SCHUBERT CALCULUS AND TORSION EXPLOSION

GEORDIE WILLIAMSON
WITH A JOINT APPENDIX WITH
ALEX KONTOROVICH AND PETER J. MCNAMARA

Abstract. We observe that certain numbers occurring in Schubert calculus for SL_n also occur as entries in intersection forms controlling decompositions of Soergel bimodules in higher rank. These numbers grow exponentially. This observation gives many counter-examples to the expected bounds in Lusztig's conjecture on the characters of simple rational modules for SL_n over fields of positive characteristic. Our examples also give counter-examples to the James conjecture on decomposition numbers for symmetric groups.

Dedicated to Meg and Gong.

1. Introduction

Let G be a connected algebraic group over an algebraically closed field. A basic question in representation theory asks for the dimensions and characters of the simple rational G-modules. Structure theory of algebraic groups allows one to assume that G is reductive. If the ground field is of characteristic zero, then the theory runs parallel to the well-understood theory for compact Lie groups. In positive characteristic p, Steinberg's tensor product theorem, the linkage principle and Jantzen's translation principle reduce this to a question about finitely many modules which occur in the same block as the trivial module (the "principal block"). For these modules Lusztig has proposed a conjecture if $p > h$, where h denotes the Coxeter number of the root system of G [Lus80].[1] He conjectures an expression for the characters of the simple modules in terms of affine Kazhdan-Lusztig polynomials and the (known) characters of standard modules.

Lusztig's conjecture has been shown to hold for p large (without an explicit bound) thanks to work of Andersen, Jantzen and Soergel [AJS94], Kashiwara and Tanisaki [KT95, KT96], Kazhdan and Lusztig [KL93, KL94a, KL94b] and Lusztig [Lus94]. Alternative proofs for large p have been given by Arkhipov, Bezrukavnikov and Ginzburg [ABG04], Bezrukavnikov, Mirkovic and Rumynin [BMR08, BM13] (in the broader context of Lie algebra representations) and Fiebig [Fie11]. Fiebig also gives an explicit enormous[2] bound [Fie12], and establishes the multiplicity one case [Fie10]. For any fixed G and "reasonable" p very little is known: the case of rank 2 groups can be deduced from Jantzen's sum formula, and intensive computational

2010 *Mathematics Subject Classification.* Primary 20C20, 20G05; Secondary 14N15, 14M15.

[1] Lusztig first proposed his conjecture under a restriction equivalent to $p \geq 2h-3$ (see [Jan08, §4] and [Jan03, §8.22] for a discussion). Kato [Kat85, §5] proved that if Lusztig's conjecture holds for restricted weights then it holds for all weights in the Jantzen region (Lusztig's original formulation). Since Kato's work $p > h$ has been widely regarded as a realistic bound [Jan08, §4].

[2] e.g. at least of the order of $p \gg n^{n^2}$ for SL_{n+1}.

efforts have checked the conjecture for small p and certain groups, all of rank ≤ 5. There is no conjecture as to what happens if p is smaller than the Coxeter number.

In [Soe00] Soergel introduced a subquotient of the category of rational representations, dubbed the "subquotient around the Steinberg weight", as a toy model for the study of Lusztig's conjecture. Whilst the full version of Lusztig's conjecture is based on the combinatorics of alcoves and the affine Weyl group, the subquotient around the Steinberg weight is controlled by the finite Weyl group, and behaves like a modular version of category $\mathcal{O}$. Lusztig's conjecture implies that the multiplicities in the subquotient around the Steinberg weight are given by finite Kazhdan-Lusztig polynomials. Thus Lusztig's conjecture implies that "the subquotient around the Steinberg weight satisfies the Kazhdan-Lusztig conjecture".

In [Soe00] Soergel goes on to explain how the subquotient around the Steinberg weight is controlled by Soergel bimodules. This allows him to relate this category to the category of constructible sheaves on the Langlands dual flag variety, with coefficients in the field of definition of G. Using Soergel's results and the theory of parity sheaves [JMW14], one can see that a part of Lusztig's conjecture for $p > h$ is equivalent to absence of p torsion in the stalks and costalks of integral intersection cohomology complexes of Schubert varieties in the flag variety. It has been known since the birth of the theory of intersection cohomology that 2-torsion occurs in type B_2, and 2- and 3-torsion occurs in type G_2. For over a decade no other examples of torsion were known. In 2002 Braden discovered 2-torsion in the stalks of integral intersection cohomology complexes on flag varieties of types D_4 and A_7 (see Braden's appendix to [WB12]). In 2011 Polo discovered 3-torsion in the cohomology of the flag variety of type E_6 and n-torsion in a flag variety of type A_{4n-1}. Polo's (as yet unpublished) results are significant, as they emphasise how little we understand in high rank (see the final lines of [Wil12]).

In general these topological calculations appear extremely difficult. Recently Elias and the author found a presentation for the monoidal category of Soergel bimodules by generators and relations [EW], building on the work of Libedinsky [Lib10], Elias-Khovanov [EK10] and Elias [Eli16]. One of the applications of this theory is that one can decide whether a given intersection cohomology complex has p-torsion in its stalks or costalks (the bridge between intersection cohomology and Soergel bimodules is provided by the theory of parity sheaves).[3] The basic idea is as follows: given any pair $(\underline{w}, x)$, where $x, w \in W$ and $\underline{w}$ is a reduced expression for $w \in W$, one has an "intersection form", an integral matrix. The stalks of the intersection cohomology complex corresponding to w are free of p-torsion if no elementary divisors of the intersection forms associated to all elements $x \leq w$ are divisible by p. In principle, this gives an algorithm to decide whether Lusztig's conjecture is correct around the Steinberg weight.[4] This algorithm (in a slightly different form) was discovered independently by Libedinsky [Lib15].

The generators and relations approach certainly makes calculations easier. However this approach still has its difficulties: the diagrammatic calculations remain

[3]One can also perform this calculation using the theory of moment graphs [FW14]. However the computations using generators and relations are generally much simpler.

[4]One can extend this to the full version of Lusztig's conjecture by using a certain subset of the affine Weyl group, thanks to the work of Fiebig [Fie11]. Although it seems likely that the converse holds, at present one only knows one implication: the absence of $p > h$ torsion implies the truth of Lusztig's conjecture in characteristic p.

extremely subtle, and the "light leaves" basis in which the intersection form is calculated depends on additional choices which seem difficult to make canonical. Recent progress in this direction has been made by Xuhua He and the author [HW], who discovered that certain entries in the intersection form (which in some important examples are all entries) are canonical and may be evaluated in terms of expressions in the nil Hecke ring.

The main result of this paper may be seen as an example of this phenomenon. We observe that one may embed certain structure constants of Schubert calculus for SL_n as the entries of 1×1 intersection forms associated to pairs $(\underline{w}, x)$ in (much) higher rank groups. In this way one can produce many new examples of torsion which grow exponentially in the rank. For example, using Schubert calculus for the flag variety of SL_4 we observe that the Fibonacci number F_{i+1} occurs as torsion in SL_{3i+5}. We deduce that there is no linear function $f(n)$ of n such that Lusztig's conjecture holds for all $p \geq f(n)$ for SL_n. In the appendix (by Kontorovich, McNamara and the author) we apply recent results of Bourgain-Kontorovich [BK14] in number theory to deduce that the torsion in SL_n grows exponentially in n.

Finally, there is a related conjecture due to James [Jam90] concerning the simple representations of the symmetric group in characteristic p. When combined with known results about the decomposition numbers for Hecke algebras at roots of unity, the James conjecture would yield the decomposition numbers for symmetric groups S_n in characteristic $p > \sqrt{n}$. In the final section of the paper we explain why our counter-examples to Lusztig's conjecture for SL_N with $p > \binom{N}{2}$ imply that the James conjecture is incorrect for $S_{p\binom{N}{2}}$. (Parts of this section were explained to me by Joe Chuang.)

1.1. **Main result.** Let $R = \mathbb{Z}[x_1, x_2, \ldots, x_n]$ be a polynomial ring in n variables. We regard R as a graded ring with $\deg x_i = 2$ (we double degrees for reasons coming from Soergel bimodules). Let $W = S_n$ denote the symmetric group on n letters viewed as a Coxeter group with simple reflections S consisting of the simple transpositions. Then W acts by permutation of variables on R. Let $s_1, \ldots, s_{n-1}$ denote the simple transpositions of S_n and let ℓ denote the corresponding length function. Let ∂_i denote the i^{th} divided difference operator:

$$\partial_i(f) = \frac{f - s_i f}{x_i - x_{i+1}} \in R.$$

For any element $w \in S_n$ we obtain well-defined operators $\partial_w = \partial_{i_1} \ldots \partial_{i_m}$ where $w = s_{i_1} \ldots s_{i_m}$ is a reduced expression for w in the generators S.

Consider elements of the form

$$\kappa = \partial_{w_m}(x_1^{a_m} x_n^{b_m} \partial_{w_{m-1}}(x_1^{a_{m-1}} x_n^{b_{m-1}} \ldots \partial_{w_1}(x_1^{a_1} x_n^{b_1}) \ldots))$$

where $w_i \in S_n$ are arbitrary. We assume that $\sum \ell(w_i) = a + b$ where $a = \sum a_i$ and $b = \sum b_i$ so that $\kappa \in \mathbb{Z}$ for degree reasons. Given a subset $I \subset \{1, \ldots, n-1\}$ let w_I denote the longest element in the parabolic subgroup $\langle s_j \rangle_{j \in I}$. Our main theorem is the following:

Theorem 1.1. *Suppose that $n \geq 1$ and $a, b \geq 0$ are as above, and that $\kappa \neq 0$. Then there exists a reduced expression $\underline{w}$ for an element of S_{a+n+b} such that the intersection form in degree zero of $\underline{w}$ at w_I, where $I = \{1, 2, \ldots, a+n+b-1\} \setminus \{a, a+n\}$, is the 1×1 matrix $((-1)^a \kappa)$.*

The construction of the expression $\underline{w}$ is explicit and combinatorial based on $w_1, \dots, w_m$, $a_1, \dots, a_m$ and $b_1, \dots, b_m$. We will also see that for certain choices of a_i, b_i, w_i the prime factors of the numbers κ grow exponentially in $h = n + a + b$.

1.2. **Schubert calculus.** We explain why "Schubert calculus" occurs in the title. Consider the coinvariant ring C for the action of $W = S_n$ on R. That is, C is equal to R modulo the ideal generated by W-invariant polynomials of positive degree. The Borel isomorphism gives a canonical identification of C with the integral cohomology of the complex flag variety of SL_n.

The divided difference operators ∂_w act on C, as do elements of R. The coinvariant ring C has a graded $\mathbb{Z}$-basis given by the Schubert classes $\{X_w \mid w \in S_n\}$ (normalised with $X_{w_0} = x_1^{n-1} x_2^{n-2} \dots x_{n-1}$ and $X_w = \partial_{ww_0} X_{w_0}$). We have:

$$\partial_i X_w = \begin{cases} X_{s_i w} & \text{if } s_i w < w, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

The action of multiplication by $f \in R$ of degree two is given as follows (the Chevalley formula):

$$f \cdot X_w = \sum_{\substack{t \in T \\ \ell(tw) = \ell(w)+1}} \langle f, \alpha_t^\vee \rangle X_{tw}. \tag{1.2}$$

(Here T denotes the set of reflections (transpositions) in S_n and if $t = (i, j) \in T$ with $i < j$ then $\alpha_t^\vee = \varepsilon_i - \varepsilon_j$ where $\{\varepsilon_i\}$ is the dual basis to $x_1, \dots, x_n$.)

