HOMOMORPHISMS OF COMMUTATOR SUBGROUPS OF BRAID GROUPS

KEVIN KORDEK AND DAN MARGALIT

Abstract. We give a complete classification of homomorphisms from the commutator subgroup of the braid group on \( n \) strands to the braid group on \( n \) strands when \( n \) is at least 7. In particular, we show that each nontrivial homomorphism extends to an automorphism of the braid group on \( n \) strands. This answers four questions of Vladimir Lin. Our main new tool is the theory of totally symmetric sets.

1. INTRODUCTION

Let \( B_n \) denote the braid group on \( n \) strands and let \( B'_n \) denote its commutator subgroup. We say that two homomorphisms \( \rho_1 : B'_n \to B_n \) and \( \rho_2 : B'_n \to B_n \) are equivalent if there is an automorphism \( \alpha \) of \( B_n \) such that \( \alpha \circ \rho_1 = \rho_2 \). The following is our main result.

Theorem 1.1. Let \( n \geq 7 \), and let \( \rho : B'_n \to B_n \) be a nontrivial homomorphism. Then \( \rho \) is equivalent to the inclusion map.

In his 1996 preprint, Vladimir Lin asks the following four questions about endomorphisms of \( B'_n \) [18, 0.9.2(b)–0.9.2(e)]:

- Is every nontrivial endomorphism of \( B'_n \) injective?
- Is every nontrivial endomorphism of \( B'_n \) equal to an automorphism of \( B'_n \)?
- Does every nontrivial endomorphism of \( B'_n \) extend to an endomorphism of \( B_n \)?
- Does every nontrivial endomorphism of \( B'_n \) extend to an automorphism of \( B_n \)?

The second and fourth questions also appear in the online problem list “Open problems in combinatorial and geometric group theory” [1, Problems B5(b) and B7(b)] and in the published version of the same problem list [4, Problems B6(b) and B8(b)]. Theorem 1.1 answers in the affirmative all four of these questions for \( n \geq 7 \). Indeed, since the inclusion map \( B'_n \to B_n \) extends to the identity map \( B_n \to B_n \), Theorem 1.1 implies an affirmative answer to the fourth question. The third question is hence answered in the affirmative because automorphisms are endomorphisms. And since an automorphism of any group restricts to an automorphism of any characteristic subgroup, this answers the first two questions in the affirmative as well.

After the first version of our paper appeared, Orevkov [21] extended Theorem 1.1 by classifying homomorphisms \( B'_n \to B_n \) for arbitrary \( n \geq 1 \).

Prior results. In 2017, Orevkov [21] showed for \( n \geq 4 \) that \( \text{Aut}(B'_n) \cong \text{Aut}(B_n) \). Another proof of Orevkov’s result for \( n \geq 7 \) was given by McLeay [20]. Theorem 1.1 gives another proof of Orevkov’s result for \( n \geq 7 \).

In his 2004 paper, Lin [19, Theorem A] proved there are no nontrivial homomorphisms from \( B'_n \) to \( B_m \) when \( n \geq 5 \) and \( m < n \). Theorem 1.1 also implies Lin’s result for \( n \geq 7 \).


Both authors were supported by the National Science Foundation under Grant No. DMS-1057874.
injective homomorphisms $B_n \to B_{n+1}$. Then in 2016 Castel [10] classified for $n \geq 6$ all homomorphisms $B_n \to B_{n+1}$. Castel’s theorem implies the theorem of Dyer–Grossman. In Section 4 we explain how our Theorem 1.1 gives a new proof of Castel’s classification of endomorphisms of $B_n$ for $n \geq 7$. In particular, our work gives a new proof of the Dyer–Grossman result for $n \geq 7$.

**New tool: totally symmetric sets.** The main new tool we use to prove Theorem 1.1 is the notion of a totally symmetric set, which we define in Section 2. Briefly, a totally symmetric set in a group $G$ is a subset $X$ of commuting elements with the property that each permutation of $X$ can be achieved by a single conjugation in $G$. Totally symmetric sets have been used in several subsequent works:

1. Chudnovsky, Li, Partin, and the first author [13] give a lower bound for the cardinality of a finite non-abelian quotient of the braid group,
2. Kordek, Li, and Partin [17] give upper bounds on the cardinalities of totally symmetric sets in various types of groups,
3. Caplinger and the first author [9] show that the smallest non-abelian finite quotients of $B_5$ and $B_6$ are the corresponding symmetric groups,
4. Chen and Mukherjea [12] classify homomorphisms from $B_n$ to the mapping class group of a surface of genus $g \leq n - 3$, and
5. Scherich and Verberne [22] improved on the aforementioned lower bound of Chudnovsky, Li, Partin, and the first author.

As such, totally symmetric sets seem to be of interest independently of our main result.

**Spaces of polynomials.** Theorem 1.1 has implications for spaces of polynomials. Let $\text{Poly}_n$ denote the space of monic, square-free polynomials of degree $n$. This is the same as the space of unordered configurations of $n$ distinct points in the plane (the $n$ points are the roots). The fundamental group $\pi_1(\text{Poly}_n)$ is isomorphic to $B_n$.

Similarly, let $\text{SPoly}_n$ denote the space of monic, square-free polynomials of degree $n$ and discriminant 1. The discriminant gives a map $\text{Poly}_n \to \mathbb{C} \setminus \{0\}$; this map is a fiber bundle with fiber $\text{SPoly}_n$. Since $\mathbb{C} \setminus \{0\}$ is a $K(G,1)$ space it follows that $\pi_1(\text{SPoly}_n)$ embeds into $\pi_1(\text{Poly}_n)$ and the isomorphism from $\pi_1(\text{Poly}_n)$ to $B_n$ induces an isomorphism from $\pi_1(\text{SPoly}_n)$ to $B'_n$. Because of these identifications, Theorem 1.1 gives constraints on maps from $\text{SPoly}_n$ to $\text{Poly}_n$.

**Outline of the paper.** In Section 2, we introduce totally symmetric sets. We also prove the following fundamental lemma: the image of a totally symmetric set under a homomorphism is either a totally symmetric set of the same cardinality or a singleton (Lemma 2.1). The section culminates with a classification of certain totally symmetric subsets of $B_n$ (Lemma 2.6). In Section 3 we prove Theorem 1.1 using the classification of totally symmetric sets and the fundamental lemma. Finally, in Section 4, we apply Theorem 1.1 to prove the aforementioned special case of Castel’s theorem.

**Acknowledgments.** We are grateful to Lei Chen and Justin Lanier for helpful discussions about totally symmetric sets. We are also grateful to Vladimir Shpilrain for pointing out the connection between our work and the questions of Vladimir Lin. We would also like to thank Noah Caplinger and an anonymous referee for comments and corrections.
2. Totally symmetric sets

In this section we introduce the main new technical tool in the paper, namely, totally symmetric sets. After giving some examples, we derive some basic properties of totally symmetric sets, in particular developing the relationship with canonical reduction systems. The main results in this section are the fundamental lemma for totally symmetric sets (Lemma 2.1) and a classification of certain totally symmetric subsets of $B_n$ (Lemma 2.6).

**Totally symmetric subsets of groups.** Let $X$ be a subset of a group $G$. We may conjugate $X$ by an element $g$ of $G$, meaning that we conjugate each element of $X$ by $g$. We say that $X$ is totally symmetric if

- the elements of $X$ commute pairwise and
- each permutation of $X$ can be achieved via conjugation by an element of $G$.

As a first example, any singleton $\{x\}$ is totally symmetric. Another example is the set of transpositions $\{(1\ 2),(3\ 4),\ldots,(m\ m+1)\}$ in the symmetric group $\Sigma_n$, where $m$ is an odd integer less than $n$.

