

POP-STACK-SORTING FOR COXETER GROUPS

Colin Defant¹

¹*Department of Mathematics, Princeton University, Princeton, New Jersey, U.S.A.
colindefant@gmail.com*

Submitted: May 27, 2021; Accepted: Sep 2, 2022; Published: Oct 15, 2022

© The author. Released under the CC BY license (International 4.0).

Abstract. Let W be an irreducible Coxeter group. We define the *Coxeter pop-stack-sorting operator* $\text{Pop} : W \rightarrow W$ to be the map that fixes the identity element and sends each nonidentity element w to the meet of the elements covered by w in the right weak order. When W is the symmetric group S_n , Pop coincides with the pop-stack-sorting map. Generalizing a theorem about the pop-stack-sorting map due to Ungar, we prove that

$$\sup_{w \in W} |O_{\text{Pop}}(w)| = h,$$

where h is the Coxeter number of W (with $h = \infty$ if W is infinite) and $O_f(w)$ denotes the forward orbit of w under a map f . When W is finite, this result is equivalent to the statement that the maximum number of terms appearing in the Brieskorn normal form of an element of W is $h - 1$. More generally, we define a map $f : W \rightarrow W$ to be *compulsive* if for every $w \in W$, $f(w)$ is less than or equal to $\text{Pop}(w)$ in the right weak order. We prove that if f is compulsive, then $\sup_{w \in W} |O_f(w)| \leq h$. This result is new even for symmetric groups.

We prove that 2-pop-stack-sortable elements in type B are in bijection with 2-pop-stack-sortable permutations in type A , which were enumerated by Pudwell and Smith. Claesson and Guðmundsson proved that for each fixed nonnegative integer t , the generating function that counts t -pop-stack-sortable permutations in type A is rational; we establish analogous results in types B and \tilde{A} .

Keywords. Pop-stack-sorting, Coxeter group, weak order, Coxeter number, compulsive map, regular language

Mathematics Subject Classifications. 05E16, 37E15, 05A05

1. Introduction

1.1. Sorting operators

Noninvertible combinatorial dynamics is the study of combinatorially-defined dynamical systems on sets of combinatorial objects, where emphasis is placed on understanding the transient (i.e., non-periodic) points. Given an arbitrary set X and a map $f : X \rightarrow X$, let f^t

denote the t^{th} iterate of f . The *forward orbit* of an element $x \in X$ under the map f is the set $O_f(x) = \{x, f(x), f^2(x), \dots\}$. It is natural to consider the quantity $\sup_{x \in X} |O_f(x)|$. Indeed, if f is invertible and all orbits are finite, this is equivalent to determining the largest size of a periodic orbit. A similar question asks for the maximum possible number of iterations needed to send every transient point to a periodic point. If f has a fixed point x_0 such that every forward orbit under f contains x_0 (as will be the case for all of the dynamical systems considered in this paper), then these two questions are essentially equivalent.

The symmetric group S_n , which is the group of permutations of the set $[n] = \{1, \dots, n\}$, provides a rich source of noninvertible combinatorial dynamical systems. We can write a permutation $w \in S_n$ in one-line notation as $w(1) \cdots w(n)$. A great amount of research in combinatorics and computer science has focused on sorting operators, which are dynamical systems on S_n that have the identity permutation $e = 123 \cdots n$ as their unique periodic (necessarily fixed) point. Some typical examples of such operators include the bubble sort map (see [17, 2] and [37, pages 106–110]), West’s stack-sorting map (see [10, 21, 22, 26, 53] and the references therein), the map *revstack* defined in [28], the pop-stack-sorting map [4, 5, 6, 29, 19, 20, 45, 52], and the Queuesort map [18, 40].

A natural condition one might wish to place on a sorting operator $f : S_n \rightarrow S_n$, which all five of the specific sorting operators listed in the previous paragraph satisfy, is that it does not remove non-inversions. This means that if $a, b \in [n]$ are such that $a < b$ and a appears to the left of b in w , then a must appear to the left of b in $f(w)$. Among the sorting operators f that do not remove non-inversions, we will be primarily interested in the ones that remove consecutive inversions, meaning that if $a, b \in [n]$ are such that $a < b$ and a appears immediately to the right of b in w , then a appears to the left of b in $f(w)$. West’s stack-sorting map, the map *revstack*, and the pop-stack-sorting map all remove consecutive inversions; the bubble sort map and Queuesort, however, do not.

We are actually interested in generalizing these notions to arbitrary Coxeter groups. The condition that $f : S_n \rightarrow S_n$ does not remove non-inversions is equivalent to the condition that $f(w) \leq_R w$ for all $w \in S_n$, where \leq_R is the right weak order on S_n . Saying that f does not remove non-inversions *and* removes consecutive inversions is equivalent to saying that $f(w) \leq_R w$ and $f(w) \leq_R ws$ for every $w \in W$ and every right descent s of w . In what follows, we let \leq_R denote the right weak order on an arbitrary Coxeter group W , and we let $D_R(w)$ denote the right descent set of an element $w \in W$ (see Section 2.1 for definitions).

Definition 1.1. Let W be a Coxeter group. We say a map $f : W \rightarrow W$ is *compulsive*¹ if $f(w) \leq_R w$ and $f(w) \leq_R ws$ for every $w \in W$ and every $s \in D_R(w)$.

Note that the condition $f(w) \leq_R w$ is necessary in Definition 1.1 in order to guarantee that f fixes the identity element e .

¹Google gives the following two definitions for the word *compulsive*: (1) resulting from or relating to an irresistible urge, especially one that is against one’s conscious wishes; (2) irresistibly interesting or exciting; compelling. Our motivation for using this word comes from the first definition since a compulsive map on S_n compulsively removes all consecutive inversions. However, we hope to convince the reader that the second definition is also appropriate.

A seminal result due to Björner [8] states that the right weak order on a Coxeter group W is a complete meet-semilattice. This means that every set $A \subseteq W$ has a greatest lower bound, called the *meet* of A , which we denote by $\bigwedge_R A$. Hence, a map $f : W \rightarrow W$ is compulsive if and only if $f(w) \leqslant_R \bigwedge_R (\{ws : s \in D_R(w)\} \cup \{w\})$ for every $w \in W$. This motivates the following definition.

Definition 1.2. Let W be a Coxeter group. The *Coxeter pop-stack-sorting operator* on W is the map $\text{Pop}_W : W \rightarrow W$ defined by

$$\text{Pop}_W(w) = \bigwedge_R (\{ws : s \in D_R(w)\} \cup \{w\})$$

for every $w \in W$.

We often write Pop instead of Pop_W if the group W is clear from context.

The Coxeter pop-stack-sorting operator is certainly compulsive; its name comes from the fact, which we will verify in Section 2.2, that the Coxeter pop-stack-sorting operator on S_n is precisely the pop-stack-sorting map. This map, which is a deterministic analogue of a pop-stack-sorting machine introduced by Avis and Newborn in [7], first appeared in a different guise in a paper of Ungar’s about discrete geometry [52]; its popularity has grown rapidly over the past few years [4, 5, 6, 29, 45, 19, 20]. Ungar [52], motivated by a question involving directions determined by points in the plane, proved that the maximum possible size of a forward orbit of a permutation in S_n under the pop-stack-sorting map is n ; this settled a conjecture due to Goodman and Pollack [33]. In other words, Ungar’s theorem states that $\text{Pop}^{n-1}(w) = e$ for every $w \in S_n$ and that there exists $v \in S_n$ such that $\text{Pop}^{n-2}(v) \neq e$. The proof requires an unexpected amount of insight. Recently, Albert and Vatter [4] provided an alternative proof of Ungar’s theorem.

Our first main result generalizes Ungar’s theorem to an arbitrary irreducible² Coxeter group. Let (W, S) be a Coxeter system. If W is finite, then a *Coxeter element* of W is an element obtained by multiplying the simple generators (the elements of S) in an arbitrary order. All Coxeter elements have the same order in the group W ; this order is called the *Coxeter number* of W and is typically denoted by h . For example, the Coxeter number of S_n is n . We make the convention that the Coxeter number of an infinite Coxeter group is ∞ .

Theorem 1.3. *If W is an irreducible Coxeter group with Coxeter number h , then*

$$\sup_{w \in W} |O_{\text{Pop}}(w)| = h.$$

The preceding theorem tells us that if W is finite, then $\text{Pop}^{h-1}(w) = e$ for all $w \in W$ and $\text{Pop}^{h-2}(v) \neq e$ for some $v \in W$. On the other hand, if W is infinite, this theorem says that there are arbitrarily large forward orbits of elements of W under Pop . Our proof of Theorem 1.3 is Coxeter-theoretic and is type-uniform for all finite Coxeter groups with even Coxeter numbers. However, we must treat symmetric groups and dihedral groups separately from the other finite irreducible Coxeter groups because they can have odd Coxeter numbers.

²The assumption of irreducibility does not limit the scope of the theorem. If $W = W_1 \times W_2$ is reducible, then Pop acts on W_1 and W_2 independently, and one can understand the dynamics of Pop on W by “piecing together” information about the dynamics of Pop on W_1 and W_2 .

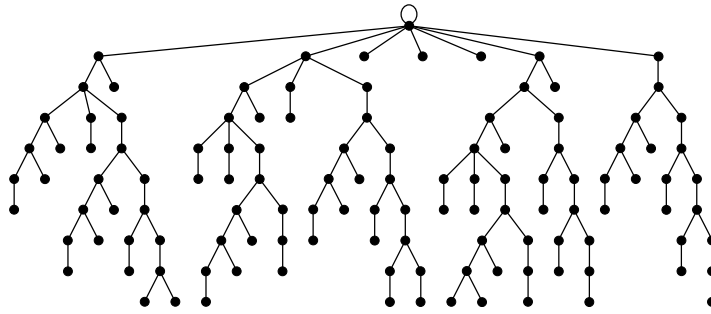


Figure 1.1: A diagram of the Coxeter pop-stack-sorting operator $\text{Pop} : H_3 \rightarrow H_3$. Each vertex in the tree represents an element of H_3 (we omit the labels). The root vertex is the identity element e ; the parent of each non-root vertex w is $\text{Pop}(w)$. As predicted by Theorem 1.3, the maximum size of a forward orbit is 10, which is the Coxeter number of H_3 .

It turns out that Pop is very closely related to *Brieskorn normal form*, which was introduced in the foundational works of Brieskorn and Brieskorn–Saito on complex reflection groups and Artin groups [14, 15]. This normal form is a specific factorization of an element of W into longest elements of parabolic subgroups. It follows from equation (2.1) below that the elements of $O_{\text{Pop}}(w)$ are exactly the prefixes of the Brieskorn normal form of w .³ More precisely, the Brieskorn normal form of w is $\beta_1 \cdots \beta_r$, where $r = |O_{\text{Pop}}(w)| - 1$ and $\beta_1 \cdots \beta_i = \text{Pop}^{r-i}(w)$ for all $0 \leq i \leq r$. Hence, the following corollary is an immediate consequence of Theorem 1.3.

Corollary 1.4. *If W is a finite irreducible Coxeter group with Coxeter number h , then the maximum number of terms that can appear in the Brieskorn normal form of an element of W is $h - 1$.*

Our second main result applies to all compulsive maps on an irreducible Coxeter group. This theorem is new even for symmetric groups.

Theorem 1.5. *Let W be an irreducible Coxeter group with Coxeter number h . If $f : W \rightarrow W$ is compulsive, then*

$$\sup_{w \in W} |O_f(w)| \leq h.$$

The conclusion of Theorem 1.5 is trivial if W is infinite, so we will only need to consider the case when W is finite. Our proof of this result is delicate and makes use of the Coxeter pop-stack-sorting operator, so we will really need to deliver it simultaneously with the proof of Theorem 1.3.

Remark 1.6. It is perhaps tempting to think that Theorem 1.5 would follow as an immediate corollary of Theorem 1.3 since a compulsive map $f : W \rightarrow W$ satisfies $f(w) \leq_R \text{Pop}(w)$ for all $w \in W$ by definition. More precisely, one might think that $f^t(w) \leq_R \text{Pop}^t(w)$ for all $w \in W$ and $t \geq 0$ so that, by Theorem 1.3, we have (assuming W is finite) $f^{h-1}(w) \leq_R \text{Pop}^{h-1}(w) = e$ for all $w \in W$. However, this is not the case. If $W = S_5$ and \mathfrak{s} is West’s stack-sorting map (see [10, 21] for the definition), then $\mathfrak{s}^3(42351) = 21345$, but $\text{Pop}^3(42351) = 12345 = e$.

³The articles [14, 15] use left weak order where we use right weak order, but this makes little difference.

Another typical approach to noninvertible combinatorial dynamical systems, especially sorting operators, concerns the enumeration of elements that require at most some fixed number t of iterations to reach a periodic point. In the case of West’s stack-sorting map, these elements are the t -stack-sortable permutations, which have been studied extensively, especially for $t \in \{1, 2, 3\}$ (see [3, 13, 10, 11, 21, 26, 27] and the references therein). For the pop-stack-sorting map, these elements are the t -pop-stack-sortable permutations investigated in [45, 19, 29]. Given a Coxeter group W , let us say an element $w \in W$ is t -pop-stack-sortable if $\text{Pop}_W^t(w) = e$. In other words, the set of t -pop-stack-sortable elements of W is $\text{Pop}_W^{-t}(e)$. At the end of Section 2.1, we will establish the following simple proposition. The *nerve* of Coxeter system is defined in that section.

Proposition 1.7. *Let (W, S) be a Coxeter system with nerve $\mathcal{N}(W, S)$. There is a bijection between the set $\text{Pop}_W^{-1}(e)$ of 1-pop-stack-sortable elements of W and $\mathcal{N}(W, S)$. In particular, if W is finite, then the number of 1-pop-stack-sortable elements of W is $2^{|S|}$.*

Pudwell and Smith [45] enumerated 2-pop-stack-sortable permutations in S_n . Our next theorem provides an analogue of this result for the hyperoctahedral groups B_n (defined in Section 4). Descending runs are defined in Section 2.2.

Theorem 1.8. *For each $n \geq 1$ and $k \geq 0$, the number of 2-pop-stack-sortable elements of B_n with $2k$ or $2k + 1$ descending runs is equal to the number of 2-pop-stack-sortable permutations in S_{n+1} with exactly $k + 1$ descending runs. In particular,*

$$\sum_{n \geq 1} |\text{Pop}_{B_n}^{-2}(e)| z^n = \sum_{n \geq 1} |\text{Pop}_{S_{n+1}}^{-2}(e)| z^n = \frac{2z(1 + z + z^2)}{1 - 2z - z^2 - 2z^3}.$$

Claesson and Guðmundsson proved that for every fixed $t \geq 0$, the generating function that counts t -pop-stack-sortable permutations in symmetric groups is rational [19]. Our final main theorems extend this result to the hyperoctahedral groups B_n and the affine symmetric groups \tilde{S}_n (see Sections 4 and 5 for the definitions).