Now consider the numbers one may obtain as coefficients in the basis of Schubert classes by multiplication by x_1 and x_n and by applying Demazure operators, starting with X_{id}. Because $\partial_w X_{w^{-1}} = X_{\mathrm{id}} = 1$, any coefficient of any Schubert class that we obtain in this way can be realised as the coefficient of X_{id}. Now Theorem 1.1 says that any such number occurs as torsion in SL_{n+a+b} where a (resp. b) counts the number of times that one has applied the operator of multiplication by x_1 (resp. x_n).

1.3. **Note to the reader.** This paper is entirely algebraic in that it relies only on Soergel (bi)modules, their diagrammatics and connections to representation theory (due to Soergel). Except in remarks, we neither explain nor use the relation to constructible sheaves and torsion. An alternative geometric proof of the main theorem (discovered a year after this paper was first circulated) is given in [Wil].

1.4. **Structure of the paper.**

§2-4: Contains background on Soergel (bi)modules and intersection forms.
§5: We prove the main theorem.
§6: We use our main theorem for $n = 4, 5$ to give examples of torsion.
§7: We explain the connection to the Lusztig conjecture.
§8: We explain the connection to the James conjecture.
§A: We (AK, PM and GW) prove exponential growth of torsion.

1.5. **Acknowledgements.** The ideas of this paper crystallised after long discussions with Xuhua He. I would like to thank him warmly for the invitation to Hong Kong and the many interesting discussions that resulted from the visit. I would also like to thank Ben Elias for countless hours (often productive, always enjoyable) getting to know Soergel bimodules. His influence is omnipresent in this paper.

Thanks to Joe Chuang for useful correspondence and explaining how to get counter-examples in the symmetric group. Finally, thanks to Henning Haahr Andersen, Ben Elias, Peter Fiebig, Anthony Henderson, Daniel Juteau, Nicolas Libedinsky, Kaneda Masaharu and especially Patrick Polo and the referees for valuable comments. These results were announced in June 2013 at ICRT VI in Zhangjiajie, China.

2. Soergel bimodules

In the first three sections we recall what we need from the theory of Soergel (bi)modules and intersection forms. This paper is not self-contained. The main references are [Soe90, Soe92, Soe07, EK10, Eli16, EW].

Fix $n \geq 1$ and let $W = S_n$ denote the symmetric group on n letters. Throughout we view W as a Coxeter group with simple reflections $S = \{(i, i+1) \mid 1 \leq i < n\}$, and denote by ℓ its length function and $\leq$ its Bruhat order. Let $\mathcal{H}$ denote the Hecke algebra of (W, S) over $\mathbb{Z}[v^{\pm 1}]$ normalised as in [Soe97]. Let $\{H_x\}_{x \in W}$ and $\{\underline{H}_x\}_{x \in W}$ denote its standard and Kazhdan-Lusztig bases.

Fix a field $\Bbbk$ of characteristic $p > 2$ and let $R = \Bbbk[x_1, \dots, x_n]$. Then W acts by permutation of variables on R (graded algebra automorphisms). The reader may easily check (see e.g. [Lib15, Lemma 7.4]) that this action is reflection faithful in the sense of [Soe07, Definition 1.5]. Given $s \in S$ we denote by $R^s \subset R$ the invariant subring.

Given a $\mathbb{Z}$-graded object (vector space, module, bimodule) $M = \bigoplus M^i$ we let $M(j)$ denote the shifted object: $M(j)^i = M^{i+j}$.

The category of *Soergel bimodules* $\mathcal{B}$ is the full additive monoidal graded Karoubian subcategory of graded R-bimodules generated by $B_s = R \otimes_{R^s} R(1)$ for all $s \in S$. In other words, the indecomposable Soergel bimodules are the shifts of the indecomposable direct summands of the *Bott-Samelson bimodules*

$$B_{\underline{w}} = B_{s_1} \otimes_R B_{s_2} \otimes_R \cdots \otimes_R B_{s_m}(m)$$

for all expressions $\underline{w} = s_1 s_2 \dots s_m$ in S. For any $w \in S_n$ let B_w denote the indecomposable self-dual Soergel bimodule which occurs as a summand of $B_{\underline{w}}$ for any reduced expression $\underline{w}$ for w, and is not isomorphic to a summand of $B_{\underline{w}'}$ for any shorter $\underline{w}'$. The set $\{B_w\}_{w \in W}$ coincides with the set of all indecomposable Soergel bimodules, up to shifts in the grading [Soe07].

Remark 2.1. In [Soe07] Soergel develops the theory of Soergel bimodules for a reflection faithful representation V over an infinite field of characteristic $\neq 2$. We have remarked above that the reflection faithful hypothesis is always satisfied. The assumption that $\Bbbk$ is infinite is made in order to identify R with the polynomial functions on V. However all the results of [Soe07] hold if one simply defines R to be the symmetric algebra on V^*, as we do. Alternatively, the reader may assume that $\Bbbk$ is infinite throughout.

We denote by $[\mathcal{B}]$ the split Grothendieck group of $\mathcal{B}$ (i.e. $[B] = [B'] + [B'']$ if $B \cong B' \oplus B''$). We make $[\mathcal{B}]$ into a $\mathbb{Z}[v^{\pm 1}]$ algebra via $v[M] := [M(-1)]$, $[B][B'] := [B \otimes_R B']$. In [Soe07] Soergel proves that there exists a unique isomorphism of $\mathbb{Z}[v^{\pm 1}]$-algebras

$$\mathrm{ch} : [\mathcal{B}] \xrightarrow{\sim} \mathcal{H}$$

such that $\mathrm{ch}(R(-1)) = v$ and $\mathrm{ch}(B_s) = \underline{H}_s$ for all $s \in S$.

Remark 2.2. Under our assumptions B_w may be realised as the equivariant intersection cohomology of the indecomposable parity sheaf [JMW14] of the Schubert variety labelled by w in the flag variety [Fie08, FW14]. In particular, if $\Bbbk$ is of characteristic zero, then B_w is the equivariant intersection cohomology of a Schubert variety. In fact the whole theory of Soergel bimodules can be seen as providing an algebraic description of the Hecke category.

Set ${}^p\underline{H}_x := \mathrm{ch}(B_x) \in \mathcal{H}$. Then $\{{}^p\underline{H}_x\}_{x\in W}$ is a basis which only depends on the characteristic p of $\Bbbk$, the *p-canonical basis* (see [Wil12, JW]). Let us write

$$\underline{H}_x = \sum h_{y,x} H_y, \qquad {}^p\underline{H}_x = \sum {}^p h_{y,x} H_y, \qquad {}^p\underline{H}_x = \sum {}^p a_{y,x} \underline{H}_y. \tag{2.1}$$

for polynomials $h_{y,x} \in \mathbb{Z}[v]$ and ${}^p h_{y,x}, {}^p a_{y,x} \in \mathbb{Z}[v^{\pm 1}]$. The polynomials $h_{y,x}$ are (normalisations of) Kazhdan-Lusztig polynomials and have non-negative coefficients. The polynomials ${}^p h_{y,x}, {}^p a_{y,x}$ also have non-negative coefficients [JW, Proposition 4.2].

Throughout this paper we will say that p *occurs as torsion in* SL_n if there exists $x \in W$ such that ${}^p\underline{H}_x \neq \underline{H}_x$.

3. Soergel modules

In this section we assume that $p > n$, so that the results of [Soe00] are available.

Let $R^W_+ \subset R$ denote the W-invariants of positive degree, $\langle R^W_+ \rangle$ the ideal they generate, and $C = R/\langle R^W_+ \rangle$ the coinvariant algebra, which inherits an (even) grading and a W-action from R. Let $\mathcal{C}$ denote the category of *Soergel modules* consisting of all

$$D_{\underline{w}} := C \otimes_{C^{s_m}} \cdots \otimes_{C^{s_2}} C \otimes_{C^{s_1}} \Bbbk(m)$$

for expressions $\underline{w} = s_1 s_2 \dots s_m$ in S, together with their shifts, direct sums and summands inside the category of graded C-modules. (Note the order of tensor factors.)

For a reduced expression $\underline{x}$ for x let D_x denote the unique summand of $D_{\underline{x}}$ which does not occur as a summand of $D_{\underline{x}'}$ for any shorter expression $\underline{x}'$. The set $\{D_x \mid x \in W\}$ is well-defined and gives a set of representatives for the isomorphism classes of indecomposable Soergel modules (up to shift) [Soe00, Theorem 2.8.1].

How to go from Soergel bimodules to Soergel modules? Given a right R-module M which is killed by R^W_+ the canonical map $M \otimes_{R^s} R \twoheadrightarrow M \otimes_{C^s} C$ is an isomorphism. Hence we have an isomorphism of graded right C-modules:

$$\Bbbk \otimes_R R \otimes_{R^s} R \otimes_{R^t} \cdots \otimes_{R^u} R \cong \Bbbk \otimes_C C \otimes_{C^s} C \otimes_{C^t} \cdots \otimes_{C^u} C.$$

It follows that if we compose the functor $M \mapsto \Bbbk \otimes_R M$ with the equivalence between right and left C-modules (C is commutative) we obtain a functor

$$c : \mathcal{B} \to \mathcal{C}$$

with $c(B_{\underline{w}}) = D_{\underline{w}}$.

Lemma 3.1. *$c(B_x) \cong D_x$.*

Proof. Step 1: We claim that the natural map provides an isomorphism:

$$\Bbbk \otimes_R \mathrm{Hom}^\bullet_{\mathcal{B}}(B_{\underline{x}}, B_{\underline{y}}) \xrightarrow{\sim} \mathrm{Hom}^\bullet_C(c(B_{\underline{x}}), c(B_{\underline{y}})). \tag{3.1}$$

(Here and in the rest of the proof, $\mathrm{Hom}^\bullet$ denotes the graded module of morphisms of all degrees.) By repeated application of the biadjunction $(\otimes_R B_s(1), \otimes_R B_s(1))$

we may assume that $\underline{x}$ is the empty sequence. The map $\phi \mapsto \phi(1)$ gives a canonical identification of $\mathrm{Hom}^\bullet_{\mathcal{B}}(R, B_{\underline{y}})$ with the submodule of invariants

$$\Gamma_{\mathrm{id}} B_{\underline{y}} := \{m \in B_{\underline{y}} \mid rm = mr \text{ for all } r \in R\}.$$

Now $\Gamma_{\mathrm{id}} B_{\underline{y}}$ is the first step in the filtration $\Gamma_{\leq 0} B_{\underline{y}} \subset \Gamma_{\leq 1} B_{\underline{y}} \subset \dots$ considered after the proof of [Soe07, Proposition 5.7], and from [Soe07, Proposition 5.9] we deduce that the subquotients of this filtration are free as left R-modules. Thus $\Gamma_{\mathrm{id}} B_{\underline{y}}$ is a summand of $B_{\underline{y}}$ as a left R-module. The injectivity of (3.1) follows.

We deduce the surjectivity of (3.1) by showing that both sides have the same (finite) dimension. If $\underline{y} = s_1 s_2 \dots s_m$ let us write $\underline{H}_{s_1} \underline{H}_{s_2} \dots \underline{H}_{s_m} = \sum g_x H_x$ for some $g_x \in \mathbb{Z}[v^{\pm 1}]$. In the notation of [Soe07] we have, by [Soe07, Proposition 5.7],

$$\sum_{m \in \mathbb{Z}} (B_{\underline{y}} : \nabla_{\mathrm{id}}[m]) v^{-m} = g_{\mathrm{id}}$$

and $(R : \Delta_x[m]) = \delta_{x,\mathrm{id}} \delta_{m,0}$ (Kronecker's δ). Now we apply [Soe07, Theorem 5.15] to deduce that $\mathrm{Hom}^\bullet(R, B_{\underline{y}})$ is free of rank $g_{\mathrm{id}}(1)$ over R. On the other hand, if $\langle -, - \rangle : \mathbb{Z}W \times \mathbb{Z}W \to \mathbb{Z}$ denotes the pairing with $\langle x, y \rangle = \delta_{x,y}$, then by [Soe00, Lemma 2.11.2], we have

$$\dim \mathrm{Hom}^\bullet_C(\Bbbk, c(B_{\underline{y}})) = \langle \mathrm{id}, \sum g_x(1) x \rangle = g_{\mathrm{id}}(1).$$

The surjectivity follows.

