

# ON FINITENESS PROPERTIES OF THE JOHNSON FILTRATIONS

MIKHAIL ERSHOV AND SUE HE

**ABSTRACT.** Let  $\Gamma$  be either the automorphism group of the free group of rank  $n \geq 4$  or the mapping class group of an orientable surface of genus  $n \geq 12$  with at most 1 boundary component, and let  $G$  be either the subgroup of IA-automorphisms or the Torelli subgroup of  $\Gamma$ , respectively. For  $N \in \mathbb{N}$  denote by  $\gamma_N G$  the  $N^{\text{th}}$  term of the lower central series of  $G$ . We prove that

- (i) any subgroup of  $G$  containing  $\gamma_2 G = [G, G]$  (in particular, the Johnson kernel in the mapping class group case) is finitely generated;
- (ii) if  $N = 2$  or  $n \geq 8N - 4$  and  $K$  is any subgroup of  $G$  containing  $\gamma_N G$  (for instance,  $K$  can be the  $N^{\text{th}}$  term of the Johnson filtration of  $G$ ), then  $G/[K, K]$  is nilpotent and hence the abelianization of  $K$  is finitely generated;
- (iii) if  $H$  is any finite index subgroup of  $\Gamma$  containing  $\gamma_N G$ , with  $N$  as in (ii), then  $H$  has finite abelianization.

## 1. INTRODUCTION

**1.1. The Johnson filtrations.** Let  $F_n$  denote the free group of rank  $n \geq 2$ , and let  $\text{IA}_n \subset \text{Aut}(F_n)$  denote the subgroup of automorphisms of  $F_n$  which act as identity on the abelianization; equivalently,  $\text{IA}_n$  is the kernel of the natural map  $\text{Aut}(F_n) \rightarrow \text{Aut}(F_n^{\text{ab}}) \cong \text{GL}_n(\mathbb{Z})$ . More generally, for each  $k \in \mathbb{N}$  define  $\text{IA}_n(k)$  to be the kernel of the natural map  $\text{Aut}(F_n) \rightarrow \text{Aut}(F_n/\gamma_{k+1}F_n)$ . The filtration  $\text{IA}_n = \text{IA}_n(1) \supset \text{IA}_n(2) \supset \dots$  was first introduced and studied by Andreadakis in [An]. It is easy to see that  $\cap_{k=1}^{\infty} \text{IA}_n(k) = \{1\}$  and  $\gamma_k \text{IA}_n \subseteq \text{IA}_n(k)$  for each  $k$  (as usual, for a group  $G$  and  $k \in \mathbb{N}$  we denote by  $\gamma_k G$  the  $k^{\text{th}}$  term of the lower central series of  $G$  defined inductively by  $\gamma_1 G = G$  and  $\gamma_{k+1} G = [\gamma_k G, G]$  for  $k \geq 1$ ). Bachmuth [Ba] proved that  $\gamma_2 \text{IA}_n = \text{IA}_n(2)$ .

The filtration  $\{\text{IA}_n(k)\}_{k \in \mathbb{N}}$  is often referred to as the *Johnson filtration* and owes its name to the corresponding filtration in mapping class groups whose study was initiated by Johnson [Jo4]. Let  $\Sigma_g^1$  be an orientable surface of genus  $g \geq 2$  with 1 boundary component and  $\text{Mod}_g^1$  its mapping class group. The fundamental group  $\pi = \pi(\Sigma_g^1)$  is free of rank  $2g$ , and for each  $k \in \mathbb{N}$  there is a natural homomorphism  $\text{Mod}_g^1 \rightarrow \text{Aut}(\pi/\gamma_{k+1}\pi)$ . Denote the kernel of this homomorphism by  $\mathcal{I}_g^1(k)$ . The subgroups  $\mathcal{I}_g^1 = \mathcal{I}_g^1(1)$  and  $\mathcal{K}_g^1 = \mathcal{I}_g^1(2)$  are known as the *Torelli subgroup* and the *Johnson kernel*, respectively, and the filtration  $\{\mathcal{I}_g^1(k)\}_{k \in \mathbb{N}}$  is called the *Johnson filtration* of  $\text{Mod}_g^1$ . Again one has  $\cap_{k=1}^{\infty} \mathcal{I}_g^1(k) = \{1\}$  and  $\gamma_k \mathcal{I}_g^1 \subseteq \mathcal{I}_g^1(k)$  for each  $k$ , but this time the inclusion is known to be strict already for  $k = 2$ . The mapping class group  $\text{Mod}_g$  of a closed orientable surface of genus  $g$  is a quotient of  $\text{Mod}_g^1$ , and the Johnson filtration  $\{\mathcal{I}_g(k)\}_{k \in \mathbb{N}}$  of  $\text{Mod}_g$  can be defined as the image of the filtration  $\{\mathcal{I}_g^1(k)\}_{k \in \mathbb{N}}$  of  $\text{Mod}_g^1$ .

**1.2. Finite generation.** The first result on finiteness properties of the Johnson filtrations is a classical theorem of Magnus [Ma] from 1935 which asserts that  $\text{IA}_n = \text{IA}_n(1)$  is finitely

---

The work of the first author was partially supported by the NSF grant DMS-1201452.

generated for all  $n \geq 2$ . In his celebrated 1983 paper, Johnson [Jo5] proved that the Torelli group  $\mathcal{I}_g^1$  (and hence also its quotient  $\mathcal{I}_g$ ) is finitely generated for  $g \geq 3$ . In 1986, McCullough and Miller [MM] proved that  $\mathcal{I}_2$  (and hence also  $\mathcal{I}_2^1$ ) is infinitely generated, thereby completely settling the finite generation question for the first terms.

The question whether the Johnson kernel  $\mathcal{I}_g(2)$  is finitely generated for  $g \geq 3$  has received a lot of attention since 1990s and was explicitly asked by Morita [Mo, Problem 2.2(i)]. For some time it was generally expected that the Johnson kernel  $\mathcal{I}_g(2)$  as well as its counterpart  $\text{IA}_n(2)$  are infinitely generated and moreover have infinite-dimensional first rational homology. The latter was refuted by the recent breakthrough results of Dimca and Papadima [DP] and Papadima and Suciu [PS2] who established that  $H_1(\mathcal{I}_g(2), \mathbb{Q})$  for  $g \geq 4$  and  $H_1(\text{IA}_n(2), \mathbb{Q})$  for  $n \geq 5$  are finite-dimensional. Our first main theorem strengthens these results by showing that the second terms of the Johnson filtrations are finitely generated in sufficiently large rank.

**Theorem 1.1.** *Let  $G$  be either  $\text{IA}_n$  for  $n \geq 4$  or  $\mathcal{I}_g^1$  for  $g \geq 12$ . If  $K$  is any subgroup of  $G$  containing  $[G, G]$ , then  $K$  is finitely generated.*

**Remark.** Theorem 1.1 is also true for  $G = \mathcal{I}_g$  since this group is a quotient of  $\mathcal{I}_g^1$ .

We will prove Theorem 1.1 by computing the *BNS invariant* of the group  $G$  in the theorem. Given an arbitrary finitely generated group  $G$ , by a *character* of  $G$  we will mean a homomorphism  $\chi : G \rightarrow \mathbb{R}$ . Two characters  $\chi$  and  $\chi'$  will be considered equivalent if they are positive scalar multiples of each other. The equivalence class of a character  $\chi$  will be denoted by  $[\chi]$ , and the set of equivalence classes of nonzero characters will be denoted by  $S(G)$ . In [BNS], Bieri, Neumann and Strebel introduced a certain subset  $\Sigma(G)$  of  $S(G)$ , now known as the *BNS invariant* of  $G$ . The BNS invariant of  $G$  completely determines which subgroups of  $G$  containing  $[G, G]$  are finitely generated:

**Theorem 1.2** ([BNS]). *Let  $G$  be a finitely generated group and let  $N$  be a subgroup of  $G$  containing  $[G, G]$ . Then  $N$  is finitely generated if and only if  $\Sigma(G)$  contains  $[\chi]$  for every non-trivial  $\chi$  which vanishes on  $N$ . In particular,  $[G, G]$  is finitely generated if and only if  $\Sigma(G) = S(G)$ .*

Thus, Theorem 1.1 is a direct consequence of Theorem 1.2 and the following result:

**Theorem 1.3.** *Let  $G$  be as in Theorem 1.1. Then  $\Sigma(G) = S(G)$ .*

The set  $\Sigma(G)$  admits many different characterizations. In order to prove Theorem 1.3 we will use the characterization in terms of actions on real trees due to Brown [Br1] (see § 2).

**1.3. Vanishing of cohomology and abelianization of the Johnson filtrations.** Our second main theorem proves the vanishing of the first cohomology of  $\text{IA}_n$  and  $\mathcal{I}_g^1$  with coefficients in a certain class of irreducible representations.

**Theorem 1.4.** *Let  $N \geq 2$  be an integer and  $G$  a group such that one of the following holds:*

- (a)  $N = 2$  and  $G = \text{IA}_n$  where  $n \geq 4$ ;
- (b)  $G = \text{IA}_n$  where  $n \geq 8N - 4$ ;
- (c)  $G = \mathcal{I}_g^1$  where  $g \geq 8N - 4$ .

Let  $\rho$  be a non-trivial irreducible (possibly infinite-dimensional) representation of  $G$  over an arbitrary field  $F$  such that  $\text{Ker } \rho \supseteq \gamma_N G$  (equivalently,  $\rho(G)$  is nilpotent of class less than  $N$ ). Then

$$H^1(G, \rho) = 0.$$

**Remark.** Theorem 1.4 remains true if  $G$  is replaced by any of its quotients. Indeed, if  $K$  is a normal subgroup of a group  $G$  and  $\rho$  is any representation of  $G/K$ , then  $H^1(G/K, \rho)$  embeds in  $H^1(G, \rho)$ , e.g., by the inflation-restriction sequence.

Our first application of Theorem 1.4 yields new structural information about the abelianization of subgroups of  $G$  containing  $\gamma_N G$ , with  $G$  and  $N$  as in Theorem 1.4. In particular, the result applies to the  $N^{\text{th}}$  term of the Johnson filtration of  $G$ .

**Corollary 1.5.** *Let  $G$  and  $N$  be as in Theorem 1.4, and let  $K$  be any subgroup of  $G$  containing  $\gamma_N G$ . Then the following equivalent conditions hold:*

- (i)  $G/[K, K]$  is nilpotent (equivalently,  $\gamma_m G \subseteq [K, K]$  for some  $m$ )
- (ii)  $K/[K, K] \cong H_1(K, \mathbb{Z})$  considered as a  $G/K$ -module is annihilated by some power of the augmentation ideal of  $\mathbb{Z}[G/K]$ .

*In particular,  $K/[K, K]$  is finitely generated.*

**Remark.** The equivalence of conditions (i) and (ii) above will be explained in the proof of Theorem 1.6. The last assertion of Corollary 1.5 holds since a group  $G$  in Corollary 1.5 is finitely generated (hence so is  $G/[K, K]$ ) and a subgroup of a finitely generated nilpotent group is finitely generated.

When  $N = 2$ , Corollary 1.5 asserts that if  $G = \text{IA}_n$  with  $n \geq 4$  or  $G = \mathcal{I}_g^1$  with  $g \geq 12$ , then any metabelian quotient of  $G$  is nilpotent, which we find quite surprising.

Corollary 1.5 is a direct consequence of Theorem 1.4 and the following general result which will be proved in § 8:

**Theorem 1.6.** *Let  $G$  be a finitely generated group, and let  $K$  be a normal subgroup of  $G$  such that  $Q = G/K$  is nilpotent. Assume that for any finite field  $F$  of prime order and any non-trivial finite-dimensional irreducible representation  $\rho$  of  $G$  over  $F$  with  $\text{Ker } \rho \supseteq K$  we have  $H^1(G, \rho) = 0$ . Then the quotient  $G/[K, K]$  is nilpotent or, equivalently,  $K/[K, K]$  is annihilated by some power of the augmentation ideal of  $\mathbb{Z}[Q]$ .*

Theorem 1.4 and Corollary 1.5 generalize two known results (with the extra restriction  $g \geq 12$  in the mapping class group case) discussed below. Both these results deal with the case  $N = 2$  and were originally obtained by rather different methods.

If  $\rho$  is a one-dimensional representation of a group  $G$ , we always have  $\text{Ker } \rho \supseteq [G, G] = \gamma_2 G$ . Thus Theorem 1.4(a) implies the following: for any field  $F$  and any  $n \geq 4$  the set

$$\{\rho \in \text{Hom}(\text{IA}_n, F^\times) : H^1(\text{IA}_n, F_\rho) \neq 0\}$$

has just one element (the trivial homomorphism)<sup>1</sup>, where  $F_\rho$  denotes  $F$  with  $g \in \text{IA}_n$  acting on  $F$  as multiplication by  $\rho(g)$ . In the case  $F = \mathbb{C}$  this answers in the affirmative [PS2, Question 10.6(2)]. In [PS2, Thm 10.5] it was proved that the above set is finite.<sup>2</sup>

<sup>1</sup>The trivial homomorphism does belong to this set since  $\text{IA}_n$  is a finitely generated group with infinite abelianization (see § 3.1).

<sup>2</sup>The question in [PS2] is stated in terms of homology, not cohomology, but this does not matter since for any group  $G$  and any representation  $\rho$  of  $G$  over a field there are isomorphisms  $H^i(G, \rho^*) \cong H_i(G, \rho)^*$  for all  $i \in \mathbb{N}$  where  $\rho^*$  is the dual representation. This can be proved by the argument from [Br2, Prop 7.1].

Corollary 1.5 and Theorem 1.4 in the case  $G = \mathcal{I}_g^1$ ,  $N = 2$  generalize a result of Dimca, Hain and Papadima [DHP, Thm A]. The latter asserts that  $H_1(\mathcal{I}_g(2), \mathbb{Q})$  is annihilated by some power of the augmentation ideal of  $\mathbb{Q}[\mathcal{I}_g/\mathcal{I}_g(2)]$  and  $H^1(\mathcal{I}_g, \mathbb{C}_\rho) = 0$  for any non-trivial  $\rho : \mathcal{I}_g \rightarrow \mathbb{C}^\times$  which vanishes on  $\mathcal{I}_g(2)$ , in both cases for all  $g \geq 4$ . The proofs of both [DHP, Thm A] and [PS2, Thm 10.5] made an essential use of ideas from algebraic geometry introduced in [DP].

**1.4. Abelianization of finite index subgroups.** Our second application of Theorem 1.4 is motivated by the following well-known question (see, e.g., [BV, Iv, PW]):

**Question 1.7.** *Let  $\Gamma = \text{Aut}(F_n)$  or  $\text{Mod}_g^1$ . If  $H$  is a finite index subgroup of  $\Gamma$ , does  $H$  always have finite abelianization?*

The answer is known to be negative (for general  $H$ ) for  $n = 2$  (since  $\text{Aut}(F_2)$  projects onto  $\text{GL}_2(\mathbb{Z})$ , which is a virtually free group),  $n = 3$  (see [GL] and [McCo]) and  $g = 2$  (see [McCa]). For  $n \geq 4$  and  $g \geq 3$ , Question 1.7 is open and appears to be difficult.

Here is a natural counterpart of Question 1.7 for finite index subgroups of  $\text{IA}_n$  or  $\mathcal{I}_g^1$ :

**Question 1.8.** *Let  $G = \text{IA}_n$  or  $\mathcal{I}_g^1$ . If  $H$  is a finite index subgroup of  $G$ , is it true that  $\dim H^1(H, \mathbb{C}) = \dim H^1(G, \mathbb{C})$ ?*

**Remark.** If  $H$  has finite index in  $G$ , the restriction map  $H^1(G, \mathbb{C}) \rightarrow H^1(H, \mathbb{C})$  is always injective since the additive group of  $\mathbb{C}$  is torsion-free. Hence  $\dim H^1(H, \mathbb{C}) = \dim H^1(G, \mathbb{C})$  if and only if the restriction map  $H^1(G, \mathbb{C}) \rightarrow H^1(H, \mathbb{C})$  is an isomorphism.

As we will explain in § 9.1, a positive answer to Question 1.8 would yield a positive answer to Question 1.7. Even if the answer to one or both of those questions ends up being negative (we expect this is the case with Question 1.8), it is still interesting to obtain positive results for some classes of subgroups  $H$ . Theorem 1.9 below establishes the latter for subgroups containing  $\gamma_N G$  with  $G$  and  $N$  as in Theorem 1.4.

**Theorem 1.9.** *Let  $G$  and  $N$  be as in Theorem 1.4, and let  $\Gamma = \text{Aut}(F_n)$  if  $G = \text{IA}_n$  and  $\Gamma = \text{Mod}_g^1$  if  $G = \mathcal{I}_g^1$ . Let  $H$  be a subgroup of  $\Gamma$  containing  $\gamma_N G$ . The following hold:*

- (1) *If  $H$  is a finite index subgroup of  $G$ , then  $\dim H^1(H, \mathbb{C}) = \dim H^1(G, \mathbb{C})$ .*
- (2) *If  $H$  is a finite index subgroup of  $\Gamma$ , then  $H$  has finite abelianization.*

To the best of our knowledge, the only previous result dealing with Question 1.8 is a theorem of Putman [Pu3, Thm B] who used a different method to establish Theorem 1.9(1) for  $G = \mathcal{I}_g^1$ , with  $g \geq 3$ , and  $H$  a subgroup containing the Johnson kernel  $\mathcal{I}_g^1(2)$ . It is interesting to note that while we will deduce Theorem 1.9 from Theorem 1.4, an opposite implication was used in [DHP] where [Pu3, Thm B] was one of the tools in the proof of [DHP, Thm A] stated in § 1.3.

The prior work on Question 1.7 is more extensive and will be summarized in § 9.2.

**1.5. On the proofs of Theorems 1.3 and 1.4.** Let  $G$  be a finitely generated group. We will show that there is a common approach to solving the following two problems:

- (1) prove that  $H^1(G, \rho) = 0$  for a given irreducible representation  $\rho$  of  $G$  and
- (2) prove that  $[\rho] \in \Sigma(G)$  for a given character  $\rho$  of  $G$ .

In both cases the desired result follows from the existence of a  $\rho$ -centralizing generating (abbreviated as  $\rho$ -CG) sequence in  $G$ . The notion of a  $\rho$ -CG sequence, introduced in § 2, will be defined with respect to an arbitrary homomorphism  $\rho$  from  $G$  to another group.

Without giving a formal definition here, we just mention that a  $\rho$ -CG sequence exists whenever  $G$  has finitely many pairwise commuting elements  $t_1, \dots, t_k$  whose centralizers generate  $G$  and such that  $\rho(t_i)$  is non-trivial and central in  $\rho(G)$  for each  $i$  (see Lemma 2.1). Note that by our hypotheses we will always deal with  $\rho$  such that  $\rho(G)$  is nilpotent (and non-trivial) and hence always contains non-trivial central elements.

In the following two paragraphs we assume that  $G$  and  $N$  are as in Theorem 1.4 and  $\rho$  is a non-trivial homomorphism from  $G$  to another group with  $\text{Ker } \rho \supseteq \gamma_N G$  (in particular,  $\rho$  can be any non-trivial character of  $G$ ). Our goal is to show that  $G$  has a  $\rho$ -CG sequence.

The group  $\text{IA}_n$  has a very simple finite generating set constructed by Magnus already in the 1930s (see § 3 for its definition). A generating set for  $\mathcal{I}_g^1$  which is in many ways analogous to Magnus' generating set for  $\text{IA}_n$  was constructed in a recent paper of Church and Putman [CP], which made use of an earlier work of Putman [Pu2] and the original work of Johnson [Jo5] (see § 7.1). We will refer to these generating sets as *standard*. We will show that if  $\rho$  is "generic" (in a suitable sense), then the existence of a  $\rho$ -CG sequence in  $G$  follows from the basic relations between the standard generators (see Lemma 4.1). For a general  $\rho$ , we will use the observation that  $G$  is a normal subgroup of  $\Gamma$ , with  $\Gamma = \text{Aut}(F_n)$  or  $\Gamma = \text{Mod}_g^1$ , and precomposing  $\rho$  by the conjugation by a fixed element of  $\Gamma$  does not affect the existence of a  $\rho$ -CG sequence.

Thus we are reduced to showing that an arbitrary  $\rho$  as above can be conjugated to a generic one. In the case  $G = \text{IA}_n$  and  $N = 2$  this can be done quite explicitly (see § 3). In the other cases an explicit calculation may be possible, but would be cumbersome, so we take a more conceptual approach. First it is not hard to see that the relevant information about the conjugation action of  $\Gamma$  on  $G$  is captured by the induced action of  $\Gamma/G$  on  $L(G) = \bigoplus_{n=1}^{\infty} \gamma_n G / \gamma_{n+1} G$ , the Lie algebra of  $G$  with respect to the lower central series.

Second, in both cases  $(\Gamma, G) = (\text{Aut}(F_n), \text{IA}_n)$  and  $(\text{Mod}_n^1, \mathcal{I}_n^1)$ , the quotient  $\Gamma/G$  contains a natural copy of  $\text{SL}_n(\mathbb{Z})$  which therefore acts on  $L(G)$ .