Theorem 1.9. *For every $t \geq 0$, the generating function $\sum_{n \geq 1} |\text{Pop}_{B_n}^{-t}(e)| z^n$ is rational.*

Theorem 1.10. *For every $t \geq 0$, the generating function $\sum_{n \geq 1} |\text{Pop}_{\tilde{S}_n}^{-t}(e)| z^n$ is rational.*

1.2. Generalizations

There are (at least) two natural ways in which one could generalize the definition of the Coxeter pop-stack-sorting operators.

First, one could replace the right weak order on a Coxeter group with an arbitrary complete meet-semilattice M to obtain a map $\text{Pop}_M : M \rightarrow M$ defined by

$$\text{Pop}_M(x) = \bigwedge (\{y \in M : y \triangleleft x\} \cup \{x\})$$

for all $x \in M$. This provides a large new class of noninvertible combinatorial dynamical systems that are ripe for investigation.⁴ We will explore this avenue in [23], with special emphasis on ν -Tamari lattices.

Remark 1.11. In light of Definition 1.1, it is natural to define a map $f : M \rightarrow M$ on an arbitrary complete meet-semilattice M to be *compulsive* if $f(x) \leq \text{Pop}_M(x)$ for all $x \in M$. Together, Theorems 1.3 and 1.5 tell us that if M is isomorphic to the right weak order on an irreducible Coxeter group and $f : M \rightarrow M$ is compulsive, then $\sup_{x \in M} |O_f(x)| \leq \sup_{x \in M} |O_{\text{Pop}_M}(x)|$. This does *not* hold if M is replaced by an arbitrary complete meet-semilattice. For example, consider the lattice M whose Hasse diagram is shown in Figure 1.2. If $f : M \rightarrow M$ is the compulsive map illustrated with green arrows, then $\sup_{x \in M} |O_f(x)| = 4 > 3 = \sup_{x \in M} |O_{\text{Pop}_M}(x)|$.

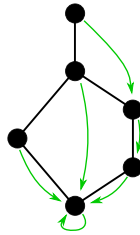


Figure 1.2: A compulsive map on a lattice.

For the second generalization, we make use of the left weak order \leq_L on a Coxeter group W . The poset (W, \leq_L) is a complete meet-semilattice that is isomorphic to (W, \leq_R) . Let \wedge_L denote the meet operation in the left weak order. A *semilattice congruence* on (W, \leq_L) is an equivalence relation \equiv on W that respects meets, meaning that $x_1 \equiv x_2$ and $y_1 \equiv y_2$ together imply $(x_1 \wedge_L y_1) \equiv (x_2 \wedge_L y_2)$. When W is finite, we define a *lattice congruence* on (W, \leq_L) to be an equivalence relation on W that respects both meets and joins, meaning $x_1 \equiv x_2$ and $y_1 \equiv y_2$ together imply $(x_1 \wedge_L y_1) \equiv (x_2 \wedge_L y_2)$ and $(x_1 \vee_L y_1) \equiv (x_2 \vee_L y_2)$. Semilattice congruences and lattice congruences on weak orders of Coxeter groups have been studied extensively (see [48, 47] and the references therein).

One of the most natural examples of a semilattice congruence on (W, \leq_L) is the *descent congruence*, which is defined by declaring two elements of W to be equivalent if they have the same right descent set. Another notable example is the *sylvester congruence* on the symmetric group S_n . There is a well-known bijection \mathcal{I} from the set of decreasing binary plane trees with label set $[n]$ to S_n ; the sylvester congruence is defined by declaring two permutations σ and σ' to be equivalent if the unlabeled binary plane trees obtained by removing the labels from $\mathcal{I}^{-1}(\sigma)$ and $\mathcal{I}^{-1}(\sigma')$ are equal (see [24, 34, 44] for more details). Further interesting examples of semilattice congruences on Coxeter groups are provided by the Cambrian congruences (see [46, 47, 49, 50] and the references therein). In the case of S_n , other notable examples

⁴Henri Mühle has informed the author that $\text{Pop}_M(x)$ is the same as what he has called the *nucleus* of x [41, 42]. However, his results do not overlap with ours because he did not view nuclei as defining a dynamical system. Instead, he was interested in the lattice-theoretic properties of the interval between $\text{Pop}_M(x)$ and x in M .

of lattice congruences include the permutree congruences [44], the k -twist congruences [43], and the Baxter congruence [38, 32].

If \equiv is a semilattice congruence on the left weak order of W , then every congruence class of \equiv has a unique minimal element. We denote by $\pi_{\downarrow} : W \rightarrow W$ the projection map that sends each element of W to the unique minimal element of its congruence class. We say \equiv is *essential* if the identity element $e \in W$ is in a singleton equivalence class.

Definition 1.12. Let W be a Coxeter group, and let \equiv be an essential semilattice congruence on the left weak order of W . Define the *Coxeter stack-sorting operator* $S_{\equiv} : W \rightarrow W$ to be the map given by $S_{\equiv}(w) = w(\pi_{\downarrow}(w))^{-1}$ for all $w \in W$.

The name *Coxeter stack-sorting operator* is motivated by two special cases. First, if \equiv is the descent congruence, then S_{\equiv} is the same as the Coxeter pop-stack-sorting operator Pop_W (we justify this claim at the end of Section 2.1). Second, if $W = S_n$ and \equiv is the sylvester congruence, then S_{\equiv} is West's stack-sorting map (this is essentially the content of [24, Corollary 16], where the projection map π_{\downarrow} goes by the name swd). Coxeter stack-sorting operators provide another broad new class of non-invertible combinatorial dynamical systems; we will investigate them more extensively in [25], with special emphasis on permutree congruences in type A and their analogues (in particular, analogues of the sylvester congruence) in types B and \tilde{A} .

Remark 1.13. A semilattice congruence on the left weak order of a Coxeter group is essential if and only if it refines the descent congruence (see [35, 25]). From this, it is not difficult to show that all Coxeter stack-sorting operators are compulsive, so they provide a wide range of maps to which one can apply Theorem 1.5 (see [25]).

Remark 1.14. A different generalization of West's stack-sorting map, introduced in [16], uses *pattern-avoiding stacks*. While these pattern-avoiding stacks are certainly interesting, we believe our Coxeter stack-sorting operators are more natural from an algebraic and lattice-theoretic point of view.

1.3. Outline

In Section 2, we recall relevant definitions and facts concerning Coxeter groups, prove Proposition 1.7, and verify that the Coxeter pop-stack-sorting operator on S_n coincides with the pop-stack-sorting map. Section 3 is devoted to proving Theorems 1.3 and 1.5. We prove Theorems 1.8 and 1.9, which concern Coxeter groups of type B , in Section 1.9. We prove Theorem 1.10, which concerns Coxeter groups of type \tilde{A} , in Section 5. Section 6 provides additional potential ideas for future work.

2. Preliminaries

2.1. Coxeter groups

We assume familiarity with basic notions and concepts from the combinatorial theory of Coxeter groups and from lattice theory; a standard reference that contains all the background information we need is [9].

A *Coxeter system* is a pair (W, S) , where W is a group generated by the set S with presentation $W = \langle S : (ss')^{m(s,s')} = e \rangle$ such that $m(s, s) = 1$ for all $s \in S$ and $m(s, s') = m(s', s) \in \{2, 3, \dots\} \cup \{\infty\}$ for all distinct $s, s' \in S$. The elements of S are called the *simple generators*. We will often refer to a Coxeter group W with the understanding that we are really referring to a Coxeter system (W, S) for some specific choice of a generating set S . The *Coxeter diagram* of W is the graph $\Gamma(W)$ with vertex set S in which two vertices s, s' are connected by an edge labeled with $m(s, s')$ if $m(s, s') \geq 3$ (and are not adjacent if $m(s, s') \leq 2$). Notice that s and s' commute if and only if they are not adjacent in $\Gamma(W)$. We say W is *irreducible* if $\Gamma(W)$ is a connected graph.

A *reduced word* for an element $w \in W$ is a word $s_1 \cdots s_k$ over the alphabet S that, when viewed as a product of elements of W , equals w . The smallest length of a reduced word for w is called the *length* of w and is denoted by $\ell(w)$. The *left weak order* on W is the partial order \leq_L on W defined by declaring that $x \leq_L y$ if $\ell(yx^{-1}) = \ell(y) - \ell(x)$. The *right weak order* on W is the partial order \leq_R on W defined by declaring that $x \leq_R y$ if $\ell(x^{-1}y) = \ell(y) - \ell(x)$. The map $W \rightarrow W$ given by $w \mapsto w^{-1}$ is an isomorphism from the left weak order to the right weak order. We will also need the *strong Bruhat order* on W , which is the partial order \leq on W defined by saying that $x \leq y$ if some (equivalently, every) reduced word for y contains a reduced word for x as a (not necessarily contiguous) subword. If $x \leq_R y$ or $x \leq_L y$, then $x \leq y$.

The poset (W, \leq_R) is a complete meet-semilattice [8, 9], meaning that every set $A \subseteq W$ has a unique meet $\bigwedge_R A$. If A has an upper bound in (W, \leq_R) , then it has a join, which we denote by $\bigvee_R A$. Similarly, we write $\bigwedge_L A$ and $\bigvee_L A$ for, respectively, the meet of A and the join of A (if it exists) in the left weak order. The right and left weak orders on W are lattices if W is finite.

A *right descent* of an element $w \in W$ is a simple generator $s \in S$ such that $\ell(ws) < \ell(w)$; the collection of all right descents of w is the *right descent set* of w , which we denote by $D_R(w)$. Similarly, the *left descent set* of w is the set $D_L(w) = \{s \in S : \ell(sw) < \ell(w)\}$ of *left descents* of w . It is a basic fact that $D_R(w) = \{s \in S : s \leq_L w\}$ and $D_L(w) = \{s \in S : s \leq_R w\}$.

Suppose $J \subseteq S$. The *parabolic subgroup* W_J is the subgroup of W generated by the elements of J . For every $w \in W$, the right coset $W_J w$ has a unique representative ${}^J w$ of minimal length. Let $w_J = w({}^J w)^{-1}$ so that $w_J \in W_J$. The factorization $w = w_J {}^J w$ is length-additive in the sense that $\ell(w) = \ell(w_J) + \ell({}^J w)$. The set ${}^J W = \{{}^J w : w \in W\}$ is called a *parabolic quotient*. The next lemma states that for any $J \subseteq S$, the map $w \mapsto {}^J w$ is order-preserving with respect to the right weak order.

Lemma 2.1. *Let (W, S) be a Coxeter system, and let $J \subseteq S$. If $y, z \in W$ are such that $y \leq_R z$, then ${}^J y \leq_R {}^J z$.*

Proof. It suffices to prove this in the case when $z = ys$ for some $s \in S \setminus D_R(y)$. Deodhar's Lemma (see, for example, [31, Lemma 2.1.2]) tells us that either ${}^J ys \in {}^J W$ or ${}^J ys = s' {}^J y$ for some $s' \in J$. Assume first that ${}^J ys \in {}^J W$. Then $z = ys = y_J ({}^J ys)$, so $z_J = y_J$ and ${}^J z = {}^J ys$. Since $s \notin D_R(y)$ and ${}^J y \leq_L y$, we must have $s \notin D_R({}^J y)$. Hence, ${}^J y \leq_R {}^J ys = {}^J z$. Now suppose ${}^J ys \notin {}^J W$. Then $W_J {}^J ys = W_J s' {}^J y = W_J {}^J y$, so $z = y_J ({}^J ys) \in W_J {}^J ys = W_J {}^J y$. It follows that ${}^J z = {}^J y$. \square

If W is finite, we write w_0 for the longest element of W . The *nerve* of (W, S) , denoted $\mathcal{N}(W, S)$, is the collection of subsets $J \subseteq S$ such that W_J is finite. If $J \in \mathcal{N}(W, S)$, we

write $w_0(J)$ for the longest element of W_J . The elements w_0 and $w_0(J)$ are involutions. It follows from [9, Lemma 3.2.4] that

$$\text{Pop}(w) = w w_0(D_R(w)) \tag{2.1}$$

for every $w \in W$. According to [9, Lemma 3.2.3], W_J is finite if and only if $\bigvee_L J$ exists; moreover, if W_J is finite, then $\bigvee_L J = w_0(J)$. Since each $w \in W$ is an upper bound for $D_R(w)$ in the left weak order, the join $\bigvee_L D_R(w)$ must exist and equal $w_0(D_R(w))$.

Proof of Proposition 1.7. If $J \in \mathcal{N}(W, S)$, then $D_R(w_0(J)) = J$ since $w_0(J) = \bigvee_L J$. For every $J \in \mathcal{N}(W, S)$, it follows from (2.1) that $\text{Pop}(w_0(J)) = (w_0(J))^2 = e$, so $w_0(J)$ is 1-pop-stack-sortable. On the other hand, if $v \in W$ is 1-pop-stack-sortable, then (2.1) implies that $v = w_0(D_R(v))$, and we know that $D_R(v) \in \mathcal{N}(W, S)$ because v is an upper bound for $D_R(v)$ in the left weak order. Hence, we have a surjective map $\mathcal{N}(W, S) \rightarrow \text{Pop}_W^{-1}(e)$ given by $J \mapsto w_0(J)$. The identity $D_R(w_0(J)) = J$ implies that this map is actually a bijection. \square

Fix $w \in W$. If $v \in W$ is such that $D_R(v) = D_R(w)$, then it follows from the above discussion that $w_0(D_R(w)) = w_0(D_R(v)) \leq_L v$. The right descent set of $w_0(D_R(w))$ is $D_R(w)$, so $w_0(D_R(w))$ is the smallest element in the left weak order that has the same right descent set as w . Since the element $w_0(D_R(w))$ is an involution, this explains why, as mentioned in Section 1.2, the Coxeter stack-sorting operator corresponding to the descent congruence on (W, \leq_L) agrees with the Coxeter pop-stack-sorting operator Pop_W .