Step 2: Because $c(B_{\underline{w}}) = D_{\underline{w}}$ we can appeal to the defining properties of B_x and D_x to see that it is enough to show: if B is indecomposable, then so is $c(B)$. By the previous step $\mathrm{End}_C(c(B)) = \Bbbk \otimes_R \mathrm{End}_{\mathcal{B}}(B)$, and so $\mathrm{End}_C(c(B))$ is a quotient of the local ring $\mathrm{End}_{\mathcal{B}}(B)$. Now the result follows as a non-zero quotient of a local ring is local. □

Remark 3.2. The above proof uses representation theory, via [Soe00, Lemma 2.11.2]. Soergel has found an algebraic proof of the above lemma, valid for any finite Coxeter group (unpublished).

Denote by f the functor of forgetting the grading on a C-module, and let $f\mathcal{C}$ denote the essential image of $\mathcal{C}$ under f. By [Soe00, Theorem 2.8.1], the indecomposable objects in $f\mathcal{C}$ are precisely the $\{fD_x\}$. We denote by $[f\mathcal{C}]$, $[\mathcal{C}]$ the split Grothendieck groups of $f\mathcal{C}$ and $\mathcal{C}$ respectively. Because $\mathcal{C}$ is graded, $[\mathcal{C}]$ is naturally a $\mathbb{Z}[v^{\pm 1}]$-module via $v[M] := [M(-1)]$ as above. These observations, together with the above lemma, show that we have a commutative diagram:

$$\begin{array}{ccccc}
[f\mathcal{C}] & \xleftarrow{\quad f \quad} & [\mathcal{C}] & \xleftarrow[\sim]{\quad c \quad} & [\mathcal{B}] \\
\beta \downarrow \sim & & \downarrow \sim & & \sim \downarrow \mathrm{ch} \\
\mathbb{Z}S_n & \xleftarrow[1 \leftarrow\!\shortmid\, v]{} & \mathcal{H} & = & \mathcal{H}
\end{array}$$

where β is defined by

$$\beta(fD_x) = \mathrm{ch}(B_x)_{|v=1} = {}^p\underline{H}_{x\,|v=1}. \tag{3.2}$$

4. Intersection forms

Let $\mathcal{B}$ denote the category of Soergel bimodules defined above. Given any ideal $I \subset W$ (i.e. $x \leq y \in I \Rightarrow x \in I$) we denote by $\mathcal{B}_I$ the ideal of $\mathcal{B}$ generated by all morphisms which factor through a Bott-Samelson bimodule $B_{\underline{y}}$, where $\underline{y}$ is a reduced expression for $y \in I$.

Given $x \in W$ we denote by $\mathcal{B}^{\geq x}$ the quotient category $\mathcal{B}/\mathcal{B}_{\not\geq x}$ where $\not\geq x := \{y | y \not\geq x\}$. We write $\mathrm{Hom}_{\geq x}$ for (degree zero) morphisms in $\mathcal{B}^{\geq x}$. All Bott-Samelson bimodules $B_{\underline{x}}$ corresponding to reduced expressions $\underline{x}$ for x become canonically isomorphic to B_x in $\mathcal{B}^{\geq x}$. We have $\mathrm{End}_{\geq x}(B_x) = \Bbbk$. Given any expression $\underline{w}$ in S the *intersection form*[5] is the canonical pairing

$$I^{\Bbbk}_{x,\underline{w},d} : \mathrm{Hom}_{\geq x}(B_x(d), B_{\underline{w}}) \times \mathrm{Hom}_{\geq x}(B_{\underline{w}}, B_x(d)) \to \mathrm{End}_{\geq x}(B_x(d)) = \Bbbk.$$

The following is standard (see e.g. [JMW14, Lemma 3.1] for a similar situation):

Lemma 4.1. *The multiplicity of $B_x(d)$ as a summand of $B_{\underline{w}}$ equals the rank of $I^{\Bbbk}_{x,\underline{w},d}$.*

In the papers [EK10, EW] the category of Soergel bimodules is presented by generators and relations. More precisely, a diagrammatic category is defined by generators and relations and it is proved that its Karoubi envelope is equivalent to Soergel bimodules, as a graded monoidal category. We will not repeat the rather complicated list of generators and relations here, see [EW, §1.4] or [HW, §2.7].

In the category $\mathcal{D}$ the intersection form is explicit and amenable to computation: see [HW, §2.10] for examples. From the diagrammatic approach it is clear that $I^{\Bbbk}_{x,\underline{w},d}$ is defined over $\mathbb{Z}$, in the sense that there exists an integral form $I_{x,\underline{w},d}$ on a pair of free $\mathbb{Z}$-modules such that $I^{\Bbbk}_{x,\underline{w},d} = I_{x,\underline{w},d} \otimes_{\mathbb{Z}} \Bbbk$ for any field $\Bbbk$.

Corollary 4.2. *The following are equivalent:*

1. *The indecomposable Soergel bimodules in characteristic p categorify the Kazhdan-Lusztig basis. That is, ${}^p\underline{H}_x = \underline{H}_x$ for all $x \in W$.*
2. *For all (reduced) expressions $\underline{w}$, all $x \in W$ and all $m \in \mathbb{Z}$ the graded ranks of $I_{x,\underline{w},m} \otimes_{\mathbb{Z}} \mathbb{Q}$ and of $I_{x,\underline{w},m} \otimes_{\mathbb{Z}} \Bbbk$ agree.*
3. *For all reduced expressions $\underline{w}$ and all $x \in W$ the graded ranks of $I_{x,\underline{w},0} \otimes_{\mathbb{Z}} \mathbb{Q}$ and of $I_{x,\underline{w},0} \otimes_{\mathbb{Z}} \Bbbk$ agree.*

Proof. Soergel's theorem [Soe01, Lemma 5] implies that the indecomposable Soergel bimodules categorify the Kazhdan-Lusztig basis if $\Bbbk$ is of characteristic zero (see [EW14] for an algebraic proof of this fact). Now Lemma 4.1 says that (2) holds if and only if $B_{\underline{w}}$ decomposes the same way over $\mathbb{Q}$ as it does over $\Bbbk$. Hence (1) and (2) are equivalent and (1) implies (3).

It remains to see that (3) implies (1). We show the contrapositive. So assume that (1) is not satisfied, and let w be of minimal length such that ${}^p\underline{H}_w \neq \underline{H}_w$. For any $s \in S$ with $ws < w$, B_w is a summand of $B_{ws}B_s$. By our minimality assumption $\mathrm{ch}(B_{ws}) = \underline{H}_{ws}$ and hence $\mathrm{ch}(B_{ws}B_s) = \underline{H}_{ws}\underline{H}_s = \sum g_x \underline{H}_x$ for some $g_x \in \mathbb{Z}_{\geq 0}$. Hence $\mathrm{ch}(B_w) = {}^p\underline{H}_w = \sum {}^p a_{x,w} \underline{H}_x$ with ${}^p a_{x,w} \in \mathbb{Z}_{\geq 0}$. By Lemma 4.1,

[5]The terminology "intersection form" comes from geometry: in de Cataldo and Migliorini's Hodge theoretic proof of the decomposition theorem, a key role is played by certain intersection forms associated to the fibres of proper maps [dCM02, dCM05]. In our setting, these intersection forms are associated to the fibres of Bott-Samelson resolutions of Schubert varieties. The relevance of these forms for the study of torsion in intersection cohomology was pointed out in [JMW14].

if $x < w$ is such that ${}^p a_{x,w} \neq 0$ then the ranks of $I_{x,\underline{w},0} \otimes_{\mathbb{Z}} \mathbb{Q}$ and $I_{x,\underline{w},0} \otimes_{\mathbb{Z}} \Bbbk$ differ, for any reduced expression $\underline{w}$ for w. Thus (3) $\Rightarrow$ (1). □

Remark 4.3. The intersection form and the above proposition is one of the tools used by Fiebig to establish his bound [Fie12].

5. Proof of the theorem

Let W denote the symmetric group on $\{1, 2, \ldots, a+n+b\}$ with Coxeter generators $S = \{s_1, s_2, \ldots, s_{a+n+b-1}\}$ the simple transpositions. Given a subset $I \subset S$ let W_I denote the corresponding standard parabolic subgroup and w_I its longest element. Consider the sets

$$A = \{s_1, s_2, \ldots, s_{a-1}\}, M = \{s_{a+1}, \ldots, s_{a+n-1}\}, B = \{s_{a+n+1}, \ldots, s_{a+n+b-1}\}.$$

Then W_A (resp. W_M, resp. W_B) is the subgroup of permutations of $\{1, \ldots, a\}$ (resp. $\{a+1, \ldots, a+n\}$, resp. $\{a+n+1, \ldots, a+n+b\}$).

We use the notation of §1.1 except we shift all indices by a. That is we regard S_n as embedded in S_{a+n+b} as the standard parabolic subgroup W_M. We rename $R = \mathbb{Z}[x_1, \ldots, x_{a+n+b}]$ and write $\alpha_i = x_i - x_{i+1}$ for the simple root corresponding to s_i. Fix

$$\kappa = \partial_{w_m}(x_{a+1}^{a_m} x_{a+n}^{b_m} \partial_{w_{m-1}}(x_{a+1}^{a_{m-1}} x_{a+n}^{b_{m-1}} \ldots \partial_{w_1}(x_{a+1}^{a_1} x_{a+n}^{b_1}) \ldots)) \tag{5.1}$$

which we assume is a non-zero integer. (Now $w_1, \ldots, w_m \in W_M$ and the fact that κ is a non-zero integer implies that $\sum \ell(w_i) = a + b$.)

We now perform some preliminary simplifications of the right hand side of (5.1). By replacing each $x_{a+1}^{a_i} x_{a+n}^{b_i}$ with $x_{a+1}^{a_i} \partial_{\mathrm{id}} x_{a+n}^{b_i}$ we may assume that for all i, either a_i or b_i is zero. Let $M' = M \setminus \{s_{a+1}, s_{a+n-1}\}$. If $w \in W_{M' \cup \{s_{a+1}\}}$ then ∂_w commutes with the operator of multiplication with x_{a+n}. Thus if a_i is zero then we may assume that w_i is minimal in its coset $w_i W_{M' \cup \{s_{a+1}\}}$. Similarly, if b_i is zero then we may assume that w_i is minimal in its coset $w_i W_{M' \cup \{s_{a+n-1}\}}$. From now on we assume that the right hand side of (5.1) has been simplified in this way. Finally, the minimal coset representatives of $W_M / W_{M' \cup \{s_{a+n-1}\}}$ are the elements:

$$\mathrm{id}, s_{a+1},\ s_{a+2}s_{a+1},\ \ldots,\ s_{a+n-1}s_{a+n-2} \ldots s_{a+2}s_{a+1}.$$

Similarly, the minimal coset representatives of $W_M / W_{M' \cup \{s_{a+1}\}}$ are the elements:

$$\mathrm{id}, s_{a+n-1},\ s_{a+n-2}s_{a+n-1},\ \ldots,\ s_{a+1}s_{a+2} \ldots s_{a+n-2}s_{a+n-1}.$$

Thus each w_i belongs to the first (resp. second) list if $b_i = 0$ (resp. $a_i = 0$).