In § 6 we will establish a sufficient condition on an abstract group  $G$  (assuming  $\text{SL}_n(\mathbb{Z})$  acts on  $L(G)$  as above) which guarantees that  $G$  has a  $\rho$ -CG sequence for any non-trivial homomorphism  $\rho$  from  $G$  to another group such that  $\rho(G)$  is nilpotent and has sufficiently small nilpotency class (see Theorem 6.1). One of the hypotheses in Theorem 6.1 is that the abelianization  $G^{ab} = G/[G, G]$  is a *regular*  $\text{SL}_n(\mathbb{Z})$ -module, a technical notion which will be introduced in § 5. While the proof of Theorem 6.1 will deal with the action of  $\text{SL}_n(\mathbb{Z})$  on the entire Lie algebra  $L(G)$ , its hypotheses only involve the action on  $G^{ab}$ , the degree 1 component of  $L(G)$ . This is very important since while the structure of  $G^{ab}$  as an  $\text{SL}_n(\mathbb{Z})$ -module is completely understood and easy to describe for  $G = \text{IA}_n$  or  $\mathcal{I}_n^1$ , this is not the case with higher degree components of  $L(G)$ .

**1.6. Some closing comments.** In a sequel paper [CEP], Theorem 1.1 will be generalized to higher terms of the lower central series; more precisely, it will be shown that any subgroup of  $\text{IA}_n$  containing  $\gamma_N \text{IA}_n$  is finitely generated whenever  $n \geq 4N - 3$  and any subgroup of  $\mathcal{I}_n^1$  containing  $\gamma_N \mathcal{I}_n^1$  is finitely generated whenever  $n \geq \max\{4, 2N - 1\}$ . This result obviously implies the last assertion of Corollary 1.5 (but not the full statement). The method from [CEP] does not seem to be applicable to Theorem 1.4.

The question whether  $[G, G]$  is finitely generated for  $G = \text{IA}_n$  or  $\mathcal{I}_n^1$  is now settled for all  $n \neq 3$ . The answer is positive for  $n \geq 4$  (by Theorem 1.1 and [CEP]) and negative for  $n = 2$ . The latter holds since  $\text{IA}_2$  is a free group of rank 2 (see § 3) and  $\mathcal{I}_2$  is a free

group of countable rank, as proved in [Me]. It seems likely that the group  $[IA_3, IA_3]$  is not finitely generated and moreover has infinite-dimensional first rational homology. Some evidence for this is provided by the fact that the assertions of Theorem 1.4 and both parts of Theorem 1.9 are false for  $G = IA_3$  and  $N = 2$  (see Proposition 9.5).

**Organization.** In § 2 we will introduce the notion of a  $\rho$ -CG sequence and reduce Theorems 1.3 and 1.4 to the problem of existence of  $\rho$ -CG sequences for suitable  $\rho$  (see Proposition 2.4). In § 3 we will prove that  $G = IA_n$ ,  $n \geq 4$ , admits a  $\rho$ -CG sequence for any  $\rho$  with  $\rho(G)$  abelian (see Theorem 3.1). Theorem 3.1 and Proposition 2.4 immediately imply Theorem 1.3 (and hence Theorem 1.1) for  $G = IA_n$  and Theorem 1.4(a).

In § 4.1 we will introduce the notion of a *partially commuting  $\mathbf{n}$ -group*, where  $\mathbf{n} = \{1, 2, \dots, n\}$ , which formalizes the key properties of basic commutation relations in  $\text{Aut}(F_n)$  and  $\text{Mod}_g^1$ . Then in § 4.2 we will describe a general method of constructing  $\rho$ -CG sequences in partially commuting  $\mathbf{n}$ -groups. In § 5 we will introduce and study the notion of a regular  $SL_n(\mathbb{Z})$ -module. In § 6 we will state and prove our main technical theorem (Theorem 6.1). Combined with Proposition 2.4, Theorem 6.1 will be used to establish the remaining cases of Theorems 1.3 and 1.4. The fact that the hypotheses of Theorem 6.1 hold for  $G = IA_n$  is easy and will be checked right after the statement of Theorem 6.1. The corresponding verification for  $G = \mathcal{I}_g^1$  is more involved and will be made in § 7.

In § 8 we will prove Theorem 1.6, thereby completing the proof of Corollary 1.5. Finally, in § 9 we will prove Theorem 1.9 and also discuss other known results of the same kind. These two sections do not use any technical machinery introduced in the paper and can be read right after the introduction.

**Acknowledgments.** We are extremely grateful to Andrei Rapinchuk who encouraged us to take on this project and suggested the general approach to the proof of Theorem 1.9. We are also indebted to Andrei Jaikin who proposed a tremendous simplification of our original proof of Theorem 1.6. We are very grateful to the anonymous referee who carefully read the paper and made a number of suggestions that helped improve the exposition. Finally, we would like to thank Thomas Church, Andrew Putman and Zezhou Zhang for useful comments on earlier versions of this paper.

**Notation.** When considering cohomology of a group  $G$  with coefficients in some representation, we will sometimes use the notation  $H^k(G, V)$  where  $V$  is the underlying space of the representation and sometimes  $H^k(G, \rho)$  when  $\rho$  is the action of  $G$  on  $V$ . We hope that this inconsistency will not cause confusion.

## 2. $\rho$ -CENTRALIZING GENERATING SEQUENCES

Let  $G$  be a group. Given a homomorphism  $\rho$  from  $G$  to another group, denote by  $C_\rho(G)$  the set of all  $g \in G$  such that  $\rho(g)$  is a non-trivial central element of  $\rho(G)$ . In particular, if  $\rho(G)$  is abelian, then  $C_\rho(G) = G \setminus \text{Ker } \rho$ .

**Definition.** Let  $g_1, g_2, \dots, g_k$  be a finite sequence of elements of  $G$ , and for each  $1 \leq i \leq k$  let  $G_i = \langle g_j : 1 \leq j \leq i \rangle$  be the subgroup generated by the first  $i$  elements of the sequence. We will say that the sequence  $\{g_n\}$  is  $\rho$ -centralizing if  $g_1 \in C_\rho(G)$  and for each  $2 \leq i \leq k$  there exists  $j < i$  such that  $g_j \in C_\rho(G)$  and  $[g_i, g_j] \in G_{i-1}$ .

A key problem we will consider in this paper is establishing the existence of a  $\rho$ -centralizing generating sequence for a given  $\rho$  (that is, a  $\rho$ -centralizing sequence whose

elements generate  $G$ ). Since this condition will arise very frequently, we will abbreviate  $\rho$ -centralizing generating as  $\rho$ -CG for the rest of the paper.

The following lemma describes a particularly simple way to produce  $\rho$ -CG sequences.

**Lemma 2.1.** *Let  $G$  be a finitely generated group, and let  $\rho$  be a homomorphism from  $G$  to another group. Suppose that there exist pairwise commuting elements  $g_1, \dots, g_k \in C_\rho(G)$  such that the union of their centralizers  $\cup_{i=1}^k \text{Cent}_G(g_i)$  generates  $G$ . Then  $G$  has a  $\rho$ -CG sequence.*

*Proof.* Since  $G$  is finitely generated, we can find a finite generating set  $h_1, \dots, h_m$  for  $G$  contained in  $\cup \text{Cent}_G(g_i)$ . It is clear that the sequence  $g_1, \dots, g_k, h_1, \dots, h_m$  is  $\rho$ -CG.  $\square$

In this section we will show that

- (i) if  $\rho$  is an irreducible representation of  $G$  over some field, then the existence of a  $\rho$ -CG sequence implies that  $H^1(G, \rho) = 0$ ;
- (ii) if  $\rho : G \rightarrow \mathbb{R}$  is a non-trivial character, the existence of a  $\rho$ -CG sequence implies that  $[\rho]$  lies in  $\Sigma(G)$ , the BNS invariant of  $G$ .

We will start with briefly recalling Brown's characterization of  $\Sigma(G)$  in terms of actions on  $\mathbb{R}$ -trees [Br1] and establishing a very simple property of group 1-cocycles.

We will need only very basic properties of group actions on  $\mathbb{R}$ -trees (see, e.g. [AB] for the proofs). Let  $X$  be an  $\mathbb{R}$ -tree with metric  $d$ , and suppose that a group  $G$  acts on  $X$  by isometries; denote the action by  $\alpha$ . For each  $g \in G$  define  $l(g) = \inf_{x \in X} d(\alpha(g)x, x)$  (it is known that the infimum is always attained), and let  $\mathcal{A}_g = \{x \in X : d(\alpha(g)x, x) = l(g)\}$ . An element  $g \in G$  is called *elliptic* (with respect to  $\alpha$ ) if  $l(g) = 0$ , in which case  $\mathcal{A}_g$  is the fixed point set of  $g$ . If  $l(g) > 0$ , then  $g$  is called *hyperbolic*. In this case  $\mathcal{A}_g$  is a line, called *the axis of  $g$* , and  $g$  acts on  $\mathcal{A}_g$  as a translation by  $l(g)$ ; moreover,  $\mathcal{A}_g$  is the unique line invariant under the action of  $g$ . An action  $\alpha$  is called

- *abelian* if there exists a character  $\chi : G \rightarrow \mathbb{R}$  such that  $l(g) = |\chi(g)|$  for all  $g \in G$ . If such  $\chi$  exists, it is unique up to sign, and we say that  $\alpha$  is associated with  $\chi$ .
- *non-trivial* if it has no (global) fixed points and no (global) invariant line.

The following characterization of the BNS invariant  $\Sigma(G)$  was obtained by Brown [Br1]:

**Theorem 2.2.** *Let  $G$  be a finitely generated group, and let  $\chi : G \rightarrow \mathbb{R}$  be a non-trivial character. Then  $[\chi] \in \Sigma(G)$  if and only if  $G$  has no non-trivial abelian actions associated with  $\chi$ . In particular, by Theorem 1.2,  $[G, G]$  is finitely generated if and only if  $G$  has no non-trivial abelian actions.*

Now let  $(\rho, V)$  be a representation of a group  $G$ . Recall that  $H^1(G, \rho) \cong Z^1(G, \rho)/B^1(G, \rho)$  where

$$Z^1(G, \rho) = \{f : G \rightarrow V : f(xy) = f(x) + \rho(x)f(y) \text{ for all } x, y \in G\};$$

$$B^1(G, \rho) = \{f : G \rightarrow V : \text{there exists } v \in V \text{ such that } f(x) = \rho(x)v - v \text{ for all } x \in G\}.$$

Also recall that  $C_\rho(G) = \{g \in G \text{ such that } \rho(g) \text{ is non-trivial and central in } \rho(G)\}$ .

**Lemma 2.3.** *Let  $G$  be a group, let  $(\rho, V)$  be a non-trivial irreducible representation of  $G$  over a field, let  $g \in C_\rho(G)$ , and let  $f \in Z^1(G, \rho)$ . The following hold:*

- (i) *There exists a coboundary  $b \in B^1(G, \rho)$  with  $f(g) = b(g)$ .*
- (ii) *Suppose that  $f(g) = 0$ . Then  $f(x) = 0$  for every  $x \in G$  such that  $f(gx) = f(xg)$ ; in particular,  $f(x) = 0$  for every  $x \in G$  which commutes with  $g$ .*

*Proof.* (i) Since  $g \in C_\rho(G)$  and  $\rho$  is irreducible, by Schur's Lemma the operator  $\rho(g) - 1$  is invertible. Thus, if  $v = (\rho(g) - 1)^{-1}f(g)$ , then the map  $b(x) = \rho(x)v - v$  is a coboundary with the required property.

(ii) We have  $f(xg) = f(x) + \rho(x)f(g) = f(x)$  and  $f(gx) = f(g) + \rho(g)f(x) = \rho(g)f(x)$ , so  $(\rho(g) - 1)f(x) = 0$  and hence  $f(x) = 0$ .  $\square$

**Proposition 2.4.** *Let  $G$  be a finitely generated group. The following hold:*

- (a) *If  $(\rho, V)$  is an irreducible representation of  $G$  over some field and  $G$  admits a  $\rho$ -CG sequence, then  $H^1(G, \rho) = 0$ .*
- (b) *Let  $\rho : G \rightarrow \mathbb{R}$  be a non-trivial character. If  $G$  admits a  $\rho$ -CG sequence, then  $[\rho] \in \Sigma(G)$ . In particular, if  $G$  admits a  $\rho$ -CG sequence for every non-trivial character  $\rho$ , then  $[G, G]$  is finitely generated.*

**Remark.** A weaker version of Proposition 2.4(b) has been proved before (see [KMM, Lemma 1.9]). That lemma covers the case when  $G$  has a generating sequence  $g_1, \dots, g_t$  such that  $g_1 \in C_\rho(g)$  and for each  $i \geq 2$  there exists  $j < i$  such that  $g_i$  and  $g_j$  commute and  $g_j \in C_\rho(g)$ .

*Proof.* In both parts of the proof we let  $g_1, \dots, g_k$  be a  $\rho$ -CG sequence and  $G_i = \langle g_1, \dots, g_i \rangle$  for  $1 \leq i \leq k$  (so that  $G_k = G$ ).

(a) Take any  $f \in Z^1(G, \rho)$ . By Lemma 2.3(i), after modifying  $f$  by a coboundary, we can assume that  $f(g_1) = 0$ . Then  $f = 0$  on  $G_1$ , and we will now prove that  $f = 0$  on  $G_i$  for all  $1 \leq i \leq k$  by induction on  $i$ .

Take  $i \geq 2$ , and assume that  $f = 0$  on  $G_{i-1}$ . By hypothesis there exists  $j < i$  and  $r \in G_{i-1}$  such that  $g_j \in C_\rho(G)$  and  $g_i g_j = g_j g_i r$ . Then  $f(g_i g_j) = f(g_j g_i) + \rho(g_j g_i)f(r) = f(g_j g_i)$ , whence  $f(g_i) = 0$  by Lemma 2.3(ii). Since  $G_i = \langle G_{i-1}, g_i \rangle$ , we have  $f(G_i) = 0$  as desired.

(b) The following argument is similar to the proof of the main theorem in [OK]. Let  $(\alpha, X)$  be an abelian action of  $G$  on an  $\mathbb{R}$ -tree  $X$  associated to  $\rho$ . By Theorem 2.2, it suffices to show that such  $\alpha$  is trivial.

First note that  $C_\rho(G) = G \setminus \text{Ker } \rho$  (since  $\rho(G)$  is abelian). Thus, if  $l : G \rightarrow \mathbb{R}_{\geq 0}$  is the length function associated to  $\alpha$ , then  $C_\rho(G) = \{g \in G : l(g) > 0\}$ . In other words,  $C_\rho(G)$  is precisely the set of hyperbolic elements (with respect to  $\alpha$ ); in particular,  $g_1$  is hyperbolic by assumption. We claim that its axis  $\mathcal{A} = \mathcal{A}_{g_1}$  is invariant under the entire group  $G$  (which will finish the proof). We will prove that  $\mathcal{A}$  is invariant under  $G_i$  for all  $1 \leq i \leq k$  by induction on  $i$ .

Take  $i \geq 2$ , and assume that  $\mathcal{A}$  is invariant under  $G_{i-1}$  (in particular,  $\mathcal{A}$  is the axis for any hyperbolic element in  $G_{i-1}$  by the uniqueness of the invariant line of a hyperbolic element). By hypothesis there exists  $j < i$  and  $r \in G_{i-1}$  such that  $g_j$  is hyperbolic and  $g_i^{-1} g_j g_i = g_j r$ . The element  $g_i^{-1} g_j g_i$  is also hyperbolic (being a conjugate of  $g_j$ ), and its axis is  $\alpha(g_i^{-1})(\mathcal{A})$ . On the other hand  $g_i^{-1} g_j g_i = g_j r \in G_{i-1}$ , so  $\mathcal{A}$  is the axis of  $g_i^{-1} g_j g_i$ . Thus,  $\alpha(g_i^{-1})(\mathcal{A}) = \mathcal{A}$ , so  $\mathcal{A}$  is  $g_i$ -invariant and hence  $G_i$ -invariant.  $\square$

### 3. PROOFS OF THEOREM 1.3 FOR $G = \text{IA}_n$ AND THEOREM 1.4(A)

In this section we will prove Theorem 1.3 for  $G = \text{IA}_n$  and Theorem 1.4(a). By Proposition 2.4, we are reduced to proving the following result:

**Theorem 3.1.** *Let  $G = \text{IA}_n$  for some  $n \geq 4$ , and let  $\rho$  be a non-trivial homomorphism from  $G$  to another group with  $\text{Ker } \rho \supseteq [G, G]$ . Then  $G$  has a  $\rho$ -CG sequence.*



The proof of Theorem 3.1 will illustrate many key ideas used in the proof of other cases of Theorems 1.3 and Theorem 1.4 without introducing the general technical machinery.

For the rest of the paper, given  $n \in \mathbb{N}$ , we set

$$\mathbf{n} = \{1, \dots, n\}.$$

In [Ma], Magnus proved that  $\text{IA}_n$  is generated by the automorphisms  $K_{ij}$ ,  $i \neq j$  and  $K_{ijk}$ ,  $i, j, k$  distinct, with  $i, j, k \in \mathbf{n}$ , defined below (see also [BBM] for another proof):

$$K_{ij} : \begin{cases} x_i \mapsto x_j^{-1} x_i x_j \\ x_k \mapsto x_k \text{ for } k \neq i \end{cases}, \quad K_{ijk} : \begin{cases} x_i \mapsto x_i[x_j, x_k] \\ x_l \mapsto x_l \text{ for } l \neq i \end{cases}.$$

Since  $K_{ij}^{-1} = K_{ikj}$ , it is enough to consider  $K_{ijk}$  with  $j < k$ . The elements  $K_{ij}$  and  $K_{ijk}$  with  $j < k$  will be referred to as the *standard generators* of  $\text{IA}_n$ .

**Remark.** If  $n = 2$ , there are no generators  $K_{ijk}$ , and  $K_{12}$  and  $K_{21}$  are simply the conjugations by  $x_2$  and  $x_1$ , respectively. Hence the group  $\text{IA}_2$  coincides with the group of inner automorphisms of  $F_2$  and thus is free of rank 2.

Below is the list of relations between the elements  $K_{ij}$  and  $K_{ijk}$  that will be used in the proof of Theorem 3.1.

**Lemma 3.2.** *Let  $a_1, a_2, a_3, b_1, b_2, b_3 \in \mathbf{n}$ , and assume that  $\{a_i\}$  are distinct and  $\{b_i\}$  are distinct. The following relations hold:*

- (R1)  $[K_{a_1 a_2}, K_{b_1 b_2}] = 1$  if  $a_1 \neq b_1, b_2$  and  $b_1 \neq a_1, a_2$ ;
- (R2)  $[K_{a_1 a_2}, K_{b_1 b_2 b_3}] = 1$  if  $a_1 \neq b_1, b_2, b_3$  and  $b_1 \neq a_1, a_2$ ;
- (R3)  $[K_{a_1 a_2 a_3}, K_{b_1 b_2 b_3}] = 1$  if  $a_1 \neq b_1, b_2, b_3$  and  $b_1 \neq a_1, a_2, a_3$ ;
- (R4)  $[K_{bcd}, K_{ab}] = [K_{ad}, K_{ac}]$  if  $a, b, c, d \in \mathbf{n}$  are distinct.

The main technical result we shall establish in this section is the following proposition:

**Proposition 3.3.** *Let  $G = \text{IA}_n$  for some  $n \geq 4$ , and let  $H$  be a proper subgroup of  $G$  containing  $[G, G]$ . Then there exists  $a \in \text{Aut}(F_n)$  such that the conjugate subgroup  $aHa^{-1}$  does not contain any of the elements  $K_{12}, K_{21}, K_{34}$  and  $K_{43}$ .*

We will start by deducing Theorem 3.1 from Proposition 3.3.

*Proof of Theorem 3.1.* Given  $a \in \text{Aut}(F_n)$ , define the homomorphism  $\rho_a$  by  $\rho_a(x) = \rho(a^{-1}xa)$ . Clearly, if  $x_1, \dots, x_t$  is a  $\rho_a$ -CG sequence in  $G$ , then  $a^{-1}x_1a, \dots, a^{-1}x_t a$  is a  $\rho$ -CG sequence, so it is enough to construct a  $\rho_a$ -CG sequence (for some  $a$ ). Since  $\text{Ker } \rho_a = a(\text{Ker } \rho)a^{-1}$ , by letting  $a$  be an element from the conclusion of Proposition 3.3 with  $H = \text{Ker } \rho$  and replacing  $\rho$  by  $\rho_a$ , we may assume that  $\text{Ker } \rho$  does not intersect the set  $S = \{K_{12}, K_{21}, K_{34}, K_{43}\}$ . Since  $\rho(G)$  is abelian, we conclude that  $S \subseteq C_\rho(G)$ .

