A DIAMOND LEMMA FOR HECKE-TYPE ALGEBRAS

BEN ELIAS

ABSTRACT. In this paper we give a version of Bergman’s diamond lemma which applies to certain monoidal categories presented by generators and relations. In particular, it applies to: the Coxeter presentation of the symmetric groups, the quiver Hecke algebras of Khovanov-Lauda-Rouquier, the Webster tensor product algebras, and various generalizations of these.

We also give an extension of Manin-Schechtman theory to non-reduced expressions.

1. INTRODUCTION

When presenting an algebra by generators and relations, it is not obvious how large this algebra will be. Even with innocuous-looking relations, it may turn out that the algebra was zero all along. It is often possible to find a set of monomials, i.e. elements of the algebra expressed as words in the generators, which span the algebra, and which one expects is a basis. However, proving the linear independence of this set of monomials can be quite a difficult task.

For example, consider the so-called Coxeter presentation of the symmetric group $S_n$ or of its group algebra: it is generated by symbols $s_i$, $1 \leq i \leq n - 1$, modulo the relations

$$s_is_i = 1,$$

$$s_is_j = s_js_i \quad \text{for } j \neq i \pm 1,$$

$$s_is_js_i = s_js_is_j \quad \text{for } j = i \pm 1.$$

Fix an arbitrary reduced expression $w = s_{i_1}s_{i_2} \cdots s_{i_d}$ for each $w \in S_n$. After a great deal of combinatorics, one can prove that these particular reduced expressions span the group algebra. Now one can ask whether these reduced expressions are linearly independent.

One way to check linear independence is to find an action of the algebra where the elements act by linearly independent operators. For example, we can consider the standard action of the symmetric group $S_n$ on a set of size $n$, and linearize it to get a representation of the group algebra. Clearly the action of distinct permutations is linearly independent, from which we deduce that our chosen reduced expressions form a basis. This is quite painless.

However, there are numerous contexts where it is not as easy to find an action of the algebra. Consider the generalized Hecke algebra $H(\alpha, \beta)$ of $S_n$ (see [11, §7.1]), which replaces (1.1a) with

$$s_is_i = \alpha s_i + \beta$$

for certain parameters $\alpha$ and $\beta$ in the base ring. The same combinatorics as above (with an additional inductive argument by the length of an expression) implies that our chosen reduced expressions will span $H(\alpha, \beta)$. It is still the case that our chosen reduced expressions will form a basis, but finding an action which illustrates this can be hard.

For another example, consider the quiver Hecke algebra defined by Khovanov-Lauda [12] and Rouquier [16]. It has a presentation similar in style to (1.1) above, with extra generators
$x_1, \ldots, x_n$ forming a polynomial algebra, and some additional relations involving them. A sample relation in the nilHecke algebra (a special case of the quiver Hecke algebra) is

$$x_i s_i - s_i x_{i+1} = s_i x_i - x_{i+1} s_i = 1.$$  

(1.3)

One expects there to be a basis of the form $\{f w\}$, where $f$ lives in a basis of the polynomial algebra, and $w \in S_n$ (and $w$ is a fixed reduced expression for $w$). That this is a basis is proven by constructing the “polynomial representation” of the quiver Hecke algebra, and confirming that the suspected basis acts in a linearly independent fashion. However, in forthcoming work of the author with Agustin Garcia [8], we will construct variations on the quiver Hecke algebra which have no polynomial representations, and as such we need a different tool to prove linear independence.

One rather painful approach is to construct an abstract version of the regular representation. If one has a set $X_{irr}$ of monomials suspected to be a basis, one creates an abstract vector space with basis given by $X_{irr}$. One works out how each generator acts on this vector space, and then checks that the relations are satisfied. It then becomes obvious that the suspected basis acts on this abstract vector space by linearly independent operators. While this approach works in theory for any presentation, computing how the generators act and checking the relations can be a prohibitive amount of work. In special cases it can be done nonetheless. This is the approach taken by Humphreys for the Hecke algebra checking the relations can be a prohibitive amount of work. In special cases it can be done nonetheless. This is the approach taken by Humphreys for the Hecke algebra proving linear independence.

 Thankfully, there is a tool designed for precisely this purpose, the Bergman diamond lemma. In some sense, one can view this diamond lemma as an efficient tool to check that the action on the abstract regular representation is well-defined. It is a valuable tool for algebraists, and if you, dear reader, have not yet seen it, it is well worth your time to learn it! See §2 for a summary.

Remark 1.1. The history of the Bergman diamond lemma vis a vis the related idea of Gröbner bases is somewhat complex, as they were pursued independently for a time. See §2.3 for a comparison, and for a discussion of previous work related to the results of this paper.

The Bergman diamond lemma fixes a direction for each relation, and only permits it to be applied in that direction. For example, in the Coxeter presentation for $S_3$, we might always send $s_1 s_1 \leftrightarrow 1$, $s_2 s_2 \leftrightarrow 1$, and $s_1 s_2 s_1 \leftrightarrow s_2 s_1 s_2$. The monomials $\{1, s_1, s_2, s_1 s_2, s_2 s_1, s_2 s_1 s_2\}$ are irreducible in that no relation can be applied to them, while every other monomial has either $s_1 s_1$, $s_2 s_2$, or $s_1 s_2 s_1$ as a subword. Then one asks which overlaps can occur between relations, and whether such an overlap can be resolved using the relations. For example, $s_1 s_2 s_1$ contains two overlapping relations, overlapping in the bold $s_1$, and hence is ambiguous because it can be hit by a relation in two different ways:

$$s_1 s_1 s_2 s_1 \leftrightarrow s_2 s_1$$

and

$$s_1 s_1 s_2 s_1 \leftrightarrow s_1 s_2 s_1 s_2 \leftrightarrow s_2 s_1 s_2 s_2 \leftrightarrow s_2 s_1.$$

Since these two ways converged to the same element $s_2 s_1$, we say that the ambiguity is resolvable. For any words $x$ and $y$, the word $x s_1 s_2 s_1 y$ also has an ambiguity, but this one is resolvable because the minimal ambiguity $s_1 s_1 s_2 s_1$ is resolvable. The minimal ambiguities in this example are

$$\{s_1 s_1 s_2 s_1, s_2 s_2 s_2, s_1 s_1 s_2 s_1, s_1 s_2 s_1 s_1 s_2 s_1, s_1 s_2 s_1 s_2 s_1 s_2, s_1 s_2 s_1 s_2 s_1 s_2 s_1\},$$

(1.4)

all of which are resolvable.
The Bergman diamond lemma states that, if the set of monomials possesses a sufficiently nice partial order compatible with our relations, then the irreducible monomials form a basis if and only if every minimal ambiguity is resolvable. This reduces the proof of linear independence to a finite (and in fact, very small) amount of work. For example, it gives a very efficient proof that the PBW basis is linearly independent, see [1, §3].

Remark 1.2. If you are defining an algebra by generators and relations, it is good practice to check whether the overlap ambiguities are resolvable, even if you do not plan to use the Bergman diamond lemma. When I referee a paper, I typically check the first few overlap ambiguities; you might be surprised how many significant errors I have found this way. Be responsible, check your ambiguities!

Unfortunately, the Bergman diamond lemma does not apply to the Coxeter presentation of \( S_n \) for \( n \geq 4 \!). There is no choice of direction on the relations which will lead to resolvable ambiguities. A thorough discussion of what goes wrong can be found in §3. Similarly, the Bergman diamond lemma does not apply to the Hecke algebra \( H(\alpha, \beta) \), or the quiver Hecke algebras of Khovanov-Lauda-Rouquier, or their generalizations in the forthcoming work [8].

It is the goal of this paper to provide a variant of the Bergman diamond lemma which does apply to these presentations. That is, it applies to Hecke-type presentations of algebras and algebroids. These are presentations of monoidal categories (and their endomorphism rings), generated by certain crossings and dots (the dots are analogous to a polynomial subalgebra), with relations like (1.1) and (1.3). See Definition 5.7 for a precise definition. The upshot is the following theorem, stated more precisely as Theorem 5.12.

**Theorem 1.3.** Given a Hecke-type presentation of an algebra with \( n \) strands, there is a basis of the form \( \{fw\} \), where \( f \) lives in (a basis for) the “polynomial” subalgebra, and \( w \) is a fixed reduced expression for \( w \in S_n \), if and only if the following types of ambiguities are resolvable:

\[
\{s_1s_1s_1, s_1s_2s_1, s_1s_1s_2s_1, s_1s_1s_2s_2s_1, s_1s_2s_1s_3s_2s_1, s_1s_1f, s_1s_2s_1f, s_1fg\}.
\]

Remark 1.4. There are many ambiguities of each type listed above. For example, \( s_1s_2s_1s_1 \) is a permutation on three strands, and in a Hecke-type algebra these strands may have different colorings, so there is one such relation for each of the possible colorings of the strands. Nonetheless, checking these ambiguities is a finite and relatively small amount of work, and this is precisely the least amount of work one needs to do.

To prove this theorem, we use Manin-Schechtmann theory (or a minor variant thereupon) in an essential way. Manin and Schechtmann [14] made a careful analysis of reduced expressions in the symmetric group, and orientations on the reduced expression graph, which we extend to non-reduced expressions. We believe this has independent interest. See §3 for more details.

Along the way we discuss some other variants of the Bergman diamond lemma, such as a version for monoidal presentations of monoidal categories, and a version over polynomial rings (or other base rings) which need not lie in the center of the algebra.

**Addendum.** This paper appeared concurrently with a related and independent paper by Dupont [6], which uses techniques established a decade earlier by Guiraud-Malbos [9]\(^1\) Guiraud and Malbos studied rewriting theory (e.g. Bergman-style arguments) in higher categories, and developed a theory of “derivations” which, in some sense, replaces the partial order used in Bergman’s diamond lemma. Their main example in [9, Chapter 5.4] is indeed

\(^1\)Indeed, the original plan was to publicize this work after [8] was complete, but the appearance of [6] prompted an early release and the addition of several addenda.
the Coxeter presentation of the symmetric group. Dupont uses this technology, and other tools from rewriting theory dealing with non-termination (a failure of the DCC property discussed in §2), to apply Bergman-style arguments to the 2-category which categorifies the quantum group. This is more impressive than our results, which apply only to the quiver Hecke algebras which categorify the positive half of the quantum group. Dupont’s thesis introduced the work of one community of mathematicians (working on rewriting theory in higher categories) to another community of mathematicians (categorification theorists, my milieu) which were mostly oblivious to it. There is a significant amount of recent literature on rewriting theory I was unaware of when I wrote this paper, but rather than recalling it here, we point the reader to [6, 9] for a more detailed account. Regardless, we feel that our methods (using Manin-Schechtman theory) are independent from and more explicit than those of Guiraud-Malbos, so are still a worthwhile addition to the literature.

Acknowledgements. The idea for this paper germinated long ago, in response to a question of Mikhail Khovanov when I was his admiring graduate student. Thanks to Agustin Garcia for the impetus to release this upon the world. The author was supported by NSF CAREER grant DMS-1553032, and by the Sloan Foundation. We would also like to thank the anonymous referee for helpful suggestions, which improved the clarity of the exposition.

2. Bergman’s diamond lemma

2.1. Statement of the lemma. Fix a commutative ring \( k \), which will be the base ring for all our constructions.

Suppose one has defined an algebra \( A \) over \( k \) by generators \( S \) and relations \( R \). Let \( X \) denote the set of monomials, i.e. words in \( S \). Given words \( A, B \in X \), we let \( AB \) denote their concatenation. Each relation \( r \in R \) is a \( k \)-linear combination inside \( k \cdot X \), and the ideal \( I \) generated by these relations is the same as the \( k \)-submodule spanned by \( ArB \) for each \( A, B \in X \). So, as a \( k \)-module, \( A \) is the quotient of \( k \cdot X \) by \( I \).

Let \( \leq \) be a partial order on \( X \). We say that it is a semigroup partial order if \( A, B, B', C \in X \) and \( B \leq B' \) implies that \( ABC \leq AB'C \). We say it is compatible with \( R \) if every relation \( r \in R \) can be rewritten as an equation \( W_r = f_r r \), where \( W_r \in X \), and \( f_r \) is a \( k \)-linear combination of monomials \( A \) satisfying \( A < W_r \). In other words, every linear combination in \( r \) has a unique maximal monomial \( W_r \), and its coefficient in \( k \) is invertible.

To apply the Bergman diamond lemma, we will need a partial order \( \leq \) which

- is a semigroup partial order,
- is compatible with \( R \), and
- satisfies the descending chain condition (DCC).

Let us fix such a partial order for the following discussion. Note that there is no requirement that the partial order respects the length of a word, or that \( B < ABC \).