2.2. The pop-stack-sorting map

The prototypical example of a Coxeter group is the symmetric group S_n . The set of simple generators for S_n is $S = \{s_1, \dots, s_{n-1}\}$, where s_i is the transposition that swaps i and $i + 1$. We often write a permutation $w \in S_n$ as a word $w(1) \cdots w(n)$ in one-line notation. A simple transposition s_i is a right descent of w if and only if $w(i) > w(i + 1)$. An *inversion* of w is a pair (i, j) such that $1 \leq i < j \leq n$ and $w(i) > w(j)$. It is well known that the number of inversions of w is $\ell(w)$. The longest element of S_n , which we denote by $w_0(S_n)$ when we wish to stress the dependence on n , is the decreasing permutation $n(n - 1) \cdots 321$.

The *direct sum* of permutations $u \in S_m$ and $v \in S_n$ is the permutation $u \oplus v \in S_{m+n}$ defined by

$$(u \oplus v)(i) = \begin{cases} u(i) & \text{if } 1 \leq i \leq m; \\ v(i - m) + m & \text{if } m + 1 \leq i \leq m + n. \end{cases}$$

A permutation is called *layered* if it can be written as $w_0(S_{n_1}) \oplus \cdots \oplus w_0(S_{n_k})$, a direct sum of decreasing permutations. Note that layered permutations are involutions. A permutation w is layered if and only if $w = w_0(D_R(w))$.

A *descending run* of a permutation $w \in S_n$ is a maximal consecutive decreasing subsequence of w . For instance, the descending runs of 42135867 are 421, 3, 5, 86, and 7. The *pop-stack-sorting map* is the operator on S_n that acts by reversing the descending runs of a permutation while keeping entries in different descending runs in the same relative order. For example, the pop-stack-sorting map sends 42135867 to 12435687. Equivalently, the pop-stack-sorting map

acts by multiplying a permutation w on the right by the unique layered permutation that has the same descent set as w . In other words, it sends w to $ww_0(D_R(w))$. In the above example, we have $w_0(D_R(42135867)) = w_0(\{s_1, s_2, s_6\}) = 32145768$, and we saw that the pop-stack-sorting map sends 42135867 to the permutation $12435687 = 42135867 \cdot 32145768$. Hence, (2.1) tells us that the Coxeter pop-stack-sorting operator $\text{Pop} : S_n \rightarrow S_n$ agrees with the pop-stack-sorting map.

3. Maximum forward orbit sizes

The goal of this section is to prove Theorems 1.3 and 1.5. Our first step will be to establish Theorem 1.3 for finite dihedral groups.

The dihedral group $I_2(m)$ has presentation $\langle s_1, s_2 : s_1^2 = s_2^2 = (s_1 s_2)^m = e \rangle$. For $1 \leq k \leq m$, let α_k denote the element $\cdots s_1 s_2 s_1$ of $I_2(m)$ obtained by multiplying the sequence of k simple generators that alternates between s_1 and s_2 and ends in s_1 . Similarly, let $\beta_k = \cdots s_2 s_1 s_2$ be the product of k simple generators that alternate between s_1 and s_2 and end in s_2 . Then $I_2(m) = \{e\} \cup \{\alpha_j : 1 \leq j \leq m\} \cup \{\beta_k : 1 \leq k \leq m-1\}$. We have $D_R(e) = \emptyset$ and $D_R(\alpha_m) = \{s_1, s_2\}$. For $1 \leq k \leq m-1$, we have $D_R(\alpha_k) = \{s_1\}$ and $D_R(\beta_k) = \{s_2\}$. Thus, by (2.1), the map $\text{Pop} : I_2(m) \rightarrow I_2(m)$ is given explicitly by $\text{Pop}(e) = \text{Pop}(\alpha_m) = \text{Pop}(s_1) = \text{Pop}(s_2) = e$, $\text{Pop}(\alpha_j) = \beta_{j-1}$ for all $2 \leq j \leq m-1$, and $\text{Pop}(\beta_k) = \alpha_{k-1}$ for all $2 \leq k \leq m-1$. Therefore, the elements of $I_2(m)$ whose forward orbits under Pop are of maximum size are α_{m-1} and β_{m-1} . Indeed, $|O_{\text{Pop}}(\alpha_{m-1})| = |O_{\text{Pop}}(\beta_{m-1})| = m$. The Coxeter number of $I_2(m)$ is m , so this proves Theorem 1.3 for finite dihedral groups.

The Coxeter numbers of finite irreducible Coxeter groups are all known [36, Section 3.18]; in particular, the only finite irreducible Coxeter groups whose Coxeter numbers are odd are the symmetric groups S_n with n odd and the dihedral groups $I_2(m)$ with m odd. Ungar already proved Theorem 1.3 for symmetric groups, and we just established this theorem for dihedral groups. Thus, we have the following lemma, which shows that we can focus much of our attention hereafter on finite irreducible Coxeter groups with even Coxeter numbers.

Lemma 3.1. *If W is a finite irreducible Coxeter group whose Coxeter number h is odd, then*

$$\max_{w \in W} |O_{\text{Pop}}(w)| = h.$$

The Coxeter diagrams of the finite irreducible Coxeter groups W have been classified (see [9, Appendix A1]); in particular, every such diagram is a tree. This means that it is possible to choose a bipartition $X \sqcup Y$ of the vertex set S of $\Gamma(W)$ (i.e., every edge in $\Gamma(W)$ has one endpoint in X and one endpoint in Y). The elements of X all commute with each other, and the elements of Y all commute with each other. The map $W \rightarrow W$ given by $w \mapsto w_0 w w_0$ is an automorphism of W such that $w_0 S w_0 = S$ [9, Section 2.3]. In particular, this map induces an automorphism of $\Gamma(W)$.

Lemma 3.2. *Let W be a finite irreducible Coxeter group with an even Coxeter number. If $X \sqcup Y$ is a bipartition of the vertex set $\Gamma(W)$, then $w_0 X w_0 = X$.*

Proof. Referring to the classification of Coxeter diagrams of finite irreducible Coxeter groups, we immediately find that every automorphism of $\Gamma(W)$ must fix the set X unless W is isomorphic to a dihedral group, a symmetric group S_n with n odd, or the exceptional group F_4 . Because W has an even Coxeter number, it is not isomorphic to a symmetric group S_n with n odd or a dihedral group $I_2(m)$ with m odd. It is known (one can even check by hand) that if W is isomorphic to F_4 or to $I_2(m)$ with m even, then the automorphism given by $w \mapsto w_0ww_0$ is trivial. \square

Lemma 3.3. *Let W be a finite Coxeter group. Choose $s \in S$, and let $J = S \setminus \{s\}$. We have $D_R({}^Jw_0) = \{w_0sw_0\}$.*

Proof. Since the map $w \mapsto w_0ww_0$ is an automorphism of W that fixes the set S , we have $\ell(w_0ww_0) = \ell(w)$ for all $w \in W$. The element $(w_0)_J = w_0({}^Jw_0)^{-1}$ is equal to the involution $w_0(J)$ [9, Section 2.5], so $w_0({}^Jw_0)^{-1} = ({}^Jw_0)w_0$. Therefore, for every $r \in S$, we have

$$\ell({}^Jw_0 \cdot w_0rw_0) = \ell(w_0({}^Jw_0)^{-1}rw_0) = \ell(w_0(r \cdot {}^Jw_0)^{-1}w_0) = \ell((r \cdot {}^Jw_0)^{-1}) = \ell(r \cdot {}^Jw_0).$$

It follows that $D_R({}^Jw_0) = w_0D_L({}^Jw_0)w_0$, so we are left to prove that $D_L({}^Jw_0) = \{s\}$. An alternative characterization of the parabolic quotient ${}^JW = \{{}^Jw : w \in W\}$ is that it is the set of elements of W whose left descent sets are contained in the set $S \setminus J = \{s\}$ (see [9, Section 2.4]). Since ${}^Jw_0 \in {}^JW$ is not the identity element e , its left descent set must be $\{s\}$. \square

The next proposition is the key to proving both Theorems 1.3 and 1.5.

Proposition 3.4. *Let W be a finite irreducible Coxeter group whose Coxeter number h is even. Choose $s \in S$, and let $J = S \setminus \{s\}$. We have $\text{Pop}^{h-1}({}^Jw_0) = e$ and $\text{Pop}^{h-2}({}^Jw_0) \neq e$. Furthermore, for every $t \geq 0$, the right descents of $\text{Pop}^t({}^Jw_0)$ all commute with each other.*

Proof. Let us choose a bipartition $X \sqcup Y$ of the vertex set S of $\Gamma(W)$. Without loss of generality, assume $s \in X$. All of the elements of X commute with each other, so it makes sense to define the product $c_X = \prod_{r \in X} r$. Similarly, we can define $c_Y = \prod_{r \in Y} r$. The element $c = c_Xc_Y$ is a Coxeter element of W . Let us fix a reduced word $r_1 \cdots r_{|S|}$ of c_Xc_Y such that $r_1 \cdots r_{|X|}$ is a reduced word for c_X and $r_{|X|+1} \cdots r_{|S|}$ is a reduced word for c_Y . Without loss of generality, we may assume $r_1 = s$. According to [12, Chapter V, Section 6.2, Proposition 2], we have $w_0 = (c_Xc_Y)^{h/2}$. It is well known [36, Section 3.18] that $\ell(w_0) = |S| \cdot h/2$, so concatenating $r_1 \cdots r_{|S|}$ with itself $h/2$ times produces a reduced word $s_1 \cdots s_{\ell(w_0)}$ for w_0 . Note that $s_1 = r_1 = s$. Let us write the word $s_1 \cdots s_{\ell(w_0)}$ as the concatenation $u_1 \cdots u_h$, where $u_i = r_1 \cdots r_{|X|}$ when i is odd and $u_i = r_{|X|+1} \cdots r_{|S|}$ when i is even. We think of the letters $s_1, \dots, s_{\ell(w_0)}$ as distinct entities, and we think of each s_j as belonging to exactly one of the words u_i . For example, s_1 belongs to u_1 , while $s_{\ell(w_0)}$ belongs to u_h .

Let $k = \ell(w_0) - \ell({}^Jw_0)$. Given indices $1 \leq p_1 < \cdots < p_k \leq \ell(w_0)$ and $j \in [k]$, we can consider the word $s_1 \cdots \widehat{s}_{p_1} \cdots \widehat{s}_{p_j} \cdots s_{\ell(w_0)}$ obtained by removing the letters in positions p_1, \dots, p_j from the word $s_1 \cdots s_{\ell(w_0)}$. Let $z^{(j)} = s_1 \cdots \widehat{s}_{p_1} \cdots \widehat{s}_{p_j} \cdots s_{\ell(w_0)}$ be the element of W represented by this word. The proof of [9, Theorem 2.5.5] shows that it is possible to choose the indices $p_1 < \cdots < p_k$ so that $\ell(z^{(j)}) = \ell(w_0) - j$ for all $j \in [k]$ and so that $z^{(k)} = {}^Jw_0$. Let us assume that we have made such a choice. Let v_i be the word obtained from u_i by removing

the letters in the sequence s_{p_1}, \dots, s_{p_k} that belong to u_i . Thus, $v_1 \cdots v_h$ is the reduced word $s_1 \cdots \widehat{s}_{p_1} \cdots \widehat{s}_{p_k} \cdots s_{\ell(w_0)}$ for ${}^J w_0$.

We claim that none of the words v_1, \dots, v_{h-1} are empty. To prove this, we first assume, by way of contradiction, that v_1 is empty. This implies that $p_1 = 1$, so the reduced word $v_1 \cdots v_h$ for ${}^J w_0$ is contained in the reduced word $s_2 \cdots s_{\ell(w_0)}$ for sw_0 . Hence, ${}^J w_0 \leq sw_0$ in the strong Bruhat order. The map $W \rightarrow W$ given by $w \mapsto ww_0$ is an antiautomorphism of the strong Bruhat order [9, Proposition 2.3.4], so $s \leq ({}^J w_0)w_0 = ((w_0)_J)^{-1}$. Since $((w_0)_J)^{-1} \in W_J$, this forces $s \in J$, which contradicts the fact that $J = S \setminus \{s\}$. Consequently, v_1 is not the empty word.

Now suppose there is some $i \in \{2, \dots, h-1\}$ such that v_i is the empty word; we may assume that this i is chosen minimally. Let j be the index such that s_{p_j} belongs to u_i and is the last letter in u_i . Then $z^{(j-1)} = v_1 \cdots v_{i-1} s_{p_j} u_{i+1} \cdots u_h$ and $z^{(j)} = v_1 \cdots v_{i-1} u_{i+1} \cdots u_h$. By the minimality of i , the word v_{i-1} is nonempty. Let s' be the last letter of v_{i-1} , and write $v_{i-1} = \tilde{v}_{i-1} s'$. Because $i-1$ and $i+1$ have the same parity, s' is an element of S that appears in u_{i+1} . Furthermore, the elements of S appearing in u_{i+1} all commute with each other, so there is a reduced word \tilde{u}_{i+1} such that $u_{i+1} = s' \tilde{u}_{i+1}$ (as elements of W). This means that $z^{(j)} = v_1 \cdots \tilde{v}_{i-1} s' \tilde{u}_{i+1} \cdots u_h = v_1 \cdots \tilde{v}_{i-1} \tilde{u}_{i+1} \cdots u_h$, so $\ell(z^{(j)}) \leq \ell(z^{(j-1)}) - 3$. This contradicts the fact that $\ell(z^{(j)}) = \ell(w_0) - j = \ell(z^{(j-1)}) - 1$. Hence, we have proven the claim that v_1, \dots, v_{h-1} are nonempty.

Now recall that $v_1 \cdots v_h = {}^J w_0$. Every letter in v_h is also a letter in u_h , and the letters in u_h are the elements of Y because h is even. It follows that if v_h is nonempty, then its last letter is in $D_R({}^J w_0) \cap Y$. Lemma 3.3 tells us that $D_R({}^J w_0) = \{w_0 s w_0\}$. Since $s \in X$, it follows from Lemma 3.2 that $w_0 s w_0 \in X$. Hence, $D_R({}^J w_0) \cap Y = \emptyset$. This demonstrates that v_h is the empty word, so ${}^J w_0 = v_1 \cdots v_{h-1}$.