Using relations (R1)-(R3) of Lemma 3.2, it is easy to show that the only standard generators of  $\text{IA}_n$  which do not commute with an element of  $S$  are  $K_{134}, K_{234}, K_{312}, K_{412}$ . However, by relations (R4) each of these 4 elements commutes with an element of  $S$  modulo the subgroup generated by all  $K_{xy}$ . Thus, if we order the standard generators of  $\text{IA}_n$  so that  $K_{12}, K_{34}, K_{21}, K_{43}$  come first (in this order) and  $K_{134}, K_{234}, K_{312}, K_{412}$  come last, then these generators form a  $\rho$ -CG sequence.  $\square$

**3.1. The abelianization of  $\text{IA}_n$ .** Since the quotient  $\text{Aut}(F_n)/\text{IA}_n$  is isomorphic to  $\text{GL}_n(\mathbb{Z})$ , the abelianization  $\text{IA}_n^{ab} = \text{IA}_n/[\text{IA}_n, \text{IA}_n]$  is a  $\text{GL}_n(\mathbb{Z})$ -module. Let  $V = \mathbb{Z}^n$ , considered as a standard  $\text{GL}_n(\mathbb{Z})$ -module, and let  $V^* = \text{Hom}(V, \mathbb{Z})$  be the dual module. It is well known that  $\text{IA}_n^{ab}$  is canonically isomorphic to  $V^* \otimes (V \wedge V)$  as a  $\text{GL}_n(\mathbb{Z})$ -module; to the best of our knowledge, the first full proof of this isomorphism was given by Formanek [Fo] (for alternative references see, e.g., [Kaw, Theorem 6.1] or [Bh, § 5]).

Let  $e_1, \dots, e_n$  be the standard basis of  $V$ , and let  $e_1^*, \dots, e_n^*$  be the corresponding dual basis of  $V^*$ , and set

$$\kappa_{ijk} = e_i^* \otimes (e_j \wedge e_k).$$

The above isomorphism between  $\text{IA}_n^{ab}$  and  $V^* \otimes (V \wedge V)$  is given by

$$(3.1) \quad [K_{ij}] \mapsto \kappa_{iij} \text{ and } [K_{ijk}] \mapsto \kappa_{ijk}$$

where  $[x]$  denotes the image of  $x \in \text{IA}_n$  in  $\text{IA}_n^{ab}$ .

From now on we will identify  $\text{IA}_n^{ab}$  with  $V^* \otimes (V \wedge V)$  via the map (3.1).

Given  $i, j \in \mathbf{n}$  with  $i \neq j$ , define  $E_{ij}, F_{ij} \in \text{SL}_n(\mathbb{Z})$  by

$$(3.2) \quad E_{ij} : \begin{cases} e_j \mapsto e_j + e_i \\ e_k \mapsto e_k \text{ for } k \neq j \end{cases} \quad F_{ij} : \begin{cases} e_i \mapsto -e_j \\ e_j \mapsto e_i \\ e_k \mapsto e_k \text{ for } k \neq i, j \end{cases}$$

Note that  $F_{ij} = E_{ij} E_{ji}^{-1} E_{ij}$ .

The following observation describes the action of  $E_{ij}^{\pm 1}$  and  $F_{ij}$  on the generators  $\kappa_{abc}$  of  $\text{IA}_n^{ab}$ .

**Observation 3.4.**

(a) *The following hold, where distinct letters denote distinct integers in  $\mathbf{n}$ :*

$$\begin{aligned} E_{ij}^{\pm 1} \kappa_{iic} - \kappa_{iic} &= \mp \kappa_{jic} & E_{ij}^{\pm 1} \kappa_{jji} - \kappa_{jji} &= 0 \\ E_{ij}^{\pm 1} \kappa_{ibc} - \kappa_{ibc} &= \mp \kappa_{jbc} & E_{ij}^{\pm 1} \kappa_{aa j} - \kappa_{aa j} &= \pm \kappa_{aa i} \\ E_{ij}^{\pm 1} \kappa_{ajc} - \kappa_{ajc} &= \pm \kappa_{aic} & E_{ij}^{\pm 1} \kappa_{iij} - \kappa_{iij} &= \pm \kappa_{jji} \\ E_{ij}^{\pm 1} \kappa_{jjc} - \kappa_{jjc} &= \pm \kappa_{jic} & E_{ij}^{\pm 1} \kappa_{ijc} - \kappa_{ijc} &= \pm (\kappa_{iic} - \kappa_{jjc}) - \kappa_{jic}. \end{aligned}$$

(b) *Let  $i, j, a, b, c \in \mathbf{n}$  with  $i \neq j$  and  $b \neq c$ . Then*

- (i)  $E_{ij} \kappa_{abc} = \kappa_{abc}$  if  $i \neq a$  and  $j \neq b, c$ ;
- (ii)  $F_{ij} \kappa_{abc} = \pm \kappa_{\sigma(a)\sigma(b)\sigma(c)}$  where  $\sigma$  is the transposition  $(i, j)$ .

*Proof.* The action of  $E_{ij}$  and  $F_{ij}$  on the basis  $\{e_k^*\}$  on  $V^*$  is given by

$$E_{ij} : \begin{cases} e_i^* \mapsto e_i^* - e_j^* \\ e_k^* \mapsto e_k^* \text{ for } k \neq i \end{cases} \quad F_{ij} : \begin{cases} e_i^* \mapsto -e_j^* \\ e_j^* \mapsto e_i^* \\ e_k^* \mapsto e_k^* \text{ for } k \neq i, j \end{cases}$$

All parts of Observation 3.4 immediately follow from these formulas and (3.2). For instance,

$$E_{ij}^{\pm 1} \kappa_{iic} = E_{ij}^{\pm 1} (e_i^* \otimes (e_i \wedge e_c)) = (e_i^* \mp e_j^*) \otimes (e_i \wedge e_c) = \kappa_{iic} \mp \kappa_{jic}.$$

□

**3.2. Proof of Proposition 3.3.** Let  $U$  be any abelian group which contains an isomorphic copy of every finitely generated abelian group, and let

$$\Lambda = \text{Hom}_{\mathbb{Z}}(\text{IA}_n^{ab}, U).$$

We will define the action of  $\text{GL}_n(\mathbb{Z})$  on  $\Lambda$  in the usual way:

$$(g\lambda)(x) = \lambda(g^{-1}x) \text{ for all } g \in \text{GL}_n(\mathbb{Z}), \lambda \in \Lambda \text{ and } x \in \text{IA}_n^{ab}.$$

Our first goal is to reformulate Proposition 3.3 in terms of the  $\text{GL}_n(\mathbb{Z})$ -module  $\Lambda$  (see Lemma 3.6 below).

Since  $\text{IA}_n^{ab}$  (and any of its quotients) is a finitely generated abelian group, every subgroup of  $\text{IA}_n^{ab}$  is the kernel of some element of  $\Lambda$ . Fix  $H$ , a proper subgroup of  $\text{IA}_n$  containing  $[\text{IA}_n, \text{IA}_n]$ , and choose  $\lambda \in \Lambda$  such that  $\text{Ker } \lambda = H/[\text{IA}_n, \text{IA}_n]$ . Note that  $\lambda \neq 0$  since  $H$  is proper.

Recall that for an element  $x \in \text{IA}_n$  we denote by  $[x]$  its image in  $\text{IA}_n^{ab}$ . Given  $a \in \text{Aut}(F_n)$ , let  $\bar{a}$  be its image in  $\text{GL}_n(\mathbb{Z})$ .

**Claim 3.5.** *Let  $a \in \text{Aut}(F_n)$  and  $x \in \text{IA}_n$ . Then  $x \in aHa^{-1}$  if and only if  $(\bar{a}\lambda)([x]) = 0$ .*

*Proof.* By definition  $(\bar{a}\lambda)([x]) = \lambda(\bar{a}^{-1}[x]) = \lambda([a^{-1}xa])$ . Thus,

$$(\bar{a}\lambda)([x]) = 0 \iff [a^{-1}xa] \in \text{Ker } \lambda = H/[\text{IA}_n, \text{IA}_n] \iff a^{-1}xa \in H \iff x \in aHa^{-1}.$$

□

Given  $i, j, k \in \mathbf{n}$  with  $j \neq k$  define

$$c_{ijk}(\lambda) = \lambda(\kappa_{ijk}).$$

In view of Claim 3.5 and (3.1), Proposition 3.3 is equivalent to the following lemma:

**Lemma 3.6.** *Assume that  $n \geq 4$ . Then for any nonzero  $\lambda \in \Lambda$  there exists  $g \in \text{GL}_n(\mathbb{Z})$  such that  $c_{112}(g\lambda), c_{221}(g\lambda), c_{334}(g\lambda)$  and  $c_{443}(g\lambda)$  are all nonzero.*

The following two results follow directly from Observation 3.4. We will use Observation 3.7 in the proof of Lemma 3.6 without an explicit reference.

**Observation 3.7.** *Let  $i, j, a \in \mathbf{n}$  be distinct and  $\lambda \in \Lambda$ . Then*

$$\begin{aligned} c_{aa j}(E_{ij}\lambda) &= c_{aa j}(\lambda) - c_{aa i}(\lambda) & c_{ii j}(E_{ij}\lambda) &= c_{ii j}(\lambda) - c_{jj i}(\lambda) \\ c_{jj a}(E_{ij}\lambda) &= c_{jj a}(\lambda) - c_{jia}(\lambda) & c_{iia}(E_{ij}\lambda) &= c_{iia}(\lambda) + c_{jia}(\lambda) \end{aligned}$$

and  $c_{xxy}(E_{ij}\lambda) = c_{xxy}(\lambda)$  for all  $(x, y) \neq (a, j), (i, j), (j, a), (i, a)$ .

**Observation 3.8.** *Let  $i, j \in \mathbf{n}$  with  $i \neq j$  and let  $\sigma$  be the transposition  $(i, j)$ . Then for any  $\lambda \in \Lambda$  and  $x, y, z \in \mathbf{n}$  we have  $c_{xyz}(F_{ij}\lambda) = \pm c_{\sigma(x)\sigma(y)\sigma(z)}(\lambda)$ .*

*Proof of Lemma 3.6.* Step 1: There exists  $g_1 \in \text{SL}_n(\mathbb{Z})$  such that  $c_{112}(g_1\lambda) \neq 0$ .

First of all, by Observation 3.8 it suffices to find distinct  $a, b \in \mathbf{n}$  such that  $c_{aab}(g_1\lambda) \neq 0$ .

If  $c_{xxy}(\lambda) \neq 0$  for some  $x \neq y$ , we are done. Suppose now that  $c_{xxy}(\lambda) = 0$  for all  $x \neq y$ . Then there must exist distinct  $a, b, c$  with  $c_{abc}(\lambda) \neq 0$ , in which case

$$c_{aab}(E_{ca}\lambda) = c_{aab}(\lambda) - c_{acb}(\lambda) = 0 + c_{abc}(\lambda) \neq 0.$$

Step 2: There exists  $g_2 \in \text{SL}_n(\mathbb{Z})$  such that  $c_{112}(g_2\lambda), c_{113}(g_2\lambda) \neq 0$ .

By Step 1, we can assume that  $c_{112}(\lambda) \neq 0$ . If  $c_{113}(\lambda) \neq 0$ , we are done. And if  $c_{113}(\lambda) = 0$ , then  $c_{113}(E_{23}\lambda) = c_{113}(\lambda) - c_{112}(\lambda) = 0 - c_{112}(\lambda) \neq 0$  and  $c_{112}(E_{23}\lambda) = c_{112}(\lambda) \neq 0$ .

Step 3: There exists  $g_3 \in \mathrm{SL}_n(\mathbb{Z})$  such that either  $c_{112}(g_3\lambda), c_{113}(g_3\lambda), c_{221}(g_3\lambda) \neq 0$  or  $c_{112}(g_3\lambda), c_{334}(g_3\lambda) \neq 0$ .

By Step 2, we can assume that  $c_{112}(\lambda), c_{113}(\lambda) \neq 0$ . If  $c_{224}(\lambda) \neq 0$  or  $c_{221}(\lambda) \neq 0$ , we are done (using Observation 3.8 in the former case), so assume that  $c_{224}(\lambda) = c_{221}(\lambda) = 0$ . We consider 3 cases.

Case 1:  $c_{223}(\lambda) \neq 0$ . Then

$$c_{224}(E_{34}\lambda) = c_{224}(\lambda) - c_{223}(\lambda) = -c_{223}(\lambda) \neq 0;$$

$$c_{113}(E_{34}\lambda) = c_{113}(\lambda) \neq 0.$$

Hence  $c_{113}(E_{34}\lambda), c_{224}(E_{34}\lambda) \neq 0$ .

Case 2:  $c_{223}(\lambda) = 0$  and  $c_{123}(\lambda) = 0$ . Then

$$c_{112}(E_{21}\lambda) = c_{112}(\lambda) \neq 0;$$

$$c_{221}(E_{21}\lambda) = c_{221}(\lambda) - c_{112}(\lambda) = -c_{112}(\lambda) \neq 0;$$

$$c_{113}(E_{21}\lambda) = c_{113}(\lambda) - c_{123}(\lambda) = -c_{123}(\lambda) \neq 0.$$

Case 3:  $c_{223}(\lambda) = 0$  and  $c_{123}(\lambda) \neq 0$ .

Again we have  $c_{112}(E_{21}\lambda) \neq 0$  and  $c_{221}(E_{21}\lambda) \neq 0$ , and this time

$$c_{223}(E_{21}\lambda) = c_{223}(\lambda) + c_{123}(\lambda) = c_{123}(\lambda) \neq 0.$$

Hence we are done by Observation 3.8.

Step 4: There exists  $g_4 \in \mathrm{SL}_n(\mathbb{Z})$  such that  $c_{112}(g_4\lambda), c_{334}(g_4\lambda) \neq 0$ .

By Step 3, we can assume that  $c_{112}(\lambda) \neq 0, c_{221}(\lambda) \neq 0$  and  $c_{113}(\lambda) \neq 0$ . If  $c_{443}(\lambda) \neq 0$  or  $c_{224}(\lambda) \neq 0$ , we are done (by Observation 3.8).

So assume that  $c_{443}(\lambda) = c_{224}(\lambda) = 0$ . Then

$$c_{224}(E_{14}\lambda) = c_{224}(\lambda) - c_{221}(\lambda) = -c_{221}(\lambda) \neq 0.$$

Case 1:  $c_{413}(\lambda) = 0$ . Then  $c_{113}(E_{14}\lambda) = c_{113}(\lambda) + c_{413}(\lambda) \neq 0$ , so  $c_{113}(E_{14}\lambda) \neq 0$  and  $c_{224}(E_{14}\lambda) \neq 0$ .

Case 2:  $c_{413}(\lambda) \neq 0$ . Then  $c_{443}(E_{14}\lambda) = c_{443}(\lambda) - c_{413}(\lambda) = -c_{413}(\lambda) \neq 0$  and  $c_{221}(E_{14}\lambda) = c_{221}(\lambda) \neq 0$ , so  $c_{443}(E_{14}\lambda) \neq 0$  and  $c_{221}(E_{14}\lambda) \neq 0$ .

In both cases we are done by Observation 3.8.

Final Step. By Step 4, we can assume that  $c_{112}(\lambda), c_{334}(\lambda) \neq 0$ . Set  $\alpha = 0$  if  $c_{221}(\lambda) \neq 0$  and  $\alpha = 1$  if  $c_{221}(\lambda) = 0$  and  $\beta = 0$  if  $c_{443}(\lambda) \neq 0$  and  $\beta = 1$  if  $c_{443}(\lambda) = 0$ . Then  $g = E_{21}^\alpha E_{43}^\beta$  satisfies the assertion of Lemma 3.6.  $\square$

It is not hard to modify the above proof of Lemma 3.6 to obtain the following generalization:

**Lemma 3.9.** *Let  $r \geq 2$  be an integer and  $n \geq 2r$ . Then for any nonzero  $\lambda \in \Lambda$  there exists  $g \in \mathrm{SL}_n(\mathbb{Z})$  such that  $c_{2i-1, 2i-1, 2i}(g\lambda)$  and  $c_{2i, 2i, 2i-1}(g\lambda)$  are nonzero for all  $1 \leq i \leq r$ .*

Using Lemma 3.9 instead of Lemma 3.6 in the proof of Theorem 3.1, it is easy to see that if  $n \geq 6$ , one can construct a  $\rho$ -CG sequence in  $G = \mathrm{IA}_n$  by using only the relations (R1)-(R3). We will not provide the details of this argument. However, in § 6 we will prove

Theorem 6.2 which generalizes Theorem 3.1 for  $n \geq 12$  and whose proof will use even fewer relations, namely only relations (R1)-(R3) with  $a_i \neq b_j$  for all  $i, j$ .

#### 4. $\mathbf{n}$ -GROUPS AND $\rho$ -CG SEQUENCES

4.1.  **$\mathbf{n}$ -groups.** Recall that given  $n \in \mathbb{N}$ , we denote by  $\mathbf{n}$  the set  $\{1, 2, \dots, n\}$ .

**Definition.** Let  $n \in \mathbb{N}$ . An  $\mathbf{n}$ -group is a group  $G$  endowed with a collection of subgroups  $\{G_I\}_{I \subseteq \mathbf{n}}$  such that

- (i)  $G_{\mathbf{n}} = G$ ;
- (ii)  $G_I \subseteq G_J$  whenever  $I \subseteq J$ .

Given  $d \in \mathbb{N}$ , we will say that  $G$  is *generated in degree  $d$*  if  $G = \langle G_I : |I| = d \rangle$ .

We will say that  $G$  is a *commuting*<sup>3</sup>  $\mathbf{n}$ -group if  $G_I$  and  $G_J$  commute (elementwise) whenever  $I$  and  $J$  are disjoint.

Note that if  $G$  is an  $\mathbf{n}$ -group, then any subgroup  $H$  of  $G$  becomes an  $\mathbf{n}$ -group if we set  $H_I = H \cap G_I$ . Clearly, if  $G$  is commuting, then so is  $H$ .

A simple example of a commuting  $\mathbf{n}$ -group is  $G = \mathrm{GL}_n(\mathbb{Z})$  with  $\{G_I\}$  defined as follows. Let  $e_1, \dots, e_n$  be the standard basis of  $\mathbb{Z}^n$ . Given  $I \subseteq \mathbf{n}$ , let  $\mathbb{Z}^I = \oplus_{i \in I} \mathbb{Z}e_i$ , and let

$$(4.1) \quad G_I = \{g \in G : g(\mathbb{Z}^I) \subseteq \mathbb{Z}^I \text{ and } g(e_j) = e_j \text{ for all } j \notin I\}.$$

Very similarly we can turn  $G = \mathrm{Aut}(F_n)$  into a commuting  $\mathbf{n}$ -group. Choose a free generating set  $x_1, \dots, x_n$ , and for each  $I \subseteq \mathbf{n}$  let  $G_I$  be the subgroup consisting of automorphisms which leave the subgroup  $\langle x_i : i \in I \rangle$  invariant and fix  $x_j$  for every  $j \notin I$ .

In the sequel we will not encounter the  $\mathbf{n}$ -group structure on  $\mathrm{Aut}(F_n)$  itself, but we will use the induced structure on  $\mathrm{IA}_n$ . Since  $\mathrm{IA}_n$  is generated by the automorphisms  $K_{ij}$  and  $K_{ijk}$  defined in § 3, it is clear that  $\mathrm{IA}_n$  is generated in degree 3 as an  $\mathbf{n}$ -group.

The fact that  $\mathrm{IA}_n$  is a commuting  $\mathbf{n}$ -group captures some of the relations (R1)-(R3) from Lemma 3.2 (namely those with  $a_i \neq b_j$  for all  $i, j$ ) while the remaining relations (R1)-(R3) and the relations (R4) use more delicate properties of  $\mathrm{IA}_n$  that seem hard to generalize to other groups.

As we will explain in § 7, the Torelli group  $\mathcal{I}_n^1$  has an analogous structure of an  $\mathbf{n}$ -group generated in degree 3, but that group is not *commuting*. Fortunately,  $\mathcal{I}_n^1$  is still *partially commuting* and *weakly commuting* according to the following definition (see § 7).

**Definition.** An  $\mathbf{n}$ -group  $G$  is called

- (i) *partially commuting* if  $G_I$  and  $G_J$  commute whenever  $I$  and  $J$  are disjoint and  $I$  consists of consecutive integers (that is,  $I = \{k, k+1, \dots, k+|I|-1\}$  for some  $k \in \mathbf{n}$ );
- (ii) *weakly commuting* if for any disjoint  $I$  and  $J$  there exists  $g \in G$  such that  $G_I^g = g^{-1}G_Ig$  commutes with  $G_J$ .

The fact that  $\mathcal{I}_n^1$  is partially commuting will be essential for the proofs of the main theorems, while the weak commutativity of  $\mathcal{I}_n^1$  will play a secondary role (it will not be used in the proof of Theorem 1.3 and could be eliminated from the proof of Theorem 1.4(c) at the expense of slightly increasing the lower bound on  $n$ ). By contrast, in [CEP] weak commutativity will play a central role, while partial commutativity will not be used at all.