Example 2.1. We follow this example throughout the section. Suppose that \( S = \{ x, y, z \} \) and \( R = \{ yx - xy - 1, zx - xz - 2, zy - yz - 3 \} \). Then we can choose the degree lexicographic order, the (total) order \( < \) where \( A < B \) if the word \( A \) is shorter than \( B \), or if they have the same length and \( A \) is smaller in the lexicographic order, with \( x < y < z \). Degree lexicographic orders are always semigroup partial orders, and this order satisfies the DCC. The order is compatible with \( R \), as we may write our relations as \( xy = xy + 1, zx = xz + 2, \) and \( zy = yz + 3 \).

We can choose to apply relations only in one direction, replacing the monomial \( W_r \) with the linear combination \( f_r \). Instead of writing \( W_r = f_r \) we write \( W_r \mapsto f_r \). Formally, for each \( A, B \in X \) and \( r \in R \), we define an elementary resolution \( \rho_{ArB} \) to be the \( k \)-linear map on \( k \cdot X \)
which sends $AW_rB$ to $Af_rB$, and fixes every other monomial. A resolution is a composition of elementary resolutions. We call a monomial irreducible if it is fixed by every resolution, i.e. it has no subwords of the form $W_r$ for any relation $r$. Let $X_{\text{irr}}$ denote the set of irreducible monomials.

**Example 2.2.** A monomial is irreducible if and only if it has the form $x^ay^bz^c$ for some $a, b, c \geq 0$. Here are two different resolutions of the monomial $zyx$:

(2.1a) \[ zyx \mapsto yzx + 3x \mapsto yxz + 2y + 3x \mapsto xyz + z + 2y + 3x, \]

(2.1b) \[ zyx \mapsto zxy + z \mapsto xyz + 3x + 2y + z. \]

By our assumptions on the partial order, a finite number of elementary resolutions will take any monomial to a linear combination of irreducible monomials. Consequently, $X_{\text{irr}}$ is a spanning set for $A$ over $k$. The question is whether it is a basis. After all, a given monomial could conceivably be resolved into a linear combination in $k\langle X_{\text{irr}} \rangle$ in two different ways, and the difference between these resolutions would give a linear relation in $X_{\text{irr}}$.

Suppose a monomial $A$ has both $W_r$ and $W_s$ appearing inside, for $r, s \in R$. We refer to the triple $(A, W_r, W_s)$ as an ambiguity; here $W_r$ and $W_s$ refer to specific subwords of $A$ (which may have many subwords of the form $W_r$), and we assume either that $W_r$ and $W_s$ are distinct subwords, or that $r \neq s$. In theory, $W_r = W_s$ for $r \neq s$ is possible, in which case $(W_r, W_r, W_s)$ is an ambiguity. There are two possible resolutions of $A$, one using the relation $r$ and the other $s$. We wish to show that these two resolutions can be further resolved until they agree. More precisely, suppose that $A = BW_sC = DW_sE$. Then we say the ambiguity is (jointly) resolvable if one can apply further resolutions to $\rho_{BrC}(A)$ and $\rho_{DsE}(A)$ to arrive at the same element of $k \cdot X$.

**Example 2.3.** There is an ambiguity $(zyx, zy, xy)$, and our computations in (2.1a) and (2.1b) verify that this ambiguity is resolvable.

The first main result of the diamond lemma says that if every ambiguity is resolvable, then there is a unique resolution of every monomial into a linear combination of irreducible monomials.

When the subwords $W_r$ and $W_s$ do not overlap, we call $(A, W_r, W_s)$ a disjoint ambiguity. These are the easy cases. Let $A = BW_rCW_sD$ for some (possibly empty) words $B, C, D$. It is clear that one can resolve both $Bf_rCW_sD$ and $BW_rCf_sD$ to the same linear combination $Bf_rCf_sD$.

**Example 2.4.** There is a disjoint ambiguity $(xxyzy, zy, yx)$. Though unclear from the notation, the first $zy$ refers to the third and fourth letter, while the second $zy$ to the fifth and sixth letter of the word $xxyzy$. We do not feel like more precise notation is called for in this paper.

There are only two ways in which subwords of a word can intersect nontrivially. The first is an overlap ambiguity, where $W_r = BC$ and $W_s = CD$, and $A = LBCDM$, so that $W_r$ and $W_s$ intersect in the bold copies of $C$. Now it is not obvious that $Lf_rDM$ and $LBf_sM$ should have a joint resolution. If the minimal such ambiguity, namely $(BCD, W_r, W_s)$, is resolvable, then any overlap ambiguity containing $BCD$ is resolvable.

The second kind of ambiguity is an inclusion ambiguity, where $W_r = C$ and $W_s = BCD$. Again, if the minimal inclusion ambiguity $(BCD, W_r, W_s)$ is resolvable, then any inclusion ambiguity containing $BCD$ is resolvable.

**Remark 2.5.** You have probably never seen a presentation with an inclusion ambiguity. One can always modify the presentation to avoid inclusion ambiguities, see [1, §5.1]. A very
stupid example of a (non-resolvable) inclusion ambiguity would be the relations \( x \mapsto 1 \) and \( x \mapsto 2 \), and the triple \((x, x, x)\).

The second main result of the diamond lemma is merely the statement that one need only check the minimal overlap and inclusion ambiguities, to determine the resolvability of all ambiguities. It is essential to observe that, while there are infinitely many overlap ambiguities, there are very few minimal overlap ambiguities, which reduces the task of checking all overlap ambiguities to a finite amount of work.

**Example 2.6.** Concluding our example from above. The ambiguity \((zyx, zy, yx)\) is the only minimal ambiguity, and it is resolvable. Hence \(\{x^ay^bz^c\}_{a,b,c \geq 0}\) forms a basis for the algebra.

**Example 2.7.** Consider again the Coxeter presentation for \(S_3\), with \(s = s_1\) and \(t = s_2\). Let us use crossing diagrams to represent words in \(s\) and \(t\), as below.

\[
s = \begin{array}{c}
| \\
| 
\end{array}
\quad t = \begin{array}{c}
| \\
| 
\end{array}
\]

Let us use the degree lexicographical order where \(t < s\). Then we get elementary resolutions \(stst \mapsto tsts\), \(ss \mapsto 1\), and \(tt \mapsto 1\). Then the reader should confirm that the only minimal ambiguities are the overlap ambiguities pictured below, and that they are all resolvable.

\begin{equation}
(2.2)
\end{equation}

Note that \(tstt\) is not an overlap ambiguity, as one never uses the relation \(stst \mapsto tsts\) in the other direction.

**Theorem 2.8 (Bergman diamond lemma).** (See [1, Theorem 1.2]) Let \(A\) be an algebra presented over a commutative (central) base ring \(k\) by generators \(S\) and relations \(R\). Suppose that one has a partial order \(\leq\) which is a semigroup partial order, is compatible with \(R\), and satisfies the DCC. Then the irreducible monomials \(X_{irr}\) form a basis for \(A\) if and only if each ambiguity is resolvable, if and only if each minimal ambiguity (overlap and inclusion) is resolvable, if and only if each minimal ambiguity is resolvable relative to \(\leq\) (see Definition 2.11 below). In this case, we say that the data \((S, R, \leq)\) is Bergman type.

**Example 2.9.** Consider a different presentation for the group algebra of \(S_3\), with the same generators \(s\) and \(t\): the relations are \(ss \mapsto 1\), \(tt \mapsto 1\), and \(ststst \mapsto 1\). There are too many irreducible monomials (like \(ststs\)), and there are unresolvable ambiguities (like \(sststst\)), so that the presentation is not Bergman type. The Bergman diamond lemma is still true, but it is not particularly useful. Having a nice presentation is extremely important.

**Remark 2.10.** In Bergman’s proof of Theorem 2.8, overlap ambiguities and inclusion ambiguities are treated in precisely the same way. The kind of ambiguity does not affect the essentially formal argument.

Bergman also introduces a more practical method to check whether ambiguities are resolvable: namely, being resolvable relative to \(\leq\).

**Definition 2.11.** Let \(A\) be a word. Define \(I_{<A}\) as the \(k\)-linear span of \(\{B(W_r - f_r)C\}\) for all relations \(r\) and words \(B, C \in X\) such that \(BW_rC < A\). This \(k\)-submodule encapsulates all the relations one has already deduced by induction, before examining the monomial \(A\). Suppose
that \( A = BW, C = DW, E \) is an ambiguity, and consider the two elementary resolutions \( \rho_{BC}(A) \) and \( \rho_{DE}(A) \). We say that the ambiguity \( (A, W_r, W_s) \) is resolvable relative to \( \leq \) if \( \rho_{BC}(A) - \rho_{DE}(A) \in I_{< A} \).

To elaborate, suppose we have an ambiguity \( A \) as in Definition 2.11. To show that it is resolvable using the original definition, one would have to apply relations \( R \) only in the oriented direction \( W_r \mapsto f_r \) until \( Bf_rC \) and \( Df_rE \) became equal. Using relative resolvability, one is permitted to apply the relations in both directions, so long as the process stays below \( A \) in the partial order. The “restricted ideal” \( I_{< A} \) does not care about the directionality of relations! In practice, the notion of relative resolvability allows one to take identities which one has already proven (possibly using relations in both directions), and use them to check whether ambiguities are resolvable.

2.2. A crucial warning. Now we must make the following warning, and make it very loudly!

The naive approach to using the Bergman diamond lemma is to do what we did in Example 2.7: choose for each relation an equation of the form \( W_r = f_r \), and check the minimal ambiguities. In checking the ambiguities, one never needs to use the partial order \( \leq \), so that it seems irrelevant. However, if one does not find an appropriate partial order \( \leq \), then one can not apply the Bergman diamond lemma.

**Definition 2.12.** This definition (which we use later in the paper) uses the notation of Example 2.7. We define a partial order \( \leq_3 \) on the set of words in the alphabet \( \{s, t\} \) as follows.

- If the word \( A \) is shorter than the word \( B \), then \( A <_3 B \).
- If \( A \) and \( B \) have the same length, and they express different elements of \( S_3 \), then they are incomparable.
- If \( A \) and \( B \) have the same length, and express the same element of \( S_3 \), but are not related by a sequence of braid relations, then they are incomparable.
- Finally, if there is still a chance of them being comparable, then \( A <_3 B \) if \( A \) occurs after \( B \) in lexicographic order.

Thus, for instance, \( tst <_3 sts \). The reader should confirm that \( \leq_3 \) is a partial order, and that it satisfies the hypotheses of Theorem 2.8.

It is not obvious that a partial order satisfying the hypotheses of Theorem 2.8 exists, even when all the ambiguities are resolvable. A counterexample can be found in [1, §5.4]. In [1, §5.4] it is stated without proof that, if the base ring \( k \) is a domain, all ambiguities are resolvable, and no infinite chain of nontrivial elementary resolutions can be applied to any monomial, then there exists a partial order satisfying the hypotheses of Theorem 2.8. This last condition corresponds to \( \leq \) satisfying the DCC, and fails for many examples.

**Example 2.13.** Suppose we wished to apply the relation \( 1 \mapsto ss \) in this direction instead. Then it is impossible for any compatible partial order to satisfy the DCC.

2.3. Comparison to Gröbner bases. When a presentation is not Bergman type, one can modify it until it is. There is a formal procedure to do this, see [1, §5.3]. In Example 2.9, the ambiguity \( ssststst \) is not resolvable because \( tstst \) and \( s \) are irreducible. We can add a new relation \( tstst \mapsto s \), which resolves the ambiguity \( ssststst \), but creates new ambiguities. Then we can force the unresolvable ambiguities to be resolvable by adding new relations, and can repeat. It is unclear if this process will terminate, but when it does, one has a Bergman type presentation, and the monomials which remain irreducible at the end are a basis. This is effectively Buchberger’s algorithm [4] for finding a Gröbner basis.
The theory of Gröbner bases, as initiated by Shirshov [18], is very similar to the Bergman diamond lemma. It fixes a set of generators, a partial order on monomials, and an ideal $I$ (i.e. the ideal generated by the relations $R$), and seeks a Gröbner basis, which is a basis for a generating subspace of $I$ (i.e. a collection of specific relations) such that all ambiguities are resolvable. Unlike the Bergman diamond lemma, Gröbner basis theory cares about the ideal itself, not a particular set of generators for the ideal (it goes on to find its own generators for the ideal).

We think of changing the relations as violating one goal of the diamond lemma: to take a fixed presentation and confirm a basis with the minimal amount of work. Instead, the procedure above adds very many new relations, and does not represent a small amount of work. Admittedly, in the age of computers, Buchberger’s algorithm can be extremely effective. Moreover, we may be interested in whether a particular set of monomials, the irreducibles for our chosen set of relations, form a basis. For the Coxeter presentation of $S_n$, this is precisely asking whether the basis of the group algebra is in bijection with $S_n$, rather than being some smaller set! For this question, adding more relations will change the set of irreducible monomials, so it will not answer the question directly.