Fix a reduced expression $\underline{w}_M$ for w_M and reduced expressions $\underline{w}_i$ for each w_i. (In fact, following the reductions of the previous paragraph each w_i has a unique reduced expression.) Let $\underline{w}$ be the sequence

$$\underline{w} = \underline{w}_m \underline{u}_m \underline{v}_m \ldots \underline{w}_2 \underline{u}_2 \underline{v}_2 \underline{w}_1 \underline{u}_1 \underline{v}_1 \underline{w}_M$$

where

$$\begin{aligned}
\underline{u}_1 &= (s_a \ldots s_{a-a_1+1}) \ldots (s_a s_{a-1})(s_a) \\
\underline{u}_2 &= (s_a \ldots s_{a-a_1-a_2+1}) \ldots (s_a \ldots s_{a-a_1-1})(s_a \ldots s_{a-a_1}) \\
&\vdots \\
\underline{u}_m &= (s_a \ldots s_1) \ldots (s_a \ldots s_{a-a_1-\cdots-a_{m-1}-1})(s_a \ldots s_{a-a_1-\cdots-a_{m-1}})
\end{aligned}$$

(subscripts fall by 1 within each parenthesis, and s_a occurs a_i times in $\underline{u}_i$) and

$$\underline{v}_1 = (s_{a+n} \dots s_{a+n+b_1-1}) \dots (s_{a+n} s_{a+n+1})(s_{a+n})$$
$$\underline{v}_2 = (s_{a+n} \dots s_{a+n+b_1+b_2-1}) \dots (s_{a+n} \dots s_{a+n+b_1+1})(s_{a+n} \dots s_{a+n+b_1})$$
$$\vdots$$
$$\underline{v}_m = (s_{a+n} \dots s_{a+n+b-1}) \dots (s_{a+n} \dots s_{a+n+b_1+\dots+b_{m-1}})$$

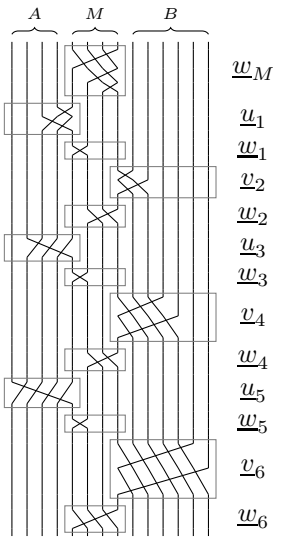
(subscripts rise by 1 within each parenthesis, and s_{a+n} occurs b_i times in $\underline{v}_i$).

Remark 5.1. The sequence $\underline{u}_m \dots \underline{u}_2 \underline{u}_1$ (resp. $\underline{v}_m \dots \underline{v}_2 \underline{v}_1$) is a reduced expression for the longest element of $W_{A \cup \{s_a\}}$ (resp. $W_{\{s_{a+n}\} \cup B}$). If we denote by $\underline{u}'_i$ (resp. $\underline{v}'_i$) the expression obtained from $\underline{u}_i$ (resp. $\underline{v}_i$) by deleting every occurrence of s_a (resp. s_{a+n}) then $\underline{u}'_m \dots \underline{u}'_2 \underline{u}'_1$ (resp. $\underline{v}'_m \dots \underline{v}'_2 \underline{v}'_1$) is a reduced expression for the longest element of W_A (resp. W_B).

Example 5.2. We give a real-life example. We take $n = 4$ and consider the operator $F : f \mapsto \partial_{23}(x_4^2(\partial_1(x_1 f)))$ on $\mathbb{Z}[x_1, x_2, x_3, x_4]$ (we write $\partial_{23} := \partial_2 \partial_3$). In the next section we will see that F is a "Fibonacci operator"; in particular

$$\partial_1 F^3(x_1) = \partial_{123}(x_4^2 \partial_1(x_1 \partial_{23}(x_4^2 \partial_1(x_1 \partial_{23}(x_4^2 \partial_1(x_1^2)))))) = 3.$$

In the notation of §1.1 we have $w_1 = w_3 = w_5 = s_1$, $w_2 = w_4 = s_2 s_3$, $w_6 = s_1 s_2 s_3$, $a_1 = 2$, $a_3 = a_5 = 1$, $a_2 = a_4 = a_6 = 0$, $b_1 = b_3 = b_5 = 0$ and $b_2 = b_4 = b_6 = 2$. Hence $a = 4, b = 6$ and $a + n + b = 14$. We can depict $\underline{w}$ as follows:



The rest of this section will be occupied with the proof of the following theorem:

Theorem 5.3. *The degree zero intersection form of $\underline{w}$ at $w_{A \cup M \cup B}$ is the 1×1-matrix $((-1)^a \kappa)$.*

In the proof we will need the notion of subexpression and defect together with the main result of [HW]. Fix a word $\underline{y} = s_{i_1} \dots s_{i_m}$ in S. A *subexpression* of $\underline{y}$ is a sequence $\underline{e} = e_1 \dots e_m$ with $e_i \in \{0, 1\}$ for all i. We set $\underline{y}^{\underline{e}} := s_{i_1}^{e_1} \dots s_{i_m}^{e_m} \in W$. Any subexpression $\underline{e}$ determines a sequence $y_0, y_1, \dots, y_m \in W$ via $y_0 := \mathrm{id}$, $y_j :=$

$s_{i_{m+1-j}}^{e_{m+1-j}} y_{j-1}$ for $1 \leq i \leq m$ (so $y_m = \underline{y}^{\underline{e}}$). Given a subexpression $\underline{e}$ we associate a sequence $d_j \in \{U, D\}$ (for Up, Down) via

$$d_j := \begin{cases} U & \text{if } s_{i_j} y_{m-j} > y_{m-j}, \\ D & \text{if } s_{i_j} y_{m-j} < y_{m-j}. \end{cases}$$

Usually we view $\underline{e}$ as the decorated sequence $(d_1 e_1, \dots, d_m e_m)$. The *defect* of $\underline{e}$ is

$$\mathrm{df}(\underline{e}) := |\{i \mid d_i e_i = U0\}| - |\{i \mid d_i e_i = D0\}|.$$

Remark 5.4. See [EW, §2.4] for examples and motivation. We warn the reader that in this paper we work from right to left to define the defect, rather than from left to right as in [EW, §2.4] and [HW, §2.3]. This change of conventions is necessary to have the operators ∂_i act on polynomials on the left. One may easily pass between the two possible choices via the symmetry on Soergel bimodules which interchanges left and right actions. (In the diagrammatic language of [EW] this corresponds to flipping diagrams about the y-axis.)

Recall that the nil Hecke ring NH is defined to be the algebra generated by R and symbols δ_i for each $s_i \in S$, and subject to the relation $\delta_i^2 = 0$ for all $s_i \in S$, the braid relations and the nil Hecke relation

$$\delta_i f = s_i(f)\delta_i + \partial_i(f) \quad \text{for all } s_i \in S \text{ and } f \in R.$$

As left R-modules NH is free with basis $\{\delta_w\}_{w \in W}$, where $\delta_w := \delta_{i_1} \dots \delta_{i_k}$ for any reduced expression $w = s_{i_1} \dots s_{i_k}$. The grading on R extends to a grading on NH with $\deg \delta_w = -2\ell(w)$ for all $w \in W$.

Equipped with this notation we can now give the proof.

Lemma 5.5. *$\underline{w}$ is reduced.*

Proof. Let us fix an element $x \in W_Q$ where $Q = \{s_p, s_{p+1}, \dots, s_{q-1}, s_q\}$ (for some p, q with $1 < p \leq q < a+n+b-1$) and a reduced expression $\underline{x}$ for x. Then for any j the expressions

$$s_j s_{j-1} \dots s_{p-1} \underline{x} \quad \text{and} \quad s_j s_{j+1} \dots s_{q+1} \underline{x}$$

are reduced. (For example, one can write a formula for how the displayed elements act on $1, 2, \dots, a+n+b$ in terms of x, and verify that their lengths differ from $\ell(x)$ by $j - p + 2$ (resp. $q - j + 2$) by counting inversions. It follows that the lengths of the displayed expressions agree with the lengths of the underlying elements, and thus they are reduced.)

From the definition of $\underline{w}$ it follows that there exists a sequence of expressions $\emptyset = \underline{x}_0, \underline{x}_1, \dots, \underline{x}_r = \underline{w}$ such that each $\underline{x}_i$ is obtained from $\underline{x}_{i-1}$ by the procedure of the previous paragraph. Thus $\underline{w}$ is reduced as claimed. □

Write $\underline{w} = s_{i_1} \dots s_{i_l}$.

Lemma 5.6. *Any subexpression $\underline{e}$ of $\underline{w}$ with $\underline{w}^{\underline{e}} = w_{A \cup M \cup B}$ has $e_j = 0$ if $s_{i_j} \in \{s_a, s_{a+n}\}$ and $e_j = 1$ if $s_{i_j} \in A \cup B$.*

Proof. Let $\underline{e}$ denote a subexpression of $\underline{w}$ with $\underline{w}^{\underline{e}} = w_{A \cup M \cup B}$.

Any expression $\underline{y}$ for w_A contains a subsequence of the form $s_{a-1} s_{a-2} \dots s_1$ (think about what happens to $1 \in \{1, \dots, a+n+b\}$). In $\underline{w}$, s_1 only occurs once. Left of s_1 there is only one occurrence of s_{a-1}, s_{a-2}, etc. We conclude that the restriction of $\underline{e}$ to $(s_a s_{a-1} \dots s_2 s_1)$ in $\underline{u}_i$ is equal to $(01 \dots 11)$, where i is the maximal index

with $\underline{u}_i \neq \emptyset$. Now any expression for w_A starting in $s_{a-1} \dots s_1$ has to contain a subsequence to the right of the form $s_{a-1} \dots s_2$ (think about what happens to $2 \in \{1, \dots, a+n+b\}$). Continuing in this way we see that the restriction of $\underline{e}$ to each $\underline{u}_j$ has the form

$$(01 \dots 1) \dots (01 \dots 1)(01 \dots 1)$$

(with the same bracketing as in the definition of each $\underline{u}_i$). Similar arguments apply to each $\underline{v}_i$ and the result follows. □

Lemma 5.7. *There is a unique subexpression $\underline{e}$ of $\underline{w}$ such that $\underline{w}^{\underline{e}} = w_{A \cup M \cup B}$ and $\underline{e}$ has defect zero.*

Proof. By the previous lemma we must have $e_j = 0$ (resp. 1) if $s_{i_j} \in \{s_a, s_{a+n}\}$ (resp. $s_{i_j} \in A \cup B$). Because each e_j with $s_{i_j} \in \{s_a, s_{a+n}\}$ is $U0$ and because $W_{A \cup B}$ and W_M commute we only have to understand subexpressions $\underline{e}'$ of

$$\underline{w}' = \underline{w}_m \underline{w}_{m-1} \cdots \underline{w}_1 \underline{w}_M$$

of defect $-(a+b) = -\sum_{i=1}^m \ell(w_i)$ such that $(\underline{w}')^{\underline{e}'} = w_M$. Now $\ell(\underline{w}') = \ell(w_M) + a + b$ and hence any subexpression $\underline{e}'$ of $\underline{w}'$ with $(\underline{w}')^{\underline{e}'} = w_M$ has at most $a+b$ zeroes. Moreover, if $\underline{e}'$ has defect $-a-b$ then $\underline{e}'$ must have exactly $a+b$ zeroes, all of which have to be $D0$. Now, using that $\underline{w}_M$ is reduced, the only subexpression of $\underline{w}'$ fulfilling these requirements is

$$(0 \dots 0)(0 \dots 0) \dots (0 \dots 0)(1 \dots 1).$$ □

Proof of Theorem 5.3. We conclude from the previous two lemmas and their proofs that the unique defect zero subexpression $\underline{e}$ of $\underline{w}$ with $\underline{w}^{\underline{e}} = w_A w_B w_M$ is

$$\underline{e} = \underline{e}_m \underline{f}_m \underline{g}_m \cdots \underline{e}_2 \underline{f}_2 \underline{g}_2 \underline{e}_1 \underline{f}_1 \underline{g}_1 \underline{e}_0$$

where $\underline{e}_0$ (resp. $\underline{e}_i$) is a subexpression of $\underline{w}_M$ (resp. $\underline{w}_i$) given by

$$\underline{e}_0 = (U1, U1, \dots, U1) \quad (\text{resp.} \quad \underline{e}_i = (D0, D0, \dots, D0))$$

and $\underline{f}_i$ (resp. $\underline{g}_i$) is a subexpression of $\underline{u}_i$ (resp. $\underline{v}_i$) given by

$$(U0, U1, \dots, U1)(U0, U1, \dots, U1) \dots (U0, U1, \dots, U1).$$

(we use the same bracketing as in the definition of $\underline{u}_i$ and $\underline{v}_i$).