Choose some $m \in [h-1]$. Because the letters in v_m all commute with each other, they must all be right descents of $v_1 \cdots v_m$. We wish to show that the set of letters appearing in v_m is actually equal to $D_R(v_1 \cdots v_m)$. Let x_m be a word obtained by multiplying the letters that appear in u_m but not v_m in some order. Then $u_m = v_m x_m$ (as elements of W), so $v_1 \cdots v_m x_m u_{m+1} \cdots u_h$ is a reduced word for some $z^{(j)}$. Consequently, the word $v_1 \cdots v_m x_m$ is reduced. All of the letters in x_m commute with each other, so none of them can be right descents of $v_1 \cdots v_m$. There is an index j' such that $v_1 \cdots v_m u_{m+1} \cdots u_h$ is a reduced word for $z^{(j')}$. It follows that the word $v_1 \cdots v_m u_{m+1}$ is reduced. All of the letters in u_{m+1} commute with each other, so none of them can be right descents of $v_1 \cdots v_m$. Since every element of S appears in exactly one of the words v_m, x_m, u_{m+1} , this proves that $D_R(v_1 \cdots v_m)$ is the set of letters appearing in v_m . Because the letters in v_m all commute with each other, the element of W represented by v_m is $w_0(D_R(v_1 \cdots v_m))$. Appealing to (2.1), we find that

$$\begin{aligned} \text{Pop}(v_1 \cdots v_m) &= v_1 \cdots v_m w_0(D_R(v_1 \cdots v_m)) \\ &= v_1 \cdots v_{m-1} (w_0(D_R(v_1 \cdots v_m)))^2 \\ &= v_1 \cdots v_{m-1}. \end{aligned}$$

It now follows by induction on t that $\text{Pop}^t({}^J w_0) = v_1 \cdots v_{h-t-1}$ for every $0 \leq t \leq h-1$. In particular, $\text{Pop}^{h-1}({}^J w_0) = e$ and $\text{Pop}^{h-2}({}^J w_0) = v_1 \neq e$. Furthermore, the right descents of

$\text{Pop}^t({}^Jw_0)$ are the letters appearing in v_{h-t-1} , which all commute with each other. If $t \geq h$, then it is vacuously true that the right descents of $\text{Pop}^t({}^Jw_0) = e$ commute with each other. \square

Proposition 3.4 tells us already that if W is finite and irreducible with an even Coxeter number h , then $\sup_{w \in W} |O_{\text{Pop}}(w)| \geq h$. To prove the reverse inequality and Theorem 1.5, we need a couple more preparatory lemmas. Recall that \leq denotes the strong Bruhat order.

Lemma 3.5. *Let W be a finite Coxeter group. Let $x, y \in W$, and assume that all of the right descents of y commute with each other. If $x \leq y$, then $\text{Pop}(x) \leq \text{Pop}(y)$.*

Proof. Let s_1, \dots, s_r be the right descents of y . Because s_1, \dots, s_r all commute with each other, we have $w_0(D_R(y)) = s_1 \cdots s_r$. Let $s'_1 \cdots s'_q$ be a reduced word for $\text{Pop}(y)$. Since $\text{Pop}(y) \leq_R y$ by definition, it follows from (2.1) and the fact that $w_0(D_R(y))$ is an involution that $s'_1 \cdots s'_q s_1 \cdots s_r$ is a reduced word for y . Assume $x \leq y$. This implies that the reduced word $s'_1 \cdots s'_q s_1 \cdots s_r$ contains a reduced word $s'_{i_1} \cdots s'_{i_a} s_{j_1} \cdots s_{j_b}$ for x . Here, $s'_{i_1} \cdots s'_{i_a}$ is a reduced word for some element z with $z \leq \text{Pop}(y)$. Note that $\ell(z) = \ell(x) - b$. Since s_{j_1}, \dots, s_{j_b} all commute with each other, they are all in $D_R(x)$. Thus, $s_{j_1} \cdots s_{j_b}$ is in the parabolic subgroup $W_{D_R(x)}$. Now, $w_0(D_R(x))$ is the unique maximal element in the left weak order on $W_{D_R(x)}$, so $s_{j_1} \cdots s_{j_b} \leq_L w_0(D_R(x))$. This means that there exists $v \in W$ such that $w_0(D_R(x)) = vs_{j_1} \cdots s_{j_b}$ and $\ell(w_0(D_R(x))) = \ell(v) + b$. Using (2.1) and the fact that $w_0(D_R(x))$ is an involution, we deduce that

$$zs_{j_1} \cdots s_{j_b} = x = \text{Pop}(x)w_0(D_R(x)) = \text{Pop}(x)vs_{j_1} \cdots s_{j_b},$$

so $z = \text{Pop}(x)v$. Since $\text{Pop}(x) \leq_R x$, the factorization $x = \text{Pop}(x)w_0(D_R(x))$ is length-additive, meaning $\ell(x) = \ell(\text{Pop}(x)) + \ell(w_0(D_R(x)))$. Thus, $\ell(z) = \ell(x) - b = \ell(\text{Pop}(x)) + \ell(w_0(D_R(x))) - b = \ell(\text{Pop}(x)) + \ell(v)$. This shows that $\text{Pop}(x) \leq_R z$, so $\text{Pop}(x) \leq z \leq \text{Pop}(y)$. \square

Lemma 3.6. *Let (W, S) be a Coxeter system, and let $J \subseteq S$. If $f : W \rightarrow W$ is compulsive, then ${}^J(f(w)) \leq_R \text{Pop}({}^Jw)$ for every $w \in W$.*

Proof. By definition, $\text{Pop}({}^Jw) = \bigwedge_R \{{}^Jwx : x \in D_R({}^Jw) \cup \{e\}\}$. Therefore, it suffices to prove that ${}^J(f(w)) \leq_R {}^Jwx$ for all $x \in D_R({}^Jw) \cup \{e\}$. Fix $x \in D_R({}^Jw) \cup \{e\}$. According to [51, Proposition 2.5], the parabolic quotient JW is an order ideal in the right weak order on W . Since ${}^Jwx \leq_R {}^Jw$, it follows that ${}^Jwx \in {}^JW$. Because $wx = w_J \cdot {}^Jwx$, we must have ${}^J(wx) = {}^Jwx$. Now, $x \in D_R({}^Jw) \cup \{e\} \subseteq D_R(w) \cup \{e\}$, so it follows from the hypothesis that f is compulsive that $f(w) \leq_R wx$. Invoking Lemma 2.1, we see that ${}^J(f(w)) \leq_R {}^J(wx) = {}^Jwx$, as desired. \square

We can now complete the proof of Theorem 1.5.

Proof of Theorem 1.5. Let W be an irreducible Coxeter group with Coxeter number h , and let $f : W \rightarrow W$ be compulsive. If W is infinite, then $h = \infty$, so the result is trivial. Therefore, we may assume W is finite. Fix $w \in W$; our goal is to prove that $f^{h-1}(w) = e$. It suffices to prove that $f^{h-1}(w) \in W_J$ for every set $J \subseteq S$ such that $|J| = |S| - 1$. Fix such a set J . Note that the desired containment $f^{h-1}(w) \in W_J$ is equivalent to the identity ${}^J(f^{h-1}(w)) = e$.

Because $w \leq_R w_0$, we have ${}^J w \leq_R {}^J w_0$ by Lemma 2.1. Therefore, ${}^J w \leq {}^J w_0$. By induction on t , we will prove that ${}^J(f^t(w)) \leq \text{Pop}^t({}^J w_0)$ for all $t \geq 0$. We have just established the base case $t = 0$, so let us assume that $t \geq 1$ and that we have already proven the inequality ${}^J(f^{t-1}(w)) \leq \text{Pop}^{t-1}({}^J w_0)$. Proposition 3.4 tells us that the right descents of $\text{Pop}^{t-1}({}^J w_0)$ all commute with each other, so we can use Lemma 3.5 with $x = {}^J(f^{t-1}(w))$ and $y = \text{Pop}^{t-1}({}^J w_0)$ to see that $\text{Pop}({}^J(f^{t-1}(w))) \leq \text{Pop}^t({}^J w_0)$. Lemma 3.6 tells us that ${}^J(f^t(w)) \leq_R \text{Pop}({}^J(f^{t-1}(w)))$, so ${}^J(f^t(w)) \leq \text{Pop}({}^J(f^{t-1}(w))) \leq \text{Pop}^t({}^J w_0)$. This completes the induction step. Now set $t = h - 1$ to find that ${}^J(f^{h-1}(w)) \leq \text{Pop}^{h-1}({}^J w_0)$. It follows from Lemma 3.1 (if h is odd) and Proposition 3.4 (if h is even) that $\text{Pop}^{h-1}({}^J w_0) = e$. We deduce that ${}^J(f^{h-1}(w)) = e$, as desired. \square

Notice how the proof of Theorem 1.5 and the arguments leading up to it have utilized the right weak order in tandem with the strong Bruhat order in a subtle manner.

Let us now wrap up the proof of Theorem 1.3.

Proof of Theorem 1.3. Let W be an irreducible Coxeter group with Coxeter number h . We first assume W is finite. If h is odd, then we are done by Lemma 3.1, so assume h is even. Choose $s \in S$, and let $J = S \setminus \{s\}$. We saw in Proposition 3.4 that $\text{Pop}^{h-2}({}^J w_0) \neq e$, so $\max_{w \in W} |O_{\text{Pop}}(w)| \geq |O_{\text{Pop}}({}^J w_0)| \geq h$. On the other hand, the inequality $\max_{w \in W} |O_{\text{Pop}}(w)| \leq h$ follows immediately from Theorem 1.5 since Pop is compulsive.

Now assume W is infinite so that $h = \infty$. Let us first suppose that S is finite. Let $\mathcal{N}(W, S) = \{J \subseteq S : |W_J| < \infty\}$ be the nerve of (W, S) . Let $K = \max_{J \in \mathcal{N}(W, S)} \ell(w_0(J))$. Choose $v \in W$. Since v is an upper bound for $D_R(v)$ in the left weak order on W , it follows from [9, Lemma 2.3.2] that $D_R(v) \in \mathcal{N}(W, S)$. Using (2.1) and the fact that $\text{Pop}(v) \leq_R v$, we obtain

$$\begin{aligned} \ell(\text{Pop}(v)) &= \ell(v) - \ell((\text{Pop}(v))^{-1}v) \\ &= \ell(v) - \ell((w_0(D_R(v)))^{-1}) \\ &= \ell(v) - \ell(w_0(D_R(v))) \\ &\geq \ell(v) - K. \end{aligned}$$

As this is true for all $v \in W$, we find that $|O_{\text{Pop}}(w)| \geq \ell(w)/K + 1$ for all $w \in W$. An infinite irreducible Coxeter group contains arbitrarily long elements, so $\sup_{w \in W} |O_{\text{Pop}}(w)| = \infty$.

Finally, suppose S is infinite. Observe that if $J \subseteq S$, then $\text{Pop}_{W_J}(w) = \text{Pop}_W(w)$ for every $w \in W_J$. If there exists a finite set $S' \subseteq S$ such that $W_{S'}$ is infinite and irreducible, then it follows from the preceding paragraph that $\sup_{w \in W} |O_{\text{Pop}_W}(w)| \geq \sup_{w \in W_{S'}} |O_{\text{Pop}_{W_{S'}}}(w)| = \infty$. Now suppose no such set S' exists. Let $S_1 \subsetneq S_2 \subsetneq \dots$ be an infinite (strictly increasing) chain of finite subsets of S such that for each $i \geq 1$, the induced subgraph of $\Gamma(W)$ on the vertex set S_i is connected. The parabolic subgroups W_{S_i} are irreducible and finite. Let h_i be the Coxeter number of W_{S_i} . For every $i \geq 1$, we have $\sup_{w \in W} |O_{\text{Pop}_W}(w)| \geq \sup_{w \in W_{S_i}} |O_{\text{Pop}_{W_{S_i}}}(w)| = h_i$. The classification of Coxeter numbers [36, Section 3.18] implies that $h_i \rightarrow \infty$ as $i \rightarrow \infty$, so $\sup_{w \in W} |O_{\text{Pop}_W}(w)| = \infty$. \square

4. t -pop-stack-sortable elements in type B

We begin this section by recalling some basic facts about Coxeter groups of type B . The map $w \mapsto w_0ww_0$ is an automorphism of S_{2n} ; the permutations fixed by this automorphism form a subgroup B_n of S_{2n} called the n^{th} *hyperoctahedral group*. The map $w \mapsto w_0ww_0$ is also a lattice automorphism of the right weak order on S_{2n} , so (B_n, \leq_R) is a sublattice of (S_{2n}, \leq_R) . Let s_i denote the simple transposition in S_{2n} that swaps i and $i + 1$. The group B_n is a Coxeter group whose simple generators are the elements s_1^B, \dots, s_n^B given by $s_i^B = s_i s_{2n-i}$ for $i \in [n - 1]$ and $s_n^B = s_n$. Let us use \wedge_R^A (respectively, \wedge_R^B) and \bigwedge_R^A (respectively, \bigwedge_R^B) to denote meets in the right weak order on S_{2n} (respectively, B_n).

Fix $w \in B_n$, and let \mathcal{D} be the set of indices $i \in [n]$ such that s_i^B is a right descent of w in B_n . The right descent set of w as an element of S_{2n} is $\{s_i : i \in \mathcal{D}\} \cup \{s_{2n-i} : i \in \mathcal{D}\}$. If $i \in \mathcal{D} \cap [n - 1]$, then $ws_i \wedge_R^A ws_{2n-i} = ws_i s_{2n-i} = ws_i^B$ because s_i and s_{2n-i} commute. On the other hand, if $n \in \mathcal{D}$, then $ws_n \wedge_R^A ws_{2n-n} = ws_n = ws_n^B$. It follows that

$$\begin{aligned} \text{Pop}_{S_{2n}}(w) &= \bigwedge_R^A (\{ws_i : i \in \mathcal{D}\} \cup \{ws_{2n-i} : i \in \mathcal{D}\} \cup \{w\}) \\ &= \bigwedge_R^A (\{ws_i \wedge_R^A ws_{2n-i} : i \in \mathcal{D}\} \cup \{w\}) \\ &= \bigwedge_R^A (\{ws_i^B : i \in \mathcal{D}\} \cup \{w\}). \end{aligned}$$

Since (B_n, \leq_R) is a sublattice of (S_{2n}, \leq_R) , this shows that $\text{Pop}_{S_{2n}}(w) = \text{Pop}_{B_n}(w)$.