There is a large literature on Gröbner bases for various algebras, including the group algebras of Coxeter groups [3] (like the symmetric group) and their braid groups [2]. It is acknowledged that (outside of dihedral type) the Coxeter presentation is not Bergman type, and (like this paper) that vast simplifications arise when one treats the commuting relations (1.1b) in different fashion to the other relations. Instead of finding Gröbner bases via an algorithm, we do something different, modifying the Bergman diamond lemma itself to let the Coxeter presentation of the symmetric group be Bergman type. Doing this requires an in-depth study of non-reduced expressions, which we have not seen in the previous works we have explored.

**Remark 2.14.** By no means am I an expert on the literature of Gröbner bases or the Bergman diamond lemma, so I would appreciate knowing of other attempts along these lines.

### 2.4. The diamond lemma for linear categories.

In this section and the next we describe some simple variants on the Bergman diamond lemma, which follow from it directly.

Let $k$ be a commutative ring. A (free) $k$-linear category is a category where the morphism spaces are (free) $k$-modules, and where composition is $k$-bilinear. Any $k$-algebra can be viewed as a $k$-linear category with one object.

Given a $k$-linear category $C$, one can consider the (possibly infinite) direct sum $A = \bigoplus_{n,m \in \text{Ob}(C)} \text{Hom}(n, m)$. This is not necessarily an algebra, as when $\text{Ob}(C)$ is infinite it will not have a unit. Instead, it is an algebroid or locally unital algebra. It has a collection $\{\text{id}_n\}_{n \in \text{Ob}(C)}$ of orthogonal idempotents, whose sum $\sum \text{id}_n$ need not exist in $A$ (when the sum is infinite), but functions like the identity element. That is, this (infinite) sum acts via a finite sum when multiplying by any element of $A$, and this finite sum acts by the identity.

Almost every result in ring theory has an immediate and obvious analog for algebroids, though they are not always written down in the literature. For example, it is easy to adapt Bergman’s diamond lemma to $k$-linear categories, constructed by generators and relations. Quiver algebras and their quotients are excellent examples to keep in mind, where the generators are thought of as arrows between two objects (i.e. vertices), words are paths, and relations are linear combinations of paths set to zero.

The main difference between algebras and algebroids is that every morphism has a source and a target: one can separate the set $X$ of monomials into sets $X(n, m)$ of monomials which are morphisms from $n$ to $m$. One may as well assume that the partial order $\leq$ is defined
separately on each set \(X(n, m)\) (or equivalently, that morphisms with different sources or targets are incomparable).

**Example 2.15.** Consider the symmetric category \(\text{Sym}\). It has one object for each \(n \in \mathbb{N}\), and \(\text{Hom}(n, m) = 0\) if \(n \neq m\). One has \(\text{End}(n) = \mathbb{k}[S_n]\). Then \(\text{Sym}\) is a \(\mathbb{k}\)-linear category. The presentations of each symmetric group individually give rise to a presentation of \(\text{Sym}\).

2.5. **When the base ring is not central.** Often one has an inclusion of rings \(\mathbb{k} \subset \mathcal{A}\), where \(\mathcal{A}\) is free over \(\mathbb{k}\) as a left module (or a right module), but \(\mathbb{k}\) does not live in the center of \(\mathcal{A}\). The inclusion of the polynomial ring inside the nilHecke algebra (or more generally, a quiver Hecke algebra) is an example of this phenomenon. One can ask whether the Bergman diamond lemma can be used to prove that \(\mathcal{A}\) is free over \(\mathbb{k}\) as a left module. We assume \(\mathbb{k}\) is commutative, for simplicity.

When presenting \(\mathcal{A}\) over \(\mathbb{k}\) with additional generators \(\mathcal{S}\), monomials in \(\mathcal{S}\) are not sufficient to describe all elements of \(\mathcal{A}\), as we may need to multiply by \(\mathbb{k}\) as well. That is, we must consider \(\mathbb{k}\)-words: sequences of the form \(A f B g C \cdots\) where \(A, B, C, \ldots\) are words in \(\mathcal{S}\), and \(f, g, \ldots \in \mathbb{k}\). The relations in \(\mathcal{R}\) will involve \(\mathbb{k}\)-words. We refer to words just in the alphabet \(\mathcal{S}\) as true words. More precisely, we are interested in whether \(\mathcal{A}\) is free over \(\mathbb{k}\) as a left module, with a basis consisting entirely of true words.

If this is to happen, then every \(\mathbb{k}\)-word must be writable (after applying relations) as a left \(\mathbb{k}\)-linear combination of true words. In particular, for each \(s \in \mathcal{S}\) and \(f \in \mathbb{k}\), there must be an equality \(s f = \sum g_i A_i\), for true words \(A_i\) and coefficients \(g_i \in \mathbb{k}\). However, the \(\mathbb{k}\)-word \(s f\) does not have many subwords, so for it to be non-irreducible in a Bergman-type presentation, there must be a relation of the form \(s f = \sum g_i A_i\), for every \(s \in \mathcal{S}\), and at least for generators \(f\) of \(\mathbb{k}\). We will return to this point shortly.

One approach is to find a presentation for \(\mathbb{k}\) over a subring \(\mathbb{k}'\) which is central in \(\mathcal{A}\) (like the ground field), and to study the “corresponding presentation” of \(\mathcal{A}\) over \(\mathbb{k}'\). This is an effective approach when, for example, \(\mathbb{k}\) is a polynomial ring. This is not the approach we are interested in, as we would rather treat the ring \(\mathbb{k}\) as a black box.

Our approach is to add to \(\mathcal{S}\) a new generator \(m_f\) (multiplication by \(f\)) for each \(f \in \mathbb{k}\). Yes, we mean every element of \(\mathbb{k}\), not just a basis. Every instance of multiplication by \(f\) in a \(\mathbb{k}\)-word in a relation \(r \in \mathcal{R}\) we replace by \(m_f\).

**Example 2.16.** Relation (1.3) in the nilHecke algebra can be written for general polynomials as

\[
(2.3) \quad s_i m_f = m_{\varphi(f)} s_i + m_{\partial(f)} \text{id},
\]

where \(\varphi\) and \(\partial\) are certain maps \(\mathbb{k} \to \mathbb{k}\). Then (2.3) can be thought of as a family of relations, one for each \(f \in \mathbb{k}\).

**Remark 2.17.** If the reader prefers, when working over a field, one can also choose a basis of \(\mathbb{k}\) over this field, and only define symbols \(m_f\) for \(f\) inside this basis. One obtains an equivalent theory, requiring a similar amount of work to resolve. One philosophical disadvantage is that one loses the wholistic approach to relations like (2.3), which are encoded in terms of maps \(\varphi\) and \(\partial\) instead of particular linear combinations for each basis element.

---

2There are many such presentations: each \(f, g, \ldots\) in a \(\mathbb{k}\)-word in a relation in \(\mathcal{R}\) would have to be expressed using linear combinations of words in the generators of \(\mathbb{k}\), and this expression is not unique.

3Unlike the previous footnote, this operation is unique.
We need to add some new relations in order for this new presentation to have the desired properties. One relation we add is

\[ mfm_g \mapsto mfg \]

for each \( f, g \in \mathbb{k} \). Now, so long as there is a relation

\[ sm_f \mapsto \sum m_g A_i \]

for each \( s \in S \) and \( f \in \mathbb{k} \), only words of the form \( m_f X \) for \( X \) a true word could possibly be irreducible. We assume that such a family of relations exists.

**Remark 2.18.** One also needs to choose a partial order on monomials compatible with (2.5). When using a degree lexicographic order or similar order for the generators \( S \), one can extend this to a partial order by declaring that all symbols \( m_f \) are incomparable with each other, and lexicographically smaller than all symbols in \( S \). This results in a partial order where \( Bm_f C < A \) implies that \( Bm_g C < A \) for all \( f, g \in \mathbb{k} \). All polynomials are treated equally.

Now we need to impose a “relation” which takes \( m_f \) appearing on the left of a word, and replaces it with multiplication by \( f \) on the remainder of the word. Let us introduce new notation for words, writing \( \bullet \) for the empty word, and placing \( \bullet \) at the beginning of every word. We now introduce the “left relation”

\[ \bullet m_f \mapsto f \bullet, \]

for each \( f \in \mathbb{k} \), where the left hand side is a word, and the right hand side is a left \( \mathbb{k} \)-linear combination. This “left relation” is only permitted to be applied at the beginning of a word. With (2.6) in place, only true words can be irreducible.

**Remark 2.19.** It is easy to modify the Bergman diamond lemma to allow for left relations. In fact, this is essentially the Bergman diamond lemma for left modules over a ring; we are viewing the ring \( A \) as a left \( \mathbb{k} \)-module.

**Remark 2.20.** The relation \( m_f + m_g = m_{f+g} \) is not necessary given the relation (2.6); this relationship between elements of \( \mathbb{k} \) is supposed to be handled by taking \( \mathbb{k} \)-linear combinations of words. It is also not desirable to have \( m_f + m_g = m_{f+g} \) as a relation, as it would make the existence of a compatible partial order \( \leq \) effectively impossible. However, one is permitted to use this relation as desired when checking the ambiguities, because one can check ambiguities relative to \( \leq \), see Definition 2.11 and following. When resolving an ambiguity \( A \), if \( Bm_f C < A \) then (for reasonable partial orders, see Remark 2.18) also \( Bm_g C < A \) and \( B(m_{f+g}) C < A \), thus

\[ Bm_f C + Bm_g C - Bm_{f+g} C \in I \leftarrow A. \]

Now (assuming the existence of a suitable partial order) we can apply the original Bergman diamond lemma to this modified presentation of \( A \). Adding the relations (2.6) and (2.4) creates (at least) two new overlap ambiguities. One is \( (m_f m_g m_h, m_f m_g m_h) \), which is always resolvable by associativity in \( \mathbb{k} \). The other is \( (\bullet m_f m_g, \bullet m_f m_g) \), which is also always resolvable. Note that the resolvability of the family of ambiguities \( (sm_f m_g, sm_f m_g) \) is an interesting and nontrivial statement about the family of relations (2.5).
3. The problem with the Coxeter presentation

3.1. Depiction of the problem. The Coxeter presentation of the symmetric group $S_n$ was given in the introduction as (1.1). We have already seen in Example 2.7 and §2.2 that this presentation is Bergman-type when $n = 3$. However, it is not Bergman-type for $n \geq 4$.

Let $s = s_1, t = s_2$, and $u = s_3$ inside $S_4$. The quadratic relations must have the form $ss \mapsto 1$, $tt \mapsto 1$, $uu \mapsto 1$, as if any of these elementary resolutions goes the other way, no compatible partial order would satisfy the DCC. Suppose, for example, that the remaining relations are $stst \mapsto tsts, tut \mapsto utu$, and $su \mapsto us$. Then $(stus, stsu)$ is an unresolvable overlap ambiguity, as both $tstu$ and $stus$ are irreducible. The longest element of $S_4$ has multiple irreducible reduced expressions, namely $utustu$ and $ustus$ and $ustustu$. The reader can confirm that no other way to “orient” the relations (i.e. replacing $stst \mapsto tsts$ with $tsts \mapsto stst$) will have resolvable ambiguities either.

Let us rephrase the problem. For each element $w \in S_n$, let $\Gamma_w$ denote its reduced expression graph. This is the graph whose vertices are reduced expressions for $w$, and the edges are applications of a single braid relation, either (1.1b) or (1.1c). The edges can be labeled by the pair $\{i, j\}$ of indices which are involved in the braid relation; there are $\binom{n-1}{2}$ possible labelings. If one chooses a partial order compatible with the relations, then for each pair $\{i, j\}$ with $j \neq i \pm 1$, either $s_is_j$ is bigger than $s_js_i$, or the other way around. This gives an orientation on the reduced expression graph of $s_is_j$ (the arrow goes from the larger word to the smaller one). If the partial order is a semigroup partial order, then this same orientation applies to every edge labeled $\{i, j\}$ in any reduced expression graph. Similar considerations can be made for pairs $\{i, j\}$ with $j = i \pm 1$. There are $2\binom{n-1}{2}$ orientations on reduced expression graphs, obtained by choosing an orientation for each pair $\{i, j\}$.

Remark 3.1. The reader who has never encountered reduced expression graphs before will be greatly served by drawing the reduced expression graph of the longest element of $S_4$. There are 16 reduced expressions; pick one and then start applying braid relations to find the rest. The answer can be found, for example, in [7, §3.1].

