JOURNAL OF ALGEBRA **195,** 241–284 (1997)
ARTICLE NO. JA966941

On Finite and Locally Finite Subgroups of Free Burnside Groups of Large Even Exponents

S. V. Ivanov*

Department of Mathematics, University of Illinois at Urbana–Champaign, 1409 West Green Street, Urbana, Illinois 61801

and

A. Yu. Ol'shanskii†

Department of Mechanics and Mathematics, Moscow State University, Moscow 119899, Russia

Communicated by Efim Zelmanov

Received October 8, 1996

The following basic results on infinite locally finite subgroups of a free m-generator Burnside group $B(m, n)$ of even exponent n, where $m > 1$ and $n \geq 2^{48}$, n is divisible by 2^9, are obtained: A clear complete description of all infinite groups that are embeddable in $B(m, n)$ as (maximal) locally finite subgroups is given. Any infinite locally finite subgroup L of $B(m, n)$ is contained in a unique maximal locally finite subgroup, while any finite 2-subgroup of $B(m, n)$ is contained in continuously many pairwise nonisomorphic maximal locally finite subgroups. In addition, L is locally conjugate to a maximal locally finite subgroup of $B(m, n)$. To prove these and other results, centralizers of subgroups in $B(m, n)$ are investigated. For example, it is proven that the centralizer of a finite 2-subgroup of $B(m, n)$ contains a subgroup isomorphic to a free Burnside group $B(\infty, n)$ of countably infinite rank and exponent n; the centralizer of a finite non-2-subgroup of $B(m, n)$ or the centralizer of a nonlocally finite subgroup of $B(m, n)$ is always finite; the centralizer of a subgroup S is infinite if and only if S is a locally finite 2-group.

*E-mail: ivanov@math.uiuc.edu.
†E-mail: olsh@nw.math.msu.su.

0021-8693/97 $25.00

INTRODUCTION

Recall a free m-generator Burnside group $B(m,n)$ of exponent n is the quotient F_m/F_m^n, where F_m is a free group of rank m and F_m^n is the subgroup of F_m generated by all nth powers of elements of F_m.

The Burnside problem (see [2, 3, 7, 14]) asks whether or not the free Burnside group $B(m,n)$ is finite. It was proven by Novikov and Adian [17, 1] that, whenever $m > 1$, n is odd, and $n \geq 665$, the group $B(m,n)$ is infinite and all of its finite and Abelian subgroups are cyclic. Much simpler proofs of these results were later given by the second author [18, 19].

However, the case of even exponent n, being especially interesting for $n = 2^k$, remained open until recently and looked much more complicated in view of noncyclicity of centralizers of elements in $B(m,n)$ and noncyclicity of finite subgroups in $B(m,n)$. For example, Held [8] had proved that any infinite 2-group contains infinite Abelian subgroups. This suggested that finite subgroups of $B(m,n)$ with even exponent $n \gg 1$ might be large, possibly large enough to ensure finiteness of the entire group $B(m,n)$.

The main result of the first author's article [11] (see also [10]) is the infiniteness of a free m-generator Burnside group $B(m,n)$ of even exponent n for $m > 1$ and $n \geq 2^{48}$.[1] However, finite subgroups of $B(m,n)$ turn out to be so important in proofs that at least a third of the article [11] is an investigation of their various properties and another third is a preparation of necessary techniques to conduct this investigation. The central result relating to finite subgroups of the $B(m,n)$ is the following: Let $n = n_1 n_2$, where n_1 is the maximal odd divisor of n, $n \geq 2^{48}$ and $n_2 \geq 2^9$. Then any finite subgroup G of $B(m,n)$ is isomorphic to a subgroup of $D(2n_1) \times D(2n_2)^l$ for some l, where $D(2k)$ denotes a dihedral group of order $2k$. The natural question on whether this description is complete [i.e., every subgroup of $D(2n_1) \times D(2n_2)^l$ can actually be found in $B(m,n)$] is not directly addressed in [11] because it is not required in proofs.

As was implied by results of [11], this is a complete description as follows at once from Theorem 1(a). Theorem 1 deals with basic properties of centralizers in $B(m,n)$. These properties will be the base of subsequent investigation of locally finite subgroups in $B(m,n)$.

Before stating Theorem 1, recall a group G is called *locally finite* if every finitely generated subgroup of G is finite.

[1] *Note added in proof*. Note this estimate $n \geq 2^{48}$ has been improved in I. G. Lysenok, Infinite Burnside groups of even period, *Izv. Ross. Akad. Nauk Ser. Mat.* **60** (1996), 3–224.

