $X^2 = a_1 X - a_2$. The pair (F, A_ψ) is a field extension of degree two iff the polynomial $f(y) = y^2 - a_1 y + a_2$ is irreducible in F, that is, does not have a root.

Assume that field F has characteristic other than two. Then for the F-algebra A_ψ there are three possibilities:

(1) A_ψ is a field. This happens exactly when $\psi(\mathcal{D}) = a_1^2 - 4a_2$ is not a square in F, and then $A_\psi \cong F[\sqrt{\psi(\mathcal{D})}]$.

(2) $A_\psi = Fe_1 \times Fe_2$ is the product of two copies of F. This happens when $f(y)$ has two distinct roots y_1, y_2 in F, with the idempotents

(51) $$ e_1 = \frac{X - y_1}{y_2 - y_1} \ \text{ and } \ e_2 = \frac{X - y_2}{y_1 - y_2}, \ \ 1 = e_1 + e_2, \ e_i e_j = \delta_{i,j} e_i, \ i, j \in \{1, 2\}. $$

Equivalently, $\psi(\mathcal{D}) \neq 0$ has a square root in F.

(3) $A_\psi \cong F[z]/(z^2)$ is a nilpotent extension of F by a nilpotent element of order 2. Equivalently, $\psi(\mathcal{D}) = 0$, that is, $a_1^2 - 4a_2 = 0$, and we can take $z = X - \frac{a_1}{2}$ for this generator.

Cases (1) and (2) happen when the homomorphism $\psi : R \longrightarrow F$ extends to localization $\psi : R_\mathcal{D} \longrightarrow F$.

Assume that $\mathrm{char}(F) = 2$. The image of the discriminant $\psi(\mathcal{D}) = a_1^2$ is a square. For the F-algebra A_ψ there are the following possibilities:

(1) A_ψ is a field, that is, $y^2 + a_1 y + a_2$ has no roots in F. There are two cases.
  (a) $a_1 \neq 0$. Replacing y by $a_1 z$ reduces the equation to $z^2 + z + c = 0$, $c = a_2 a_1^{-2}$, with no solutions in F (which requires c not to be in the image of the map $F \to F, z \mapsto z^2 + z$). Extension (F, A_ψ) is separable.
  (b) $a_1 = 0$ and a_2 is not a square in F. Extension (F, A_ψ) is inseparable.

(2) $A_\psi = Fe_1 \times Fe_2$ is the product of two copies of F. This happens when $f(y) = y^2 + a_1 y + a_2$ has two distinct roots y_1, y_2 in F. The idempotents for the direct product decomposition are

(52) $$ e_1 = \frac{X - y_1}{y_2 - y_1}, \ \text{ and } \ e_2 = \frac{X - y_2}{y_1 - y_2}. $$

(3) $A_\psi \cong F[z]/(z^2)$ is a nilpotent extension of F by a nilpotent element of order 2. For this we need $a_1 = 0$ and $a_2 = a^2$ to have a square root in F. One can then take $z = X + a$ for the above isomorphism.

Among these four cases in characteristic two, (1a) and (2) correspond to $\psi(\mathcal{D}) \neq 0$, giving a separable field extension and a direct product decomposition, respectively. When $\psi(\mathcal{D}) = 0$, we are either in (1b) or (3), that is, A_ψ is an inseparable field extension or a nilpotent extension of F.

Looking across all characteristics, the case $\psi(\mathcal{D}) = 0$ corresponds to a nilpotent extension (3) or an inseparable extension (1b).

Remark: As is well-known, separability property in Definition 3.1 for extensions $S \subset T$ with the Galois group $\mathbb{Z}/2 \cong \{1, \tau\}$ is equivalent to the following: $S = T^\tau$ and there exist $x_1, x_2, y_1, y_2 \in T$ such that

$$(53) \qquad x_1 g(y_1) + x_2 g(y_2) = \left\{ \begin{array}{ll} 1 & \text{if } g = 1 \\ 0 & \text{if } g = \tau \end{array} \right.$$

It is similar but not the same as our neck-cutting relations for (R, A), see equations (12) and (13). For comparisons and analogies between Frobenius and separable extensions we refer to [17, 9] and references therein.

3.2. **Involution σ and defect lines.** For a TQFT interpretation, it's convenient to see σ as an involution of the R-module A, not of A as an algebra. Denote by $\sigma_+ = \sigma$ this R-module map and by σ_- the R-module map $-\sigma$. We have

$$\sigma_+(1) = 1, \quad \sigma_+(X) = E_1 - X, \qquad \sigma_-(1) = -1, \quad \sigma_-(X) = X - E_1.$$

Also, $\sigma_\pm(a) = \pm(\epsilon(Xa) - X\epsilon(a))$, for $a \in A$. Pictorially, this reads

We introduce a diagrammatic notation for $\sigma_\pm$ as a defect line on an annulus with a choice of co-orientation, that is, a preferred side:

Co-orientation at this defect line can be reversed at the cost of adding a minus sign.

One can also call these defect or seam lines σ-*defect lines*. Notice that the diagrams for σ_+ and σ_- are not diffeomorphic rel boundary. One can be taken to the other by reflection in a

horizontal plane, but that map is not the identity on the boundary. Sliding a dot through a σ-defect line converts it to the dual dot. Namely, as endomorphisms of A,

$$\sigma_+ \, X_1 = X_2 \, \sigma_+, \quad \sigma_- \, X_1 = X_2 \, \sigma_-,$$

and likewise with indices $1, 2$ transposed. Diagrammatically,

Sliding a star dot through a σ-defect line flips its co-orientation. As endomorphisms of A,

$$\sigma_+ \, X_\star = X_\star \, \sigma_-, \quad \sigma_- \, X_\star = X_\star \, \sigma_+.$$

Alternatively, one can add a minus sign as star dot is slid:

(54)

The involutory property of $\sigma_\pm$ says that two parallel defect circles, with co-orientation pointing in the same direction, can be removed. Equivalently, if two parallel defect circles both point either into or out of the annulus region separating them, they can be removed with a minus sign, see equation (58) below.

Although σ_+ is an algebra involution, while σ_- is a coalgebra involution, so that

(55) $$\sigma_\pm(1) = \pm 1, \quad \sigma_\pm \circ m \circ (\sigma_\pm \otimes \sigma_\pm) = \pm m,$$

(56) $$\epsilon \circ \sigma_\pm = \mp \epsilon, \quad (\sigma_\pm \otimes \sigma_\pm) \circ \Delta \circ \sigma_\pm = \mp \Delta,$$

it's more natural to think of σ_+ and σ_- as R-module isomorphisms only. Relations between $\sigma_\pm$, multiplication and comultiplication can be rewritten as the equations for the removal of three defect circles around the three holes of a thrice-punctured 2-sphere, see equation (59). Likewise, the relations on $\sigma_\pm$ and the unit and the counit maps, see equations (55) and (56) left, are the removal relations for a defect circle that bounds a disk, see equation (57).

(57)

(58)

(59)

Together with the involutory property, this allows to remove a collection of n distinct defect circles bounding an n-punctured 2-sphere facet without dots, perhaps with a sign, depending on co-orientations. When all circles are co-oriented into a dotless facet, the sign is (-1). When all circles are co-oriented out of a dotless facet, the sign is $(-1)^{n+1}$. In general, if k circles are co-oriented out of the facet, the sign is $(-1)^{k+1}$. In the special case $n = 0$ there is still consistency since the facet is a dotless 2-sphere, evaluating to zero. If a region bounding an n-punctured sphere contains dots, they can be reduced to a linear combination of the same diagrams with at most one dot in the region. A dot can be flipped over a defect line to a dual dot on the other side, see equation (54) above.

To prove that a cobordism equipped with σ-defect lines gives a well-defined map between tensor powers of A (which is implicitly assumed in this discussion), it's convenient to use evaluation of cobordisms with σ-defect lines. We do this a little later, in Section 4.1.
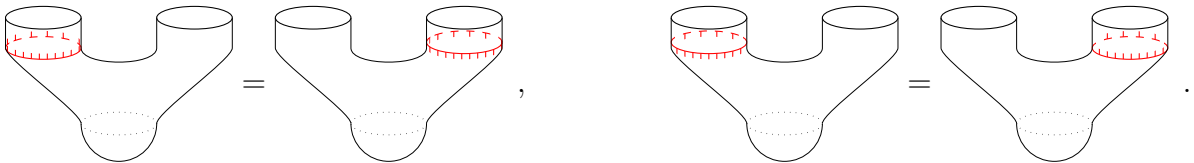
Maps $\sigma_\pm : A \longrightarrow A$ also decompose via the dual dot as

(60) $$\sigma_\pm(a) \ = \pm \left(\epsilon(X_2\, a) - \ X_2\, \epsilon(a) \right).$$

Furthermore, we can interpret the equation

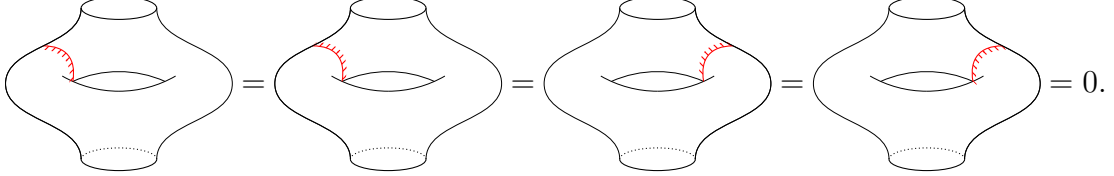(61) $$(\sigma_\pm \otimes 1)\Delta(1) = (1 \otimes \sigma_\mp)\Delta(1)$$

via the movement of a defect line from the left to the right side of a tube. The tube has both boundary circles at the top, and the co-orientation goes from pointing up to pointing down as the defect line is isotoped, as illustrated below.