Hence the intersection form of $\underline{w}$ at $w_{A \cup M \cup B}$ for degree $d = 0$ is indeed a 1×1 matrix. Applying [HW, Theorem 5.1] its unique entry is given by the coefficient of $\delta_{w_{A \cup M \cup B}} = \delta_{w_A} \delta_{w_M} \delta_{w_B}$ in [6]

$$E := E_m F_m G_m \dots E_2 F_2 G_2 E_1 F_1 G_1 E_0$$

[6] Actually, as noted in Remark 5.4, here we use a "right to left" convention, rather than the "left to right" convention of [HW]. One can check that [HW, Theorem 5.1] holds in either convention. Alternatively one can proceed as follows. Let $\underline{w}^r = s_{i_\ell} \dots s_{i_1}$ denote the reversed sequence, and let $\iota : NH \to NH$ denote the anti-involution with $\iota(f) = f$ for $f \in R$ and $\iota(\delta_x) = \delta_{x^{-1}}$ for $x \in W$. Then [HW, Theorem 5.1] implies that the intersection form of $\underline{w}^r$ at $w_{A \cup M \cup B}$ is the 1×1 matrix given by the coefficient of $\delta_{w_A} \delta_{w_M} \delta_{w_B}$ in $\iota(E)$. This implies the statement because the intersection forms of $\underline{w}$ and $\underline{w}^r$ at $w_{A \cup M \cup B}$ agree.

where $E_0 = \delta_{w_M}$, $E_i = \delta_{w_i}$ for $1 \leq i \leq m$ and the F_i, G_i are given by:

$$\begin{aligned}
F_1 &= (\alpha_a \delta_{a-1} \dots \delta_{a-a_1+1}) \dots (\alpha_a \delta_{a-1})(\alpha_a) \\
F_2 &= (\alpha_a \delta_{a-1} \dots \delta_{a-a_1-a_2+1}) \dots (\alpha_a \delta_{a-1} \dots \delta_{a-a_1-1})(\alpha_a \delta_{a-1} \dots \delta_{a-a_1}) \\
&\vdots \\
F_m &= (\alpha_a \delta_{a-1} \dots \delta_1) \dots (\alpha_a \delta_{a-1} \dots \delta_{a-a_1-\dots-a_{m-1}-1})(\alpha_a \delta_{a-1} \dots \delta_{a-a_1-\dots-a_{m-1}}) \\
G_1 &= (\alpha_{a+n} \delta_{a+n+1} \dots \delta_{a+n+b_1-1}) \dots (\alpha_{a+n} \delta_{a+n+1})(\alpha_{a+n}) \\
&\vdots \\
G_m &= (\alpha_{a+n} \delta_{a+n+1} \dots \delta_{a+n+b-1}) \dots (\alpha_{a+n} \delta_{a+n+1} \dots \delta_{a+n+b_1+\dots+b_{m-1}}).
\end{aligned}$$

In NH we can write $E = \sum_{y \in W_{A \cup M \cup B}} f_y \delta_y$. After noting that

$$\deg E = 2(-\ell(w_A) - \ell(w_B) + a + b - \sum \ell(w_i) - \ell(w_M)) = -2\ell(w_{A \cup M \cup B})$$

we see that in fact $E = \kappa' \delta_{w_{A \cup M \cup B}}$ for some $\kappa' \in \mathbb{Z}$. In particular, whenever we apply a nil Hecke relation $f\delta_i = \delta_i s_i(f) + \partial_i(f)$ with $s_i \in A \cup B$ to reduce E the term involving $\partial_i(f)$ does not contribute. (It would lead to a term which is zero for degree reasons.) Hence we can write

$$E = \delta_{w_A} \delta_{w_B} (\delta_{w_m} \gamma_m \delta_m) \dots (\delta_{w_2} \gamma_2 \delta_2)(\delta_{w_1} \gamma_1 \delta_1) \delta_{w_M}$$

where each γ_i (resp. δ_i) is a product of a_i (resp. b_i) roots of the form $x_k - x_{a+1}$ with $k < a+1$ (resp. $x_{a+n} - x_k$ for $k > a+n$). Hence we have

$$\begin{aligned}
E &= \delta_{w_A w_B} (\delta_{w_m} (-x_{a+1})^{a_m} x_{a+n}^{b_m}) \dots (\delta_{w_2} (-x_{a+1})^{a_2} x_{a+n}^{b_2})(\delta_{w_1} (-x_{a+1})^{a_1} x_{a+n}^{b_1}) \delta_{w_M} \\
&= (-1)^a \kappa \cdot \delta_{w_{A \cup M \cup B}}
\end{aligned}$$

where the first (resp. second) equality follows from Lemma 5.8 (resp. 5.9) below. The theorem follows. □

Lemma 5.8. *Let $w_1, \dots, w_m \in W_M$ and $\zeta_1, \dots, \zeta_m \in R$. Assume that for some $1 \leq i \leq m$ we can write $\zeta_i = \zeta_i^M \zeta_i'$ for some W_M-invariant ζ_i^M of positive degree, and that $\sum \deg \zeta_i = \sum \ell(w_i)$. Then $\delta_{w_m} \zeta_m \dots \delta_{w_1} \zeta_1 \delta_{w_M} = 0$.*

Proof. We have $\delta_{w_m} \zeta_m \dots \delta_{w_i} \zeta_i' \dots \delta_{w_1} \zeta_1 \delta_{w_M} \in \bigoplus_{y \in W_M} R\delta_y$, and hence

$$\delta_{w_m} \zeta_m \dots \delta_{w_i} \zeta_i' \dots \delta_{w_1} \zeta_1 \delta_{w_M} = 0$$

because it is of degree $< -2\ell(w_M)$. As ζ_i^M is W_M-invariant:

$$0 = \zeta_i^M (\delta_{w_m} \zeta_m \dots \delta_{w_i} \zeta_i' \dots \delta_{w_1} \zeta_1 \delta_{w_M}) = \delta_{w_m} \zeta_m \dots \delta_{w_1} \zeta_1 \delta_{w_M}. \qquad \square$$

Lemma 5.9. *With w_i, a_i, b_i, κ as above we have:*

$$(\delta_{w_m} x_{a+1}^{a_m} x_{a+n}^{b_m})(\delta_{w_{m-1}} x_{a+1}^{a_{m-1}} x_{a+n}^{b_{m-1}}) \dots (\delta_{w_1} x_{a+1}^{a_1} x_{a+n}^{b_1}) \delta_{w_M} = \kappa \cdot \delta_{w_M}.$$

Proof. It is well known that $\delta_i \mapsto \partial_i, f \mapsto (f\cdot)$ makes R into an NH-module. In NH we can write

$$(\delta_{w_m} x_{a+1}^{a_m} x_{a+n}^{b_m})(\delta_{w_{m-1}} x_{a+1}^{a_{m-1}} x_{a+n}^{b_{m-1}}) \dots (\delta_{w_1} x_{a+1}^{a_1} x_{a+n}^{b_1}) = K + \sum_{\mathrm{id} \neq w \in W_M} f_w \delta_w$$

where $K \in \mathbb{Z}$ for degree reasons. By applying this identity to $1 \in R$ we deduce that $K = \kappa$. The lemma now follows because if $w \in W_M$ then $\delta_w \delta_{w_M} = 0$ unless $w = \mathrm{id}$. □

6. (Counter)-examples

We use the notation of §1.2 and write $\partial_{12} := \partial_1\partial_2$, $X_1 := X_{s_1}, X_{12} := X_{s_1 s_2}$ etc.

6.1. $n < 4$**:** One checks easily using (1.1) and (1.2) that for $n = 2, 3$ one can only obtain $\kappa = \pm 1$.

6.2. $n = 4$**:** Using (1.1) and (1.2) we see that in C we have

$$X_1 = x_1 \qquad \text{and} \qquad X_3 = -x_4. \tag{6.1}$$

Consider the (degree zero) operator $F : C \to C$ given by

$$F : h \mapsto \partial_{23}(x_4^2(\partial_1(x_1 h))).$$

Using (1.1) and (1.2) one checks that F preserves the submodule $\mathbb{Z}X_1 \oplus \mathbb{Z}X_3$ and in the basis X_1, X_3 is given by

$$F = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

This matrix determines the Fibonacci recursion! Hence for $i \geq 1$ we have

$$F^i(x_1) = F_{i+1}X_1 + F_i X_3$$

where $F_1 = 1$, $F_2 = 1$, $F_3 = 2$, $F_4 = 3$ etc. denote the Fibonacci numbers. In particular,

$$\partial_1(F^i(x_1)) = F_{i+1}.$$

We conclude from the main theorem that any prime dividing the Fibonacci number F_{i+1} occurs as torsion in SL_{3i+5}. By Carmichael's theorem [Car14] the first $n \gg 1$ Fibonacci numbers have at least n distinct prime factors. By the prime number theorem we conclude that the torsion in SL_n grows at least as fast as some constant times $n \log n$. Hence no linear bound is sufficient for Lusztig's conjecture.

It is a well-known conjecture that infinitely many Fibonacci numbers are prime. By the above results, this conjecture would immediately imply that the torsion in SL_n grows exponentially in n. Unfortunately, little seems to be known about the rate of growth of prime factors of Fibonacci numbers.

In the appendix we work with different operators in order to establish exponential growth of torsion. If U_l (resp. U_u) denotes the operator $h \mapsto \partial_{21}(x_1^2(\partial_1(x_1 h)))$ (resp. $h \mapsto \partial_{23}(x_4^2(\partial_3(x_4 h)))$) then U_l and U_u preserve the submodule $\mathbb{Z}X_1 \oplus \mathbb{Z}X_3$ and in the basis X_1, X_3 are given by

$$U_l = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad U_u = \begin{pmatrix} -1 & -1 \\ 0 & -1 \end{pmatrix}.$$

It follows from our main theorem that any prime dividing any matrix coefficient of any word of length ℓ in the generators $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ occurs as torsion in $\mathrm{SL}_{3\ell+5}$. Indeed, given any word $\omega_1\omega_2\dots\omega_r$ in the operators U_l and U_u we may obtain all four coefficients (up to sign) of the corresponding product of the matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ as $\partial_i(\omega_1(\dots(\omega_r(x_j))\dots))$ for $i \in \{1, 3\}$ and $j \in \{1, 4\}$ (use (6.1) and the fact that $\partial_1(X_1) = 1 = \partial_3(X_3)$).

