Hindawi Publishing Corporation
International Journal of Mathematics and Mathematical Sciences
Volume 2010, Article ID 978635, 58 pages
doi:10.1155/2010/978635

Research Article

Diagrammatics for Soergel Categories

Ben Elias and Mikhail Khovanov

Department of Mathematics, Columbia University, New York, NY 10027, USA

Correspondence should be addressed to Ben Elias, belias@math.columbia.edu

Received 4 May 2010; Accepted 31 December 2010

Academic Editor: Alistair Savage

The monoidal category of Soergel bimodules can be thought of as a categorification of the Hecke algebra of a finite Weyl group. We present this category, when the Weyl group is the symmetric group, in the language of planar diagrams with local generators and local defining relations.

1. Introduction

In this paper [1], Soergel gave a combinatorial description of a certain category of Harish-Chandra bimodules over a simple Lie algebra $\mathfrak{g}$. This category was and continues to be of primary interest in the infinite-dimensional representation theory of simple Lie algebras. Soergel discovered a functor from this category to a full subcategory of bimodules over a certain ring R, the objects of which are now commonly called *Soergel bimodules*. The category of Soergel bimodules is additive and monoidal, unlike the original category which is abelian, but it still has sufficient information to describe the original category. Soergel constructed an isomorphism between the Grothendieck ring of his category and the integral form of the Hecke algebra of the Weyl group W of $\mathfrak{g}$. Hence, Soergel's construction provides a categorification of the Hecke algebra.

Given a k-dimensional $\mathbb{C}$-vector space V and a generic $q \in \mathbb{C}$, there are commuting actions of the quantum group $U_q(\mathfrak{sl}_k)$ and the Hecke algebra H of the symmetric group S_n on $V^{\otimes n}$. These actions turn the quotient $U_q(\mathfrak{sl}_k)/J_1$ of the quantum group and H/J_2 of the Hecke algebra by the kernels of these action into a dual pair. A categorical realization of the triple $(V^{\otimes n}, U_q(\mathfrak{sl}_k)/J_1, H/J_2)$ was given by Grojnowski and Lusztig [2] via categories of perverse sheaves on products of flag and partial flag varieties, also see [3–6].

Many foundational ideas about categorification were put forward by Igor Frenkel in the early 90s (a small fraction of these ideas formed a part of the paper [7]). In particular, Frenkel conjectured [8] that quantum groups and not just their finite-dimensional quotients $U_q(\mathfrak{sl}_k)/J_1$ can be categorified. These conjectures remained open until recently, when categorifications of quantum $\mathfrak{sl}_2$ and $\mathfrak{sl}_k$ were discovered in [9, 10], with a related but different approach developed in [11, 12]. In the categorifications [9, 10] of quantum groups, 2 morphisms are given by linear combinations of planar diagrams, modulo local relations.

The parallel objective would be to categorify the Hecke algebra in the same spirit, using planar diagrams. Soergel had already provided a categorification, so it remains to ask whether his category can be rephrased diagrammatically. Diagrammatics should also provide a presentation of the category by generators and relations. A similar question was recently posed by Libedinsky [13], who essentially produced such a description for categorifications of Hecke algebras associated to "right-angled" Coxeter systems.

Here, we answer this question positively in the case of the Hecke algebra associated to the symmetric group. This is, of course, the Hecke algebra that appears in the Schur-Weyl duality for $V^{\otimes n}$. For notational convenience, we use $n+1$, not n, as our parameter and define a diagrammatical version $\mathfrak{DC}$ of the category of Soergel bimodules that categorifies the Hecke algebra of the symmetric group S_{n+1}.

In some sense, diagrammatic categorifications are very "low-tech," in that they can be described easily and do not rely on heavy machinery. While one can prove that Soergel bimodules categorify the Hecke algebra using only elaborate commutative algebra (see [14], although it is never stated explicitly), showing that indecomposable bimodules descend to the Kazhdan-Lusztig basis of the Hecke algebra utilizes Kazhdan-Lusztig theory [15, 16]. This, in turn, is related to fundamental developments in geometric representation theory like D-modules on flag manifolds [17, 18] and perverse sheaves [19]. One hopes that a diagrammatic approach will help to visualize and work with these sophisticated constructions, in the same way that the categorifications of quantum groups [9, 10] have led to an improved understanding of perverse sheaves on quiver varieties (see [20]). One also hopes that this approach can shed light on categorifications of representations of the Hecke algebra coming from the context of category $\mathcal{O}$ (see [21]).

We start with an intermediate category $\mathfrak{DC}_1$ whose objects are finite sequences $\underline{\mathbf{i}} = i_1 \cdots i_d$ of numbers between 1 and n. An object is represented graphically by marking d points in the standard position (say, having coordinates $1, \ldots, d$) on the x-axis and assigning labels $i_1, \ldots, i_d$ to marked points from left to right. Morphisms between $\underline{\mathbf{i}}$ and $\underline{\mathbf{j}}$ are given by linear combinations (with coefficients in a ground field $\Bbbk$) of planar diagrams embedded in the strip $\mathbb{R} \times [0,1]$. These diagrams are decorated planar graphs, where edges may extend to the boundary $\mathbb{R} \times \{0,1\}$. Each edge carries a label between 1 and n, so that the induced labellings of the lower and upper boundaries are $\underline{\mathbf{i}}$ and $\underline{\mathbf{j}}$, respectively. In the interior of the strip, we allow

(1) vertices of valence 1,

(2) vertices of valence 3 with all 3 edges carrying the same label,

(3) vertices of valence 4 seen as intersections of i and j-labelled lines with $|i-j| > 1$,

(4) vertices of valence 6 with the edge labelling $i, i+1, i, i+1, i, i+1$, reading clockwise around the vertex,

(5) boxes labelled by numbers between 1 and $n+1$ which float in the regions of the graph.

We impose a set of local relations on linear combinations of these diagrams, including invariance of diagrams under all isotopies. A subset of the relations says that i is a Frobenius object in the category $\mathfrak{DC}_1$.

The space of morphisms in $\mathfrak{DC}_1$ between $\underline{\mathbf{i}}$ and $\underline{\mathbf{j}}$ is naturally a graded vector space. Allowing grading shifts and direct sums of objects, then restricting to grading-preserving morphisms, and finally passing to the Karoubian closure of the category results in a graded $\Bbbk$-linear additive monoidal category $\mathfrak{D}\mathcal{C}$. Our main result (Theorem 4.22 in Section 4) is an explicit equivalence between this category and the category $\mathcal{SC}$ of Soergel bimodules.

The category $\mathfrak{D}\mathcal{C}$ is monoidal, and can be viewed as a 2-category with a single object. It may be easier to tackle the diagrammatics after reading an introductory reference on diagrammatics for 2-categories. Such an introduction can be found in [9]. This may make it easier to explore similarities with the categorifications of quantum groups in [9, 10], where regions of diagrams are labelled by integers in [9] and integral weights of $\mathfrak{sl}_n$ in [10]. Boxes floating in regions are superficially analogous to floating bubbles of [9, 10]. Unlike the diagrammatic categorifications in [9, 10], our lines do not carry dots and are not oriented.

There is another way to view our diagrammatics, which is not developed in this paper. Rouquier [22, 23] defined an action of the Coxeter braid group associated to W on the category of complexes of Soergel bimodules up to homotopies, which is related to a braid group action using Harish-Chandra bimodules that had been known for some time. These complexes were later used in an alternative construction [24] of a triply graded link homology theory [25] categorifying the HOMFLY-PT polynomial [25–28]. In this approach, a product Soergel bimodule $B_{i_1} \otimes \cdots \otimes B_{i_d}$ is depicted by a planar diagram given by concatening elementary planar diagrams lying in the xy-plane that consist of $n+1$ strands going up, with i and $i+1$-st strands merging and splitting, see [24, Figures 2 and 3]. Morphisms between product bimodules can be realized by linear combinations of foams—decorated two-dimensional CW-complexes embedded in $\mathbb{R}^3$ with suitable boundary conditions. Foams have been implicit throughout papers on triply graded link homology (see [29] for instance, where various arrows between planar diagrams can be implemented by foams), and explicitly appear in the papers on their doubly-graded cousins, see [30, 31] and references therein.

Foams are 3-dimensional objects; they are two-dimensional CW-complexes embedded in $\mathbb{R}^3$ that produce cobordisms along the z-axis direction between planar objects corresponding to product Soergel bimodules. The planar diagrams of our paper are two-dimensional encodings of these foams, essentially projections of the foams onto the yz-plane along the x-axis.

It was shown in [32] that the action of the braid group on the homotopy category of Soergel bimodules extends to a (projective) action of the category of braid cobordisms. Thus, the homotopy category of $\mathfrak{D}\mathcal{C}$ produces invariants of braid cobordisms, so that our planar diagrammatics carry information about four-dimensional objects. This informational density indicates the efficiency of such encodings.

Addendum 1. *Since this paper first appeared, the diagrammatics developed here have led to several developments which we briefly mention here. In [33] it is shown that Rouquier's braid group action lifts functorially to the braid cobordism category. In [34, 35], a functor is given from the category* $\mathfrak{D}\mathcal{C}$ *to categories of foams used in link homology. Together, these papers show that the encodings mentioned above are more than simply heuristic. In [36], the Temperley-Lieb algebra is categorified as a quotient of* $\mathfrak{D}\mathcal{C}$*. Additional statements relating this paper to either newer papers or to previous versions of this paper are found sparsely under a similar "Addendum" heading.*

2. Preliminaries

Henceforth, we will fix a positive integer n. Indices i, j, and k will range over $1, \ldots, n$ if not otherwise specified. Finite ordered sequences of such indices (allowing repetition) will be denoted $\underline{\mathbf{i}} = i_1 \cdots i_d$, as well as $\underline{\mathbf{j}}$ and $\underline{\mathbf{k}}$. The length of the sequence will be denoted $d = d(\underline{\mathbf{i}})$.

For sequences of length $d = 1$ where the single entry is i, we use i and $\underline{\mathbf{i}}$ interchangeably. Occasionally $i+1$ will also be used as an index, and whenever this occurs we make the tacit assumption that $i \leq n-1$, so that all indices used remain between 1 and n. The same goes for $i-1$, $i+2$, and the like. We denote the length 0 sequence by the empty set symbol $\emptyset$.

We work over a field $\Bbbk$, usually assuming that $\operatorname{Char} \Bbbk \neq 2$, and sometimes specializing it to $\mathbb{C}$.

Given a noetherian ring R, the category R-molf-R is the full subcategory of R-bimodules consisting of objects which are finitely generated as left R-modules. If R is graded, the category R-molf$_{\mathbb{Z}}$-R is the analogous subcategory of graded R-bimodules and grading-preserving homomorphisms.

2.1. Hecke Algebra

Let $(W, \mathcal{S})$ be a Coxeter system of a finite Weyl group W, with length function $l : W \to \mathbb{N} = \{0, 1, 2, \ldots\}$, and let $e \in W$ be the identity. The Hecke algebra H is an algebra over $\mathbb{Z}[v, v^{-1}]$ (we follow Soergel's use [37] of the variable v; related variables are denoted in the literature by $t = v^{-1}$ and $q = t^2$), which is free as a module with basis T_w, $w \in W$. Multiplication in this basis is given by $T_x T_w = T_{xw}$ when $l(v) + l(w) = l(vw)$, and $T_s^2 = (v^{-2} - 1)T_s + v^{-2}T_e$ for $s \in \mathcal{S}$. T_e is the identity element in H and will often be written as 1.

In the case we are interested in presently, $W = S_{n+1}$, and $\mathcal{S}$ consists of the transpositions $s_i = (i, i+1)$ for $i = 1, \ldots, n$. The element T_{s_i} will be denoted T_i. The Hecke algebra has a presentation over $\mathbb{Z}[v, v^{-1}]$, being generated by T_i subject to the relations

$$\begin{gathered} T_i^2 = \left(v^{-2} - 1\right)T_i + v^{-2}, \\ T_i T_j = T_j T_i \quad \text{for } \left|i - j\right| \geq 2, \\ T_i T_{i+1} T_i = T_{i+1} T_i T_{i+1}. \end{gathered} \tag{2.1}$$

Clearly then, H is also generated as an algebra by $b_i \overset{\text{def}}{=} C'_{s_i} = v(T_i + 1)$, $1 \leq i \leq n$, and the relations above transform into

$$\begin{gathered} b_i^2 = \left(v + v^{-1}\right)b_i, \\ b_i b_j = b_j b_i \quad \text{for } \left|i - j\right| \geq 2, \\ b_i b_{i+1} b_i + b_{i+1} = b_{i+1} b_i b_{i+1} + b_i. \end{gathered} \tag{2.2}$$

We often write the monomial $b_{i_1} b_{i_2} \cdots b_{i_d}$ as $b_{\underline{\mathbf{i}}}$ where $\underline{\mathbf{i}} = i_1 \cdots i_d$. Notice that $b_\emptyset = 1$.

Let $a \mapsto \overline{a}$ be the involution of $\mathbb{Z}[v, v^{-1}]$ determined by $\overline{v} = v^{-1}$. It extends to an involution of H given by

$$\overline{\sum a_w T_w} = \sum \overline{a_w} T_{w^{-1}}^{-1}. \tag{2.3}$$

In particular, $\overline{T_i} = T_i^{-1} = v^2 T_i + v^2 - 1$.

Kazhdan and Lusztig [15] defined a pair of bases $\{C_w\}_{w \in W}$ and $\{C'_w\}_{w \in W}$ for H, which immediately proved to be of fundamental importance for representation theory and combinatorics. The two bases are related via a suitable involution of H, and the elements of the second Kazhdan-Lusztig basis $\{C'_w\}_w$ are determined by the two properties

$$\overline{C'_w} = C'_w,$$
$$C'_w = v^{l(w)} \sum_{y \leq w} P_{y,w} T_y, \tag{2.4}$$

where $P_{y,w} \in \mathbb{Z}[v^{-2}]$ has negative v-degree strictly less than $l(w) - l(y)$ for $y < w$ and $P_{w,w} = 1$. There is no simple formula expressing C'_w in terms of T_y, but observe that $C'_e = 1$ and $C'_{s_i} = b_i = v(T_i + 1)$. For a good introduction to the Kazhdan-Lusztig basis, see [37].

Let $\varepsilon : H \to \mathbb{Z}[v, v^{-1}]$ be the $\mathbb{Z}[v, v^{-1}]$-linear map given by $\varepsilon(T_e) = 1$ and $\varepsilon(T_w) = 0$ if $w \neq e$. Thus, ε simply picks up the coefficient of T_e in x. The easily checked property $\varepsilon(T_i x) = \varepsilon(x T_i)$ for any $x \in H$ implies that $\varepsilon(xy) = \varepsilon(yx)$, for all $x, y \in H$, so that ε is a trace map and turns H into a symmetric Frobenius $\mathbb{Z}[v, v^{-1}]$-algebra.

Denote by ω a v-antilinear anti-involution $\omega : H \to H$ defined uniquely by $\omega(b_i) = b_i$. The anti-involution and v-antilinearity conditions say that $\omega(xy) = \omega(y)\omega(x)$ and $\omega(ax) = \overline{a}\omega(x)$, for $x, y \in H$ and $a \in \mathbb{Z}[v, v^{-1}]$.

Consider the pairing $(,) : H \times H \to \mathbb{Z}[v, v^{-1}]$ of $\mathbb{Z}$-modules given by

$$(x, y) = \varepsilon(\omega(x) y). \tag{2.5}$$

It satisfies the following properties:

(1) the pairing is *semi-linear*, that is, $(ax, y) = \overline{a}(x, y)$ while $(x, ay) = a(x, y)$, for $a \in \mathbb{Z}[v, v^{-1}]$,

(2) b_i is self-adjoint, that is, $(x, b_i y) = (b_i x, y)$ and $(x, y b_i) = (x b_i, y)$,

(3) if $\underline{\mathbf{i}} = i_1 \cdots i_d$ with $i_1 < i_2 < \cdots < i_d$ then $(1, b_{\underline{\mathbf{i}}}) = v^d$. Such a monomial $b_{\underline{\mathbf{i}}}$ is called an *increasing monomial*, and $\underline{\mathbf{i}}$ an *increasing sequence*. When $d = 0$, the sequence $\underline{\mathbf{i}}$ is empty and $(1, 1) = 1$.

Remark 2.1. It is not difficult to observe that $(,)$ is the unique form satisfying these three properties. This is because the Hecke algebra has a spanning set over $\mathbb{Z}[t, t^{-1}]$ consisting of monomials $b_{\mathbf{i}}$, and every monomial may be reduced, by cycling the last b_i to the beginning and by applying the Hecke algebra relations, to an increasing monomial. This is a simple combinatorial argument that we leave to the reader.

2.2. Soergel Bimodules

In [1], Soergel introduced a category of bimodules which categorified the Hecke algebra, and later generalized his construction to any Coxeter group W [14]. Within the category R-molf$_{\mathbb{Z}}$-R, for R a certain graded $\Bbbk$-algebra ($\Bbbk$ an infinite field of characteristic $\neq 2$), he identified indecomposable modules B_w for $w \in W$, such that the only indecomposable summands of tensor products of B_w's are $B_{w'}$ for $w' \in W$. Thus, the subcategory of R-molf$_{\mathbb{Z}}$-R generated additively by the B_w has a tensor product, and its Grothendieck ring is isomorphic to H, under the isomorphism sending C'_s to $[B_s]$. Moreover, every B_w shows up as a summand of some tensor product of various B_s for $s \in \mathcal{S}$. While the general B_w may be difficult to describe, B_s has an easy description.

It is conjectured in [14] that this isomorphism sends $[B_w]$ to C'_w for all $w \in W$, and it is proven for $\Bbbk = \mathbb{C}$ and W a Weyl group in [1], using geometric methods.

Henceforth we specialize to the case where $W = S_{n+1}$ and $\mathcal{S} = \{s_i\}$. We also make one additional change from Soergel's conventions.

Remark 2.2. Soergel defines R to be the coordinate ring of the n-dimensional reflection representation V of S_{n+1}, while we find it easier to consider R', the coordinate ring of the $n+1$-dimensional standard representation V'. This is akin to treating $\mathfrak{gl}_n$ instead of $\mathfrak{sl}_n$, and a similar convention is adopted in [24]. The bimodules B_w are defined in [14] to be the coordinate rings of unions of "twisted diagonals" in $V \times V$, and B'_w can be defined analogously for V'. Now $R' \cong R \otimes_{\Bbbk} \Bbbk[y]$ and the entire story of B_w translates to B'_w by base extension. Conversely, R is a quotient of R' by the first elementary symmetric polynomial e_1, which is a central element of our category, so that the entire story of B'_w translates easily to B_w under the quotient. We will interest ourselves with the B'_w story below because the ring R' is slightly more intuitive, and mention briefly the changes required to deal with R in Section 4.6. Since we only use B'_w and R' below, we will denote them as B_w and R instead to avoid an apostrophe catastrophe.

With these conventions, we now make the story explicit.

Notation 2.3. Let $R = \Bbbk[x_1, x_2, \ldots, x_{n+1}]$ be the ring of polynomials in $n+1$ variables, with the natural action of S_{n+1}. The ring R is graded, with $\deg(x_i) = 2$. If W is the subgroup of S_{n+1} generated by transpositions $\{s_{i_1}, \ldots, s_{i_r}\}$, then we denote the ring of invariants under W as $R^{i_1,\ldots,i_r}$ or R^W. Thus R^i are the invariants under the transposition $(i, i+1)$.

Since R is an R^W-algebra, $\otimes_{R^W}$ is a bifunctor sending two R-bimodules to an R-bimodule. Henceforth, $\otimes$ with no subscript denotes tensoring over R, while $\otimes_{i_1,\ldots,i_r}$ denotes tensoring over the subring $R^{i_1,\ldots,i_r}$. Most commonly we will just use $\otimes_i$ for various indices i.

Definition 2.4. The Soergel bimodule B_i is $R \otimes_i R\{-1\}$, where $\{m\}$ denotes a grading shift.

Notation 2.5. We denote by $B_{\underline{i}}$ the tensor product $B_{i_1} \otimes B_{i_2} \otimes \cdots \otimes B_{i_d}$.

Note that $B_{\emptyset} = R$ and

$$B_{\underline{i}} = (R \otimes_{i_1} R\{-1\}) \otimes (R \otimes_{i_2} R\{-1\}) \otimes \cdots = R \otimes_{i_1} R \otimes_{i_2} \cdots \otimes_{i_d} R\{-d\}. \tag{2.6}$$

We reiterate this important point: a typical element of a tensor product of d generators B_i can be expressed (up to linear combination) by a choice of $d+1$ polynomials, one in each slot separated by the tensors. Multiplication by an element of R in any particular slot is an R-bimodule endomorphism.

For each i there is a map of graded vector spaces $\partial_i : R \to R^i\{2\}$, called the *Demazure operator*, sending $f \mapsto (f - s_i(f))/(x_i - x_{i+1})$. This map is R^i-linear. Since $\partial_i(f)$ is s_i-invariant, it is not hard to see that $P_i(f) = f - x_i \partial_i(f)$ is also s_i-invariant. Since $f = P_i(f) + x_i \partial_i(f)$, this implies that R is a free graded R^i-module of rank two, with homogeneous generators 1 and x_i. In other words, there is an isomorphism $R \cong R^i \oplus R^i\{2\}$ of graded R^i-modules, sending $f \mapsto (P_i(f), \partial_i(f))$, with inverse $(f, g) \mapsto f + g x_i$.

From the isomorphism $R \cong R^i \oplus R^i\{2\}$ of graded R^i-modules just illustrated, one can deduce other isomorphisms. For instance, $B_i = R \otimes_i R\{-1\} \cong R\{-1\} \oplus R\{1\}$ as graded left (or right) R-modules. Repeating this, we see that $B_{\underline{i}}$ is a free left R-module of rank $2^{d(\underline{i})}$, and properly belongs in R-molf$_{\mathbb{Z}}$-R. Finally, we can deduce an isomorphism $B_i \otimes_i B_i \cong B_i\{1\} \oplus B_i\{-1\}$, which unlike the previous isomorphisms is actually an isomorphism of R-bimodules.