We have

(62) $$m \circ (\sigma_{\pm} \otimes 1) \circ \Delta = 0 = m \circ (1 \otimes \sigma_{\pm}) \circ \Delta.$$

Pictorially, this reads:

$$= 0.$$

In these cobordisms, the defect circle bounds the same region on both sides. We'll see in the next section that any cobordism with this property defines the zero map. More generally, if a cobordism admits a circle that intersects defect lines odd number of times, the map associated to such cobordism is zero, see Section 4.1.

4. Two evaluations of seamed surfaces

In this section we continue to use variables α_1, α_2 generating the ring R_α as above, with symmetric functions in α_1, α_2 giving us the subring $R = \mathbb{Z}[E_1, E_2]$. To connect the evaluation formulas below with those in [31, 25] one should replace our α_1, α_2 by X_1, X_2 as in [31, 25] and replace our X_1, X_2, heavily used in Section 2, by a different notation, for example Y_1, Y_2.

4.1. **An evaluation over R with defect lines.**

Definition 4.1. A *closed seamed surface* is a closed compact surface F equipped with a PL embedding in $\mathbb{R}^3$ and with finitely many disjoint simple closed curves on F and dots. Each of these curves comes with a co-orientation, that is, a preferred side. Such a curve is also called a *seam* or a *defect*. The set of seams of F is denoted $\Theta(F)$.

The set $\Theta(F)$ of seams may be empty. Denote by $f(F)$ the set of connected components of F with all seams removed. Its elements are called *facets*, or, more precisely, *open facets*. We consider both open facets and their closures in F. The closure of a facet may contain one or more seams. For simplicity, by a facet we usually mean a closed facet.

Dots on F, if any, are placed away from the seams. A dot may float freely in its facet but cannot cross a seam.

Preferred side at a seam may be indicated by a short interval or an arrow pointing from a point on the seam into the corresponding side. Alternatively, we can draw a whole comb of spaced out intervals from the seam and into that side.

Using the identities for neck-cutting (13), dot reduction (15), removal of contractible seam (57), and evaluation of a seamless sphere with at most one dot (10), one can define an *algorithmic evaluation* of closed seamed surfaces as follows:

- Perform neck-cutting on both sides of each seams;
- Use even more neck-cutting to remove handles, we end up with a formal linear combination of disjoint unions of spheres with seams and dots;
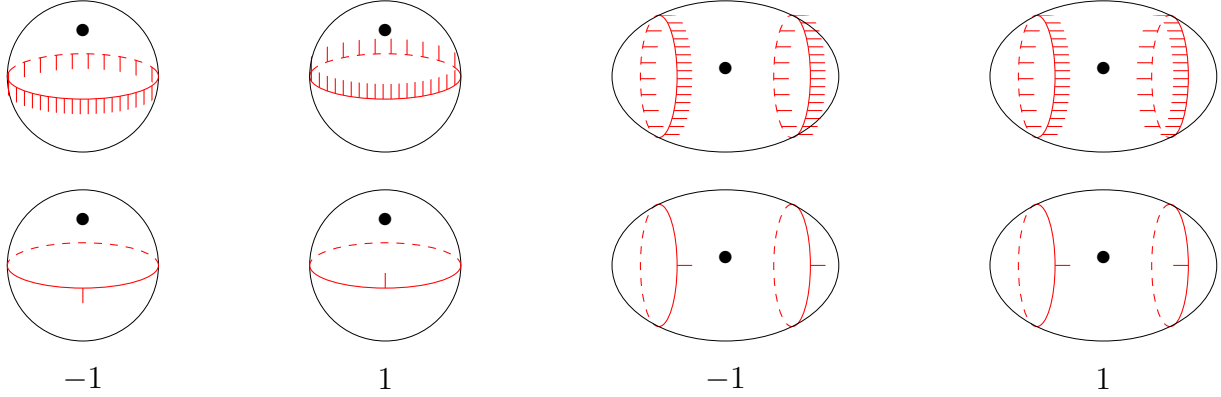
FIGURE 4. Four examples of seamed and dotted spheres represented with both comb and segment notation. Their $\langle\cdot\rangle$-evaluation is given at the bottom. See example 4.4 for details in evaluating the fourth diagram.

- Move the dots and remove seams using (15) and (57) in order to have a formal linear combination of disjoint unions of sphere with at most one dot;
- Finally, these sphere with (10).

That is a map from the set of closed seamed surfaces to the ring R. One would need to check that this is well-defined, which follows from a straightforward computation.

A feature of such reductions is that they may not be implementable while keeping all surfaces in the intermediate steps embedded in $\mathbb{R}^3$. One example is to take a torus embedded in $\mathbb{R}^3$ as the boundary of a knotted solid torus and add several seamed circles in parallel along the longitude of the torus. The first step in the combinatorial evaluation is to do surgeries along annuli separating seamed circles on the torus. This step cannot be done in a natural local way inside $\mathbb{R}^3$ and requires forgetting the embedding but, at least in some similar cases, making sure that all intermediate surfaces are orientable. This is not a serious obstacle, though.

Instead, to show that the map is well-defined, we now exhibit a closed formula for this combinatorial evaluation.

Two facets f_1, f_2 are called *adjacent* if they share at least one seam. A facet is called *self-adjacent* if it comes to some seam from both sides.

We call a seamed surface *even* iff the union of its seams represents zero in $\mathrm{H}_1(F, \mathbb{Z}/2)$. Otherwise a surface is called *odd* (and will evaluate to zero). A seamed surface with a self-adjacent facet is necessarily odd. We can refine this terminology and call each connected component of F *even* or *odd* depending on this property of its seams. Note that a component F_k is *even* if each circle in generic position to the seamed circles in F_k has even number of intersections with the union of seamed circles.

Definition 4.2. A *checkerboard coloring* of a closed seamed surface F is a map $c : f(F) \to \{1, 2\}$ from the set of facets of F to the two-element set $\{1, 2\}$ such that, along every seam, its two sides have different colors. The set of checkerboard colorings of F is denoted $\mathrm{adm}(F)$.

A coloring c of the facets induces a coloring of the seams, by assigning to a seam the color of the facet into which its co-orientation points. For a seamed circle γ denote its induced coloring by $c(\gamma) \in \{1, 2\}$.

Note that F with a self-adjacent facet has no checkerboard coloring, $\mathrm{adm}(F) = \varnothing$. More generally, F admits a checkerboard coloring iff it is an even seamed surface.

Denote by $|F|$ the number of connected components of F. An even surface admits $2^{|F|}$ checkerboard colorings. Denote by $\theta(F) = |\Theta(F)|$ the number of seams of F.

For $i = 1, 2$ denote by $F_i(c)$ the union of facets of F colored by i. Boundary circles of facets are included in both $F_1(c), F_2(c)$. Denote by $\Theta_i(c)$ the set of seams of F colored by i and by $\theta_i(c)$ its cardinality, $\theta_i(c) = |\Theta_i(c)|$. We have $\Theta(c) = \Theta_1(c) \sqcup \Theta_2(c)$ and

$$\theta(F) = \theta_1(c) + \theta_2(c)$$

is the number of singular circles of F, which does not depend on the coloring.

Denote by $d_i(c)$ the number of dots on facets of color i for a coloring c and by $d(F)$ the number of dots on F. Necessarily, $d(F) = d_1(c) + d_2(c)$ for any coloring c. Define the degree

$$(63) \qquad\qquad \deg(F) = -\chi(F) + 2d(F).$$

Since $F_i(c) \subset F \subset \mathbb{R}^3$, and F is a closed surface in $\mathbb{R}^3$, both F and $F_i(c)$ are orientable. Hence, the Euler characteristic of $F_i(c)$ has the same parity as the number of connected component of its boundary. Define

$$(64) \qquad s(F, c) = \theta_1(c) + \frac{\chi(F_1(c)) + \theta(F)}{2} = \theta_1(c) + \frac{\chi(\overline{F}_1(c))}{2} \in \mathbb{Z},$$

where $\chi(\overline{F}_1(c)) = \chi(F_1(c)) + \theta(F)$ is the Euler characteristic of the closed surface $\overline{F}_1(c)$ given by attaching 2-disks to $F_1(c)$ along all boundary circles.

Recall the rings $R = \mathbb{Z}[E_1, E_2]$, $R_\alpha = \mathbb{Z}[\alpha_1, \alpha_2]$, and $R_{\alpha\mathcal{D}} = R_\alpha[(\alpha_1 - \alpha_2)^{-1}]$, the localization of R_α at $\alpha_1 - \alpha_2$. There are ring inclusions

$$R \subset R_\alpha \subset R_{\alpha\mathcal{D}}.$$

Finally, the *c-evaluation* $\langle F, c \rangle$ (or the evaluation of F at c) and the *evaluation* $\langle F \rangle$ are given by the following formulas:

$$(65) \qquad\qquad \langle F, c \rangle = (-1)^{s(F,c)} \frac{\alpha_1^{d_1(c)} \alpha_2^{d_2(c)}}{(\alpha_2 - \alpha_1)^{\chi(F)/2}} \in R_{\alpha\mathcal{D}},$$

$$(66) \qquad\qquad \langle F \rangle = \sum_{c \in \mathrm{adm}(F)} \langle F, c \rangle.$$

The denominator term does not depend on a choice of coloring.