THEOREM 1. *Let $B(m,n)$ be a free m-generator Burnside group of exponent n, where $m > 1$ and $n \geq 2^{48}$, $n = n_1 n_2$, n_1 is odd, n_2 is a power of 2, $n_2 \geq 2^9$. Then the following hold*:

(a) *Suppose G is a finite 2-subgroup of $B(m,n)$. Then the centralizer $C_{B(m,n)}(\mathsf{G})$ of G in $B(m,n)$ contains a subgroup B isomorphic to a free Burnside group $B(\infty, n)$ of infinite countable rank such that $\mathsf{G} \cap \mathsf{B} = \{1\}$. In particular, $\langle \mathsf{G}, \mathsf{B} \rangle = \mathsf{G} \times \mathsf{B}$.*

(b) *The centralizer $C_{B(m,n)}(\mathsf{H})$ of a subgroup H of $B(m,n)$ is infinite if and only if H is a locally finite 2-subgroup. In particular, $C_{B(m,n)}(\mathsf{H})$ is finite provided H is not locally finite.*

(c) *If G is a finite 2-subgroup of $B(m,n)$, then $C_{B(m,n)}(C_{B(m,n)}(\mathsf{G})) = \mathsf{G}$.*

COROLLARY 1. *A finite group G embeds in $B(m,n)$ if and only if G is isomorphic to a subgroup of the direct product $D(2n_1) \times D(2n_2)^l$ for some $l > 0$, where $D(2k)$ is a dihedral group of order $2k$.*

COROLLARY 2. *Every infinite locally finite subgroup of $B(m,n)$ belongs to the quasivariety* qvar $D(2n_2)$ *of groups generated by $D(2n_2)$ and, being so, is isomorphic to a subgroup of the Cartesian product $D(2n_2)^\infty$ of countably many copies of $D(2n_2)$.*

Recall a group G is called *locally normal* if every finite subset of G is contained in a finite normal subgroup of G. A group G is termed an *FC-group* if every conjugacy class in G is finite. It is well known [5, 22] (and easy to show) that a periodic group is locally normal if and only if it is an *FC*-group.

COROLLARY 3. *Suppose a countable subgroup G of the Cartesian product $D(2n_2)^\infty$ is locally normal. Then G embeds in $B(m,n)$.*

To give a complete description of infinite groups that are embeddable in $B(m,n)$ as locally finite subgroups, let us introduce the following notation. Let D_i, $i = 1, 2, \ldots,$ be groups isomorphic to $D(2n_2)$, D be the Cartesian product of D_i, $i = 1, 2, \ldots, C_i$ be the normal cyclic subgroup of D_i of order n_2, and $b_i \in D_i$ be an element of order 2 that together with C_i generate $D_i = \langle C_i, b_i \rangle$. By B denote the subgroup of D that consists of all elements whose projection on every D_i is either b_i or 1. By C denote the direct product of groups C_i naturally embedded in D. At last, let $\mathsf{E} = \langle \mathsf{B}, \mathsf{C} \rangle$. Clearly, $\mathsf{E} = \mathsf{B}\mathsf{C}$ is a semidirect product of B and C.

THEOREM 2. *Let $B(m,n)$, m, and n be defined as in Theorem 1. Then an arbitrary infinite group G embeds in $B(m,n)$ as a locally finite subgroup if and only if G is isomorphic to a countable subgroup of E.*

The following becomes obvious from Theorem 2.

COROLLARY 4. *The group $B(m, n)$ contains locally finite subgroups that are not FC-groups.*

Nevertheless, locally finite subgroups of $B(m, n)$ are rather close to FC-groups.

COROLLARY 5. *Suppose a locally finite group G embeds in $B(m, n)$. Then the square g^2 of every element $g \in G$ lies in a finite conjugacy class of G and every finite subset of $G^2 = \langle g^2 \mid g \in G \rangle$ is contained in a finite normal subgroup of G.*

Several natural questions about the inclusion relation on the set of finite and locally finite subgroups of $B(m, n)$ are answered subsequently.