6.3. $n = 5$**:** In the following table we list some examples of p torsion in SL_N found using $n = 5$. The entries in the list were found by random computer searches:

N	14	17	20	22	25	30	40	50	70	100
p	3	7	13	23	53	197	2 237	34 183	4 060 219	470 858 183

(These entries were found as follows. Consider the eight degree zero operators:

$$\partial_{4321}x_1^4, \partial_{321}x_1^3, \partial_{21}x_1^2, \partial_1 x_1, \partial_{1234}x_5^4, \partial_{234}x_5^3, \partial_{34}x_5^2, \partial_4 x_5.$$

It is not difficult to calculate the matrices of these operators acting on any homogenous component of C in the Schubert basis. The above entries were obtained as prime factors of coefficients obtained by repeated application of these operators to x_1^3 and $x_1^2 x_5 \in C^6$.)

7. Lusztig conjecture

This section consists of connections and complements to [Soe00], with which we assume the reader is familiar. In keeping with the setting of this paper, we work with $G = \mathrm{SL}_n$ throughout, however analogous statements are true (with the same proofs) for any connected reductive group.

As in §3 we assume in this and the following section that $p > n$. Let $\mathcal{O}$ denote the "regular subquotient around the Steinberg weight" as defined in [Soe00, §2.3]. We denote by $\Delta(x), P(x)$ the standard and projective objects in $\mathcal{O}$ and by $\theta_s : \mathcal{O} \to \mathcal{O}$ for $s \in S$ the translation functor [Soe00, §2.5]. Let $[\mathcal{O}]$ denote the Grothendieck group of $\mathcal{O}$ and

$$\alpha : [\mathcal{O}] \xrightarrow{\sim} \mathbb{Z}[W]$$

the isomorphism with $\alpha([\Delta(x)]) = x$ for all $x \in W$ (α is denoted A in [Soe00, §2.10]). As observed in [Soe00], Lusztig's conjecture implies that

$$\alpha(P(x)) = \underline{H}_{x|v=1} \qquad \text{for all } x \in W.$$

Remark 7.1. This observation should be compared with a much earlier theorem of Jantzen [Jan79, Anhang, Corollar] matching multiplicities of simple modules in Weyl modules in sufficiently large characteristic p and multiplicities of simple modules in Verma modules in characteristic 0. This observation, together with Jantzen's calculations in rank 2, were the main ingredients that led to the formulation of the Lusztig conjecture.

Proposition 7.2. *We have $\alpha(P(x)) = {}^p\underline{H}_{x|v=1}$. In particular, if ${}^p\underline{H}_x \neq \underline{H}_x$ with $p > n$ then Lusztig's conjecture fails for SL_n in characteristic p.*

Remark 7.3. Recall that ${}^p a_{y,x}, {}^p h_{y,x} \in \mathbb{Z}_{\geq 0}[v^{\pm 1}]$ (see §2). In particular:

$${}^p\underline{H}_x = \underline{H}_x \Leftrightarrow {}^p\underline{H}_{x|v=1} = \underline{H}_{x|v=1}.$$

Proof. Let $p\mathcal{O}$ denote the full subcategory of projective objects in $\mathcal{O}$, and $[p\mathcal{O}]$ its split Grothendieck group. Because $\mathcal{O}$ has finite homological dimension, the map $[p\mathcal{O}] \to [\mathcal{O}]$ induced by the inclusion is an isomorphism. Recall the commutative

diagram

$$\begin{array}{ccc} [p\mathcal{O}] & \xrightarrow{\mathbb{V}} & [f\mathcal{C}] \\ \alpha \downarrow & & \beta' \downarrow \\ \mathbb{Z}S_n & = & \mathbb{Z}S_n \end{array}$$

from [Soe00, §2.11] (see §3 for the definition of $f\mathcal{C}$).

We claim that β' above agrees with the β defined in §3. If $\underline{w} = st \dots u$ we have, by [Soe00, Theorem 2.6.2]:

$$\mathbb{V}(\theta_u \dots \theta_t \theta_s M_{\mathrm{id}}) \cong fD_{\underline{w}}.$$

Thus, by [Soe00, §2.5 and §2.10]:

$$\beta'([fD_{\underline{w}}]) = \alpha(\theta_u \dots \theta_t \theta_s M_{\mathrm{id}}) = (1+s)(1+t)\dots(1+u).$$

By the commutativity of the diagram in §3:

$$\beta([fD_{\underline{w}}]) = \mathrm{ch}([B_s B_t \dots B_u])_{|v=1} = (\underline{H}_s \underline{H}_t \dots \underline{H}_u)_{|v=1} = (1+s)(1+t)\dots(1+u).$$

Hence $\beta = \beta'$ as claimed, as $[f\mathcal{C}]$ is generated by $[fD_{\underline{w}}]$ over all expressions $\underline{w}$.

Now we are done: by [Soe00, Theorem 2.8.2] we have $\mathbb{V}P(x) = fD_x$ and the proposition follows from (3.2). □

8. James conjecture

In this section we explain why the results of the previous section yield counter-examples to the James conjecture [Jam90] on the decomposition numbers of Schur algebras and the symmetric group.

Fix positive integers N and r. Let $\Lambda^+(N,r)$ denote the set of partitions of r into at most N parts; that is, sequences $(\lambda_1, \dots, \lambda_N)$ such that $\lambda_1 \geq \dots \geq \lambda_N \geq 0$ with $r = \sum \lambda_i$. Then $\Lambda^+(N,r)$ is a partially ordered set with respect to the dominance order $\leq$.

Let $S(N,r)$ denote the Schur algebra over $\mathbb{Z}$ (see e.g. [Gre81]). Its category of representations is equivalent to the category of polynomial representations of the group scheme GL_N of fixed degree r. Fix a field $\Bbbk$ of characteristic $p > N$ and let $S_{\Bbbk}(N,r)$ denote the Schur algebra over $\Bbbk$. The category $\mathrm{Rep}\, S_{\Bbbk}(N,r)$ of finitely generated $S_{\Bbbk}(N,r)$-modules is a highest weight category with simple modules indexed by $\Lambda^+(N,r)$. Given $\lambda \in \Lambda^+(N,r)$ we denote by $L(\lambda)$ (resp. $\Delta(\lambda)$, $\nabla(\lambda)$, $P(\lambda)$, $T(\lambda)$) the simple (resp. standard, costandard, indecomposable projective, indecomposable tilting) module indexed by λ.

Let $S_q(N,r)$ denote the q-Schur algebra and $S_\varepsilon(N,r)$ its specialisation at a fixed primitive p^{th}-root of unity $\varepsilon \in \mathbb{C}$ (see e.g. [Don98]). Then the category of finitely generated $S_\varepsilon(N,r)$-modules is highest weight. As above we write $L_\varepsilon(\lambda)$ (resp. $\Delta_\varepsilon(\lambda)$ etc.) for the simple (resp. standard etc.) module corresponding to $\lambda \in \Lambda^+(N,r)$. Given a module M for $S_\varepsilon(N,r)$ we may choose a stable $\mathbb{Z}[\varepsilon]$-lattice and reduce to obtain a module over $S_{\Bbbk}(N,r)$. In this way we obtain the decomposition map on Grothendieck groups

$$d : [\mathrm{Rep}\, S_\varepsilon(N,r)] \to [\mathrm{Rep}\, S_{\Bbbk}(N,r)].$$

One has $d([\nabla_\varepsilon(\lambda)]) = [\nabla(\lambda)]$. The James conjecture [Jam90] predicts that

$$d([L_\varepsilon(\lambda)]) = [L(\lambda)] \tag{8.1}$$

if $p > \sqrt{r}$.[7]

Let $\rho = (N-1, \dots, 1, 0) \in \Lambda^+(N, \binom{N}{2})$ and let $\mathrm{st} := (p-1)\rho$ denote the "Steinberg weight". Let S_N denote the symmetric group of N letters, acting by permutation on $\mathbb{Z}^N$. Lusztig's quantum character formula gives (see (2.1) for notation)

$$[\nabla_\varepsilon(\mathrm{st} + x\rho) : L_\varepsilon(\mathrm{st} + y\rho)] = h_{x,y}(1). \tag{8.2}$$

as modules for $S_\varepsilon(N, p\binom{N}{2})$. (Actually, Lusztig's quantum character formula gives the multiplicity for the quantum group of $\mathfrak{sl}_N$ specialised at ε in terms of an affine Kazhdan-Lusztig polynomial. The translation of his formula to yield the above multiplicity is standard but technical. Alternatively, one can appeal to [AJS94] and the $p \gg 0$ version of (8.3) below.)

Similarly, [Soe00, Theorem 1.2.2] gives (again see (2.1) for notation)

$$[\nabla(\mathrm{st} + x\rho) : L(\mathrm{st} + y\rho)] = {}^p h_{x,y}(1) \tag{8.3}$$

as modules for $S_{\Bbbk}(N, p\binom{N}{2})$. (Actually, Soergel's result gives this multiplicity for rational modules for $\mathrm{SL}_N(\Bbbk)$. The translation to $GL_N(\Bbbk)$ and hence to modules for the Schur algebra is standard.)

We conclude that whenever ${}^p\underline{H}_x \neq \underline{H}_x$ for some x the characters of the simple modules for $S_\varepsilon(N, p\binom{N}{2})$ and $S_{\Bbbk}(N, p\binom{N}{2})$ are different, because the simple and costandard modules both give bases for the Grothendieck group. In particular, there exists λ such that $d([L_\varepsilon(\lambda)]) \neq [L(\lambda)]$. Hence any p appearing on the table in Section 6 with $p > \binom{N}{2}$ contradicts the James conjecture for $S(N, p\binom{N}{2})$.

Remark 8.1. A straightforward computation in $[\mathrm{Rep}\, S_\varepsilon(N, r)]$ and $[\mathrm{Rep}\, S_{\Bbbk}(N, r)]$ shows that

$$d([L_\varepsilon(\mathrm{st} + x\rho)]) = \sum a_{x,y}(1)[L(\mathrm{st} + y\rho)]$$

and so the $a_{x,y}$ evaluated at 1 give part of James's "adjustment matrix".

However, amongst weights of the form $\mathrm{st} + x\rho$ for $x \in S_N$ only $\mathrm{st} + w_0\rho$ is p-restricted (w_0 denotes the longest element of S_N). Hence the above non-trivial decomposition numbers are invisible to the symmetric group, as all simples corresponding to non p-restricted weights are killed by the Schur functor.

To get counter-examples in the symmetric group we can use the Ringel self-duality of the Schur algebra and modular category $\mathcal{O}$. (I thank Joe Chuang for explaining this to me.) Given any $N' \geq N$ we have an obvious embedding $\Lambda^+(N, r) \hookrightarrow \Lambda^+(N', r)$ obtained by appending 0's to the partition. There is a quotient functor $f : \mathrm{Rep}\, S(N', r) \to \mathrm{Rep}\, S(N, r)$ which preserves simple, standard, costandard modules and indecomposable tilting modules corresponding to λ in $\Lambda^+(N, r) \subset \Lambda^+(N', r)$ (see [Gre81, §6.5] and [Don98, §A4.5]).

Now consider a variant of the subquotient around the Steinberg weight discussed in the previous section. Consider the Serre subquotient $\mathcal{O} := \mathcal{A}/\mathcal{N}$ of $\mathrm{Rep}\, S_{\Bbbk}(N, p\binom{N}{2})$ where $\mathcal{A}$ is the Serre subcategory generated by simple modules $L(\lambda)$ such that $\lambda \leq st + \rho$ and λ lies in the same block as $st + \rho$, and $\mathcal{N}$ denotes the Serre subcategory generated by those simples $L(\lambda) \in \mathcal{A}$ which are not of the form $L(\mathrm{st} + x\rho)$ for some $x \in S_N$. The definition of $\mathcal{O}_\varepsilon$ is obtained by replacing $L(\lambda)$ by $L_\varepsilon(\lambda)$ in the definition of $\mathcal{O}$.