Remark 2.6. Let us make this slightly more explicit. To give the isomorphism of left R-modules $B_i \cong R\{-1\} \oplus R\{1\}$, note that

$$f \otimes g = f \otimes P_i(g) + f \otimes x_i \partial_i(g) = P_i(g) f \otimes 1 + \partial_i(g) f \otimes x_i. \tag{2.7}$$

Rewriting a term $1 \otimes g$ as a sum of terms like above will happen often, and we refer to it as *forcing* the polynomial g across the tensor. If g is s_i-invariant then it may be slid across leaving nothing behind, while an arbitrary g when forced leaves terms with either 1 or x_i behind (alternatively, we may choose to leave 1 and x_{i+1} behind, if it is more convenient). We consistently use the term "slide" instead of "force" when the polynomial g is invariant so it can be moved across without any ado.

Inside $B_i \otimes B_i$, taking an element $f \otimes g \otimes h$ and forcing g to the left (or right) is now an R-bilinear operation, since multiplication on the left or right will only affect f or h, not g. This gives the isomorphism $B_i \otimes B_i \cong B_i\{1\} \oplus B_i\{-1\}$ of R-bimodules, via $f \otimes g \otimes h \mapsto (f P_i(g) \otimes h, f \partial_i(g) \otimes h)$ with inverse $(f_1 \otimes h_1, f_2 \otimes h_2) \mapsto f_1 \otimes 1 \otimes h_1 + f_2 \otimes x_i \otimes h_2$. These maps are R-bimodule morphisms, since the only polynomials which can slide from f_i to h_i in both the source and the target of the map are those polynomials in R^i.

Remark 2.7. We also remark on spanning sets for $B_{\underline{i}}$ as R-bimodules. For instance, we've seen that $B_i \otimes B_i$ has a spanning set $\{1 \otimes 1 \otimes 1,\ 1 \otimes x_i \otimes 1\}$. The bimodule $B_i \otimes B_j$ for $j \neq i$ has a spanning set $\{1 \otimes 1 \otimes 1\}$, since any polynomial in the middle can be forced to the left leaving at most x_i behind (or x_{i+1}, which we choose when $j = i - 1$), and that can be slid to the right; thus $1 \otimes 1 \otimes 1$ generates it as a R-bimodule. The bimodule $B_i \otimes B_j \otimes B_k \otimes B_i$ has a spanning set $\{1 \otimes 1 \otimes 1 \otimes 1 \otimes 1,\ 1 \otimes x_i \otimes 1 \otimes 1 \otimes 1\}$. This is because all polynomials anywhere between the two i tensors may be slid out, leaving x_i somewhere in-between. As an exercise, the reader may generalize this argument to an arbitrary $B_{\underline{i}}$ and find a spanning set as a R-bimodule, consisting of $2^{m(\underline{i})}$ terms, where $m(\underline{i})$ is the number of pairs $1 \leq r < s \leq d$ such that $i_r = i_s$ and $i_t \neq i_s$ for t between r and s. Between such a pair, one either places a linear "unslideable" term like x_{i_r}, or just 1. Note that $m(\underline{i})$ is equal to $d(\underline{i})$ minus the number of distinct indices in $\underline{\mathbf{i}}$.

For more information on Soergel bimodules and their applications we refer the reader to [1, 22, 24, 38–44] and references therein.

2.3. The Soergel Categorification

Let us summarize the Soergel categorification of the Hecke algebra H of the symmetric group S_{n+1} and the various structures on it, following [1, 14, 22, 38, 41].

Several subcategories of R-molf-R and R-molf$_{\mathbb{Z}}$-R will play a role in what follows. Let $\mathcal{SC}_1$ be the full subcategory of R-molf-R whose objects consist of $B_{\underline{i}}$ for all sequences of indices $\underline{\mathbf{i}}$; these are called *Bott-Samelson bimodules*. Since R is a commutative ring, the Hom spaces in $\mathcal{SC}_1$ are in fact enriched in R-molf$_{\mathbb{Z}}$-R. Let $\mathcal{SC}_2$ be the subcategory of R-molf$_{\mathbb{Z}}$-R whose objects are finite direct sums of various graded shifts of objects in $\mathcal{SC}_1$ and the morphisms are all grading-preserving bimodule homomorphisms. Finally, let $\mathcal{SC}$ be the Karoubi envelope of $\mathcal{SC}_2$, a category equivalent to the full subcategory of R-molf$_{\mathbb{Z}}$-R which contains all summands of objects of $\mathcal{SC}_2$

$$\mathcal{SC}_1 \xrightarrow{\text{Grading shifts and direct sums}} \mathcal{SC}_2 \xrightarrow{\text{Karoubi envelope}} \mathcal{SC}. \tag{2.8}$$

In general, the Karoubi envelope is the category which formally includes all "summands," where a summand of an object M is identified by an idempotent $p \in \mathrm{End}(M)$ corresponding to projection to that summand. Since R-molf$_{\mathbb{Z}}$-R is abelian, it is idempotent-closed, and the Karoubi envelope of $\mathcal{SC}_2$ can be realized as a subcategory in R-molf$_{\mathbb{Z}}$-R. We refer the reader to [45] for basic information about Karoubi envelopes.

This final category $\mathcal{SC}$ (the category of *Soergel bimodules*) is a $\Bbbk$-linear additive monoidal category with the Krull-Schmidt property. Soergel showed that, when $\Bbbk$ is an infinite field of characteristic other than 2, the indecomposable bimodules in this category are enumerated by elements of the Weyl group and grading shifts (Theorem 6.16 in [14]). They are denoted by $B_w\{j\}$ for $w \in W$ and $j \in \mathbb{Z}$. An indecomposable B_w is determined by the condition that it appears as a direct summand of $B_{\underline{\mathbf{i}}}$, where $\underline{\mathbf{i}} = i_1 \cdots i_d$ and $s_{i_1} \cdots s_{i_d}$ is a reduced presentation of w, and does not appear as direct summand of any $B_{\underline{\mathbf{i}}}$, for sequences $\underline{\mathbf{i}}$ of length less than $l(w)$.

The Hecke algebra H is canonically isomorphic to $K_0(\mathcal{SC})$, the Grothendieck group of $\mathcal{SC}$. Multiplication by v corresponds to the grading shift: $[M\{d\}] = v^d[M]$. Multiplication in the Hecke algebra corresponds to the tensor product of bimodules:

$$[M] \cdot [N] := [M \otimes_R N]. \tag{2.9}$$

The isomorphism takes $[B_i]$ to b_i and $[B_{\underline{\mathbf{i}}}]$ to $b_{\underline{\mathbf{i}}}$.

Remark 2.8. Nothing prevents one from defining a category $\mathcal{SC}_{\mathbb{Z}}$ where the field $\Bbbk$ is replaced with $\mathbb{Z}$ in the definitions of the previous section. Thus one could define the category for any ring. However, one does not have control over the size of the Grothendieck group in this instance. When defined over a field $\Bbbk$ of characteristic $\neq 2$, we may use Theorem 6.16 of [14] to classify indecomposables and get results about the Grothendieck group. When $\Bbbk = \mathbb{C}$, Soergel has shown that the Kazhdan-Lusztig basis $\{C'_w\}$ satisfies $C'_w = [B_w]$. This is unknown in general.

Relation (2.2) lifts to isomorphisms of graded bimodules in $\mathcal{SC}$

$$B_i \otimes B_i \cong B_i\{1\} \oplus B_i\{-1\}, \tag{2.10}$$

$$B_i \otimes B_j \cong B_j \otimes B_i \quad \text{for } |i-j| \geq 2, \tag{2.11}$$

$$(B_i \otimes B_{i+1} \otimes B_i) \oplus B_{i+1} \cong (B_{i+1} \otimes B_i \otimes B_{i+1}) \oplus B_i. \tag{2.12}$$

It is important to note that these isomorphisms take place in $\mathcal{SC}_2$, not $\mathcal{SC}_1$, since the latter does not have grading shifts or direct sums of objects. However, the same information can be encapsulated in inclusion and projection morphisms of various degrees, which do live in $\mathcal{SC}_1$. This will be explored in Section 4.5.

The first isomorphism has already been made explicit in Remark 2.6. We have chosen a specific isomorphism; other choices were possible. The second and third isomorphisms come from the following isomorphisms in R-molf$_{\mathbb{Z}}$-R:

$$B_i \otimes B_j \cong R \otimes_{i,j} R\{-2\} \cong B_j \otimes B_i \quad \text{for } |i-j| \geq 2, \tag{2.13}$$

$$B_i \otimes B_{i+1} \otimes B_i \cong B_i \oplus (R \otimes_{i,i+1} R\{-3\}), \tag{2.14}$$

$$B_{i+1} \otimes B_i \otimes B_{i+1} \cong B_{i+1} \oplus (R \otimes_{i,i+1} R\{-3\}). \tag{2.15}$$

These isomorphisms will be made explicit in Section 4.5.

The ring $\mathbb{Z}[v,v^{-1}]$ is canonically isomorphic to the Grothendieck group of the category R-fmod of finitely generated graded free R-modules. Under this isomorphism

$$K_0(R\text{-fmod}) \cong \mathbb{Z}\left[v,v^{-1}\right] \tag{2.16}$$

$[R]$ goes to 1 and $[R\{d\}]$ to v^d. In particular, given any graded free R-module M, its image in the Grothendieck group will be its *graded rank*, calculated by choosing a homogeneous R-basis $\{y_j\}$ of M and letting $\operatorname{grk} M \overset{\text{def}}{=} \sum_j v^{\deg y_j}$. The bar involution on $\mathbb{Z}[v,v^{-1}]$ lifts to the contravariant equivalence that takes $M \in R$-fmod to $\operatorname{Hom}_R(M,R)$, the latter naturally viewed as an R-module.

In the category $\mathcal{SC}_1$, given any objects M,N, the space $\operatorname{Hom}_{\mathcal{SC}_1}(M,N)$ is a graded free finitely generated left R-module. By extension, the same is true of the module $\oplus_{m\in\mathbb{Z}}\operatorname{Hom}_{\mathcal{SC}}(M\{m\},N)$. Shifting the grading of N will shift the grading of this Hom space in the same direction, while shifting M will shift the Hom space in the opposite direction. Therefore, the bifunctor

$$\operatorname{Hom}_{\mathcal{SC}}(\cdot,\cdot) : \mathcal{SC}^{\text{op}} \times \mathcal{SC} \longrightarrow R\text{-fmod} \tag{2.17}$$

categorifies a semilinear form $H \times H \to \mathbb{Z}[v,v^{-1}]$ which sends

$$([M],[N]) \longmapsto \operatorname{grk}(\oplus_{m\in\mathbb{Z}}\operatorname{Hom}_{\mathcal{SC}}(M\{m\},N)). \tag{2.18}$$

The bimodule B_i self-biadjoint, that is, that $\operatorname{Hom}_{\mathcal{SC}_1}(M\otimes B_i,N) = \operatorname{Hom}_{\mathcal{SC}_1}(M,N\otimes B_i)$ and $\operatorname{Hom}_{\mathcal{SC}_1}(B_i\otimes M,N) = \operatorname{Hom}_{\mathcal{SC}_1}(M,B_i\otimes N)$ via some adjunction maps. This will become explicit in Section 3.1. In fact, every bimodule M in $\mathcal{SC}$ has a biadjoint bimodule $\Omega(M)$ such that tensoring with M on the left (resp., right) is biadjoint to tensoring with $\Omega(M)$ on the left (resp., right). Due to a cyclicity property (see the next section) any homomorphism $f : M \to N$ of bimodules dualizes to a canonical homomorphism $\Omega(f) : \Omega(N) \to \Omega(M)$, so that Ω can be made into an antiequivalence of $\mathcal{SC}$, lifting the anti-involution ω. Notice that Ω takes B_i to itself and $B_{\underline{\mathbf{i}}}$ to $B_{\underline{\mathbf{j}}}$ where $\underline{\mathbf{j}}$ is given by reading $\underline{\mathbf{i}}$ from right to left.

Unsurprisingly, the semilinear product on H above (induced by the Hom bifunctor) agrees with the one defined in Section 2.1. To check this, following Remark 2.1 and using the self-biadjointness of B_i, we only need to show the following claim.

Claim 2.9. When $\underline{\mathbf{i}}$ is an increasing sequence, $\operatorname{Hom}_{\mathcal{SC}_1}(R,B_{\underline{\mathbf{i}}})$ is a free left R-module of rank 1, generated by a morphism of degree d, the length of $\underline{\mathbf{i}}$.

Proof. We only sketch this result. An R-bimodule map from R to B_i is clearly determined by an element of B_i on which right and left multiplication by polynomials in R are identical. Any element of B_i is of the form $m = f\otimes 1 + g\otimes x_i$, and clearly $x_j m = m x_j$ for $j \neq i, i+1$, and $(x_i + x_{i+1})m = m(x_i + x_{i+1})$. Hence m can be the image of 1 under a bimodule map from R if and only if $x_i m = m x_i$.

$$m x_i = f\otimes x_i + g\otimes x_i^2 = f\otimes x_i + g(x_i + x_{i+1})\otimes x_i - g(x_i x_{i+1})\otimes 1. \tag{2.19}$$

Thus m can be an image if and only if $g(x_i x_{i+1}) = -f x_i$ and $f + g(x_i + x_{i+1}) = g x_i$, if and only if $f = -g x_{i+1}$. Such elements form a left R-module generated by the case $g = 1$, $f = -x_{i+1}$, or in other words by $m = 1 \otimes x_i - x_{i+1} \otimes 1$. The element m has degree 1 in B_i, so we deduce that $(R, B_i) = v$. Let us call φ_i the corresponding map $R \to B_i$, $1 \mapsto m$.

One may use the same argument for the general case. Suppose $\underline{\mathbf{i}}$ is increasing and has length d, then $B_{\underline{\mathbf{i}}}$ is a free left R-module of degree 2^d, with generators $\{1 \otimes_{i_i} f_1 \cdots \otimes_{i_k} f_k\}$ ranging over terms where either $f_l = 1$ or $f_l = x_{i_l}$. As an exercise, the reader may find the criteria for a general element to be ad-invariant under x_i, and verify that the only possible bimodule maps $R \to B_{\underline{\mathbf{i}}}$ are R-multiples of the following iterated version of φ_i:

$$R \longrightarrow B_{i_1} \cong B_{i_1} \otimes R \longrightarrow B_{i_1} \otimes B_{i_2} \cong \cdots . \tag{2.20}$$

The first map is φ_{i_1}, the second map is $\mathrm{Id} \otimes \varphi_{i_2}$, and so forth. This generator is a map of degree d, so that $\mathrm{Hom}_{\mathcal{SC}_1}(R, B_{\underline{\mathbf{i}}}) = R\{d\}$ and $([R], [B_{\underline{\mathbf{i}}}]) = v^d$. □

Remark 2.10. We have swept the calculation under the rug, so the dependence of this claim on the fact that $\underline{\mathbf{i}}$ is increasing is unclear. In general, when $\underline{\mathbf{i}}$ has a repeated index there will be additional maps from R to $B_{\underline{\mathbf{i}}}$. Roughly speaking, certain symmetry conditions are placed upon polynomials in order for them to slide across certain tensors. The duplication of an index will yield a redundant symmetry condition that places fewer constraints on an ad-invariant element than would be expected from the length of the sequence. We suggest the reader try to find all the maps from R to $B_{\underline{\mathbf{i}}}$ in the length 2 case, first when $\underline{\mathbf{i}} = ij$ has no repeated index and then when $\underline{\mathbf{i}} = ii$ has a repeated index. This should illustrate the main idea.

Because it is a crucial statement which we use again and again, we restate the overall result and give a reference.

Proposition 2.11. *Given two sequences $\underline{\mathbf{i}}$ and $\underline{\mathbf{j}}$, $\mathrm{Hom}_{\mathcal{SC}}(B_{\underline{\mathbf{i}}}, B_{\underline{\mathbf{j}}})$ is a free graded left (or right) R-module of rank $\varepsilon(\omega(b_{\underline{\mathbf{i}}})b_{\underline{\mathbf{j}}})$, where ε is the standard trace on H defined in Section 2.1.*

Proof. This is deduced from the above discussion. For Soergel's proof, see [14], although it is somewhat obscured. Theorem 5.15 in that paper and especially its proof together state this result, once one unravels exactly what h_Δ means. Propositions 5.7 and 5.9 state that $h_\Delta(B_{\underline{\mathbf{i}}}) = b_{\underline{\mathbf{i}}}$, since h_Δ sends $B_i \otimes \cdot$ to $b_i \cdot$ and sends R to 1. □

The facts below will not be used in this paper.

For a Soergel bimodule M the space of bimodule homomorphisms $\mathrm{Hom}_{\mathcal{SC}_1}(R, M)$ is just the 0th Hochschild cohomology $\mathrm{HH}^0(R, M)$ of M. Thus, unraveling the definitions,

$$\varepsilon([M]) := \mathrm{grk}\Big(\mathrm{HH}^0(R, M)\Big). \tag{2.21}$$

Calculations in Hochschild cohomology can be used to provide a proof of the claim above. One could also define a trace map $\tau(x) = \overline{\varepsilon(\omega(x))}$ which is the decategorification of the functor HH_0 of taking the 0th Hochschild homology.

Hecke algebra H has a trace more sophisticated than ε or τ, called the Ocneanu trace [46], which describes the HOMFLY-PT polynomial. The categorification of the Ocneanu trace utilizes all Hochschild homology groups rather than just HH_0, see [24, 43, 47].

The Rouquier complexes mentioned in the introduction are described here. Invertible elements T_i that satisfy the braid relations become [22] invertible complexes

$$0 \longrightarrow R \longrightarrow B_i\{-1\} \longrightarrow 0, \tag{2.22}$$

in the homotopy category of the Soergel category (with R sitting in cohomological degree -1). This aligns with the fact that $T_i = v^{-1}b_i - 1$ in the Hecke quotient of the braid group. Their inverses T_i^{-1} become inverse complexes

$$0 \longrightarrow B_i\{1\} \longrightarrow R \longrightarrow 0, \tag{2.23}$$

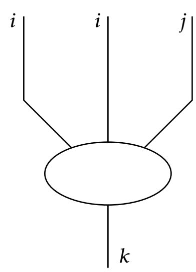
with R in cohomological degree 1, agreeing with $T_i^{-1} = vb_i - 1$. The homomorphism from the braid group into the Hecke algebra is categorified by a projective functor from the category of braid cobordisms between $(n+1)$-stranded braids to the category of endofunctors of the homotopy category of the Soergel category [32].

Remark 2.12. Note that this convention for Rouquier complexes is opposite that found in [22], which is to say that we have flipped T_i with T_i^{-1}. Presumably this arises because we are using v as a parameter for the grading shift, and not $t = v^{-1}$. The choice of v is more natural for the calculation of graded ranks of Hom spaces.

2.4. Diagrammatic Calculus for Bimodule Maps

We follow the standard rules for the diagrammatic calculus of bimodules, or more generally for the diagrammatic calculus of a monoidal category. An excellent and thorough explanation of these rules can be found in [9], so we will provide a quick summary. A planar diagram will represent a morphism of R-bimodules, with the following conventions. A horizontal slice or line segment in this diagram will represent an object (an R-bimodule). A rectangle inside the plane will represent a morphism from its bottom horizontal line segment to its top horizontal line segment.

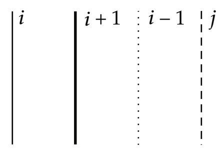
The R-bimodule B_i is denoted by a point (on a horizontal line segment) labelled i. The tensor product of bimodules is depicted by a sequence of labelled points on a horizontal line segment, so that tensor products are formed "horizontally". A vertical line labelled i denotes the identity endomorphism of B_i, and similarly labelled lines placed side by side denote the identity endomorphism of the tensor product. More general bimodule maps are represented by some symbols connecting the appropriate lines, and are composed "vertically", and tensored "horizontally". All diagrams are read from bottom to top, so that the following diagram represents a bimodule map from B_k to $B_i \otimes B_i \otimes B_j$:



A horizontal line segment which does not contain any marked points represents R as a bimodule over itself, the monoidal identity. The empty rectangle represents the identity endomorphism of R. Planar diagrams without top and bottom endpoints (without boundary) represent more general endomorphisms of R.

The structure of bimodule categories (or more generally strict 2-categories) guarantees that a planar diagram will unambiguously denote a morphism of bimodules.

We will be using so many such pictures that it will become cumbersome to continuously label each line by an index. Generally, the calculations we do will work independently for each i, and can be expressed with diagrams that use lines labelled i, $i+1$ and the like. In these circumstances, when there is no ambiguity, we will fix an index i and draw a line labelled i with one style, a line labelled $i+1$ with a different style, and so forth, maintaining the same conventions throughout the paper. We use different styles of lines because most printers are black and white, but we recommend that you do your calculations at home in colored pen or pencils instead; we even refer to the labels as "colors" throughout this paper.



We use the styles above when referring to indices i, $i{+}1$, $i{-}1$, and j, where j will be used unambiguously for any index which is "far away" from any other indices in the picture (in other words, when drawing a picture only involving i-colored strands, we require $|i-j|>1$, while for a picture involving both i and $i+1$ we require $j<i-1$ or $j>i+2$).

2.5. Methodology

Proposition 2.13. *Suppose one chooses the subset of the morphisms in $\mathcal{SC}_1$, including the identity morphism of each object, as well as the following morphisms:*

(1) *the generating morphism from R to B_i,*

(2) *some isomorphisms that yield the Hecke algebra relations, as well as the respective projections to and inclusions from each summand in* (2.10) *and* (2.12)*,*

(3) *the unit and counit of adjunction that make B_i into a self-biadjoint bimodule.*

Consider $\mathcal{C}$ the subcategory generated monoidally over the left action of R by these morphisms, that is, it includes left R-linear combinations, compositions, and tensors of all its morphisms, then $\mathcal{C}$ is a full subcategory, and thus it is actually $\mathcal{SC}_1$.

Proof. For any objects M, N in $\mathcal{SC}_1$, there is an inclusion $\mathrm{Hom}_{\mathcal{C}}(M,N) \subset \mathrm{Hom}_{\mathcal{SC}_1}(M,N)$ of graded left R-modules (since it is clearly an inclusion of R-modules, and all generating morphisms are homogeneous). One can define grk for any graded left R-module M by choosing generators of $M/(R^+M)$, where R^+ is the ideal of positively graded elements, and it is a simple argument that a submodule of a free graded R-module with the same graded rank is in fact the entire module. So we need only show that Hom spaces in $\mathcal{C}$ have the same graded rank.