THEOREM 3. *Let $B(m, n)$, m, and n be defined as in Theorem 1. Then the following are true:*

(a) *An infinite locally finite subgroup L of $B(m, n)$ is contained in a unique maximal locally finite subgroup. That is, the intersection of two distinct maximal locally finite subgroups of $B(m, n)$ is always finite.*

(b) *Given a finite 2-subgroup G of $B(m, n)$ there are continuously many pairwise nonisomorphic maximal locally finite subgroups that contain G.*

(c) *If a finite subgroup G of $B(m, n)$ contains a nontrivial element of odd order, then G is contained in a unique maximal finite subgroup. In particular, the intersection of two distinct maximal finite subgroups of $B(m, n)$ is always a 2-group.*

A description of maximal locally finite subgroups in $B(m, n)$ will be derived from Theorem 2 by making use of the following definition and Theorem 4. Let A, B be some subgroups of a group G. The subgroups A, B of G are called *locally conjugate* if there is an isomorphism $\alpha : A \to B$ such the restriction $\alpha|_M$ of α on every finite subset M of A equals the restriction $\tau|_M$, where τ is an inner automorphism of G depending on M. Note this property is weaker than that of local conjugacy in the sense of Kurosh's book [13]. (For example, if α is a locally inner automorphism of G such $\alpha(A) = B$, then the maximality of locally finite subgroup A implies that of B, but this is not the case for our definition in which α is an arbitrary isomorphism between A and B.)

THEOREM 4. *Let $B(m, n)$, m, and n be defined as in Theorem 1. Then an infinite locally finite subgroup L of $B(m, n)$ is locally conjugate to a maximal locally finite subgroup.*

A description of infinite maximal locally finite subgroups of $B(m, n)$ is now immediate from Theorem 2.

COROLLARY 6. *An infinite group G embeds in $B(m,n)$ as a maximal locally finite subgroup if and only if G is isomorphic to a countable subgroup of the group E of Theorem* 2.

Section 1 contains several technical lemmas whose terminology and proofs depend heavily upon those of [11]. In Section 2, we will prove Theorems 1–4 and Corollaries 1–6 on the basis of lemmas in Section 1 and lemmas of [11].

Let us conclude with several remarks. First, the mutual disposition of infinite maximal locally finite subgroups stated in Theorem 3(a) and (b) is reminiscent of a known puzzle-type problem: Find, in a countably infinite set, continuously many subsets whose pairwise intersections are all finite (note this is impossible if the cardinalities of the intersections are bounded).

A theorem of Mal'cev (see [15, Sect. 11.V]) claims that the quasivariety qvar G of groups generated by a finite group G (i.e., the class of groups that satisfy the quasi-identities that hold in G) consists of subgroups of Cartesian products of copies of G. Hence Corollary 1 means that a finite 2-group embeds in $B(m,n)$ if and only if it satisfies all quasi-identities of $D(2n_2)$. So it would be of interest to express this condition in terms of quasi-identities, that is, to find a basis of quasi-identities of qvar $D(2n_2)$. Note that, according to a result of the second author [20], the quasivariety qvar G, where G is finite, has a finite basis of quasi-identities if and only if all Sylow subgroups of G are Abelian. In particular, qvar $D(2n_2)$ has no finite basis of quasi-identities provided $n_2 > 2$.

Let us emphasize that the class of finite 2-subgroups in $B(m,n)$ may not be described by identities, since it is not closed under taking homomorphic images. For example, any free group in the variety var $D(2n_2)$ belongs to qvar $D(2n_2)$ by the Birkhoff theorem on varieties (see [15, 16]). In particular, by Corollary 1, a free group $F(r, D(2n_2))$ of any finite rank r in the variety var $D(2n_2)$ embeds in $B(m,n)$. However, the quaternion group Q_8 of order 8, being a homomorphic image of $F(2, D(2n_2))$, is not embeddable in the group $B(m,n)$ of Theorems 1–4, because Q_8 does not satisfy the following quasi-identity which holds in $D(2n_2)$:

$$x^2 = y^2 = [x,y] \quad \Rightarrow \quad [x,y] = 1.$$

It is also of interest to point out that the free group $F(\infty, D(2n_2))$ of countable rank in var $D(2n_2)$ is not embeddable in the group $B(m,n)$ of Theorems 1–4 either, because, otherwise, it would follow from Corollary 5 that the set $[x_1^2, x_j] \subseteq F(\infty, D(2n_2))$ was finite, whence $[x^2, y] \equiv 1$ were identity in $F(\infty, D(2n_2))$. However, this identity does not hold in $D(2n_2)$ when $n_2 > 2$.

Finally, Corollaries 2–5 suggest a conjecture that for a countable subgroup G of the Cartesian power $D(2n_2)^\infty$ the condition that G^2 consists of

FC-elements might be sufficient for embedding of G in $B(m, n)$. However, there is an example showing that this is not the case.