For $w \in S_{2n}$, the one-line notation of w_0ww_0 is

$$(2n + 1 - w(2n))(2n + 1 - w(2n - 1)) \cdots (2n + 1 - w(1)).$$

More geometrically, one can consider the *plot* of a permutation $v \in S_{2n}$, which is the set of points $(i, v(i)) \in \mathbb{R}^2$ for $i \in [2n]$. The plot of w_0ww_0 is obtained by rotating the plot of w by 180° about the point $(\frac{n+1}{2}, \frac{n+1}{2})$. The previous paragraph shows that Pop_{B_n} is the restriction of $\text{Pop}_{S_{2n}}$ to B_n . The main conclusion we wish to draw from this is that the t -pop-stack-sortable elements of B_n are precisely the t -pop-stack-sortable elements of S_{2n} that are fixed by the automorphism $w \mapsto w_0ww_0$. In other words, they are the t -pop-stack-sortable permutations in S_{2n} whose plots are invariant under 180° rotation.

4.1. Enumerating 2-pop-stack-sortable elements in type B

In order to enumerate the 2-pop-stack-sortable permutations in B_n , we first state a simple characterization of 2-pop-stack-sortable permutations in S_n from [45]; this characterization also follows easily from the description of the pop-stack-sorting map given in Section 2.2.

Lemma 4.1 ([45]). *Suppose $w \in S_n$ is a permutation whose descending runs, read from left to right, are $\delta_1, \dots, \delta_r$. Then w is 2-pop-stack-sortable if and only if for every $j \in [r - 1]$, the largest (equivalently, the first) entry in δ_j is at most 1 more than the smallest (equivalently, the last) entry in δ_{j+1} .*

For example, 62135847 is not 2-pop-stack-sortable because the largest entry in the first descending run is 6, the smallest entry in the second descending run is 3, and $6 > 3 + 1$.

In what follows, we will find it convenient to view permutations more generally as orderings of arbitrary finite subsets of \mathbb{Z} . For example, 3617 is the one-line notation of a permutation of size 4. Our convention is that S_n is the set of permutations of $[n]$. The *standardization* of a permutation w of size n is the permutation in S_n obtained by replacing the i^{th} -smallest entry in w with i for all $i \in [n]$. For example, the standardization of 3617 is 2314.

For $n \geq 1$, let $L(n, k)$ be the set of 2-pop-stack-sortable permutations in B_{n-1} that have either $2k$ or $2k + 1$ descending runs. Let $M(n, k)$ be the set of permutations $w \in L(n, k)$ such that the last descending run of w contains exactly 1 entry. We make the convention $L(1, 0) = M(1, 0) = \{\varepsilon\}$, where ε is the empty permutation. Furthermore, $L(1, k) = M(1, k) = \emptyset$ for $k \neq 0$. The only permutation in B_{n-1} that has 0 or 1 descending runs is the decreasing permutation $w_0(S_{2n-2}) = (2n-2)(2n-3) \cdots 321$, so $L(n, 0) = \{w_0(S_{2n-2})\}$. All of the entries in a decreasing permutation are in the same descending run, so $M(n, 0) = \emptyset$ whenever $n \geq 2$.

Proposition 4.2. *Preserve the notation from above. If $n \geq 2$ and $1 \leq k \leq n - 1$, then*

$$|L(n, k)| = 2 \sum_{i=1}^{n-1} |L(i, k-1)| - |M(n-1, k-1)|$$

and

$$|M(n, k)| = 2|L(n-1, k-1)| - |M(n-1, k-1)|.$$

Proof. This follows from Lemma 4.1 and the fact that a permutation is in a hyperoctahedral group if and only if its plot is invariant under 180° rotation. Indeed, suppose we are given $w \in L(n, k)$. If we remove the first and last descending runs and standardize the resulting permutation, we obtain a permutation $\alpha(w) \in \bigcup_{i=1}^{n-1} L(i, k-1)$. Hence, we have a map $\alpha : L(n, k) \rightarrow \bigcup_{i=1}^{n-1} L(i, k-1)$. Suppose $v \in L(i, k-1)$ for some $1 \leq i \leq n-1$. If $v \in M(n-1, k-1)$, then, upon inspecting the characterization of 2-pop-stack-sortable permutations in Lemma 4.1, we find that $|\alpha^{-1}(v)| = 1$. Indeed, the only element of $\alpha^{-1}(v)$ is (in the notation of Section 2.2) $1 \oplus v \oplus 1$. Now suppose $v \notin M(n-1, k-1)$. In this case, there are exactly 2 elements of $\alpha^{-1}(v)$. The first is $w_0(S_{n-i}) \oplus v \oplus w_0(S_{n-i})$ (recall from Section 2.2 that $w_0(S_{n-i})$ is the decreasing permutation of size $n-i$). The second is the permutation whose plot is obtained from that of $w_0(S_{n-i}) \oplus v \oplus w_0(S_{n-i})$ by sliding the rightmost point down so that it is immediately below the highest point in the second-to-last descending run and sliding the leftmost point up so that it is immediately above the lowest point in the second descending run. For example, suppose $n = 7$, $i = 4$, and $v = 264315$. The first element of $\alpha^{-1}(v)$ is $w_0(S_3) \oplus v \oplus w_0(S_3) = 321597648121110$. The other element of $\alpha^{-1}(v)$ is 521410763912118 . The latter permutation was obtained from 321597648121110 by decreasing the last entry so that it is 1 less than the largest entry in the second-to last descending run (i.e., $8 = 9 - 1$) and increasing the first entry so that it is 1 more than the smallest entry in the second descending run (i.e., $5 = 4 + 1$). It follows that $|L(n, k)|$ is equal to

$$|M(n-1, k-1)| + 2 \left| \bigcup_{i=1}^{n-1} L(i, k-1) \setminus M(n-1, k-1) \right| = 2 \sum_{i=1}^{n-1} |L(i, k-1)| - |M(n-1, k-1)|.$$

Note that if $w \in M(n, k)$, then $\alpha(w) \in L(n - 1, k - 1)$. Let $\beta : M(n, k) \rightarrow L(n - 1, k - 1)$ be the restriction of α to $M(n, k)$. If $v \in M(n - 1, k - 1)$, then we saw above that the unique element of $\alpha^{-1}(v)$ is $1 \oplus v \oplus 1$, which is in $M(n, k)$. Thus, $|\beta^{-1}(v)| = 1$ in this case. If $v \in L(n - 1, k - 1) \setminus M(n - 1, k - 1)$, then we saw above how to construct the two elements of $\alpha^{-1}(v)$. Both of these elements are actually in $M(n, k)$, so $|\beta^{-1}(v)| = 2$ in this case. This shows that $|M(n, k)|$ is equal to

$$2|L(n-1, k-1) \setminus M(n-1, k-1)| + |M(n-1, k-1)| = 2|L(n-1, k-1)| - |M(n-1, k-1)|. \quad \square$$

Proof of Theorem 1.8. In [45], Pudwell and Smith define $a(n, k)$ to be the number of 2-pop-stack-sortable permutations in S_n that have exactly $k + 1$ descending runs. They also define $b(n, k)$ to be the number of 2-pop-stack-sortable permutations $w \in S_n$ with exactly $k + 1$ descending runs such that the last descending run of w has exactly one entry. They prove (see [45, Proposition 1]) that $a(n, k)$ and $b(n, k)$ satisfy the same recurrence relation that we found for $|L(n, k)|$ and $|M(n, k)|$ in Proposition 4.2, even with the same initial conditions. It follows that $|L(n, k)| = a(n, k)$ and $|M(n, k)| = b(n, k)$. This proves the first statement in Theorem 1.8 and shows that

$$\sum_{n \geq 1} |\text{Pop}_{S_n}^{-2}(e)| z^n = \sum_{n \geq 1} |\text{Pop}_{S_{n+1}}^{-2}(e)| z^n.$$

Pudwell and Smith also found that the generating function $\sum_{n \geq 0} |\text{Pop}_{S_n}^{-2}(e)| z^n$ that counts 2-pop-stack-sortable permutations is equal to $\frac{1 - z - z^2 - z^3}{1 - 2z - z^2 - 2z^3}$ (see [45, Corollary 1]). Consequently,

$$\sum_{n \geq 1} |\text{Pop}_{S_{n+1}}^{-2}(e)| z^n = \frac{1}{z} \left(\frac{1 - z - z^2 - z^3}{1 - 2z - z^2 - 2z^3} - 1 \right) = \frac{2z(1 + z + z^2)}{1 - 2z - z^2 - 2z^3}. \quad \square$$

4.2. Generating functions for t -pop-stack-sortable elements in type B

We now proceed to prove Theorem 1.10, which states that for each fixed $t \geq 0$, the generating function that counts t -pop-stack-sortable permutations in hyperoctahedral groups is rational. Our proof makes heavy use of the tools that Claesson and Guðmundsson developed in [19].

Fix an integer $t \geq 1$. Consider a t -pop-stack-sortable permutation $w \in S_n$. Let us write out the one-line notations of $w, \text{Pop}(w), \text{Pop}^2(w), \dots, \text{Pop}^{t-1}(w)$, with $\text{Pop}^k(w)$ directly below $\text{Pop}^{k-1}(w)$ for each $k \in [t - 1]$. Draw boxes around the descending runs of each of these permutations. The resulting array of numbers and boxes is called the *sorting trace* of w . If we delete the numbers in this sorting trace, we obtain an array of boxes called the *sorting plan* of w . We call n and t the *length* and *order*, respectively, of the sorting trace and the sorting plan. For a concrete example, see Figure 4.1, which shows the sorting trace and sorting plan of order 4 of the permutation 6524713. Let $\text{SP}_n(t)$ denote the set of all sorting plans of length n and order t (i.e., the set of all sorting plans of t -pop-stack-sortable permutations in S_n). Let $\text{SP}(t) = \bigcup_{n \geq 1} \text{SP}_n(t)$. Each box in the sorting trace or sorting plan of w is called a *block*; the *length* of a block is the number of entries that it contains in the sorting trace.

6	5	2	4	7	1	3
2	5	6	4	1	7	3
2	5	1	4	6	3	7
2	1	5	4	3	6	7

Figure 4.1: The sorting trace (left) and sorting plan (right) of order 4 of 6524713.

Notice that the sorting trace of w is completely determined by the sorting plan of w . Indeed, for each $k \in [t]$, the permutation $\text{Pop}^k(w)$ is obtained by reversing the descending runs of $\text{Pop}^{k-1}(w)$. The lengths of the descending runs of $\text{Pop}^{k-1}(w)$ are precisely the lengths of the blocks in the k^{th} row of the sorting plan of w . Thus, $\text{Pop}^{k-1}(w)$ can be obtained by inserting $\text{Pop}^k(w)$ into the k^{th} row of the sorting plan of w and then reversing the entries within each block. Since $\text{Pop}^t(w) = e$, this shows that we can determine w , and, hence, the sorting trace of w , from the sorting plan of w . It follows that t -pop-stack-sortable permutations in S_n are in bijection with sorting plans of length n and order t .

Define a *bar code* of order t to be a sequence of t vertical bars and blank spaces, each of height 1, arranged vertically from top to bottom. If we associate each vertical bar with the digit 0 and associate each blank space with the digit 1, then a bar code corresponds to an element of $\{0, 1\}^t$. We associate each element (x_1, \dots, x_t) of $\{0, 1\}^t$ with the integer

$$x_t + 2x_{t-1} + 2^2x_{t-2} + \dots + 2^{t-1}x_1 \in \Sigma_t,$$

where $\Sigma_t = \{0, \dots, 2^t - 1\}$. Hence, we have a bijective correspondence between bar codes of order t and elements of Σ_t . For example, with $t = 4$, the bar codes in Figure 4.2 correspond, from left to right, to the vectors $(0, 0, 0, 0)$, $(0, 0, 1, 0)$, and $(1, 1, 0, 1)$. These, in turn, correspond to the numbers 0, 2, and 13 in Σ_4 .



Figure 4.2: Three bar codes of order 4.

A *segment* of length n and order t is a sequence of $n + 1$ bar codes of order t that are arranged from left to right and separated by spaces of width 1, along with $t + 1$ separating horizontal lines of length n . For instance, the left side of Figure 4.3 depicts a segment of length 4 and order 4. Being (essentially) a sequence of bar codes, a segment σ corresponds to a word $\psi(\sigma)$ over the alphabet Σ_t . For example, if σ is the segment on the left of Figure 4.3, then $\psi(\sigma)$ is the word 9 10 5 5 10. Notice that ψ is an injection from the set of segments of order t into the set of words over Σ_t . Every sorting plan is a segment; we will be primarily interested in the set $\psi(\text{SP}(t))$ of words over Σ_t that correspond to sorting plans of order t . The number of t -pop-stack-sortable permutations in S_n is equal to $|\psi(\text{SP}_n(t))|$, so the generating function whose rationality was demonstrated by Claesson and Guðmundsson in [19] is $\sum_{n \geq 1} |\psi(\text{SP}_n(t))| z^n$.

We say a segment σ *contains* a segment σ' if the word $\psi(\sigma)$ contains the word $\psi(\sigma')$ as a factor (i.e., a contiguous subword). For example, if σ is the sorting plan on the right side of Figure 4.1 and σ' is the segment on the left side of Figure 4.3, then σ contains σ' because $\psi(\sigma) = 0910551040$ and $\psi(\sigma') = 9105510$.



Figure 4.3: On the left is a segment of length 4 and order 4. On the right, the left-open (respectively, right-open) blocks of this segment have been shaded in pink (respectively, blue).

Imagine enclosing a segment σ of length n and order t inside a rectangle R of width n and height t . The connected components of the complement of σ in the interior of R are called the *blocks* of σ . A block is *left-open* if it touches R on its left side but does not touch σ on its left side. A block is *right-open* if it touches R , but not σ , on its right side. On the right side of Figure 4.3, the left-open and right-open blocks of the segment have been shaded pink and blue, respectively. We say a segment is *bounded* if each of its blocks has length at most 3.

An *operation array* is a segment σ such that $\psi(\sigma)$ starts and ends with 0 (equivalently, σ has no left-open or right-open blocks). Suppose we are given an operation array σ of length n and order t . Place the identity permutation $123 \cdots n$ below σ . Now fill the rows of σ , one at a time, from bottom to top. At each step, copy the numbers in the $(k + 1)^{\text{th}}$ row into the k^{th} row and then reverse the numbers within each block in the k^{th} row. Let T be the resulting array of blocks and numbers, excluding the identity permutation at the bottom. The array T is called the *semitrace* of σ (see Figure 4.4). Consider a pair of integers (a, b) with $1 \leq a < b \leq n$. Let $X_{a,b}(T)$ be the collection of blocks in T that contain a or b and do not lie in the first row of T ; let $X_{a,b}(\sigma)$ be the corresponding collection of blocks in σ . Let $\sigma_{a,b}$ be the smallest segment contained in σ that includes all of the blocks in $X_{a,b}(\sigma)$. We say the pair (a, b) is a *violating pair* of the semitrace T if there is some row of T that either contains a and b in the same block with a immediately before b or contains a and b in different blocks with b immediately before a . We say the segment $\sigma_{a,b}$ is *forbidden* if (a, b) is a violating pair in T . Claesson and Guðmundsson showed that whether or not the segment $\sigma_{a,b}$ is forbidden only depends on the segment $\sigma_{a,b}$ itself; it does not depend on the pair (a, b) or the operation array σ in which $\sigma_{a,b}$ is embedded.