<sup>3</sup>The authors thank Thomas Church and Andrew Putman for suggesting this terminology.

**4.2. Constructing  $\rho$ -CG sequences in partially commuting  $\mathbf{n}$ -groups.** In this subsection we fix an  $\mathbf{n}$ -group  $G$  and an integer  $m \geq 1$ , and denote by  $\text{Hom}_m(G)$  the class of all homomorphisms  $\rho$  from  $G$  to another group such that  $\rho(G)$  is *nilpotent of class  $m$* , that is,  $\gamma_{m+1}\rho(G) = \{1\}$  while  $\gamma_m\rho(G) \neq \{1\}$  (equivalently,  $\gamma_{m+1}G \subseteq \text{Ker } \rho$ , but  $\gamma_m G \not\subseteq \text{Ker } \rho$ ).

Our goal is to find a sufficient condition on  $G$  and  $m$  that ensures the existence of a  $\rho$ -CG sequence for any  $\rho \in \text{Hom}_m(G)$ . The precise formulation of this condition involves the notion of a regular  $\text{SL}_n(\mathbb{Z})$ -module and will be given in § 6. We will use the proof of Theorem 3.1 as a general guide to finding the suitable restrictions on  $G$  and  $m$ .

Let  $L(m) = \gamma_m G / \gamma_{m+1} G$ . As in § 3.2, let  $U$  be any abelian group which contains an isomorphic copy of every finitely generated abelian group, and let

$$\Lambda_m = \text{Hom}_{\mathbb{Z}}(L(m), U).$$

Since  $G$  is a finitely generated group,  $L(m)$  is a finitely generated abelian group, whence every subgroup of  $L(m)$  is the kernel of some element of  $\Lambda_m$ .

Now take any  $\rho \in \text{Hom}_m(G)$ . Then  $\gamma_{m+1}G \subseteq \text{Ker } \rho$  while  $\gamma_m G \not\subseteq \text{Ker } \rho$ , so the subgroup  $\text{Ker } \rho \cap \gamma_m G$  contains  $\gamma_{m+1}G$  and is strictly contained in  $\gamma_m G$ . Thus we can find a nonzero element  $\lambda_\rho \in \Lambda_m$  such that

$$\text{Ker } \lambda_\rho = (\text{Ker } \rho \cap \gamma_m G) / \gamma_{m+1} G.$$

For each  $I \subseteq \mathbf{n}$  let  $L(m)_I$  be the image of  $\gamma_m G \cap G_I$  in  $L(m)$ . The following technical definition will be very convenient for the subsequent discussion.

**Definition.** Let  $d \in \mathbb{N}$ . An element  $\lambda \in \Lambda_m$  will be called  *$d$ -generic* if there exist disjoint subsets  $I_1, \dots, I_{d+1}$  of  $\mathbf{n}$  each of which consists of consecutive integers such that  $\lambda$  does not vanish on  $L(m)_{I_k}$  for each  $1 \leq k \leq d+1$ .

**Remark.** Note that whether a given  $\lambda \in \Lambda_m$  is  $d$ -generic is completely determined by  $\text{Ker } \lambda$ .

**Lemma 4.1.** *Let  $G$  be a partially commuting  $\mathbf{n}$ -group generated in degree  $d$ , and let  $\rho \in \text{Hom}_m(G)$  be such that  $\lambda_\rho$  is  $d$ -generic. Then  $G$  admits a  $\rho$ -CG sequence.*

*Proof.* Since  $\lambda_\rho$  is  $d$ -generic, there exist elements  $g_k \in \gamma_m G \cap G_{I_k}$  for  $1 \leq k \leq d+1$  such that  $\rho(g_k)$  is non-trivial. The assumption  $\rho \in \text{Hom}_m(G)$  guarantees that  $\rho(\gamma_m G) = \gamma_m \rho(G)$  lies in the center of  $\rho(G)$ . Thus,  $\rho(g_k)$  is central, and so  $g_k \in C_\rho(G)$ .

The hypotheses on  $I_1, \dots, I_{d+1}$  imply that  $g_1, \dots, g_{d+1}$  commute with each other. Now take any  $I \subseteq \mathbf{n}$  with  $|I| = d$ . There exists  $1 \leq k \leq d+1$  such that  $I \cap I_k = \emptyset$ . Then every element of  $G_I$  commutes with  $g_k$ . Since  $G$  is generated in degree  $d$ , the centralizers of the elements  $g_1, \dots, g_{d+1}$  generate  $G$ . Hence by Lemma 2.1  $G$  has a  $\rho$ -CG sequence.  $\square$

Assume now that  $G$  is a normal subgroup of a group  $\Gamma$ , and let  $\bar{\Gamma} = \Gamma/G$ . The conjugation action of  $\Gamma$  on  $\gamma_m G$  and  $\gamma_{m+1} G$  induces an action of  $\bar{\Gamma}$  on  $L(m) = \gamma_m G / \gamma_{m+1} G$ . Define the corresponding action of  $\bar{\Gamma}$  on  $\Lambda_m$  in the usual way (as in § 3.2):

$$(4.2) \quad (g\lambda)(x) = \lambda(g^{-1}x) \text{ for all } g \in \bar{\Gamma}, \lambda \in \Lambda_m \text{ and } x \in L(m).$$

**Claim 4.2.** *Let  $G$  be a partially commuting  $\mathbf{n}$ -group generated in degree  $d$ , and suppose that for every nonzero  $\lambda \in \Lambda_m$  the  $\bar{\Gamma}$ -orbit of  $\lambda$  contains a  $d$ -generic element. Then  $G$  admits a  $\rho$ -CG sequence for any  $\rho \in \text{Hom}_m(G)$ .*

*Proof.* Take any  $\rho \in \text{Hom}_m(G)$ . By assumption there exists  $a \in \Gamma$  such that  $\bar{a}\lambda_\rho$  is  $d$ -generic, where  $\bar{a}$  is the image of  $a$  in  $\bar{\Gamma} = \Gamma/G$ . Define the homomorphism  $\rho_a$  by  $\rho_a(x) = \rho(a^{-1}xa)$ . Then

$$\begin{aligned} \text{Ker } \lambda_{\rho_a} &= (\text{Ker } \rho_a \cap \gamma_m G) / \gamma_{m+1} G = a(\text{Ker } \rho \cap \gamma_m G) a^{-1} / \gamma_{m+1} G \\ &= \bar{a} ((\text{Ker } \rho \cap \gamma_m G) / \gamma_{m+1} G) = \bar{a} \text{Ker } \lambda_\rho = \text{Ker } (\bar{a}\lambda_\rho). \end{aligned}$$

Since  $\bar{a}\lambda_\rho$  is  $d$ -generic, so is  $\lambda_{\rho_a}$ . By Lemma 4.1,  $G$  admits a  $\rho_a$ -CG sequence and hence also admits a  $\rho$ -CG sequence.  $\square$

**4.3. How to apply Claim 4.2.** This subsection contains a somewhat informal discussion of how to ensure that the hypotheses of Claim 4.2 hold. This discussion will not be formally used in any of the proofs in this paper, but it will help motivate the definition of a regular  $\text{SL}_n(\mathbb{Z})$ -module given in the next section. We will also mention one of the key ideas from [CEP] used in the proof of the generalization of Theorem 1.1 stated in § 1.6.

First, let us assume that for any permutation  $\sigma \in S_n$  there exists  $F_\sigma \in \bar{\Gamma}$  such that

$$(4.3) \quad F_\sigma(L(m)_I) = L(m)_{\sigma(I)} \text{ for every } I \subseteq \mathbf{n}.$$

We did not formally record this property in the proof of Theorem 3.1, but it is closely related to Observation 3.4(b)(ii).

Suppose now that  $G$  is generated in degree  $d$  with  $n \geq md(d+1)$ . An easy argument shows that  $L(m) = \sum_{|I|=md} L(m)_I$ . Thus, given any nonzero  $\lambda \in \Lambda_m$ , there exists at least one set  $I$  with  $|I| = md$  such that  $\lambda$  does not vanish on  $L(m)_I$ . Condition (4.3) then implies that for any  $J$  with  $|J| = md$  there exists  $g \in \bar{\Gamma}$  such that  $g\lambda$  does not vanish on  $L(m)_J$ . At this point it is natural to ask if we can guarantee (under a suitable extra assumption) that there exists  $g \in \bar{\Gamma}$  such that

$$(4.4) \quad g\lambda \text{ does not vanish on } L(m)_J \text{ for every } J \text{ with } |J| = md.$$

Since  $n \geq md(d+1)$ , if such  $g$  exists, then  $g\lambda$  is clearly  $d$ -generic.

There are two important cases where  $g$  satisfying (4.4) exists:

*Case 1:*  $\lambda = \lambda_\rho$  for some homomorphism  $\rho \in \text{Hom}_m(G)$  which can be extended to a homomorphism defined on  $\Gamma$ . In this case  $\rho_a$  defined by  $\rho_a(x) = \rho(a^{-1}xa)$ , with  $a \in \Gamma$ , has the same kernel as  $\rho$  itself, and the proof of Claim 4.2 shows that  $g\lambda_\rho = \lambda_\rho$  for all  $g \in \Gamma/G$ . Hence any  $g \in \bar{\Gamma}$  satisfies (4.4). This case, which may seem rather trivial, is actually sufficient for the proof of Theorem 1.9(2), as we will explain in § 9.1.

*Case 2:*  $\text{Im } \lambda$  is torsion-free. In this case we can embed  $\text{Im } \lambda$  into the additive group of  $\mathbb{R}$  (since  $L(m)$  is finitely generated) and hence extend  $\lambda$  to a homomorphism  $\lambda_{\mathbb{R}} : L(m) \otimes \mathbb{R} \rightarrow \mathbb{R}$  such that  $\lambda(L(m)_I) \neq 0 \iff \lambda_{\mathbb{R}}(L(m)_I \otimes \mathbb{R}) \neq 0$ . As will be shown in [CEP], one can find an element  $g$  satisfying (4.4) whenever there exists a subgroup  $\bar{\Gamma}'$  of  $\bar{\Gamma}$  which contains the elements  $F_\sigma$  from (4.3) and whose image in  $\text{GL}(L(m) \otimes \mathbb{R})$  has connected closure in the Zariski topology. The latter holds in the cases we are interested in: for instance, if  $G = \text{IA}_n$ , we can take  $\Gamma = \text{Aut}(F_n)$ , so that  $\bar{\Gamma} = \text{GL}_n(\mathbb{Z})$  and let  $\bar{\Gamma}' = \text{SL}_n(\mathbb{Z})$ . It is not hard to show that the closure of the image of  $\text{SL}_n(\mathbb{Z})$  in  $\text{GL}(L(m) \otimes \mathbb{R})$  is isomorphic to a quotient of  $\text{SL}_n(\mathbb{R})$  and hence Zariski-connected. Note that the torsion-free case is sufficient for the proof of Theorem 1.3.

In order to prove Theorem 1.4 and its consequences Corollary 1.5 and Theorem 1.9(1), we have to deal with  $\lambda$  not covered by either of the above cases, so we shall make more specific assumptions on the groups  $\Gamma$  and  $G$ . We shall assume that

- (i)  $\bar{\Gamma}$  contains a copy of  $\mathrm{SL}_n(\mathbb{Z})$ ;
- (ii)  $L(m)$  admits a graded structure  $\{L(m)^I\}$  which is compatible with the filtered structure  $\{L(m)_I\}$  in some sense (see condition (iii) in Theorem 6.1 and Lemma 6.3);
- (iii) the graded structure  $\{L(m)^I\}$  is “well-behaved” with respect to the action of the elements  $E_{ij}, F_{ij}$  defined by (3.2).

Property (iii) will formally say that  $L(m)$  is a regular  $\mathrm{SL}_n(\mathbb{Z})$ -module, as defined in the next section. A nice feature of this definition is that we will only need to construct the grading and check (iii) for  $m = 1$ ; once this is done, we will automatically obtain the grading on  $L(m)$  satisfying (iii) for arbitrary  $m$ . The assumptions (i)-(iii) will not be sufficient to ensure the existence of  $g \in \mathrm{SL}_n(\mathbb{Z})$  satisfying (4.4), but they will allow us to find  $g$  such that  $g\lambda$  is  $d$ -generic, again assuming  $n \geq md(d+1)$ .

## 5. REGULAR $\mathrm{SL}_n(\mathbb{Z})$ -MODULES

**Definition.** Let  $V$  be an abelian group (written additively) and  $n \in \mathbb{N}$ . An **n-grading** of  $V$  is a collection of additive subgroups  $\{V^I : I \subseteq \mathbf{n}\}$  of  $V$  such that

$$\sum_{I \subseteq \mathbf{n}} V^I = V$$

(note that the sum is not required to be direct). The *degree* of  $V$  (with respect to the grading  $\{V^I\}$ ), denoted by  $\deg(V)$ , is the smallest integer  $k$  such that  $V^I = 0$  whenever  $|I| > k$  (this automatically implies that  $\sum_{|I| \leq k} V^I = V$ ).

Let  $E_{ij}, F_{ij} \in \mathrm{SL}_n(\mathbb{Z})$  be defined by (3.2) in § 3.1.

**Definition.** Let  $V$  be an  $\mathrm{SL}_n(\mathbb{Z})$ -module endowed with an **n-grading**  $\{V^I\}$ . We will say that  $V$  is *regular* if the following properties hold:

- (1)  $E_{ij}$  acts trivially on  $V^I$  if  $I \cap \{i, j\} = \emptyset$ .
- (2) If  $I \subseteq \mathbf{n}$ ,  $i \in I$ ,  $j \notin I$ , then for any  $g \in \{E_{ij}^{\pm 1}, E_{ji}^{\pm 1}\}$  and  $v \in V^I$  we have

$$gv - v \in V^{I \cup \{j\} \setminus \{i\}} + V^{I \cup \{j\}}.$$

- (3) If  $I \subseteq \mathbf{n}$ ,  $i, j \in \mathbf{n}$  with  $i \neq j$ , then

$$F_{ij}V^I = V^{(i,j)I}$$

where  $(i, j)I$  is the image of  $I$  under the transposition  $(i, j)$ . In particular, if  $i \in I$  and  $j \notin I$ , then  $F_{ij}V^I = V^{I \cup \{j\} \setminus \{i\}}$ .

Informally speaking, the above definition generalizes some properties of the action of  $\mathrm{SL}_n(\mathbb{Z})$  on  $\mathrm{IA}_n^{ab}$  that we used in the proof of Theorem 3.1. More precisely, (1), (2) and (3) correspond to parts (b)(i), (a) and (b)(ii) of Observation 3.4, respectively. There are variations of (2) that we could consider; we have chosen condition (2) above primarily because it is preserved under many natural operations (see Lemma 5.2 below). A key reason why we assume the existence of a graded structure (rather than a filtered structure) in the above definition is that otherwise it seems hard to find a simple counterpart of (2).

The following result provides our starting examples of regular  $\mathrm{SL}_n(\mathbb{Z})$ -modules.



**Lemma 5.1.** *Let  $R$  be a commutative ring with 1,  $n \geq 2$ , let  $R^n$  be a free  $R$ -module of rank  $n$  with basis  $e_1, \dots, e_n$ , and let  $(R^n)^* = \text{Hom}_R(R^n, R)$  be the dual module, with dual basis  $e_1^*, \dots, e_n^*$ . Consider  $R^n$  and  $(R^n)^*$  as  $\text{SL}_n(\mathbb{Z})$ -modules with standard actions. Define  $(R^n)^{\{i\}} = Re_i$ ,  $((R^n)^*)^{\{i\}} = Re_i^*$  for  $1 \leq i \leq n$  and  $(R^n)^I = 0$ ,  $((R^n)^*)^I = 0$  for  $|I| \neq 1$ . Then with respect to these gradings  $R^n$  and  $(R^n)^*$  are regular of degree 1.*

*Proof.* The only non-obvious part is condition (2) in the definition of a regular  $\text{SL}_n(\mathbb{Z})$ -module. Condition (2) is vacuous if  $|I| \neq 1$ , so assume that  $I = \{i\}$  and  $j \neq i$ , in which case  $I \cup \{j\} \setminus \{i\} = \{j\}$  and  $I \cup \{j\} = \{i, j\}$ .

We have  $E_{ij}^{\pm 1}(re_i) = re_i$ ,  $E_{ji}^{\pm 1}(re_i) - re_i = \pm re_j \in (R^n)^{\{j\}}$ ,  $E_{ji}^{\pm 1}(re_i^*) = re_i^*$ ,  $E_{ij}^{\pm 1}(re_i^*) - re_i^* = \mp re_j^* \in ((R^n)^*)^{\{j\}}$ , so (ii) holds (note that in this case terms from  $V^{I \cup \{j\}}$  do not arise).  $\square$

**Lemma 5.2.** *The following hold:*

- (a) *Let  $V$  and  $W$  be regular  $\text{SL}_n(\mathbb{Z})$ -modules.*
  - (i) *Define  $(V \oplus W)^I = V^I \oplus W^I$ . Then  $V \oplus W$  is regular and  $\deg(V \oplus W) = \max\{\deg(V), \deg(W)\}$ .*
  - (ii) *Define  $(V \otimes_{\mathbb{Z}} W)^I = \sum_{I=I_1 \cup I_2} V^{I_1} \otimes W^{I_2}$  (here the union  $I_1 \cup I_2$  need not be disjoint). Then  $V \otimes_{\mathbb{Z}} W$  is regular and  $\deg(V \otimes_{\mathbb{Z}} W) \leq \deg(V) + \deg(W)$ .*
- (b) *Let  $V$  be a regular  $\text{SL}_n(\mathbb{Z})$ -module and  $U$  a submodule. For  $I \subseteq \mathbf{n}$  set  $U^I = U \cap V^I$  and  $(V/U)^I = V^I + U$ . Then  $V/U$  is always regular with  $\deg(V/U) \leq \deg(V)$ , and  $U$  is regular with  $\deg(U) \leq \deg(V)$  provided  $U = \sum_{I \subseteq \mathbf{n}} U^I$ .*
- (c) *Suppose that  $\text{SL}_n(\mathbb{Z})$  acts on an  $\mathbb{N}$ -graded Lie ring  $L = \bigoplus_{k=1}^{\infty} L(k)$  by graded automorphisms. Suppose that  $L$  is generated by  $L(1)$  (as a Lie ring) and that  $L(1)$  is a regular  $\text{SL}_n(\mathbb{Z})$ -module of degree  $d$ . For each  $k > 1$  and  $I \subseteq \mathbf{n}$  define*

$$(5.1) \quad L(k)^I = \sum_{I=I_1 \cup I_2 \cup \dots \cup I_k} [L(1)^{I_1}, L(1)^{I_2}, \dots, L(1)^{I_k}].$$

*Then  $L(k)$  is a regular  $\text{SL}_n(\mathbb{Z})$ -module of degree at most  $kd$ .*

- (d) <sup>4</sup> *In the setting of (c) assume in addition that  $[L(1)^I, L(1)^J] = 0$  whenever  $I$  and  $J$  are disjoint. Then for every  $m \in \mathbb{N}$  the degree of  $L(m)$  is at most  $m(d-1) + 1$ .*

*Proof.* Parts (a)(i) and (b) are straightforward.

(a)(ii) As in the proof of Lemma 5.1, we only need to check condition (2) in the definition of a regular module. So take  $I \subseteq \mathbf{n}$ ,  $i \in I$  and  $j \notin I$ . It is enough to check the condition for  $v$  of the form  $v = x \otimes y$  where  $x \in V^{I_1}$ ,  $y \in W^{I_2}$  and  $I_1 \cup I_2 = I$ . We will consider the case when  $i$  belongs to both  $I_1$  and  $I_2$  (the case when  $i$  only lies in one of those sets is analogous). Thus,  $I_1 = K_1 \cup \{i\}$  and  $I_2 = K_2 \cup \{i\}$  with  $i \notin K_1, K_2$ .

Let  $g \in \{E_{ij}^{\pm 1}, E_{ji}^{\pm 1}\}$ . Since  $V$  and  $W$  are regular, we have  $gx = x + x_1 + x_2$  and  $gy = y + y_1 + y_2$  where  $x_1 \in V^{K_1 \cup \{j\}}$ ,  $x_2 \in V^{K_1 \cup \{i, j\}}$ ,  $y_1 \in W^{K_2 \cup \{j\}}$ ,  $y_2 \in W^{K_2 \cup \{i, j\}}$ . Then  $x_1 \otimes y_1 \in (V \otimes W)^{K_1 \cup K_2 \cup \{j\}} = (V \otimes W)^{I \cup \{j\} \setminus \{i\}}$  and each of the 7 terms  $x \otimes y_1$ ,  $x \otimes y_2$ ,  $x_1 \otimes y$ ,  $x_2 \otimes y$ ,  $x_2 \otimes y_1$  and  $x_2 \otimes y_2$  lies in  $(V \otimes W)^{K_1 \cup K_2 \cup \{i, j\}} = (V \otimes W)^{I \cup \{j\}}$ . Hence  $gv - v = gx \otimes gy - x \otimes y \in (V \otimes W)^{I \cup \{j\} \setminus \{i\}} + (V \otimes W)^{I \cup \{j\}}$ , as desired.