1. SEVERAL LEMMAS

Let us fix an alphabet $\mathsf{A} = \{a_1, \dots, a_m\}$. A word W in $\mathsf{A}^{\pm 1}$ will be called an A-*word*. As in [18, 11], define the group $B(i) = B(m, n)(i)$ by a presentation which is constructed by induction on i as follows. Let $B(0) = F(\mathsf{A})$, where $F(\mathsf{A})$ is the free group over A. Suppose $B(i-1)$, $i \geq 1$, is already given by defining relations and A_i is a shortest A-word (if any) such that its image has infinite order in the group $B(i-1)$. Then $B(i)$ is the quotient group of $B(i-1)$ by the normal closure of the image of A_i^n. Clearly, $B(i)$ has the presentation

$$B(i) = \langle \mathsf{A} \parallel A_1^n, \dots, A_i^n \rangle .$$

It is shown in [11] (and for odd exponents $n > 10^{10}$ in [18]) that if m and n are as in Theorem 1, then A_i exists for every $i \geq 1$ and

$$B(m, n) = \langle \mathsf{A} \parallel A_1^n, \dots, A_i^n, \dots \rangle .$$

The terminology of statements and proofs of Lemmas 1–12 are those of [11] (see also [18, 19]), and all undefined terms and notions are found in [11]. The proofs will involve many references to lemmas in [11]. When making these references we will often drop "[11]" providing only lemma numbers from [11].

Recall a *k-aperiodic* word is a word with no nonempty subwords of the form E^k.

LEMMA 1. *Let Δ be a reduced diagram of rank i and let p be a section of $\partial\Delta$ such that $\varphi(p)$ is a reduced* 9-*aperiodic word. Then the degree of contiguity of any cell π in Δ to p does not exceed β and p is geodesic in Δ.*

Proof. The first claim easily follows from the definition of a contiguity subdiagram and the 9-aperiodicity of $\varphi(p)$ (for details see the proof of Lemma 18.1 [11]). Assume that p is not geodesic. Then there is a subdiagram Γ in Δ with $\partial\Gamma = pq$, where $|q| < |p|$ and q is geodesic in Γ. Since $\varphi(p)$ is reduced, Γ contains cells. Then, by Lemmas 5.7 and 9.2 [11], there is a θ-cell Π in Γ. Denote contiguity subdiagrams of Π to p and q by Γ_p and Γ_q, respectively. By Lemma 3.3, $|\Gamma_q \wedge \Pi| < \alpha|\partial\Pi|$, whence $|\Gamma_p \wedge \Pi| > (\theta - \alpha)|\partial\Pi| > \beta|\partial\Pi|$. A contradiction to the first claim completes the proof. ∎

LEMMA 2. *Suppose* Δ *is a reduced diagram of rank* i *with* $\partial\Delta = p_1 t p_2 q$, *where* $\varphi(p_1)$, $\varphi(p_2)$ *are* 9-*aperiodic words, the path* $p_2 q p_1$ *is reduced,* t *is geodesic in* Δ, q *is geodesic or a smooth section of rank* $j \le i$, *and there are no* 0-*bonds between* p_1 *and* p_2 *in* Δ. *Then* $|p_1|, |p_2| \le 1.04|t|$ *and* $|q| \le 4|t|$.

Proof. This lemma is obvious if Δ has no cells. So assume Δ has cells. Recall p_- and p_+ denote the initial and terminal vertices, respectively, of a path p. Let s_1 be a simple path in Δ such that $(s_1)_- = t_-$, $(s_1)_+ \in q$, and $|s_1|$ is minimal. Let us show that

$$|s_1| < 1.04|t|. \tag{1}$$

By Lemmas 5.7 and 9.2, there is a θ-cell Π in Δ. Let $\Gamma_1, \Gamma_t, \Gamma_2, \Gamma_q$ be contiguity subdiagrams between Π and p_1, t, p_2, q, respectively, and $\psi_1, \psi_t, \psi_2, \psi_q$, denote the degrees of contiguity of Π to p_1, t, p_2, q, respectively. By Lemma 1, $\psi_1, \psi_2 \le \beta$, whence $\psi_t + \psi_q > \theta - 2\beta$.