A sink in the oriented reduced expression graph is an irreducible monomial. Thus, if the presentation is Bergman-type, then there must be an orientation where each reduced expression graph has a unique sink. As noted, the oriented reduced expression graph of the longest element of $S_4$ always has multiple sinks.

3.2. Avoidance of the problem. Now we point out the real mathematical meat of this paper: the only problem is choosing an orientation for the relation (1.1b), i.e. for pairs $\{i, j\}$ with $j \neq i \pm 1$.

Let $\Gamma_w$ denote the quotient of $\Gamma_w$ by all edges with a label $\{i, j\}$ for $j \neq i \pm 1$. That is, a vertex of $\Gamma_w$ is an equivalence class $[w]$ of reduced expressions of $w$, where reduced expressions $w$ and $w'$ are equivalent if they differ only by commuting relations like $su = us$. There is an edge from one vertex to another, labeled by $\{i, j\}$ with $j = i \pm 1$, if there is a representative of each of the two equivalence classes which is related by (1.1c).

The following is a theorem of Manin and Schechtmann, and (in this particular form) is implicit but not explicitly stated in [14]. For this reformulation, see [7, §1.3, §3.1].

**Proposition 3.2.** Place the antilexicographic orientation on $\Gamma_w$, where each edge is oriented $s_is_{i+1}s_i \mapsto s_{i+1}s_is_{i+1}$. Then $\Gamma_w$ has a unique source and a unique sink. In fact, $\Gamma_w$ is a graded graph: every vertex $[w]$ has a height $h([w]) \in \mathbb{N}$ and oriented edges increase the height by one.
We recall Manin-Schechtmann theory and prove this proposition in §3.3, see Corollary 3.16.

Remark 3.3. When \( w \) is the longest element of \( S_n \), \( \overline{\Gamma}_w \) is the Hasse diagram of the second higher Bruhat order. In [14] they define a family of higher Bruhat orders, starting with the familiar Bruhat order, then \( \overline{\Gamma}_w \), and beyond. Each higher Bruhat order is a graded graph with unique source and sink, whose vertices are equivalence classes of paths from source to sink in the previous Bruhat order.

In particular, this indicates that, if there were some way to apply a modified Bergman diamond lemma where one was allowed to apply the relation \( s_i s_j = s_j s_i \) in either direction, then there might be a chance that the Coxeter presentation is “modified Bergman-type.” This is what we accomplish in this paper.

3.3. Manin-Schechtmann theory: triples and inversions. The results in this section are due to Manin-Schechtmann. More precisely, they all effectively come from the \( k = 1 \) and \( k = 2 \) cases of [14, §2, Theorem 3] and [14, §2, Lemma 8], though this should not be obvious to the reader at first glance. Because we use crossing diagrams to represent permutations, we call indices \( 1 \leq i \leq n \) by the name strands.

Definition 3.4. Given a permutation \( w \in S_n \), and inversion of \( w \) is a pair of strands \((i|j)\) with \( w(j) > w(i) \). We always write pairs \((i|j)\) in order, with \( i < j \). We write \( I(w) \) for the inversion set of \( w \), the set of all its inversions.

Note that the length of \( w \) is the size of the inversion set. For example, the inversion set of the longest element includes all \( \binom{n}{2} \) pairs \((i|j)\).

Definition 3.5. To any given crossing in an expression (not necessarily reduced), we associate the pair \((i|j)\) of strands which is crossed. The following example, where each crossing is labeled with its pair, should make this definition clear. Note that we read a diagram from bottom to top.

(3.1)

The string diagram from (3.1) corresponds to the expression \( s_1 s_2 s_1 s_2 s_1 \), or \( ststs \) where \( s = s_1 \) and \( t = s_2 \) as before. Because diagrams are read from bottom to top, and permutations act from right to left, it is the leftmost \( s_1 \) which corresponds to the top crossing labeled \((1|3)\), and the rightmost \( s_1 \) which corresponds to the bottom crossing labeled \((1|2)\).

Let \( w \) be a reduced expression for an element \( w \in S_n \). Since the length of \( w \) agrees with the size of \( I(w) \), and since each crossing adds at most one inversion, it must be the case that every pair \((i|j)\) in \( I(w) \) labels precisely one crossing in \( w \).

Definition 3.6. Let \( w \) be a reduced expression for \( w \in S_n \). To this reduced expression we associate a total order \( \prec \) on \( I(w) \) where \((i|j) \prec (k|l)\) if the crossing labeled \((i|j)\) appears below \((k|l)\). We call this the crossing sequence of the reduced expression. This map from reduced expressions to crossing sequences is injective.
In contrast to the notation \((i|j) < (k,l)\), which depends on a choice of reduced expression \(w\), we will continue to use \((i|j) < (k|l)\) to indicate the usual lexicographic order. Manin-Schechtman theory comes in when one compares the two orders \(<\) and \(\leq\).

**Example 3.7.** In the reduced expression \(s_1s_2s_3\) in \(S_3\), the rightmost \(s\) crosses \((1|2)\), then \(t\) crosses \((1|3)\), and finally \(s\) crosses \((2|3)\). This yields the lexicographic order on \(I(s_1s_2s_3)\): \((1|2) < (1|3) < (2|3)\). Meanwhile, the reduced expression \(tst\) yields the antilexicographic order \((2|3) < (1|3) < (1|2)\).

![Diagram of strand crossings]

Not every set of pairs of strands is the inversion set of a permutation. For example, one can not cross \((1|3)\) before getting the strand \(2\) out of the way somehow. Moreover, not every total order on an inversion set corresponds to a reduced expression: there are six orders on \(\{(1|2), (1|3), (2|3)\}\), but only two correspond to reduced expressions.

**Definition 3.8.** A **triple** of strands is \((i|j|k)\) with \(1 \leq i < j < k \leq n\). The packet \(P(i|j|k)\) of a triple is the set \(((i|j) < (i|k) < (j|k))\), given its lexicographic ordering. Let \(I\) be a set of pairs of strands \(\{i|j\}\) (for example, \(I\) might be an inversion set). We call a triple \((i|j|k)\) a **prefix triple** for \(I\) if \(P(i|j|k) \cap I\) is a prefix of \(P(i|j|k)\). That is, the intersection is either the empty set, \(\{i|j\} \cap \{i|j, i|k\}\), or all of \(P(i|j|k)\). We call \((i|j|k)\) a **suffix triple** for \(I\) if \(P(i|j|k) \cap I\) is a suffix of \(P(i|j|k)\). We call \((i|j|k)\) a **full triple** for \(I\) if \(P(i|j|k) \subseteq I\). In particular, a full triple is both prefix and suffix. When \(I = I(w)\), we say instead that \((i|j|k)\) is a **prefix triple** for \(w\), etcetera.

**Example 3.9.** For the set \(I = \{(1|3)\}\), the triple \((1|2|3)\) is neither prefix nor suffix.

**Example 3.10.** For \(i < j < k\), the triple \((i|j|k)\) is full for \(w\) if and only if \(w(k) < w(j) < w(i)\).

The key point to make about triples is that, whenever a subword \(s_is_{i+1}s_i\) or \(s_{i+1}s_is_{i+1}\) appears inside an expression, it is because the packet of some triple \((i|j|k)\) appears in consecutive order inside the crossing sequence of the expression. Applying the braid relation \(s_is_{i+1}s_i \rightarrow s_{i+1}s_is_{i+1}\) to a reduced expression will flip the order induced on this packet from lexicographic to antilexicographic. On the other hand, applying a braid relation \(s_is_js_j = s(js_i)\) for \(j \neq i+1\) will flip the order on two adjacent pairs \((i|j)\) and \((k|l)\) with \(\{i, j\} \cap \{k, l\} = \emptyset\). In particular, \((i|j)\) and \((k|l)\) appear in no common packets, so that swapping their order is not detected by any triple.

We now state a number of theorems, and we strongly encourage the reader to test out these theorems on (all) the reduced expressions of the longest element of \(S_n\).

**Theorem 3.11.**

1. A set \(I\) of pairs \(\{i|j\}\) is the inversion set \(I(w)\) of some permutation \(w \in S_n\) if and only if each triple \((i|j|k)\) is either a prefix or a suffix triple for \(I\).
2. A total order \(\prec\) on the inversion set \(I(w)\) corresponds to some reduced expression \(w\) for \(w\) if and only if, for each triple \((i|j|k)\):
   - if \((i|j|k)\) is prefix for \(w\), but not full, then the ordering induced on \(I(w) \cap P(i|j|k)\) by \(\prec\) is the lexicographic ordering;
   - if \((i|j|k)\) is suffix for \(w\), but not full, then the ordering induced on \(I(w) \cap P(i|j|k)\) by \(\prec\) is the antilexicographic ordering;
• if \((i|j|k)\) is full for \(w\), then the ordering induced on \(P(i|j|k) \subset I(w)\) by \(<\) is either the lexicographic or the antilexicographic ordering.

**Definition 3.12.** For a reduced expression \(w\) of \(w\), define its higher inversion set \(J(w)\) to be the set of full triples \((i|j|k)\) for \(w\) such that \(<\) induces the antilexicographic order on \(P(i|j|k)\).

**Theorem 3.13.** (1) If \(w \rightarrow w'\) are reduced expressions related by a single braid relation \(s_is_{i+1}s_i \rightarrow s_{i+1}s_is_i\), then \(J(w) \subset J(w')\), and the only triple in \(J(w')\) not in \(J(w)\) is the triple \((i|j|k)\) of strands involved in this braid relation. Consequently, applying a braid relation is sometimes called a packet flip, or flipping the packet \(P(i|j|k)\).

(2) For reduced expressions \(w\) and \(w'\) of \(w \in S_w\), one has \(J(w) = J(w')\) if and only if \(w\) and \(w'\) are equivalent.

For part (b) above and the results below, we recall that two reduced expressions are equivalent if they are related by commuting braid relations \(s_is_j = s_js_i\) for \(j \neq i \pm 1\), and that \(\Gamma_w\) is the graph of equivalence classes of reduced expressions.

**Remark 3.14.** Manin and Schechtman also have a result indicating which sets of triples \(J\) are the higher inversion sets of a reduced expression, by examining packets of quadruples, analogous to part (1) of Theorem 3.11. This question relates to the third higher Bruhat order.

Finally we have the result which proves Proposition 3.2.

**Proposition 3.15.** Suppose that \(w\) is a reduced expression which has at least one full triple not in \(J(w)\). Then there is some full triple \((i|j|k)\) \(\notin J(w)\) whose packet is flippable, i.e. there exists a reduced expression \(w'\) equivalent to \(w\) in \(\Gamma_w\), and a braid relation \(s_is_{i+1}s_i \rightarrow s_{i+1}s_is_i\) which can be applied to \(w'\) which adds \((i|j|k)\) to the higher inversion set.

**Corollary 3.16.** The function \(h(w)\) given by the size of \(J(w)\) is invariant on equivalence classes of reduced expressions, and descends to a height function on \(\Gamma_w\). There is a unique sink in \(\Gamma_w\), whose higher inversion set is the set of full triples for \(w\). (There is also a unique source whose higher inversion set is empty, which follows from an “backwards” version of Proposition 3.15.)

The paper [14] only proves Proposition 3.15 for the longest element of \(S_n\), so for completeness we include a proof for an arbitrary permutation \(w\). Our proof is different than the one in [14], though a proof along their lines would also work.

**Proof.** Let \(w = ((i_1|j_1) < (i_2|j_2) < \cdots < (i_d|j_d))\) be a reduced expression for \(w\), for which we record its crossing sequence. Amongst the set of full triples \((i|j|k)\) for \(w\) induces the lexicographic order on the packet \(P(i|j|k)\), choose one where the subword starting with \((i|j)\) (then passing through \((i|k)\)) and ending with \((j|k)\) has minimal length. We will use induction on this minimal length to prove the result. If this length is 3, i.e. the packet appears consecutively in \(w\), then the packet can obviously be flipped. If not, then some other pair \((x|y)\) intervenes, either between \((i|j)\) and \((i|k)\), or between \((i|k)\) and \((j|k)\), or both. We will assume an intervening pair between \((i|j)\) and \((i|k)\), as the other case is treated similarly.

Suppose that \(w = (\cdots < (i|j) < (x|y) < \cdots )\), with \((x|y) \neq (i|k)\). If \(\{i, j\} \cap \{x, y\} = \emptyset\) then a commuting relation will reverse the order, yielding \(w' = (\cdots < (x|y) < (i|j) < \cdots )\).

So \(w'\) is equivalent to \(w\) and the minimal length in \(w'\) is shorter. If \(\{i, j\} \cap \{x, y\}\) is nonempty, then there are several cases.
If \( i = y \) so that \( x < i < j < k \), then consider the triple \((x|i|j)\). The pairs \((i|j)\) and \((x|i)\) appear consecutively in this order in \( w \), but \((x|i) < (i|j)\) does not appear in either the lexicographic or the antilexicographic order on \( P(x|i|j) \), a contradiction.