We can define a semilinear form on the free $\mathbb{Z}[v, v^{-1}]$-algebra generated by b_i by the formula $(b_{\underline{\mathbf{i}}}, b_{\underline{\mathbf{j}}}) = \operatorname{grk} \operatorname{Hom}_{\mathcal{C}}(B_{\underline{\mathbf{i}}}, B_{\underline{\mathbf{j}}})$. The existence of isomorphisms and projection maps will give us the direct sum decompositions (2.10)–(2.12) in $\mathcal{C}$, with the resulting implications for Hom spaces. Therefore the Hecke algebra relation (2.2) are in the kernel of this semilinear form, so it descends to a form on the Hecke algebra. Each b_i will be self-adjoint. When $\underline{\mathbf{i}}$ is increasing, $\operatorname{Hom}_{\mathcal{C}}(R, B_{\underline{\mathbf{i}}})$ contains the generator of the free rank one R-module $\operatorname{Hom}_{\mathcal{SC}_1}(R, B_{\underline{\mathbf{i}}})$, since that generator is the tensor of the generating morphisms from R to various B_i (see the proof of Claim 2.9). Hence it is in fact the entire module, so $\operatorname{Hom}_{\mathcal{C}}(R, B_{\underline{\mathbf{i}}}) \cong R\{d\}$, and $(1, b_{\underline{\mathbf{i}}}) = v^d$.

By unicity, this inner product agrees with our earlier inner product on the Hecke algebra. In particular, the graded ranks agree, and the inclusion is full. □

Below we will construct a category $\mathfrak{D}\mathcal{C}_1$ of diagrams via generators and local relations, where the Hom spaces will be graded R-bimodules. We will construct a functor F_1 from $\mathfrak{D}\mathcal{C}_1$ to $\mathcal{SC}_1$, showing that our diagrams give graphical presentation of morphisms in $\mathcal{SC}_1$. The morphisms in the image of F_1 will include all the morphisms enumerated in Proposition 2.13, hence the functor will be full. Calculating the Hom spaces in $\mathfrak{D}\mathcal{C}_1$ between certain objects (corresponding to R, $B_{\underline{\mathbf{i}}}$ for $\underline{\mathbf{i}}$ increasing), we may use a similar argument to the above proposition to show that they are free R-modules of the same graded rank as the Hom spaces in $\mathcal{SC}_1$, then the functor F_1 will be faithful, and an equivalence of categories. This describes $\mathcal{SC}_1$ in terms of generators and relations.

Let $\mathfrak{D}\mathcal{C}_2$ be the category whose objects are finite direct sums of formal grading shifts of objects in $\mathfrak{D}\mathcal{C}_1$, but whose morphisms only include degree 0 maps. Finally, let $\mathfrak{D}\mathcal{C} = \operatorname{Kar}(\mathfrak{D}\mathcal{C}_2)$ be the Karoubi envelope of $\mathfrak{D}\mathcal{C}_2$. The functor F_1 lifts to functors F_2 and F, as in the picture below, with all three horizontal arrows being equivalences of categories.

$$\begin{tikzcd}
\mathfrak{D}\mathcal{C}_1 \arrow[r, "\mathcal{F}_1"] \arrow[d] & \mathcal{SC}_1 \arrow[d, "\text{Grading shifts and direct sums}"] \\
\mathfrak{D}\mathcal{C}_2 \arrow[r, "\mathcal{F}_2"] \arrow[d] & \mathcal{SC}_2 \arrow[d, "\text{Karoubi envelope}"] \\
\mathfrak{D}\mathcal{C} \arrow[r, "\mathcal{F}"] & \mathcal{SC}
\end{tikzcd} \tag{2.24}$$

We will define the category $\mathfrak{D}\mathcal{C}_1$ originally without reference to isotopy, in order to make the definition of the functor F_1 entirely straightforward, using the standard rules for diagrammatics for bimodules. The category would be entirely unchanged if one used different pictures to represent each morphism. However, when the "correct" pictures are chosen for the generators, then every morphism can actually be viewed as a *planar graph*, and moreover two embedded graphs linked by isotopy represent the same morphism. One could very well define $\mathfrak{D}\mathcal{C}_1$ originally using graphs, but this would obscure the definition of F_1.

The most difficult part of the proof will be showing the faithfulness of F_1, which involves a calculation of certain Hom spaces in the diagrammatic category. This calculation will be made possible by the planar graphs interpretation of $\mathfrak{D}\mathcal{C}_1$, wherein some relatively simple graph theory can be applied to simplify pictures.

3. Definition of $\mathfrak{D}\mathcal{C}$

This section contains a piecemeal definition of $\mathfrak{D}\mathcal{C}$ and $\mathfrak{D}\mathcal{C}_1$. For pedagogical reasons, we prefer to provide commentary as we go, instead of defining the category all at once (in fact, some relations do not make sense without the commentary). We also provide some redundant relations in the first pass, because they help make the category more intuitive. However, we repeat the definition all in one place in Section 3.4, without redundant relations, where we also explicitly define the functor F_1.

3.1. The Category $\mathfrak{D}\mathcal{C}_1$: Zero Colors and One Color

This section and the next several will hold the definition of the category $\mathfrak{D}\mathcal{C}_1$, which will be a $\Bbbk$-linear additive monoidal category, with $\mathbb{Z}$-graded Hom spaces. Shortly it will become clear that Hom spaces are actually graded R-bimodules. It is generated monoidally by n objects $i = 1, \dots, n$, whose tensor products will be denoted $\underline{\mathbf{i}} = i_1 \cdots i_d$.

Morphisms will be given by (linear combinations of) diagrams inside the strip $\mathbb{R} \times [0, 1]$, constructed out of lines colored by an index i, and certain other planar diagrams, modulo local relations. The intersection of the diagram with $\mathbb{R} \times \{0\}$, called the *lower boundary*, will be a sequence $\underline{\mathbf{i}}$ of colored endpoints, the source of the map, and the *upper boundary* $\underline{\mathbf{j}}$ will be the target. A vertical line colored i represents the identity map from i to i. The monoidal structure consists of placing diagrams side by side, and composition consists of placing diagrams one above the other, in the standard fashion for diagrammatic categories.

We present the generators and relations in an order based on the number of colors they use. The one-color generators and relations will be sufficient to describe the category for $n = 1$, the two-color ones for $n = 2$, and the three-color ones for the general case. The set of all relations is invariant under all color changes that preserve adjacency, so we only display each generator for a single color i, using the conventions described in Section 2.4. However, the generator exists for each index i.

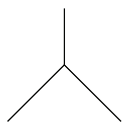
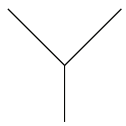
All the relations we will give are homogeneous with respect to the grading on generators stated.

The first class of generators, which use no colors, are the following endomorphisms of the monoidal identity $\emptyset$:

$$\boxed{i}$$

There is one such generator for each $i = 1, \dots, n+1$. It is a map of degree 2, which we call *multiplication by* x_i. After we apply the functor F_1, this will actually correspond to the endomorphism of R given by multiplication by x_i. Together, these generators are called *boxes*. A morphism from $\emptyset$ to $\emptyset$ consisting of a sum of disjoint unions of boxes will be called a *polynomial*. Since the composition of multiplication by x_i and multiplication by x_j is multiplication by $x_i x_j$, such a sum of products of boxes will obviously correspond under the functor to multiplication by an element $f \in R$. As a shorthand we draw such a morphism as a box with the corresponding element f inside. As a map from $\emptyset$ to $\emptyset$, and thus a closed diagram, a polynomial may be placed in any region of another diagram. Placing boxes in the rightmost and leftmost regions of a diagram will define the R-bimodule structure on Hom spaces in $\mathfrak{D}\mathcal{C}_1$.

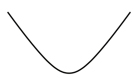
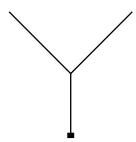
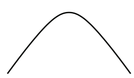
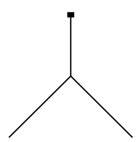
The generating morphisms which use only one color are

Symbol	Degree	Name
	1	EndDot
	1	StartDot
	–1	Merge
	–1	Split

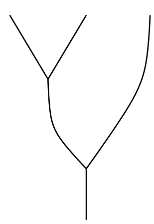
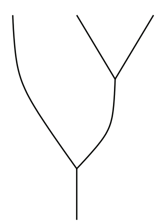
For the beginner, the maps are, respectively, a map from i to $\emptyset$, a map from $\emptyset$ to i, a map from ii to i, a map from i to ii.

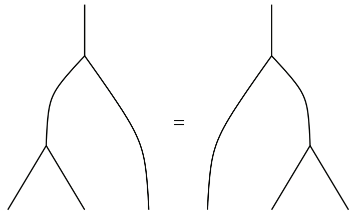
Remember, there is one such set of generators for each color i. We give these maps names, but the names are temporary. Once we explore the meaning of isotopy invariance, we will stop distinguishing between Merge and Split, and call them both *trivalent vertices*. Similarly we will stop distinguishing between StartDot and EndDot, and call them both *dots*.

We also use a shorthand for the following compositions:

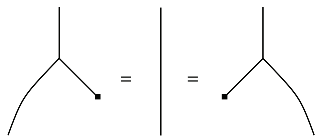
	Symbol	Degree	Name
	$\stackrel{\text{def}}{=}$ 	0	Cup
	$\stackrel{\text{def}}{=}$ 	0	Cap

We now list a series of relations using only one color, the *one-color relations*, dividing them into several types of relations for ease of reference. The first set we refer to as the *Frobenius relations*, since they imply that i is a Frobenius object in $\mathfrak{D}\mathcal{C}_1$ (see [48, 49] for more on Frobenius algebras). Once we define the functor F_1, this will imply that B_i is a Frobenius object in $\mathcal{SC}_1$. Remember that the cups and caps appearing below can actually be rewritten in terms of the generators.

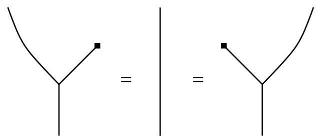
 =  (3.1)



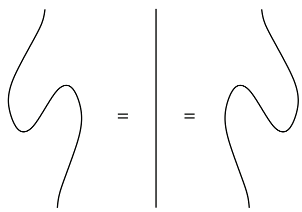
(3.2)



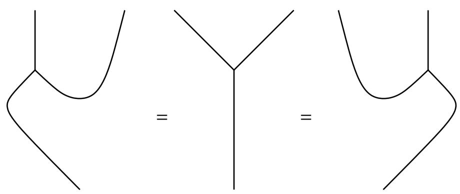
(3.3)



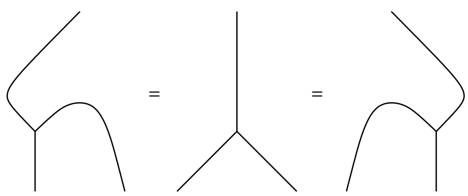
(3.4)



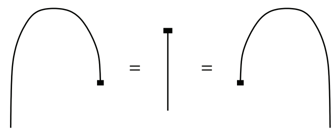
(3.5)



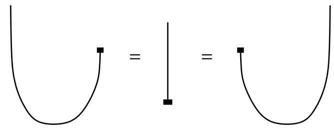
(3.6)



(3.7)



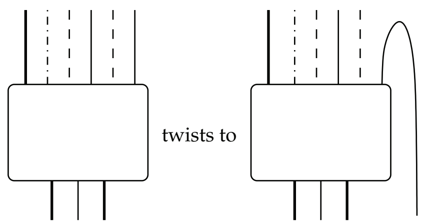
(3.8)



(3.9)

For quick reference, we refer to these relations by their Frobenius algebra names. The first two are the associativity of Merge and the coassociativity of Split. The next two are the unit and counit relations. Relation (3.5) is the biadjunction relation, and the final four are cyclicity relations.

Remark 3.1. For readers not well versed in cyclicity properties and their implications towards isotopy invariance, let us quickly discuss the topic, using the easily visualized notion of a *twist*. Given a morphism, one can twist it by taking a line which goes to the upper boundary and adding a cap, letting the line go to the other boundary instead. An example is given below,



One can also twist a downward line back up, or twist lines on the left as well. Two morphisms are twists of each other if they are related by a series of these simple twists, using cups and caps on the right and left side. For instance, relations (3.6) and (3.7) state that the Merge is a simple twist of the Split, twisting on the left or right. If one applies the same twist to every term in a relation, one gets a twist of that relation. For instance, relation (3.4) is actually a twist of the definition of the cup.

Because of biadjointness (3.5), twisting a line down and then back up will do nothing to the morphism. Once biadjointness is shown, all twists of a relation are equivalent, because twisting in the reverse direction we get the original relation back. When a morphism has a total of m inputs and outputs, twisting a single strand will often be referred to as rotation by $180/m$ degrees.

The above relations imply that twisting any of the above generators by 360 degrees will do nothing. A morphism is said to be *cyclic* with respect to a fixed set of adjunctions (i.e., cups and caps) if 360 rotation does not change the morphism. Cyclicity is useful because of the following proposition.

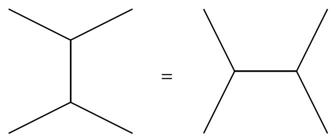
Proposition 3.2. *Fix adjunctions of each object, which are drawn as caps and cups. If every generating morphism in a diagram is cyclic with respect to those adjunctions, then so is the entire diagram, and the morphism represented by that diagram is invariant under isotopy of the diagram.*

For more on diagrammatics of biadjointness and the cyclicity property we refer the reader to [9, 48, 50–52].

Merge and Split are 60 rotations of each other, and each is invariant under 120 degree rotation, so we may represent them isotopy-unambiguously with pictures that satisfy the same properties. A similar statement holds for StartDot and EndDot. We will refer to these morphisms as dots and trivalent vertices from now on, because these terms encapsulate the picture up to rotation.

Remark 3.3. Henceforth, we can take more liberties in our drawings. We can draw a horizontal line colored i, and even though this can not be constructed using our generators, it is isotopy equivalent to a cup or cap which can be so constructed. We can allow a diagram to have a boundary not just on the top or bottom, but also on the side. While this does not represent a morphism in our category, the line running to the side boundary can be twisted either up or down to represent a genuine morphism. A relation drawn using diagrams with side boundaries does unambiguously give a relation in $\mathfrak{D}\mathcal{C}_1$.

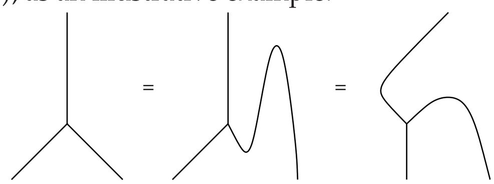
Associativity and coassociativity are twists of each other. This relation is written in a rotation-invariant form below, and will be crucial in the sequel. We refer to this relation, which permits one to "slide" one trivalent vertex over another, as *one-color associativity*.



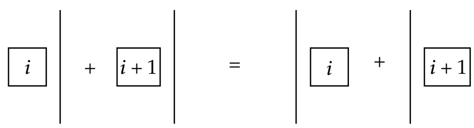
(3.10)

We refer to either picture above as an "H". The horizontal line in the right picture is exactly such a liberty as in Remark 3.3.

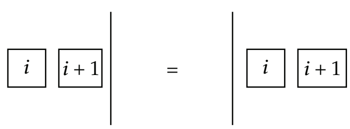
Note that relations (3.1), (3.5), (3.6), and (3.8) are sufficient to imply the other Frobenius relations, because of the remarks about twisting made above. Here is the proof of half of (3.7) using (3.6), as an illustrative example.



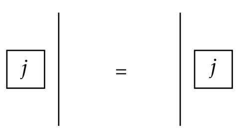
The next set of relations are known as *polynomial slides*, which have obvious analogies in the definitions of the modules B_i.



(3.11)



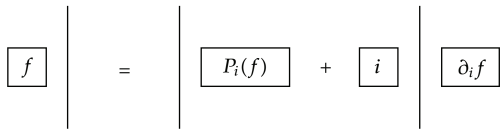
(3.12)



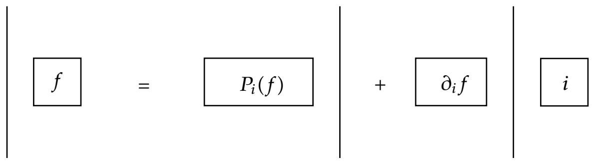
(3.13)

The j appearing in the box in the last relation can be any index not equal to i or $i+1$. Together, these relations imply precisely that any polynomial which is invariant under s_i can be slid across a line colored i, since R^i is generated by $x_i + x_{i+1}$, $x_i x_{i+1}$, and x_j for $j \neq i, i+1$. Therefore, for an arbitrary polynomial f, we have the following immediate consequence (see Remark 2.6).

Proposition 3.4. *One may force a polynomial to the other side of a line, leaving at most x_i behind, as follows:*



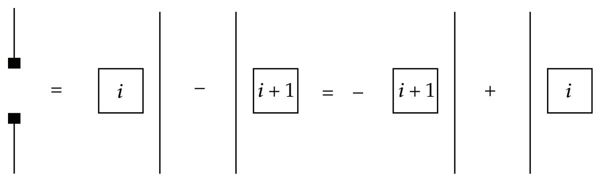
(3.14)



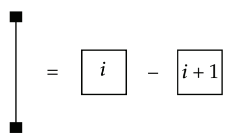
(3.15)

Proof. This is proven in the same way as (2.7). □

Now these are the final one-color relations. First, the *dot relations*

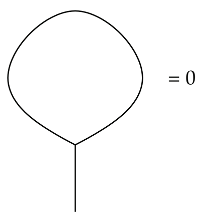


(3.16)



(3.17)

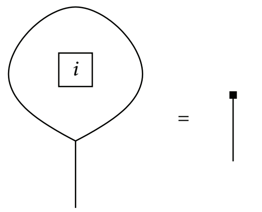
The second equality in (3.16) is just the relation (3.11). Now, is the *needle relation*:



(3.18)

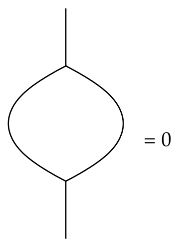
It is important to realize that such a relation does *not* apply if there is anything inside the eye of the needle, as can be seen in the following examples.

Example 3.5. Combining these relations, we have a number of simple but important consequences, which we leave as easy exercises to get the reader used to the diagrammatic calculus



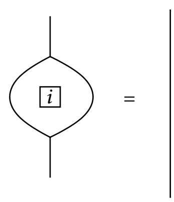
(3.19)

Use the first dot relation, then the needle relation and the unit relation



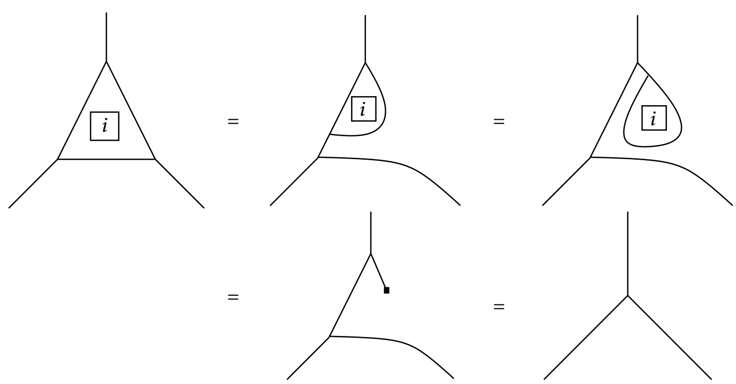
(3.20)

Use the needle relation and associativity



(3.21)

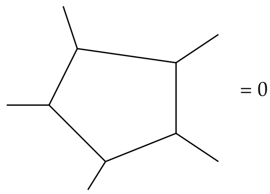
As above, with the unit relation



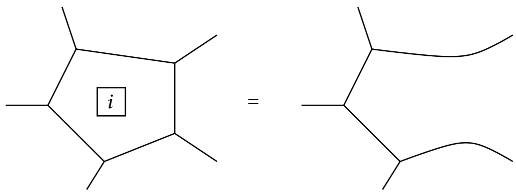
This is effectively the same example again, with more uses of associativity.

As the examples demonstrate, and the following proposition proves, we may remove cycles of this nice form from a one-color graph.

Proposition 3.6. *The following relations hold:*

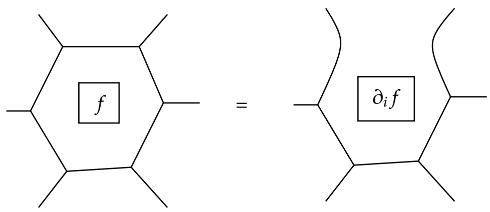


(3.22)



(3.23)

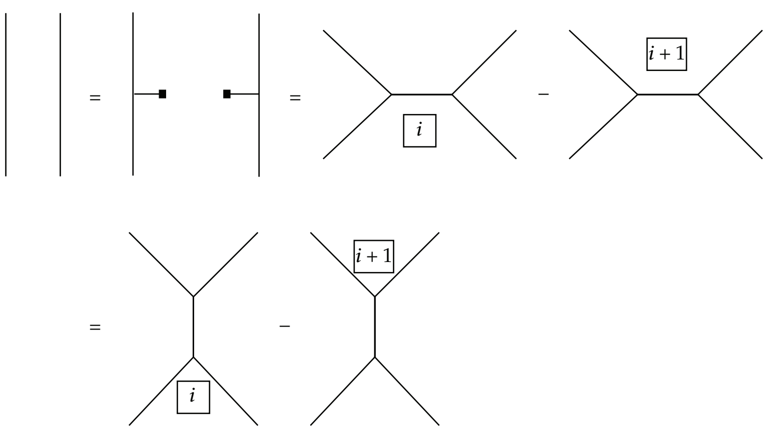
More generally, for any polynomial f, one has



(3.24)

Proof. This is a simple consequence of (3.4), along with the needle, associativity, and unit relations. □

There is another relation which is equivalent (given the others) to the first equality in (3.16)



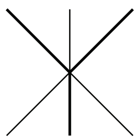
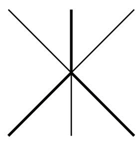
(3.25)

This relation quickly leads to the decomposition $B_i \otimes B_i = B_i\{1\} \oplus B_i\{-1\}$, see Section 4.5.