We say that a totally symmetric set $X \subseteq G$ is totally symmetric with respect to a subgroup $H$ of $G$ if $X$ satisfies the above definition, with the additional constraint that the conjugating elements can be chosen to lie in $H$. We observe that if $X \subseteq G$ is totally symmetric with respect to $H \leq G$, and $X \subseteq H$, then $X$ is a totally symmetric subset of $H$.

The definition of a totally symmetric set is inspired by the work of Aramayona–Souto, who studied a particular example of a totally symmetric set in their work on homomorphisms between mapping class groups [2, Section 5].

**New totally symmetric sets from old.** Let $X = \{x_1,\ldots,x_m\}$ be a totally symmetric subset of $G$. There are several ways of obtaining new totally symmetric sets from $X$. Let $k, \ell \in \mathbb{Z}$ and let $z$ be an element of $G$ that commutes with $z$ (for example $z$ can lie in $Z(G)$). Also, for each $i$ let $x_i^*$ denote the product of the $x_j \in X$ with $j \neq i$. Starting from $X$, we may create the following totally symmetric sets:

\[
X^k = \{x_1^k,\ldots,x_m^k\} \\
X^* = \{x_1^*,\ldots,x_m^*\} \\
X^{k,\ell} = \{x_1^k(x_1^*)^\ell,\ldots,x_m^k(x_m^*)^\ell\} \\
X' = \{x_1x_2^{-1},\ldots,x_1x_m^{-1}\} \\
X^z = \{x_1z,\ldots,x_mz\}.
\]

We can combine these constructions, for instance $(X^k)^z$ and $(X^*)^z$ are totally symmetric. Also, if all permutations of $X$ are achievable by elements of a subgroup $H$ of $G$, then the same is true for $X^k$, $X^*$, $X'$, and $X^z$.

**The fundamental lemma.** We have the following fundamental fact about totally symmetric sets. It is an analog of Schur’s lemma from representation theory.

**Lemma 2.1.** Let $X$ be a totally symmetric subset of a group $G$ and let $\rho : G \to H$ be a homomorphism of groups. Then $\rho(X)$ is either a singleton or a totally symmetric set of cardinality $|X|$.
Proof. It is clear from the definition of a totally symmetric set that \( \rho(X) \) is totally symmetric and that its cardinality is at most \( |X| \). Suppose that the restriction of \( \rho \) to \( X \) is not injective; say \( \rho(x_1) = \rho(x_2) \). For any \( x_i \in X \) there is (by the definition of a totally symmetric set) a \( g \in G \) so that \( (gx_1g^{-1}, gx_2g^{-1}) = (x_1, x_2) \). Thus
\[
\rho(x_1x_2^{-1}) = \rho((gx_1g^{-1})(gx_2^{-1}g^{-1})) = \rho(g)\rho(x_1x_2^{-1})\rho(g)^{-1} = 1.
\]
The lemma follows. \( \square \)

Totally symmetric sets in braid groups. In the braid group \( B_n \) the most basic example of a totally symmetric set is
\[
X_n = \{\sigma_1, \sigma_3, \sigma_5, \ldots, \sigma_m\}
\]
where \( m \) is the largest odd integer less than \( n \). (Here the \( \sigma_i \) are the standard Artin generators for \( B_n \). Also, when writing elements of \( B_n \) we compose elements right to left.) As above, the sets \( X_n^k, X_n^*, X_n^\prime \), and \( X^\delta \) are totally symmetric. In the following lemma, let \( z \in B_n \) be a generator for the center \( Z(B_n) \); the signed word length of \( z \) is \( n(n-1) \). We also define
\[
Y_n = X_n', \quad \text{and} \quad \quad Z_n = \left( X_n^{n(n-1)} \right)^{z^{-1}}.
\]

Lemma 2.2. Let \( n \geq 2 \). The set \( X_n \subseteq B_n \) is totally symmetric with respect to \( B_n' \). In particular, \( Y_n \) and \( Z_n \) are totally symmetric subsets of \( B_n' \).

Proof. Suppose some permutation \( \tau \) of \( X_n \) is achieved by \( g \in B_n \). Since \( \sigma_1 \) commutes with each element of \( X_n \), the permutation \( \tau \) is also achieved by \( g\sigma_1^k \) for all \( k \in \mathbb{Z} \). If we take \( k \) to be the negative of the signed word length of \( g \), then \( g\sigma_1^k \) lies in \( B_n' \). The first statement follows. The second statement follows similarly, once we observe that each element of \( Y_n \) and \( Z_n \) lies in \( B_n' \). \( \square \)

The goal of the remainder of the section is to classify certain totally symmetric subsets of size \( \lfloor n/2 \rfloor \) in \( B_n \). The end result is Lemma 2.6 below. The proof requires three auxiliary results, Lemmas 2.3, 2.4, and 2.5.

Totally symmetric multicurves. The first tool is a topological version of total symmetry. Let \( Y \) be a set and let \( S \) be a surface. We say that a multicurve \( M \) is \( Y \)-labeled if each component of \( M \) is labeled by a non-empty subset of \( Y \). The symmetric group \( \Sigma_Y \) acts on the set of \( Y \)-labeled multicurves by acting on the labels. The mapping class group \( \text{Mod}(S) \)—the group of homotopy classes of orientation-preserving homeomorphisms of \( S \) fixing the boundary of \( S \)—also acts on the set of \( Y \)-labeled multicurves via its action on the set of multicurves.

Let \( M \) be a \( Y \)-labeled multicurve in \( S \). We say that \( M \) is totally symmetric if for every \( \sigma \in \Sigma_Y \) there is an \( f \in \text{Mod}(S) \) so that \( \sigma \cdot M = f \cdot M \). As in the case of totally symmetric sets, we say that \( M \) is totally symmetric with respect a subgroup \( H \) of \( \text{Mod}(S) \) if the elements \( f \) from the definition can all be chosen to lie in \( H \).

We say that a \( Y \)-labeled multicurve has the trivial labeling if each component of the multicurve has the label \( Y \) (recall that empty labels are not allowed). Every such multicurve is totally symmetric (with respect to any subgroup \( H \) of \( \text{Mod}(S) \)). We also say that a component of a \( Y \)-labeled multicurve has the trivial label if its label is \( Y \).

We can describe a \( Y \)-labeled multicurve in a surface \( S \) as a set of pairs \( \{(d_i, A_i)\} \) where each \( d_i \) is a curve in \( S \), where each \( A_i \) is a subset of \( Y \), and where \( \{d_i\} \) is a multicurve in \( S \).
Let \( Y \) be a set. If \( M = \{(d_1, A_1), \ldots, (d_m, A_m)\} \) is a totally symmetric \( Y \)-labeled multicurve in a surface \( S \), and all of the labels \( A_i \) are nontrivial, then we may create new totally symmetric multicurves from \( Y \) as follows. For a subset \( A \) of \( Y \), we denote by \( A^c \) the complement \( Y \setminus A \). The new totally symmetric multicurve is

\[
M^* = \{(d_1, A_1^c), \ldots, (d_m, A_m^c)\}
\]

Let \( N = \lfloor n/2 \rfloor \), and let \([N]\) denote the set \( \{1, \ldots, N\} \). For \( 1 \leq i \leq N \) let \( c_i \) be the curve in \( D_n \) with the property that the (left) half-twist \( H_{c_i} \) about \( c_i \) is \( \sigma_{2i-1} \). The \([N]\)-labeled multicurve

\[
M_n = \{(c_1, \{1\}), (c_2, \{2\}), \ldots, (c_N, \{N\})\}
\]

is totally symmetric.

For the statement of the next lemma, we require several further definitions. First, for \( H \) a subgroup of \( \text{Mod}(S) \), we say that two labeled multicurves in \( S \) are \( H \)-equivalent if they lie in the same orbit under \( H \).