The idea to extend the Robert–Wagner foam evaluation [31] to the case when $F_i(c)$ are not closed surfaces by capping their boundary circles with disks and taking the Euler characteristic of the resulting surfaces $\overline{F}_i(c)$ was proposed by Yakov Kononov [26], who also pointed out that such closure constructions are used implicitly in the physics TQFT literature. The alternative is to use the Euler characteristic of $F_1(c)$, which may be an odd integer. This requires adding $\sqrt{-1}$ to the ground ring, see Section 4.2. A deformation of evaluation (65) appears in [24], with much larger state spaces associated to collections of marked circles in the plane.

Remark: Equation (65) has $\alpha_2 - \alpha_1$ in the denominator, compared to $X_i - X_j$ for $i < j$ in [31]. This is done to make the 2-sphere with one dot evaluate to 1 rather than -1. The same answer can be achieved by changing to $\alpha_1 - \alpha_2$ in the denominator of (65) and to $\overline{F}_2(c)$ in place of $\overline{F}_1(c)$ in formula (64), so that

$$(67) \qquad \langle F, c \rangle = (-1)^{\theta_1(c) + \chi(\overline{F}_2(c))/2} \frac{\alpha_1^{d_1(c)} \alpha_2^{d_2(c)}}{(\alpha_1 - \alpha_2)^{\chi(F)/2}}.$$

If F and G are closed seamed surfaces, it follows directly from the definitions that

$$(68) \qquad \langle F \sqcup G \rangle = \langle F \rangle \cdot \langle G \rangle.$$

Lemma 4.3 (Compare with [31, Proposition 2.18] and [25, Theorem 2.17]). *$\langle F \rangle$ is an homogeneous symmetric polynomial in α_1 and α_2 of degree $\deg(F)$, for any seamed surface F.*

Proof. Let us show that $\langle F \rangle$ is a polynomial. Because of (68), we can assume that F is connected. Since F is embedded in $\mathbb{R}^3$, it is orientable. In particular, $\chi(F) \le 0$ unless F is a sphere, and this is the only case for which the statement is non-trivial. Since $H_1(\mathbb{S}^2, \mathbb{Z}_2)$ is trivial, F is even regardless of its seams and admits exactly two checkerboard colorings. Denote them c and c'. One has $d_i(c') = d_{3-i}(c)$, $\theta_i(c') = \theta_{3-i}(c)$ and $F_i(c') = F_{3-i}(c)$ for $i = 1, 2$. Hence,

$$\begin{aligned}
\langle F \rangle &= \frac{(-1)^{s(F,c)}}{\alpha_2 - \alpha_1} \left(\alpha_1^{d_1(c)} \alpha_2^{d_2(c)} + (-1)^{s(F,c') + s(F,c)} \alpha_1^{d_1(c')} \alpha_2^{d_2(c')} \right) \\
&= \frac{(-1)^{s(F,c)}}{\alpha_2 - \alpha_1} \left(\alpha_1^{d_1(c)} \alpha_2^{d_2(c)} + (-1)^{\theta_1(c) + \theta_2(c) + \frac{\chi(F_1(c) + \chi(F_2(c)))}{2} + \theta(F)} \alpha_1^{d_2(c)} \alpha_2^{d_1(c)} \right) \\
&= \frac{(-1)^{s(F,c)}}{\alpha_2 - \alpha_1} \left(\alpha_1^{d_1(c)} \alpha_2^{d_2(c)} + (-1)^{\frac{\chi(F)}{2}} \alpha_1^{d_2(c)} \alpha_2^{d_1(c)} \right) \\
&= (-1)^{s(F,c)} \frac{\alpha_1^{d_1(c)} \alpha_2^{d_2(c)} - \alpha_1^{d_2(c)} \alpha_2^{d_1(c)}}{\alpha_2 - \alpha_1},
\end{aligned}$$

and $\langle F \rangle$ is a symmetric polynomial in α_1 and α_2. The statement about the degree follows directly from the definition of degree of a seamed surfaces and of the c-evaluation.

That $\langle F \rangle$ is symmetric follows directly from the identity $\tau(\langle F, c \rangle) = \langle F, \tau(c) \rangle$ for the transposition $\tau = (1, 2)$ acting on both the set of colorings and on the ring of rational functions in α_1, α_2. $\qquad \square$

Thus, $\langle F \rangle \in R$ for any seamed surface F.

Remark: If F has no checkerboard colorings (F has an odd component), then $\langle F \rangle = 0$.

Example 4.4. In the following computation, colorings are depicted directly on seamed surfaces: facets with color 1 are hashed and facets with color 2 are plain white.

$$\left\langle \text{⬭} \right\rangle = \left\langle \text{⬭} \right\rangle + \left\langle \text{⬭} \right\rangle$$

$$= (-1)^{2+\frac{0+2}{2}} \frac{\alpha_1}{(\alpha_2 - \alpha_1)^{\frac{2}{2}}} + (-1)^{0+\frac{2+2}{2}} \frac{\alpha_2}{(\alpha_2 - \alpha_1)^{\frac{2}{2}}}$$

$$= \frac{-\alpha_1}{\alpha_2 - \alpha_1} + \frac{\alpha_2}{\alpha_2 - \alpha_1} = 1,$$

using formulas (64)-(66). In more details, $\chi(F) = \chi(\mathbb{S}^2) = 2$, so the denominator is $\alpha_2 - \alpha_1$ in both terms. For the first coloring c_1 (the annulus facet has color 1),

$$\theta_1(c_1) = 2, \ \chi(\overline{F}_1(c_1)) = \chi(\mathbb{S}^2) = 2, \ s(F, c_1) = 2 + 1 = 3, \ d_1(c_1) = 1, \ d_2(c_1) = 0.$$

For the second coloring c_2 (the annulus facet has color 2),

$$\theta_1(c_2) = 0, \ \chi(\overline{F}_1(c_2)) = \chi(\mathbb{S}^2 \sqcup \mathbb{S}^2) = 4, \ s(F, c_2) = 0 + 2 = 2, \ d_1(c_2) = 0, \ d_2(c_2) = 1,$$

so the contributions of $\langle F, c_1 \rangle$ and $\langle F, c_2 \rangle$ to the sum are as above, and $\langle F \rangle = 1$.

Lemma 4.5. *Evaluation of seamed surfaces satisfies the following local relations:*

(69)
$$\left\langle \text{⬭} \right\rangle = \left\langle \text{⬭} \right\rangle + \left\langle \text{⬭} \right\rangle - E_1 \left\langle \text{⬭} \right\rangle,$$

(70)
$$\left\langle \text{⬭} \right\rangle = \left\langle \text{⬭} \right\rangle - \left\langle \text{⬭} \right\rangle,$$

(71)
$$\left\langle \text{⬭} \right\rangle + \left\langle \text{⬭} \right\rangle = E_1 \left\langle \text{⬭} \right\rangle,$$

(72)
$$\left\langle \text{⬭} \right\rangle = E_2 \left\langle \text{⬭} \right\rangle,$$

(73)
$$\left\langle \text{⬭} \right\rangle = E_1 \left\langle \text{⬭} \right\rangle - E_2 \left\langle \text{⬭} \right\rangle$$

(74)
$$\left\langle \text{⬭} \right\rangle = - \left\langle \text{⬭} \right\rangle = - \left\langle \text{⬭} \right\rangle,$$

(75)
$$\left\langle \ominus \right\rangle = 0, \qquad \left\langle \underset{\bullet}{\ominus} \right\rangle = 1.$$

Proof. The only identities which is not straightforward are (69) and (70). The proof of these identities are similar to that of [23, Proposition 4.7] and [25, Proposition 2.22]. We only prove (69), the other one is analogous.

Let us denote by F the seam surface on the left-hand side of the identity and by G_t, G_b and G the seamed surfaces on the right-hand side, where t and b stand for dot on **t**op and dot on **b**ottom. As surfaces, G_t and G_b and G_t are identical, they only differ by their dot distributions. There is a canonical one-to-one correspondence between $\mathrm{adm}(G)$, $\mathrm{adm}(G_b)$ and $\mathrm{adm}(G_t)$. Let c be an element of $\mathrm{adm}(G)$. There are 4 possible local models for c. They are given in Figure 5.

$$\quad\text{(a)}\qquad\qquad\text{(b)}\qquad\qquad\text{(c)}\qquad\qquad\text{(d)}$$
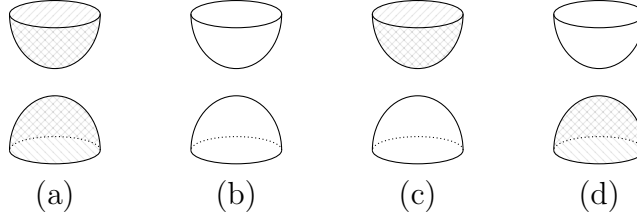
FIGURE 5. The four possible local models for an coloring c of G. Hash means facet has color 1 while solid white means facet has color 2.