Lemma 4.3 ([19, Proposition 3.9]). *A segment σ of order t is a sorting plan if and only if the following conditions hold:*

- σ is an operation array;
- every block of σ that is not in the first row of σ has length at most 3;
- σ does not contain any bounded forbidden segments.

Lemma 4.4 ([19, Lemma 3.12]). *There are finitely many bounded forbidden segments of order t .*

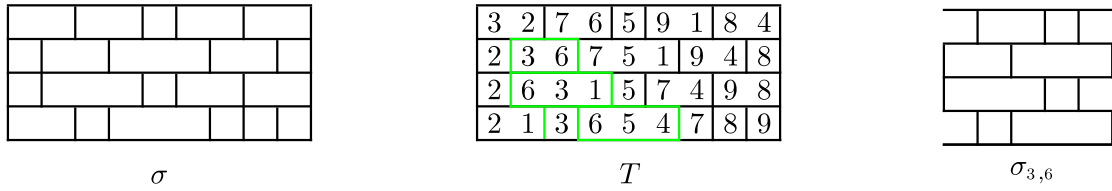


Figure 4.4: An operation array σ (left) and its corresponding semitrace T (middle). The blocks in $X_{3,6}(T)$ are shown in green; the corresponding blocks in σ form the set $X_{3,6}(\sigma)$. The segment $\sigma_{3,6}$ (right) is the smallest segment contained in σ that includes all of the blocks in $X_{3,6}(\sigma)$. The violating pairs of T are $(3, 6)$ (because of the second row) and $(5, 6)$ (because of the first row). Since $(3, 6)$ is a violating pair of T , the segment $\sigma_{3,6}$ is forbidden.

To prove their main result, Claesson and Guðmundsson employed the theory of formal languages; we will do the same. We recall the basic notions from this theory, referring the reader to [39] for more information.

Let \mathcal{A} be a nonempty finite alphabet. A *language* over \mathcal{A} is a collection of finite (possibly empty) words over \mathcal{A} . Given a language \mathcal{L} , let \mathcal{L}^* denote the set of all finite words, including the empty word, that can be obtained by concatenating words from \mathcal{L} . The *concatenation* of two languages \mathcal{L}_1 and \mathcal{L}_2 is the language $\mathcal{L}_1\mathcal{L}_2 = \{xy : x \in \mathcal{L}_1, y \in \mathcal{L}_2\}$. The *reverse* of a language \mathcal{L} is $\text{rev}(\mathcal{L}) = \{\text{rev}(x) : x \in \mathcal{L}\}$, where, for $a_1, \dots, a_k \in \mathcal{A}$, we write $\text{rev}(a_1 \cdots a_k) = a_k \cdots a_1$.

A language is *regular* if it is the set of words accepted by a deterministic finite automaton. The following lemma lists several standard properties of the collection of regular languages; we refer to [39, Chapter 4] for its proof.

Lemma 4.5. *Let \mathcal{A} be an alphabet. Every finite language over \mathcal{A} is regular. If $\mathcal{L}, \mathcal{L}_1, \dots, \mathcal{L}_k$ are regular languages over \mathcal{A} , then*

$$\mathcal{L}^*, \quad \bigcup_{i=1}^k \mathcal{L}_i, \quad \bigcap_{i=1}^k \mathcal{L}_i, \quad \mathcal{L}_1\mathcal{L}_2, \quad \mathcal{A}^* \setminus \mathcal{L}, \quad \text{rev}(\mathcal{L})$$

are all regular.

Let \mathcal{A}_n denote the set of words over \mathcal{A} of length n . The crucial fact that we need states that if \mathcal{L} is a regular language, then the generating function $\sum_{n \geq 1} |\mathcal{L} \cap \mathcal{A}_n| z^n$ is rational (see, e.g., [30, Proposition I.2]). Claesson and Guðmundsson showed that $\psi(\text{SP}(t))$ is a regular language over the alphabet Σ_t . We have seen that the words in $\psi(\text{SP}(t))$ are in bijection with t -pop-stack-sortable permutations, so the regularity of $\psi(\text{SP}(t))$ implies the rationality of the generating function that counts t -pop-stack-sortable permutations in symmetric groups.

We are almost ready to prove Theorem 1.9. We just need to discuss one additional concept. Say a segment σ is *symmetric* if the word $\psi(\sigma)$ is a palindrome. In other words, a segment is symmetric if it is left unchanged when we reflect it through a central vertical axis. Observe that a segment is symmetric if and only if each of its rows is a symmetric segment of order 1. Define a *type-B sorting plan* to be a symmetric sorting plan of even length.

Lemma 4.6. *Let σ be the sorting plan of order t of a t -pop-stack-sortable permutation $w \in S_{2n}$. Then σ is a type- B sorting plan if and only if $w \in B_n$.*

Proof. Suppose $w \in B_n$. We have seen that Pop_{B_n} is the restriction of $\text{Pop}_{S_{2n}}$ to B_n . Therefore, the permutations in the rows of the sorting trace of w are

$$w, \text{Pop}_{B_n}(w), \text{Pop}_{B_n}^2(w), \dots, \text{Pop}_{B_n}^{t-1}(w).$$

If $k \in [t]$ and $\delta_1, \dots, \delta_r$ are the descending runs (from left to right) of $\text{Pop}_{B_n}^{k-1}(w)$, then the length of δ_i is equal to the length of δ_{r+1-i} for each $i \in [r]$ (because $\text{Pop}_{B_n}^{k-1}(w) \in B_n$). In other words, the k^{th} row of σ is a symmetric segment of length $2n$ and order 1. As this is true for every $k \in [t]$, σ must be a type- B sorting plan.

Conversely, suppose σ is a type- B sorting plan. We will prove by backward induction on k that $\text{Pop}^{k-1}(w) \in B_n$ for all $k \in [t + 1]$; setting $k = 1$ will then show that $w \in B_n$. We certainly have $\text{Pop}^t(w) = e \in B_n$, so let us choose $k \in [t]$ and assume we have already proven that $\text{Pop}^k(w) \in B_n$. Let m_1, \dots, m_r be the lengths of the descending runs (from left to right) of $\text{Pop}^{k-1}(w)$. Let λ be the unique layered permutation in S_{2n} whose descending runs (from left to right) have lengths m_1, \dots, m_r . By the discussion in Section 2.2, we have $\text{Pop}^k(w) = \text{Pop}^{k-1}(w)\lambda$. Since σ is symmetric and m_1, \dots, m_r are the lengths of the blocks in the k^{th} row of σ , we have $m_i = m_{r+1-i}$ for all $i \in [r]$. It follows that $\lambda \in B_n$. Thus, $\text{Pop}^{k-1}(w) = \text{Pop}^k(w)\lambda^{-1} \in B_n$. \square

Proof of Theorem 1.9. If $t = 0$, then $\sum_{n \geq 1} |\text{Pop}_{B_n}^{-t}(e)| z^n = \frac{z}{1-z}$ is rational, so we may assume

$t \geq 1$. Given a palindromic word $x = x_1 \cdots x_{2n} \in \Sigma_t^{2n}$, let $\text{half}(x) = x_1 \cdots x_n \in \Sigma_t^n$. Note that x is uniquely determined by $\text{half}(x)$ since $x = \text{half}(x) \text{rev}(\text{half}(x))$. Let $\text{SP}^B(t)$ be the set of type- B sorting plans of order t . Each of the words in $\psi(\text{SP}^B(t))$ is a palindrome of even length, so it makes sense to define the language $\text{half}(\psi(\text{SP}^B(t))) = \{\text{half}(\psi(\sigma)) : \sigma \in \text{SP}^B(t)\}$.

We have seen that the map sending each t -pop-stack-sortable permutation to its sorting plan is a bijection from the set of t -pop-stack-sortable permutations in S_{2n} to $\text{SP}_{2n}^B(t)$. It follows from Lemma 4.6 that the number of t -pop-stack-sortable elements of B_n is equal to the number of type- B sorting plans of length $2n$ and order t . This is also equal to the number of words of length n in the language $\text{half}(\psi(\text{SP}^B(t)))$. Therefore, in order to prove that the generating function $\sum_{n \geq 1} |\text{Pop}_{B_n}^{-t}(e)| z^n$ is rational, it suffices to show that $\text{half}(\psi(\text{SP}^B(t)))$ is a regular language.

Let U be the set of words $u \in \Sigma_t^*$ such that $\psi^{-1}(u)$ and $\psi^{-1}(\text{rev}(u))$ do not contain any bounded forbidden segments while $\psi^{-1}(u \text{rev}(u))$ does contain a bounded forbidden segment. Consider the following six properties that a word $x \in \Sigma_t^*$ may or may not have:

- (i) x begins with the letter 0;
- (ii) every block of $\psi^{-1}(x)$ that is not in the first row has length at most 3;
- (iii) every right-open block of $\psi^{-1}(x)$ that is not in the first row has length at most 1;
- (iv) $\psi^{-1}(x)$ does not contain any bounded forbidden segments;

- (v) $\psi^{-1}(\text{rev}(x))$ does not contain any bounded forbidden segments;
- (vi) no suffix of x is in U .

It follows from Lemma 4.3 that a word $x \in \Sigma_t^*$ satisfies all six of these properties if and only if $\psi^{-1}(x \text{ rev}(x))$ is a sorting plan. Indeed, saying x satisfies (i) is equivalent to saying that $\psi^{-1}(x \text{ rev}(x))$ is an operation array. Saying x satisfies (ii) and (iii) is equivalent to saying that every block of $\psi^{-1}(x \text{ rev}(x))$ that is not in the first row of $\psi^{-1}(x \text{ rev}(x))$ has length at most 3. Finally, saying that x satisfies (iv), (v), and (vi) is equivalent to saying that $\psi^{-1}(x \text{ rev}(x))$ does not contain any bounded forbidden segments. Consequently, $\text{half}(\psi(\text{SP}^B(t)))$ is the set of words satisfying all six of the above properties. For each $P \in \{\text{i, ii, iii, iv, v, vi}\}$, let \mathcal{L}_P be the language of words over Σ_t^* that satisfy the property (P). Since, by Lemma 4.5, the intersection of a finite collection of regular languages is regular, our proof will be complete once we demonstrate that each of the languages \mathcal{L}_P is regular.

We make repeated tacit use of Lemma 4.5 in this paragraph. We have $\mathcal{L}_i = \{0\}\Sigma_t^*$, so \mathcal{L}_i is regular. Let Z_i be the set of letters $m \in \Sigma_t$ such that, when m is written in binary as $x_t + 2x_{t-1} + 2^2x_{t-2} + \cdots + 2^{t-1}x_1$, we have $x_i = 1$. Equivalently, Z_i is the set of elements m of Σ_t such that the i^{th} entry of the bar code $\psi^{-1}(m)$ is a blank space. We have

$$\mathcal{L}_{\text{ii}} = \bigcap_{i=2}^t (\Sigma_t^* \setminus (\Sigma_t^* Z_i Z_i Z_i \Sigma_t^*)) \quad \text{and} \quad \mathcal{L}_{\text{iii}} = \bigcap_{i=2}^t (\Sigma_t^* \setminus (\Sigma_t^* Z_i Z_i)),$$

so \mathcal{L}_{ii} and \mathcal{L}_{iii} are regular. Let \mathcal{F}_t be the set of bounded forbidden segments of order t , which is finite by Lemma 4.4. For each $f \in \mathcal{F}_t$, the set of words x such that $\psi^{-1}(x)$ does not contain f is $\Sigma_t^* \setminus (\Sigma_t^* \{\psi(f)\} \Sigma_t^*)$. Therefore, the language $\mathcal{L}_{\text{iv}} = \bigcap_{f \in \mathcal{F}_t} (\Sigma_t^* \setminus (\Sigma_t^* \{\psi(f)\} \Sigma_t^*))$ is regular. Furthermore, \mathcal{L}_{v} is regular because it is the reverse of \mathcal{L}_{iv} .

We are left with the task of proving that \mathcal{L}_{vi} is regular. Let $|x|$ denote the length of a word x . Since the set \mathcal{F}_t of bounded forbidden segments of order t is finite by Lemma 4.4, there exists a positive integer K such that $|\psi(f)| \leq K$ for all $f \in \mathcal{F}_t$. Let U' be the set of words $u \in U$ such that no proper suffix of u is in U . Consider $u \in U'$. Let a be the first letter of u , and write $u = au'$. Note that $u' \notin U$ because $u \in U'$. Since $u \in U$, we can write $u = vv'$ and $\text{rev}(u) = y'y$ so that v' and y' are nonempty and $\psi^{-1}(v'y') \in \mathcal{F}_t$. Either v or y must be empty since, otherwise, u' would be in U . Therefore, $|u| = \max\{|v'|, |y'|\} < |v'y'| \leq K$. This proves that every word in U' has length at most K , so U' is finite. In particular, U' is a regular language by Lemma 4.5. Note that a word x is in \mathcal{L}_{vi} if and only if no suffix of x is in U' . In symbols, this says that $\mathcal{L}_{\text{vi}} = \Sigma_t^* \setminus (\Sigma_t^* U')$, so \mathcal{L}_{vi} is regular by Lemma 4.5. \square

5. t -pop-stack-sortable elements in type \tilde{A}

For $n \geq 1$, an *affine permutation* of size n is a bijection $w : \mathbb{Z} \rightarrow \mathbb{Z}$ such that

$$w(i+n) = w(i) + n \quad \text{for all } i \in \mathbb{Z} \tag{5.1}$$

and

$$\sum_{i=1}^n w(i) = \binom{n+1}{2}. \tag{5.2}$$

The set \tilde{S}_n of affine permutations of size n forms a group under composition; it is a Coxeter group of type A_{n-1} . The simple generators are $\tilde{s}_1, \dots, \tilde{s}_n$, where \tilde{s}_i is the affine permutation that swaps $i + mn$ and $i + mn + 1$ for all $m \in \mathbb{Z}$ and fixes all other elements of \mathbb{Z} . The simple generator \tilde{s}_i is a right descent of an affine permutation w if and only if $w(i) > w(i + 1)$. Furthermore, $\ell(w)$ is equal to the number of pairs $(i, j) \in [n] \times \mathbb{Z}$ such that $i < j$ and $w(i) > w(j)$.