<sup>4</sup>This property as well as its application in the proof of Lemma 6.4(b) have been discovered during the writing of the follow-up paper [CEP].

(c) Consider the map  $\varphi : L(1)^{\otimes k} \rightarrow L(k)$  given by  $\varphi(v_1 \otimes \dots \otimes v_k) = [v_1, \dots, v_k]$  (where the commutator on the right-hand side is left-normed). Since  $\mathrm{SL}_n(\mathbb{Z})$  acts on  $L$  by graded automorphisms,  $\varphi$  is a homomorphism of  $\mathrm{SL}_n(\mathbb{Z})$ -modules, and since  $L$  is generated by  $L(1)$ ,  $\varphi$  is surjective. Thus,  $L(k)$  is a quotient of  $L(1)^{\otimes k}$  as an  $\mathrm{SL}_n(\mathbb{Z})$ -module, and the grading (5.1) coincides with the quotient grading defined in (b). Thus, (c) follows from (a)(ii) and (b).

(d) follows from the proof of (c) combined with the observation that  $[L(1)^{I_1}, \dots, L(1)^{I_m}] = 0$  unless  $I_k$  has non-empty intersection with  $I_1 \cup I_2 \cup \dots \cup I_{k-1}$  for all  $2 \leq k \leq m$ , and the latter condition implies that  $|\cup I_i| \leq (\sum |I_i|) - (m - 1)$ .  $\square$

## 6. THE MAIN TECHNICAL THEOREM

Given a group  $G$ , let  $L(G) = \bigoplus_{i=1}^{\infty} \gamma_i G / \gamma_{i+1} G$ . The graded abelian group  $L(G)$  has a natural structure of a graded Lie ring with the bracket on homogeneous elements defined by

$$[g\gamma_{i+1}G, h\gamma_{j+1}G] = [g, h]\gamma_{i+j+1}G \text{ for all } g \in \gamma_i G \text{ and } h \in \gamma_j G,$$

where  $[g, h] = g^{-1}h^{-1}gh$ . The bracket operation is well defined since  $[\gamma_i G, \gamma_j G] \subseteq \gamma_{i+j} G$ . Moreover,  $\gamma_{i+1}G = [\gamma_i G, G]$  for each  $i$ , which implies that  $L(G)$  is generated as a Lie ring by its degree 1 component  $G^{ab} = G/\gamma_2 G$ .

Suppose now we are given another group  $\Gamma$  which contains  $G$  as a normal subgroup and a homomorphism  $\varphi : \mathrm{SL}_n(\mathbb{Z}) \rightarrow \Gamma/G$ . Define the action of  $\mathrm{SL}_n(\mathbb{Z})$  on  $L(G)$  by composing  $\varphi$  with the conjugation action of  $\Gamma/G$  on  $L(G)$ . It is easy to see that the obtained action is by Lie ring automorphisms and explicitly given by

$$(6.1) \quad g \cdot (x\gamma_{i+1}G) = \left( \widehat{\varphi(g)} x \widehat{\varphi(g)}^{-1} \right) \gamma_{i+1}G \text{ for all } g \in \mathrm{SL}_n(\mathbb{Z}), \quad i \in \mathbb{N} \text{ and } x \in \gamma_i G$$

(where  $\widehat{\varphi(g)}$  is any lift of  $\varphi(g)$  to  $\Gamma$ ).

The following is our main technical result which will be used to prove Theorems 1.3 and Theorem 1.4.

**Theorem 6.1.** *Let  $d, n \in \mathbb{N}$ , and let  $G$  be a finitely generated group with the structure of an  $\mathbf{n}$ -group generated in degree  $d$ . Let  $\Gamma$  and  $\varphi$  be as above, and consider the following four conditions:*

- (i)  $G$  is partially commuting (as defined in § 4.1).
- (ii)  $L(1) = G/[G, G] = G^{ab}$  has the structure of a regular  $\mathrm{SL}_n(\mathbb{Z})$ -module of degree at most  $d$  (with respect to the action of  $\mathrm{SL}_n(\mathbb{Z})$  given by (6.1)).
- (iii) For every  $I \subseteq \mathbf{n}$ , with  $|I| \geq d$ , the image of  $G_I$  in  $G^{ab} = L(1)$  contains  $\sum_{J \subseteq I} L(1)^J$ .
- (iv)  $G$  is weakly commuting (as defined in § 4.1).

Let  $N \in \mathbb{N}$ , and assume that either

- (i)-(iii) hold and  $n \geq d(d+1)(N-1)$  or
- (i)-(iv) hold and  $n \geq (d+1) + (d^2-1)(N-1)$ .

If  $\rho$  is any non-trivial homomorphism from  $G$  to another group with  $\mathrm{Ker}(\rho) \supseteq \gamma_N G$ , then  $G$  has a  $\rho$ -CG sequence.

**Remark.** Since  $L(1)^J = 0$  for  $|J| > d$  (by (ii)), it suffices to know (iii) when  $|I| = d$ .

Before proving Theorem 6.1, we shall use it to prove Theorem 1.4(b). Theorem 1.4(b) is a direct consequence of Proposition 2.4 and the following result.

**Theorem 6.2.** *Let  $N \geq 2$  and  $n \geq 8N - 4$  be integers and  $G = \text{IA}_n$ . If  $\rho$  is a non-trivial homomorphism from  $G$  to another group with  $\text{Ker } \rho \supseteq \gamma_N G$ , then  $G$  has a  $\rho$ -CG sequence.*

*Proof.* We just need to check that the pair  $(G, N)$  satisfies conditions (i)-(iv) in Theorem 6.1 for  $d = 3$ . We shall use the  $\mathbf{n}$ -group structure on  $\text{IA}_n$  defined in § 4.1. Recall that this  $\mathbf{n}$ -group is commuting and generated in degree 3, so (i) and (iv) hold. Also recall that  $L(1) = G^{ab} \cong V^* \otimes (V \wedge V)$  as  $\text{SL}_n(\mathbb{Z})$ -modules where  $V = \mathbb{Z}^n$  with the standard  $\text{SL}_n(\mathbb{Z})$ -action. Using Lemmas 5.1 and 5.2 we deduce that  $L(1)$  is a regular  $\text{SL}_n(\mathbb{Z})$ -module of degree at most 3, so (ii) holds. Moreover, if  $\{e_i\}$  is the standard basis of  $V$  and  $\{e_i^*\}$  the dual basis, then for every  $I \subseteq \mathbf{n}$  we have

$$L(1)^I = \sum_{I=\{i,j,k\}} \mathbb{Z}e_i^* \otimes e_j \wedge e_k \text{ where } i, j, k \text{ need not be distinct.}$$

(so in particular  $L(1)^I = 0$  if  $|I| \neq 2, 3$ ). Let us now take any  $I$  with  $|I| \geq 3$ . The subgroup  $G_I$  contains the element  $K_{ij}$  for any distinct  $i, j \in I$  and  $K_{ijk}$  for any distinct  $i, j, k \in I$ . Since  $K_{ij}$  and  $K_{ijk}$  project to  $e_i^* \otimes (e_i \wedge e_j)$  and  $e_i^* \otimes (e_j \wedge e_k)$ , respectively, we conclude that (iii) holds as well.  $\square$

Note that Theorem 6.2 yields not only Theorem 1.4(b), but also an alternative proof of Theorem 1.3 for  $G = \text{IA}_n$ ,  $n \geq 12$ . By the same logic, to prove Theorem 1.3 for the Torelli group and Theorem 1.4(c), we need to check conditions (i)-(iv) of Theorem 6.1 for  $G = \mathcal{I}_n^1$  and  $d = 3$ . This verification requires more work and will be postponed till § 7.

The remainder of this section will be devoted to the proof of Theorem 6.1.

**6.1. Proof of Theorem 6.1.** Throughout this subsection we fix the notations introduced in Theorem 6.1. As before, we set  $L(k) = \gamma_k G / \gamma_{k+1} G$  for all  $k \in \mathbb{N}$ .

Take any non-trivial homomorphism  $\rho$  from  $G$  to another group with  $\text{Ker } \rho \supseteq \gamma_N G$ . Let  $m$  be the nilpotency class of  $\rho(G)$ , so that  $1 \leq m \leq N - 1$ .

The proof of Theorem 6.1 will start with a sequence of simple reductions; in fact, some steps in that sequence have already been completed in § 4.2 when we established Claim 4.2. We will eventually reduce Theorem 6.1 to Proposition 6.8, a combinatorial statement about the regular  $\text{SL}_n(\mathbb{Z})$ -module  $L(m)$ , and it is only the proof of that proposition where the definition of a regular  $\text{SL}_n(\mathbb{Z})$ -module will be fully used. The main steps in the proof of Theorem 6.1 are indicated in the following diagram:

$$\text{Proposition 6.8} \implies \text{Claim 6.6} \xrightarrow{\text{Lemma 6.7}} \text{Claim 6.5} \xrightarrow{\text{Claim 4.2}} \text{Theorem 6.1.}$$

As in § 4.2, for each  $k \in \mathbb{N}$  let  $L(k)_I$  be the image of  $\gamma_k G \cap G_I$  in  $L(k)$ . Let  $\{L(k)^I\}$  be the  $\mathbf{n}$ -grading on  $L(k)$  obtained from the grading  $\{L(1)^I\}$  via (5.1) (see Lemma 5.2(c)).

Condition (iii) of Theorem 6.1 simply says that  $L(1)^J \subseteq L(1)_I$  whenever  $J \subseteq I$  and  $|I| \geq d$ . We claim that the inclusion remains true if we replace  $L(1)$  by  $L(m)$ :

**Lemma 6.3.** *For any subsets  $J \subseteq I \subseteq \mathbf{n}$  with  $|I| \geq d$  we have  $L(m)^J \subseteq L(m)_I$ .*

*Proof.* By definition  $L(m)^J$  is spanned by (some, not necessarily all) elements of the form  $[x_1, \dots, x_m]$  where  $x_k \in L(1)^{J_k}$  for some  $J_k \subseteq J$  (and thus  $J_k \subseteq I$ ). Since  $|I| \geq d$ , condition (iii) implies that  $x_k \in L(1)_I$ , so there exist elements  $\{g_k \in G_I\}_{k=1}^d$  such that  $x_k = g_k[G, G]$ . Then  $[x_1, \dots, x_m]$  is the image in  $L(m)$  of the element  $[g_1, \dots, g_m]$  which lies in  $\gamma_m G_I \subseteq \gamma_m G \cap G_I$ , so by definition  $[x_1, \dots, x_m] \in L(m)_I$ .  $\square$

Now set  $M = 1 + m(d - 1)$  if  $G$  is weakly commuting and  $M = md$  otherwise.

**Lemma 6.4.** *The following hold:*

- (a)  $n \geq (d+1)M$ ;
- (b)  $L(m)$  is a regular  $\mathrm{SL}_n(\mathbb{Z})$ -module of degree at most  $M$ .

*Proof.* (a) holds by the hypotheses on  $n$  in Theorem 6.1 (since  $m \leq N-1$ ).

(b) Recall that  $L(1)$  is a regular  $\mathrm{SL}_n(\mathbb{Z})$ -module of degree at most  $d$ , so by Lemma 5.2(c),  $L(m)$  is regular of degree at most  $md$ .

Suppose now that  $G$  is weakly commuting, and let  $I, J \subseteq \mathbf{n}$  be disjoint. By definition there exists  $g \in G$  such that  $G_I^g$  and  $G_J$  commute. Since for every  $x \in G$  the elements  $x$  and  $x^g = x[x, g]$  have the same image in  $G^{ab} = L(1)$ , it follows that the images of  $G_I$  and  $G_J$  in  $L(1)$  commute, that is,

$$[L(1)_I, L(1)_J] = 0. \quad (***)$$

We claim that  $[L(1)^I, L(1)^J] = 0$  for any disjoint  $I, J$ . Once this is proved, we deduce from Lemma 5.2(d) that  $L(m)$  has degree at most  $1 + m(d-1)$ .

So let  $I, J \subseteq \mathbf{n}$  be disjoint. If  $|I| > d$  or  $|J| > d$ , then  $L(1)^I = 0$  or  $L(1)^J = 0$ , and there is nothing to prove. If  $|I| \leq d$  and  $|J| \leq d$ , we can choose disjoint  $I'$  and  $J'$  with  $|I'| = |J'| = d$ ,  $I \subseteq I'$  and  $J \subseteq J'$  (this is possible since  $n \geq d^2 + d \geq 2d$ ). Then  $[L(1)_{I'}, L(1)_{J'}] = 0$  by  $(***)$ ; on the other hand  $L(1)^I \subseteq L(1)_{I'}$  and  $L(1)^J \subseteq L(1)_{J'}$  by condition (iii) in Theorem 6.1 and hence  $[L(1)^I, L(1)^J] = 0$ .  $\square$

As in § 4.2, let  $U$  be any abelian group containing a copy of every finitely generated abelian group and  $\Lambda_m = \mathrm{Hom}_{\mathbb{Z}}(L(m), U)$ . Define the action of  $\mathrm{SL}_n(\mathbb{Z})$  on  $\Lambda_m$  by

$$(g\lambda)(x) = \lambda(g^{-1} \cdot x) \text{ for all } g \in \mathrm{SL}_n(\mathbb{Z}), \lambda \in \Lambda_m \text{ and } x \in L(m)$$

where the action of  $\mathrm{SL}_n(\mathbb{Z})$  on  $L(m)$  is given by (6.1). Equivalently, this action of  $\mathrm{SL}_n(\mathbb{Z})$  on  $\Lambda_m$  is the composition of  $\varphi : \mathrm{SL}_n(\mathbb{Z}) \rightarrow \Gamma/G$  with the action of  $\Gamma/G$  on  $\Lambda_m$  given by (4.2). In particular, for every  $\lambda \in \Lambda_m$  its  $\mathrm{SL}_n(\mathbb{Z})$ -orbit is contained in its  $\Gamma/G$ -orbit.

Thus, by Claim 4.2, Theorem 6.1 is implied by the following statement:

**Claim 6.5.** *For any nonzero  $\lambda \in \Lambda_m$  there exists an element  $\mu$  in the  $\mathrm{SL}_n(\mathbb{Z})$ -orbit of  $\lambda$  and disjoint subsets  $I_1, \dots, I_{d+1}$  each of which consists of consecutive integers such that  $\mu$  does not vanish on  $L(m)_{I_k}$  for each  $k$ .*

**Definition.** Let  $\lambda \in \Lambda_m$ . Define  $\mathrm{supp}(\lambda)$ , the *support* of  $\lambda$ , to be the set of all subsets  $I \subseteq \mathbf{n}$  such that  $\lambda$  does not vanish on  $L(m)^I$ .

Here is our next reduction.

**Claim 6.6.** *For any nonzero  $\lambda \in \Lambda_m$  the  $\mathrm{SL}_n(\mathbb{Z})$ -orbit of  $\lambda$  contains an element  $\mu$  such that  $\mathrm{supp}(\mu)$  contains  $d+1$  disjoint subsets.*

The following lemma shows that Claim 6.5 follows from Claim 6.6.

**Lemma 6.7.** *Take any  $\lambda \in \Lambda_m$  whose support contains  $d+1$  disjoint subsets. Then the  $\mathrm{SL}_n(\mathbb{Z})$ -orbit of  $\lambda$  contains an element  $\mu$  satisfying the conclusion of Claim 6.5.*

*Proof.* For  $1 \leq k \leq d+1$  define

$$I_k = \{(k-1)M+1, (k-1)M+2, \dots, kM\}.$$

The sets  $I_1, \dots, I_{d+1}$  are disjoint and contained in  $\mathbf{n}$  since  $n \geq (d+1)M$ .

Let  $J_1, \dots, J_{d+1}$  be disjoint subsets in  $\mathrm{supp}(\lambda)$ . By Lemma 6.4 we have  $|J_k| \leq M = |I_k|$ . Hence there exists a permutation  $\sigma \in S_n$  such that  $\sigma(J_k) \subseteq I_k$  for each  $k$ . Write  $\sigma$  as

a product of transpositions  $\sigma = \prod (i_t, j_t)$  and let  $g = \prod F_{i_t j_t}$ . By condition (3) in the definition of a regular module we have

$$(g\lambda)(L(m)^{\sigma(J_k)}) = \lambda(g^{-1} \cdot L(m)^{\sigma(J_k)}) = \lambda(L(m)^{\sigma^{-1}(\sigma(J_k))}) = \lambda(L(m)^{J_k}) \neq 0.$$

Since  $|I_k| = M \geq d$  and  $\sigma(J_k) \subseteq I_k$ , Lemma 6.3 implies that  $g\lambda$  does not vanish on  $L(m)_{I_k}$ . Since  $I_1, \dots, I_{d+1}$  consist of consecutive integers,  $\mu = g\lambda$  has the desired property.  $\square$

Claim 6.6 easily follows from Proposition 6.8 below.

**Definition.** Let  $0 \neq \lambda \in \Lambda_m$  and  $s \in \mathbb{N}$ . If  $I_1, \dots, I_s$  are elements of  $\text{supp}(\lambda)$ , define

$$D(I_1, \dots, I_s) = \sum_{i=1}^s |I_i| - \sum_{i \neq j} |I_i \cap I_j|.$$

The maximum possible value of  $D(I_1, \dots, I_s)$  will be denoted by  $D_s(\lambda)$ , and any  $s$ -tuple in  $\text{supp}(\lambda)$  on which this maximum is achieved will be called *maximally disjoint* for  $\lambda$ .

**Proposition 6.8.** Fix  $0 \neq \lambda \in \Lambda_m$  and  $s \in \mathbb{N}$ . Let  $\mu$  be an element in the  $\text{SL}_n(\mathbb{Z})$ -orbit of  $\lambda$  such that  $D_s(\mu)$  is maximum possible, and let  $I_1, \dots, I_s \in \text{supp}(\mu)$  be maximally disjoint for  $\mu$ . Then one of the following holds:

- (i)  $I_1, \dots, I_s$  are disjoint;
- (ii)  $\cup I_i = \mathbf{n}$ .

Before proving Proposition 6.8, we shall deduce Claim 6.6 from it.

*Proof of Claim 6.6.* Let us apply Proposition 6.8 to  $\lambda$  (from the statement of Claim 6.6) and  $s = d+1$ , and let  $I_1, \dots, I_{d+1}$  be as in the statement of Proposition 6.8. By Lemma 6.4 we have  $|I_k| \leq M \leq \frac{n}{d+1}$ . If  $I_1, \dots, I_{d+1}$  are not disjoint, then  $|\cup_{i=1}^{d+1} I_i| < (d+1) \cdot M \leq n$ , so  $\cup I_i \neq \mathbf{n}$ , which contradicts Proposition 6.8.  $\square$

We will now prove Proposition 6.8, thereby completing the proof of Theorem 6.1.

*Proof of Proposition 6.8.* Assume that neither (i) nor (ii) holds. Without loss of generality we can assume that  $|I_1 \cap I_2| \neq \emptyset$ , and choose  $i \in I_1 \cap I_2$  and  $j \notin \cup_{l=1}^s I_l$  (with  $j \in \mathbf{n}$ ). Since  $(I_1, \dots, I_s)$  is maximally disjoint for  $\mu$ , its support  $\text{supp}(\mu)$  contains neither  $I_k \cup \{j\} \setminus \{i\}$  nor  $I_k \cup \{j\}$  for any  $k$ . Indeed, if  $\text{supp}(\mu)$  contained one of those sets, replacing  $I_k$  by  $I_k \cup \{j\} \setminus \{i\}$  or  $I_k \cup \{j\}$  in  $(I_1, \dots, I_s)$  would produce an  $s$ -tuple in  $\text{supp}(\mu)$  with a larger value of the function  $D$ , which is impossible.

Now take any  $g \in \{E_{ij}^{\pm 1}, E_{ji}^{\pm 1}\}$ . We claim that  $\text{supp}(g\mu)$  contains  $I_k$  for any  $k$ . Indeed, by assumption  $I_k \in \text{supp}(\mu)$ , so  $\mu(x) \neq 0$  for some  $x \in L(m)^{I_k}$ . We know that  $j \notin I_k$ . If  $i \in I_k$ , then by condition (2) in the definition of a regular module,

$$g^{-1}x - x \in L(m)^{I_k \cup \{j\} \setminus \{i\}} + L(m)^{I_k \cup \{j\}}.$$

The latter containment is also true if  $i \notin I_k$  (in which case  $g^{-1}x = x$  by condition (1)). Since  $\text{supp}(\mu)$  contains neither  $I_k \cup \{j\} \setminus \{i\}$  nor  $I_k \cup \{j\}$ , we have  $\mu(g^{-1}x - x) = 0$ . Hence  $(g\mu)(x) = \mu(g^{-1}x) = \mu(x) \neq 0$ , so  $I_k \in \text{supp}(g\mu)$ .