In cases (c) (resp. (d)), one has

$$\langle G_t, c \rangle = \alpha_1 \langle G, c \rangle \quad \text{and} \quad \langle G_b, c \rangle = \alpha_2 \langle G, c \rangle$$
$$(\text{resp. } \langle G_t, c \rangle = \alpha_1 \langle G, c \rangle \quad \text{and} \quad \langle G_b, c \rangle = \alpha_2 \langle G, c \rangle)$$

therefore, in both of these cases, $\langle G_t, c \rangle + \langle G_b, c \rangle - E_1 \langle G, c \rangle = 0$. The admissible colorings of G of types (a) and (b) are in a one-to-one correspondence with admissible colorings of F. Let c be a coloring of type (a) of G, with the corresponding coloring of F still denoted by c. On the one hand, one has:

$$\theta_1(G, c) = \theta_1(F, c), \quad \chi(G) = \chi(F) + 2, \quad \chi(G_1(c)) = \chi(F_1(c)) + 2, \quad \theta(G) = \theta(F),$$

On the other hand, $\langle G_t, c \rangle = \langle G_b, c \rangle = \alpha_1 \langle G, c \rangle$. Hence

$$\langle G_t, c \rangle + \langle G_b, c \rangle - E_1 \langle G, c \rangle = (\alpha_1 - \alpha_2) \langle G, c \rangle = -\frac{\alpha_1 - \alpha_2}{\alpha_2 - \alpha_1} \langle F, c \rangle = \langle F, c \rangle.$$

If the coloring c is of type (d), a similar computation gives as well $\langle F, c \rangle = \langle G_t, c \rangle + \langle G_b, c \rangle - E_1 \langle G, c \rangle$. Summing over all admissible colorings, one obtains $\langle F \rangle = \langle G_t \rangle + \langle G_t \rangle - E_1 \langle G \rangle$. $\square$

Formulas (69), (73), and (75) coincide with the corresponding formulas for the 2-dimensional TQFT assigned to the Frobenius pair (R, A), see Section 2.1.

As in Section 2.1, one can introduce the hollow and star dots, which will satisfy the same relations as earlier.

We'll see in Section 4.4 that the state spaces associated to collections of planar circles in this theory coincide with those for the (R, A) TQFT. Seams correspond to σ-defect lines in Section 3.2. Allowing them to end on the boundary enriches (R, A)-TQFT, see examples at the end of Section 4.5.

4.2. **An evaluation over R_ω with defect lines.** Instead of completing the surface $F_1(c)$ to a closed surface $\overline{F}_1(c)$, which has Euler characteristic divisible by two, one can keep $\chi(F_1(c))$ in the evaluation formulas (64), (65) at the cost of adding $\sqrt{-1}$ to the ground ring R to make sense of $(-1)^{\chi(F_1(c))/2}$.

Denote by $\mathbb{Z}_\omega = \mathbb{Z}[\omega]/(\omega^2 + 1)$ the ring of Gaussian integers and let

$$(76) \qquad R_\omega = \mathbb{Z}_\omega \otimes_{\mathbb{Z}} R = \mathbb{Z}_\omega[E_1, E_2]$$

be the ring obtained from R by formally adding a square root ω of -1.

Furthermore, denote $A_\omega = R_\omega \otimes_R A$. The pair (R_ω, A_ω) is a Frobenius extension of rank two, with $\{1, X\}$ a basis of A_ω as a free graded R_ω-module.

We give an alternative evaluation of seamed surfaces in this new algebraic context, using the ring R_ω, together with the rings

$$R_{\omega\alpha} = \mathbb{Z}_\omega \otimes_{\mathbb{Z}} R_\alpha = \mathbb{Z}[\alpha_1, \alpha_2, \omega]/(\omega^2 + 1)$$

and

$$R_{\omega\alpha\mathcal{D}} = R_{\alpha\mathcal{D}}[\omega]/(\omega^2 + 1) = \mathbb{Z}_\omega[\alpha_1, \alpha_2, (\alpha_1 - \alpha_2)^{-1}].$$

There are ring inclusions

$$R_\omega \subset R_{\omega\alpha} \subset R_{\omega\alpha\mathcal{D}}.$$

Define the evaluation of F at a coloring c and the overall evaluation of F by

$$(77) \qquad \langle F, c \rangle_\omega := \omega^{2\theta_1(c) + \chi(F_1(c))} \frac{\alpha_1^{n_1(c)} \alpha_2^{n_2(c)}}{(\alpha_2 - \alpha_1)^{\chi(F)/2}} = \omega^{-\theta(F)} \langle F, c \rangle,$$

$$(78) \qquad \langle F \rangle_\omega := \sum_{c \in \mathrm{adm}(F)} \langle F, c \rangle_\omega = \omega^{-\theta(F)} \langle F \rangle.$$

Note that $\langle F, c \rangle_\omega \in R_{\omega\alpha\mathcal{D}}$ while $\langle F \rangle_\omega \in R_\omega$ for any seamed surface F. Thus, $\langle F \rangle$ is an homogeneous symmetric polynomial in α_1 and α_2 of degree $\deg(F)$ with coefficients in the ring of Gaussian integers $\mathbb{Z}_\omega = \mathbb{Z}[\omega]/(\omega^2 + 1)$.

Lemma 4.6. *Evaluation of seamed surfaces satisfies the following local relations:*

$$(79)$$

$$(80)$$

$$(81) \qquad \left\langle \begin{array}{c}\bullet\end{array}\right\rangle_\omega + \left\langle \begin{array}{c}\bullet\end{array}\right\rangle_\omega = E_1 \left\langle \begin{array}{c}\end{array}\right\rangle_\omega ,$$

$$(82) \qquad \left\langle \begin{array}{c}\bullet\quad\bullet\end{array}\right\rangle_\omega = E_2 \left\langle \begin{array}{c}\end{array}\right\rangle_\omega ,$$

$$(83) \qquad \left\langle \begin{array}{c}\bullet\bullet\end{array}\right\rangle_\omega = E_1 \left\langle \begin{array}{c}\bullet\end{array}\right\rangle_\omega - E_2 \left\langle \begin{array}{c}\end{array}\right\rangle_\omega ,$$

$$(84) \qquad \left\langle \begin{array}{c}\bigcirc\end{array}\right\rangle_\omega = - \left\langle \begin{array}{c}\bigcirc\end{array}\right\rangle_\omega = -\omega \left\langle \begin{array}{c}\end{array}\right\rangle_\omega ,$$

$$(85) \qquad \left\langle \begin{array}{c}\end{array}\right\rangle_\omega = \omega \left\langle \begin{array}{c}\end{array}\right\rangle_\omega ,$$

$$(86) \qquad \left\langle \begin{array}{c}\end{array}\right\rangle_\omega = 0, \qquad \left\langle \begin{array}{c}\bullet\end{array}\right\rangle_\omega = 1.$$

From relation (80) we see that the seam line is now $\sigma_\pm$ scaled by ω rather than just $\sigma_\pm$ as in equation (70). Multiplicative factor $\omega^{-\theta(F)}$ in the equation (78) tells us how other formulas, including equations (58) and (59), will modify in this evaluation.

4.3. Universal construction.

The *universal construction* constructs functors from a cobordism category to an algebraic category. It was introduced by Blanchet–Habegger–Masbaum–Vogel [8] and used to build foam state spaces in [20, 31, 25]. For this construction, one needs an evaluation of closed objects, such as cobordisms, that is, a map from the isomorphism classes of closed n-manifolds to some commutative ring. In the cases we consider the evaluation is $\langle \cdot \rangle$ or $\langle \cdot \rangle_\omega$.

In favorable situations, functors obtained by the universal construction are TQFTs. However, this is not always the case. In particular these functors can fail to be monoidal.

Suppose Cob is a category of cobordisms, S is a commutative unital ring and

$$\tau \colon \mathrm{End}_{\mathsf{Cob}}(\varnothing) \to S$$

a monoid homomorphism. Here $\mathrm{End}_{\mathsf{Cob}}(\varnothing)$ is the monoid of isomorphism classes of cobordisms with the empty boundary as elements and the disjoint union as the composition. The map τ should take composition of cobordisms to the product of corresponding elements in S. In particular it maps the empty n-cobordism $\varnothing_n = \mathrm{Id}_{\varnothing_{n-1}}$ to 1_S. In these notations we distinguish between the empty n-cobordism $\varnothing_n$ and the empty $(n-1)$-cobordism $\varnothing_{n-1}$. The former is the identity endomorphism of the latter.