Next, let \( c_0 \) denote the standard curve in \( D_n \) that surrounds the first \( n - 1 \) marked points (so that \( c_0 \) is disjoint from \( c_1, \ldots, c_N \) ). The multicurve in \( D_n \) (with \( n \) odd) whose components are \( c_0, \ldots, c_N \) is depicted in Figure 1. For \( n \) odd, let \( \hat{M}_n \) and \( \hat{M}_n^* \) be the labeled multicurves \( \hat{M}_n = M_n \cup \{c_0, [N]\} \) and \( \hat{M}_n^* = M_n^* \cup \{c_0, [N]\} \); these are depicted in Figure 2.

**Lemma 2.3.** Let \( n \geq 1 \), let \( N = \lfloor n/2 \rfloor \).

1. If \( n \) is even, then every totally symmetric \([N]\)-labeled multicurve in \( D_n \) with nontrivial labeling is \( B_n \)-equivalent to \( M_n \) or \( M_n^* \).
2. If \( n \) is odd, then every totally symmetric \([N]\)-labeled multicurve in \( D_n \) with nontrivial labeling is \( B_n \)-equivalent to \( M_n, M_n^*, \hat{M}_n, \) or \( \hat{M}_n^* \).

**Proof.** Say that \( M \) is a totally symmetric \([N]\)-labeled multicurve in \( D_n \) with nontrivial label. Let \( c \) be a curve in \( M \) with nontrivial label \( \emptyset \neq A \subseteq [N] \). The symmetric group \( \Sigma_N \) acts on the power set of \([N]\) and the orbit of \( A \) under this action has \( \ell \geq N \) elements. Since \( M \) is totally symmetric, there must be for each element \( A' \) of this orbit a curve \( d \) in \( M \) so that: (1) \( d \) lies in the same \( B_n \)-orbit as \( c \) and (2) the label for \( d \) is \( A' \). In particular, \( M \) contains distinct curves \( d_1, \ldots, d_\ell \) that all lie in the same \( B_n \)-orbit. It follows that \( \ell = N \), that each \( d_i \) surrounds exactly two marked points, and that the labels are either of the form \( \{i\} \) or \( \{i\}^c \). We can further conclude that there are no other curves in \( M \) with nontrivial
Figure 2. Left: the labeled multicurve $\hat{M}_n$; Right: the labeled multicurve $\hat{M}_n^*$

label besides $d_1, \ldots, d_N$. Up to $B_m$-equivalence, we may therefore assume that the labeled multicurve $\{d_1, \ldots, d_N\}$ is exactly $M_n$ or $M_n^*$.

Let $T = \{b_1, \ldots, b_k\}$ be the set of curves of $M$ with trivial label. The curves of $T$ induce a partition of the set $\{d_1, \ldots, d_N\}$: the curves $d_i$ and $d_j$ are in the same subset of the partition if and only if they are not separated by an element of $T$. Since the $b_i$ are essential and distinct from the $d_i = c_i$, it must be that either (1) $k = 1$ and (up to $B_m$-equivalence) $b_1 = c_0$ or (2) the partition is nontrivial, meaning that it contains more than one subset. The second possibility violates the assumption that $M$ is totally symmetric. The lemma follows.

From totally symmetric sets to totally symmetric multicurves. Associated to each element $f$ of Mod($S$) is its canonical reduction system $\Gamma(f)$, which is a multicurve. We will make use of several basic facts about canonical reduction systems. First, we have $\Gamma(f) = \emptyset$ if and only if $f$ is periodic or pseudo-Anosov. Next, if $f$ and $g$ commute then $\Gamma(f) \cap \Gamma(g) = \emptyset$. Also, for any $f$ and $g$ we have $\Gamma(gfg^{-1}) = g\Gamma(f)$. See the paper by Birman–Lubotzky–McCarthy for background on canonical reduction systems [7].

Given a totally symmetric subset $X = \{x_1, \ldots, x_m\}$ of Mod($S$) we obtain an $[m]$-labeled multicurve as follows: the underlying multicurve $M$ is obtained from the disjoint union of the $\Gamma(x_i)$ by identifying homotopic curves, and the label of a curve $c \in M$ is the set of $i$ with $c$ a component of $\Gamma(x_i)$. We denote this $[m]$-labeled multicurve by $\Gamma(X)$. We have the following lemma, which follows immediately from the definitions and the stated facts about canonical reduction systems.

Lemma 2.4. If $X$ is a totally symmetric subset of Mod($S$) then $\Gamma(X)$ is totally symmetric.

The totally symmetric multicurves associated to $X_n$, $Y_n$, and $Z_n$ are

\[
\Gamma(X_n) = \{(c_1, \{1\}), (c_2, \{2\}), \ldots, (c_N, \{N\})\} = M_n,
\]

\[
\Gamma(Y_n) = \{(e_1, \{N\})\} \cup \{(c_2, \{1\}), \ldots, (c_N, \{N-1\})\}, \text{ and}
\]

\[
\Gamma(Z_n) = \{(c_1, \{1\}), \ldots, (c_N, \{N\})\} = M_n.
\]

Classification of derived totally symmetric subsets. Let $X = \{x_1, \ldots, x_m\}$ be a totally symmetric subset of a group $G$. In this case, we say that a totally symmetric set $Y$ in $G$ is derived
from $X$ if $Y$ lies in the free abelian subgroup $\langle X \rangle$ of $G$. We have already seen examples of derived totally symmetric sets, such as $X^k$, $X^*$, $X^{k,\ell}$, and $X'$.

Let $X$ be a totally symmetric subset of a group $G$, and let $Y$ be a derived totally symmetric subset. We consider the action of $G$ on itself by conjugation and write $\text{Stab}_G(X)$ and $\text{Stab}_G(Y)$ for the stabilizers of the sets $X$ and $Y$. We say that the derived totally symmetric set $Y$ is robust if

$$\text{Stab}_G(Y) \subseteq \text{Stab}_G(X).$$

As an example in $G = B_n$, the totally symmetric set $Y = X^{k,\ell}_n$ is a robust totally symmetric set in $X_n$ as long as at least one of $k$ and $\ell$ is nonzero. Indeed, since each $c_j$ lies in the canonical reduction system for some element of $X^{k,\ell}_n$, any element of $\text{Stab}_G(X^{k,\ell}_n)$ preserves the set of curves $\{c_1, \ldots, c_N\}$ and hence lies in $\text{Stab}_G(X_n)$.

**Lemma 2.5.** Let $X = \{x_1, \ldots, x_m\}$ be a totally symmetric subset of a group $G$, and let $Y$ be a robust derived totally symmetric set with $m$ elements. Then $Y$ is equal to some $X^{k,\ell}$.

**Proof.** We may assume that $m \geq 2$, for otherwise the lemma is trivial. Say that the elements of $Y$ are $y_1, \ldots, y_m$ and that the elements of $X$ have order $d$. Then the elements of $Y$ can be written as

$$y_i = x_1^{a_{i,1}} \cdots x_m^{a_{i,m}},$$

where each $a_{i,j}$ lies in $\mathbb{Z}/d\mathbb{Z}$ (when $d = \infty$ we interpret $\mathbb{Z}/d\mathbb{Z}$ as $\mathbb{Z}$). Let $A$ be the $m \times m$ matrix $(a_{i,j})$. As such, the $i$th row of $A$ records the exponents on the $x_j$ in the expression for $y_i$.