Thus  $\text{supp}(g\mu)$  contains  $I_1, \dots, I_s$ . Since  $(I_1, \dots, I_s)$  is maximally disjoint for  $\mu$ , we must have  $D_s(g\mu) \geq D(I_1, \dots, I_s) = D_s(\mu)$ . On the other hand,  $D_s(g\mu) \leq D_s(\mu)$  by the choice of  $\mu$ , so  $D_s(g\mu) = D_s(\mu)$  and  $(I_1, \dots, I_s)$  is also maximally disjoint for  $g\mu$ .

Applying the same argument to  $E_{ij}\mu$ , we conclude that  $(I_1, \dots, I_s)$  is also maximally disjoint for  $E_{ji}^{-1}E_{ij}\mu$  and likewise maximally disjoint for  $E_{ij}E_{ji}^{-1}E_{ij}\mu$ . In particular,  $I_1 \cup$

$\{j\} \setminus \{i\} \notin \text{supp}(E_{ij}E_{ji}^{-1}E_{ij}\mu)$ . On the other hand,  $E_{ij}E_{ji}^{-1}E_{ij} = F_{ij}$ , and condition (3) in the definition of a regular module implies that  $\text{supp}(F_{ij}\mu)$  contains  $(i, j)I_1 = I_1 \cup \{j\} \setminus \{i\}$ , a contradiction.  $\square$

## 7. VERIFYING THE HYPOTHESES OF THEOREM 6.1 FOR THE TORELLI GROUPS.

Let  $g \geq 0$ , and let  $\Sigma_g^1$  be an orientable surface of genus  $g$  with 1 boundary component. The mapping class group  $\text{Mod}_g^1 = \text{Mod}(\Sigma_g^1)$  is defined as the subgroup of orientation preserving homeomorphisms of  $\Sigma_g^1$  which fix the boundary  $\partial\Sigma_g^1$  pointwise modulo the isotopies which fix  $\partial\Sigma_g^1$  pointwise.

Choose a base point  $p_0 \in \partial\Sigma_g^1$ . Since  $\Sigma_g^1$  has a bouquet of  $2g$  circles as a deformation retract, the fundamental group  $\pi = \pi_1(\Sigma_g^1, p_0)$  is free of rank  $2g$ . The action of  $\text{Mod}_g^1$  on  $\Sigma_g^1$  induces an action on  $\pi$  and thus we obtain a homomorphism  $\iota : \text{Mod}_g^1 \rightarrow \text{Aut}(\pi)$ . It is well known that  $\iota$  is injective and  $\text{Im } \iota$  is equal to the stabilizer of the element  $\prod_{i=1}^g [x_{2i-1}, x_{2i}] \in \pi$ , for a suitable choice of basis  $\{x_1, \dots, x_{2g}\}$  of  $\pi$ . For each  $k \in \mathbb{N}$  define  $\mathcal{I}_g^1(k)$  to be the kernel of the induced map  $\text{Mod}_g^1 \rightarrow \text{Aut}(\pi/\gamma_{k+1}\pi)$ .

The filtration  $\{\mathcal{I}_g^1(k)\}_{k=1}^\infty$  is called the *Johnson filtration*, and the subgroups  $\mathcal{I}_g^1 = \mathcal{I}_g^1(1)$  and  $\mathcal{K}_g^1 = \mathcal{I}_g^1(2)$  are known as the *Torelli subgroup* and the *Johnson kernel*, respectively. Note that  $\mathcal{I}_g^1$  can also be defined as the set of all elements of  $\text{Mod}_g^1$  which act trivially on the integral homology group  $H_1(\Sigma_g^1)$ .

For the rest of this section we set  $\Sigma = \Sigma_g^1$ ,  $\mathcal{M} = \text{Mod}_g^1$ ,  $\mathcal{I} = \mathcal{I}_g^1 = \mathcal{I}_g^1(1)$  and  $\mathcal{K} = \mathcal{K}_g^1 = \mathcal{I}_g^1(2)$ . Starting in § 7.3, we will assume that  $g \geq 3$ .

Our goal is to verify the hypotheses of Theorem 6.1 for  $\Gamma = \mathcal{M}$ ,  $G = \mathcal{I}$ ,  $n = g$  and  $d = 3$ . Combined with Proposition 2.4, this will prove Theorem 1.3 for the Torelli group and Theorem 1.4(c) and thus will complete the proofs of Theorems 1.3 and 1.4.

The homomorphism  $\varphi : \text{SL}_g(\mathbb{Z}) \rightarrow \mathcal{M}/\mathcal{I}$  to be used in Theorem 6.1 will be defined in § 7.5 (assuming the canonical isomorphism  $\mathcal{M}/\mathcal{I} \cong \text{Sp}(V)$  defined in § 7.2).

**7.1. The  $\mathbf{g}$ -group structure on  $\mathcal{M}$ .** We start by recalling the construction from [CP, §4.1]. It will be convenient to think of  $\Sigma$  as a (closed) disk with  $g$  handles attached; call the handles  $H_1, \dots, H_g$ . Choose disjoint subsurfaces  $X_1, \dots, X_g$ , each homeomorphic to  $\Sigma_1^1$ , such that  $X_i$  contains  $H_i$ , and let  $Y = \Sigma \setminus \bigcup_{i=1}^g \text{Int}(X_i)$  where  $\text{Int}(X_i)$  is the interior of  $X_i$ . Fix a point  $p \in \text{Int}(Y)$ , and for each  $1 \leq i \leq g$  choose a simple path  $\delta_i$  in  $Y$  which connects  $p$  to a point on  $\partial X_i$  such that the following hold (see Figure 1):

- (i) the paths  $\delta_i$  are disjoint apart from  $p$ ;
- (ii) if  $C$  is a small circle around  $p$ , then the intersection points  $C \cap \delta_1, \dots, C \cap \delta_g$  appear in the clockwise order.

Now for each  $I \subseteq \mathbf{g}$  define  $S_I$  to be a closed regular neighborhood of  $\bigcup_{i \in I} (\delta_i \cup X_i)$ . Then  $S_I$  is well defined up to isotopy (see Figure 2).

The following properties (1)-(4) are straightforward. Given two subsurfaces  $R$  and  $S$ , we will say that certain relation between  $R$  and  $S$  holds up to isotopy if there exist subsurfaces isotopic to  $R$  and  $S$ , respectively, satisfying the stated relation:

- (1)  $S_{\mathbf{g}}$  is isotopic to  $\Sigma$ ;
- (2) if  $I \subseteq J$ , then  $S_I$  is contained in  $S_J$  up to isotopy;

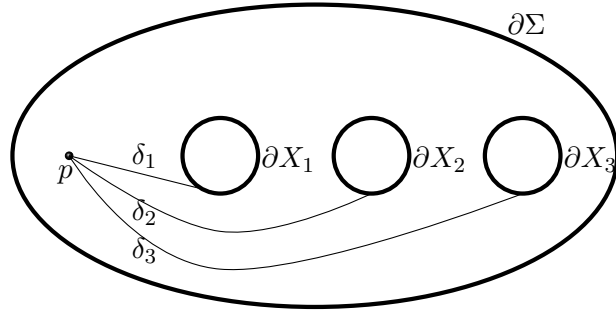
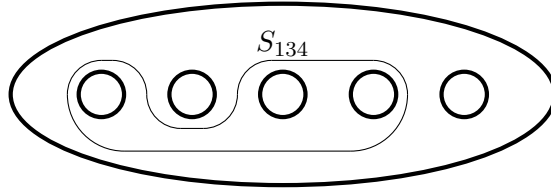


FIGURE 1.

FIGURE 2. The surface  $S_{134}$  perturbed by an isotopy (only part of the surfaces inside  $Y$  is shown)

- (3) if  $I$  and  $J$  are disjoint and  $I$  consists of consecutive integers, then  $S_I$  and  $S_J$  are disjoint up to isotopy;
- (4) for any disjoint  $I, J$  there exists a subsurface  $S'_I$  homeomorphic to  $S_I$  which is disjoint from  $S_J$  and such that  $H_1(S_I)$  and  $H_1(S'_I)$  have the same image in  $H_1(\Sigma)$ .

Now define  $\mathcal{M}_I$  to be the subgroup of  $\mathcal{M}$  consisting of mapping classes which have a representative supported on  $S_I$ , and let  $\mathcal{I}_I = \mathcal{M}_I \cap \mathcal{I}$ . Properties (1)-(3) above clearly imply that  $\mathcal{M}$  and  $\mathcal{I}$  are partially commuting  $\mathbf{g}$ -groups (here we use the obvious fact that homeomorphisms supported on disjoint surfaces commute). The fact that  $\mathcal{I}$  is weakly commuting follows from (4) and [CP, Lemma 4.2(ii)].

A deep result of Church and Putman [CP, Proposition 4.5] based on earlier work of Putman [Pu2] asserts that if  $g \geq 3$ , then  $\mathcal{I} = \langle \mathcal{I}_I : |I| = 3 \rangle$ , so  $\mathcal{I}$  is generated in degree 3 as a  $\mathbf{g}$ -group.

In order to check the remaining conditions in Theorem 6.1, namely conditions (ii) and (iii), we need some preparations.

**7.2.  $H_1(\Sigma)$  as an  $\mathrm{Sp}_{2g}(\mathbb{Z})$ -module.** Let  $V = H_1(\Sigma)$ . Then  $V$  is a free abelian group of rank  $2g$  endowed with the canonical symplectic form

$$([\alpha], [\beta]) \mapsto [\alpha] \cdot [\beta]$$

where  $[\alpha] \cdot [\beta]$  is the algebraic intersection number between the closed curves  $\alpha$  and  $\beta$  on  $\Sigma$ . Clearly the action of  $\mathcal{M}/\mathcal{I}$  on  $V$  preserves this form, so there is a canonical group homomorphism  $\mathcal{M}/\mathcal{I} \rightarrow \mathrm{Sp}(V)$  where  $\mathrm{Sp}(V)$  is the group of automorphisms of  $V$  preserving the above form. It is well known that this homomorphism is an isomorphism, so from now on we will identify  $\mathcal{M}/\mathcal{I}$  with  $\mathrm{Sp}(V)$ .

For the rest of this section we will assume that  $g \geq 3$ .

In the next two subsections (§ 7.3, § 7.4) we will describe the abelianization  $\mathcal{I}^{ab}$  as an  $\mathrm{Sp}(V)$ -module. This description was obtained by Johnson in [Jo6] using earlier works of Birman and Craggs [BC] and Johnson [Jo2, Jo3].

**7.3. Johnson and Birman-Craggs-Johnson homomorphisms.** Before describing  $\mathcal{I}^{ab}$ , we need to introduce two smaller abelian quotients of  $\mathcal{I}$ , the largest torsion-free quotient and the largest quotient of exponent 2. These quotients are the images of the *Johnson homomorphism* (denoted by  $\tau$ ) and the *Birman-Craggs-Johnson homomorphism* (denoted by  $\sigma$ ), respectively. We refer the reader to Johnson's papers [Jo3] and [Jo2] for the conceptual definition of  $\tau$  and  $\sigma$ . For our purposes it will be sufficient to know the values of  $\tau$  and  $\sigma$  on the generating set for  $\mathcal{I}$  consisting of 1-BP maps.

For each simple closed curve  $\gamma$  on  $\Sigma$  denote by  $T_\gamma \in \mathcal{M}$  the Dehn twist about  $\gamma$ .

**Definition.** Let  $(\gamma, \delta)$  be a pair of disjoint non-separating simple closed curves on  $\Sigma$ . It is called a *bounding pair of genus  $k$*  if

- (i)  $\gamma$  and  $\delta$  are homologous to each other and non-homologous to zero;
- (ii) the union  $\gamma \cup \delta$  separates  $\Sigma$ ; moreover, if  $\Sigma_{\gamma, \delta}$  is the connected component of  $\Sigma \setminus (\gamma \cup \delta)$  which does not contain  $\partial\Sigma$ , then  $\Sigma_{\gamma, \delta}$  has genus  $k$ .

If  $(\gamma, \delta)$  is a bounding pair of genus  $k$ , then the map  $T_\gamma T_\delta^{-1}$  always lies in  $\mathcal{I}$  and is called a  *$k$ -BP map*.

Johnson [Jo1] proved that  $\mathcal{I}$  is generated by 1-BP maps.

The *Johnson homomorphism* is a surjective homomorphism  $\tau : \mathcal{I} \rightarrow \wedge^3 V$  with  $\mathrm{Ker}(\tau) = \mathcal{K}$  (the Johnson kernel) introduced in [Jo3]. The induced map  $\mathcal{I}^{ab} \rightarrow \wedge^3 V$  (which we will also denote by  $\tau$ ) is a homomorphism of  $\mathrm{Sp}(V)$ -modules. In [Jo6], Johnson proved that  $\mathcal{I}/\mathcal{K}$  is the largest torsion-free abelian quotient of  $\mathcal{I}$ .

We will now give an explicit formula for the values of  $\tau$  on 1-BP maps. Let  $(\gamma, \delta)$  be a bounding pair of genus 1, and assume that  $\gamma$  is oriented in such a way that  $\Sigma_{\gamma, \delta}$  is on its left. Let  $c \in V = H_1(\Sigma)$  be the homology class of  $\gamma$ , and choose any  $a, b \in H_1(\Sigma_{\gamma, \delta}) \subset H_1(\Sigma)$  such that  $a \cdot b = 1$ . Then

$$(7.1) \quad \tau(T_\gamma T_\delta^{-1}) = a \wedge b \wedge c.$$

For the justification of (7.1) and the related formula (7.2) below see [Jo6, § 2] and references therein.

We now turn to the Birman-Craggs-Johnson homomorphism. Let  $V_{\mathbb{F}_2} = V \otimes \mathbb{F}_2 \cong H_1(\Sigma, \mathbb{F}_2)$ . Let  $B$  be the ring of polynomials over  $\mathbb{F}_2$  in formal variables  $X = \{\bar{v} : v \in H_{\mathbb{F}_2} \setminus \{0\}\}$  subject to relations

- (R1)  $\overline{v+w} = \bar{v} + \bar{w} + v \cdot w$  for all  $\bar{v}, \bar{w} \in X$
- (R2)  $\bar{v}^2 = \bar{v}$  for all  $\bar{v} \in X$ .

The group  $\mathrm{Sp}(V)$  has a natural action on  $B$  by ring automorphisms such that  $g(\bar{v}) = \bar{g}v$  for all  $\bar{v} \in X$ . Let  $B_n$  be the subspace of  $B$  consisting of elements representable by a polynomial of degree at most  $n$ . Then each  $B_n$  is an  $\mathrm{Sp}(V)$ -submodule, and it is easy to see that  $B_n/B_{n-1} \cong \wedge^n V_{\mathbb{F}_2}$  for each  $n \geq 1$  and  $B_0 \cong \wedge^0 V_{\mathbb{F}_2}$  (as  $\mathrm{Sp}(V)$ -modules).

In [Jo2], Johnson constructed a surjective homomorphism  $\sigma : \mathcal{I} \rightarrow B_3$  which induces an  $\mathrm{Sp}(V)$ -module homomorphism  $\mathcal{I}^{ab} \rightarrow B_3$ . Following [BF], we will refer to  $\sigma$  as the *Birman-Craggs-Johnson (BCJ) homomorphism*. Here is the formula for  $\sigma$  on 1-BP maps:

$$(7.2) \quad \sigma(T_\gamma T_\delta^{-1}) = \bar{a}\bar{b}(\bar{c} + 1).$$



where  $a, b, c$  are defined as in (7.1) except that this time they are mod 2 homology classes.

**7.4. The full abelianization.** Let  $\alpha : \wedge^3 V \rightarrow \wedge^3 V_{\mathbb{F}_2}$  be the natural reduction map, and let  $\beta : B_3 \rightarrow \wedge^3 V_{\mathbb{F}_2}$  be the unique linear map such that  $\beta(B_2) = 0$  and  $\beta(\bar{u} \bar{v} \bar{w}) = u \wedge v \wedge w$  for all  $u, v, w \in V_{\mathbb{F}_2}$ . Clearly  $\alpha$  and  $\beta$  are both  $\mathrm{Sp}(V)$ -module homomorphisms. Let

$$(7.3) \quad W = \{(u, v) \in \wedge^3 V \oplus B_3 : \alpha(u) = \beta(v)\}.$$

The following result is a reformulation of [Jo6, Thm 3].

**Theorem 7.1** ([Jo6]). *The map  $(\tau, \sigma) : \mathcal{I}^{ab} \rightarrow \wedge^3 V \oplus B_3$  given by  $g[G, G] \mapsto (\tau(g), \sigma(g))$  is injective and  $\mathrm{Im}((\tau, \sigma)) = W$ .*

Now let  $I \subseteq \mathbf{g}$  be any subset with  $|I| \geq 3$ , and let  $S_I$  be the corresponding subsurface of  $\Sigma$  introduced in § 7.1. Since  $S_I$  is itself a closed orientable surface of genus  $\geq 3$  with one boundary component, we can repeat the entire construction described in this section starting with  $S_I$  instead of  $\Sigma$ , so in particular we can define the modules  $V(I) = H_1(S_I)$  and  $B_3(I)$  and the homomorphisms  $\tau_I : \mathcal{I}_I \rightarrow \wedge^3 V(I)$  and  $\sigma_I : \mathcal{I}_I \rightarrow B_3(I)$  (recall that  $\mathcal{I}_I$  was defined as the subgroup of  $\mathcal{I}$  consisting of mapping classes supported on  $S_I$ , but this group is canonically isomorphic to the Torelli subgroup of  $\mathrm{Mod}(S_I)$  – see [Pu3, § 2.1]).

**Proposition 7.2.** *The following diagrams are commutative:*

$$(7.4) \quad \begin{array}{ccc} (\mathcal{I}_I)^{ab} & \xrightarrow{\tau_I} & \wedge^3 V(I) \\ \downarrow & & \downarrow \\ \mathcal{I}^{ab} & \xrightarrow{\tau} & \wedge^3 V \end{array} \quad \begin{array}{ccc} (\mathcal{I}_I)^{ab} & \xrightarrow{\sigma_I} & B_3(I) \\ \downarrow & & \downarrow \\ \mathcal{I}^{ab} & \xrightarrow{\sigma} & B_3 \end{array}$$

where the vertical maps are induced by the natural inclusions  $\mathcal{I}_I \rightarrow \mathcal{I}$  and  $H_1(S_I) \rightarrow H_1(\Sigma)$ .

*Proof.* For both diagrams, it is enough to check commutativity for the values on 1-BP maps. This follows immediately from (7.1) and (7.2) since if  $(\gamma, \delta)$  is a genus 1 bounding pair on  $S_I$ , it is also a genus 1 bounding pair on  $\Sigma$ , and replacing  $S_I$  by  $\Sigma$  does not change the surface  $\Sigma_{\gamma, \delta}$  or the homology classes  $a, b$  and  $c$  in (7.1) and (7.2).  $\square$

**7.5.  $\mathcal{I}^{ab}$  as a regular  $\mathrm{SL}_g(\mathbb{Z})$ -module.** For each  $1 \leq i \leq g$  choose any basis  $\{a_i, b_i\}$  for  $H_1(S_i) \subset V$  s.t.  $a_i \cdot b_i = 1$ . Then  $V = \bigoplus_{i=1}^g (\mathbb{Z}a_i \oplus \mathbb{Z}b_i)$ , and  $\{a_i, b_i\}_{i=1}^g$  is a symplectic basis for  $V$ , that is,  $a_i \cdot a_j = b_i \cdot b_j = 0$  for all  $i, j$ , and  $a_i \cdot b_j = \delta_{ij}$ . Now define  $\varphi : \mathrm{SL}_g(\mathbb{Z}) \rightarrow \mathrm{Sp}(V)$  by

$$(7.5) \quad \varphi(A) = \begin{pmatrix} A & 0 \\ 0 & (A^t)^{-1} \end{pmatrix},$$

where the above matrix is with respect to the ordered basis  $(a_1, \dots, a_g, b_1, \dots, b_g)$ . Thus, all the  $\mathrm{Sp}(V)$ -modules defined in this section can also be considered as  $\mathrm{SL}_g(\mathbb{Z})$ -modules.

Note that the map  $(\tau, \sigma) : \mathcal{I}^{ab} \rightarrow W$  from Theorem 7.1 is an isomorphism of  $\mathrm{SL}_g(\mathbb{Z})$ -modules, and the  $\mathrm{SL}_g(\mathbb{Z})$ -module structure on  $\mathcal{I}^{ab}$  coincides with the one given by (6.1) with  $\varphi$  as above (recall that we identified  $\mathrm{Sp}(V)$  with  $\mathcal{M}/\mathcal{I}$ ).

We could proceed working directly with  $W$ , but things can be simplified further thanks to the following observations. Let  $\pi : B_3 \rightarrow \bigoplus_{i=0}^3 \wedge^i V_{\mathbb{F}_2}$  be the unique linear map such that  $\pi(1) = 1$ ,  $\pi(\bar{x}) = x$ ,  $\pi(\bar{x} \bar{y}) = x \wedge y$  and  $\pi(\bar{x} \bar{y} \bar{z}) = x \wedge y \wedge z$  where  $x, y, z$  are distinct elements

of the basis  $\{a_i, b_i\}_{i=1}^g$ . Clearly,  $\pi$  is bijective, and it is straightforward to check that  $\pi$  is an isomorphism of  $\mathrm{SL}_g(\mathbb{Z})$ -modules<sup>5</sup> (but not an isomorphism of  $\mathrm{Sp}(V)$ -modules!)