Let M be an object of Cob. Define $\widetilde{\mathcal{F}_\tau}(M)$ to be the free S-module generated by $\mathrm{Hom}_{\mathsf{Cob}}(\varnothing, M)$. For any W in $\mathrm{Hom}(M, \varnothing)$, define the S-linear map $\varphi_W \colon \widetilde{\mathcal{F}_\tau}(M) \to S$ on basis elements

$V \in \mathrm{Hom}_{\mathsf{Cob}}(\varnothing, M)$ by $\varphi_W(V) = \tau(W \circ V)$. Finally, let

$$\mathcal{F}_\tau(M) = \widetilde{\mathcal{F}}_\tau(M) \bigg/ \bigcap_{W \in \mathrm{Hom}_{\mathsf{Cob}}(\varnothing, M)} \mathrm{Ker}(\varphi_W)$$

be the quotient module, $\mathcal{F}_\tau(M) \in S\text{-}\mathsf{mod}$. Thus, $\mathcal{F}_\tau(M)$ is the quotient of the free S-module $\widetilde{\mathcal{F}}_\tau(M)$ by the kernel of a suitable bilinear form. For a cobordism W representing a generator of $\widetilde{\mathcal{F}}_\tau(M)$, denote by $[W]$ its equivalence class in $\mathcal{F}_\tau(M)$. One extends $\mathcal{F}$ to a functor by defining for any $W \in \mathrm{Hom}_{\mathsf{Cob}}(M_1, M_2)$ and $V \in \widetilde{\mathcal{F}}_\tau(M_1)$:

$$\mathcal{F}_\tau([W])([V]) = [W \circ V].$$

Bilinear pairing. For every object M in Cob, evaluation τ induces an R-bilinear pairing $(\cdot, \cdot)_\tau$ on $\mathcal{F}_\tau(M)$: If W_1 and W_2 are two elements of $\mathrm{Hom}_{\mathsf{Cob}}(\varnothing, M)$, define

(87) $([W_1], [W_2])_\tau = \tau(\overline{W_1} \circ W_2) = \varphi_{\overline{W_1}}(W_2),$

where $\overline{W_1} \in \mathrm{Hom}_{\mathsf{Cob}}(M, \varnothing)$ is the mirror image of W_1. This pairing is non-degenerate at least on the right. We did not require skew-invariance of τ under the flip (often related to orientation-reversal of closed cobordisms), where ring S would carry a bar involution, with $\overline{\tau(W)} = \tau(\overline{W})$ for closed cobordisms W. This condition would make the bilinear form skew-invariant as well. In our examples, the involution on the ring is the identity and $\tau(W) = \tau(\overline{W})$ for closed cobordisms W, making the bilinear form $(\cdot, \cdot)_\tau$ symmetric.

4.4. Category of seamed surfaces.

Seamed surfaces can be extended to a category whose objects are finite disjoint unions of *marked circles*. Marks (or defects) on circles are endpoints of seams ending on the boundary of a surface.

A *marked circle* is a circle equipped with a PL-embedding in $\mathbb{R}^2$. It carries a finite number (possibly none) of marked points, also called *seam points*. Each marked point carries a co-orientation, that is a preferred direction in a circle at this point. Equivalently, we say that a co-orientation is a preferred side of a circle near a marked point. Marked points and their co-orientations are depicted by red solid arrows on circles. If one chooses an orientation o of a seam circle, co-orientation at each marked point either agrees or disagrees with o. If the orientation agrees, respectively disagrees, with co-orientation, we label the point as $+$ point (plus point), respectively, $-$ point (minus point), relative to this orientation. Reversing the orientation of the circle flips plus and minus points. If there are as many plus as minus points, the circle is called *balanced*. This notion does not depend on the choice of orientation. If a circle has no marked point it is *unmarked*.

An object of the category SeSu is a finite collection C of disjoint marked circles in the plane. We call such C a *marked embedded one-manifold* or *meom*, for short.

Let us orient circles in C so that outermost circles are oriented clockwise. When all outermost circles are removed, the outermost circles in the remaining one-manifold must have anticlockwise orientation, and so on. Iterating this condition, we come to a canonical choice of orientation $o(C)$ for circles in any C.
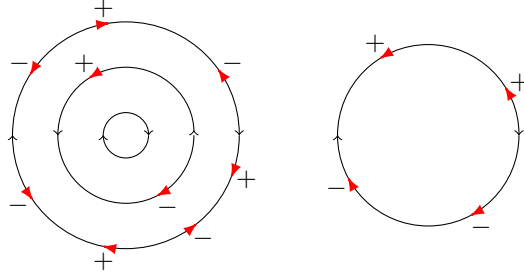
FIGURE 6. Example of a meom. Orientations are depicted by thin black arrows ($\rightarrowtail$) and marks are indicated by triangular red arrows ($\blacktriangleright$). The outer left circle is not balanced, the others are.

Since marked points in C carry co-orientations, each point is either compatibly oriented relative to $o(C)$ (a + point) or oppositely oriented (a − point). On a given circle in C, we can encode the sequence of orientations as a sequence of signs $\underline{\ell} = (\ell_1, \ldots, \ell_k)$, $\ell_i \in \{+, -\}$, up to cyclic order, as we go along the circle following its orientation.

If an object C is a single circle in the plane, necessarily clockwise oriented, with the sequence of signs $\underline{\ell}$, we denote it by $\mathbb{S}^1_{\underline{\ell}}$.

We say that C is *balanced* if it has as many plus signs as minus signs in the collection of sequences for its circles. For instance, if C has three circles, with cyclic sequences $\{(+ + +), (- - +), (--)\}$, it is balanced. Note that C may be balanced without individual circles having this property. Notice also that the assignment of pluses and minuses to marked points depends not only on their co-orientations but also on the parity of the circle in its nesting in C, that is, whether its orientation in the plane is clockwise or anticlockwise.

More generally, to C we associate its *weight* $w(C)$, the difference between the number of pluses and minuses on its circles. A meom C is balanced if and only if $w(C) = 0$.

A circle is called *odd*, respectively *even*, if it has an odd, respectively even, number of seam points. A meom C is called *even* if each circle in it is even.

If C_0 and C_1 are two meoms, a *seamed cobordism from C_0 to C_1* is a compact surface F equipped with a proper PL-embedding in $\mathbb{R}^2 \times [0, 1]$ and with finitely many disjoint simple curves with co-orientations in F (*seams*) and dots such that:

- The embedding of F is transverse to $\mathbb{R}^2 \times \{0, 1\}$.
- The seams of F are properly embedded and transverse to the boundary.
- The boundary of F is equal to $C_0 \times \{0\} \sqcup C_1 \times \{1\}$.
- Marked points of C_0 and C_1 coincide with the intersection of seams of F and the boundary. Their co-orientations agree with the ones induced by seams.

Our orientation conventions for surfaces and their top and bottom boundaries are the following. Orientation of a surface F induces an orientation of its top boundary $\partial_1 F$ by sticking the first vector of an orientation basis out of the surface, see Figure 7. The second vector then shows the direction for the boundary orientation. For the bottom boundary $\partial_0 F$ the convention is the opposite: when the second vector of an orientation basis points out, the first vector shows
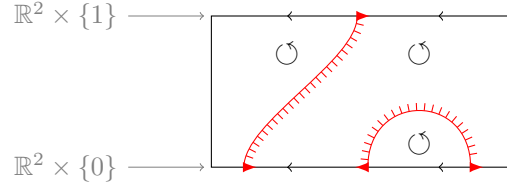
FIGURE 7. Diagrammatic summary of conventions for orientation and co-orientations.

induced orientation of the boundary. We also adopt the convention that

$$\partial F = \partial_1 F \sqcup (-\partial_0 F).$$

In Figure 7 top and bottom horizontal lines indicate parts of circles in $\mathbb{R}^2 \times \{1\}$ and $\mathbb{R}^2 \times \{0\}$, respectively. Marked points, shown as red triangles at these boundary lines, inherit co-orientations from those of seamed arcs. Top marked point is a minus point, the three marked points at the bottom edge have signs $(- + -)$, reading from left to right. The seam with both endpoints on the bottom connects a $+$ and a $-$ endpoints (different signs). An edge betweeen the top and the bottom boundary connects two $-$ enpoints (same sign on both).

Meoms and seamed cobordisms between them, up to rel boundary isotopies in $\mathbb{R}^2 \times [0,1]$, form a category denoted SeSu. Composition is given by superposition and rescaling.

In SeSu a morphism from C_0 to C_1 exists iff $w(C_0) = w(C_1)$, that is, if they have the same weight. Indeed, such a morphism may have several seamed arcs connecting marked points on the same boundary $\partial_i F$, $i = 0, 1$, and connecting marked points of $\partial_0 F$ to the points of $\partial_1 F$. Seamed arcs of the first type connect a plus point and a minus point, each contributing zero to $w(C_i)$. Seamed arcs of the second type connect points of $\partial_0 F$ and $\partial_1 F$ of the same sign, contributing zero to $w(C_0) - w(C_1)$. Consequently, $w(C_0) = w(C_1)$.

Remark: It's important to point out that our setup with orientations of circles and surfaces carries a secondary role, and these orientations can almost be ignored. We are orienting collections C of circles in $\mathbb{R}^2$ so that any surface F embedded in $\mathbb{R}^2 \times [0,1]$ can be compatibly oriented, together with its boundary. These orientations also allow to assign signs to all marked points on the boundary of F. In fact, we need less data. What's useful is the relative index of two marked points on ∂F carrying co-orientations to see if they can be boundary points of a single seam. Marked points of ∂F, with co-orientations, decompose as a disjoint union of two sets. For any two points p_0 and p_1 in the opposite sets it's possible to replace F by a surface F' in $\mathbb{R}^2 \times [0,1]$ with the same boundary and boundary co-orientations as in F, such that p_0 and p_1 are connected by a seam in F'. Equivalently, F can be completed by a surface F'' with $\partial F = \partial F''$ and F'' lying 'outside' of F with $F \cup F''$ a closed surface in $\mathbb{R}^3$, such that p_0 and p_1 are connected by a seam in F'' and, consequently, belong to the same defect circle (seamed circle) in $F \cup F''$. Points p_0 and p_1 from the same subset cannot be connected by a seam in F or in any such replacement F' or complement F''.