For a single color and two variables x_1, x_2, the category above, modulo the relation $x_1 = x_2$, is equivalent to the category considered by Libedinsky [13] in the case of a single label r. Morphisms given by dot, Merge, Split, and Cap correspond to morphisms $\widehat{\varepsilon}_r$, $\widehat{m}_r$, $\widehat{p}_r$, $\widehat{j}_r$, and $\widehat{\alpha}_r$ in [13, section 2.4]. Planar graphical notation, of paramount importance to us, is implicit in [13]. From here on, we diverge from Libedinsky's work, by generalizing to the case of the Weyl group S_{n+1}, while Libedinsky [13] investigates the right-angled case.

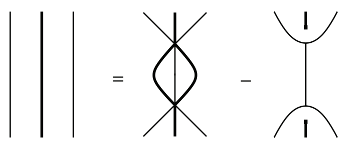
3.2. The Category $\mathfrak{D}\mathcal{C}_1$: Adjacent Colors

We now add some generators which mix adjacent colors, which we call *6-valent vertices*. Remember that the thick lines represent $i+1$, and the thin lines represent i,

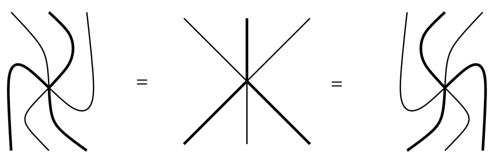
Symbol	Degree	Symbol	Degree
	0		0

For the beginner, these maps are, respectively, a map from $i(i+1)i$ to $(i+1)i(i+1)$, a map from $(i+1)i(i+1)$ to $i(i+1)i$.

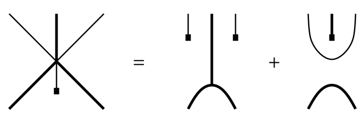
Below are the relations which deal with our new generators. In addition to the relations below, we also impose the same relations with the colors switched. The two color variants in general do not imply each other. However, it is better to think of the two colors as being arbitrary adjacent colors, rather than one being i and the other $i+1$; then one views these relations as generic for adjacent colors



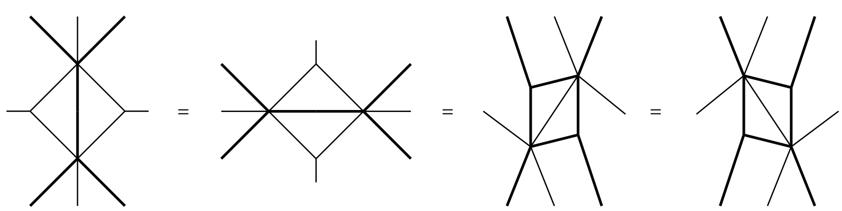
(3.26)



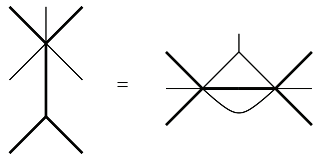
(3.27)



(3.28)



(3.29)



(3.30)

It will be shown in Section 4.5 that the first relation is related to the isomorphism $(B_i \otimes B_{i+1} \otimes B_i) \oplus B_{i+1} \cong (B_{i+1} \otimes B_i \otimes B_{i+1}) \oplus B_i$.

The second relation shows that the 6-valent vertex is cyclic, that drawing it as a 6-valent vertex is unambiguous, and that isotopy classes of diagrams built out of our local generators will still unambiguously designate a morphism. See Remark 3.1 for details. Because of this, we have used the liberties of Remark 3.3 when writing the last two relations. Note that (3.27) does in fact imply the color-switched version of that same relation, using (3.5).

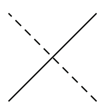
The relation (3.29) contains a number of equalities, and it is clear that the last equality is merely a rotation of the color switch of the first equality. In fact, there are numerous redundancies amongst (3.29) and (3.30). It is a worthwhile exercise for the reader at this point to check the following statement.

Example 3.7. Assume the relation (3.28) and those before it, then any pair of equalities from (3.29) will imply both color variants of (3.30) as well as the remainder of the equalities from (3.29). Hint: adding a dot to the relation (3.29) allows one to recover (3.30), while the latter may be applied twice within the former.

An important feature to notice is that the 6-valent vertex can be visualized as two trivalent vertices, one of each color, that overlap. If one takes a graph constructed out of dots, trivalent vertices, and 6-valent vertices (our generators so far), then the subgraph formed by all edges of a specific color i will have only univalent and trivalent vertices. We use the term *two-color (overlap) associativity for* $i+1$ to refer to the transformation performed by either (3.30) or the first equality of (3.29), because when viewed as an operation on the "thick"-colored graph, these operations mimic one-color associativity (3.10). Note that, under the same transformations, the "thin"-colored graph (labelled i) is transformed in a different way. However, the color-switched relations will give two-color associativity for i instead.

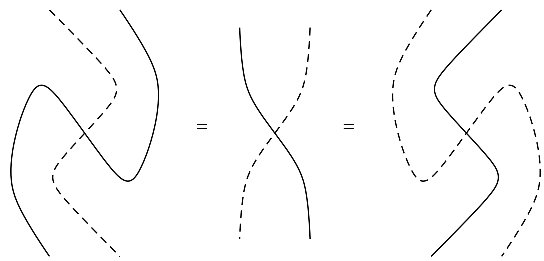
3.3. The Category $\mathfrak{D}\mathcal{C}_1$: Distant Colors

Fix j, an index which is not adjacent to i. In pictures involving both i and $i+1$, we also assume j is not adjacent to $i+1$. Remember that j is represented by a dashed line. This new generator is called a *4-valent vertex*, or a *crossing*.

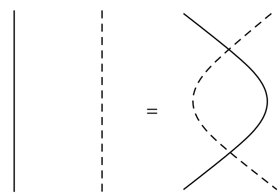
Symbol	Degree
	0

Note that this definition also covers the same picture with the colors reversed. The colors i and j can be switched freely since the only requirement was that they were distant from each other.

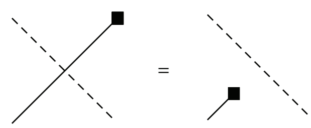
Now, for relations involving the new generator.



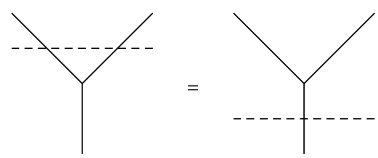
(3.31)



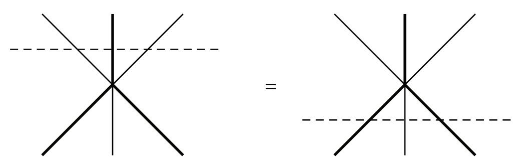
(3.32)



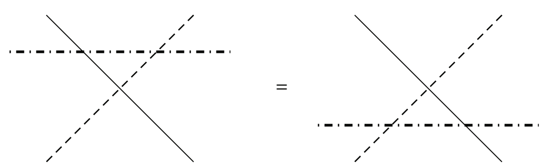
(3.33)



(3.34)



(3.35)



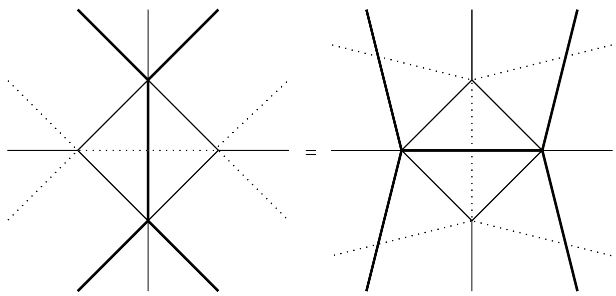
(3.36)

Relation (3.35) holds when you switch i and $i+1$, but one color variant will follow quickly from the other by twisting and applying the first relation. In the final relation, the new color represents an index k which is distant from both i and j. We will refer to (3.32) and (3.36) as the R2 and R3 moves, respectively, because of the obvious analogy to knot theory. The R2 relation is essentially the isomorphism $B_i \otimes B_j \cong B_j \otimes B_i$, see Section 4.5.

The same statements about cyclicity and drawing diagrams with sideways boundaries apply from before (see Remarks 3.1 and 3.3). Once again, the 4-valent vertices are drawn so that morphisms are isotopy invariant.

The relations (3.32)–(3.36) imply that a j-colored strand can just be pulled underneath any morphism only using colors distant from j, since it can be pulled under any generating morphism, whether it be a line, a dot, a trivalent vertex, or a 6-valent vertex. In fact, thanks to (3.4), the R2 move follows from (3.33) and (3.34).

We have now listed all the generators of our subcategory: trivalent, 4-valent, and 6-valent vertices, and dots. There is one final relation, coming from the fact that $i+1$ and $i-1$ may not interact, but they do jointly interact with i. The final relation will be called *three-color (overlap) associativity for* $i \pm 1$:



(3.37)

In the above diagram, dotted lines carry label $i-1$, thick lines $i+1$ and thin solid lines i. Rotating this relation by 90 degrees, we get the same relation except with $i+1$ and $i-1$ switched, so that only color variant is needed to imply both. The "thick"-colored graph undergoes the associativity transformation. The same is true (symmetrically) with the "dotted"-colored graph.

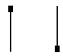
This concludes the definition of the category $\mathfrak{D}\mathcal{C}_1$.

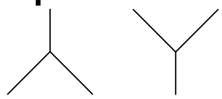
3.4. The Complete Definition and the Functor $\mathcal{F}_1$

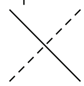
In order to put everything in one place with no redundancy, let us define the category again.

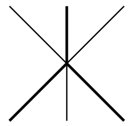
Definition 3.8. The category $\mathfrak{D}\mathcal{C}_1$ has objects given by sequences $\underline{\mathbf{i}}$ of indices in $\{1,\ldots,n\}$, with a monoidal structure given by concatenation. Fix two sequences $\underline{\mathbf{i}}$ and $\underline{\mathbf{j}}$. Consider the set of all diagrams in $\mathbb{R} \times [0,1]$, constructed out of vertical lines colored by indices, and out of the generating pictures below, such that the intersection of the diagram with $\mathbb{R} \times \{0\}$ is the sequence of points $\underline{i}$, and the intersection with $\mathbb{R} \times \{1\}$ is $\underline{\mathbf{j}}$. This set is graded, where the generators have the degree indicated, then the space $\mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1}(\underline{\mathbf{i}},\underline{\mathbf{j}})$ is defined to be the $\Bbbk$-linear span of this set of diagrams, modulo the homogeneous local relations below.

Generators:

(1) For each color, pictures of degree 1: 

(2) For each color, pictures of degree −1: 

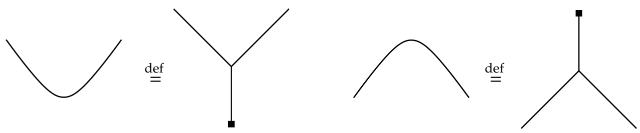
(3) For each pair of distant colors, a picture of degree 0: 

(4) For each pair of adjacent colors, a picture of degree 0: 

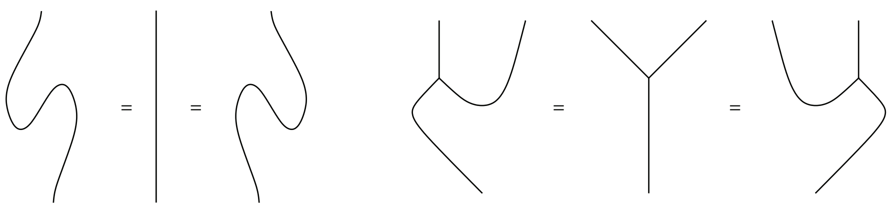
(5) For each $i \in \{1, \ldots, n+1\}$, a picture of degree 2: $\boxed{i}$

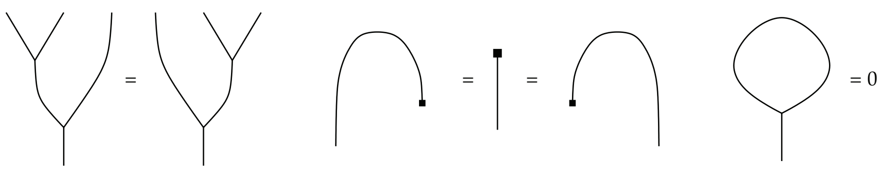
Relations:

Some relations are drawn using the liberties of Remark 3.3. We also use the definition of the cap and cup,

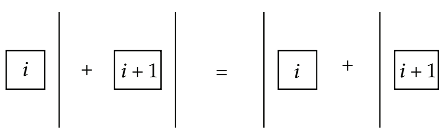


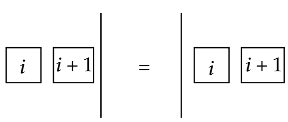
For each color,

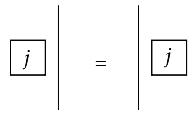


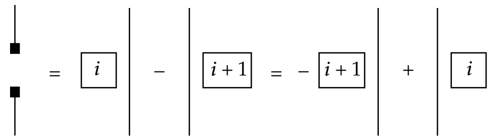


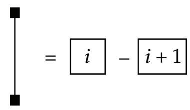
Here, i is the color of the line, and j is a color $\neq i, i+1$,



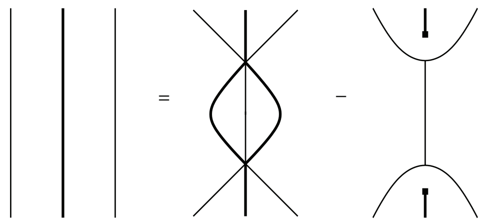


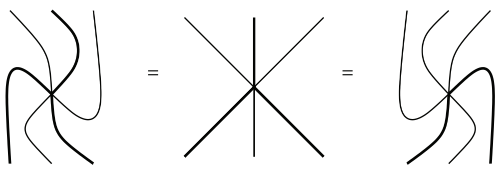


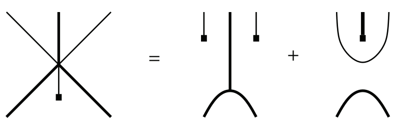


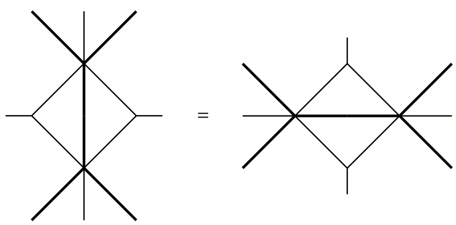


For any two adjacent colors,

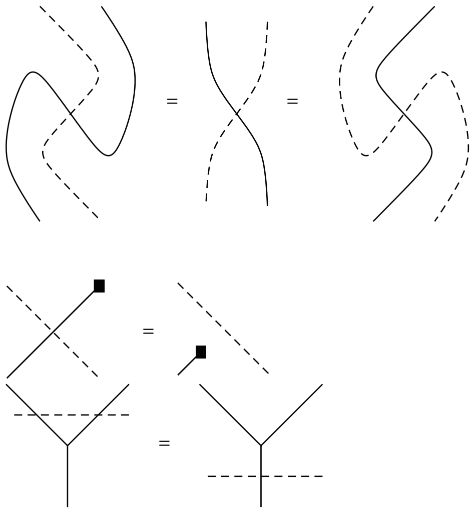




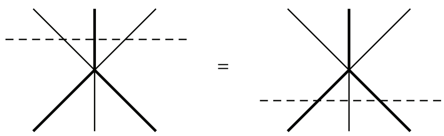




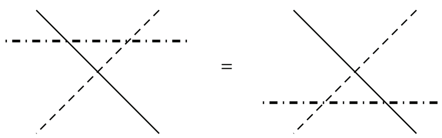
For any two distant colors,



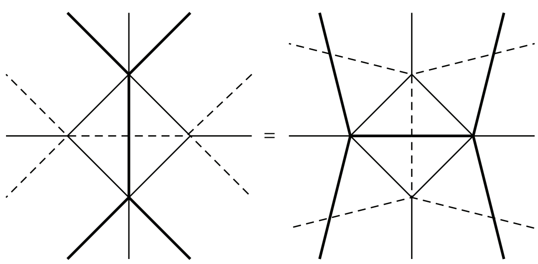
For two adjacent colors and a third, distant to both,



For three mutually distant colors,



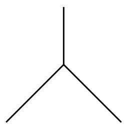
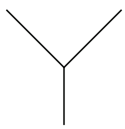
For three colors with the same adjacency as $\{1, 2, 3\}$,



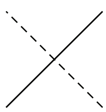
Definition 3.9. Let F_1 be the functor from $\mathfrak{D}\mathcal{C}_1$ to $\mathcal{SC}_1$ specified as follows. On objects, $F_1(\underline{\mathbf{i}}) = B_{\underline{\mathbf{i}}}$. We define the functor on generating morphisms and extend it monoidally to all morphisms.

In doing so, we *always* use the isomorphism (2.6) to identify $B_{\underline{\mathrm{i}}}$ with the R-bimodule spanned by a choice of $d(\underline{\mathrm{i}}) + 1$ polynomials. If one thinks of $B_{\underline{\mathrm{i}}}$ diagrammatically as d vertical lines, then a spanning element of $B_{\underline{\mathrm{i}}}$ is a choice of polynomial for each empty region delineated by the lines (and polynomials with the appropriate symmetry may slide across the lines). We write the map explicitly for a general element when it is easy enough to do so, or we write it for a spanning set as an R-bimodule (see Remark 2.7).

For a line colored i,

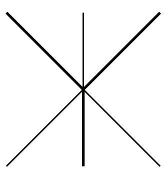
Symbol	$\mathcal{F}_1$
	$f \otimes g \mapsto fg$
	$1 \mapsto x_i \otimes 1 - 1 \otimes x_{i+1}$
	$f \otimes 1 \otimes g \mapsto 0$ $\quad$ $f \otimes x_i \otimes g \mapsto f \otimes g$
	$f \otimes g \mapsto f \otimes 1 \otimes g$

For lines colored i and j distant,

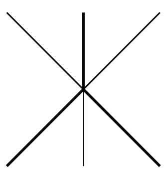


$f \otimes 1 \otimes g \mapsto f \otimes 1 \otimes g$

For a thin line colored i and a thick line colored $i+1$,



$1 \otimes 1 \otimes 1 \otimes 1 \mapsto 1 \otimes 1 \otimes 1 \otimes 1$ $\quad$ $1 \otimes x_i \otimes 1 \otimes 1 \mapsto (x_i + x_{i+1}) \otimes 1 \otimes 1 \otimes 1 - 1 \otimes 1 \otimes 1 \otimes x_i + 2$



$1 \otimes 1 \otimes 1 \otimes 1 \mapsto 1 \otimes 1 \otimes 1 \otimes 1$ $\quad$ $1 \otimes x_{i+2} \otimes 1 \otimes 1 \mapsto 1 \otimes 1 \otimes 1 \otimes (x_{i+1} + x_{i+2}) - x_i \otimes 1 \otimes 1 \otimes 1$

For any $1 \leq i \leq n+1$,

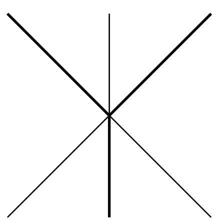
$$\boxed{i} \qquad 1 \mapsto x_i$$

Claim 3.10. The above maps are R-bimodule maps.

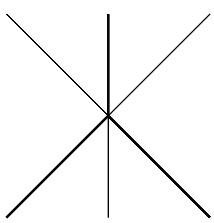
Proof. This is obviously true for EndDot, since the resulting map is no more than multiplication. StartDot is sent precisely to the generator φ_i of $\mathrm{Hom}(R, B_i)$ discussed in Section 2.3. Split and Merge have already been seen as inclusion and projection maps in the isomorphism $B_i \otimes B_i \cong B_i\{1\} \oplus B_i\{-1\}$, see Remark 2.6. For the 4-valent vertex, the only polynomials which slide all the way across $B_i \otimes B_j$ or $B_j \otimes B_i$ are in $R^{i,j}$, so that the map $f \otimes 1 \otimes g \to f \otimes 1 \otimes g$ is a bimodule map (that $f \otimes 1 \otimes g$ spans it was observed in Remark 2.7).

Only the 6-valent vertices remain to be checked. Consider the first of the two variants. The generating set $\{1 \otimes 1 \otimes 1 \otimes 1, 1 \otimes x_i \otimes 1 \otimes 1\}$ as an R-bimodule was chosen because x_i can be slid freely between the second and third slots. We have defined the R-bimodule map on generators before showing that the map is an R-bimodule map at all, which is akin to putting the cart before the horse. Let us explicitly define the map on a $\Bbbk$ spanning set by the following algorithm: given $f \otimes g \otimes h \otimes k \in B_i \otimes B_{i+1} \otimes B_i$, first we force h to the right and slide the "remainder" to the left, that is $f \otimes g \otimes h \otimes k = f \otimes g \otimes 1 \otimes kP_i(h) + f \otimes gx_i \otimes 1 \otimes k\partial_i(h)$; then we force the terms in the second slot to the left, yielding $fP_i(g) \otimes 1 \otimes 1 \otimes kP_i(h) + f\partial_i(g) \otimes x_i \otimes 1 \otimes kP_i(h) + fP_i(gx_i) \otimes 1 \otimes 1 \otimes k\partial_i(h) + f\partial_i(gx_i) \otimes x_i \otimes 1 \otimes k\partial_i(h)$. Finally, each term can be evaluated using the given definition of F_1 on generators. This gives an explicit formula for the image of $f \otimes g \otimes h \otimes k$, which we only need check is invariant under: sliding an element of R^i from f to g, or from h to k; sliding an element of R^{i+1} from g to h. Sliding elements of R^i does not pose a problem, since we defined the map by forcing h to k and g to f, which fully respects such slides. Checking invariance under slides from g to h is nontrivial. However, the bulk of the work is encapsulated in the following discussion, which is useful for calculations in general.