If \( j = x \) so that \( i < j < y \), then consider the triple \((i|j|y)\). The pairs \((i|j)\) and \((j|y)\) appear consecutively in this order, which is again a contradiction.

If \( j = y \) so that \( x < j < k \), then consider the triple \((x|j|k)\). The pairs \((x|j)\) and \((j|k)\) appear in this order (though not consecutively). This means that \((x|k)\) must be somewhere in between, and hence \((x|j|k)\) is a full triple, appearing in lexicographic order. But then this triple \((x|j|k)\) has shorter length than \((i|j|k)\), so induction implies that some packet can be flipped.

Finally, if \( i = x \), then there are several cases depending on where \( y \) falls. Recall that

\[
w = (\cdots < (i|j) < (i|y) < \cdots < (i|k) < \cdots < (j|k) < \cdots).
\]

If \( i < y < j < k \) then consider the triple \((i|y|j)\). Since \((i|j)\) appears before \((i|y)\), the pair \((y|j)\) is also in \( I(w) \) and occurs somewhere before \((i|j)\). Now consider the triple \((y|j|k)\). Since \((y|j)\) and \((j|k)\) both appear, so must \((y|k)\), at some point before \((j|k)\). Finally, consider the triple \((i|y|k)\). It is a full triple, and \((i|y)\) appears before \((i|k)\) so it must be in lexicographic order. Then the full triple \((i|y|k)\) has shorter length than \((i|j|k)\).

If \( i < j < y < k \) then consider the triple \((j|y|k)\). Since \((j|k)\) does appear, either \((j|y)\) or \((y|k)\) must appear before it. If \((j|y)\) appears then the triple \((i|j|y)\) is full, with \((i|j)\) appearing before \((i|y)\), so it is lexicographic, and has shorter length. If \((y|k)\) appears then the triple \((i|y|k)\) is full, with \((i|y)\) appearing before \((i|k)\), so it is lexicographic, and has shorter length.

If \( i < j < k < y \) then consider the triple \((i|k|y)\). Since \((i|y)\) appears before \((i|k)\), the \((k|y)\) is also in \( I(w) \) and occurs somewhere before \((i|k)\). Now consider the triple \((j|k|y)\). Since \((k|y)\) and \((j|k)\) both appear, so must \((j|y)\), at some point before \((j|k)\). Finally, consider the triple \((i|j|y)\). It is a full triple, and \((i|j)\) appears before \((i|y)\) so it must be in lexicographic order. Then the full triple \((i|j|y)\) has shorter length than \((i|j|k)\).

This concludes the enumeration of cases, and the proof that some flippable packet exists. \( \square \)

3.4. Non-reduced expressions. For each \( w \in S_n \), let \( X_w \) denote the (infinite) set of all expressions for \( w \), not necessarily reduced. Let \( \bar{X}_w \) denote the set of equivalence classes in \( X_w \), identifying \( w \) and \( w' \) if they differ only by commuting relations (1.1b). Then \( X_w \) and \( \bar{X}_w \) can be given the structure of graphs, analogous to \( \Gamma_w \) and \( \bar{\Gamma}_w \), where one also has an edge for each relation \( s_is_i \rightarrow 1 \). In fact, \( \Gamma_w \) will be a subgraph of \( X_w \), and \( \bar{\Gamma}_w \) of \( \bar{X}_w \).

There is an additional structure on the graph of all expressions, which is not present when one restricts to reduced expressions. Namely, if \( w \in X_w \) and \( y \in X_y \) then the concatenation \( w\bar{y} \) lives in \( X_{wy} \). This concatenation operation descends to equivalence classes

\[
\bar{X}_w \times \bar{X}_y \rightarrow \bar{X}_{wy}.
\]

There is one more operation, which we could have discussed previously for the graphs \( \Gamma_w \) and \( \bar{\Gamma}_w \). For each \( 1 \leq n \leq m \) and \( 0 \leq k \leq m - n \) there is an inclusion \( \iota = \iota_{m-n:k} : S_n \rightarrow S_m \), via the maps \( S_n \rightarrow S_k \times S_n \times S_{m-k} \rightarrow S_m \). That is, \( \iota \) sends \( s_i \) to \( s_{i+k} \) for each \( 1 \leq i \leq m - 1 \). This map \( \iota \) induces inclusions of graphs \( X_w \rightarrow X_{\iota(w)} \), which also descends to equivalence classes.

**Definition 3.17.** Place an orientation on \( \bar{X}_w \) for each \( w \), where the edges are oriented \( s_is_{i+1}s_i \rightarrow s_{i+1}s_is_{i+1} \) and \( s_is_i \rightarrow 1 \).
This is the only reasonable orientation extending Manin-Schechtmann’s orientation on $\tilde{\Gamma}_w$. We will prove later in this section that $\bar{X}_w$ has a unique sink.

Note that this orientation gives $\bar{X}_w$ the structure of a poset, where a sink is a minimal element. We now define a preorder on $X_w$, which will descend to a partial order on $\bar{X}_w$, and which is stronger than the one induced by the orientation. This partial order will be one of the inputs to our generalized Bergman diamond lemma for Hecke-type algebras. Recall the partial order $\leq_3$ on words in $S_3$, built from the generators $s$ and $t$, given in Definition 2.12.

**Notation 3.18.** If $w$ is an expression for $w \in S_n$, then for any pair $(i|j)$ of strands, let $n_{ij}$ denote the number of $(i|j)$ crossings in $w$. Clearly $\sum_{1 \leq i < j \leq n} n_{ij}$ is the length of $w$. If $w'$ is another expression, we write $n_{ij}'$ for the same statistic.

**Lemma 3.19.** If $w$ and $w'$ are words in $S_3$ and have the same length, and they are comparable in $\leq_3$, then $n_{12} = n_{12}'$, $n_{13} = n_{13}'$, and $n_{23} = n_{23}'$.

**Proof.** This is obvious, as braid relations never change the number of crossings of each type $(i|j)$.

**Definition 3.20.** If $w$ is an expression and $(i|j|k)$ is a triple, let $w_{ijk}$ denote the expression inside $S_3$ obtained by tracing out the triple of strands $(i|j|k)$. An example is pictured below inside $S_8$, where the irrelevant strands are drawn without color:

For two expressions $w$ and $w'$ in $X_w$, we say that $w < w'$ if $w$ is shorter than $w'$, and that $w \leq w'$ if they have the same length and if $w_{ijk} \leq_3 w_{ijk}'$ for all triples $(i|j|k)$.

Clearly $\leq$ is a transitive and reflexive on $X_w$, so it is a preorder.

**Remark 3.21.** If desired, one can also extend this relation to the entire set $X$ of all expressions for any element. If $w$ and $y$ are expressions for different elements, then $w < y$ if and only if $w$ is shorter. In this section, we consider each $X_w$ separately.

**Lemma 3.22.** The preorder $\leq$ on $X_w$ is compatible with concatenation in that, for any words $A$ and $B$, $w \leq w'$ if and only if $AwB \leq Aw'B$. It is compatible with inclusions in that, for any $n, m, k$ such that $i = i_k^{n-m}$ is defined, $w \leq w'$ implies $\iota(w) \leq \iota(w')$. Finally, $\leq$ satisfies the DCC.

**Proof.** We check compatibility with concatenation, as the remaining properties are a simple exercise. We claim that this follows from the fact that the partial order $\leq_3$ on words in $S_3$ is compatible with concatenation. For any $i < j < k$, tracing three strands through the composition $AwB$ is the composition of tracing three strands through $B$, then through $w$, then through $A$, although the labels on these strands change. More precisely, if $w$ expresses $w \in S_m$ then

\begin{equation}
(AwB)_{ijk} = A_{i''j''k''}w_{i'j'k'}B_{ijk},
\end{equation}

where $B((i,j,k)) = \{i',j',k'\}$ and $w((i',j',k')) = \{i'',j'',k''\}$. If $w \leq w'$ and both have the same length, then they express the same element $w$, so that $(AwB)_{ijk}$ and $(Aw'B)_{ijk}$
both obey (3.2) with the same primed indices. Then \((AwB)_{ijk} \leq (Aw'B)_{ijk}\) follows because 
\[w_{ij'k'} \leq w'_{ij'k'}\]

**Lemma 3.23.** For \(w, w' \in X_w\), we have \(w \leq w' \leq w\) if and only if 
\[w_{ijk} = w'_{ijk}\] for all triples \((i|j|k)\).

If so, \(w\) and \(w'\) have the same length, and 
\[n_{ij} = n'_{ij}\] for all pairs \((i|j)\).

**Proof.** Most of this lemma is obvious, since \(\leq_3\) is a partial order so 
\(A \leq_3 B \leq_3 A\) implies \(A = B\). The only thing to show is that if 
\[w_{ijk} = w'_{ijk}\] for all \((i|j|k)\) then \(w\) and \(w'\) have the same length.

If \[w_{ijk} = w'_{ijk}\] then \[n_{ij} = n'_{ij}\]. Summing over all pairs \((i|j)\), we see that they 
have the same overall length. \(\square\)

We also prove, but do not use, the following observation.

**Lemma 3.24.** If \[w_{ijk} \leq_3 w'_{ijk}\] for all triples \((i|j|k)\), then the length of \(w\) is at most the length of \(w'\).

**Proof.** Suppose \(w\) is a permutation in \(S_m\). Then the statement is obvious for \(m \leq 3\).

Because \[w_{ijk} \leq_3 w'_{ijk}\], either \(w_{ijk}\) has shorter length and
\[n_{ij} + n_{ik} + n_{jk} < n'_{ij} + n'_{ik} + n'_{jk}\]
or \(w_{ijk}\) and \(w'_{ijk}\) have the same length, and
\[n_{ij} = n'_{ij}, \quad n_{ik} = n'_{ik}, \quad n_{jk} = n'_{jk}\]
If (3.4) holds for all triples \((i|j|k)\) then clearly the lengths of \(w\) and \(w'\) are equal. So suppose (3.3) holds for some triple. Taking the sum over all triples, we obtain
\[\sum_{(i|j)k} n_{ij} + n_{ik} + n_{jk} < \sum_{(i|j)k} n'_{ij} + n'_{ik} + n'_{jk}\]
However, the left hand side is precisely \((m - 2) \sum_{i|j} n_{ij}\), which is \((m - 2)\) times the length of \(w\). Thus the length of \(w\) is smaller than the length of \(w'\). \(\square\)

**Proposition 3.25.** One has \(w \leq w' \leq w\) if and only if these two expressions are equivalent in \(\bar{X}_w\). Therefore, \(\leq\) descends to a partial order on \(\bar{X}_w\).

**Proof.** One direction is easy. The commuting braid relations \(s_is_j = s_js_i\) are not detected by any triples, and do not change the length of the expression. Thus two expressions equivalent in \(\bar{X}_w\) are equivalent under \(\leq\). We now prove the other direction. Let \(w \leq w' \leq w\). By Lemma 3.23, we know that \[n_{ij} = n'_{ij}\] for all \((i|j)\), and \[w_{ijk} = w'_{ijk}\] for all \((i|j|k)\).

Suppose that \(s_i\) is the first simple reflection in \(w'\) (i.e. it appears on the far right of the word, reading right to left, or on the bottom of a crossing diagram). Clearly \(s_i\) appears somewhere in \(w\), or the crossing \((i|i + 1)\) could never happen. We claim that an instance of \(s_i\) appears before any \(s_{i+1}\) or \(s_{i-1}\) in \(w\). Suppose that \(s_{i+1}\) appears before the first \(s_i\) in \(w\). Before the first instance of \(s_{i+1}\), neither \(s_i\) nor \(s_{i+1}\) appears, so the strands \(\{i + 2, i + 3, \ldots, m\}\) are permuted amongst themselves (where \(w \in S_m\)), and strand \(i + 1\) is fixed. Therefore, the first \(s_{i+1}\) has crossing type \((i + 1, j)\) for some \(j \in \{i + 2, \ldots, m\}\). But then a crossing \((i + 1, j)\) appears before \((i|i + 1)\). This contradicts the fact that \[w_{(i|i+1)j} = w'_{(i|i+1)j}\]. A similar contradiction arises if \(s_{i-1}\) appears before \(s_i\).