The statement of the lemma is equivalent to the statement that there exist $k, \ell \in \mathbb{Z}/d\mathbb{Z}$ such that, up to reordering the rows of $A$, we have

$$A = \begin{pmatrix}
k & \ell & \cdots & \ell \\
\ell & k & \cdots & \ell \\
s & s & \cdots & s \\
\ell & \ell & \cdots & k
\end{pmatrix}$$

We claim that any permutation of the rows of $A$ is achieved by a permutation of the columns of $A$. Let $\sigma \in \Sigma_m$ and let $g \in G$ be such that $gy_ig^{-1} = y_{\sigma(i)}$. This determines a permutation of the rows of $A$. By the total symmetry of $Y$, every permutation of the rows arises in this way. On the other hand, since $Y$ is robust, the conjugating element $g$ also permutes $X$, and hence determines a permutation of the columns of $A$. As both permutations have the same effect on the set $Y$, the claim follows.

We next claim that if $v$ is a column of $A$, and $w$ is an element of $(\mathbb{Z}/d\mathbb{Z})^m$ obtained by permuting the entries of $v$, then $w$ is also a column of $A$. The claim follows from the previous claim and the fact that any permutation of the entries of $v$ can be achieved by a permutation of the rows of $A$.

It must be that some column of $A$ has at least two distinct entries; if not, then the rows of $A$ are equal, violating that assumption that the $y_i$ are distinct. Let $v$ be such a column of $A$. It must be that, up to reordering the rows of $A$, we have $v = (k, \ell, \ldots, \ell)$ for some $k$ and $\ell$. Indeed, otherwise, there would be more than $m$ distinct permutations of the entries of $v$, violating the previous claim. It further follows from the previous claim that the $m$ columns of $A$ are the $m$ distinct permutations of the entries of $v$. After reordering the rows, $A$ has the desired form. \qed
Lemma 2.6. Let $n \geq 1$, let $N = \lfloor n/2 \rfloor$, and let $X = \{x_1, \ldots, x_N\}$ be a totally symmetric subset of $B_n$. Assume that $\Gamma(X)$ is nonempty and has nontrivial labeling.

1. If $n$ is even, $X$ is $B_n$-equivalent to $(X_0^n)_{\ell}^z$ or $((X_0^n)^s)^{\ell}z^s$ for some $\ell, s \in \mathbb{Z}$ with $\ell \neq 0$.

2. If $n$ is odd, $X$ is $B_n$-equivalent to $(X_0^n)^{T_{c_0}^{z^s}}$ or $((X_0^n)^{s})^{T_{c_0}^{z^s}}$ for some $\ell, r, s \in \mathbb{Z}$ with $\ell \neq 0$.

Proof of Lemma 2.6. We also restrict to the case $n \geq 4$, since in the other cases both the statement and the proof degenerate (for instance when $n = 3$ we have that $T_{c_0} = \sigma_1^2$ and so the $T_{c_0}$ term is not needed). We further restrict to the case where $n$ is odd. The other case is similar (and simpler). By Lemma 2.4, the multicurve $\Gamma(X)$ is totally symmetric. It then follows from Lemma 2.3 that $\Gamma(X)$ is $B_n$-equivalent to $M_n, M_n^*, \hat{M}_n$, or $\hat{M}_n^*$. We may assume then that $\Gamma(X)$ is in fact equal to $M_n, M_n^*, \hat{M}_n$, or $\hat{M}_n^*$. We discuss the cases $\hat{M}_n$ and $\hat{M}_n^*$ in turn, the other cases again being similar.

Suppose first that $\Gamma(X)$ is equal to $\hat{M}_n$. This means that $\Gamma(x_1)$ is equal to $\{c_0, c_1\}$. The multicurve $\Gamma(x_1)$ divides $D_n$ into three regions, each corresponding to a periodic or pseudo-Anosov Nielsen–Thurston component of $x_1$. To prove the lemma in this case it suffices to show that the outer two Nielsen–Thurston components of $x_1$ are trivial. Indeed, it follows from this that $x_1$ is of the desired form $H_{c_1}^{\ell} T_{c_0}^{r} z^s$ with $\ell \neq 0$, and then by the total symmetry that each $x_i$ is of the desired form $H_{c_i}^{\ell} T_{c_0}^{r} z^s$.

The outermost Nielsen–Thurston component of $x_1$ (exterior to $c_0$) is necessarily trivial since this outermost region is a pair of pants (after collapsing the boundary components to points), and since $x_1$ fixes the three marked points.

It remains to show that the Nielsen–Thurston component of $x_1$ corresponding to the region lying between $c_0$ and $c_1$ is trivial. Since $n \geq 4$, we have $N \geq 2$. And since $x_1$ commutes with $x_2$, it follows that $x_1$ fixes $\Gamma(x_2) = \{c_2\}$. From this, it immediately follows that the Nielsen–Thurston component in question is not pseudo-Anosov. Therefore it is periodic. If we collapse $c_0$ and $c_1$ to marked points, the region lying between $c_0$ and $c_1$ becomes a sphere with $n - 1$ marked points. A periodic mapping class of this sphere is a rotation fixing the marked points coming from $c_0$ and $c_1$. Since $x_1$ fixes $c_2$ (which surrounds two marked points) it follows that the rotation is trivial, completing the proof of the lemma in this case.

We now address the case where $\Gamma(X)$ is equal to $\hat{M}_n^*$. In this case, it follows as in the previous case that each $x_i$ is of the form

$$x_i = P_i T_{c_0}^{r} z^s$$

where each $P_i$ is a product of nonzero powers of half-twists about the elements of the set $\{c_1, \ldots, c_N\} \setminus \{c_i\}$ (it follows from the total symmetry that $r$ and $s$ are independent of $i$). Let $Y$ be the totally symmetric set $Y = X^{T_{c_0}^{r} z^s}$. This $Y$ has $N$ elements and is a robust totally symmetric set derived from $X_n$ (the argument is the same as the above argument that $X_{n}^{k,\ell}$ is robust in $X_n$). By Lemma 2.5, $Y$ is equal to $X_{k,\ell}$ for some $k, \ell \in \mathbb{Z}$. It then follows from the fact that $\Gamma(X) = \hat{M}_n$ that $k = 0$ and $\ell \neq 0$. This implies that $Y = (X_0^n)^{\ell}$. It follows that

$$X = Y^{T_{c_0}^{r} z^s} = ((X_0^n)^{s})^{T_{c_0}^{r} z^s},$$

as desired. \qed
3. Proof of Theorem 1.1

In this section, we prove Theorem 1.1, which states that, for \( n \geq 7 \), every homomorphism \( \rho : B'_n \to B_n \) is equivalent to the inclusion map. We will require two tools, a direct product decomposition for certain subgroups of \( B_n \), and a lemma about uniqueness of roots of certain types of braids.

The cabling decomposition. Let \( M \) be a multicurve in \( D_n \). Lemma 3.1 below gives a semi-direct product decomposition for \( \text{Stab}_{B_n}(M) \), the stabilizer of \( M \) in \( B_n \). In order to state it we require some setup.

Let \( \Gamma \) be the graph defined as follows: there is a distinguished vertex corresponding to the boundary of \( D_n \), there are vertices for each of the marked points in \( D_n \), and there are vertices corresponding to the components of \( M \); the edges correspond to immediate nesting, meaning that two vertices are connected by an edge if one of the corresponding curves or marked points is nested in the other and there are no components of \( M \) in between.

The graph \( \Gamma \) is a rooted tree, with the distinguished vertex as the root. For each vertex \( v \) there is a corresponding rooted subtree \( \Gamma_v \) with \( v \) as the root. The tree \( \Gamma \) has exactly \( n \) leaves, one for each marked point of \( D_n \). Similarly, the leaves of \( \Gamma_v \) correspond to the marked points of \( D_n \) contained in the interior of the curve corresponding to \( v \).

Let \( M_0 \) be the multicurve whose components are the outermost components of \( M \). These correspond to the vertices of \( \Gamma \) adjacent to the root. The multicurve \( M_0 \) is un-nested, meaning that no component is contained in the disk bounded by another component.