By adding and subtracting x_{i+2}, the image of $1 \otimes x_i \otimes 1 \otimes 1$ under the first 6-valent vertex (see above) can be written more symmetrically as $(x_i + x_{i+1} + x_{i+2})(1 \otimes 1 \otimes 1 \otimes 1) - x_{i+2} \otimes 1 \otimes 1 \otimes 1 - 1 \otimes 1 \otimes 1 \otimes x_{i+2}$. The first term is a polynomial symmetric in all the relevant variables and thus can be slid anywhere. In the other two terms, x_{i+2} can not be slid freely under a line labelled $i+1$, so it is stuck in its respective position. In contrast, $1 \otimes x_{i+1} \otimes 1 \otimes 1$ and $1 \otimes 1 \otimes x_{i+1} \otimes 1$ are not equal, since x_{i+1} can not be slid over a line labelled $i+1$, but the images of both these elements are easier to remember, and are shown below,



$$1 \otimes x_{i+1} \otimes 1 \otimes 1 \mapsto 1 \otimes 1 \otimes 1 \otimes x_{i+2} \qquad 1 \otimes 1 \otimes x_{i+1} \otimes 1 \mapsto x_{i+2} \otimes 1 \otimes 1 \otimes 1$$



$$1 \otimes x_{i+1} \otimes 1 \otimes 1 \mapsto 1 \otimes 1 \otimes 1 \otimes x_i \qquad 1 \otimes 1 \otimes x_{i+1} \otimes 1 \mapsto x_i \otimes 1 \otimes 1 \otimes 1$$

The way to remember these formulae is that the variable which cannot be slid is sent to the variable which cannot be slid, from the middle on one side to the exterior on the other. It is easy to see that these calculations were done according to the algorithm above, forcing x_{i+1} to the outside first and then evaluating on the leftover x_i.

Now, we do the consistency check for the simplest cases. We wish to show that $1 \otimes (x_{i+1} + x_{i+2}) \otimes 1 \otimes 1$ and $1 \otimes 1 \otimes (x_{i+1} + x_{i+2}) \otimes 1$ are sent to the same element by the algorithm. However, this is rather easy, for in both cases, the x_{i+2} term slides immediately to the exterior, and the x_{i+1} term is evaluated as above, so both are sent to $x_{i+2} \otimes 1 \otimes 1 \otimes 1 + 1 \otimes 1 \otimes 1 \otimes x_{i+2}$. Similarly, both $1 \otimes x_{i+1}x_{i+2} \otimes 1 \otimes 1$ and $1 \otimes 1 \otimes x_{i+1}x_{i+2} \otimes 1$ are sent to $x_{i+2} \otimes 1 \otimes 1 \otimes x_{i+2}$. The general case is not significantly more difficult than this; we leave the details to the reader. □

Proposition 3.11. *The functor F_1 is well-defined. That is, the relations of $\mathfrak{D}\mathcal{C}_1$ hold between morphisms of R-bimodules in $\mathcal{SC}_1$.*

Checking that the relations hold is a series of simple but tedious calculations that is postponed until Section 5.1. We assume this result henceforth. In addition, we note once and for all that (as one can easily check) all relations in the definition of $\mathfrak{D}\mathcal{C}_1$ are homogeneous, and F_1 preserves the degree of the generators.

Addendum 2. *It may strike the reader as unusual that the definition of the functor seems lopsided, while the definition of $\mathfrak{D}\mathcal{C}_1$ is invariant under right-left reflection, or under reversing the order of the colors $n, n-1, \ldots, 1$. For instance, F_1 applied to StartDot yields the element $x_i \otimes 1 - 1 \otimes x_{i+1}$, which is actually invariant under right-left reflection but not immediately so. Had this element been rewritten as $(x_i - x_{i+1})/2 \otimes 1 + 1 \otimes (x_i - x_{i+1})/2$, perhaps the calculations would be more natural despite having more fractions. A worse offender is the forcing rule $1 \otimes g = (g - \partial_i(g)x_i) \otimes 1 + \partial_i(g) \otimes x_i$, which should be rewritten $1 \otimes g = (g + s_i(g))/2 \otimes 1 + \partial_i(g) \otimes (x_i - x_{i+1})/2$. In general, the elements $x_i - x_{i+1}$ are more natural than x_i or x_{i+1}, coming from the reflection representation rather than the standard representation of S_{n+1} (see Remark 2.2 and Section 4.6).*

Previous versions of this paper, however, used the style above, and so we feel compelled to stick with it to maintain consistency. Also, checking that F_1 is a functor may be easier with the current notation.

4. Consequences

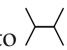
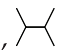
4.1. Terminology

We will spend the next few sections classifying the homomorphisms in $\mathfrak{D}\mathcal{C}_1$. For many of the results, proofs will be postponed until Section 5.2.

We will we using the fact, extensively discussed in the previous sections, that a morphism can be viewed unambiguously as an isotopy class of graphs with polynomials in the regions (or rather, a linear combination of these). Henceforth, the term *graph* only refers to colored finite graphs with boundary (embedded in the planar strip) which can be constructed out of univalent, trivalent, 4-valent, and 6-valent vertices as above. Remember that these graphs do have edges which run to the boundary, which we call *boundary lines*, and may have edges which meet neither the boundary nor any vertex, and thus must necessarily form a circle. We say a graph *has a boundary* if it has at least one boundary line. A graph divides the planar strip into regions, and there are two distinguished regions: the *lefthand* and *righthand* regions, which contain $-\infty$ and ∞, respectively.

We call a *boundary dot* any connected component of a graph which consists entirely of an edge starting at the boundary and ending in a dot. We call a *double dot* any connected component of a graph which consists entirely of an edge with a dot on both ends. Cutting an edge in a diagram and replacing it with two dots we call *breaking* the edge (see, for instance, relation (3.16)).

Given a set S of graphs and a morphism φ in $\mathfrak{D}\mathcal{C}_1$, we say that S *underlies* φ if φ can be written as a linear combination of morphisms, each of which is given by a graph $\Gamma \in S$ with polynomials in regions. We say that a graph Γ*reduces* to S if S underlies every morphism that Γ underlies. Clearly reduction is transitive, in that if Γ reduces to S, and every graph in S reduces to S', then Γ reduces to S'. Our goal will be to find a nice set of graphs to which all other graphs reduce. We will do this by finding *reduction moves*, which are local moves on graphs, sending a graph to a set of graphs to which it reduces.

Example 4.1. The relation (3.25) implies that $||$ reduces to . In other words,  underlies both the terms on the right side of (3.25). This can be applied as a local reduction move within any graph.

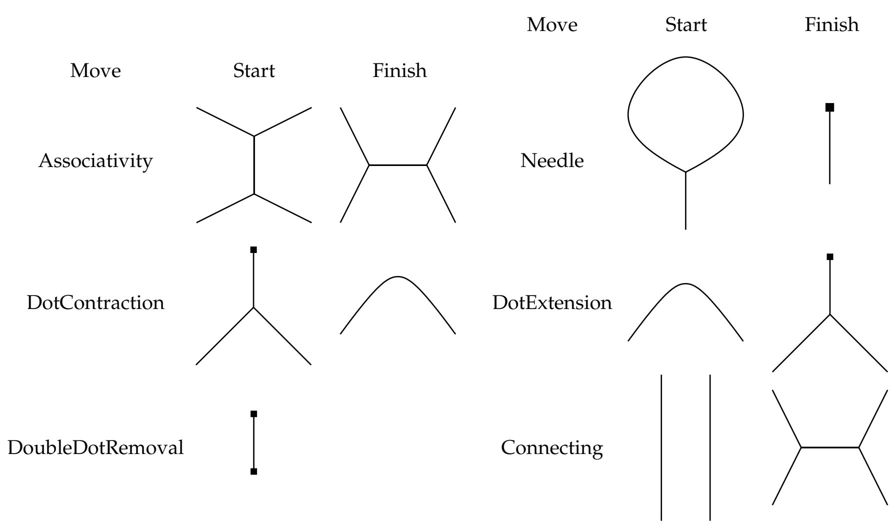
Let T be a subset of $\{1, \ldots, n\}$. The T-*graph* of a graph will be the subgraph consisting of all edges colored i for $i \in T$. Some 6-valent vertices in the original graph may become trivalent vertices in the T-graph. Similarly, some 4-valent vertices in the original graph may become 2-valent vertices in the T-graph, which we ignore, connecting the incoming edges into a single edge. The T-graph is itself a graph by our above definition. Most often we will just consider the i-graph for a single color (i.e., $T = \{i\}$). Typically, our reduction moves will be designed to simplify the i-graph for a particular i, allowing us to simplify the graph one color at a time.

Remark 4.2. The rest of this paper will have numerous calculations, but they will mostly be calculations with the underlying graphs, not keeping track of polynomials, so they do not reflect how morphisms actually behave in $\mathfrak{D}\mathcal{C}_1$. For lots of examples of computations in the graphical calculus, see [33].

4.2. One Color Reductions

In this section, we assume all graphs consist of a single color i.

Definition 4.3. Consider the following "moves", or transformations. They take a subdiagram looking like Start, and replace that subdiagram with Finish. We call these the *basic moves*.



Remember, these are moves on graphs, not graphs with polynomials. Note that the needle move, by adding a dot on the bottom, yields a reduction from the circle to a double dot. The only moves which change the connectivity of a graph are double dot removal, which deletes a connected component, and the connecting move, which has the potential to link two components into one.

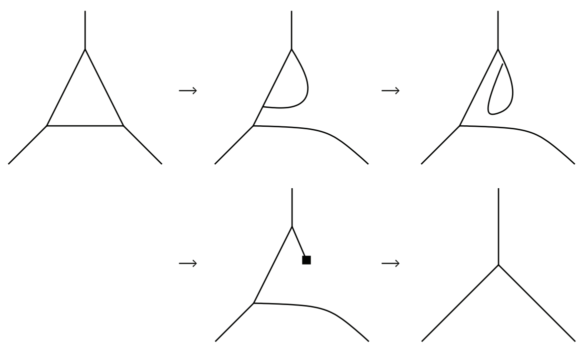
Claim 4.4. All of these moves are reduction moves in $\mathfrak{D}\mathcal{C}_1$.

Proof. The associativity move follows from (3.10). That is, even if there are polynomials in the regions of the graph, the relation (3.10) can still be applied. These polynomials, being in external regions, do not interfere with the application of relations. Similarly, dot contraction/extension follow from (3.4), dot removal follows from (3.17), and the connecting move follows from (3.25).

The needle move remains. Suppose we have an arbitrary polynomial f in the eye of the needle. We may use (3.24), generalizing (3.19), to replace the diagram with a dot accompanied by $\partial_i(f)$. □

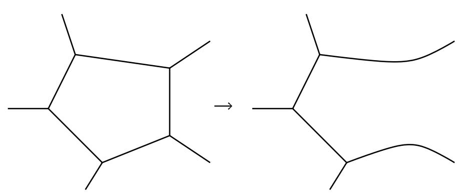
The following example of reduction should be familiar.

Example 4.5.



In order, this is done with associativity, associativity, needle, and dot contraction moves.

Example 4.6.



This is meant to indicate an arbitrary length cycle of this form, and reduction is done with associativity, needle, and dot contraction moves.

Definition 4.7. A *simple tree* T with m boundary lines is a connected one-color graph with boundary, whose form depends on m:

(1) if $m \geq 2$, then T is a trivalent tree with $m - 2$ vertices connecting all the boundary lines. Note that any two such trees are equivalent under the associativity move;

(2) if $m = 1$, then T is a single boundary dot;

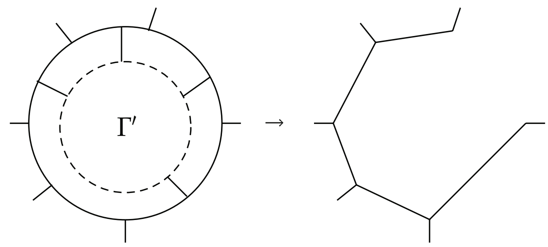
(3) if $m = 0$, then T is the empty graph.

Definition 4.8. A *cycle* in a one-color graph is either a circle or a path from a vertex to itself which does not repeat any edges. Any cycle splits the plane into two parts, the inside and outside of the cycle. A cycle is *minimal* if the inside of the cycle consists of a single region.

By counting vertices, it is clear that any connected purely trivalent graph with no cycles is a simple tree. Any graph with a cycle has a minimal cycle.

We will now give a precise inductive algorithm to reduce a graph to a disjoint union of simple trees, by reducing minimal cycles.

Proposition 4.9. *Consider a minimal cycle in a one-color graph* Γ*. Using the associativity, needle, dot removal, and dot contraction/extension moves, we may reduce a neighborhood of the cycle (including the inside region) to a simple tree (see* (4.1)*).*



(4.1)

Proposition 4.10. *Using the associativity, needle, dot removal and dot contraction/extension moves, one can reduce any one-color graph* Γ *to a disjoint union of simple trees. During this process, each component with no boundary lines will be replaced with the empty graph, and after this is done, no further connected components are created or destroyed or merged.*

We prove both propositions together, in several steps.

Proof of Proposition 4.10 for a Graph with No Cycles. Suppose there are no cycles in Γ. If there is a dot, the edge coming from that dot must connect to either the boundary, another dot, or a trivalent vertex. Our simplification algorithm is as follows.

(1) Remove any dot connected to a trivalent vertex, using dot contraction. Repeat Step 1 until no such dots remain. This does not alter the connected components.

(2) Replace any double dot with the empty set, using double dot removal. Because Step 1 did not alter the connected components, this could only be applied to double dots which arose from components which had no boundary lines.

Boundary dots are in their own connected component. Any other connected component is purely trivalent and has no cycles, so it must be a simple tree. □

Proof of Proposition 4.9, Assuming Proposition 4.10 for Graphs with No Cycles. Let γ denote a minimal cycle of Γ. Consider a neighborhood of γ in Γ, and let Γ' be the subgraph consisting of the interior of the cycle, as in (4.1). Let us call the boundary lines of Γ'*spokes*, since they run into γ from the inside, like spokes hitting the wheel of a bicycle. Now Γ' may not have any cycles, or else γ would not be a minimal cycle. Moreover, the spokes of Γ' must be in distinct connected components of Γ', or else they would create additional regions and γ would not be a minimal cycle. Our simplification algorithm is as follows:

(1) Apply Proposition 4.10 to Γ', replacing Γ' with a disjoint union of boundary dots, one for each spoke.

(2) Use dot contraction to remove all the spokes. We are now in the situation of Example 4.6.

(3) Apply associativity as in Example 4.5, reducing the length of γ by one. Repeat until γ is a needle or a circle.

(4) If γ is a needle, apply the needle move to replace it with a dot. If associativity moves were performed in Step 3, use dot contraction to contract this dot into one of the trivalent vertices, as in Example 4.5.

(5) If γ is a circle, apply dot extension to replace it with a needle attached to a dot, then apply needle reduction to obtain a double dot, and double dot removal to obtain the empty graph.

It is a simple observation that the result of this procedure is a simple tree, and that the only alteration of connected components which occurred was the removal of components which had no boundary lines to begin with. □

Proof of Proposition 4.10 in the General Case. Suppose we have an arbitrary graph Γ. Our simplification algorithm is as follows:

(1) If Γ has a cycle, apply Proposition 4.9 to replace its neighborhood with a simple tree. Repeat this process until Γ has no cycles.

(2) Apply the procedure for the case of no cycles above.

Note that Step 1 will terminate, which can be shown by induction on the number of internal regions (regions which do not meet the boundary of the planar strip). Each application of Proposition 4.9 reduces the number of internal regions by 1. □

Corollary 4.11. *Using the connecting move in addition, we can reduce the graph to a single simple tree.*

Proof. It is an easy observation that when one uses the connecting move on a simple tree with m boundary lines and a simple tree with m' boundary lines, one gets a simple tree with $m+m'$ boundary lines, after possibly removing extraneous dots if either m or m' equals 1. □

Remark 4.12. There are two useful sets of one-color graphs with m boundary lines, to which all others reduce. The first set just contains the simple tree with m boundary lines, and the latter is the collection of all disjoint unions of simple trees whose number of boundary lines add up to m. The former is useful because there is a single graph, so we have fewer cases to deal with. The latter is useful because it does not require the connecting move.

More importantly, these sets behave differently when we introduce polynomials into the equation. Let us assume that all m boundary lines are on the top boundary, so that we are looking at a morphism in $\mathrm{Hom}(\emptyset, \underline{\mathbf{i}})$ where $\underline{\mathbf{i}}$ is *iiiiiiii* (m times). The following statements will not be used in this paper, and can be more easily proven after the calculation of Hom spaces.

Claim 4.13. Consider diagrams which are a simple tree, with an arbitrary monomial in the lefthand region, and either 1 or x_i in each other region. This is a basis for $\mathrm{Hom}(\emptyset, \underline{\mathbf{i}})$ over $\Bbbk$.

Claim 4.14. Consider diagrams which are disjoint unions of simple trees, with an arbitrary monomial in the lefthand region, and no other polynomials. These are a spanning set for $\mathrm{Hom}(\emptyset, \underline{\mathbf{i}})$ over $\Bbbk$.

The second claim is easy to see, given Proposition 4.10. Given a disjoint union of simple trees, with arbitrary polynomials, we may use relation (3.16) and the polynomial slides to force all the polynomials to the left, at the cost of potentially breaking some lines. This breakage is not a problem, since one can reduce again to a simple tree without adding more polynomials, using (3.10) and (3.4). We do not get a basis this way: consider the three different ways to break a line in a trivalent vertex diagram; there is a linear dependence relation between these diagrams and the trivalent vertex with a polynomial in the lefthand region.

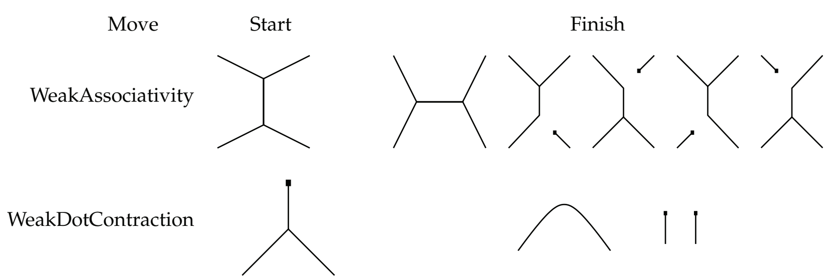
Relations (3.25) and (3.11) essentially allow us to get from the second spanning set to the first, showing that the first is at least a spanning set. That it is a basis is immediate from counting the graded dimension of the Hom space, one we prove that the dimension of Hom spaces in $\mathfrak{D}\mathcal{C}_1$ conforms with a certain semilinear form.

The connecting move is less important than the others in the proofs, and was introduced primarily to make these remarks. It can generally be ignored below.

4.3. Broken One-Color Reductions

The reductions of the previous section do apply, as stated, to any one color graph. However, we would like to apply these moves to the i-colored graph of a multicolor graph, where the moves above do not extend trivially to reduction moves in $\mathfrak{D}\mathcal{C}_1$. In this section, we quickly generalize the results of the previous section to a weaker set of moves.

Definition 4.15. Consider the following reductions for one-color graphs, which take the graph on the left and replace it with the set of graphs on the right:



We call the moves above (together with the basic moves that have no weak analog) the *broken* or *weak basic moves*.

These moves behave like the basic moves of the same name, except that they may also replace the original diagram with a broken version of itself, that is, a version with some edges broken. To distinguish the original basic moves, we may call them the *strict basic moves*.

What is important is that we have analogs of Propositions 4.9 and 4.10, despite only being able to use weak moves.

Proposition 4.16. *Consider a minimal cycle in a one-color graph* Γ*. Using the weak associativity, needle, dot removal, weak dot contraction, and dot extension moves, we may reduce a neighborhood of the cycle (including the inside region) to a disjoint union of simple trees.*

Proposition 4.17. *Using the weak associativity, needle, dot removal, weak dot contraction, and dot extension moves, one can reduce any one-color graph* Γ *to a disjoint union of simple trees.*

Of course, two distinct simple trees are no longer equivalent under the weak associativity move, but this is really irrelevant for us.

Proof. Breaking a line will never create a cycle or increase the number of trivalent vertices. Because of this, the proofs of the previous section go through almost verbatim (ignoring any statements about connected component, and occasionally replacing "a simple tree" with "a disjoint union of simple trees"). The only significant alterations that need to be made come in the proof of Proposition 4.9. In Step 4 or Step 5 one may need to remove an additional double dot. In Step 2 or Step 3, weak dot contraction and weak associativity have multiple outcomes, but each outcome that does not agree with strict dot contraction or strict associativity will have broken the cycle already, allowing us to complete the proof using the no-cycle algorithm of Proposition 4.10.

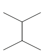
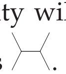
Alternatively, one could also prove these statements by induction on the number of trivalent vertices. Each weak move is equivalent to a strong move modulo diagrams with fewer trivalent vertices. The only part of the proof that ever created additional trivalent vertices was the single use of dot extension in Step 5 of the proof of Proposition 4.9. It is easy to see how Step 5 does not actually cause a problem, however, since after dot extension is applied, the needle move and double dot removal will do the trick in the same way regardless. ☐

Addendum 3. *The overall proof using weak one-color moves is slightly different than the treatment in previous versions of this paper, but it is cleaner and more straightforward.*

4.4. i-Colored Moves

Now we list the moves which allow us to simplify multicolor graphs.

Definition 4.18. Consider a graph Γ whose i-graph looks like one of the pictures in the start column of Definition 4.3. Let S be the set of all graphs whose i-graph looks like the corresponding picture in the finish column. Let W be the set of all graphs whose i-graph looks like any of the corresponding pictures in the finish column of definition weakbasicmoves. The *strict* i*-colored move* replaces Γ with the set S. The *weak* i*-colored move* replaces Γ with the set W.

For instance, strict i-colored associativity will replace any graph Γ whose i-graph is  with the set S of graphs whose i-graph is . This set S is enormous, for other colors can interfere, and the j-graph for some other j can be arbitrarily complicated. The i-colored vertices, seemingly trivalent, could come from 6-valent vertices in Γ. In general, an i-colored move will behave nicely on the i-graph, but may significantly complicate the full graph.

Proposition 4.19. *The weak* i*-colored basic moves are reduction moves in* $\mathfrak{DC}_1$*, so long as they are applied to graphs which do not contain either the color* $i-1$ *or* $i+1$*.*

A color i will be called *extremal* for a graph Γ if it appears in Γ but either $i-1$ or $i+1$ does not appear. Clearly, any nonempty graph will have an extremal color, such as the minimal color present.