Thus, a sequence of commuting braid relations \(s_is_j = s_js_i\) will send \(s_i\) to the start of \(w\). Up to equivalence, we can assume \(w\) also started with \(s_i\). Removing this index \(s_i\) from the start of both expressions, the inequality \(\leq\) still holds by Lemma 3.22. We can then use induction on the length of expressions to prove that \(w\) and \(w'\) are equivalent in \(\bar{X}_w\). \(\square\)

**Lemma 3.26.** The partial order \(\leq\) on \(\bar{X}_w\) respects the orientation of Definition 3.17.
Proof. Applying the relation $ss \rightarrow 1$ will shorten the length of a word. Applying the relation $s_\ell s_{\ell+1} s_\ell \rightarrow s_{\ell+1} s_\ell s_{\ell+1}$ will affect precisely one triple $(i|j|k)$, and will send it to something smaller in $\leq 3$. □

**Theorem 3.27.** The oriented graph $\bar{X}_w$ has a unique sink, which is a reduced expression.

**Proof.** Suppose that the equivalence class $[w]$ of $w$ is a sink. If $w$ is a reduced expression, then $[w]$ must be the unique sink of $\Gamma_w$ by Proposition 3.2. Thus we suppose that $w$ is not reduced, and derive a contradiction using the following lemma. □

**Lemma 3.28.** Suppose that $\underline{w} \in X_w$ is not reduced. Then $[\underline{w}]$ is not a sink.

**Proof.** We show this by induction on the length of $\underline{w}$. The base case is trivial. If any proper subword of $\underline{w}$ is not a reduced expression, then by induction it is not a sink. Hence (using the compatibility of the orientation with concatenation), $\underline{w}$ is also not a sink. Thus we can assume each subword of $\underline{w}$ is a reduced expression. Moreover, using Proposition 3.2, we can assume that each subword of $\underline{w}$ is (a representative of the equivalence class of) the unique sink in its reduced expression graph.

So suppose $\underline{w} = s_i \underline{y} s_j$, for some reduced expression $\underline{y}$ for $y \in S_m$. Both $s_i \underline{y}$ and $\underline{y} s_j$ are reduced. By the exchange condition for the symmetric group (see [11, §1.7]), we see that $s_i$ and $s_j$ must cross the same pair of strands $(j|j+1)$. In particular, $y(j) = i$ and $y(j+1) = i+1$. Therefore, $y s_j = s_i y$ and $y = w$.

\begin{center}
\begin{tikzpicture}
\draw (0,0) -- (0,4);
\draw (1,0) -- (1,4);
\draw (2,0) -- (2,4);
\draw (3,0) -- (3,4);
\draw (4,0) -- (4,4);
\draw (5,0) -- (5,4);
\draw (3,1) -- (3,3);
\draw (4,1) -- (4,3);
\node at (1,2) {\underline{s_i}};
\node at (3,2) {\underline{y}};
\node at (5,2) {\underline{s_j}};
\end{tikzpicture}
\end{center}

(The reader who prefers a non-trivial base case will content themself with the case where $\underline{y}$ is the empty word, and $i = j$.)

Suppose that $k > j+1$. We claim that $y(k) > i+1$. Otherwise, $y(k) < i$ and the triple $(j|j+1|k)$ is full in $y s_j$. However, the reduced expression $y s_j$ is the sink in its reduced expression graph, so by Corollary 3.16 the packet of $(j|j+1|k)$ must appear in antilexicographic order. However, $(j|j+1)$ appears first, a contradiction.

Suppose that $k < j$. We claim that $y(k) < i$. Otherwise, $y(k) > i+1$ and the triple $(k|j+1)$ is full in $s_i y$. However, the reduced expression $s_i y$ is the sink in its reduced expression graph, so by Corollary 3.16 the packet of $(k|j+1)$ must appear in antilexicographic order. However, $(j|j+1)$ appears last, a contradiction.

By counting, we see that the number of strands to the right of $j+1$ (resp. the left of $j$) must equal the number of strands to the right of $i+1$ (resp. the left of $i$). Thus $i = j$. Then $y$ is some permutation in $S_{j-1} \times S_1 \times S_1 \times S_{n-i-1} \subset S_m$. Every simple reflection $s_k$ in the reduced expression $y$ satisfies $k \neq i \pm 1$, and thus commuting braid relations will bring $\underline{w}$ to the form $y s_i s_i$.

Finally, the oriented edge $s_i s_i \rightarrow 1$ indicates that $\underline{w}$ is not a sink. □

**Remark 3.29.** The proof of this lemma actually gives an algorithm for taking a non-reduced expression and shortening it, using only the moves $s_i s_{i+1} s_i \rightarrow s_{i+1} s_i s_{i+1}$, $s_i s_j = s_j s_i$, and $s_i s_i \rightarrow 1$. Namely, look at some non-reduced subword $\underline{w} = s_i \underline{y} s_j$ where $s_i \underline{y}$ and $\underline{y} s_j$ are...
reduced. Shorten $y$ as much as possible by using commuting braid relations to bring simple reflections to the left past $s_i$ or the right past $s_j$. If $y$ is empty, then $i = j$ and $s_i s_i \mapsto 1$ suffices. Otherwise, either some $k > j + 1$ has $y(k) < i$, in which some packet of $ys_j$ is flippable, or some $k < j$ has $y(k) > i + 1$, in which case some packet of $s_j y$ is flippable. Repeat this operation for the win.

4. THE DIAMOND LEMMA FOR MONOIDAL CATEGORIES

A natural setting wherein the set of expressions modulo equivalence $\bar{X}_w$ is more natural than the set of expressions $X_w$ is the setting of monoidal categories. This shall be explained further below. Thus, we seek a diamond lemma for monoidal categories.

It would be very surprising if this had not been done before, but we could not find it. See [5] for another extension of Gröbner bases to a more complicated setup.

Addendum. Diamond lemmas for monoidal categories have appeared in the literature, see the references in [6].

4.1. Monoidal categories. We turn our attention to $\mathbb{k}$-linear (strict) monoidal categories. These are equipped with an additional structure, horizontal composition $\otimes$, which is also $\mathbb{k}$-bilinear like ordinary composition (which we call vertical composition). String diagrams are an effective tool for describing elements of a strict monoidal category. See [13, 15] for more background.

The compatibility laws between horizontal and vertical multiplication in a monoidal category state that, when $f : n \to n'$ and $g : m \to m'$, one has $(f \otimes \text{id}_{m'}) \circ (\text{id}_n \otimes g) = (f \otimes g) = (\text{id}_{n'} \otimes g) \circ (f \otimes \text{id}_{m})$ as maps $n \otimes m \to n' \otimes m'$. The effect of this law on string diagrams is to state that two string diagrams which are equivalent under rectilinear isotopy represent morphisms which are equal.

$$f = g$$

One can present a monoidal algebra by generators and relations. One thinks of the generating morphisms as elementary string diagrams, and relations as local rules which can be applied to any subdiagram of a diagram.

Example 4.1. The category $\text{Sym}$ from Example 2.15 is monoidal. Horizontal composition is induced by the map $S_n \times S_m \to S_{n+m}$, which gives a map $\mathbb{k}[S_n] \otimes \mathbb{k}[S_m] \to \mathbb{k}[S_{n+m}]$.

Its presentation as a monoidal category is much simpler than its presentation as a $\mathbb{k}$-linear category. There is one generating object $1 \in \mathbb{N}$, and one generating morphism $\sigma \in \text{End}(2)$, drawn as a crossing. There are two relations, corresponding to (1.1a) and (1.1c): $\sigma \sigma = \text{id}_2$, and $(\sigma \otimes \text{id}_1) \circ (\text{id}_1 \otimes \sigma) \circ (\sigma \otimes \text{id}_1) = (\text{id}_1 \otimes \sigma) \circ (\sigma \otimes \text{id}_1) \circ (\text{id}_1 \otimes \sigma)$ in $\text{End}(3)$. These relations are drawn as follows.

\begin{equation}
(4.1)
\end{equation}

By convention, when discussing a morphism in $\text{End}(2)$ we write $\sigma$, but when discussing a morphism in $\text{End}(3)$ we write $s = \sigma \otimes \text{id}_1$ and $t = \text{id}_1 \otimes \sigma$, so that the second relation is $sts = tst$. We also write $s, t, u$ for the diagrams corresponding to the three simple reflections in $S_4$, viewed as elements of $\text{End}(4)$. For example, the equality $su = us$ is a consequence of rectilinear isotopy, so is not needed as a relation.
A presentation of a monoidal category does give rise to a presentation of its endomorphism algebras as ordinary categories. For instance, one can recover the Coxeter presentation of $S_n$ from the presentation above. The point of this discussion is that, for many monoidal categories, the monoidal presentation is adapted to a variant of the diamond lemma, while the ordinary presentation is not adapted to the original diamond lemma.

4.2. The diamond lemma for monoidal presentations. We seek a version of Bergman’s diamond lemma which functions for monoidal presentations of $k$-linear monoidal categories. Thankfully, Bergman’s original proof functions almost verbatim once one changes the notation, replacing “linear” descriptions of concatenation with “planar” ones. One replaces monomials with isotopy classes of diagrams, and replaces concatenation (the composition of words) with plugging diagrams into each other (the composition of diagrams).

Let $C$ be a strict monoidal category with generating objects $N$, generating morphisms $S$ and local relations $R$. In particular, the objects of $C$ are words in $N$, but the objects will not play a major role in this discussion. Let $X$ denote the set of all diagrams built from the generating morphisms $S$, and $X(n,m)$ the diagrams with source $n$ and target $m$, for words $n,m$ in the alphabet $N$. Let $\bar{X}$ denote the quotient of $X$ by the equivalence relation given by rectilinear isotopy.

Let us fix a partial order $\leq$ on $\bar{X}$. We assume that $X(n,m)$ and $X(n',m')$ are incomparable unless $n = n'$ and $m = m'$, i.e. the partial order respects morphism spaces. It is a monoidal partial ordering if the following condition holds: if $\varphi$ is obtained by plugging $B$ into a larger diagram, and $\psi$ is obtained by plugging $B'$ into the same larger diagram, and $B < B'$, then $\varphi < \psi$. This is equivalent to saying that $\leq$ respects vertical concatenation on either top or bottom (c.f. the previous notion of a semigroup partial order), and respects horizontal concatenation on either left or right with the identity map of any object. We say that $\leq$ is compatible with $R$ if every relation $r \in R$ can be rewritten as an equation $W_r = f_r$, where $W_r \in \bar{X}$, and $f_r$ is a $k$-linear combination of various $A \in \bar{X}$ satisfying $A < W_r$. In the statement of the diamond lemma, we also need that $\leq$ satisfies the DCC.

Given such a partial order, we can define elementary resolutions, resolutions, and irreducible elements of $\bar{X}$, just as in §2.1. We can also define ambiguities $(A,W_r,W_s)$, which appear when a diagram $A$ has two distinct subdiagrams $W_r$ and $W_s$ for $r,s \in R$, or the same subdiagram when $W_r = W_s$ and $r \neq s$. This leads to two possible elementary resolutions of $A$, and the ambiguity is resolvable if they can be jointly resolved to the same linear combination of isotopy classes of diagrams. One possibility is that the subdiagrams $W_r$ and $W_s$ are disjoint, meaning that they do not intersect in the plane. As before, disjoint ambiguities can be easily jointly resolved.

There are a vast number of ways in which non-disjoint ambiguities can overlap! After all, there are many interesting ways in which two simply-connected domains in the plane can intersect (as in the picture below), and there is some crazy monoidal category with an ambiguity of each kind.

It is not worth classifying how diagrams can meet (as we did in the linear case, e.g. overlap vs. inclusion), as every overlap will be treated in an identical way, regardless of how complicated the picture is.
However, there is a major complication that comes with planar diagrams rather than linear words. In the original setting there are certain minimal examples of each ambiguity, which give rise to all possible ambiguities. For example, when $W_r = BC$ and $W_s = CD$ overlap, then the minimal ambiguity $BCD$ appears inside every other ambiguity $LBCDM$. Consequently, if the minimal ambiguity is resolvable, then so is any overlap ambiguity of this type. For general monoidal categories, it is far less easy to find a collection of minimal ambiguities which account for every possible ambiguity! In the picture above, any diagram with the appropriate boundary could go in the “gaps” between the subdiagrams $W_r$ and $W_s$. Let us give a relevant example which illustrates the difficulty.

**Example 4.2.** Consider the presentation of $Sym$ in Example 4.1, and suppose we plan to apply the relations in the direction $\sigma \sigma \mapsto 1$ and $sts \mapsto tst$. There is the potential that $sts$ overlaps with $sts$ in the bold copies of $s$, but the ways in which this overlap can occur can be complicated, because the “loose ends” in the relations can interact with other diagrams.

(4.2)

In this picture, $B$ can be any diagram with the appropriate boundary. There is no single minimal diagram which contains both copies of $sts$, and is contained in any such ambiguity $A$! Nonetheless, below we will find two “minimal ambiguities” associated with this overlap which, if resolvable, imply that any such overlap is resolvable.