Let \( \Delta_0 \) be the disk with marked points obtained from \( D_n \) by crushing to a marked point each disk bounded by a component of \( M_0 \). We label each of the marked points by the corresponding rooted tree \( \Gamma_v \). Let \( B_{\Delta_0} \) denote the subgroup of the mapping class group of \( \Delta_0 \) that preserves the labels. If \( \Delta_0 \) has \( \ell \) marked points, then \( B_{\Delta_0} \) is a subgroup of \( B_\ell \).

There is an induced map

\[
\Pi : \text{Stab}_{B_n}(M) \to B_{\Delta_0}.
\]

We would like to describe the kernel.

Say that \( M_0 \) has \( q \) components, which surround \( n_1, \ldots, n_q \) marked points, respectively. Let \( v_1, \ldots, v_q \) be the corresponding vertices of \( \Gamma \). Let \( \Delta_1, \ldots, \Delta_q \) be the disks bounded by the components of \( M_0 \). Each \( \Delta_i \) inherits a multicurve \( M_i \) from \( M \), which is given by the components of \( M \) that are nested inside \( M_i \); the corresponding rooted tree is \( \Gamma_{v_i} \). Let \( \text{Stab}_{B_{\Delta_i}}(M_i) \) denote the stabilizer of \( M_i \) in the braid group \( B_{n_i} \), which we identify with the mapping class group of \( \Delta_i \).

The kernel of \( \Pi \) is isomorphic to the direct product

\[
\text{Stab}_{B_{n_1}}(M_1) \times \cdots \times \text{Stab}_{B_{n_q}}(M_q).
\]

The map \( \Pi \) is split. Indeed, Bell and the second author [6] constructed a map from the braid group \( B_\ell \) to the mapping class group of the surface obtained from \( D_\ell \) by blowing up the \( \ell \) marked points to boundary components (this is essentially the map \( \iota : L_k \to \text{Mod}(S_k) \) from Figure 7 in their paper). By including this surface into \( D_n \), we obtain the desired splitting. We thus have the following lemma, which summarizes the above discussion.

Lemma 3.1. Let \( M \) be a multicurve in \( D_n \). Let \( B_{\Delta_0}, \Pi, \text{Stab}_{B_{n_1}}(M_1), \ldots, \text{Stab}_{B_{n_q}}(M_q) \) be defined as above. The map \( \Pi \) induces a semi-direct product decomposition

\[
\text{Stab}_{B_n}(M) \cong B_{\Delta_0} \rtimes \left( \text{Stab}_{B_{n_1}}(M_1) \times \cdots \times \text{Stab}_{B_{n_q}}(M_q) \right)
\].
Lemma 3.1 can be iteratively applied in order to decompose each of the Stab_{B_n}(M_j), giving an iterated semi-direct product decomposition of Stab_{B_n}(M). In the final decomposition, each factor is a subgroup of some braid group, specifically, a subgroup preserving some partition of the strands.

Roots of powers of differences of half-twists. Our proof of Theorem 1.1 also uses the following.

**Lemma 3.2.** Let n ≥ 1, let c and d be disjoint curves in D_n that surround exactly 2 marked points each, and suppose the braid (H_c H_d)^{-1} has a pth root f. Then ℓ is divisible by p and

$$f = (H_c H_d)^{-1}^{\ell/p}.$$ 

**Proof.** Since canonical reduction systems are invariant under taking powers, Γ(f) = {c, d}. In particular f acts on the set {c, d}. Since an element of a group commutes with its powers, f commutes with (H_c H_d)^{-1}. Since the signs on H_c and H_d differ, it follows that f acts trivially on {c, d}. By collapsing the disks bounded by c and d to marked points, we obtain a disk with n - 2 marked points and f induces a mapping class ı of this disk, hence an element of B_{n-2}. Since f^p = (H_c H_d)^{-1}^{\ell}, it must be that f^p is the identity. Since B_{n-2} is torsion free, ı is trivial. Thus f = H_c^{r_1} H_d^{r_2} for some r_1 and r_2. The lemma follows. □

**Proof of Theorem 1.1.** As in the statement, assume n ≥ 7 and let ρ : B_n' → B_n be a nontrivial homomorphism. Let Z_n be the totally symmetric subset of B_n' defined in Section 2, and let M denote the labeled multicurve Γ(ρ(Z_n)). For 1 ≤ i ≤ [n/2] let x_i be the element

$$x_i = ρ(σ_i^n z^{-1}),$$

so that ρ(Z_n) is the set of all x_i (they may not be all distinct).

By Lemma 2.1, the cardinality of ρ(Z_n) is either 1 or [n/2]. We thus have four cases:

1. ρ(Z_n) is a singleton,
2. M is empty,
3. M is non-empty and is trivially labeled, and
4. M is non-empty and is not trivially labeled.

In the first two cases we will show that ρ is trivial, in the third case we will derive a contradiction, and in the last case we will show that ρ is equivalent to the inclusion map.

**Case 1.** If ρ(Z_n) is a singleton then x_1 = x_3 and so

$$ρ(σ_1 σ_3^{-1})^p = ρ((σ_1 σ_3^{-1})^p) = ρ((σ_1^{-1} z^{-1})(σ_3^p z^{-1})) = x_1 x_3^{-1} = 1.$$ 

Since B_n is torsion-free, we therefore have that ρ(σ_1 σ_3^{-1}) = 1, and since the normal closure of σ_1 σ_3^{-1} in B_n' is equal to B_n' for n ≥ 5 (this follows from [19, Remark 1.10] and the fact that there are elements of B_n of arbitrary word length that commute with any given σ_i σ_j^{-1}, namely, σ_i^j) it follows that ρ is trivial.

**Case 2.** In this case we will prove that ρ is trivial. To say that M is empty is to say that the x_i are periodic or they are pseudo-Anosov.

Assume first that the x_i are periodic. Since the image of ρ is contained in B_n' and since the only periodic element of B_n' is the identity, we must have that x_i = 1 for all i. This implies that ρ(Z_n) is a singleton, in which case we can apply Case 1 in order to conclude that ρ is trivial.
Now assume that the $x_i$ are all pseudo-Anosov. Let $\bar{B}_n$ denote the quotient $B_n/Z(B_n)$ and let $\bar{x}_i$ denote the image of $x_i$ in this quotient. Since the $x_i$ are pseudo-Anosov, so too are the $\bar{x}_i$. By a theorem of McCarthy [8], there is a short exact sequence
\[ 1 \to F \to C_{\bar{B}_n}(\bar{x}_1) \to \mathbb{Z} \to 1, \]
where $C_{\bar{B}_n}(\bar{x}_1)$ is the centralizer of $x_1$ in $\bar{B}_n$ and $F$ is a finite subgroup of $\bar{B}_n$. Since $\bar{x}_3$ commutes with $\bar{x}_1$ and is conjugate, it follows that $(\bar{x}_1\bar{x}_3^{-1})^p$ lies in $F$ for some $p$. In particular $\bar{x}_1\bar{x}_3^{-1}$, hence $x_1x_3^{-1}$, is periodic. We again use the fact that $B_n'$ contains no nontrivial periodic elements to conclude that $x_1x_3^{-1} = 1$. By the fundamental lemma of totally symmetric sets (Lemma 2.1), the set $\rho(Z_n)$ is a singleton and we may again apply Case 1 in order to conclude that $\rho$ is trivial, contradicting the assumption that the $x_i$ are pseudo-Anosov.

**Case 3.** The basic strategy is to show that $\rho(B_n')$ lies in $\text{Stab}_{\bar{B}_n}(M)$, to decompose the latter using Lemma 3.1, and then to inductively apply the argument of Case 2 to the resulting factors in order to derive a contradiction.