[7] A stronger version requires that p be larger than the weight of λ. It reduces to the condition $p > \sqrt{r}$ for the principal block, which will be the only case considered below.

Let us denote the images of $L(\mathrm{st}+x\rho)$, $\Delta(\mathrm{st}+x\rho)$, etc. in $\mathcal{O}$ as $L(x)$, $\Delta(x)$, etc. and similarly for $\mathcal{O}_\varepsilon$. Then $\mathcal{O}$ is a highest weight category with simple, standard, etc. and tilting objects $L(x)$, $\Delta(x)$, etc., and similarly for $\mathcal{O}_\varepsilon$. It is known that both categories are Ringel self-dual. (The proof of this fact seems not to be explicit in the literature. However a proof may be obtained by adapting ideas of [BBM04]. One shows that one has a braid group action on $D^b(\mathcal{O})$ (resp. $D^b(\mathcal{O}_\varepsilon)$) and a lift of the longest element interchanges a projective and tilting generator.) Applying BGG reciprocity then Ringel self-duality for $\mathcal{O}$ and $\mathcal{O}_\varepsilon$ we obtain:

$$h_{x,y}(1) = [\nabla_\varepsilon(x) : L_\varepsilon(y)] = (P_\varepsilon(y) : \Delta_\varepsilon(x)) = (T_\varepsilon(yw_0) : \nabla_\varepsilon(xw_0)),$$
$$^p h_{x,y}(1) = [\nabla(x) : L(y)] = (P(y) : \Delta(x)) = (T(yw_0) : \nabla(xw_0)).$$

Applying Ringel self-duality of $S(p\binom{N}{2}, p\binom{N}{2})$ [Don93] and of $S_\varepsilon(p\binom{N}{2}, p\binom{N}{2})$ [Don98] we have:

$$h_{x,y}(1) = (T_\varepsilon(st + yw_0\rho) : \nabla_\varepsilon(st + xw_0\rho)) = (P_\varepsilon((st + yw_0\rho)') : \Delta_\varepsilon((st + xw_0\rho)')),$$
$$^p h_{x,y}(1) = (T(st + yw_0\rho) : \nabla(st + xw_0\rho)) = (P((st + yw_0\rho)') : \Delta((st + xw_0\rho)')).$$

Finally, again by BGG reciprocity

$$h_{x,y}(1) = [\nabla_\varepsilon((st + xw_0\rho)') : L_\varepsilon((st + yw_0\rho)')],$$
$$^p h_{x,y}(1) = [\nabla((st + xw_0\rho)') : L((st + yw_0\rho)')].$$

The partitions $(\mathrm{st}+xw_0\rho)'$ and $(\mathrm{st}+yw_0\rho)'$ are p-restricted. Hence, after applying the Schur functor the first (resp. second) number can be interpreted as a decomposition number for the Hecke algebra specialised at ε (resp. the symmetric group in characteristic p). It follows that the results of the previous section also produce counter-examples for the symmetric group.

Remark 8.2. Consulting the table of counter-examples in Section 6 we see that the smallest counter-example produced by the above methods occurs in $S_{N'}$ with

$$N' = p\binom{N}{2} = 2237\binom{40}{2} = 1\ 744\ 860.$$

The size of this number is a relic of our method (in particular the fact that we cannot say anything about p-restricted weights). It is an important question as to where the first counter-examples occur.

Appendix A. Exponential Growth of Torsion

by Alex Kontorovich, Peter J. McNamara and Geordie Williamson. [8]

A.1. **Statement of the theorem.** Let

$$\Gamma := \left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\rangle^+ \tag{A.1}$$

be the sub-*semi*-group of $\mathrm{SL}(2,\mathbb{Z})$ generated (freely) by the matrices displayed. For a matrix $\gamma \in \Gamma$, let $\ell(\gamma)$ be its wordlength in the generators of Γ. In the main

[8]Kontorovich is partially supported by an NSF CAREER grant DMS-1455705 and an Alfred P. Sloan Research Fellowship.

body of the paper, the third-named author proves that any prime p dividing any coefficient γ_{ij} of any matrix

$$\gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \in \Gamma$$

occurs as torsion in $\mathrm{SL}_{3\ell+5}$, where $\ell = \ell(\gamma)$ is the wordlength (see §6.2). The purpose of this appendix is to show the existence of exponentially large (relative to wordlength) prime divisors of matrix coefficients in Γ, thus giving exponentially large counterexamples to the expected bounds in Lusztig's conjecture.

In fact, the stated purpose can be accomplished by an almost[9] direct application of the Affine Sieve [BGS06, BGS10, SGS13]; see also, e.g., [Kon14]. It turns out that one can do much more using recent progress on "local-global" problems in "thin orbits" (see, e.g., the discussion in [Kon13]); namely, one can produce not just prime divisors but actual primes in the entries of Γ, and moreover give explicit estimates for their exponential growth rates (which are far superior compared to those which would come from an Affine Sieve analysis). Our main result is the following

Theorem A.1. *There are absolute constants $\tau > 0$ and $c > 1$ so that, for all L large, there exists $\gamma \in \Gamma$ of wordlength $\ell(\gamma) \leq L$ and top-left entry $\gamma_{11} = p$ prime with $p > \tau c^L$. In fact, there are many primes arising this way:*

$$\#\left\{p > \tau c^L \ : \ \exists \gamma \in \Gamma \text{ with } \ell(\gamma) \leq L \text{ and } \gamma_{11} = p\right\} \gg \frac{c^L}{L}. \tag{A.2}$$

The implied constant above is absolute and effective.

Throughout this appendix, p always denotes a prime. The notation $f(L) \gg g(L)$ means that $g = O(f)$, i.e. $|g(L)| \leq M|f(L)|$ for a fixed $M > 0$ and all large L. In this case, M is the *implied constant* referred to above.

Exact estimates for τ and c can be readily determined; the value coming from our proof is $c = (\frac{1+\sqrt{5}}{2})^{1/5} \approx 1.101\ldots$ and we can take $\tau = 5/7$; see (A.7). It turns out that Theorem A.1 is a nearly immediate consequence of recent advances towards Zaremba's conjecture on continued fractions with bounded partial quotients.

Given $A \geq 1$, let Γ_A be the sub-semigroup:

$$\Gamma_A := \left\langle \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix} \ : \ 1 \leq a, b \leq A \right\rangle^+. \tag{A.3}$$

(In fact, Γ_A is freely generated by the displayed elements.)

Theorem A.2 (Bourgain-Kontorovich [BK14])**.** *There exists A_0 and an absolute constant $\mathfrak{c} < \infty$ so that, for $A \geq A_0$ and all N large,*

$$\#\{n \leq N \ : \ \exists \gamma \in \Gamma_A \text{ with } \gamma_{11} = n\} \ = \ N\left(1 + O\left(e^{-\mathfrak{c}\sqrt{\log N}}\right)\right),$$

where the implied constant and $\mathfrak{c} > 0$ are both absolute.

That is, almost all *integers* (not just primes) arise as top-left entries in the semigroup Γ_A. Bourgain-Kontorovich give $A_0 = 50$ as an allowable value for A, and this has since been reduced to $A_0 = 5$ [FK14, Hua15]; furthermore, Hensley [Hen96] has conjectured that $A_0 = 2$ is allowable, and that the error rate $O(e^{-\mathfrak{c}\sqrt{\log N}})$ can

[9]For the application to be immediate, Γ would need to be a Zariski-dense *group* and not just a semi-group; minor modifications are needed to handle this case.

be replaced by $O(1/N)$. What is most important to our application is that the error rate is asymptotically $o(1/\log N)$. This, together with the Prime Number Theorem, has the following immediate

Corollary A.3. *Let notation be as above and set $A = 5$. Then for any fixed constant $\theta < 1$,*

$$\#\{p \in (\theta N, N] \ : \ \exists \gamma \in \Gamma_A \text{ with } \gamma_{11} = p\} \ = \ (1-\theta)\frac{N}{\log N}\Big(1 + o(1)\Big), \tag{A.4}$$

as $N \to \infty$.

Equipped with this estimate, is it a simple matter to give the

A.2. **Proof of Theorem A.1.** Fix constants $c > 1$ and $\tau > 0$ to be chosen later, and let $\mathcal{S}_1$ denote the set of primes on the left hand side of (A.2),

$$\mathcal{S}_1 \ := \ \left\{p > \tau c^L \ : \ \exists \gamma \in \Gamma \text{ with } \ell(\gamma) \leq L \text{ and } \gamma_{11} = p\right\}.$$

We seek a lower bound on the cardinality of $\mathcal{S}_1$.

For a parameter A (which we will soon set to $A = 5$) and a matrix $\gamma \in \Gamma_A$, let $\ell_A(\gamma)$ denote the wordlength in the generators of Γ_A given in (A.3). We make the pleasant observation that

$$\begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix} \ = \ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix},$$

and hence Γ_A is a sub-semigroup of Γ. Moreover, if $\gamma \in \Gamma_A \subset \Gamma$, then the wordlengths in the two semigroups are related by

$$\ell(\gamma) \ \leq \ 2A \cdot \ell_A(\gamma),$$

since each generator in Γ_A has wordlength at most $2A$ in the generators of Γ. We decrease $\mathcal{S}_1$ to a smaller set $\mathcal{S}_2 \subset \mathcal{S}_1$ of primes coming from the top-left entries of Γ_A instead of Γ:

$$\mathcal{S}_2 \ := \ \left\{p > \tau c^L \ : \ \exists \gamma \in \Gamma_A \text{ with } \ell_A(\gamma) \leq L/(2A) \text{ and } \gamma_{11} = p\right\}.$$

Next we define the archimedean sup-norm

$$\|\gamma\|_\infty \ := \ \max(\gamma_{ij}),$$

which for $\gamma \in \Gamma_A$ is easily seen to be the top left entry

$$\|\gamma\|_\infty \ = \ \gamma_{11}. \tag{A.5}$$

Let

$$\varphi \ := \ \frac{1+\sqrt{5}}{2} \qquad \text{and} \qquad \overline{\varphi} \ := \ \frac{1-\sqrt{5}}{2}$$

denote the eigenvalues of $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

For any

$$\gamma = \prod_{i=1}^{n} \left(\begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_i & 1 \\ 1 & 0 \end{pmatrix} \right) \in \Gamma_A,$$

we have

$$\|\gamma\|_\infty = \begin{pmatrix} 1 & 0 \end{pmatrix} \gamma \begin{pmatrix} 1 \\ 0 \end{pmatrix} \geq \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{2n} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = F_{2n+1}$$

where F_m is the m-th Fibonacci number. Because $F_{2n+1} = (\varphi^{2n+1} - \overline{\varphi}^{2n+1})/\sqrt{5}$, if we set $d := \varphi/\sqrt{5}$, then for all $\gamma \in \Gamma_A$,

$$\|\gamma\|_\infty \ \geq \ d \cdot \varphi^{2\ell_A(\gamma)}.$$

That is, the logarithm of the archimedean norm is controlled (up to a constant) by the wordlength. Define the "archimedean" parameter N (with respect to L) by

$$N \ := \ d \cdot \varphi^{L/A}. \tag{A.6}$$

Replacing the wordlength condition $\ell_A(\gamma) \leq L/(2A)$ in $\mathcal{S}_2$ by the stronger restriction that $\|\gamma\|_\infty \leq N$ decreases $\mathcal{S}_2$ to a subset $\mathcal{S}_3$ defined by

$$\mathcal{S}_3 \ := \ \left\{p > \tau c^L \ : \ \exists \gamma \in \Gamma_A \text{ with } \|\gamma\|_\infty \leq N \text{ and } \gamma_{11} = p\right\}.$$

Since $\gamma_{11} = p = \|\gamma\|_\infty$, the condition $\|\gamma\|_\infty \leq N$ can be replaced by $p \leq N$; hence

$$\mathcal{S}_3 \ = \ \left\{\tau c^L < p \leq N \ : \ \exists \gamma \in \Gamma_A \text{ with } \gamma_{11} = p\right\}.$$

Make the choice

$$c \ = \ \varphi^{1/A}, \tag{A.7}$$

which is $(\frac{1+\sqrt{5}}{2})^{1/5} \approx 1.101\ldots$ when $A = 5$. Then for any $\theta < 1$, take $\tau = \theta d$. With these choices of parameters, we see that

$$\mathcal{S}_3 \ = \ \left\{\theta N < p \leq N \ : \ \exists \gamma \in \Gamma_A \text{ with } \gamma_{11} = p\right\}.$$

Now we are done: combining the above with (A.6) and (A.4) gives

$$\#\mathcal{S}_1 \ \geq \ \#\mathcal{S}_3 \ \gg \ \frac{N}{\log N} \ \gg \ \frac{c^L}{L},$$

as claimed in (A.2). This completes the proof of Theorem A.1.