The proof of this proposition is found in Section 5.2, as well as more precise details on what can be done. The power of the proposition can be seen immediately:

Corollary 4.20. *Any graph* Γ *without boundary lines can be reduced to the empty graph.*

Proof. We induct on the set of colors present in the graph Γ. If no colors are present, then Γ is the empty graph and we are done. Else, choose an extremal color i. By Proposition 4.19 we may apply the weak i-colored basic moves and use Proposition 4.17 to replace every connected component of the i-graph with a disjoint union of simple trees with no boundary lines. Since a simple tree with no boundary lines is the empty set, Γ reduces to a set of graphs which do not include the color i. By induction, Γ now reduces to the empty graph. □

One can apply a similar procedure to a graph Γ with boundary lines. Choose an extremal color i, and reduce the i-graph to a disjoint union of simple trees, then, within each region delimited by the i-graph, the colors $i+1$ and $i-1$ are now extremal, and we can reduce those. One can repeat this procedure, however, it will not produce a very simple graph in all cases. If the color i has at least 3 boundary lines, the i-graph may have trivalent vertices, and the graph Γ itself may have 6-valent vertices in their place. These 6-valent vertices will produce more $i+1$ or $i-1$ colored boundary lines inside the regions delimited by the i-graph. Nonetheless, we have the following simple case.

Corollary 4.21. *Any graph whose boundary has at most one line of each color can be reduced to a disjoint union of boundary dots.*

Proof. We know we can reduce the i-graph, for i an extremal color, to a disjoint union of simple trees. A simple tree with at most one boundary line is either the empty set or a boundary dot. Therefore the i-graph is now either the empty set or a boundary dot, depending on whether or not i appears in the boundary. The dot need not be a boundary dot in the entire graph Γ, but it can encounter only 4-valent vertices en route to the boundary. Since a dot can be slid under a 4-valent vertex by relation (3.33), we may turn the dot into a boundary dot (its own connected component). The remaining connected components form a subgraph (also viewable in the planar strip) without the color i. Induction now concludes the proof. □

4.5. $\mathcal{F}_1$ *Is Fully Faithful*

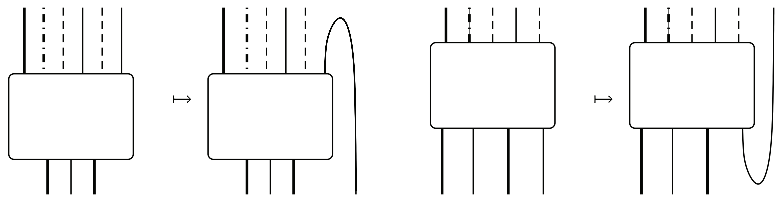
In this section, modulo the proofs of previous sections which were delayed until Section 5, we prove our main theorem.

Theorem 4.22. *The functor* F_1 *from* $\mathfrak{DC}_1$ *to* $\mathcal{SC}_1$ *is an equivalence of* $\Bbbk$*-linear monoidal categories with Hom spaces enriched in* R*-molf*$_{\mathbb{Z}}$*-*R*.*

We know F_1 is a functor by Proposition 3.11, and inspection of the objects in both categories shows immediately that it is essentially surjective. To show F_1 is full, we use Proposition 2.13, which motivates the next few statements.

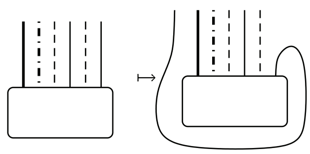
Corollary 4.23. *For any index i, the object i in $\mathfrak{D}\mathcal{C}_1$ is self-biadjoint. This means that for any sequences $\underline{\mathbf{j}}$ and $\underline{\mathbf{k}}$, there are natural isomorphisms $Hom_{\mathfrak{D}\mathcal{C}_1}(\underline{\mathbf{k}}, \underline{\mathbf{j}}i) \rightarrow Hom_{\mathfrak{D}\mathcal{C}_1}(\underline{\mathbf{k}}i, \underline{\mathbf{j}})$ and $Hom_{\mathfrak{D}\mathcal{C}_1}(\underline{\mathbf{k}}, i\underline{\mathbf{j}}) \rightarrow Hom_{\mathfrak{D}\mathcal{C}_1}(i\underline{\mathbf{k}}, \underline{\mathbf{j}})$.*

Proof. The first isomorphism and its inverse are shown below. That these maps compose to be the identity is exactly the relation (3.5)



The second isomorphism and its inverse are the left-right mirror of the maps above. ☐

Note that when $\underline{\mathbf{k}} = \emptyset$ in the corollary above, these isomorphisms combine to yield an isomorphism $\mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1}(\emptyset, \underline{\mathbf{j}}i) \rightarrow \mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1}(\emptyset, i\underline{\mathbf{j}})$, drawn as below



At this point, one could construct a semilinear product on the free algebra generated by b_i, $i = 1, \ldots, n$, via $(b_{\underline{\mathbf{i}}}, b_{\underline{\mathbf{j}}}) = \operatorname{grk} \mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1}(\underline{\mathbf{i}}, \underline{\mathbf{j}})$, and b_i would be self-adjoint. If we show that the Hecke algebra relations are in the kernel of this semilinear product then it will descend to the Hecke algebra. We have several methods by which we could do this.

(1) Look in $\mathfrak{D}\mathcal{C}_2$, where we have direct sums and grading shifts, and prove the isomorphisms (2.10)–(2.12).

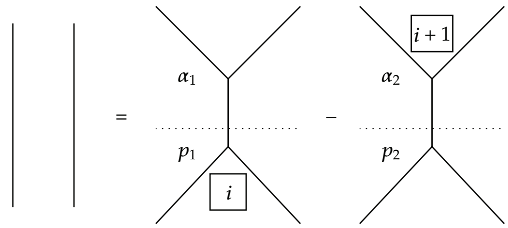
(2) Look in the Karoubi envelope $\mathfrak{D}\mathcal{C}$, find idempotents corresponding to the auxiliary modules in (2.13) and friends, and show those isomorphisms.

(3) Work entirely within $\mathfrak{D}\mathcal{C}_1$ and show the isomorphisms (2.12) only after applying the Hom functor. For instance, showing that $\mathrm{Hom}(ii, \underline{\mathbf{j}}) \cong \mathrm{Hom}(i, \underline{\mathbf{j}})\{1\} \oplus \mathrm{Hom}(i, \underline{\mathbf{j}})\{-1\}$ will be sufficient.

All these tactics are primarily the same. We illustrate the third method, although we do explore the auxiliary modules of the second method.

The relation (3.25) precisely descends to $B_i \otimes B_i \cong B_i\{1\} \oplus B_i\{-1\}$. We decompose the identity of ii into the sum of two idempotents, and obtain orthogonal projections from ii to i

of degrees 1 and −1, respectively,

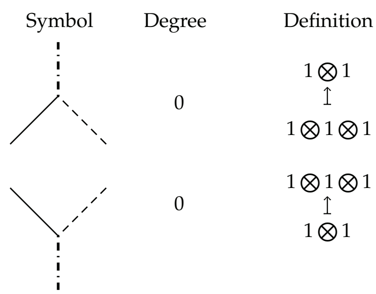


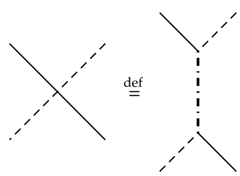
(4.2)

Stated very explicitly, we have two projection maps $p_1, p_2 : ii \longrightarrow i$, and two inclusion maps $\alpha_1, \alpha_2 : i \longrightarrow ii$, as indicated in the diagram above, which is just relation (3.25) divided in half. Include the minus sign on the right picture into the map α_2, then one can quickly check $p_1\alpha_1 = 1_i$, $p_2\alpha_2 = 1_i$, $p_1\alpha_2 = 0$, $p_2\alpha_1 = 0$, and $1_{ii} = \alpha_1 p_1 + \alpha_2 p_2$. So these must be projections to and inclusions of summands, and they are maps of the correct degree.

Now, we look at i and j which are distant. We have the isomorphism $B_i \otimes B_j \cong R\otimes_{R^{i,j}} R\{-2\} \cong B_j \otimes B_i$. It is clear that we may construct for any $\underline{\mathbf{k}}$ isomorphisms $\operatorname{Hom}(ij, \underline{\mathbf{k}}) \longrightarrow \operatorname{Hom}(ji, \underline{\mathbf{k}})$ or vice versa merely by precomposing with the appropriate 4-valent vertex, and (3.32) shows that these maps are inverses.

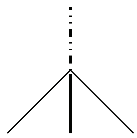
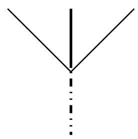
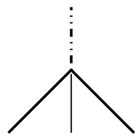
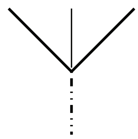
Because we can, let us discuss how we might extend our diagrammatic calculus to include additional modules from R-molf-R, like $R\otimes_{R^{i,j}} R\{-2\}$. Let us (temporarily) allow a new color of line in our pictures, call it w, and extend the functor F_1 so that it sends w to $R\otimes_{R^{i,j}} R\{-2\}$. We add new generators to our diagrammatics, and specify the image under the extended functor, as below

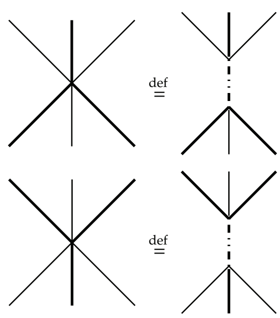




The definition of these bimodule maps, and the proof that they are in fact bimodule maps, is exactly akin to the discussion in Section 3.4. It is clear that composing these morphisms to get an endomorphism of ij will yield the identity map, and composing them to get an endomorphism of w will also yield the identity map (these would be relations in the extended diagrammatic calculus). Thus we get isomorphisms $ij \cong w \cong ji$ in $\mathfrak{D}\mathcal{C}$.

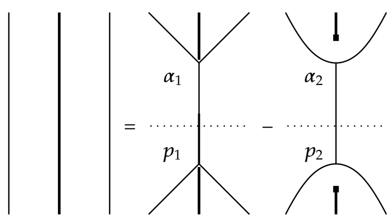
We now combine these techniques to deal with adjacent colors. We have isomorphisms $B_i \otimes B_{i+1} \otimes B_i \cong B_i \oplus (R\otimes_{R^{i,i+1}} R\{-3\})$ and $B_{i+1} \otimes B_i \otimes B_{i+1} \cong B_{i+1} \oplus (R\otimes_{R^{i,i+1}} R\{-3\})$. If we allow the new color of line, again called w, to represent the bimodule $R\otimes_{R^{i,i+1}} R\{-3\}$, then we may define the following maps:

Symbol	Degree	Definition	
	0	$\begin{array}{c} 1\otimes 1 \\ \uparrow \\ 1\otimes 1\otimes 1\otimes 1 \end{array}$	$\begin{array}{c} (x_i + x_{i+1})\otimes 1 - 1\otimes 1\otimes x_{i+2} \\ \uparrow \\ 1\otimes x_i\otimes 1\otimes 1 \end{array}$
	0	$\begin{array}{c} 1\otimes 1\otimes 1\otimes 1 \\ \uparrow \\ 1\otimes 1 \end{array}$	
	0	$\begin{array}{c} 1\otimes 1 \\ \uparrow \\ 1\otimes 1\otimes 1\otimes 1 \end{array}$	$\begin{array}{c} 1\otimes (x_{i+1} + x_{i+2}) - x_i\otimes 1 \\ \uparrow \\ 1\otimes x_{i+2}\otimes 1\otimes 1 \end{array}$
	0	$\begin{array}{c} 1\otimes 1\otimes 1\otimes 1 \\ \uparrow \\ 1\otimes 1 \end{array}$	



Again, checking that we have well-defined bimodule maps is akin to Section 3.4. Composing the two maps that go through $i(i+1)i$ and w to get an endomorphism of w will yield the identity map of w, and the same is true, respectively, of $(i+1)i(i+1)$.

Then the equation (3.26), which is a decomposition of the identity of $i(i+1)i$, actually follows from this relation in $\mathfrak{D}\mathcal{C}$



(4.3)

There is a more explicit statement to be derived from this relation, completely analogous to the two-line case, with projection and inclusion maps $p_1, \alpha_1, p_2, \alpha_2$ (include the minus sign on the right picture in α_2). Similarly we get a decomposition of $(i+1)i(i+1)$ via the same relation with the colors switched. The auxiliary module used here is in fact the indecomposable Soergel bimodule B_w where $w = s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$.

To state the result without extending the calculus to the Karoubi envelope, one may merely observe that for any $\underline{\mathbf{k}}$ there is a map $\mathrm{Hom}(i(i+1)i, \underline{\mathbf{k}}) \to \mathrm{Hom}((i+1)i(i+1), \underline{\mathbf{k}})$ and vice versa given by precomposing with the appropriate 6-valent vertex; there is another map $\mathrm{Hom}(i(i+1)i, \underline{\mathbf{k}}) \to \mathrm{Hom}(i, \underline{\mathbf{k}})$ given by precomposing with α_2, and a map backwards given by precomposing with p_2, then (3.26) exactly yields that $\mathrm{Hom}(i(i+1)i, \underline{\mathbf{k}}) \oplus \mathrm{Hom}((i+1), \underline{\mathbf{k}}) \cong \mathrm{Hom}((i+1)i(i+1), \underline{\mathbf{k}}) \oplus \mathrm{Hom}(i, \underline{\mathbf{k}})$.

Thus we have shown the following claim:

Claim 4.24. The isomorphisms (2.10) through (2.12) hold in $\mathfrak{D}\mathcal{C}_1$ after applying any Hom functor.

We have now satisfied the requirements of Proposition 2.13, which implies the fullness of F_1.

Claim 4.25. The functor F_1 is full.

Finally, our graphical reductions give us a classification of certain Hom spaces.

Corollary 4.26. *The space $\mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1}(\emptyset, \underline{\mathbf{i}})$, where $\underline{\mathbf{i}}$ is a length d increasing sequence, is a free left (or right) R-module of rank 1, generated by a homogeneous morphism of degree d.*

Proof. Let φ be the morphism consisting entirely of boundary dots. This corresponds, under the functor, to the generator of degree d discussed in Claim 2.9. By Corollary 4.21, φ viewed as a graph underlies every morphism in $\mathfrak{D}\mathcal{C}_1$. This graph has a single region, so any morphism that it underlies will be generated by φ under the action of R on the right or left, which puts a polynomial into that single region. This gives a surjective map from R to the Hom space. After applying F_1, we know that the morphisms must surject onto $\mathrm{Hom}_{\mathcal{SC}_1}(R, B_{\underline{\mathbf{i}}})$, which is a free R-module (see Claim 2.9). Therefore, the Hom space in $\mathfrak{D}\mathcal{C}_1$ must also be free. □

Corollary 4.27. *The semilinear form on H induced by Hom spaces in $\mathfrak{D}\mathcal{C}_1$ agrees with the form defined in Section 2.1, so it agrees with the form induced by $\mathcal{SC}_1$. Therefore F_1 is faithful.*

Proof. This is now immediate, from Remark 2.1. □

In conclusion, F_1 is an equivalence of categories, and Theorem 4.22 is proven.

We also get for free the following corollary, which is difficult to prove purely diagrammatically.

Corollary 4.28. *Hom spaces in $\mathfrak{D}\mathcal{C}_1$ are free as left or right R-modules.*

Remark 4.29. It is worth reiterating what is proven diagrammatically, and what is proven using the functor F_1 to Soergel's category. Diagrammatically, we can prove Proposition 4.19 and Corollary 4.21, which implies that R surjects onto $\mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1}(\emptyset, \underline{\mathbf{i}})$ for $\underline{\mathbf{i}}$ increasing. However, without the full functor F_1 and the nondiagrammatic knowledge of Hom spaces in Soergel's category, we do not know how to prove that the Hom space is free. We do not have a fully diagrammatic proof that the Hom spaces in $\mathfrak{D}\mathcal{C}_1$ are what they are, we only have a diagrammatic proof of an upper bound on the size of the Hom spaces.

4.6. The e_1 Quotient

In the remainder of this chapter, we provide some sketched statements about generalizations and variations of $\mathfrak{D}\mathcal{C}_1$. We do not use these results elsewhere in the paper, and while the proofs are only sketched, they are fairly obvious and can be fleshed out without too much work.

As described in Remark 2.2, one usually constructs Soergel bimodules with respect to the n-dimensional fundamental or geometric representation, instead of the $n+1$-dimensional standard representation. This amounts to working over the ring $R_1 = \Bbbk[x_1, x_2, \ldots, x_{n+1}]/(x_1 + x_2 + \cdots + x_{n+1})$, which is a quotient of R by the first elementary symmetric function e_1. Since e_1 is symmetric, it is in the center of the category $\mathfrak{D}\mathcal{C}_1$ (it slides freely under all lines, tensor-commuting with all morphisms), and one may easily take the quotient category, setting $e_1 = 0$, without changing any of the diagrammatics. There is a functor from this quotient category to the appropriate category of Soergel bimodules in R_1-molf-R_1, which is an equivalence of categories, so this diagrammatical quotient category also categorifies the Hecke algebra.

One advantage to passing to the quotient $e_1 = 0$ is that, after inverting a suitable integer, we may remove the boxes from our list of generators. According to relation (3.17), the double dot colored i is equal to $x_i - x_{i+1}$. Thus linear combinations of double dots of colors $i = 1, \ldots, n$ will give us the $\Bbbk$-span of $x_1 - x_2, x_2 - x_3, \ldots$ inside the space of linear polynomials (these are the simple roots). This span will not include x_i in R, which is why we require at least one box as an additional generator. However, it is easy to check that, if $n+1$ is invertible in $\Bbbk$, then x_i is in the $\Bbbk$-span after passing to the quotient R_1. As an example, when $n = 1$, the double dot is equal to $x_1 - x_2$, and passing to $x_1 + x_2 = 0$, we get $x_1 = (1/2)(x_1 - x_2)$. Thus, if $n+1$ is invertible in $\Bbbk$, one can eliminate boxes from the quotient calculus altogether, replacing them with linear combinations of double dots.

Addendum 4. *All the relations which involve boxes can be rewritten so that they only involve double dots. This is not done here, but can be found in any of the papers [33, 34, 36].*

Replacing boxes with double dots is much more natural, since it emphasizes that the boxes themselves should have a color, and that the set of polynomials depends in a natural way upon the coxeter group S_{n+1}. Viewing them as boxes may help make the proofs in this paper more intuitive, however.

4.7. Color Elimination

We have already shown that $\mathfrak{D}\mathcal{C}_1 \cong \mathcal{S}\mathcal{C}_1$, without needing to investigate in depth any morphisms except those from $\emptyset$ to $\underline{\mathbf{i}}$ an increasing sequence. Because this pins down the size of all Hom spaces, we can deduce some additional facts about general morphisms. The following result is not used elsewhere in this paper, but may come in handy when constructing the analogous category for arbitrary Coxeter systems.

Proposition 4.30. *Let Γ be a graph where the color j does not appear in the boundary, then Γ can be reduced to graphs not containing the color j at all.*

We can already show this when j is extremal, since by Proposition 4.19 we may apply the weak j-colored basic moves and reduce the color j to the empty graph. Unfortunately, as in Remark 4.29, we do not have a diagrammatic result of this proposition in general, but use the known size of Hom spaces to prove it.

For $X \subset \{1, \ldots, n\}$, (i.e., for a Coxeter subgraph of A_n) we let $\mathfrak{D}\mathcal{C}_1[X]$ be the category defined analogously but where edges can only be labelled by colors in X. If colors which do not appear on the boundary are not needed in the graph after reduction, then the natural inclusion of $\mathfrak{D}\mathcal{C}_1[X]$ into $\mathfrak{D}\mathcal{C}_1$ is fully faithful, which one would expect. The full faithfulness of this inclusion is equivalent to the above proposition.

However, as discussed in the previous section, the boxes which are allowed to appear should also depend on X. For the rest of this section, we assume we are working in the e_1 quotient. Let $\mathfrak{D}\mathcal{C}_1(X)$ be the category defined analogously, where edges are labelled in X, and where the only polynomials appearing are in the subring $R(X)$ generated by double dots colored in X. This is really the category we should look at. Given a morphism in $\mathfrak{D}\mathcal{C}_1[X]$, we can use polynomial forcing rules to guarantee that the only polynomials not in $R(X)$ appear in the lefthand region. We do not prove it here, $\mathfrak{D}\mathcal{C}_1[X]$ is simply $\mathfrak{D}\mathcal{C}_1(X)\otimes_{R(X)}R$; that is, the objects are unchanged, and morphism spaces undergo base change.

Proposition 4.31. *Let X be a Coxeter subgraph of A_n, and let $W(X)$, $H(X)$, $\mathcal{SC}_1(X)$ designate the corresponding constructions for this Coxeter graph, then there is a functor $F_1(X)$ from $\mathfrak{D}\mathcal{C}_1(X)$ to $\mathcal{SC}_1(X)$ which is an equivalence of categories.*

Proof. We will only sketch this result. The set X will be a disjoint union of various subgraphs X_l, each isomorphic to A_{m_l} for some $m_l \leq n$. For each X_l we know that $\mathfrak{D}\mathcal{C}_1(X_l) \cong \mathcal{SC}_1(X_l)$, and we have all the results aforementioned. Moreover, for any $l \neq l'$ and objects $M \in \mathfrak{D}\mathcal{C}_1(X_l)$ and $N \in \mathfrak{D}\mathcal{C}_1(X'_l)$, we have natural isomorphisms $M \otimes N \cong N \otimes M$ constructed with 4-valent crossings. Thanks to the distant sliding rules, these natural isomorphisms commute in the proper way with all morphisms in $\mathfrak{D}\mathcal{C}_1(X_l)$ or $\mathfrak{D}\mathcal{C}_1(X'_l)$. The category $\mathfrak{D}\mathcal{C}_1(X)$ can be constructed via some universal "symmetric monoidal product" construction, by taking the product over the categories $\mathfrak{D}\mathcal{C}_1(X_l)$. The same holds true for $\mathcal{SC}_1(X)$, which will yield the result. □

Given the sketchy result above, we can calculate the rank of Hom spaces in $\mathfrak{D}\mathcal{C}_1(X)$ by using the standard trace map on $H(X)$. This trace map commutes with the standard trace on H under the inclusion $H(X) \subset H$. Thus we see that, for objects $\underline{\mathbf{i}}$ and $\underline{\mathbf{j}}$ in $\mathfrak{D}\mathcal{C}_1(X)$, the space $\mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1}(\underline{\mathbf{i}}, \underline{\mathbf{j}})$ is a free R module of some rank r, and the space $\mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1(X)}(\underline{\mathbf{i}}, \underline{\mathbf{j}})$ is a free $R(X)$ module of the same rank r. Therefore $\mathrm{Hom}_{\mathfrak{D}\mathcal{C}_1[X]}(\underline{\mathbf{i}}, \underline{\mathbf{j}})$ is a free R module of rank r as well, and the inclusion of $\mathfrak{D}\mathcal{C}_1[X]$ in $\mathfrak{D}\mathcal{C}_1$ is fully faithful.