For $n \geq 5$, the group $B_n'$ is generated by the elements of the form $\sigma_1\sigma_i^{-1}$ where $2 \leq i \leq n - 1$; see [19, p. 7] or [16, Prop. 3.1]. For each such generator $\sigma_1\sigma_i^{-1}$, there exists a $\sigma_j^p z^{-1} \in Z_n$ that commutes with it. This implies that each $\rho(\sigma_1\sigma_i^{-1})$ preserves $\Gamma(\rho(\sigma_j^p z^{-1}))$. Since $M$ is trivially labeled, the latter is equal to $M$. Thus $\rho(B_n')$ lies in $\text{Stab}_{\bar{B}_n}(M)$.

Let $\Pi$ be the map from Lemma 3.1. The composition $\Pi \circ \rho : B_n' \to B_{\Delta_0}$ satisfies the hypothesis of Case 2, in that $\Gamma(\Pi \circ \rho(Z_n))$ is empty. By the argument of Case 2, $\Pi \circ \rho$ is trivial (we cannot apply Case 2 verbatim since $\Delta_0$ has fewer than $n$ marked points). Thus, the image of $\rho$ lies in the second factor of the decomposition of $\text{Stab}_{\bar{B}_n}(M)$ given by Lemma 3.1. We may now iterate the argument on each factor. After finitely many steps, we conclude that $\rho$ is trivial, a contradiction.

**Case 4.** In this case we will prove that $\rho$ is equivalent to the inclusion map. The proof has five steps. In the fourth step, we say that a sequence of curves $d_1, \ldots, d_k$ in $D_n$ forms a chain if each $d_i$ surrounds two marked points, if $i(d_i, d_{i+1}) = 2$ for $1 \leq i \leq k - 1$, and if the $d_i$ are disjoint otherwise. Also, let $a_1, \ldots, a_{n-1}$ be the curves in $D_n$ with the property that $\sigma_i = H_{a_i}$ for $1 \leq i \leq n - 1$. Note that $c_i = a_{2i-1}$ for $1 \leq i \leq [n/2]$ and that $a_1, \ldots, a_{n-1}$ form a chain.

**Step 1.** Up to equivalence, we have $\rho(\sigma_1\sigma_j^{-1}) = \sigma_1\sigma_j^{-1}$ for all odd $j$.

**Step 2.** For each even $i \geq 6$ there exists a curve $b_i$ such that $\rho(\sigma_1\sigma_i^{-1}) = \sigma_1 H_{b_i}^{-1}$.

**Step 3.** For $i \in \{2, 4\}$ there exists a curve $b_i$ such that $\rho(\sigma_1\sigma_i^{-1}) = \sigma_1 H_{b_i}^{-1}$.

**Step 4.** The curves $a_1, b_2, a_3, b_4, a_5, b_6, \ldots$ form a chain.

**Step 5.** The homomorphism $\rho$ is equivalent to the inclusion map.

We complete the five steps in turn. Step 3 is the only part of the proof that uses the assumption $n \geq 7$; Step 4 uses the assumption $n \geq 6$, and the rest of the proof only uses the assumption $n \geq 5$.

**Step 1.** Up to equivalence, we have $\rho(\sigma_1\sigma_j^{-1}) = \sigma_1\sigma_j^{-1}$ for all odd $j$.

When $n$ is even, Lemma 2.6 implies there is a nonzero $\ell$ so that $\rho(Z_n)$ is $B_n$-equivalent to one of the following totally symmetric sets:
\[ (X_n^0, \ell) \tau_0 z \quad \text{or} \quad (X_n^0, \ell) T_\alpha z. \]
When $n$ is odd, Lemma 2.6 implies that there is a nonzero $\ell$ so that $\rho(Z_n)$ is $B_n$-equivalent to one of the following totally symmetric sets:

$$(X_n^{0,\ell})^{z^s} (X_n^{\ell,0})^{z^s} (X_n^{0,\ell})^{T_{c_0} z^s} \text{ or } (X_n^{\ell,0})^{T_{c_0} z^s}.$$ 

Therefore, up to replacing $\rho$ with $g$ and such that $\rho(1) = 1$, let $b = (X_n^{\ell,0})^{T_{c_0} z^s}$. By Lemma 3.2, $(\sigma_1 \sigma_j^{-1})^\ell$ has a $p$th root if and only if $p$ divides $\ell$ and in this case there is a unique root, namely, $(\sigma_1 \sigma_j^{-1})^{\ell/p}$. Setting $q = \pm \ell/p$ then gives the claim.

We next claim that $q = \pm 1$. Let $g$ be an element of $B'_n$ such that

$$g \sigma_1 g^{-1} = \sigma_2$$

and such that $g$ commutes with $\sigma_5$ (for example $g = \sigma_2^{-2} \sigma_1 \sigma_2$). Then $g$ conjugates $\sigma_1 \sigma_5^{-1}$ to $\sigma_2 \sigma_5^{-1}$. Define $b$ to be the curve with

$$H_b = \rho(g) \sigma_1 \rho(g)^{-1}.$$ 

We then have that

$$\rho(\sigma_2 \sigma_5^{-1}) = \rho(g(\sigma_1 \sigma_5^{-1}) g^{-1}) = \rho(g) \rho(\sigma_1 \sigma_5^{-1}) \rho(g)^{-1} = \rho(g) \rho(\sigma_1 \sigma_5^{-1})^\ell \rho(g)^{-1} = H_b^q \sigma_5^{-q}.$$ 

The element $\sigma_1 \sigma_5^{-1}$ satisfies a braid relation with $\sigma_2 \sigma_5^{-1}$, and so $\rho(\sigma_1 \sigma_5^{-1})$ satisfies a braid relation with $\rho(\sigma_2 \sigma_5^{-1})$. It follows that $(\sigma_1 \sigma_5^{-1})^q$ satisfies a braid relation with $H_b^q \sigma_5^{-q}$. Since $\sigma_5$ commutes with both $\sigma_1$ and $H_b$, we further have that $\sigma_1 \sigma_5^{-1}$ satisfies a braid relation with $H_b^q$. A result of Bell and second author [5, Lemma 4.9] states that if two half-twists $H_a'$ and $H_b'$ satisfy a braid relation, then $r = s = \pm 1$. The claim follows.

If $q = 1$, then $\rho(\sigma_1 \sigma_j^{-1}) = \sigma_1 \sigma_j^{-1}$ for $j$ odd, as desired. If $q = -1$, then we may further postcompose $\rho$ with the inversion automorphism of $B_n$ to obtain again that $\rho(\sigma_1 \sigma_j^{-1}) = \sigma_1 \sigma_j^{-1}$. This completes the first step.

**Step 2.** For each even $i \geq 6$ there exists a curve $b_i$ such that $\rho(\sigma_i \sigma_1^{-1}) = \sigma_1 H_{b_i}^{-1}$.

Fix some $i \geq 5$. There exists $g_i \in B'_n$ such that

$$g_i \sigma_5 g_i^{-1} = \sigma_i$$

and such that $g_i$ commutes with each of $\sigma_1$ and $\sigma_3$; for instance we may take

$$g_i = \sigma_i^{9-2i} (\sigma_{i-1} \cdots \sigma_5)(\sigma_i \cdots \sigma_3).$$

Let $b_i$ be the curve such that

$$H_{b_i} = \rho(g_i) \sigma_5 \rho(g_i)^{-1}.$$ 

Since $g_i$ commutes with $\sigma_1$ and $\sigma_3$, and hence with $\sigma_3 \sigma_1^{-1}$, it follows that $\rho(g_i)$ commutes with $\rho(\sigma_1 \sigma_3^{-1}) = \sigma_1 \sigma_3^{-1}$. It follows further that $\rho(g_i)$ commutes with each of $\sigma_1$ and $\sigma_3$ (it
cannot be that \( \rho(g_i) \) interchanges \( a_1 \) and \( a_3 \) because the signs of the half-twists differ). Hence \( H_{b_i} \) commutes with \( \sigma_1 \).