References

[ABG04] S. Arkhipov, R. Bezrukavnikov, and V. Ginzburg. Quantum groups, the loop Grassmannian, and the Springer resolution. *J. Amer. Math. Soc.*, 17(3):595–678, 2004.

[AJS94] H. H. Andersen, J. C. Jantzen, and W. Soergel. Representations of quantum groups at a pth root of unity and of semisimple groups in characteristic p: independence of p. *Astérisque*, (220):321, 1994.

[BBM04] A. Beilinson, R. Bezrukavnikov, and I. Mirković. Tilting exercises. *Mosc. Math. J.*, 4(3):547–557, 782, 2004.

[BGS06] J. Bourgain, A. Gamburd, and P. Sarnak. Sieving and expanders. *C. R. Math. Acad. Sci. Paris*, 343(3):155–159, 2006.

[BGS10] J. Bourgain, A. Gamburd, and P. Sarnak. Affine linear sieve, expanders, and sum-product. *Invent. Math.*, 179(3):559–644, 2010.

[BK14] J. Bourgain and A. Kontorovich. On Zaremba's conjecture. *Ann. of Math. (2)*, 180(1):137–196, 2014.

[BM13] R. Bezrukavnikov and I. Mirković. Representations of semisimple Lie algebras in prime characteristic and the noncommutative Springer resolution. *Ann. of Math. (2)*, 178(3):835–919, 2013.

[BMR08] R. Bezrukavnikov, I. Mirković, and D. Rumynin. Localization of modules for a semisimple Lie algebra in prime characteristic. *Ann. of Math. (2)*, 167(3):945–991, 2008. With an appendix by Bezrukavnikov and Simon Riche.

[Car14] R. D. Carmichael. On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$. *Ann. of Math. (2)*, 15(1-4):49–70, 1913/14.

[dCM02] M. A. A. de Cataldo and L. Migliorini. The hard Lefschetz theorem and the topology of semismall maps. *Ann. Sci. École Norm. Sup. (4)*, 35(5):759–772, 2002.

[dCM05] M. A. A. de Cataldo and L. Migliorini. The Hodge theory of algebraic maps. *Ann. Sci. École Norm. Sup. (4)*, 38(5):693–750, 2005.

[Don93] S. Donkin. On tilting modules for algebraic groups. *Math. Z.*, 212(1):39–60, 1993.
[Don98] S. Donkin. *The q-Schur algebra*. Cambridge: Cambridge University Press, 1998.
[EK10] B. Elias and M. Khovanov. Diagrammatics for Soergel categories. *Int. J. Math. Math. Sci.*, 2010:58, 2010.
[Eli16] B. Elias. The two-color Soergel calculus. *Compos. Math.*, 152(2):327–398, 2016.
[EW] B. Elias and G. Williamson. Soergel calculus. Represent. Theory, to appear. arXiv:1309.0865.
[EW14] B. Elias and G. Williamson. The Hodge theory of Soergel bimodules. *Ann. of Math. (2)*, 180(3):1089–1136, 2014.
[Fie08] P. Fiebig. The combinatorics of Coxeter categories. *Trans. Am. Math. Soc.*, 360(8):4211–4233, 2008.
[Fie10] P. Fiebig. The multiplicity one case of Lusztig's conjecture. *Duke Math. J.*, 153(3):551–571, 2010.
[Fie11] P. Fiebig. Sheaves on affine Schubert varieties, modular representations, and Lusztig's conjecture. *J. Amer. Math. Soc.*, 24(1):133–181, 2011.
[Fie12] P. Fiebig. An upper bound on the exceptional characteristics for Lusztig's character formula. *J. Reine Angew. Math.*, 673:1–31, 2012.
[FK14] D. A. Frolenkov and I. D. Kan. A strengthening of a theorem of Bourgain-Kontorovich II. *Mosc. J. Comb. Number Theory*, 4(1):78–117, 2014.
[FW14] P. Fiebig and G. Williamson. Parity sheaves, moment graphs and the p-smooth locus of Schubert varieties. *Ann. Inst. Fourier (Grenoble)*, 64(2):489–536, 2014.
[Gre81] J. A. Green. Polynomial representations of GL_n. Algebra, Proc. Conf., Carbondale 1980, Lect. Notes Math. 848, 124-140, 1981.
[Hen96] D. Hensley. A polynomial time algorithm for the Hausdorff dimension of continued fraction Cantor sets. *J. Number Theory*, 58(1):9–45, 1996.
[Hua15] S. Huang. An improvement to Zaremba's conjecture. *Geom. Funct. Anal.*, 25(3):860–914, 2015.
[HW] X. He and G. Williamson. Soergel calculus and Schubert calculus. Bull. Inst. Math. Acad. Sin. (N.S.), to appear. arXiv:1502.04914.
[Jam90] G. James. The decomposition matrices of $GL_n(q)$ for $n \leq 10$. *Proc. Lond. Math. Soc. (3)*, 60(2):225–265, 1990.
[Jan79] J. C. Jantzen. *Moduln mit einem höchsten Gewicht*, volume 750 of *Lecture Notes in Mathematics*. Springer, Berlin, 1979.
[Jan03] J. C. Jantzen. *Representations of algebraic groups*, volume 107 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, second edition, 2003.
[Jan08] J. C. Jantzen. Character formulae from Hermann Weyl to the present. In *Groups and analysis*, volume 354 of *London Math. Soc. Lecture Note Ser.*, pages 232–270. Cambridge Univ. Press, Cambridge, 2008.
[JMW14] D. Juteau, C. Mautner, and G. Williamson. Parity sheaves. *J. Amer. Math. Soc.*, 27(4):1169–1212, 2014.
[JW] L. T. Jensen and G. Williamson. The p-canonical basis for Hecke algebras. To appear in Perspectives in categorification. arXiv:1304.1448.
[Kat85] S.-i. Kato. On the Kazhdan-Lusztig polynomials for affine Weyl groups. *Adv. in Math.*, 55(2):103–130, 1985.
[KL93] D. Kazhdan and G. Lusztig. Tensor structures arising from affine Lie algebras. I, II. *J. Amer. Math. Soc.*, 6(4):905–947, 949–1011, 1993.
[KL94a] D. Kazhdan and G. Lusztig. Tensor structures arising from affine Lie algebras. III. *J. Amer. Math. Soc.*, 7(2):335–381, 1994.
[KL94b] D. Kazhdan and G. Lusztig. Tensor structures arising from affine Lie algebras. IV. *J. Amer. Math. Soc.*, 7(2):383–453, 1994.
[Kon13] A. Kontorovich. From Apollonius to Zaremba: local-global phenomena in thin orbits. *Bull. Amer. Math. Soc. (N.S.)*, 50(2):187–228, 2013.
[Kon14] A. Kontorovich. Levels of distribution and the affine sieve. *Ann. Fac. Sci. Toulouse Math. (6)*, 23(5):933–966, 2014.
[KT95] M. Kashiwara and T. Tanisaki. Kazhdan-Lusztig conjecture for affine Lie algebras with negative level. *Duke Math. J.*, 77(1):21–62, 1995.
[KT96] M. Kashiwara and T. Tanisaki. Kazhdan-Lusztig conjecture for affine Lie algebras with negative level. II. Nonintegral case. *Duke Math. J.*, 84(3):771–813, 1996.

[Lib10] N. Libedinsky. Presentation of right-angled Soergel categories by generators and relations. *J. Pure Appl. Algebra*, 214(12):2265–2278, 2010.

[Lib15] N. Libedinsky. Light leaves and Lusztig's conjecture. *Adv. Math.*, 280:772–807, 2015.

[Lus80] G. Lusztig. Some problems in the representation theory of finite Chevalley groups. In *The Santa Cruz Conference on Finite Groups (Univ. California, Santa Cruz, Calif., 1979)*, volume 37 of *Proc. Sympos. Pure Math.*, pages 313–317. Amer. Math. Soc., Providence, R.I., 1980.

[Lus94] G. Lusztig. Monodromic systems on affine flag manifolds. *Proc. Roy. Soc. London Ser. A*, 445(1923):231–246, 1994.

[SGS13] A. Salehi Golsefidy and P. Sarnak. The affine sieve. *J. Amer. Math. Soc.*, 26(4):1085–1105, 2013.

[Soe90] W. Soergel. Kategorie $\mathcal{O}$, perverse Garben und Moduln über den Koinvarianten zur Weylgruppe. *J. Amer. Math. Soc.*, 3(2):421–445, 1990.

[Soe92] W. Soergel. The combinatorics of Harish-Chandra bimodules. *J. Reine Angew. Math.*, 429:49–74, 1992.

[Soe97] W. Soergel. Kazhdan-Lusztig polynomials and a combinatoric[s] for tilting modules. *Represent. Theory*, 1:83–114 (electronic), 1997.

[Soe00] W. Soergel. On the relation between intersection cohomology and representation theory in positive characteristic. *J. Pure Appl. Algebra*, 152(1-3):311–335, 2000.

[Soe01] W. Soergel. Langlands' philosophy and Koszul duality. In *Algebra—representation theory (Constanta, 2000)*, volume 28 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 379–414. Kluwer Acad. Publ., Dordrecht, 2001.

[Soe07] W. Soergel. Kazhdan-Lusztig-Polynome und unzerlegbare Bimoduln über Polynomringen. *J. Inst. Math. Jussieu*, 6(3):501–525, 2007.

[WB12] G. Williamson and T. Braden. Modular intersection cohomology complexes on flag varieties. *Math. Z.*, 272(3-4):697–727, 2012.

[Wil] G. Williamson. On torsion in the intersection cohomology of Schubert varieties. J. Algebra, to appear. arXiv:1512.08295.

[Wil12] G. Williamson. Some applications of parity sheaves. *Oberwolfach reports*, 2012.

Rutgers University, New Brunswick, NJ
E-mail address: alex.kontorovich@rutgers.edu

University of Queensland, Brisbane, QLD, Australia.
E-mail address: p.mcnamara@uq.edu.au

Max-Planck-Institut für Mathematik, Vivatsgasse 7, 53111, Bonn, Germany.
E-mail address: geordie@mpim-bonn.mpg.de