We quickly sketch an alternative proof of Proposition 4.30, which assumes that we know that Hom spaces in $\mathfrak{D}\mathcal{C}_1$ conform to the standard trace. We can think of any graph with boundary as a morphism in $\mathfrak{D}\mathcal{C}_1$ from $\emptyset$ to $\underline{\mathbf{j}}$. We have already "inductively calculated" the space of such morphisms, along the lines of Remark 2.1 and Proposition 2.13. Namely, if $\underline{\mathbf{j}}$ has no repeated colors, then every morphism is a polynomial with boundary dots. If $\underline{\mathbf{j}}$ has repeated colors, we use idempotent decompositions to consider the Hom space as the sum of the Hom spaces of its summands, or apply biadjunction to cycle the sequence $\underline{\mathbf{j}}$, until we have expressed the Hom space in terms of $\mathrm{Hom}(\emptyset, \underline{\mathbf{k}})$ for various $\underline{\mathbf{k}}$ nonrepeating. Since neither biadjunctions nor idempotent decompositions add any new colors to the graph, we have just constructed a generating set of morphisms in a way which does not involve any colors which were not in $\underline{\mathbf{j}}$ to begin with.

5. Proofs

It remains to prove Proposition 3.11, which we do in the first section, and Proposition 4.19, which we do in the second.

5.1. $\mathcal{F}_1$ Is a Functor

We need now to check that each of our relations holds true for Soergel bimodules, after application of F_1. The relations are listed in full in Section 3.4. These checks are entirely tedious and straightforward, and require little imagination. With a little imagination, however, there are some tricks which make checking most relations trivial.

Again, we always use relation (2.6) to identify elements of $B_{\underline{i}}$ with $d(\underline{i}) + 1$ tensors of polynomials.

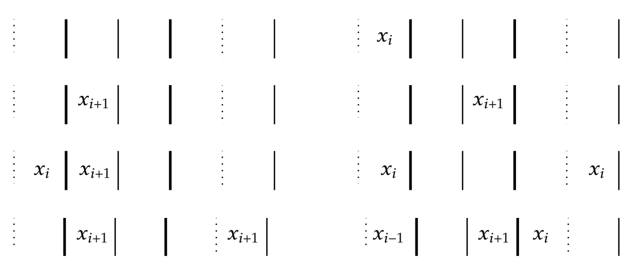
For brevity, we will call any element which is a tensor product $1 \otimes 1 \otimes 1 \otimes \cdots$ a *1-tensor*. EndDot, Split, and the 4-valent and 6-valent vertices all send 1-tensors to 1-tensors. This simple fact is already enough to prove relations (3.1), (3.32), (3.33), (3.34), and (3.36), since the bottom bimodule in each of these pictures is generated by the 1-tensor; we call this the *1-tensor trick*. Caps and Merges kill a 1-tensor, while Cups and StartDots send it to a sum of two-linear terms.

There are several choices to make when checking the various relations. Once the twisting relations are shown, one is free to prove any twist of the other relations. Also, one may choose which set of generators of the source bimodule to check equality on. Finally, whenever a Cup or a StartDot appears, there are two equivalent ways to write the result, since $x_i \otimes 1 - 1 \otimes x_{i+1} = 1 \otimes x_i - x_{i+1} \otimes 1$ in B_i, and one might be easier than the other to evaluate.

The first trick is to choose the set of generators which makes the calculation the easiest. Let us remind ourselves of the arguments of Remark 2.7. An arbitrary module $B_{\underline{i}}$ of length d will have 2^{d-m} generators as an R-bimodule, where m is the number of different colors that appear. Let us use the term *i-pair* to denote two instances of the index i in $\underline{\mathbf{i}}$, separated only by colors $\neq i$. Let X denote the set of all i-pairs for all i; the size of X is $d - m$. As we force polynomials to the right or left, a variable x_i (or alternatively x_{i+1}) might get stuck between the two i-colored lines of an i-pair, and this independently of each other i-pair or j-pair for $j \neq i$. The following claim is easy to show from the forcing rules.

Claim 5.1. Each $B_{\underline{i}}$ will be generated as an R-bimodule by any set Y of 2^{d-m} linearly independent tensors, for which we have a bijection between Y and the power set of X, satisfying the following property: If a tensor $y \in Y$ corresponds to a subset $X_y \subset X$, then each i-pair in X_y corresponds to a distinct linear factor of y which is either x_i or x_{i+1} somewhere inside the i-pair.

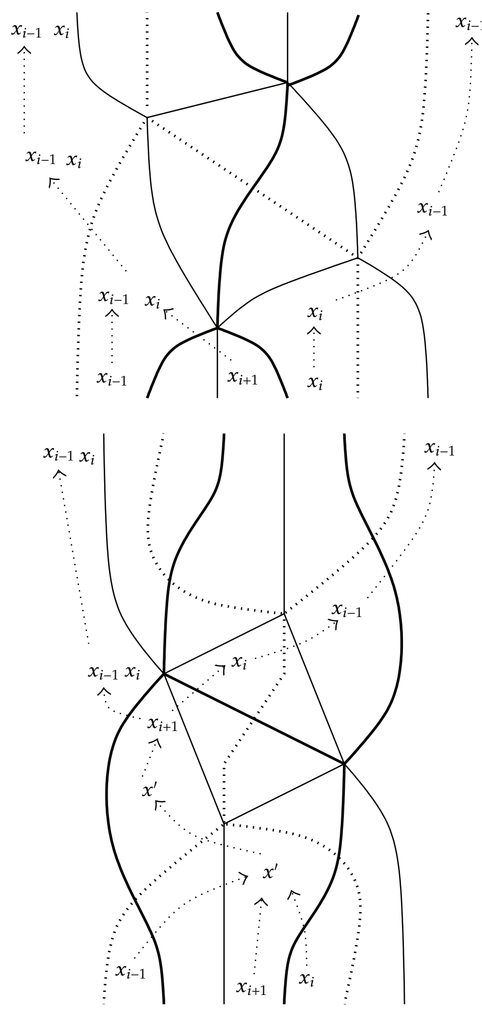
It is a mouthful, but an example clarifies it. The following is an example of a set of generators for $B_{(i-1)(i+1)i(i+1)(i-1)i}$, where $d - m = 3$ so we need 8 generators:



Each picture represents a tensor, where by convention a blank area is filled with a 1-tensor. Reading across, the first picture corresponds to the empty subset of X. The second picture corresponds to the $i-1$-pair, the third to the $i+1$-pair, and the fourth to the i-pair. Since these three are linearly independent, they take care of all the linear generators. Clearly the fourth picture could have also worked for the $i+1$-pair, but if we had chosen it as such, we would have had to choose a new linearly independent vector for the i-pair. The fifth generator corresponds to the i − 1-pair and the $i+1$-pair, the sixth generator to the $i-1$-pair and the i-pair, and the seventh generator to the i-pair and the $i+1$-pair. These three are also linearly independent. The final generator corresponds to the entire set X.

A clever choice of which generators to use may greatly simplify a calculation by reducing the number of terms in intermediate steps. There are two main instances when this occurs. Either 6-valent vertex with strands i and $i+1$ will send $1 \otimes x_{i+1} \otimes 1 \otimes 1$ or $1 \otimes 1 \otimes x_{i+1} \otimes 1$ to a single tensor, while it may send $1 \otimes x_i \otimes 1 \otimes 1$ to the sum of two tensors, thus doubling the work we need to do in the remainder of the calculation. Also, $x_i x_{i+1}$ entering a 6-valent vertex $i(i+1)i$ in the second slot may by moved across the i-strand to the left for a simple calculation, while x_i^2 leads to a more complicated solution.

The second trick will be a useful diagrammatic way of evaluating homomorphisms in $\mathcal{SC}_1$, which really only works well when applied to a well-chosen set of generators. In particular, the choice of generators above makes the verification of triple overlap associativity (3.37) rather straightforward. We demonstrate the graphical method in the most difficult case, the highest degree generator

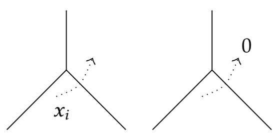


Here $x' = x_{i-1}x_i x_{i+1}$.

We keep track of the image at every stage in the calculation, which is one term as shown, with all blank spaces being filled by 1-tensors. An arrow indicates either bringing a symmetric polynomial through a line, or applying a 6-valent vertex to a tensor. The lack of an arrow indicates that a 1-tensor is sent to a 1-tensor. This works well only because every intermediate term is a single tensor, not a sum of tensors.

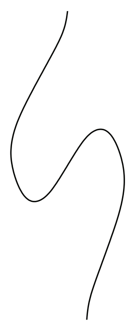
We leave it as an exercise to the reader to verify, using this graphical method of calculation, that both sides of the triple overlap associativity relation, as displayed above, send the other 7 generators to the same elements. Precisely, both sides send the 1-tensor to a 1-tensor, and the other generators to 1-tensors with polynomials in various slots: the 2nd and 4th generators to x_{i+2} in the last slot, the third generator to x_{i-1} in the last slot, the 5th to $x_{i-1}x_{i+2}$ in the last slot, the 6th to $x_i x_{i-1}$ in the first slot, the 7th to $x_{i+1}x_{i+2}$ in the first slot, and the 8th as shown above to $x_i x_{i-1}$ in the first slot and x_{i-1} in the last slot. We promise that the graphical method will work for this set of generators.

One can extend this graphical method easily to handle other morphisms. As an example, we show the Merge map below



When x_{i+1} enters a Merge, it is sent to -1. Almost the same pictures can be drawn for the Cap. Split and EndDot send 1-tensors to 1-tensors, and no arrows are needed.

When there is a cup or a StartDot, a 1-tensor is sent to a sum of two terms, each of which must be evaluated separately. However, the sum can be written as either $x_i \otimes 1 - 1 \otimes x_{i+1}$ or $1 \otimes x_i - x_{i+1} \otimes 1$, so we may choose the one whichever is more convenient. Often the problem of having two terms is very temporary. For example, consider what happens to the 1-tensor under the map



The cup creates two terms, the first with a 1 under the cap and the second with x_i under the cap. But the first term is annihilated immediately by the cap, and the cap eats the x_i from the other term, to return back a 1-tensor. So long as a cup or a StartDot appears right next to a cap or a merge, one of the two terms is always immediately annihilated. This is the case for (3.5), (3.6), (3.8), which all send 1-tensors to 1-tensors as a result. After a quick polynomial slide, the same is true for (3.31), and using the more convenient choice so the x_{i+1} ends up underneath the 6-valent vertex, a quick calculation shows it for (3.27) as well. We call this the *cupcap trick*.

We now list and prove the necessary relations, following the same order as in Section 3.4. One could print out this list and hold it next to that section, where the relations are and the functor is defined, to make this ordeal easier.

(1) Biadjointness (3.5) follows from the cupcap trick.

(2) Trivalent twisting (3.6) follows from the cupcap trick.

(3) Associativity (3.1) follows from the 1-tensor trick.

(4) Dot twisting (3.8) follows from the cupcap trick.

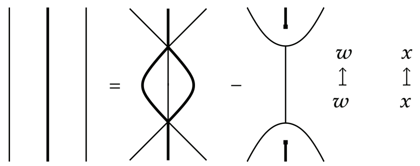
(5) The needle consists of a split, sending a 1-tensor to a 1-tensor, and a cap, killing a 1-tensor. Hence the needle relation (3.18) follows.

(6) The polynomial slide relations clearly hold for Soergel bimodules, by definition.

(7) For the broken line relation (3.16), both sides clearly send the 1-tensor to $x_i \otimes 1 - 1 \otimes x_{i+1}$.

(8) For the double dot relation (3.17), both sides clearly send the 1-tensor to $x_i - x_{i+1}$.

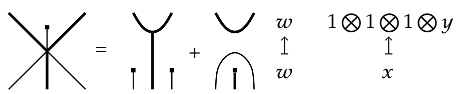
(9) The three-line decomposition relation (3.26) is a calculation. We need to show it for both color variants, whether thin is i and thick $i+1$ or vice versa. Let w be the 1-tensor and $x = 1 \otimes x_{i+1} \otimes 1 \otimes 1$, regardless of which variant we are in, then it is easy to check that



Consider what happens to x, under the first color variant. The double 6-valent vertex sends x to $1 \otimes 1 \otimes 1 \otimes x_{i+2} = 1 \otimes 1 \otimes x_{i+2} \otimes 1$, while the rightmost picture sends x to $1 \otimes x_{i+1} \otimes 1 \otimes 1 - 1 \otimes 1 \otimes x_{i+2} \otimes 1$. The sum of these two is just x again. We leave similarly easy calculations to the reader in the future.

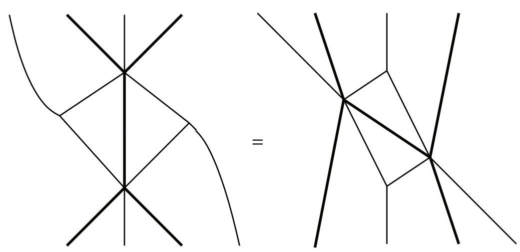
(10) 6-valent twisting (3.27) follows from the cupcap trick.

(11) Adding a dot to a 6-valent vertex (3.28) needs to be checked for both color variants. Define w and x as before, and it is easy to check that



Here $y = x_i$ for one color variant, $y = x_{i+2}$ for the other.

(12) Two-color associativity (3.29) is a calculation. We recommend using the following twist:



and using the four generators: a 1-tensor, $1 \otimes 1 \otimes x_{i+1} \otimes 1 \otimes 1$, $1 \otimes x_{i+1} \otimes 1 \otimes 1 \otimes 1$ and $1 \otimes x_{i+1} \otimes 1 \otimes x_{i+1} \otimes 1$. Evaluate using the graphical calculus. The first two generators are killed by both maps thanks to merge maps. The third generator is send to a 1-tensor by both maps. The final generator is also killed by both maps, as symmetric polynomials are slid out of the way and merge maps eat the remaining 1-tensor.

(13) 4-valent twisting (3.31) follows from the cupcap trick.

(14) Sliding a dot or a trivalent vertex through a distant line both follow from the 1-tensor trick.

(15) Sliding a 6-valent vertex through a distant line is easily checked on both generators.

(16) Sliding a 4-valent vertex through a distant line follows from the 1-tensor trick.

(17) Three-color associativity will be a calculation, using the generators and pictures described above, that we leave to the reader.

5.2. Graphical Proofs

In this section, we provide graphical proofs for a series of propositions which collectively prove Proposition 4.19. Recall the definitions of the strict and weak i-colored moves from Section 4.4. Remember that an extremal color i is one which appears in a graph, but for which either $i-1$ or $i+1$ does not appear. Each of the results states what reduction moves the relations of $\mathfrak{D}\mathcal{C}_1$ allow. Because more than three colors may be required for some proofs, we use extra line styles that designate other arbitrary colors, and are very explicit about what colors they are allowed to be. A thin line will still always represent i and a thick line $i+1$, but we will be lax about other colors.

Lemma 5.2. *One may strictly pull a distant line under any other graph, like in the relations* (3.32)–(3.36). *That is, these relations can be viewed as graph reduction moves.*

Proof. The only significant part of this lemma is that we can still apply the relations mentioned, even in the presence of arbitrary polynomials in each region. However, it is an immediate observation that any polynomial inside a region bounded by at least two-lines of distant colors may be slid entirely out of the region, which does not affect the underlying graph. □

The above lemma highlights the fact that we must deal with polynomials when they appear. However, the only significant polynomials are those which appear in internal regions (regions not touching the boundary of the planar strip) since all our computations will be local, so polynomials in external regions can be moved outside the calculation. We will mention whenever a polynomial is relevant, and the reader can check that whenever we do not mention it, there are only external regions.

Proposition 5.3. *One may apply the (strict) i-colored double dot removal move to any graph Γ.*

Proof. Suppose the i-graph of Γ contains two dots connected by an edge, then in Γ, the "edge" connecting them can only meet series of 4-valent vertices with various j-strands for j distant from i. Since dots can be slid across distant-colored strands by (3.33), we may slide the dots until they are connected directly by an edge, and just use (3.17) to reduce the graph. □

Proposition 5.4. *One may apply the weak i-colored dot contraction move to any graph Γ.*

Proof. Suppose the i-graph of Γ contains a trivalent vertex connected to a dot. In Γ, the trivalent vertex is either a trivalent or 6-valent vertex v, and the edge connected the dot to the vertex may have numerous 4-valent vertices. Again, in Γ the dot may be slid across distant-colored strands until it is connected directly to v. If v is trivalent then (3.4) allows strict dot contraction, while if v is 6-valent, (3.28) sends Γ to the sum of two graphs allowed by weak dot contraction. □

Proposition 5.5. *One may apply the (strict) i-colored dot extension move to any graph Γ.*

Proof. This is trivial, by (3.4). □

The following lemma is needed to prove the remainder of the basic i-colored moves.

Lemma 5.6. *Let Γ be a sequence of lines $\underline{\mathbf{j}}$ (i.e., the identity map of $\underline{\mathbf{j}}$), and let i be an extremal color in $\underline{\mathbf{j}}$, then Γ is equal in $\mathfrak{D}\mathcal{C}_1$ to a sum of idempotents, such that each idempotent factors through a sequence of lines $\underline{\mathbf{k}}$ where i appears no more than once, and such that the idempotents do not introduce any new colors not present in $\underline{\mathbf{j}}$.*

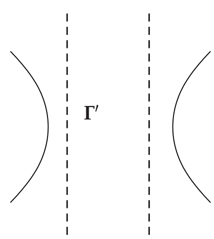
Proof. This proof is akin to the combinatorial argument that the relations defining the Hecke algebra or the symmetric algebra are enough to take any complicated word in b_j or s_j to a reduced form. We use induction on the length of $\underline{\mathbf{j}}$. If $k(k+1)k$ appears in $\underline{\mathbf{j}}$ for some k, then using (3.26), we may factor through $(k+1)k(k+1)$ instead, modulo morphisms that factor through shorter length sequences. If any color appears twice consecutively in $\underline{\mathbf{j}}$ then we may apply (3.25) to replace the two adjacent lines with an "H", which factors through a sequence of lines of shorter length. If jk appears for j distant from k, then applying the R2 move we can factor through kj. None of these procedures added any new colors. So if we consider the sequence $\underline{\mathbf{j}}$ as a word in the symmetric group $s_{j_1}s_{j_2}\cdots s_{j_d}$, then any nonreduced words will reduce to smaller length sequences, and any two reduced words for the same element will factor through each other, modulo smaller length sequences.

We may now use the easily observable fact that any element of the symmetric group can be represented by a reduced word which uses s_1 at most once, or s_n at most once (although not both!). It follows from this that any element in a parabolic subgroup of S_{n+1} can be represented by a reduced word which uses an extremal index at most once. Thus, by the process above, we can replace $\underline{\mathbf{j}}$ by idempotents factoring through reduced expressions $\underline{\mathbf{k}}$ which use i at most once. □

The usefulness of this lemma can be seen in the following proof.

Proposition 5.7. *One may apply the (strict) i-colored connecting move to any graph Γ, so long as i is extremal.*

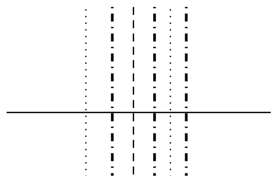
Proof. Assume $i-1$ does not appear, although the $i+1$ case is analogous. Induct on the number of colors present in the graph: the base case of one color is clear, then we are attempting to apply the i-colored connecting move to some region of the graph which looks like this:



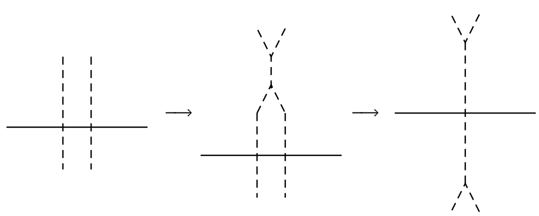
Γ' is some graph composed entirely of the colors $1, \ldots, i-2$ and $i+1, \ldots, n$, so that $i+1$ is extremal or nonexistent. In some neighborhood Γ' is just a sequence of lines $\mathbb{1}_{\underline{\mathbf{j}}}$. By the previous lemma, we may replace this sequence of lines with a sum of idempotents factoring through sequences $\underline{\mathbf{k}}$ where $i+1$ appears at most once, and $i-1$ appears not at all. Localizing to a neighborhood around $\mathbb{1}_{\underline{\mathbf{k}}}$, we may as well assume that $\underline{\mathbf{j}} = \underline{\mathbf{k}}$. Now apply the R2 move to bring the i-strands past all the other colors so that they are separated by either nothing or a single $i+1$ strand. We have not yet altered the i-graph at all, then equation (3.25) or equation (3.26) (respectively) will allow the strict i-colored connecting move. □

When we have an edge in an i-graph, a neighborhood of that edge in the full graph will be nothing more than a sequence of 4-valent vertices as the i-strand crosses distant strands. The next corollary allows us to place restrictions on which distant strands the i-strand need cross.

Corollary 5.8. *Suppose we are given a graph Γ which is a neighborhood of an i-colored strand consisting only of 4-valent vertices crossing the i-strand (see the picture below), then Γ is equivalent in $\mathfrak{D}\mathcal{C}_1$ to a linear combination of graphs Γ' which have the same i-graph, and for which the sequence of lines crossing the i-strand in Γ' contains $i+2$ and $i-2$ each at most once.*

$\Gamma =$ 

Proof. Let $\underline{\mathbf{j}}$ be the sequence of lines crossing the i-strand in Γ; clearly i, $i-1$, and $i+1$ do not appear in $\underline{\mathbf{j}}$. According to the above lemma, we may replace $\mathbb{1}_{\underline{\mathbf{j}}}$ with a sum of idempotents which factor through "nice" sequences $\underline{\mathbf{k}}$ where $i+2$ and $i-2$ each appear at most once. Moreover, for any idempotent appearing in this sum, i is distant from every color appearing in the idempotent (as well as the polynomials which may appear in the idempotent). So replacing $\mathbb{1}_{\underline{\mathbf{j}}}$ with this sum immediately above the i-strand, we may then slide the i-strand in each idempotent so that it passes through $\underline{\mathbf{k}}$. An example is given below, ignoring polynomials which can also be slid.