It will be useful to consider the one-line notation of a bijection $w : \mathbb{Z} \rightarrow \mathbb{Z}$, which is simply the bi-infinite word $\dots w(-2)w(-1)w(0).w(1)w(2)\dots$. The decimal point between $w(0)$ and $w(1)$ is meant to indicate which letters are indexed by which integers. For example, $\dots(-1)0.123\dots$ represents the identity element e of \tilde{S}_n , while $\dots(-1)01.23\dots$ represents the bijection given by $i \mapsto i + 1$, which is not an affine permutation because it fails to satisfy (5.2).

A *descending run* of an affine permutation w is a maximal consecutive decreasing subsequence of w . We say an affine permutation $w \in \tilde{S}_n$ is *layered* if there exists $k \in \{0, \dots, n - 1\}$ such that $(w(k + 1) - k)(w(k + 2) - k)\dots(w(k + n) - k)$ is a layered permutation in S_n . In other words, we can think of a layered affine permutation as an infinite direct sum of decreasing permutations. Just as for symmetric groups, one can show that an affine permutation $w \in \tilde{S}_n$ is layered if and only if $w = w_0(D_R(w))$. Hence, for an arbitrary $w \in \tilde{S}_n$, we can compute $\text{Pop}_{\tilde{S}_n}(w)$ (using (2.1)) by multiplying w on the left by the unique layered affine permutation that has the same right descent set as w . Using this description, it is straightforward to check that $\text{Pop}_{\tilde{S}_n}(w)$ is obtained by reversing all of the descending runs of w while keeping entries in different descending runs in the same relative order.

Lemma 5.1. *Let $w \in \tilde{S}_n$. Every descending run of $\text{Pop}(w)$ has length at most 3.*

Proof. Let $\text{Pop}(w) = v$, and suppose, by way of contradiction, that there exists $i \in \mathbb{Z}$ such that $v(i) > v(i + 1) > v(i + 2) > v(i + 3)$. For $j \in \{0, 1, 2, 3\}$, let δ_j be the descending run of w that contains the entry $v(i + j)$. Since v is obtained by reversing the descending runs of w , the descending runs $\delta_0, \delta_1, \delta_2, \delta_3$ are distinct and appear consecutively (in this order) in w . This implies that the only entry in δ_1 is $v(i + 1)$ and that the only entry in δ_2 is $v(i + 2)$. Since $v(i + 1) > v(i + 2)$, this contradicts the fact that δ_1 and δ_2 are distinct descending runs. \square

Our goal in this section is to prove Theorem 1.10, which states that the generating function $\sum_{n \geq 1} |\text{Pop}_{\tilde{S}_n}^{-t}(e)| z^n$ is rational. First, we should check that this generating function is even well-defined! In other words, we should verify that for each $t \geq 0$, there are only finitely many t -pop-stack-sortable affine permutations in \tilde{S}_n . This follows from our proof of Theorem 1.3. In that proof, we showed that $|O_{\text{Pop}_{\tilde{S}_n}}(w)| \geq \ell(w)/K + 1$ for all $w \in \tilde{S}_n$, where $K = \max_{J \in \mathcal{N}(\tilde{S}_n, S)} \ell(w_0(J))$. Since there are only finitely many elements of \tilde{S}_n of each fixed length, there are only finitely many t -pop-stack-sortable affine permutations in \tilde{S}_n .

We are going to make use of the ideas from the previous section concerning sorting traces, sorting plans, segments, and semitraces, but we need to modify them for the affine setting. Fix $t \geq 1$, and consider a t -pop-stack-sortable affine permutation $w \in \tilde{S}_n$. Write out the one-line notations of $w, \text{Pop}(w), \text{Pop}^2(w), \dots, \text{Pop}^{t-1}(w)$, with $\text{Pop}^k(w)$ directly below $\text{Pop}^{k-1}(w)$ for each $k \in [t - 1]$. Draw boxes around the descending runs of each of these affine permutations.

The resulting bi-infinite array of numbers and boxes is the *affine sorting trace* of w . Deleting the numbers in the sorting trace produces a bi-infinite array of boxes called the *affine sorting plan* of w . We call n and t the *period* and *order*, respectively, of the affine sorting trace and the affine sorting plan. Note that an affine sorting plan of period n also has period dn for each positive integer d . The columns of an affine sorting plan are indexed by \mathbb{Z} (so an affine sorting plan can change when it is shifted). Figure 5.1 shows the affine sorting trace and affine sorting plan of order 3 of the affine permutation $w \in \widetilde{S}_5$ with $w(1) = 0, w(2) = 3, w(3) = 2, w(4) = 6, w(5) = 4$. Denote by $\widetilde{SP}_n(t)$ the set of all affine sorting plans of period n and order t (i.e., the set of all affine sorting plans of t -pop-stack-sortable permutations in \widetilde{S}_n). Let $\widetilde{SP}(t) = \bigcup_{n \geq 1} \widetilde{SP}_n(t)$. As in the non-affine case, boxes in affine sorting traces and affine sorting plans are called *blocks*.

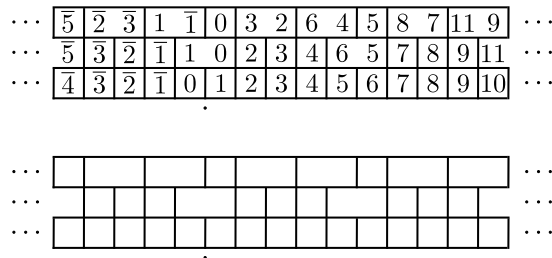


Figure 5.1: The affine sorting trace (top) and affine sorting plan (bottom) of order 3 of the affine permutation $w \in \widetilde{S}_5$ satisfying $w(1) = 0, w(2) = 3, w(3) = 2, w(4) = 6, w(5) = 4$. In each diagram, notice the decimal point between the column indexed by 0 and the column indexed by 1.

Bar codes have the same meaning that they did in the previous section, and we still have a natural bijection ψ between bar codes of order t and elements of the alphabet $\Sigma_t = \{0, \dots, 2^t - 1\}$. Segments also have the same meaning as in the previous section, except we now allow them to be bi-infinite (in which case, the columns are indexed by \mathbb{Z}). Blocks of segments are defined in the obvious way (blocks are now permitted to be infinitely long). In this setting, ψ extends to an injection from the set of segments of order t to the set of finite or bi-infinite words over Σ_t . We say a bi-infinite segment σ is *n-periodic* if $\psi(\sigma)$ is an n -periodic word. As before, we say a segment σ *contains* a segment σ' if $\psi(\sigma)$ contains $\psi(\sigma')$ as a factor. Finally, we say a segment σ is *non-Escher*⁵ if it does not contain any infinitely long blocks.

Now suppose we are given a non-Escher bi-infinite segment σ of order t , which we assume has its columns indexed by the integers. Place the identity affine permutation below σ so that for each $i \in \mathbb{Z}$, the number i is below column i . Fill the rows of σ from bottom to top. At each step, copy the numbers in the $(k + 1)$ th into the k th row and then reverse the numbers within each block in the k th row. The resulting array T , excluding the identity affine permutation at the very bottom, is called the *affine semitrace* of σ . Given a pair of integers (a, b) with $a < b$, we define

⁵This clever terminology is stolen from [1], where it was used in a similar but different manner. The fact that an affine sorting plan is non-Escher corresponds to the fact that an affine permutation in \widetilde{S}_n cannot have the entire set $S = \{\widetilde{s}_1, \dots, \widetilde{s}_n\}$ as its right descent set, contrary to M. C. Escher's classical pieces of art that portray staircases perpetually descending and somehow looping back on themselves.

$\sigma_{a,b}$ just as in the previous section. As before, we say the pair (a, b) is a *violating pair* of the affine semitrace T if there is some row of T that either contains a and b in the same block with a immediately before b or contains a and b in different blocks with b immediately before a . We say the segment $\sigma_{a,b}$ is *forbidden* if (a, b) is a violating pair in T .

Note that whether or not a finite-length segment is forbidden is a local property; this means that a segment of finite length is forbidden in the affine setting if and only if it is forbidden in the sense of the preceding section. Thus, the set \mathcal{F}_t of bounded forbidden segments of order t is the same as it was in the previous section. In particular, bounded forbidden segments are of finite length, and the set \mathcal{F}_t is finite by Lemma 4.4.

The map sending each affine permutation to its affine sorting plan is a bijection from $\text{Pop}_{\tilde{S}_n}^{-t}(e)$ to $\widetilde{\text{SP}}_n(t)$. Indeed, each $w \in \tilde{S}_n$ is determined by its affine sorting plan σ because w is the sequence of numbers in the first row of the affine semitrace of σ .

The following lemma serves as an affine version of Lemma 4.3.

Lemma 5.2. *A segment σ of order t is in $\widetilde{\text{SP}}_n(t)$ if and only if the following conditions hold:*

- σ is n -periodic and non-Escher;
- every block of σ that is not in the first row of σ has length at most 3;
- σ does not contain any bounded forbidden segments.

Proof. Suppose first that σ is in $\widetilde{\text{SP}}_n(t)$. This means that σ is the sorting plan of a t -pop-stack-sortable affine permutation $w \in \tilde{S}_n$. The lengths of the blocks in the k^{th} row of σ are the lengths of the descending runs of $\text{Pop}_{\tilde{S}_n}^{k-1}(w)$. The condition (5.1) ensures that, for each $k \in [t]$, the right descent set of $\text{Pop}_{\tilde{S}_n}^{k-1}(w)$ is not the entire set $S = \{\tilde{s}_1, \dots, \tilde{s}_n\}$, so σ is non-Escher. The condition (5.1) also guarantees that σ is n -periodic. Furthermore, Lemma 5.1 implies that every block of σ that is not in the first row has length at most 3. Because w is t -pop-stack-sortable, the affine trace of w is the same as the affine semitrace of σ . The affine trace of w cannot contain a violating pair since its blocks are constructed by putting boxes around the descending runs of the affine permutations in each row. It follows that σ cannot contain a forbidden segment. In particular, σ does not contain any bounded forbidden segments.

To prove the converse, assume σ satisfies the three bulleted conditions in the statement of the lemma. Construct the affine semitrace T of σ . Consider the step when we fill the k^{th} row of the semitrace by copying the numbers in the $(k + 1)^{\text{th}}$ row into the k^{th} row and then reversing the numbers within each block in the k^{th} row. If we already know that the entries in the $(k + 1)^{\text{th}}$ row form an affine permutation in \tilde{S}_n , then the entries in the k^{th} row must also form an affine permutation in \tilde{S}_n . Indeed, it follows from the first two bulleted conditions that the entries in the k^{th} row are obtained by multiplying the affine permutation in the $(k + 1)^{\text{th}}$ row by a layered affine permutation. Let w be the bounded affine permutation in the first row of T . It follows from the second bulleted condition that every segment contained in σ is bounded, so it follows from the third bulleted condition that σ does not contain any forbidden segments. Consequently, T does not contain any violating pairs. This implies that w is t -pop-stack-sortable and that T is the affine trace of w . Hence, $\sigma \in \widetilde{\text{SP}}_n(t)$. □

Before proving our final theorem, we need one additional lemma about regular languages. Given a word $x = x_1 \cdots x_n$ over a finite alphabet \mathcal{A} , define $\text{cyc}(x)$ to be the cyclic shift $x_2 \cdots x_n x_1$. For a language $\mathcal{L} \subseteq \mathcal{A}^*$, define $\text{cyc}(\mathcal{L}) = \{\text{cyc}(x) : x \in \mathcal{L}\}$.

Lemma 5.3 ([39, Chapter 4, Exercise 20]). *If \mathcal{L} is a regular language over a finite alphabet \mathcal{A} , then so is $\text{cyc}(\mathcal{L})$.*

Proof of Theorem 1.10. If $t = 0$, then $\sum_{n \geq 1} \left| \text{Pop}_{\widetilde{S}_n}^{-t}(e) \right| z^n = \frac{z}{1-z}$ is rational, so we may assume

$t \geq 1$. Let $K = \max\{K' + 1, 5\}$, where K' is the maximum length of a bounded forbidden segment of order t (this is finite by Lemma 4.4). To prove the theorem, it suffices to show that

$\sum_{n \geq K} \left| \text{Pop}_{\widetilde{S}_n}^{-t}(e) \right| z^n$ is rational.

Assume $n \geq K$. Suppose $w \in \widetilde{S}_n$, and let T and σ be the affine sorting trace and affine sorting plan, respectively, of w . The i^{th} column of T is the column of numbers whose top entry is $w(i)$. Recall the injection ψ that sends segments of order t to words over Σ_t . Also, recall that σ can be seen as a bi-infinite sequence of bar codes; let b_i be the bar code in T that is immediately to the left of the i^{th} column of T . Let $a_i \in \Sigma_t$ be the letter corresponding to b_i (via its binary expansion). Define $\alpha_n(\sigma)$ to be the segment of length $n - 1$ such that $\psi(\alpha_n(\sigma)) = a_1 a_2 \cdots a_n$. For example, if σ is the affine sorting plan in Figure 5.1, then

$$\alpha_5(\sigma) = \psi^{-1}(20404) = \begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} .$$

Notice that σ can be reconstructed from $\alpha_n(\sigma)$ because it is n -periodic by Lemma 5.2. Thus, α_n is an injection from $\widetilde{\text{SP}}_n(t)$ into the set of segments of length $n - 1$ and order t . The number of t -pop-stack-sortable affine permutations in \widetilde{S}_n is equal to $|\widetilde{\text{SP}}_n(t)|$, which is equal to $|\alpha_n(\widetilde{\text{SP}}_n(t))|$, which is equal to $|\psi(\alpha_n(\widetilde{\text{SP}}_n(t)))|$.

Consider the language $\mathcal{L} = \bigcup_{n \geq K} \psi(\alpha_n(\widetilde{\text{SP}}_n(t)))$ over Σ_t . In order to complete the proof of the theorem, it suffices to show that \mathcal{L} is regular. To do this, we define Y to be the set of finite words y over Σ_t such that:

- (I) y has length at least K ;
- (II) the first row of $\psi^{-1}(y)$ has at least one vertical bar;
- (III) every block of $\psi^{-1}(y)$ that is not in the first row of $\psi^{-1}(y)$ has length at most 3;
- (IV) $\psi^{-1}(y)$ does not contain any bounded forbidden segments.