Using the above properties of \( g_i \) and the fact (from Step 1) that \( \rho(\sigma_1 \sigma_5^{-1}) = \sigma_1 \sigma_5^{-1} \), we have

\[
\rho(\sigma_1 \sigma_i^{-1}) = \rho(g_i(\sigma_1 \sigma_5^{-1})g_i^{-1}) = \rho(g_i)\rho(\sigma_1 \sigma_5^{-1})\rho(g_i)^{-1} = \rho(g_i)\sigma_1 \sigma_5^{-1} \rho(g_i)^{-1} = \sigma_1 H_{b_i}^{-1}.
\]

This completes the second step.

**Step 3.** For \( i \in \{2, 4\} \) there exists a curve \( b_i \) such that \( \rho(\sigma_1 \sigma_i^{-1}) = \sigma_1 H_{b_i}^{-1} \).

First we treat the case \( i = 4 \). Choose \( g_4 \in B_n \) such that \( g_4 \sigma_3 g_4^{-1} = \sigma_4 \) and such that \( g_4 \) commutes with \( \sigma_1 \) and each \( \sigma_i \) with \( i \geq 6 \) (for instance \( g_4 = \sigma_4^{-2} \sigma_3 \sigma_4 \)). The second condition implies that \( g_4 \) commutes with \( \sigma_1 \sigma_6^{-1} \) and hence that \( \rho(g_4) \) commutes with \( \rho(\sigma_1 \sigma_6^{-1}) = \sigma_1 H_{b_4}^{-1} \). Equivalently, \( \rho(g_4) \) commutes with \( \sigma_1 \) and \( H_{b_4} \). Define \( b_4 \) to be the curve such that

\[
H_{b_4} = \rho(g_4)\sigma_3 \rho(g_4)^{-1}.
\]

We then have that

\[
\rho(\sigma_1 \sigma_4^{-1}) = \rho(g_4(\sigma_1 \sigma_3^{-1})g_4^{-1}) = \rho(g_4)\sigma_3 \rho(g_4)^{-1} = \sigma_1 H_{b_4}^{-1}.
\]

This completes the \( i = 4 \) case.

We now address the \( i = 2 \) case; this is similar to the \( i = 4 \) case, but more complicated. Choose \( g_2 \in B_n' \) such that \( g_2 \sigma_1 g_2^{-1} = \sigma_2 \) and such that \( g_2 \) commutes with each \( \sigma_i \) with \( i \geq 4 \). The second condition implies that \( g_2 \) commutes with \( \sigma_5 \sigma_6^{-1} \) and hence that \( \rho(g_2) \) commutes with \( \rho(\sigma_5 \sigma_6^{-1}) \). The latter is equal to \( \sigma_5 H_{b_6}^{-1} \); indeed,

\[
\rho(\sigma_5 \sigma_6^{-1}) = \rho((\sigma_1 \sigma_5^{-1})^{-1}(\sigma_1 \sigma_6^{-1})) = (\sigma_1 \sigma_5^{-1})^{-1} \sigma_1 H_{b_6}^{-1} = \sigma_5 H_{b_6}^{-1}.
\]

Thus, \( \rho(g_2) \) commutes with \( \sigma_5 H_{b_6}^{-1} \).

Next, the element \( g_2 \) commutes with \( \sigma_4 \sigma_6^{-1} \) so \( \rho(g_2) \) commutes with \( \rho(\sigma_4 \sigma_6^{-1}) \). We have

\[
\rho(\sigma_4 \sigma_6^{-1}) = \rho((\sigma_1 \sigma_4^{-1})^{-1}(\sigma_1 \sigma_6^{-1})) = (\sigma_1 H_{b_4}^{-1})^{-1}(\sigma_1 H_{b_6}^{-1}) = H_{b_4} H_{b_6}^{-1}.
\]

Thus \( \rho(g_2) \) also commutes with \( H_{b_4} H_{b_6}^{-1} \).

We next show that \( i(b_4, b_6) = 0 \). Since \( \sigma_4 \sigma_6^{-1} \) is conjugate to \( \sigma_1 \sigma_3^{-1} \) in \( B_n' \), it follows that \( \rho(\sigma_4 \sigma_6^{-1}) \) is conjugate to \( \rho(\sigma_1 \sigma_3^{-1}) \). By Step 1, the latter is equal to \( \sigma_3 \sigma_1^{-1} \). In particular, \( \rho(\sigma_4 \sigma_6^{-1}) \) is equal to a difference of two commuting half-twists. It follows from the Thurston construction of pseudo-Anosov mapping classes [15, Theorem 14.1] that if \( i(b_4, b_6) \) were nonzero then \( H_{b_4} H_{b_6}^{-1} \) would be a partial pseudo-Anosov mapping class and hence would not be conjugate to \( \sigma_1 \sigma_3^{-1} \). (Strictly speaking the Thurston construction applies to Dehn twists, not half-twists, but this can be remedied by using the standard embedding of the braid group to the mapping class group [15, Section 9.4] of the hyperelliptic double cover of \( D_n \), which maps half-twists to Dehn twists.) We thus have \( i(b_4, b_6) = 0 \), as desired.

By the previous two paragraphs we have that \( \rho(g_2) \) commutes with \( H_{b_4} H_{b_6}^{-1} \) and that the latter is the difference of two commuting half-twists. As in Step 2, we can use the fact that the signs on \( H_{b_4} \) and \( H_{b_6} \) differ in the product \( H_{b_4} H_{b_6}^{-1} \) to conclude that \( \rho(g_2) \) commutes with \( H_{b_6} \).

We have shown that \( \rho(g_2) \) commutes with \( \sigma_5 H_{b_6}^{-1} \) and \( H_{b_6} \). It follows that \( \rho(g_2) \) commutes with \( \sigma_5 \).
Define $b_2$ to be the curve such that
\[ H_{b_2} = \rho(g_2)\sigma_1\rho(g_2)^{-1} \]
We then have that
\[ \rho(\sigma_5\sigma_5^{-1}) = \rho(g_2(\sigma_1\sigma_5^{-1})g_2^{-1}) = \rho(g_2)(\sigma_1\sigma_5^{-1})\rho(g_2)^{-1} = \rho(g_2)\sigma_1\rho(g_2)^{-1}\sigma_5^{-1} = H_{b_2}\sigma_5^{-1}. \]
Now
\[ \rho(\sigma_1\sigma_2^{-1}) = \rho((\sigma_1\sigma_5^{-1})(\sigma_2\sigma_5^{-1})^{-1}) = (\sigma_1\sigma_5^{-1})(H_{b_2}\sigma_5^{-1})^{-1} = \sigma_1H_{b_2}^{-1}, \]
as desired.

**Step 4.** The curves $a_1, b_2, a_3, b_4, a_5, b_6, \ldots$ form a chain.

For $i$ odd let $b_i$ be the standard curve $a_i$; so we need to show that $b_1, \ldots, b_{n-1}$ form a chain. It follows from the definition of the $b_i$ and the fact that each $\sigma_i\sigma_j^{-1}$ is conjugate in $B'_n$ to $\sigma_1\sigma_3^{-1}$ that each $b_i$ surrounds exactly two marked points. To complete this step we must show that $i(b_i, b_j) = 0$ if $j - i \geq 2$ and that $i(b_i, b_{i+1}) = 2$ for each $1 \leq i \leq n - 2$.