**Example 4.3.** Sometimes life is not so bad. Any ambiguity of the form $(A, ss, sts)$, where the bold copies of $s$ overlap, will contain the following subdiagram, because all loose ends are accounted for.

This ambiguity $(ssts, ss, sts)$ is therefore a minimal such ambiguity.

We say that a family $\{(A_i, W_r, W_s)\}$ is a minimal set of ambiguities if, when they are all resolvable, every ambiguity is resolvable. The word “minimal” should be taken with a grain of salt: for example, one could choose the set of all ambiguities. Typically, though, one can do better.

**Theorem 4.4** (Diamond Lemma for monoidal categories). Let $C$ be a $k$-linear monoidal category for a commutative base ring $k$, presented by generators and relations. Suppose that one has a partial order $\leq$, respecting the morphism spaces, which is a monoidal ordering, is compatible with $R$, and satisfies the DCC. Then the irreducible diagrams $X_{irr}$ form a basis for $C$ if and only if each ambiguity is resolvable, if and only if each ambiguity is resolvable relative to $\leq$. If there is a minimal set of ambiguities for each type of ambiguity, then these conditions are equivalent to all the minimal ambiguities being resolvable, or resolvable relative to $\leq$. If these conditions are true, we say that the data of the presentation and the partial order is monoidal Bergman type.
Proof. The proof of [1, Theorem 1.2] translates almost verbatim to this setting, replacing $X$ with $\bar{X}$, words with diagrams, etcetera.

The main difference between Theorems 2.8 and 4.4 is the lack of any automatic and reasonably small minimal set of ambiguities. This makes the adaptation to any given setting somewhat trickier than the original Bergman Diamond Lemma, which can be applied almost without thinking (although with the caveat that one must first find a partial order $\leq$). The point of the diamond lemma is that it enables one to check if $\bar{X}_{\text{irr}}$ is a basis with only a finite amount of work, and until one finds finitely many minimal ambiguities, the advantage of the diamond lemma is not really present. So there is one additional piece of work which needs to be done for a monoidal presentation: find a reasonably small minimal set of ambiguities for each possible overlap.

5. Applications to Hecke-type categories

Let us give a concrete application of Theorem 4.4, where a small minimal set of ambiguities is found.

5.1. Applications to the symmetric group.

Lemma 5.1. For the presentation of $S_{\text{Sym}}$ in Example 4.1, the list below gives a minimal set of ambiguities.

(5.1)

Proof. Consider an ambiguity $(A, B, C)$. Obviously, if a proper subdiagram $A'$ of $A$ contains both $B$ and $C$, then using the fact that $A' < A$ (because it has fewer crossings) we see that the difference between the two elementary resolutions lies within $I_{<A}$. As discussed in Example 4.3, whenever the ambiguity has $B = ss$ and $C = stss$, the corresponding ambiguity $sstss$ in (5.1) is a proper subdiagram of $A$. The same is true for the ambiguities $s\sigma\sigma\sigma$ and $stss$. All that remains is to discuss the ambiguities of the form in (4.2).

Let us rewrite such an ambiguity as follows.

(5.2)

Up to taking subdiagrams, we can assume that what appears in regions $X$ and $Y$ is trivial, and what appears in region $Z$ is an expression for a permutation in $S_k$ for some $k$. Within the ideal $I_{<A}$ we can apply any known reductions to $Z$. More precisely, each diagram in the two elementary resolutions of $A$ will be strictly less than $A$, so that applying relations to the copy
of $Z$ in each such diagram is adding an element of $I_{< A}$. Thus, we can assume inductively that $Z$ is a reduced expression, and even a reduced expression of our choice!

Any permutation in $S_k$ has a reduced expression where the first simple reflection $s_1$ appears either once or not at all. That is, $S_k = S_{k-1} \cup S_{k-1} s_1 S_{k-1}$, where $S_{k-1}$ refers to the permutations of all but the first strand. Any permutation in $S_{k-1}$ may be assumed to be in region $X$ or $Y$; that is, up to taking subdiagrams, we can assume that $Z$ is either the identity or is $s_1$. This leads to the two remaining minimal ambiguities in (5.1), $sts$ and $stsuts$.

Remark 5.2. To be very clear, not every diagram with two overlapping copies of $sts$ will have $stsuts$ as a subdiagram. After all, $Z$ can be any expression in $S_k$. The point is that, if the ambiguities of (5.1) are resolvable, then so is any other ambiguity.

Of course, the ambiguities in (5.1) are resolvable. Because it will be useful to keep in mind, we demonstrate the resolvability of $stsuts$.

Thus, one follows the graph $\Gamma_{stsuts}$ from source to sink along its two different paths, and confirms that the answer is the same.

However, to apply Theorem 4.4, we still need the existence of a suitable partial order $\leq$. We have already constructed this order in Definition 3.20, and proven that $\bar{X}_{\text{irr}} = \bigsqcup \bar{X}_{w, \text{irr}}$ has precisely one sink for each $w \in S_n$. Thus we have the following very unsurprising theorem.

**Theorem 5.3.** The Coxeter presentation is actually a presentation of $\text{Sym}$ as claimed. That is, $\text{End}(n)$ has a basis $\{w\}$ indexed by $w \in S_n$, obtained by taking an arbitrary reduced expression $w$ for each $w$.

**Proof.** Theorem 4.4 implies that one has a basis $\{w\}$, where $w$ is a very particular reduced expression for each $w$, namely the sink of $\bar{X}_w$. However, from here it is easy to see that any arbitrary reduced expression will do. \qed

5.2. **Applications to modified symmetric categories.** This same exact proof generalizes to a broader class of presentations.
Definition 5.4. Consider a $k$-linear monoidal category $C$ with the following type of presentation, which we call a modified symmetric category. It has one generating object $1$, and one generating morphism $\sigma \in \text{End}(2)$, drawn as a crossing. The relations are:

\begin{align}
(5.4a) \quad \sigma \sigma &= \alpha \sigma + \beta \text{id}_2 \quad \text{for some } \alpha, \beta \in k. \\
(5.4b) \quad stst &= tst + ast + bts + cs + dt + e \text{id}_3 \quad \text{for some } a, b, c, d, e \in k.
\end{align}

Here, $s = \sigma \otimes \text{id}_1$ and $t = \text{id}_1 \otimes \sigma$, both in $\text{End}(3)$. Using diagrams, these are:

\begin{align}
(5.4c) \quad \sigma \sigma &= \alpha + \beta, \\
(5.4d) \quad t &= \alpha + b + c + d + e.
\end{align}

Proposition 5.5. Let $C$ be a modified symmetric category. Then $C$ has a basis indexed by $w \in S_n$, obtained by taking an arbitrary reduced expression for each $w$, if and only if each of the ambiguities in (5.1) is resolvable.

Proof. The only thing which needs proving is the existence of a suitable partial order, with the correct list of irreducible diagrams. The partial order $\leq$ of Definition 3.20, modified as in Remark 3.21, will clearly suffice, as each relation is identical to the Coxeter relations modulo words of shorter length. \qed

Let us examine what checking the ambiguities entails.

The ambiguity $\sigma \sigma \sigma$ is always resolvable, as both elementary resolutions yield $\alpha \sigma \sigma + \beta \sigma$.

Checking the ambiguity $stst$ directly is a chore. One can resolve it into linear combinations of $\{1, s, t, st, ts, tst\}$ in two ways. Let us examine the coefficient of $tst$ which occurs (ignoring any terms of shorter length).

\begin{align}
(5.5a) \quad ssts \mapsto \alpha stst \mapsto \alpha tst. \\
(5.5b) \quad ssts \mapsto stst + bsts \mapsto tstt + btst + btst \mapsto (\alpha + 2b)tst.
\end{align}

Thus, for (5.5) to agree, one requires $2b = 0$. Let us assume that $k$ has no 2-torsion for the rest of the calculation, so that $b = 0$.

Checking the $ts$ coefficient in the same way yields the equality $\beta + \alpha b = \beta + b^2$, which holds when $b = 0$.

Similarly, checking the $tst$ coefficient in the ambiguity $stst$ yields $2a = 0$. In general, checking the $w$ coefficient of $stst$ is the same as checking the $w^{-1}$ coefficient of $ssts$, after swapping $a$ and $b$, and swapping $c$ and $d$ (this is clear from the word-reversing symmetry). So we set $a = 0$ for the rest of the calculation.

Let us now check the $st$ coefficient of $ssts$, assuming that $a = b = 0$.

\begin{align}
(5.6a) \quad ssts \mapsto \alpha sts \mapsto 0. \\
(5.6b) \quad ssts \mapsto stst + dst \mapsto tstt + cst + dst \mapsto (c + d)st.
\end{align}

Thus the agreement of (5.6) implies $c + d = 0$.

We leave the reader to verify that the $s$ coefficient of $ssts$ yields the equality $e = 0$, as does the $t$ coefficient. The 1 coefficient yields the equality $\beta(c + d) = 0$, which is nothing new.
In conclusion, the ambiguities $ssts$ and $stss$ are resolvable if and only if $a = b = e = c + d = 0$, which we assume henceforth. We leave the reader to confirm that the ambiguity $ststs$ is resolvable under these assumptions.

To check if the ambiguity $stsuts$ is resolvable, one follows the two different paths from source to sink in $\Gamma_{stsuts}$. This time, lower terms are produced by (5.4b), and one needs to check if the lower terms agree. We leave the reader to verify that the lower terms in (5.3a) are

$$$(5.7a) \quad csuts + dtuts + ctsts + dtsus + ctusu + ctst + dusu,$$$$while the lower terms in (5.3b) are

$$$(5.7b) \quad cstus + dstut + ctsst + dsust + cusut + dutut + cutst + dusu.$$$$Resolving each of these expressions further, taking the difference, and using $d = -c$, one eventually obtains

$$$(5.7c) \quad c\beta(st - ts + tu - ut).$$$

Thus $stsuts$ is resolvable (given the assumptions above) if and only if $c\beta = 0$.

**Theorem 5.6.** Assume that $k$ has no 2-torsion. Then a modified symmetric category $C$ with the partial order $\leq$ is monoidal Bergman type if and only if $a = b = e = c + d = 0$ and $c\beta = 0$.

**Proof.** This follows from Proposition 5.5. The computations above checked all the ambiguities of (5.1). \qed

In particular, there are very few isomorphism classes of Bergman type modified symmetric categories.

5.3. **Applications to Hecke-type categories.** Once more, the same proof as in §5.1, combined with the observations of §2.5, generalizes easily to a much broader class of categories. Checking the ambiguities is more of a chore, and will be done (under some minor additional assumptions) in [8], but we should emphasize that there are many more possibilities than in §5.2.

The following definition encapsulates many categories in the literature, like Khovanov-Lauda-Rouquier algebras [12, 16] and Webster algebras [19].

**Definition 5.7.** A Hecke-type category is a strict monoidal category $C$ with the following type of presentation. The generating objects are a set $I$, so that objects are words $i$ in the alphabet $I$. We think of the elements of $I$ as colors. The morphisms are generated by:

- For each $i \in I$, some commutative ring $R_i$ of maps in $\text{End}(i)$. We think of these maps as “dots,” analogous to the dots in the Khovanov-Lauda calculus. Here is the picture for multiplication by $f \in R_i$, when $i$ is the color red.

$$$(5.8) \quad \begin{array}{c} j \\ f \\ \end{array}$$$

- For some subset $C$ of ordered pairs $(i,j) \in I \times I$, there is a crossing $\sigma_{ij} : ij \to ji$. Here is the picture, when $i$ is red and $j$ is blue.

$$$(5.9) \quad \begin{array}{c} i \\ j \\ \end{array}$$$
If \((i, j)\) is not in the subset \(C\), then \(\sigma_{ij}\) is not a generator, and we say that crossing \(i\) over \(j\) is not permitted.

With these generators, a diagram is a crossing diagram, where the strands are each colored by an element of \(I\), and only certain colors are permitted to cross. In addition, the diagram is decorated with dots in various places; here is an example.

We refer to a diagram with the dots removed as the underlying crossing diagram or underlying expression.

The relations will be discussed below.

**Remark 5.8.** Note that we have not decided to give a presentation for the rings \(R_i\). Instead, we prefer to think of the rings \(R_i\) as a black box, along the lines of the discussion in §2.5.

Let \(i\) and \(j\) be two sequences of colors of length \(n\). We call \(w \in S_n\) a permissible permutation if \(w(i) = j\), and the crossings in a reduced expression of \(w\) are all permissible. This does not depend on the reduced expression, only on the inversion set \(I(w)\); the colors which cross in an expression form a subset of \(I \times I\) which is not changed by any braid relation. More generally, we refer to a (non-reduced) expression of a permutation as permissible if the crossings are all permissible. It is clear that every diagram representing a morphism from \(i\) to \(j\) has underlying crossing diagram given by a permissible expression, and every permissible expression gives such a diagram.