Applying the universal construction to $\langle \cdot \rangle$ and $\langle \cdot \rangle_\omega$ and the category SeSu yields functors, denoted $\langle \cdot \rangle$ and $\langle \cdot \rangle_\omega$, respectively. These are functors

$$\langle \cdot \rangle : \mathsf{SeSu} \longrightarrow R-\mathrm{gmod}, \quad \langle \cdot \rangle_\omega : \mathsf{SeSu} \longrightarrow R_\omega-\mathrm{gmod},$$

from SeSu to the category of graded R-modules, respectively graded R_ω-modules, and homogeneous module homomorphisms. The images of objects of SeSu by these functors, that is, $\langle C \rangle$ and $\langle C \rangle_\omega$, are called *state spaces* of C.

Suppose $F \colon C_0 \to C_1$ is a seam surface. It follows from the definition of evaluation of seam surface that the R-module map $\langle F \rangle$ (resp. R_ω-module map $\langle F \rangle_\omega$) is homogenous and its degree is given by the formula:

$$\deg \langle F \rangle = \deg \langle F \rangle_\omega = -\chi(F) + 2d(F).$$

4.5. **State spaces for functors $\langle \cdot \rangle$ and $\langle \cdot \rangle_\omega$.** The state space $\langle \varnothing_1 \rangle \cong R$ of the empty collection of circles is a rank one free R-module generated by $[\varnothing_2]$, where $\varnothing_2$ is the empty surface. If a meom C is unbalanced (that is, $w(C) \neq 0$), there are no cobordisms from the empty meom $\varnothing_1$ into it and $\langle C \rangle = 0$, $\langle C \rangle_\omega = 0$.

Given meoms C and C', there is a natural injective graded R-module homomorphism

$$(88) \qquad\qquad \langle C \rangle \otimes_R \langle C' \rangle \longrightarrow \langle C \sqcup C' \rangle$$

intertwining monoidal structures on the category SeSu and the category of graded R-modules. This functor is not monoidal, though. The above homomorphism is not an isomorphism, in general, since we can take C and C' unbalanced, with zero state spaces, but make $C \sqcup C'$ balanced, with $\langle C \sqcup C' \rangle \neq 0$. A simple example is choosing an unbalanced circle for C, with even number of marked points, and taking $C' = C^!$, the mirror image of C. The natural tube cobordism from $\varnothing_1$ to $C \sqcup C^!$, composed with its reflection, evaluates to ± 2. This implies nontriviality of the state space $C \sqcup C^!$.

Lemma 4.7. *The state space $\langle \mathbb{S}^1 \rangle$ of a single unmarked circle is isomorphic to A as a graded R-module.*

Proof. We first construct a map from A to $\langle \mathbb{S}^1 \rangle$. Recall that A is a free R-module and that $1, X$ is a basis of A. Define $\Phi : A \longrightarrow \langle \mathbb{S}^1 \rangle$ as the R-module map which maps 1 to the class of ⌣ and X to the class of ⌣•.

Let us first show that the map Φ is surjective. By R-linearity, it is enough to show that any element of the form $[F]$ with $F \in \mathrm{Hom}_{\mathsf{SeSu}}(\varnothing, \mathbb{S}^1)$ has a preimage by Φ. Using the neck-cutting relation (69) one obtains that $[F]$ equals an R-linear combination of three surfaces. Two of them are disjoint unions of ⌣ and closed seamed surfaces. One of them is the disjoint union of ⌣• and a closed seamed surface. Evaluate closed seam surfaces, we see that $[F]$ is the an R-linear combination of ⌣ and ⌣•. Hence, Φ is surjective.

Let us now show that Φ is injective. It is enough to show that $\left[\begin{array}{c}\bigcirc\end{array}\right]$ and $\left[\begin{array}{c}\bigcirc\bullet\end{array}\right]$ are linearly independent. We can compute the matrix of the R-bilinear pairing (87) on this set:

$$\begin{pmatrix} \left\langle \begin{array}{c}\bigcirc\end{array} \right\rangle & \left\langle \begin{array}{c}\bigcirc\bullet\end{array} \right\rangle \\ \left\langle \begin{array}{c}\bullet\bigcirc\end{array} \right\rangle & \left\langle \begin{array}{c}\bullet\bigcirc\bullet\end{array} \right\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & E_1 \end{pmatrix}.$$

This matrix is invertible, proving that $\left[\begin{array}{c}\bigcirc\end{array}\right]$ and $\left[\begin{array}{c}\bigcirc\bullet\end{array}\right]$ are linearly independent. $\square$

We remind the reader that in this lemma A is viewed as a module, not as a ring, and its generator 1 lives in degree -1.

One can easily adapt the previous proof to obtain the following proposition.

Proposition 4.8. *Let C be a meom in* SeSu *obtained by inserting an unmarked circle $\mathbb{S}^1$ into one of the regions of a meom C' as an innermost circle. Then $\langle C \rangle \simeq A \otimes_R \langle C' \rangle$.*

Corollary 4.9. *The state space $\left\langle \sqcup_{i=1}^{k} \mathbb{S}^1 \right\rangle$ of a collection of k unmarked circles is isomorphic to $A^{\otimes k}$ as a graded R-module.*

The same isomorphism holds for arbitrarily nested collections of unmarked circles.

Each circle corresponds to a tensor factor. Using relations in Lemma 4.5, maps associated to generating cobordisms between unmarked circles can be identified with the following maps:

$$\left\langle \begin{array}{c} \end{array} \right\rangle = \sigma_+ \colon A \to A, \qquad \left\langle \begin{array}{c} \end{array} \right\rangle = \sigma_- \colon A \to A,$$

$$\left\langle \begin{array}{c} \end{array} \right\rangle = \epsilon \colon A \to R, \qquad \left\langle \begin{array}{c} \end{array} \right\rangle = \begin{array}{ccc} R & \to & A \\ 1_R & \mapsto & 1_A \end{array}, \qquad \left\langle \begin{array}{c}\bullet\end{array} \right\rangle = \begin{array}{ccc} R & \to & A \\ 1_R & \mapsto & X \end{array},$$

$$\left\langle \begin{array}{c} \end{array} \right\rangle = \Delta \colon A \to A \otimes A, \qquad \left\langle \begin{array}{c} \end{array} \right\rangle = m \colon A \to A \otimes A.$$

The maps are those of the (1+1)-dimensional TQFT (R, A) and the σ-defect circles, so the theory 1 of Sections 2 and 3 appears out of evaluation $\langle \cdot \rangle$.

Let us denote by $\mathbb{S}^1_{+-}$ the balanced seamed circle in the plane with two seam points. It is unique up to planar isotopy.

Lemma 4.10. *Suppose that γ is a separating closed curve on a seamed surface F that intersects seams transversely in exactly two points with opposite co-orientations. Then the following*

relation holds:

(89)
$$\left\langle \gamma \rightarrow \vcenter{\hbox{\includegraphics{cylinder}}} \right\rangle \overset{\triangle}{=} \left\langle \vcenter{\hbox{\includegraphics{dot1}}} \right\rangle - \left\langle \vcenter{\hbox{\includegraphics{dot2}}} \right\rangle.$$

Notation △ above the equal sign emphasizes the fact this local *relation is valid provided that a* global *condition (γ being a separating curve) is satisfied.*

Proof. Let us denote by F the seam surface on the left-hand side of the identity and by G_1 and G_2 the seamed surfaces on the right-hand side. As surfaces G_1 and G_2 are identical, they only differ by their dot distributions. Denote by G the seam surface which is identical to G_1 and G_2 but with the visible dot removed. There is a canonical one-to-one correspondence between $\mathrm{adm}(G)$, $\mathrm{adm}(G_1)$ and $\mathrm{adm}(G_2)$. Let c be an element of $\mathrm{adm}(G)$. There are 4 possible local types of c. They are given in Figure 8.
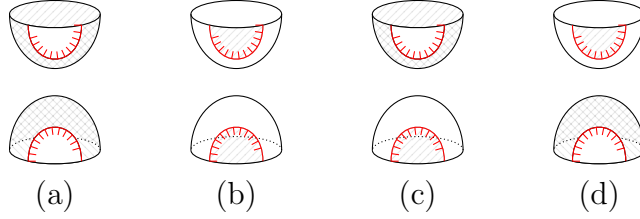

(a)   (b)   (c)   (d)

FIGURE 8. The four possible local types of a coloring c of G. Hash means facet has color 1 while solid white means facet has color 2.

In cases (c) (resp. (d)), one has

$$\langle G_1, c \rangle = \langle G_2, c \rangle = \alpha_2 \langle G, c \rangle$$
$$(\text{resp.} \quad \langle G_1, c \rangle = \langle G_2, c \rangle = \alpha_1 \langle G, c \rangle),$$

therefore, in both of these cases, $\langle G_2, c \rangle - \langle G_1, c \rangle = 0$.