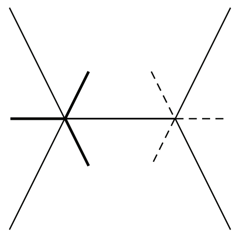


□

The last two moves will take the rest of the paper to prove.

Proposition 5.9. *One may apply the (strict) i-colored needle move to any graph Γ.*

Proposition 5.10. *One may apply the weak* i*-colored associativity move to any* i*-colored "H" in any graph* Γ*, so long as the "H" does not look like the following picture:*

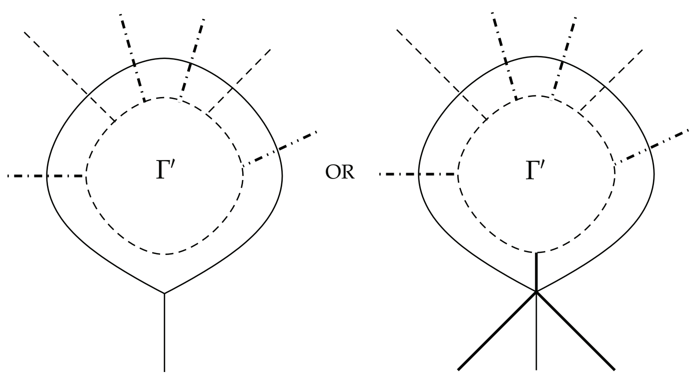


where the two vertices are 6-valent in Γ *with different colors* $i+1$ *and* $i-1$*.*

In particular, we may apply the i*-colored associativity move to any graph* Γ *when* i *is extremal.*

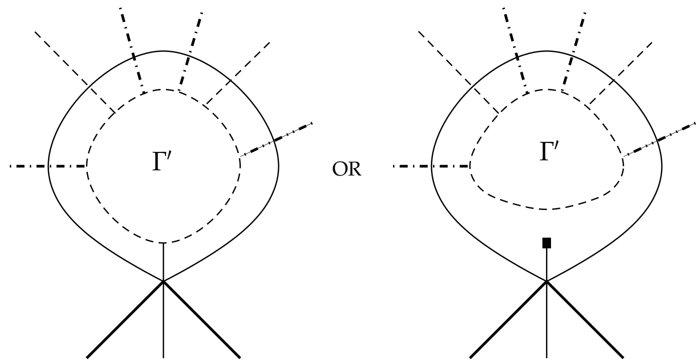
Proof Setup. We apply induction on the set of colors present in the graph to prove these two propositions simultaneously. Both propositions hold when i is the only color in the graph, by (3.10), (3.18), (3.24), where the latter is needed because an arbitrary polynomial may be within the eye of the needle. The inductive hypothesis then implies, with the previous propositions, that for any graph with fewer colors, we may apply *all* the weak k-colored basic moves for an extremal color k. In this case Proposition 4.17 says that we may reduce the k-graph to a disjoint union of simple trees. Within each region of Γ delimited by the i-graph the color i is absent, so there are fewer total colors and we may reduce both the $i-1$-graph and the $i+1$-graph inside this region to a union of simple trees. This application of the inductive hypothesis was the only reason to consider the two propositions together; we now treat them separately.

Needle Move: Modulo Induction Hypothesis. Suppose the i-colored graph contains a needle. The trivalent vertex of the needle is either trivalent or 6-valent in Γ, and the edge of the needle may contain a series of 4-valent vertices. So a neighborhood of the needle in Γ looks like one of the pictures below.

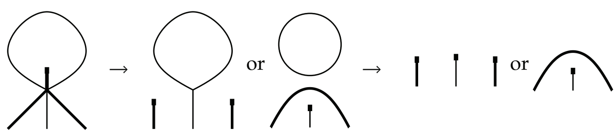


The lines entering the needle around the top are colored with various j all distant from i. We assume without loss of generality that, in the second case, the 6-valent vertex has second color $i+1$.

We may now reduce the $i+1$ and $i-1$ graph of Γ', since Γ' does not contain the color i. In the first case, the $i \pm 1$ graph has no boundary, so both reduce to the empty graph. In the second case, $i+1$ has a single boundary line so it reduces to a boundary dot, and $i-1$ has none so it reduces to the empty graph. Within the reduced Γ', the possible $i+1$ dot may be slid under other lines until no 4-valent vertices separate it from the 6-valent vertex (as in the proof of weak dot contraction). Thus we may assume that, after reduction, our neighborhood looks like

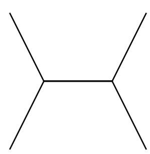


where now Γ' does not contain any of the colors $i-1$, i, $i+1$. Note that Γ' may have arbitrary polynomials in its various regions. But then by Lemma 5.2 we can pull the line forming the needle through all of Γ', thus completely ignoring Γ' from the picture! There still may be a polynomial in the eye of the needle. We have effectively reduced to the 2-color case on the right, or the one color case on the left. We know the one color case works. To check the 2-color case, we use (3.28) followed by other reduction moves.

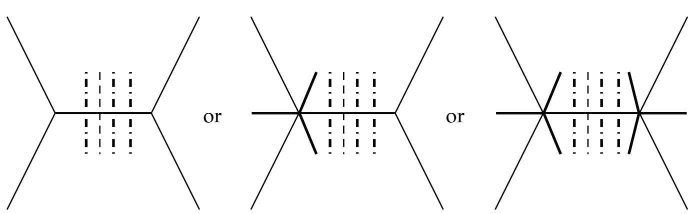


In these final graphs, the i-graph is a dot, as desired. So we may apply the strict i-colored needle move.

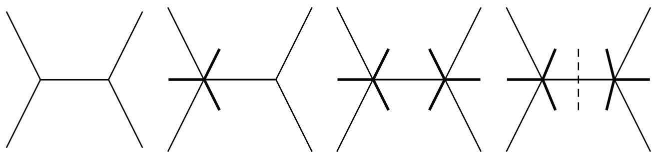
Associativity Move: Modulo Induction Hypothesis. We would like to apply associativity to the following subgraph of the i-graph.



Each trivalent vertex of the i-graph may be either trivalent or 6-valent in Γ. We are forbidding the case when one is 6-valent with $i+1$ and the other with $i-1$, so without loss of generality with assume that a 6-valent vertex has $i+1$ as the other color. The neighborhood of this subgraph in Γ looks like one of the following cases:



All polynomials may be assumed to lie outside the neighborhood. We may use Corollary 5.8 to assume that the sequence of lines passing through the middle strand contains at most one instance of $i+2$ and $i-2$. In the first two cases, all such lines may be slid one by one to the right over the trivalent vertex, removing them from the neighborhood. In the third case, all lines not labelled $i+2$ can be slid to the right or to the left, and since there is at most one line labelled $i+2$, then at most one line remains. (If there had been multiple lines labelled $i+2$, then additional lines may have been stuck between them, but thankfully Corollary 5.8 eliminates this possibility.) Four cases remain:



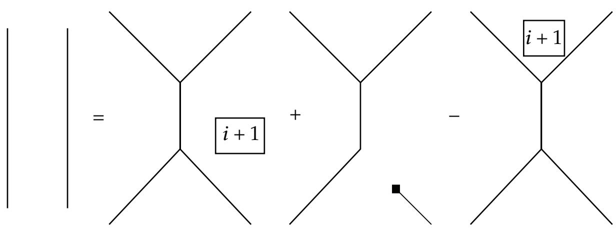
In the fourth case, the additional line is $i+2$. So far, the i-graph has been unchanged.

Case 1. One color associativity allows the strict i-colored associativity move.

Case 2. Double overlap associativity (3.30) allows the strict i-colored associativity move.

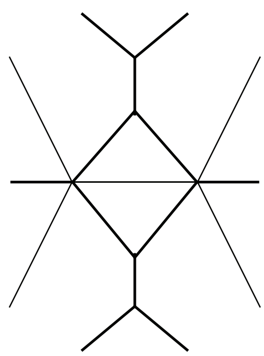
The remaining two cases use the same trick: they replace the interior lines on the top and bottom of the graph with the corresponding sum of idempotents, and then resolve each one with double or triple overlap associativity. The remaining two cases will not allow the strict associativity move, only the weak one.

Case 3. We rewrite equation (3.25), using (3.16), so that there is no polynomial on the bottom.



Applying this to the thick edges on top of Case 3, and symmetrically to the bottom, we get a difference of graphs of two kinds.

For terms which do not involve a dot, we get a graph which looks like the following, with *no* polynomial in any interior region:

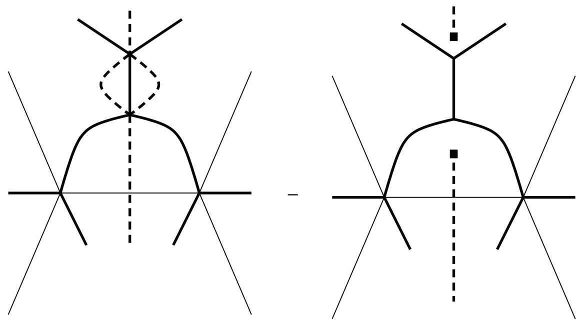


We may apply double overlap associativity (3.29) to a subgraph of this diagram (ignoring the top and bottom thick trivalent vertices), which will apply associativity to the i-graph.

Consider a term where a dot connects to a 6-valent vertex. Resolving the dot using (3.28), we get a difference of two further terms: one which looks like Case 2, and the other which is one of the alternate graphs allowed by the weak associativity move.

Thus we may apply weak i-colored associativity to each term.

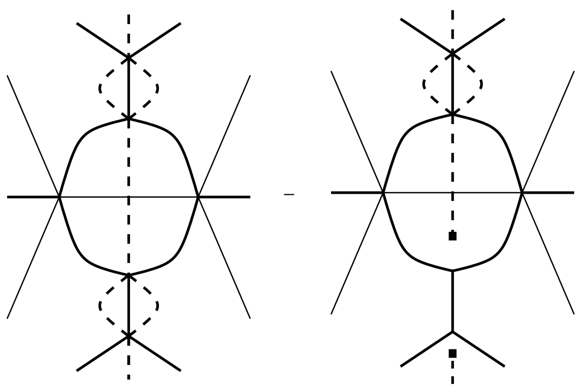
Case 4. Applying equation (3.26) to the $i+1$, $i+2$, $i+1$ sequence on top of the graph, we get the difference



Now we need to show that we can apply associativity to each of these.

For the second picture, we may drag the dot on $i+2$ through the i-strand, and then a smaller neighborhood of the X looks like Case 3.

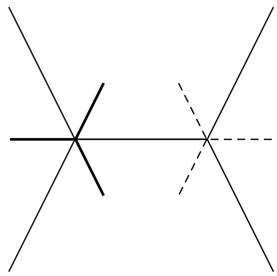
For the first picture, we once again apply (3.26) to the bottom of the graph to get the difference



Again, for the second picture we may drag the dot through and reduce using Case 3.

For the first picture, there are *no* polynomials in any internal region, because these internal regions were just created by relations. We may now apply triple overlap associativity (3.37) to the subgraph of the this picture which ignores the top and bottom 6-valent vertices. This has the effect of applying weak i-colored associativity.

Remark 5.11. It does not seem possible to apply i-colored associativity to the following graph.



Consider switching the i-graph to the other version of an "H", and look at the triangular region delimited by the i-graph on top. It has at most one 6-valent vertex on its perimeter, so that its boundary lines consist of an $i-1$ line, an $i+1$ line, and at most one other $i \pm 1$ line. This means that one of the two colors $i-1$ or $i+1$ must reduce to a boundary dot. It is not possible to obtain a degree 0 morphism if this happens.

Because of the failure of i-colored associativity in this context, our algorithm towards reducing graphs has always been to treat extremal colors and use induction. Despite this pessimism, Proposition 4.30 (color elimination) implies that we can still do a lot. However, any purely graphical proof of color elimination is complicated by the failure of i-colored associativity for nonextremal colors.

Acknowledgments

The authors were supported by the NSF Grants nos. DMS-0706924 and DMS-0739392. They are grateful to Geordie Williamson and Catharina Stroppel for remarks and suggestions on an earlier version of this paper.

References

[1] W. Soergel, "The combinatorics of Harish-Chandra bimodules," *Journal für die Reine und Angewandte Mathematik*, vol. 429, pp. 49–74, 1992.
[2] I. Grojnowski and G. Lusztig, "On bases of irreducible representations of quantum GL_n," in *Kazhdan-Lusztig Theory and Related Topics*, vol. 139 of *Contemporary Mathematics*, pp. 167–174, American Mathematical Society, Providence, RI, USA, 1992.
[3] A. A. Beĭlinson, G. Lusztig, and R. MacPherson, "A geometric setting for the quantum deformation of GL_n," *Duke Mathematical Journal*, vol. 61, no. 2, pp. 655–677, 1990.
[4] I. Grojnowski, "The coproduct for quantum GL_n," preprint 1992.
[5] J. Bernstein, I. Frenkel, and M. Khovanov, "A categorification of the Temperley-Lieb algebra and Schur quotients of $U(sl_2)$ via projective and Zuckerman functors," *Selecta Mathematica*, vol. 5, no. 2, pp. 199–241, 1999.
[6] I. Frenkel, M. Khovanov, and C. Stroppel, "A categorification of finite-dimensional irreducible representations of quantum sl_2 and their tensor products," *Selecta Mathematica*, vol. 12, no. 3-4, pp. 379–431, 2006.
[7] L. Crane and I. B. Frenkel, "Four-dimensional topological quantum field theory, Hopf categories, and the canonical bases," *Journal of Mathematical Physics*, vol. 35, no. 10, pp. 5136–5154, 1994.
[8] I. B. Frenkel, "Unpublished notes," Circa 1994.
[9] A. Lauda, "A categorification of quantum sl_2," *Advances in Mathematics*, vol. 225, no. 6, pp. 3327–3424, 2010.
[10] M. Khovanov and A. Lauda, "A diagrammatic approach to categorification of quantum groups III," *Quantum Topology*, vol. 1, no. 1, pp. 1–92, 2010.

[11] J. Chuang and R. Rouquier, "Derived equivalences for symmetric groups and sl_2-categorification," *Annals of Mathematics*, vol. 167, no. 1, pp. 245–298, 2008.
[12] R. Rouquier, "2-Kac-Moody Lie algebras," http://arxiv.org/abs/0812.5023.
[13] N. Libedinsky, "Presentation of right-angled Soergel categories by generators and relations," *Journal of Pure and Applied Algebra*, vol. 214, no. 12, pp. 2265–2278, 2010.
[14] W. Soergel, "Kazhdan-Lusztig-Polynome und unzerlegbare Bimoduln "uber Polynomringen," *Journal of the Institute of Mathematics of Jussieu*, vol. 6, no. 3, pp. 501–525, 2007.
[15] D. Kazhdan and G. Lusztig, "Representations of Coxeter groups and Hecke algebras," *Inventiones Mathematicae*, vol. 53, no. 2, pp. 165–184, 1979.
[16] D. Kazhdan and G. Lusztig, "Schubert varieties and Poincaré duality," in *Geometry of the Laplace operator*, vol. 36 of *Proceedings of Symposia in Pure Mathematics*, pp. 185–203, American Mathematical Society, Providence, RI, USA, 1980.
[17] A. Beĭlinson and J. Bernstein, "Localisation de g-modules," *Comptes Rendus des Séances de l'Académie des Sciences*, vol. 292, no. 1, pp. 15–18, 1981.
[18] J.-L. Brylinski and M. Kashiwara, "Kazhdan-Lusztig conjecture and holonomic systems," *Inventiones Mathematicae*, vol. 64, no. 3, pp. 387–410, 1981.
[19] A. A. Beĭlinson, J. Bernstein, and P. Deligne, "Faisceaux pervers," in *Analysis and Topology on Singular Spaces*, vol. 100 of *Astérisque*, pp. 5–171, Société Mathématique de France, Paris, France, 1982.
[20] M. Varagnolo and E. Vasserot, "Canonical bases and Khovanov-Lauda algebras," http://arxiv.org/abs/0901.3992.
[21] V. Mazorchuk and C. Stroppel, "Categorification of (induced) cell modules and the rough structure of generalised Verma modules," *Advances in Mathematics*, vol. 219, no. 4, pp. 1363–1426, 2008.
[22] R. Rouquier, "Categorification of the braid groups," http://arxiv.org/abs/math/0409593.
[23] R. Rouquier, "Categorification of sl_2 and braid groups," in *Trends in Representation Theory of Algebras and Related Topics*, vol. 406 of *Contemporary Mathematics*, pp. 137–167, American Mathematical Society, Providence, RI, USA, 2006.
[24] M. Khovanov, "Triply-graded link homology and Hochschild homology of Soergel bimodules," *International Journal of Mathematics*, vol. 18, no. 8, pp. 869–885, 2007.
[25] M. Khovanov and L. Rozansky, "Matrix factorizations and link homology. II," *Geometry & Topology*, vol. 12, no. 3, pp. 1387–1425, 2008.
[26] N. M. Dunfield, S. Gukov, and J. Rasmussen, "The superpolynomial for knot homologies," *Experimental Mathematics*, vol. 15, no. 2, pp. 129–159, 2006.
[27] J. Rasmussen, "Some differentials on Khovanov-Rozansky homology," http://arxiv.org/abs/math/0607544.
[28] B. Webster, "Khovanov-Rozansky homology via a canopolis formalism," *Algebraic & Geometric Topology*, vol. 7, pp. 673–699, 2007.
[29] M. Mackaay, M. Stosic, and P. Vaz, "The 1,2-coloured HOMFLY-PT link homology," *Transactions of The American Mathematical Society*, vol. 363, pp. 2091–2124, 2011.
[30] M. Mackaay, M. Stosic, and P. Vaz, "Sl(N) link homology using foams and the Kapustin-Li formula," http://arxiv.org/abs/0708.2228.
[31] S. Morrison and A. Nieh, "On Khovanov's cobordism theory for su(3) knot homology," *Journal of Knot Theory and Its Ramifications*, vol. 17, no. 9, pp. 1121–1173, 2008.
[32] M. Khovanov and R. Thomas, "Braid cobordisms, triangulated categories, and flag varieties," *Homology, Homotopy and Applications*, vol. 9, no. 2, pp. 19–94, 2007.
[33] B. Elias and D. Krasner, "Rouquier complexes are functorial over braid cobordisms," *Homology, Homotopy and Applications*, vol. 12, no. 2, pp. 109–146, 2010.
[34] P. Vaz, "The diagrammatic Soergel category and sl_2 and sl_3 foams," *International Journal of Mathematics and Mathematical Sciences*, vol. 2010, Article ID 612360, 23 pages, 2010.
[35] M. Mackaay and P. Vaz, "The diagrammatic Soergel category and sl(N) foams, for $N > 3$," *International Journal of Mathematics and Mathematical Sciences*, vol. 2010, Article ID 468968, 20 pages, 2010.
[36] B. Elias, "A diagrammatic Temperley-Lieb categorification," *International Journal of Mathematics and Mathematical Sciences*, vol. 2010, Article ID 530808, 47 pages, 2010.
[37] W. Soergel, "Kazhdan-Lusztig polynomials and a combinatoric[s] for tilting modules," *Representation Theory*, vol. 1, pp. 83–114, 1997.
[38] W. Soergel, "Gradings on representation categories," in *Proceedings of the International Congress of Mathematicians (ICM '95)*, pp. 800–806, Birkhäuser, Zürich, Switzerland, 1995.

[39] W. Soergel, "Combinatorics of Harish-Chandra modules," in *Representation Theories and Algebraic Geometry*, A. Broer, Ed., vol. 514 of *NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences*, pp. 401–412, Kluwer Academic, Boston, Mass, USA, 1998.
[40] P. Fiebig, "The combinatorics of Coxeter categories," *Transactions of the American Mathematical Society*, vol. 360, no. 8, pp. 4211–4233, 2008.
[41] N. Libedinsky, "Sur la catégorie des bimodules de Soergel," *Journal of Algebra*, vol. 320, no. 7, pp. 2675–2694, 2008.
[42] N. Libedinsky, "Équivalences entre conjectures de Soergel," *Journal of Algebra*, vol. 320, no. 7, pp. 2695–2705, 2008.
[43] B. Webster and G. Williamson, "A geometric model for Hochschild homology of Soergel bimodules," *Geometry & Topology*, vol. 12, no. 2, pp. 1243–1263, 2008.
[44] G. Williamson, *Singular Soergel bimodules*, Doctoral thesis, Albert-Ludwigs-Universität, 2008.
[45] D. Bar-Natan and S. Morrison, "The Karoubi envelope and Lee's degeneration of Khovanov homology," *Algebraic & Geometric Topology*, vol. 6, pp. 1459–1469, 2006.
[46] V. F. R. Jones, "Hecke algebra representations of braid groups and link polynomials," *Annals of Mathematics*, vol. 126, no. 2, pp. 335–388, 1987.
[47] B. Webster and G. Williamson, "Colored HOMFLYPT homology via equivariant sheaves and the weight filtration," http://arxiv.org/pdf/0905.0486.
[48] A. D. Lauda, "Frobenius algebras and ambidextrous adjunctions," *Theory and Applications of Categories*, vol. 16, no. 4, pp. 84–122, 2006.
[49] M. Müger, "From subfactors to categories and topology. I. Frobenius algebras in and Morita equivalence of tensor categories," *Journal of Pure and Applied Algebra*, vol. 180, no. 1-2, pp. 81–157, 2003.
[50] B. Bartlett, "The geometry of unitary 2-representations of finite groups and their 2-characters," *Applied Categorical Structures*. In press.
[51] B. Bartlett, *On unitary 2-representations of finite groups and topological quantum field theory*, Ph.D. thesis, University of Sheffield, 2008.
[52] J. R. B. Cockett, J. Koslowski, and R. A. G. Seely, "Introduction to linear bicategories," *Mathematical Structures in Computer Science*, vol. 10, no. 2, pp. 165–203, 2000.