**Example 5.9.** Consider diagrams \(ij \to ij\) with \(i \neq j\). Only the identity is a permissible permutation. If either the crossing \(\sigma_{ij}\) or the crossing \(\sigma_{ji}\) is not permitted, then only the empty word is a permissible expression. If both are permitted, then the permissible expressions are \(\sigma^k\) for \(k\) even, where \(\sigma\) is the generator of \(S_2\).

For any sequence \(i\), the commutative ring \(R_i = R_{i_1} \boxtimes R_{i_2} \boxtimes \cdots \boxtimes R_{i_d}\) is a subring of \(\text{End}(i)\), given by decorations of the identity crossing diagram. We tend to use the rectangle notation from (5.8), rather than the dot notation, to more efficiently encode an element of \(R_i\) (see the examples below). We are interested in imposing relations upon these generators, such that the morphism spaces \(\text{Hom}(i, j)\) are free as \(R_j\)-modules on the left (and as \(R_i\)-modules on the right), with a basis given a fixed reduced expression for each permissible permutation. For this to occur, we need certain relations: a quadratic relation like (5.4a), a braid relation like (5.4b) (that is, we need the Coxeter relations modulo shorter expressions), and a dot-moving relation like (2.5). However, in each of these relations, the coefficients that appear (like \(a, b, c, d, e\) in (5.4b)) can be elements of \(R_j\). Moreover, only those shorter expressions which are permissible can appear in such a relation.

**Definition** (Definition 5.7 Continued). The relations in a Hecke-type category have the following form. Once the source and target of a morphism are understood, we may write \(s\) and \(t\) for the corresponding elements of \(S_3\), viewed as crossing diagrams (independent of color). In the drawings below, \(i\) is red, \(j\) is blue, and \(k\) is green.

If \(\sigma_{ii}\) is permitted then one has the following relation in \(\text{End}(ii)\).

\[
(5.10a) \quad \sigma_{ii} \sigma_{ii} = \alpha_i \sigma_{ii} + \beta_i \text{id}_{ii}
\]
for some $\alpha_i, \beta_i \in R_{ii}$.
If $\sigma_{ij}$ and $\sigma_{ji}$ are permitted then one has the following relation in $\text{End}(ji)$.

\begin{equation}
\sigma_{ij} \sigma_{ji} = Q_{ji} \text{id}_{ji}
\end{equation}

for some $Q_{ji} \in R_{ji}$.
If $\sigma_{ii}$ is permitted then one has the following relation in $\text{End}(iii)$.

\begin{equation}
sts = \lambda_{iii} tst + a_i st + b_i ts + c_i s + d_i t + e_i
\end{equation}

for some $a_i, b_i, c_i, d_i, e_i \in R_{iii}$ and some invertible scalar $\lambda_{iii}$.
If $\sigma_{ii}$ and $\sigma_{ij}$ are permitted then one has the following relation in $\text{Hom}(iij, jii)$.

\begin{equation}
sts = \lambda_{jii} tst + p_{jii} st
\end{equation}

for some $p_{jii} \in R_{jii}$ and some invertible scalar $\lambda_{jii}$.
If $\sigma_{ii}$ and $\sigma_{ji}$ are permitted then one has the following relation in $\text{Hom}(jii, iij)$.

\begin{equation}
sts = \lambda_{iij} tst + p_{iij} ts
\end{equation}

for some $p_{iij} \in R_{iij}$ and some invertible scalar $\lambda_{iij}$.
If $\sigma_{ii}$ and $\sigma_{ij}$ and $\sigma_{ji}$ are permitted then one has the following relation in $\text{End}(iji)$.

\begin{equation}
sts = \lambda_{iji} tst + q_{iji} \text{id}_{iji}
\end{equation}

for some $q_{iji} \in R_{iji}$ and some invertible scalar $\lambda_{iji}$.
If $\sigma_{ij}, \sigma_{ik}$, and $\sigma_{jk}$ are all permitted, then one has the following relation in $\text{Hom}(ijk, kji)$.

\begin{equation}
sts = \lambda_{kji} tst.
\end{equation}
for some invertible scalar $\lambda_{kji}$.

If $\sigma_{ii}$ is permitted, then one has the following relation in $\text{End}(ii)$ for each $f \in R_{ii}$.

$$(5.10h) \quad \sigma_{ii}f = \varphi_i(f)\sigma_{ii} + \partial_i(f)\text{id}_{ii}$$

for some linear maps $\varphi_i : R_{ii} \to R_{ii}$ and $\partial_i : R_{ii} \to R_{ii}$.

If $\sigma_{ij}$ is permitted, then one has the following relation in $\text{Hom}(ij, ji)$ for each $f \in R_{ij}$.

$$(5.10i) \quad \sigma_{ij}f = \varphi_{ij}(f)\sigma_{ij}$$

for some linear map $\varphi_{ij} : R_{ij} \to R_{ji}$.

This concludes the list of relations.

**Example 5.10.** The Khovanov-Lauda-Rouquier algebras $[12, 16]$ attached to a Cartan datum (and a matrix $Q$ of polynomials) are an example of a Hecke-type category. In this example, every crossing $\sigma_{ij}$ is permissible. The dot rings $R_i$ are each polynomial rings in one variable (with different gradings).

**Example 5.11.** The Webster algebras $[19]$ (for simplicity, let us do the case of $\text{sl}_2$) are an example of a Hecke-type category. The set $I$ of colors consists of one kind of black strand, and one kind of red strand for each $\lambda \in \mathbb{Z}_{\geq 0}$. It is not permissible to cross two red strands, but other crossings are permissible. $R_{\text{black}}$ is a polynomial ring in one variable, while $R_{\text{red}}$ is just the ground field, for any $\lambda \geq 0$.

**Theorem 5.12.** Let $C$ be a Hecke-type category. Then $\text{Hom}(i, j)$ has a basis (over dots $R_j$ acting on the left) indexed by permissible $w \in S_n$, obtained by taking an arbitrary reduced expression for each $w$, if and only if:

- Each of the ambiguities in (5.1) is resolvable, for each permissible coloring.
- Each of the ambiguities below is resolvable, for each possible coloring.

$$(5.11)$$

**Proof.** One can easily modify the relation of Definition 3.20 to account for the extra data of the coloring. Given Lemma 5.1, it is easy to argue that the ambiguities in (5.1) and (5.11) combine to give a minimal set of ambiguities. Then Theorem 4.4, Theorem 3.27, and the discussion of §2.5 combine to give this theorem. $\square$
As stated earlier, in this paper we will not compute what compatibilities between these coefficients are required in order that the minimal ambiguities of (5.1) and (5.11) are resolvable. Under some simplifying assumptions, this will be done in [8]. Note that, for Webster algebras and Khovanov-Lauda-Rouquier algebras, checking the ambiguities is greatly simplified by the fact that \( a_i = b_i = c_i = d_i = e_i = 0 \) in (5.10c), so that it becomes a relatively easy exercise (and worthwhile practice for the neophyte).

Remark 5.13. Checking the ambiguity \( \sigma_i m_f m_g \) will imply that \( \varphi_i \) is a homomorphism of rings. Similarly, so is \( \varphi_{ij} \). The map \( \partial_i \) is not a homomorphism, instead satisfying some kind of modified Leibniz rule.

If the maps \( \varphi_i \) and \( \varphi_{ij} \) are isomorphisms, then one can use these same relations to move dots to the right rather than the left. Then \( C \) will have the same basis over dots acting on the right.

5.4. A further direction. We end this paper by posing a question for the interested student. Manin-Schechtman theory is a statement only about symmetric groups, which are the Coxeter groups of type \( A \). It has begun to be generalized to type \( B \), see [17], though one does not expect that it will generalize to other types, even type \( D \) (see [7] for further discussion). Hence, we are no closer to a Bergman diamond lemma which works for (modified) Coxeter group algebras in general.

The Bergman diamond lemma relies upon a partial order, and upon choosing a direction for each relation. However, it is reasonable to ask whether there might be a modification of the diamond lemma which works for preorders, where certain relations can be applied in both directions. For example, in a Coxeter presentation, perhaps one wishes to apply the braid relations in either direction, but requires the quadratic relation to be applied as \( \sigma \sigma \mapsto 1 \). Working with preorders creates new “circular ambiguities,” as one uses relations to travel in circular paths around equivalent words in this preorder. There are other new kinds of ambiguities as well: two different words in the same equivalence class might admit a relation, giving two possibly different ways of resolving the equivalence class.

To resolve the possible (infinite) chains of resolutions which can now occur, one needs a fairly strong understanding of the equivalence classes. For Coxeter groups, we know a lot about reduced expression graphs: all their monodromy (i.e. generators of \( \pi_1 \) of the graph) comes from finite parabolic subgroups of rank 3, and appearances of their longest elements as subwords. For example, we have already seen the loop in the reduced expression graph of \( s t s u t s \): follow (5.3a) forwards then (5.3b) backwards. Perhaps there is a variant on the diamond lemma which applies to this preorder, and for which the minimal ambiguities are classifiable.

Remark 5.14. There is also monodromy in reduced expression graphs coming from disjoint ambiguities (i.e. applying braid relations to disjoint parts of a word), which are always resolvable.

Having such a variant on the diamond lemma would be more philosophically satisfying, as choosing particular reduced expressions (like the Manin-Schechtman sink) seems like an unnatural thing to be forced to do.

Addendum. Circular ambiguities should be treatable using the same “non-terminating” techniques as were showcased in [6].

REFERENCES

[1] George M. Bergman. The diamond lemma for ring theory. Adv. in Math., 29(2):178–218, 1978.
[2] L. A. Bokut. Gröbner-Shirshov basis for the braid group in the Artin-Garside generators. J. Symbolic Comput., 43(6-7):397–405, 2008.

[3] L. A. Bokut and L.-S. Shiao. Gröbner-Shirshov bases for Coxeter groups. Comm. Algebra, 29(9):4305–4319, 2001. Special issue dedicated to Alexei Ivanovich Kostrikin.

[4] Bruno Buchberger. An algorithmic criterion for the solvability of a system of algebraic equations [MR0268178 (42 #3077)]. In Gröbner bases and applications (Linz, 1998), volume 251 of London Math. Soc. Lecture Note Ser., pages 535–545. Cambridge Univ. Press, Cambridge, 1998. Translated from the German by Michael Abramson and Robert Lumbert.

[5] Vladimir Dotsenko and Anton Khoroshkin. Gröbner bases for operads. Duke Math. J., 153(2):363–396, 2010.

[6] Benjamin Dupont. Rewriting modulo isotopies in Khovanov-Lauda-Rouquier’s categorification of quantum groups. Preprint, 2019. arXiv:1907.09901.

[7] Ben Elias. Thicker Soergel calculus in type A. Proc. Lond. Math. Soc. (3), 112(5):924–978, 2016.

[8] Ben Elias and Agustin Garcia. Variations on a theme by Khovanov-Lauda-Rouquier. In preparation.

[9] Yves Guiraud and Philippe Malbos. Higher-dimensional categories with finite derivation type. Theory Appl. Categ., 22:No. 18, 420–478, 2009.

[10] James E. Humphreys. Introduction to Lie algebras and representation theory, volume 9 of Graduate Texts in Mathematics. Springer-Verlag, New York-Berlin, 1978. Second printing, revised.

[11] James E. Humphreys. Reflection groups and Coxeter groups, volume 29 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1990.

[12] Mikhail Khovanov and Aaron D. Lauda. A diagrammatic approach to categorification of quantum groups. I. Represent. Theory, 13:309–347, 2009.

[13] Aaron D. Lauda. An introduction to diagrammatic algebra and categorified quantum sl_2. Bull. Inst. Math. Acad. Sin. (N.S.), 7(2):165–270, 2012.

[14] Yu. I. Manin and V. V. Schechtman. Arrangements of hyperplanes, higher braid groups and higher Bruhat orders. In Algebraic number theory, volume 17 of Adv. Stud. Pure Math., pages 289–308. Academic Press, Boston, MA, 1989.

[15] Daniel Marsden. Category theory using string diagrams. Preprint, 2014. arXiv:1401.7220.

[16] Raphaël Rouquier. 2-Kac-Moody algebras. Preprint, 2008. arXiv:0812.5023.

[17] Seth Shelley-Abrahamson and Suhas Vijaykumar. Higher Bruhat orders in type B. Electron. J. Combin., 23(3):Paper 3.13, 19, 2016.

[18] A. I. Shirshov. Some algorithm problems for Lie algebras. Sibirsk. Mat. Ž., 3:292–296, 1962.

[19] Ben Webster. Knot invariants and higher representation theory. Preprint, 2013. arXiv:1309.3796.

UNIVERSITY OF OREGON, USA