The admissible colorings of G of types (a) and (b) are in a one-to-one correspondence with admissible colorings of F. Let c be a coloring of type (a) of G, with the corresponding coloring of F still denoted by c. On the one hand, one has:

$$\theta_1(G, c) = \theta_1(F, c), \quad \chi(G) = \chi(F) + 2, \quad \chi(G_1(c)) = \chi(F_1(c)) + 1, \quad \theta(G) = \theta(F) + 1.$$

The last identity comes from γ being a separating curve of F and implies

$$\langle F, c \rangle = -\frac{\langle G, c \rangle}{\alpha_2 - \alpha_1}.$$

On the other hand, $\langle G_1, c \rangle = \alpha_1 \langle G, c \rangle$ and $\langle G_2, c \rangle = \alpha_2 \langle G, c \rangle$ so that, finally, $\langle F, c \rangle = \langle G_1, c \rangle - \langle G_2, c \rangle$.

If the coloring c is of type (b), a similar computation gives as well $\langle F, c \rangle = \langle G_1, c \rangle - \langle G_2, c \rangle$. Summing over all admissible colorings, one obtains $\langle F \rangle = \langle G_1 \rangle - \langle G_2 \rangle$. □

Corollary 4.11. *The state space $\left\langle \mathbb{S}^1_{+-} \right\rangle$ is isomorphic to A as a graded R-module.*

Sketch of the proof. Same argument as for Lemma 4.7 shows that $\left\langle \mathbb{S}^1_{+-} \right\rangle$ is generated by and . One shows similarly that these elements are R-linearly independent. $\square$

The behavior of the functor $\langle \cdot \rangle$ on disjoint unions of circles with marked points is not fully monoidal, even if the circles are balanced.

First, consider the seamed surface $\mathbb{S}^1_{+-} \times \mathbb{S}^1$ given by taking meom $\mathbb{S}^1_{+-}$, a balanced circle with two marks in the plane, and multiplying by the circle $\mathbb{S}^1$ to get a standardly embedded torus in $\mathbb{R}^3$ with two seamed circles. The evaluation of this surface

$$\left\langle \mathbb{S}^1_{+-} \times \mathbb{S}^1 \right\rangle = -2 \neq 2 = \mathrm{rk}_R \left\langle \mathbb{S}^1_{+-} \right\rangle$$

is minus two, different from two, which is the rank of $\left\langle \mathbb{S}^1_{+-} \right\rangle$ as a free R-module. Usually, in the TQFT land, multiplying an $(n-1)$-manifold M by a circle and evaluating gives the dimension of the state space associated to the manifold M. In the example above, the evaluation is -2 rather than 2, perhaps implying the need to make the latter a super-module sitting in odd degree and hinting that we cannot expect monoidality property on the nose without further modifications.

Similarly, the evaluation $\left\langle \mathbb{S}^1_{++} \times \mathbb{S}^1 \right\rangle = 2$, but the circle $\mathbb{S}^1_{++}$ is unbalanced, with the trivial state space of zero dimension. Of course, a cross-section $\mathbb{S}^1_{++} \sqcup \mathbb{S}^1_{--}$ of $\mathbb{S}^1_{++} \times \mathbb{S}^1$ is balanced, while individual circles in it are not, with the state space $\left\langle \mathbb{S}^1_{++} \sqcup \mathbb{S}^1_{--} \right\rangle$ nontrivial, showing a failure of monoidality for trivial reasons.

Next, consider the following six elements of $\left\langle \mathbb{S}_{+-} \sqcup \mathbb{S}_{+-} \right\rangle$:
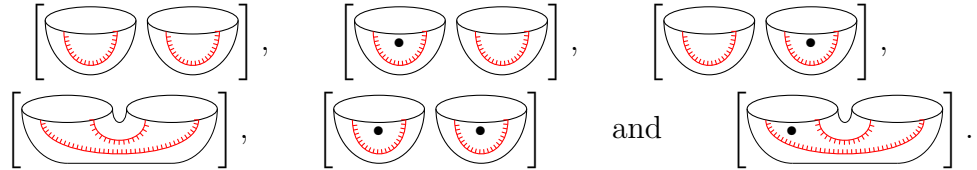
The surfaces are listed in the order of increasing degree, which is $-2, 0, 0, 0, 2, 2$ in this order. The matrix of the R-bilinear pairing (Gram matrix) on these elements is:

(90)
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & E_1 & E_1 \\ 0 & 1 & 0 & 1 & E_1 & E_1 \\ 0 & 1 & 1 & -2 & E_1 & -E_1 \\ 1 & E_1 & E_1 & E_1 & E_1^2 & E_1^2 - E_2 \\ 1 & E_1 & E_1 & -E_1 & E_1^2 - E_2 & -E_1^2 + 2E_2 \end{pmatrix}$$

Its determinant is $4E_1^2 - 16E_2 = 4(E_1^2 - 4E_2) = 4\mathcal{D}$, which is not a zero divisor in R. Consequently, these six elements are R-linearly independent in $\left\langle \mathbb{S}^1_{+-} \sqcup \mathbb{S}^1_{+-} \right\rangle$ and span a graded submodule isomorphic to R^6, of graded rank $(q + q^{-1})^2 + 1 + q^2$.

The natural map of state spaces

$$\left\langle \mathbb{S}^1_{+-} \right\rangle \otimes_R \left\langle \mathbb{S}^1_{+-} \right\rangle \longrightarrow \left\langle \mathbb{S}^1_{+-} \sqcup \mathbb{S}^1_{+-} \right\rangle$$

given by putting two seamed surfaces with boundary $\mathbb{S}^1_{+-}$ next to each other is injective. It takes products of standard basis vectors for $\mathbb{S}^1_{+-}$ to the first, second, third and fifth surfaces, among the six elements above. In particular, this map is not surjective, missing the free R-module generated by the two other surfaces, given by a seamed tube, either with a dot or undecorated.

In particular, the state spaces $\langle \mathbb{S}^1_{+-} \sqcup \mathbb{S}^1_{+-} \rangle$ and $\langle \mathbb{S}^1_{+-} \rangle \otimes \langle \mathbb{S}^1_{+-} \rangle$ are not isomorphic. With more work, one can check that the above six surfaces are a basis of $\langle \mathbb{S}^1_{+-} \sqcup \mathbb{S}^1_{+-} \rangle$.

We see that the functor $\langle \cdot \rangle$, even restricted to meoms with every circle balanced, is not a monoidal functor from the category of seamed surfaces with boundary embedded in $\mathbb{R}^3$ to the category of graded R-modules, strengthening non-monoidality property observed at the beginning of this section.

The matrix (90) becomes unimodular if we invert 2 and $\mathcal{D}$, but we don't explore this here. Note also that the value of the closed genus three surface is $2\mathcal{D}$, see (20).

Finally, let us inspect the state space $\langle \mathbb{S}^1_4 \rangle$, where $\mathbb{S}^1_4 = \mathbb{S}^1_{+-+-}$ is the circle with four seam points with alternating co-orientations, see Figure 9.
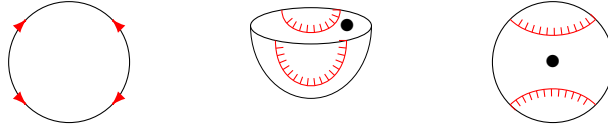

FIGURE 9. $\mathbb{S}^1_4$ is on the left; in the middle is a seamed surface from the empty web $\varnothing$ to $\mathbb{S}^1_4$, and on the right, its flat representation.

One can show using the neck-cutting relation that the following four elements generate $\langle \mathbb{S}^1_4 \rangle$:

The matrix of the R-bilinear pairing on these vectors is:

$$\begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & E_1 & 0 \\ -1 & 1 & 0 & E_1 \end{pmatrix}$$

The determinant of this matrix is 4, which is not a zero divisor. Hence, $\langle \mathbb{S}^1_4 \rangle$ is a free R-module of rank 4 and graded rank $2(q + q^{-1})$. At the same time, $\langle \mathbb{S}^1_{+-} \rangle$ and $\langle \mathbb{S}^1 \rangle$ are free R-modules of rank 2 and graded rank $q + q^{-1}$. We see that rank and graded rank of the module $\langle \mathbb{S}^1_{\underline{\ell}} \rangle$ assigned to a circle with a balanced sequence $\underline{\ell} = (\ell_1, \ldots, \ell_{2k})$ of signs may depend on k. This is a more subtle phenomenon than the observation that $\langle \mathbb{S}^1_{\underline{\ell}} \rangle = 0$ for an unbalanced $\underline{\ell}$ due to absence of seamed embedded surfaces bounding $\mathbb{S}^1_{\underline{\ell}}$.

Even more substantial dependence of $\langle \mathbb{S}^1_{\underline{\ell}} \rangle$ on the choice of a balanced sequence $\underline{\ell}$ is investigated in [24] for a deformation of evaluation $\langle \cdot \rangle$.

Sequence $\underline{\ell}$	rank	graded rank
$\varnothing$	2	$q + q^{-1}$
$+-$	2	$q + q^{-1}$
$++--$	2	$q + q^{-1}$
$+-+-$	4	$2(q + q^{-1})$

TABLE 2. Ranks and graded ranks of the space $\langle \mathbb{S}_{\underline{\ell}}^1 \rangle$ for balanced $\underline{\ell}$ of short length.