We begin by showing that $i(b_i, b_j) = 0$ if $j - i \geq 2$. In this case we have
\[ \rho(\sigma_i\sigma_j^{-1}) = \rho(\sigma_i\sigma_j^{-1})(\sigma_1\sigma_j^{-1}) = H_{b_i}H_{b_j}^{-1}. \]
Since $\sigma_i\sigma_j^{-1}$ is conjugate in $B'_n$ to $\sigma_1\sigma_3^{-1}$ and since $\rho$ fixes the latter, it follows that $H_{b_i}H_{b_j}^{-1}$ is conjugate to $\sigma_1\sigma_3^{-1}$. As in Step 3, it follows from the Thurston construction that $b_i$ and $b_j$ are disjoint, as desired.

We now proceed to show that $i(b_i, b_{i+1}) = 2$ for each $1 \leq i \leq n - 2$. Here, it suffices to show that $H_{b_i}$ and $H_{b_{i+1}}$ satisfy the braid relation. We already showed in Step 1 that $\sigma_1$ satisfies a braid relation with $H_{b_2}$ and so the $i = 1$ case is settled. It remains to treat the cases $i \geq 3$ and $i = 2$.

First, fix some $i \geq 3$. Since $\sigma_i$ satisfies a braid relation with $\sigma_{i+1}$, the element $\sigma_1\sigma_i^{-1}$ satisfies a braid relation with $\sigma_1\sigma_{i+1}^{-1}$. It follows that $\rho(\sigma_1\sigma_i^{-1}) = \sigma_1H_{b_i}^{-1}$ satisfies a braid relation with $\rho(\sigma_1\sigma_{i+1}^{-1}) = \sigma_1H_{b_{i+1}}^{-1}$. Since both $H_{b_i}$ and $H_{b_{i+1}}$ commute with $\sigma_1$, this implies that $H_{b_i}$ satisfies a braid relation with $H_{b_{i+1}}$, as desired.

Finally, we show that $H_{b_2}$ and $H_{b_3}$ satisfy the braid relation. Similar to the previous paragraph, $\rho(\sigma_2\sigma_5^{-1})$ satisfies a braid relation with $\rho(\sigma_3\sigma_5^{-1})$. In Step 3 we showed that $\rho(\sigma_2\sigma_5^{-1}) = H_{b_2}\sigma_5^{-1}$. Since $n \geq 6$ we also have
\[ \rho(\sigma_3\sigma_5^{-1}) = \rho((\sigma_1\sigma_3^{-1})(\sigma_5\sigma_5^{-1})) = (\sigma_1\sigma_3^{-1})(\sigma_5\sigma_5^{-1}) = \sigma_3\sigma_5^{-1}. \]
It follows that $H_{b_2}$ and $\sigma_3 = H_{b_3}$ satisfy the braid relation. The completes the fourth step.

**Step 5.** The homomorphism $\rho$ is equivalent to the inclusion map.

Since the curves $b_1, b_2, \ldots, b_{n-1}$ from Step 4 form a chain, there is an element $\alpha$ of $B_n$ such that the curves $\alpha(b_1), \ldots, \alpha(b_{n-1})$ are equal to $a_1, \ldots, a_{n-1}$, respectively (this is an instance of the change of coordinates principle for mapping class groups [15, Section 1.3]).

After replacing $\rho$ by its post-composition with the inner automorphism of $B_n$ induced by $\alpha$, we have that
\[ \rho(\sigma_1\sigma_i^{-1}) = \sigma_1\sigma_i^{-1} \]
Since (as in Case 3 above) the elements $\sigma_1\sigma_i^{-1}$ generate $B'_n$, it follows that $\rho$ is equal to the standard inclusion. This completes Step 5, and the theorem is proven. \(\square\)
4. Homomorphisms between braid groups

In this section we classify homomorphisms $B_n \to B_n$ for $n \geq 7$. As discussed in the introduction, this is a special case of a theorem of Castel.

Let $\rho : B_n \to B_n$ be a homomorphism, and let $k$ be an integer. The transvection of $\rho$ by $z^k$ is the homomorphism given by

$$\rho^z(k)(\sigma_i) = \rho(\sigma_i)z^k$$

for all $1 \leq i \leq n - 1$. There is an equivalence relation on the set of homomorphisms $B_n \to B_n$ whereby $\rho_1 \sim \rho_2$ if $\rho_2 = \alpha \circ \rho_1^z$ for some automorphism $\alpha$ of $B_n$ and some $k \in \mathbb{Z}$. This notion of equivalence is more complicated than the one we defined for homomorphisms $B'_n \to B_n$ in the introduction, in that it involves the transvections. There are no analogous transvections of homomorphisms $B'_n \to B_n$, since the image of $B'_n$ must lie in $B'_n$, and the only power of $z^k$ in $B'_n$ is the identity.

The following theorem represents the special case of Castel’s theorem that we will prove.

**Theorem 4.1** (Castel). Let $n \geq 7$, and let $\rho : B_n \to B_n$ be a homomorphism whose image is not cyclic. Then $\rho$ is equivalent to the identity.

**Proof.** Assume that $\rho : B_n \to B_n$ is a homomorphism with non-cyclic image. This is equivalent to the assumption that restriction $\rho'$ of $\rho$ to $B'_n$ is nontrivial. Theorem 1.1 then implies that there is an automorphism $\alpha$ of $B_n$ such that $\alpha \circ \rho'$ is the identity. Thus replacing $\rho$ by $\alpha \circ \rho$, we may assume that $\rho'$ is the inclusion map.

We claim that $\rho(z)$ is equal to $z^k$ for some integer $k$. As in Section 3 let $p = n(n - 1)$. Since $z \in B_n$ is central, we have that $\rho(z)$ commutes with each $\rho(\sigma_i^pz^{-1})$. Since

$$\rho(\sigma_i^pz^{-1}) = \rho'(\sigma_i^pz^{-1}) = \sigma_i^pz^{-1},$$

it follows that $\rho(z)$ commutes with each $\sigma_i^p$, hence with each $\sigma_i$. The claim follows.

We next claim that $\rho(\sigma_i)^p = \sigma_i^pz^{k-1}$ for each $i$. By the previous claim we indeed have

$$\sigma_i^pz^{-1} = \rho'(\sigma_i^pz^{-1}) = \rho(\sigma_i^p)\rho(z)^{-1} = \rho(\sigma_i^p)z^{-k} = \rho(\sigma_i)^pz^{-k},$$

whence the claim.

We now claim that there exist integers $r$ and $s$ such that for all $1 \leq i \leq n - 1$ we have

$$\rho(\sigma_i) = \sigma_i^rz^s.$$ 

By the previous claim, the canonical reduction system of $\rho(\sigma_i)$ is equal to $a_i$. Since $\sigma_i$ commutes with $\sigma_j$ for $j \geq 3$, we have that $\rho(\sigma_1)$ fixes each $a_j$ with $j \geq 3$. Thus the Nielsen–Thurston component of $\rho(\sigma_1)$ corresponding to the region between $a_1$ and the boundary of $D_n$ cannot be pseudo-Anosov or a nontrivial periodic element. It follows that $\rho(\sigma_1) = \sigma_1^rz^s$ for some $r$ and $s$. Since $\rho(\sigma_1)$ is conjugate to $\rho(\sigma_1)$ for $1 \leq i \leq n - 1$, the claim follows.

Next we claim that $r = 1$. We have

$$\sigma_1\sigma_3^{-1} = \rho(\sigma_1\sigma_3^{-1}) = \sigma_1^rz^s\sigma_3^{-r}z^{-s} = \sigma_1^r\sigma_3^{-r},$$

whence the claim.

We now have that $\rho(\sigma_i) = \sigma_i^rz^s$ for $1 \leq i \leq n - 1$. The transvection of $\rho$ by $z^{-s}$ is then equal to the identity. This completes the proof of the theorem. \qed
References


Kevin Kordek, School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, GA 30332
Email address: kevin.a.kordek@gmail.com

Dan Margalit, School of Mathematics, Georgia Institute of Technology, 686 Cherry St., Atlanta, GA 30332
Email address: margalit@math.gatech.edu