Let $\mathcal{L}' = \bigcap_{k=0}^{K-1} \text{cyc}^k(Y)$, where $\text{cyc}^k(Y) = \{\text{cyc}^k(y) : y \in Y\}$.

We claim that $\mathcal{L}' = \mathcal{L}$. To see this, first suppose $x = x_1 \cdots x_n \in \mathcal{L}'$. Since $x \in Y$, we have $n \geq K$ by (I). Let $\sigma = \psi^{-1}(\cdots x x x x \cdots)$ be the bi-infinite segment obtained by concatenating $\psi^{-1}(x)$ with itself infinitely many times. We want to show that σ satisfies the three bulleted conditions in Lemma 5.2; it will then follow that $\sigma \in \widetilde{\text{SP}}_n(t)$ so that $x = \psi(\alpha_n(\sigma))$.

The n -periodicity of σ is clear from its definition, and the non-Escher property of σ follows from (II), (III), and the fact that $n \geq K \geq 5$. If σ has a block of length at least 4 that is not in its first row, then there is some $k \in \{0, 1, 2, 3, 4\}$ such that $\psi^{-1}(\text{cyc}^{-k}(x))$ has a block of length at least 4 not in its first row. However, this contradicts property (III) of $\text{cyc}^{-k}(x)$, which is in Y because $K \geq 5$. Finally, if σ contains a bounded forbidden segment, then there is some $k \in \{0, 1, \dots, K'\}$ such that $\psi^{-1}(\text{cyc}^{-k}(x))$ contains a bounded forbidden segment. This contradicts property (IV) of $\text{cyc}^{-k}(x)$, which is in Y because $K \geq K' + 1$. Hence, $\mathcal{L}' \subseteq \mathcal{L}$.

Now suppose $x \in \mathcal{L}$. Then $x = \psi(\alpha_n(\sigma))$, where σ is the affine sorting plan of order t of some $w \in \tilde{S}_n$ with $n \geq K$. Choose $k \geq 0$, and let $y = \text{cyc}^{-k}(x)$. The word y has length n , so it satisfies (I). Notice that $\psi(\sigma)$ is the bi-infinite word $\dots xx.xxxx \dots$, which is obtained by shifting $\dots yy.yyyy \dots$ by k . Lemma 5.2 tells us that σ is non-Escher, so y must satisfy (II). The second bulleted item in Lemma 5.2 implies that y satisfies (III). Similarly, the third bulleted item in Lemma 5.2 implies that y satisfies (IV). This proves that $y \in Y$. As k was arbitrary, $x \in \bigcap_{k \geq 0} \text{cyc}^k(Y) \subseteq \mathcal{L}'$.

We have now established that $\mathcal{L} = \mathcal{L}'$. We are left with the task of proving that \mathcal{L}' is a regular language. We make tacit use of Lemma 4.5. For $P \in \{I, II, III, IV\}$, let \mathcal{L}_P be the set of words y in Σ_t^* satisfying property (P). Note that \mathcal{L}_I is regular because it is equal to $(\Sigma_t)_K \Sigma_t^*$, where $(\Sigma_t)_K$ is the regular language consisting of all words in Σ_t^* of length K . Let Z_1 be the set of letters $m \in \Sigma_t$ such that, when m is written in binary as $x_t + 2x_{t-1} + 2^2x_{t-2} + \dots + 2^{t-1}x_1$, we have $x_1 = 1$. Equivalently, $m \in Z_1$ if and only if the bar code corresponding to m starts with a blank space. Then $\mathcal{L}_{II} = \Sigma_t^* \setminus (Z_1^*)$ is regular. The languages \mathcal{L}_{III} and \mathcal{L}_{IV} are equal to the languages \mathcal{L}_{ii} and \mathcal{L}_{iv} , respectively, from the proof of Theorem 1.9 in Section 4; we saw in that proof that these languages are regular. We conclude that $\mathcal{L}_I, \mathcal{L}_{II}, \mathcal{L}_{III}, \mathcal{L}_{IV}$ are regular languages, so their intersection Y is regular as well. It follows from Lemma 5.3 that $\text{cyc}^k(Y)$ is regular for each $k \in \{0, \dots, K - 1\}$. Hence, \mathcal{L}' is regular. \square

6. Further directions

In Section 1.2, we mentioned two extensions of Coxeter pop-stack-sorting operators: one to other complete meet-semilattices and one to other semilattice congruences on weak orders of Coxeter groups. We explore the first of these extensions in [23] and explore the second in [25]. Here, we mention some other potential avenues for future work.

The authors of [5] suggested considering the average size of the forward orbit of a permutation in S_n under the pop-stack-sorting map. In [22], the current author conjectured that this average number of iterations is asymptotically equal to n , which is the maximum possible size of a forward orbit by Ungar's theorem. We believe that the same statement should hold for Coxeter groups of other classical types as well. In these cases, the maximum possible size of the forward orbit of an element is the Coxeter number of the group by Theorem 1.3. The Coxeter number of B_n is $2n$, and the Coxeter number of D_n is $2n - 2$.

Conjecture 6.1. As $n \rightarrow \infty$, we have

$$\frac{1}{|B_n|} \sum_{w \in B_n} |O_{\text{Pop}_{B_n}}(w)| \sim 2n \quad \text{and} \quad \frac{1}{|D_n|} \sum_{w \in D_n} |O_{\text{Pop}_{D_n}}(w)| \sim 2n - 2.$$

In Section 1.2, we defined a map $\text{Pop}_M : M \rightarrow M$, where M is an arbitrary complete meet-semilattice. In Remark 1.11, we defined the notion of a *compulsive* map $f : M \rightarrow M$. We also exhibited a 6-element lattice M and a compulsive map $f : M \rightarrow M$ such that $\sup_{x \in M} |O_f(x)| > \sup_{x \in M} |O_{\text{Pop}_M}(x)|$. It could be interesting to investigate which complete meet semilattices M have the property that $\sup_{x \in M} |O_f(x)| \leq \sup_{x \in M} |O_{\text{Pop}_M}(x)|$ for every compulsive map $f : M \rightarrow M$; Theorems 1.3 and 1.5 tell us that weak orders of Coxeter groups have this property. It is also natural to consider this question only for finite meet-semilattices, or even just for finite lattices.

Acknowledgments

The author thanks Henri Mühle for pointing out the connection between this work and the articles [41, 42], and he thanks Nathan Williams for pointing out the connection between Pop and Brieskorn normal form. He thanks the anonymous referees for valuable suggestions. The author was supported by the National Science Foundation under Award No. DGE-1656466 and Award No. 2201907, by a Fannie and John Hertz Foundation Fellowship, and by a Benjamin Peirce Fellowship at Harvard University.

References

- [1] R. M. Adin, V. Reiner, and Y. Roichman, On cyclic descents for tableaux. *Int. Math. Res. Not. IMRN*, **2020** (2020), 10231–10276.
- [2] M. H. Albert, M. D. Atkinson, M. Bouvel, A. Claesson, and M. Dukes, On the inverse image of pattern classes under bubble sort. *J. Comb.*, **2** (2011), 231–243.
- [3] M. Albert, M. Bouvel, and V. Féray, Two first-order logics of permutations. *J. Combin. Theory Ser. A*, **171** (2020).
- [4] M. Albert and V. Vatter, How many pop-stacks does it take to sort a permutation? *Comput. J.*, (2021).
- [5] A. Asinowski, C. Banderier, and B. Hackl, Flip-sort and combinatorial aspects of pop-stack sorting. *Discrete Math. Theor. Comput. Sci.*, **22** (2021).
- [6] A. Asinowski, C. Banderier, S. Billey, B. Hackl, and S. Linusson, Pop-stack sorting and its image: permutations with overlapping runs. *Acta. Math. Univ. Comenian.*, **88** (2019), 395–402.
- [7] D. Avis and M. Newborn. On pop-stacks in series. *Util. Math.*, **19** (1981), 129–140.
- [8] A. Björner, Orderings of Coxeter groups. In *Combinatorics and algebra (Boulder, Colo., 1983)*, vol. 34 of Contemp. Math., American Mathematical Society, 1984.
- [9] A. Björner and F. Brenti, *Combinatorics of Coxeter groups*, vol. 231 of Graduate Texts in Mathematics. Springer, 2005.
- [10] M. Bóna, A survey of stack-sorting disciplines. *Electron. J. Combin.*, **9** (2003).

- [11] M. Bóna, Symmetry and unimodality in t -stack sortable permutations. *J. Combin. Theory Ser. A*, **98.1** (2002), 201–209.
- [12] N. Bourbaki, Lie groups and Lie algebras. Springer–Verlag, 2002.
- [13] P. Brändén, On linear transformations preserving the Pólya frequency property. *Trans. Amer. Math. Soc.*, **358** (2006), 3697–3716.
- [14] E. Brieskorn, Die fundamentalgruppe des raumes der regulären orbits einer endlichen komplexen spiegelungsgruppe. *Invent. Math.*, **12** (1971), 57–61.
- [15] E. Brieskorn and K. Saito, Artin-gruppen und Coxeter-gruppen. *Invent. Math.*, **17** (1972), 245–271.
- [16] G. Cerbai, A. Claesson, and L. Ferrari, Stack sorting with restricted stacks. *J. Combin. Theory Ser. A*, **173** (2020).
- [17] F. Chung, A. Claesson, M. Dukes, and R. Graham, Descent polynomials for permutations with bounded drop size. *European J. Combin.*, **31** (2010), 1853–1867.
- [18] L. Cioni and L. Ferrari, Preimages under the Queuesort algorithm. *Discrete Math.*, **344** (2021).
- [19] A. Claesson and B. Á. Guðmundsson, Enumerating permutations sortable by k passes through a pop-stack. *Adv. Appl. Math.*, **108** (2019), 79–96.
- [20] A. Claesson, B. Á. Guðmundsson, and J. Pantone, Counting pop-stacked permutations in polynomial time. arXiv:1908.08910.
- [21] C. Defant, Counting 3-stack-sortable permutations. *J. Combin. Theory Ser. A*, **172** (2020).
- [22] C. Defant, Fertility monotonicity and average complexity of the stack-sorting map. *European J. Combin.*, **93** (2021).
- [23] C. Defant, Meeting covered elements in ν -Tamari lattices. *Adv. Appl. Math.*, **134** (2022).
- [24] C. Defant, Polyurethane toggles. *Electron. J. Combin.*, **27** (2020).
- [25] C. Defant, Stack-sorting for Coxeter groups. *Comb. Theory*, **2** (2022).
- [26] C. Defant, Troupes, cumulants, and stack-sorting. *Adv. Math.*, **399** (2022).
- [27] C. Defant, A. Elvey Price, and A. J. Guttmann, Asymptotics of 3-stack-sortable permutations. *Electron. J. Combin.*, **28** (2021).
- [28] M. Dukes, Revstack sort, zigzag patterns, descent polynomials of t -revstack sortable permutations, and Steingrímsson’s sorting conjecture. *Electron. J. Combin.*, **21** (2014).
- [29] M. Elder and Y. K. Goh, k -pop stack sortable permutations and 2-avoidance. *Electron. J. Combin.*, **28** (2021).
- [30] P. Flajolet and R. Sedgewick, Analytic combinatorics. Cambridge University Press, 2009.
- [31] M. Geck and G. Pfeiffer, Characters of finite Coxeter groups and Iwahori–Hecke algebras, vol. 21 of London Mathematical Society Monographs, New Series, The Clarendon Press/Oxford University Press, 2000.
- [32] S. Giraudo. Algebraic and combinatorial structures on pairs of twin binary trees. *J. Algebra*, **360** (2012), 115–157.

- [33] J. E. Goodman and R. Pollack, A combinatorial perspective on some problems in geometry. *Congr. Numer.*, **32** (1981), 383–394.
- [34] F. Hivert, J.-C. Novelli, and J.-Y. Thibon, The algebra of binary search trees. *Theoret. Comput. Sci.*, **339** (2005), 129–165.
- [35] H. P. Hoang and T. Mütze, Combinatorial generation via permutation languages. II. Lattice congruences. To appear in *Israel J. Math.*, (2021).
- [36] J. E. Humphreys, Reflection groups and Coxeter groups. Cambridge University Press, 1990.
- [37] D. E. Knuth, The Art of Computer Programming, volume 3, Sorting and Searching. Addison–Wesley, Reading, 2nd ed., 1998.
- [38] S. Law and N. Reading. The Hopf algebra of diagonal rectangulations. *J. Combin. Theory Ser. A*, **119** (2012), 788–824.
- [39] P. Linz, An introduction to formal languages and automata. Jones & Bartlett Learning, 5th ed., 2012.
- [40] H. Magnusson, Sorting operators and their preimages, M.Sc. Thesis, Reykjavík University, 2013.
- [41] H. Mühle, The core label order of a congruence-uniform lattice. *Algebra Universalis*, **80** (2019).
- [42] H. Mühle, Hochschild lattices and shuffle lattices. *European J. Combin.*, **103** (2022).
- [43] V. Pilaud. Brick polytopes, lattice quotients, and Hopf algebras. *J. Combin. Theory Ser. A*, **155** (2018), 418–457.
- [44] V. Pilaud and V. Pons, Permutrees. *Algebr. Comb.*, **1** (2018), 173–224.
- [45] L. Pudwell and R. Smith, Two-stack-sorting with pop stacks. *Australas. J. Combin.*, **74** (2019), 179–195.
- [46] N. Reading, Cambrian lattices. *Adv. Math.*, **205** (2006), 313–353.
- [47] N. Reading. Finite Coxeter groups and the weak order. In *Lattice theory: special topics and applications*. vol. 2, pages 489–561. Birkhäuser/Springer, 2016.
- [48] N. Reading, Lattice congruences, fans and Hopf algebras. *J. Combin. Theory Ser. A*, **110** (2005), 237–273.
- [49] N. Reading and D. Speyer, Cambrian frameworks for cluster algebras of affine type. *Trans. Amer. Math. Soc.*, **370** (2018), 1429–1468.
- [50] N. Reading and D. Speyer, Sortable elements in infinite Coxeter groups. *Trans. Amer. Math. Soc.*, **363** (2011), 699–761.
- [51] J. R. Stembridge, On the fully commutative elements of Coxeter groups. *J. Algebraic Combin.*, **5** (1996), 353–385.
- [52] P. Ungar, $2N$ noncollinear points determine at least $2N$ directions. *J. Combin. Theory Ser. A*, **33** (1982), 343–347.
- [53] J. West, Permutations with restricted subsequences and stack-sortable permutations, Ph.D. Thesis, MIT, 1990.