The rank of the state space $\langle \mathbb{S}_{\underline{\ell}}^1 \rangle$ depends not only on the length of $\underline{\ell}$ but also on its cyclic ordering. Consider $\mathbb{S}_{4'}^1 := \mathbb{S}_{++--}^1$. One can show that $\langle \mathbb{S}_{4'}^1 \rangle$ is generated by

and
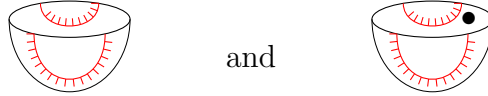
and that these two elements are linearly independent, thus $\langle \mathbb{S}_{4'}^1 \rangle$ is a free R-module of graded rank $q + q^{-1}$. Table 2 collects rank and graded rank information for state spaces $\langle \mathbb{S}_{\underline{\ell}}^1 \rangle$, for shortest balanced sequences $\underline{\ell}$. Cyclic permutation of the sequence does not change the state space.

This somewhat unusual behavior improves in the alternative evaluation $\langle \cdot \rangle_\omega$. Our original reason for considering evaluation $\langle \cdot \rangle_\omega$ was to rederive some of the constructions of Caprau [10, 11] and Clark, Morrison, and Walker [13] via an evaluation framework. We encourage the reader to match the relations in [11, Section 3.2] with those in Lemma 4.6 above. Orienting a seamed line, as done in [11], is equivalent to co-orienting it, as in [13] and the present paper, as long as the ambient surface comes with an orientation.

Using the neck-cutting relation for $\langle \cdot \rangle_\omega$ one obtains an analogue of Proposition 4.8.

Proposition 4.12. *Let C be an object of* SeSu *obtained by inserting an unmarked circle $\mathbb{S}^1$ into one of the regions of a meom C' as an innermost circle. Then $\langle C \rangle_\omega \simeq A \otimes_R \langle C' \rangle_\omega$.*

Lemma 4.13. *Let C be an object of* SeSu *with (at least) one balanced circle S. Let C' be obtained from C by changing marked points on S to obtain another balanced circle S'. Then $\langle C \rangle_\omega \simeq \langle C' \rangle_\omega$.*

Proof. It is enough to consider the case where S' and S differ by a pair of adjacent marked points with opposite co-orientation:

Relations (84) and (85) imply that the morphisms

$$\left\langle \begin{array}{c} \end{array} \right\rangle_\omega \; : \; \langle S'\rangle_\omega \to \langle S\rangle_\omega \qquad \text{and} \qquad \omega\left\langle \begin{array}{c} \end{array} \right\rangle_\omega \; : \; \langle S\rangle_\omega \to \langle S'\rangle_\omega$$

are mutually inverse isomorphisms. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Corollary 4.14. *Let C be an object of SeSu which consists of k balanced circles, possibly nested. Then $\langle C\rangle_\omega$ is isomorphic to $A_\omega^{\otimes k}$ as a graded R_ω-module. In particular, it is free of graded rank $(q + q^{-1})^k$.*

References

[1] R. Akhmechet. Equivariant annular khovanov homology, 2020. arXiv:2008.00577.

[2] M. Auslander and O. Goldman. The Brauer group of a commutative ring. *Trans. Amer. Math. Soc.*, 97:367–409, 1960. doi:10.2307/1993378.

[3] D. Bar-Natan. On Khovanov's categorification of the Jones polynomial. *Algebr. Geom. Topol.*, 2:337–370, 2002. doi:10.2140/agt.2002.2.337.

[4] D. Bar-Natan. Khovanov's homology for tangles and cobordisms. *Geom. Topol.*, 9:1443–1499, 2005. doi:10.2140/gt.2005.9.1443.

[5] D. Bar-Natan and S. Morrison. The Karoubi envelope and Lee's degeneration of Khovanov homology. *Algebr. Geom. Topol.*, 6:1459–1469, 2006. doi:10.2140/agt.2006.6.1459.

[6] A. Beliakova, M. Hogancamp, K. K. Putyra, and S. M. Wehrli. On the functoriality of sl(2) tangle homology, 2019. arXiv:1903.12194.

[7] C. Blanchet. An oriented model for Khovanov homology. *J. Knot Theory Ramifications*, 19(2):291–312, 2010. doi:10.1142/S0218216510007863.

[8] C. Blanchet, N. Habegger, G. Masbaum, and P. Vogel. Topological quantum field theories derived from the Kauffman bracket. *Topology*, 34(4):883–927, 1995. doi:10.1016/0040-9383(94)00051-4.

[9] S. Caenepeel, B. Ion, and G. Militaru. The structure of Frobenius algebras and separable algebras. *K-Theory*, 19(4):365–402, 2000. doi:10.1023/A:1007849203555.

[10] C. L. Caprau. sl(2) tangle homology with a parameter and singular cobordisms. *Algebr. Geom. Topol.*, 8(2):729–756, 2008. doi:10.2140/agt.2008.8.729.

[11] C. L. Caprau. The universal $\mathfrak{sl}(2)$ cohomology via webs and foams. *Topology Appl.*, 156(9):1684–1702, 2009. doi:10.1016/j.topol.2009.02.001.

[12] S. U. Chase, D. K. Harrison, and A. Rosenberg. Galois theory and Galois cohomology of commutative rings. *Mem. Amer. Math. Soc.*, 52:15–33, 1965.

[13] D. Clark, S. Morrison, and K. Walker. Fixing the functoriality of Khovanov homology. *Geom. Topol.*, 13(3):1499–1582, 2009. doi:10.2140/gt.2009.13.1499.

[14] T. J. Ford. *Separable algebras*, volume 183 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2017.

[15] G. Hochschild. Cohomology and representations of associative algebras. *Duke Math. J.*, 14:921–948, 1947. URL: http://projecteuclid.org/euclid.dmj/1077474484.

[16] M. Jacobsson. An invariant of link cobordisms from Khovanov homology. *Algebr. Geom. Topol.*, 4:1211–1251, 2004. doi:10.2140/agt.2004.4.1211.

[17] L. Kadison. *New examples of Frobenius extensions*, volume 14 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1999. doi:10.1090/ulect/014.

[18] M. Khovanov. A categorification of the Jones polynomial. *Duke Math. J.*, 101(3):359–426, 2000. doi:10.1215/S0012-7094-00-10131-7.

[19] M. Khovanov. A functor-valued invariant of tangles. *Algebr. Geom. Topol.*, 2:665–741, 2002. doi:10.2140/agt.2002.2.665.

[20] M. Khovanov. sl(3) link homology. *Algebr. Geom. Topol.*, 4:1045–1081, 2004. doi:10.2140/agt.2004.4.1045.

[21] M. Khovanov. An invariant of tangle cobordisms. *Trans. Amer. Math. Soc.*, 358(1):315–327, 2006. doi:10.1090/S0002-9947-05-03665-2.

[22] M. Khovanov. Link homology and Frobenius extensions. *Fund. Math.*, 190:179–190, 2006. doi:10.4064/fm190-0-6.

[23] M. Khovanov and N. Kitchloo. A deformation of Robert–Wagner foam evaluation and link homology, 2020. arXiv:2004.14197.

[24] M. Khovanov, N. Kitchloo, and Y. Kononov. In preparation.

[25] M. Khovanov and L.-H. Robert. Foam evaluation and Kronheimer–Mrowka theories, 2018. arXiv:1808.09662.

[26] Y. Kononov. Private Communication, 2019.

[27] E. S. Lee. An endomorphism of the Khovanov invariant. *Adv. Math.*, 197(2):554–586, 2005. doi:10.1016/j.aim.2004.10.015.

[28] J.-L. Loday. *Cyclic homology*, volume 301 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1992. Appendix E by María O. Ronco. doi:10.1007/978-3-662-21739-9.

[29] J. H. Przytycki. When the theories meet: Khovanov homology as Hochschild homology of links. *Quantum Topol.*, 1(2):93–109, 2010. doi:10.4171/QT/2.

[30] J. Rasmussen. Khovanov homology and the slice genus. *Invent. Math.*, 182(2):419–447, 2010. doi:10.1007/s00222-010-0275-6.

[31] L.-H. Robert and E. Wagner. A closed formula for the evaluation of $\mathfrak{sl}_n$-foams. *To appear in Quantum Topology*, 2017. arXiv:1702.04140.

[32] T. Sano. Fixing the functoriality of khovanov homology: a simple approach, 2020. arXiv:2008.02131.

[33] P. Vogel. Functoriality of Khovanov homology. *J. Knot Theory Ramifications*, online ready(0):2050020, 2020. arXiv:1505.04545, doi:10.1142/S0218216520500200.

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY, NEW YORK, NY 10027, USA
Email address: khovanov@math.columbia.edu

UNIVERSITÉ DE GENÈVE, SECTION DE MATHÉMATIQUES, RUE DU LIÈVRE 2–4, 1227 GENÈVE, SWITZERLAND

UNIVERSITÉ DU LUXEMBOURG, RMATH, MAISON DU NOMBRE, 6 AVENUE DE LA FONTE, L-4365 ESCH-SUR-ALZETTE, LUXEMBOURG
Email address: louis-hadrien.robert@unige.ch
Email address: louis-hadrien.robert@uni.u