Now define  $\pi' : \wedge^3 V \oplus B_3 \rightarrow \wedge^3 V \oplus \bigoplus_{i=0}^3 \wedge^i V_{\mathbb{F}_2}$  by  $\pi'(u, v) = (u, \pi(v) - \alpha(u))$ . Clearly,  $\pi'$  is an isomorphism of  $\mathrm{SL}_g(\mathbb{Z})$ -modules, and it is easy to show that  $\pi'(W) = \wedge^3 V \oplus \bigoplus_{i=0}^2 \wedge^i V_{\mathbb{F}_2}$ . Thus, in view of Theorem 7.1, we have the following isomorphism of  $\mathrm{SL}_g(\mathbb{Z})$ -modules:

$$\lambda = \pi' \circ (\tau, \sigma) : \mathcal{I}^{ab} \rightarrow \wedge^3 V \oplus \left( \bigoplus_{i=0}^2 \wedge^i V_{\mathbb{F}_2} \right).$$

It is clear from (7.5) that  $V$  is a direct sum of the natural  $\mathrm{SL}_g(\mathbb{Z})$ -module  $\mathbb{Z}^g$  and its dual. Using Lemmas 5.1 and 5.2 and the isomorphism  $\lambda$ , we endow  $\mathcal{I}^{ab}$  with the structure of a regular  $\mathrm{SL}_g(\mathbb{Z})$ -module of degree 3, so condition (ii) in Theorem 6.1 holds. Here is an explicit description of the obtained grading on  $\mathcal{I}^{ab}$ . Below  $e_t$  stands for  $a_t$  or  $b_t$ .

- (1) If  $I = \{i < j < k\}$ , then  $(\mathcal{I}^{ab})^I = \lambda^{-1}(\oplus \mathbb{Z}e_i \wedge e_j \wedge e_k)$ .
- (2) If  $I = \{i < j\}$ , then  $(\mathcal{I}^{ab})^I = \lambda^{-1}((\oplus \mathbb{Z}e_i \wedge a_j \wedge b_j) \oplus (\oplus \mathbb{Z}e_j \wedge a_i \wedge b_i) \oplus (\oplus \mathbb{F}_2 e_i \wedge e_j))$ .
- (3) If  $I = \{i\}$ , then  $(\mathcal{I}^{ab})^I = \lambda^{-1}(\mathbb{F}_2 a_i \oplus \mathbb{F}_2 b_i)$ .
- (4) If  $I = \emptyset$ , then  $(\mathcal{I}^{ab})^I = \lambda^{-1}(\mathbb{F}_2)$ .

It remains to check condition (iii), i.e., that the  $\mathbf{g}$ -group structure on  $\mathcal{I}$  defined in § 7.1 is compatible with the grading defined above. Let  $I \subseteq \mathbf{g}$ , with  $|I| \geq 3$ , and let  $\iota_I : (\mathcal{I}_I)^{ab} \rightarrow \mathcal{I}^{ab}$  be the natural map. We need to show that

$$(7.6) \quad \iota_I((\mathcal{I}_I)^{ab}) \supseteq \sum_{J \subseteq I} (\mathcal{I}^{ab})^J$$

or, equivalently, that  $\lambda \circ \iota_I((\mathcal{I}_I)^{ab})$  contains  $\sum_{J \subseteq I} \lambda((\mathcal{I}^{ab})^J)$ . We claim that the latter two

sets are both equal to  $Z_I := \wedge^3 V(I) \oplus \left( \bigoplus_{i=0}^2 \wedge^i V(I)_{\mathbb{F}_2} \right)$  (where we identify  $V(I) = H_1(S_I)$  with its image in  $V = H_1(\Sigma)$ ). Indeed,  $\sum_{J \subseteq I} \lambda((\mathcal{I}^{ab})^J) = Z_I$  directly from (1)-(3) above,

while the equality  $\lambda \circ \iota_I((\mathcal{I}_I)^{ab}) = Z_I$  follows from Proposition 7.2, as we now explain.

For simplicity of notation we identify  $\wedge^3 V(I)$  and  $B_3(I)$  with their canonical images in  $\wedge^3 V$  and  $B_3$ . With this identification, Proposition 7.2 asserts that  $(\tau_I, \sigma_I) = (\tau, \sigma) \circ \iota_I$  where each side of the equality is a map from  $(\mathcal{I}_I)^{ab}$  to  $\wedge^3 V \oplus B_3$ . Thus

$$\lambda \circ \iota_I((\mathcal{I}_I)^{ab}) = \pi' \circ (\tau, \sigma) \circ \iota_I((\mathcal{I}_I)^{ab}) = \pi' \circ (\tau_I, \sigma_I)((\mathcal{I}_I)^{ab}).$$

Since  $|I| \geq 3$ , we can apply Theorem 7.1 (with  $\mathcal{I}_I$  playing the role of  $\mathcal{I}$ ) to conclude that  $(\tau_I, \sigma_I)((\mathcal{I}_I)^{ab}) = W_I$  where (using the notations of § 7.4)

$$W_I = \{(u, v) \in \wedge^3 V(I) \oplus B_3(I) : \alpha(u) = \beta(v)\} = W \cap (\wedge^3 V(I) \oplus B_3(I)).$$

Thus,  $\lambda \circ \iota_I((\mathcal{I}_I)^{ab}) = \pi'(W_I)$ , and a straightforward computation shows that  $\pi'(W_I) = Z_I$ .

**Remark.** The hypothesis  $|I| \geq 3$  was needed twice in the proof of the inclusion (7.6) – first to apply Proposition 7.2 and then to use Theorem 7.1. We do not know if Proposition 7.2 remains valid for all  $I$ , but the equality  $(\tau_I, \sigma_I)((\mathcal{I}_I)^{ab}) = W_I$  and the inclusion (7.6) are

<sup>5</sup>This is true since  $\mathrm{SL}_g(\mathbb{Z})$  preserves both  $V_A = \oplus \mathbb{Z}a_i$  and  $V_B = \oplus \mathbb{Z}b_i$  and the intersection form vanishes on both  $V_A$  and  $V_B$ .

definitely false at least when  $|I| = 0$  (that is,  $I = \emptyset$ ) since  $\mathcal{I}_\emptyset$  is the trivial subgroup while  $(\mathcal{I}^{ab})^\emptyset = \lambda^{-1}(\mathbb{F}_2)$  and  $W_\emptyset = (0, \mathbb{F}_2)$ . Thus it is essential that condition (iii) in Theorem 6.1 is only required to hold when  $|I| \geq d$  rather than for all  $I$ .

## 8. THE ABELIANIZATION OF THE KERNEL OF A HOMOMORPHISM TO A NILPOTENT GROUP

In this short section we prove Theorem 1.6 using the following results of Roseblade and Robinson. A major part of the argument below (including the use of Theorem 8.2) was suggested to us by Andrei Jaikin who substantially simplified our original proof. Our original argument was inspired by the proof of [PS1, Thm 3.6].

**Theorem 8.1** (Roseblade). *Let  $Q$  be a virtually polycyclic group. Then*

- (a) *Every simple  $\mathbb{Z}[Q]$ -module is finite and thus is an  $\mathbb{F}_p[Q]$ -module for some prime  $p$ .*
- (b) *Assume now that  $Q$  is nilpotent, let  $\Omega$  be the augmentation ideal of  $\mathbb{Z}[Q]$  and  $V$  a finitely generated  $\mathbb{Z}[Q]$ -module. If  $\Omega M = 0$  for every simple quotient  $M$  of  $V$ , then  $\Omega^N V = 0$  for some  $N \in \mathbb{N}$ .*

Parts (a) and (b) of Theorem 8.1 are special cases of [Ro, Cor A] and [Ro, Thm B], respectively.

**Theorem 8.2** (Robinson [Rb]). *Let  $Q$  be a nilpotent group, and let  $M$  be a  $Q$ -module. Assume that either  $H_0(Q, M) = 0$  and  $M$  is Noetherian or  $H^0(Q, M) = 0$  and  $M$  is Artinian. Then  $H^i(Q, M) = 0$  and  $H_i(Q, M) = 0$  for all  $i > 0$ .*

We are now ready to prove Theorem 1.6; in fact, we will prove a slightly more general statement:

**Theorem 8.3.** *Let  $G$  be a finitely generated group, let  $K$  be a normal subgroup of  $G$  such that  $Q = G/K$  is nilpotent, and let  $\Omega$  be the augmentation ideal of  $\mathbb{Z}[Q]$ . The following hold:*

- (a)  *$K^{ab} = K/[K, K]$  is finitely generated as a  $\mathbb{Z}[Q]$ -module (equivalently, as a  $\mathbb{Z}[G]$ -module)*
- (b)  *$G/[K, K]$  is nilpotent if and only if  $K/[K, K]$  is annihilated by some power of  $\Omega$ .*
- (c) *Assume that for any finite field  $F$  of prime order and any non-trivial finite-dimensional irreducible representation  $\rho$  of  $G$  over  $F$  with  $\text{Ker } \rho \supseteq K$  we have  $H^1(G, \rho) = 0$ . Then the equivalent conditions from (b) hold.*

**Remark.** (a) and (b) are well known, but we are not aware of a reference in the literature.

*Proof.* We start with an observation that will be used for both (a) and (b). Given  $c \in K$ , let  $\bar{c} = c[K, K]$  denote its image in  $K^{ab}$ . Then

$$(8.1) \quad (g^{-1} - 1)\bar{c} = \overline{[c, g]} \text{ for all } c \in K \text{ and } g \in G$$

(recall that  $[c, g] = c^{-1}g^{-1}cg$ ).

(a) We know that  $K$  contains  $\gamma_n G$  for some  $n$ . Since  $G$  is finitely generated,  $K/\gamma_n G$  is also finitely generated (being a subgroup of a finitely generated nilpotent group). Hence to prove that  $K^{ab}$  is a finitely generated  $\mathbb{Z}[G]$ -module, it suffices to do the same for  $(\gamma_n G)^{ab}$ .

Let  $X$  be a finite generating set for  $G$ . Then  $\gamma_n G$  is generated as a group by left-normed commutators in  $X$  of length at least  $n$ . But if  $c$  is such a commutator and  $x \in X$ , then  $\overline{[c, x]} = (x^{-1} - 1)\bar{c}$  by (8.1). Hence by straightforward induction  $(\gamma_n G)^{ab}$  is generated as a  $\mathbb{Z}[G]$ -module by the images of left-normed commutators in  $X$  of length exactly  $n$ .

(b) Given  $m \in \mathbb{N}$ , we set  $[K, {}_mG] = [K, \underbrace{G, \dots, G}_{m \text{ times}}]$ . First, we claim that  $\Omega^m K^{ab} = \{0\}$

if and only if  $[K, {}_mG] \subseteq [K, K]$ . The forward direction is immediate by (8.1) while the opposite direction follows from (8.1) and the fact that  $\Omega^m$  is generated as a left ideal of  $\mathbb{Z}[Q]$  by elements of the form  $(g_1 - 1) \dots (g_m - 1)$  with  $g_i \in Q$ .

Thus, (b) reduces to showing that  $[K, {}_mG] \subseteq [K, K]$  for some  $m$  if and only if  $G/[K, K]$  is nilpotent. If  $G/[K, K]$  is nilpotent and  $m$  is its nilpotency class, then  $\gamma_{m+1}G \subseteq [K, K]$ , so in particular  $[K, {}_mG] \subseteq [K, K]$ .

Conversely, suppose that  $[K, {}_mG] \subseteq [K, K]$ . Since  $K \supseteq \gamma_n G$  for some  $n$ , we have  $[K, K] \supseteq [\gamma_n G, \underbrace{G, \dots, G}_{m \text{ times}}] = \gamma_{n+m} G$  and so  $G/[K, K]$  is nilpotent.

(c) We will prove that  $\Omega^m K^{ab} = \{0\}$  for some  $m \in \mathbb{N}$  by contradiction. Let  $V = K^{ab}$ , and assume that  $\Omega^m K^{ab} \neq \{0\}$  for any  $m \in \mathbb{N}$ . By Theorem 8.1(b) there is a simple  $Q$ -module  $M$  which is a quotient of  $V$  such that  $\Omega M \neq 0$ , so  $Q$  acts on  $M$  non-trivially. Moreover, by Theorem 8.1(a),  $M$  is a finite  $F[Q]$ -module for some finite field  $F$  of prime order, so  $H^1(G, M) = 0$  by the hypotheses of Theorem 1.6.

Since  $K$  acts trivially on  $V$  and hence on  $M$ , we have a natural isomorphism

$$H^1(K, M) \cong \text{Hom}(K, M) \cong \text{Hom}(V, M).$$

Under this isomorphism  $H^1(K, M)^Q$ , the subspace of  $Q$ -invariant elements of  $H^1(K, M)$ , maps to  $\text{Hom}_Q(V, M)$ , the subspace of  $Q$ -module homomorphisms from  $V$  to  $M$ . Since  $M$  is a quotient of  $V$  as a  $Q$ -module, we have  $\text{Hom}_Q(V, M) \neq 0$  and hence  $H^1(K, M)^Q \neq 0$ .

On the other hand, we have the inflation-restriction sequence

$$0 \rightarrow H^1(Q, M) \rightarrow H^1(G, M) \rightarrow H^1(K, M)^Q \rightarrow H^2(Q, M) \rightarrow H^2(G, M).$$

Since  $M$  is finite, simple and non-trivial, it is both Artinian and Noetherian, and both  $H^0(Q, M)$  and  $H_0(Q, M)$  are trivial. Thus, using either condition in Theorem 8.2, we get  $H^1(Q, M) = H^2(Q, M) = 0$ , so  $H^1(G, M) \cong H^1(K, M)^Q \neq 0$ , a contradiction.  $\square$

## 9. ABELIANIZATION OF FINITE INDEX SUBGROUPS

**9.1. Proof of Theorem 1.9.** In this subsection we will prove Theorem 1.9 whose statement is recalled below.

**Theorem 1.9.** *Let  $G$  and  $N$  be as in Theorem 1.4, and let  $\Gamma = \text{Aut}(F_n)$  if  $G = \text{IA}_n$  and  $\Gamma = \text{Mod}_g^1$  if  $G = \mathcal{I}_g^1$ . Let  $H$  be a subgroup of  $\Gamma$  containing  $\gamma_N G$ . The following hold:*

- (1) *If  $H$  is a finite index subgroup of  $G$ , then  $\dim H^1(H, \mathbb{C}) = \dim H^1(G, \mathbb{C})$ .*
- (2) *If  $H$  is a finite index subgroup of  $\Gamma$ , then  $H$  has finite abelianization.*

We will deduce Theorem 1.9 from Theorem 1.4 and the following two standard results.

**Lemma 9.1.** *Let  $G$  be a group,  $H$  a normal subgroup of  $G$  of finite index and  $Q = G/H$ . Then*

$$(9.1) \quad H^1(H, \mathbb{C}) \cong H^1(G, \mathbb{C}) \oplus \bigoplus_{V \in \text{Irr}(Q) \setminus V_0} H^1(G, V)^{\dim(V)}$$

where  $\text{Irr}(Q)$  is the set of equivalence classes of irreducible complex representations of  $Q$  and  $V_0$  is the trivial representation of  $Q$ .

*Proof.* By Shapiro's Lemma  $H^1(H, \mathbb{C}) \cong H^1(G, \text{Coind}_H^G(\mathbb{C}))$  where  $\text{Coind}_H^G(\mathbb{C})$  is the coinduced module. Since  $H$  is a normal finite index subgroup of  $G$ , we have an isomorphism of  $G$ -modules  $\text{Coind}_H^G(\mathbb{C}) \cong \mathbb{C}[Q] \cong \bigoplus_{V \in \text{Irr}(Q)} V^{\dim(V)}$ , which clearly implies Lemma 9.1.  $\square$

**Definition.** A representation  $(\rho, V)$  of a group  $G$  will be called *finite-image* if  $\rho(G)$  is finite.

**Lemma 9.2.** *Let  $G$  be a finitely generated group and  $K$  a normal subgroup of  $G$ . The following are equivalent:*

- (i)  $H^{ab}$  is finite for any finite index subgroup  $H$  of  $G$  which contains  $K$ .
- (ii)  $H^1(G, V) = 0$  for any irreducible finite-image complex representation  $V$  of  $G$  on which  $K$  acts trivially.

*Proof.* First note that since  $G$  is finitely generated, a finite index subgroup  $H$  of  $G$  is also finitely generated. Hence such  $H$  has finite abelianization if and only if  $H^1(H, \mathbb{C}) = 0$ .

“(ii) $\Rightarrow$ (i)” Assume that (ii) holds, and let  $H$  be a finite index subgroup of  $G$  containing  $K$ . If  $H$  is normal, then  $H^1(H, \mathbb{C}) = 0$  directly by Lemma 9.1. In general, we can find a finite index subgroup  $H'$  of  $H$  which is normal in  $G$  and still contains  $K$ . Then  $H^1(H', \mathbb{C}) = 0$ . Since  $H'$  has finite index in  $H$ , the restriction map  $H^1(H, \mathbb{C}) \rightarrow H^1(H', \mathbb{C})$  is injective and therefore  $H^1(H, \mathbb{C}) = 0$  as well.

“(i) $\Rightarrow$ (ii)” Now assume that (i) holds, and let  $(\rho, V)$  be a representation satisfying the hypotheses of (ii). Then  $H = \text{Ker } \rho$  is a finite index normal subgroup of  $G$  containing  $K$ , so  $H^1(H, \mathbb{C}) = 0$  by (i). Since  $V$  is also an irreducible representation of  $G/H$ , Lemma 9.1 implies that  $H^1(G, V) = 0$  as desired.  $\square$

*Proof of Theorem 1.9.* Part (1) follows directly from Theorem 1.4 and Lemma 9.1.

Let us now prove (2). By Lemma 9.2 we need to show that  $H^1(\Gamma, V) = 0$  for any irreducible finite-image complex representation  $V$  of  $\Gamma$  on which  $\gamma_N G$  acts trivially. The exact sequence of groups  $1 \rightarrow G \rightarrow \Gamma \rightarrow \Gamma/G \rightarrow 1$  yields the following inflation-restriction sequence:

$$(9.2) \quad 0 \rightarrow H^1(\Gamma/G, V^G) \rightarrow H^1(\Gamma, V) \rightarrow H^1(G, V)^{\Gamma/G}.$$

Since  $V$  is irreducible and  $G$  is normal in  $\Gamma$ , we either have  $V^G = 0$  or  $V^G = V$ .

*Case 1:*  $V^G = 0$ . Then  $V$  is a direct sum of non-trivial irreducible finite-image complex representations of  $G$  on which  $\gamma_N G$  acts trivially. Thus  $H^1(G, V) = 0$  by Theorem 1.4 and so  $H^1(\Gamma, V) = 0$  by (9.2).

*Case 2:*  $V^G = V$ . In this case we can deduce that  $H^1(\Gamma, V) = 0$  directly from Lemma 9.2 and the previously known result that  $\Lambda^{ab}$  is finite for every finite index subgroup  $\Lambda$  of  $\Gamma$  containing  $G$  (see § 9.2). For completeness we will give an argument which does not use the latter result.

Recall that the group  $\Gamma/G$  is isomorphic to either  $\text{GL}_n(\mathbb{Z})$  or  $\text{Sp}_{2g}(\mathbb{Z})$ . It is well known that finite index subgroups of  $\text{GL}_n(\mathbb{Z})$ ,  $n \geq 3$ , and  $\text{Sp}_{2g}(\mathbb{Z})$ ,  $g \geq 2$ , have finite abelianization (for instance, because these groups have Kazhdan's property (T) – see, e.g., [BHV, § 1.3, 1.4, 1.7]). Now we can apply Lemma 9.2 to  $\Gamma/G$  with  $K$  being the trivial subgroup to conclude that  $H^1(\Gamma/G, V) = 0$ .

In view of (9.2), it remains to prove that  $H^1(G, V)^{\Gamma/G} = 0$ . Let  $\Lambda$  be the kernel of the action of  $\Gamma$  on  $V$ . By our assumptions  $\Lambda$  is a finite index subgroup of  $\Gamma$  containing  $G$ . Clearly, it suffices to show that  $H^1(G, V)^{\Lambda/G} = 0$ .

Since  $G$  acts trivially on  $V$ , the space  $H^1(G, V)$  is a direct sum of finitely many copies of  $H^1(G, \mathbb{C})$ . Moreover, since  $\Lambda$  acts trivially on  $V$ , each of those copies is  $\Lambda/G$ -invariant, and the action of  $\Lambda/G$  on each copy is the standard action of  $\Lambda/G$  on  $H^1(G, \mathbb{C})$ . Thus, it suffices to show that  $H^1(G, \mathbb{C})^{\Lambda/G} = 0$ .

Let  $\mathcal{G} = \mathrm{SL}_n(\mathbb{R})$  if  $\Gamma = \mathrm{Aut}(F_n)$  and  $\mathcal{G} = \mathrm{Sp}_{2g}(\mathbb{R})$  if  $\Gamma = \mathrm{Mod}_g^1$ . Let  $\bar{\Gamma}' = \mathrm{SL}_n(\mathbb{Z})$  or  $\mathrm{Sp}_{2g}(\mathbb{Z})$ , respectively. The description of  $G^{ab}$  given in § 3.1, § 7 clearly implies that the natural action of  $\bar{\Gamma}'$  on  $H^1(G, \mathbb{C}) \cong (G^{ab} \otimes \mathbb{C})^*$  extends to a polynomial representation of  $\mathcal{G}$ , and it is well known (and easy to check) that the obtained representation does not contain a copy of the trivial representation. Hence  $H^1(G, \mathbb{C})^{\mathcal{G}} = 0$ .

Now let  $\bar{\Lambda}' = (\Lambda/G) \cap \bar{\Gamma}'$ . Then  $\bar{\Lambda}'$  is a finite index subgroup of  $\bar{\Gamma}'$  and hence  $\bar{\Lambda}'$  is Zariski dense in  $\mathcal{G}$  (one possible argument is that  $\bar{\Lambda}'$  has infinite intersection with each of the root subgroups of  $\mathcal{G}$ , so the Zariski closure of  $\bar{\Lambda}'$  contains all the root subgroups which, in turn, generate  $\mathcal{G}$ ). Thus, the equality  $H^1(G, \mathbb{C})^{\mathcal{G}} = 0$  implies that  $H^1(G, \mathbb{C})^{\bar{\Lambda}'} = 0$ , and so  $H^1(G, \mathbb{C})^{\Lambda/G} = 0$  as well.  $\square$

**Remark.** (a) The above argument shows that in order to prove Theorem 1.9(2), we only need to know Theorem 1.4 for representations  $(\rho, V)$  of  $G$  which are extendable to  $\Gamma$ . The proof of Theorem 1.4 in this special case is much easier (see Case 1 in § 4.3).

(b) As mentioned in the introduction, a positive answer to Question 1.8 (for all  $H$ ) would imply the same for Question 1.7. Indeed, let  $G = \mathrm{IA}_n$  or  $\mathrm{Mod}_g^1$  and assume that  $\dim H^1(G, \mathbb{C}) = \dim H^1(H, \mathbb{C})$  for every finite index subgroup  $H$  of  $G$ . By Lemma 9.2,  $H^1(G, V) = 0$  for every non-trivial irreducible finite-image complex representation  $V$  of  $G$ . The proof of Theorem 1.9(2) then shows that  $H^1(\Gamma, V) = 0$  for every irreducible finite-image complex representation  $V$  of  $\Gamma$ , so applying Lemma 9.2 again we conclude that every finite index subgroup of  $\Gamma$  has finite abelianization.

Let us now explain why the class of subgroups  $H$  satisfying the hypotheses of Theorem 1.9 is quite large. Let  $\Gamma$ ,  $G$  and  $N$  be as in Theorem 1.9. The group  $G/\gamma_N G$  is finitely generated nilpotent and hence residually finite. Thus,  $G/\gamma_N G$  has plenty of finite index subgroups whose preimages in  $G$  satisfy the hypotheses of Theorem 1.9(1).

For part (2) we can use, for instance, appropriate congruence subgroups. Recall that  $\Gamma$  in Theorem 1.9 is either equal to  $\mathrm{Aut}(F)$ , where  $F$  is free of rank  $n$ , or is isomorphic to a subgroup of  $\mathrm{Aut}(F)$ , where  $F$  is free of rank  $2g$ . Given a normal subgroup  $S$  of  $F$ , the corresponding congruence subgroup  $\Gamma(S)$  is defined to be the set of all  $f \in \Gamma$  which leave  $S$  invariant and act trivially on  $F/S$ . Note that if  $S = \gamma_{k+1} F$ , then  $\Gamma(S) = G(k)$ , the  $k^{\mathrm{th}}$  term of the Johnson filtration. If  $S$  has finite index in  $F$ , it is clear that  $\Gamma(S)$  has finite index in  $\Gamma$ . If in addition  $F/S$  is nilpotent of class at most  $N$ , then  $\Gamma(S)$  contains  $G(N)$  and hence also contains  $\gamma_N G$ .

**9.2. Earlier work on Question 1.7.** In this subsection we will summarize previously known positive results on Question 1.7 which is restated below:

**Question 1.7.** *Let  $\Gamma = \mathrm{Aut}(F_n)$  or  $\mathrm{Mod}_g^1$ . If  $H$  is a finite index subgroup of  $\Gamma$ , does  $H$  always have finite abelianization?*

We will start with the mapping class group case. Hain [Ha, Prop 5.2] positively answered Question 1.7 for  $\Gamma = \mathrm{Mod}_g^1$ ,  $g \geq 3$  and  $H$  containing the Torelli subgroup  $\mathcal{I}_g^1$ . Putman [Pu1,

Thm B] generalized this result to all  $H$  containing the Johnson kernel  $\mathcal{I}_g^1(2)$ .<sup>6</sup> Slightly earlier Boggi [Bo, Cor 3.11(b)] proved by a different method that  $H^{ab}$  is finite for all finite index subgroups of  $\text{Mod}_g$  containing  $\mathcal{I}_g(2)$ . Finally, an interesting general approach to Question 1.7 was suggested by Putman and Wieland [PW].

In the case  $\Gamma = \text{Aut}(F_n)$  a positive answer to Question 1.7 for  $H$  containing  $\text{IA}_n$ ,  $n \geq 3$ , was given by elementary arguments by Bhattacharya [Bh]<sup>7</sup> and independently by Bogopolski and Vikentiev [BV]. This result also follows from the earlier work of Lubotzky and Pak [LP] (see Theorem 9.4 below).

The following theorem was established in the Ph.D. thesis of Kielak:

**Theorem 9.3.** ([Ki, Thm 5.3.2]). *Let  $n \geq 3$  and  $H$  a finite index subgroup of  $\text{Aut}(F_n)$  such that  $[H \cap \text{IA}_n, H \cap \text{IA}_n]$  contains  $\gamma_k \text{IA}_n$  for some  $k$ . Then  $H$  has finite abelianization.*

A very similar result follows from [LP, Thm 3.8] which asserts that  $\text{Aut}(F_n)/\text{IA}_n(k)$  has Kazhdan's property (T) for all  $n \geq 3, k \geq 1$ :

**Theorem 9.4.** Let  $n \geq 3$  and  $H$  a finite index subgroup of  $\text{Aut}(F_n)$  such that  $[H, H]$  contains  $\text{IA}_n(k)$  (the  $k^{\text{th}}$  term of the Johnson filtration) for some  $k$ . Then  $H$  has finite abelianization.

*Proof.* Let  $\Gamma = \text{Aut}(F_n)$  and  $f : \text{Aut}(F_n) \rightarrow \text{Aut}(F_n)/\text{IA}_n(k)$  the natural projection. Since  $H$  has finite index in  $\Gamma$ , its image  $f(H)$  has finite index in  $f(\Gamma)$ . Since  $f(\Gamma)$  has property (T),  $f(H)$  has finite abelianization, whence  $f([H, H]) = [f(H), f(H)]$  has finite index in  $f(\Gamma)$ . Finally, since  $[H, H] \supseteq \text{IA}_n(k) = \text{Ker}(f)$ , we have  $[H, H] = f^{-1}(f([H, H]))$ , so  $[H, H]$  has finite index in  $\Gamma$  and thus  $H^{ab}$  is finite.  $\square$

Corollary 1.5 implies that Theorem 9.3 applies to any subgroup  $H$  satisfying the hypotheses of Theorem 1.9(2) with  $\Gamma = \text{Aut}(F_n)$ . However this does not give a substantially new proof of Theorem 1.9(2) for  $\Gamma = \text{Aut}(F_n)$  since the proof of Corollary 1.5 also uses Theorem 1.4. It would be interesting to find other classes of subgroups which can be explicitly constructed and satisfy the hypotheses of Theorem 9.3 or Theorem 9.4.

**9.3. Some negative results for  $\text{IA}_3$ .** In this final subsection we explain why the assertions of Theorem 1.4 and both parts of Theorem 1.9 do not hold for  $G = \text{IA}_3$  and  $N = 2$ . More explicitly, the following is true:

**Proposition 9.5.** *The following hold:*

- (a) *There exists a finite index subgroup  $\Delta$  of  $\text{Aut}(F_3)$  such that  $\Delta$  has infinite abelianization and  $[\text{IA}_3 : \Delta \cap \text{IA}_3] = 2$ .*
- (b) *There exists a subgroup  $L$  of  $\text{IA}_3$  with  $[\text{IA}_3 : L] = 2$  (so in particular,  $L \supseteq [\text{IA}_3, \text{IA}_3]$ ) such that  $\dim H^1(L, \mathbb{C}) > \dim H^1(\text{IA}_3, \mathbb{C})$ .*
- (c) *There exists a one-dimensional complex representation  $(\rho, V)$  of  $\text{IA}_3$  with  $|\rho(\text{IA}_3)| = 2$  such that  $H^1(\text{IA}_3, V) \neq 0$ .*

<sup>6</sup>[Pu1, Thm B] actually covers all subgroups containing a “large chunk” of the Johnson kernel. The results of [Ha] and [Pu1] apply to mapping class groups of surfaces with an arbitrary number of punctures and boundary components.

<sup>7</sup>Our argument in Case 2 of the proof of Theorem 1.9(2) is similar (but not identical) to the argument from [Bh].

*Proof.* (a) follows immediately from the work of Grunewald and Lubotzky [GL, Prop 6.2] (see also [BV]). We note that the existence of some finite index subgroup of  $\text{Aut}(F_3)$  with infinite abelianization was established earlier by McCool [McCo] by a different method.

(c) Let  $\Delta$  be as in (a), and let  $\Delta' = \text{IA}_3 \cdot \Delta$ . Then  $\Delta'$  is a finite index subgroup of  $\text{Aut}(F_3)$  containing  $\text{IA}_3$ , so  $\Delta'$  has finite abelianization by Theorem 9.3 or 9.4. Note that  $[\Delta' : \Delta] = [\text{IA}_3 : \Delta \cap \text{IA}_3] = 2$ , so in particular  $\Delta$  is normal in  $\Delta'$ . Applying Lemma 9.1 with  $G = \Delta'$  and  $H = \Delta$ , we get  $H^1(\Delta, \mathbb{C}) \cong H^1(\Delta', \mathbb{C}) \oplus H^1(\Delta', V)$  where  $(\rho, V)$  is the unique non-trivial irreducible complex representation of  $\Delta'/\Delta$  (in particular  $\dim V = 1$  and  $|\rho(\text{IA}_3)| = |\rho(\Delta')| = 2$ ).

Since  $\Delta'$  has finite abelianization while  $\Delta$  does not, we have  $H^1(\Delta', \mathbb{C}) = 0$  and  $H^1(\Delta, \mathbb{C}) \neq 0$ , whence  $H^1(\Delta', V) \neq 0$ . Since  $V$  is irreducible and non-trivial, we have  $V^{\text{IA}_3} = V^{\Delta'} = 0$ . Thus, by inflation-restriction sequence  $H^1(\Delta', V)$  embeds in  $H^1(\text{IA}_3, V)$ , so  $H^1(\text{IA}_3, V) \neq 0$  and  $(\rho, V)$  has the desired properties.

(b) Let  $(\rho, V)$  be as in (c) and  $L = \text{Ker } \rho$ . Then  $(\rho, V)$  is the unique non-trivial irreducible complex representation of  $\text{IA}_3/L$ . By Lemma 9.1 with  $G = \text{IA}_3$  and  $H = L$  we have  $H^1(L, \mathbb{C}) \cong H^1(\text{IA}_3, \mathbb{C}) \oplus H^1(\text{IA}_3, V)$ , which implies (b).  $\square$

## REFERENCES

- [AB] Roger Alperin and Hyman Bass, *Length functions of group actions on  $\Lambda$ -trees*. Combinatorial group theory and topology (Alta, Utah, 1984), 265-378, Ann. of Math. Stud., 111, Princeton Univ. Press, Princeton, NJ, 1987.
- [An] Stylianos Andreadakis, *On the automorphisms of free groups and free nilpotent groups*. Proc. London Math. Soc. (3) 15 (1965), 239-268.
- [Ba] Seymour Bachmuth, *Induced automorphisms of free groups and free metabelian groups*. Trans. Amer. Math. Soc. 122 (1966), 1-17.
- [BHV] Bachir Bekka, Pierre de la Harpe and Alain Valette, *Kazhdan's property (T)*. New Math. Monogr. 11, Cambridge Univ. Press, Cambridge, 2008.
- [BBM] Mladen Bestvina, Kai-Uwe Bux and Dan Margalit, *Dimension of the Torelli group for  $\text{Out}(F_n)$* . Invent. Math. 170 (2007), no. 1, 1-32.
- [Bh] Chirashree Bhattacharya, *Free groups and their automorphisms*. Thesis (Ph.D.)-University of Virginia. 2004. 47 pp.
- [BNS] Robert Bieri, Walter D. Neumann and Ralph Strebel, *A geometric invariant of discrete groups*. Invent. Math. 90 (1987), no. 3, 451-477.
- [BC] Joan S. Birman and Robert Craggs, *The  $\mu$ -invariant of 3-manifolds and certain structural properties of the group of homeomorphisms of a closed, oriented 2-manifold*. Trans. Amer. Math. Soc. 237 (1978), 283-309.
- [Bo] Marco Boggi, *Fundamental groups of moduli stacks of stable curves of compact type*. Geom. Topol. 13 (2009), no. 1, 247-276.
- [BV] Oleg Bogopolski and R. Vikentiev, *Subgroups of small index in  $\text{Aut}(F_n)$  and Kazhdan's property (T)*. Combinatorial and geometric group theory, 1-17, Trends Math., Birkhäuser/Springer Basel AG, Basel, 2010.
- [BF] Tara E. Brendle and Benson Farb, *The Birman-Craggs-Johnson homomorphism and abelian cycles in the Torelli group*. Math. Ann. 338 (2007), no. 1, 33-53.
- [Br1] Kenneth S. Brown, *Trees, valuations, and the Bieri-Neumann-Strebel invariant*. Invent. Math. 90 (1987), no. 3, 479-504.
- [Br2] Kenneth S. Brown, *Cohomology of groups. Corrected reprint of the 1982 original*. Graduate Texts in Mathematics, 87. Springer-Verlag, New York, 1994. x+306 pp.
- [CEP] Thomas Church, Mikhail Ershov and Andrew Putman, *On finite generation of the Johnson filtrations*, preprint (2017), arXiv:1711.04779.
- [CP] Thomas Church and Andrew Putman, *Generating the Johnson filtration*. Geom. Topol. 19 (2015), no. 4, 2217-2255.



- [DHP] Alexandru Dimca, Richard Hain and Stefan Papadima, *The abelianization of the Johnson kernel*. J. Eur. Math. Soc. (JEMS) 16 (2014), no. 4, 805-822.
- [DP] Alexandru Dimca and Stefan Papadima, *Arithmetic group symmetry and finiteness properties of Torelli groups*. Ann. of Math. (2) 177 (2013), no. 2, 395-423.
- [FM] Benson Farb and Dan Margalit, *A primer on mapping class groups*. Princeton Mathematical Series, 49. Princeton University Press, Princeton, NJ, 2012. xiv+472 pp.
- [Fo] Edward Formanek, *Characterizing a free group in its automorphism group*. J. Algebra 133 (1990), no. 2, 424-432.
- [GL] Fritz Grunewald and Alex Lubotzky, *Linear representations of the automorphism group of a free group*. Geom. Funct. Anal. 18 (2009), no. 5, 1564-1608.
- [Ha] Richard Hain, *Torelli groups and geometry of moduli spaces of curves. (English summary)* Current topics in complex algebraic geometry (Berkeley, CA, 1992/93), 97-143, Math. Sci. Res. Inst. Publ., 28, Cambridge Univ. Press, Cambridge, 1995.
- [Jo1] Dennis Johnson, *Homeomorphisms of a surface which act trivially on homology*. Proc. Amer. Math. Soc. 75 (1979), no. 1, 119-125.
- [Jo2] Dennis Johnson, *Quadratic forms and the Birman-Craggs homomorphisms*. Trans. Amer. Math. Soc. 261 (1980), no. 1, 235-254.
- [Jo3] Dennis Johnson, *An abelian quotient of the mapping class group  $\mathcal{I}_g$* . Math. Ann. 249 (1980), no. 3, 225-242.
- [Jo4] Dennis Johnson, *A survey of the Torelli group*. Low-dimensional topology (San Francisco, Calif., 1981), 165-179, Contemp. Math., 20, Amer. Math. Soc., Providence, RI, 1983.
- [Jo5] Dennis Johnson, *The structure of the Torelli group. I. A finite set of generators for  $\mathcal{I}$* . Ann. of Math. (2) 118 (1983), no. 3, 423-442.
- [Jo6] Dennis Johnson, *The structure of the Torelli group. III. The abelianization of  $\mathcal{T}$* . Topology 24 (1985), no. 2, 127-144.
- [Iv] Nikolai V. Ivanov, *Fifteen problems about the mapping class groups*. Problems on mapping class groups and related topics, 71-80, Proc. Sympos. Pure Math., 74, Amer. Math. Soc., Providence, RI, 2006.
- [Kaw] Nariya Kawazumi, *Cohomological Aspects of Magnus Expansions*, preprint (2005), arXiv:math/0505497.
- [Ki] Dawid Kielak, *Free and linear representations of outer automorphism groups of free groups*, Ph.D. Thesis (2012), available at <https://www.math.uni-bielefeld.de/~dkielak/DPhil.pdf>
- [KMM] Nic Koban, Jon McCammond and John Meier, *The BNS-invariant for the pure braid groups*, Groups Geom. Dyn. 9 (2015), no. 3, 665-682.
- [LP] Alexander Lubotzky and Igor Pak, *The product replacement algorithm and Kazhdan's property (T)*. J. Amer. Math. Soc. 14 (2001), no. 2, 347-363.
- [Ma] Wilhelm Magnus, *Über  $n$ -dimensional Gittertransformationen*. Acta Math. 64 (1934), 353-367.
- [McCa] John D. McCarthy, *On the first cohomology group of cofinite subgroups in surface mapping class groups*, Topology 40 (2001), no. 2, 401-418.
- [McCo] James McCool, *A faithful polynomial representation of  $\text{Out}F_3$* . Math. Proc. Cambridge Philos. Soc. 106 (1989), no. 2, 207-213.
- [MM] Daryl McCullough and A. Miller, *The genus 2 Torelli group is not finitely generated*, Topology Appl. 22 (1986), 43-49.
- [Me] Geoffrey Mess, *The Torelli groups for genus 2 and 3 surfaces*. Topology 31 (1992), no. 4, 775-790.
- [Mo] Shigeyuki Morita, *Structure of the mapping class groups of surfaces: a survey and a prospect*. Proceedings of the Kirbyfest (Berkeley, CA, 1998), 349-406, Geom. Topol. Monogr., 2, Geom. Topol. Publ., Coventry, 1999.
- [OK] Lisa A. Orlandi-Korner, *The Bieri-Neumann-Strebel invariant for basis-conjugating automorphisms of free groups*. Proc. Amer. Math. Soc. 128 (2000), no. 5, 1257-1262.
- [PS1] Stefan Papadima and Alexander Suciu, *Bieri-Neumann-Strebel-Renz invariants and homology jumping loci*. Proc. Lond. Math. Soc. (3) 100 (2010), no. 3, 795-834.
- [PS2] Stefan Papadima and Alexander Suciu, *Homological finiteness in the Johnson filtration of the automorphism group of a free group*. J. Topol. 5 (2012), no. 4, 909-944.
- [Pu1] Andrew Putman, *A note on the abelianizations of finite-index subgroups of the mapping class group*. Proc. Amer. Math. Soc. 138 (2010), no. 2, 753-758.

- [Pu2] Andrew Putman, *Small generating sets for the Torelli group*. Geom. Topol. 16 (2012), no. 1, 111-125.
- [Pu3] Andrew Putman, *The Johnson homomorphism and its kernel*, to appear in J. Reine Angew. Math., arXiv:0904.0467.
- [PW] Andrew Putman and Ben Wieland, *Abelian quotients of subgroups of the mappings class group and higher Prym representations*. J. Lond. Math. Soc. (2) 88 (2013), no. 1, 79-96.
- [Rb] Derek J. S. Robinson, *The vanishing of certain homology and cohomology groups*. J. Pure Appl. Algebra 7 (1976), no. 2, 145-167.
- [Ro] James E. Roseblade, *Group rings of polycyclic groups*. J. Pure Appl. Algebra 3 (1973), 307-328.

UNIVERSITY OF VIRGINIA

*E-mail address:* `ershov@virginia.edu`

UNIVERSITY OF VIRGINIA

*E-mail address:* `suzy.he@gmail.com`