By Lemmas 3.3 and 3.4, $\psi_t, \psi_q < \alpha$, whence $\psi_t, \psi_q > \theta - \alpha - 2\beta$. Then, by Lemmas 3.1 and 6.1 applied to Γ_t, we have

$$\rho(\theta - \alpha - 2\beta)|\partial\Pi| < \rho|\Gamma_t \wedge \Pi| < |t| + 2\gamma|\partial\Pi|.$$

Therefore,

$$|\partial\Pi| < \big(\rho(\theta - \alpha - 2\beta) - 2\gamma\big)^{-1}|t|. \tag{2}$$

Considering the bond between t and q consisting of Π, Γ_t, Γ_q, we see that there are vertices $o_t \in t$, $o_q \in q$ and a simple path $u = o_t - o_q$ such that

$$|u| < (1 - \theta + 2\beta + 2\gamma)|\partial\Pi| < \frac{1 - \theta + 2\beta + 2\gamma}{\rho(\theta - \alpha - 2\beta) - 2\gamma}|t| < 0.04|t|,$$

following from Lemma 3.1 and (2). The existence of such u obviously implies inequality (1).

Let $q = q_1 q_2$ be the factorization of q defined be $(s_1)_+$. Consider a subdiagram Δ_1 of Δ given by $\partial\Delta_1 = p_1 s_1 q_2$. If Δ_1 has no cells, then $|p_1| \le |s_1| < 1.04|t|$ for $q_2 p_1$ is a reduced path.

Assume Δ_1 has cells. Applying Lemmas 5.7 and 9.2, we find a θ-cell Π_1 in Δ_1. Let $\Gamma_p, \Gamma_s, \Gamma_{q_2}$ be contiguity subdiagrams of Π_1 to p_1, s_1, q_2, respectively, and $\psi_p, \psi_s, \psi_{q_2}$ denote the contiguity degrees of Π_1 to p_1, s_1, q_2, respectively. By Lemma 1, $\psi_p \le \beta$. By Lemmas 3.3 and 3.4 applied to Γ_{q_2}, $\psi_{q_2} \le \alpha$. Hence $\psi_s > \theta - \beta - \alpha$ for $\psi_p + \psi_s + \psi_{q_2} > \theta$. Therefore, by Lemmas 3.1 and 6.1 applied to Γ_s, we have

$$|\Gamma_s \wedge s_1| > (\rho\psi_s - 2\gamma)|\partial\Pi_1| > 0.3|\partial\Pi_1|.$$

On the other hand, considering the bond consisting of $\Pi_1, \Gamma_s, \Gamma_{q_2}$, we can construct a path s' with $s'_- = t_-$ and $s'_+ \in q$ such that $s' = s_{11}u'$, where s_{11} is a beginning of s_1 not containing $\Gamma_s \wedge s_1$ and

$$|u'| < (1 - \theta + \beta + 2\gamma)|\partial \Pi_1| < 0.02|\partial \Pi_1|,$$

following from Lemma 3.1.

However, the existence of such a path s' contradicts the choice of s_1. Hence, Δ_1 contains no cells and inequality $|p_1| < 1.04|t|$ is proven. Quite analogously, $|p_2| < 1.04|t|$.

Finally, by Lemma 6.1 applied to Δ,

$$|q| \le \rho^{-1}(|p_1| + |p_2| + |t|) < 4|t|$$

and Lemma 2 is proven. ∎

Let T_i be a word in the alphabet $\mathcal{A}$ (with no occurrences of $a_1^{-1}, \dots, a_m^{-1}$) such that $|T_i| = i$ and T_i is 3-aperiodic (e.g., see Lemma 1.5). Let us construct words $S_1, S_2, \dots$ as follows:

$$S_k \equiv a_2 a_1^3 a_2 T_{kL\Omega} a_2 a_1^3 a_2 T_{kL\Omega+1} a_2 a_1^3 a_2 T_{kL\Omega+2} \cdots a_2 a_1^3 a_2 T_{kL\Omega+h} a_2 a_1^3 a_2, \tag{3}$$

where $\Omega > 0$ is a fixed integer, $h = \beta^{-1} = 2^{14}$, and $L = 2h^2$.

Recall by writing $X \equiv YZ$ we mean the graphical (letter-by-letter) equality of words.

LEMMA 3. (a) *Suppose $S_k \equiv XYZ$, where $Y > 2h^{-1}|S_k|$, and $S_l \equiv X'YZ'$. Then $k = l$ and $X \equiv X'$, $Y \equiv Y'$.*

(b) *For every k the word S_k is 5-aperiodic.*

Proof. (a) Note $|S_k| > (h+1)kL\Omega$. Hence

$$2h^{-1}|S_k| > 2(1 + h^{-1})kL\Omega > 2(kL\Omega + h) + 15,$$

following from choice of L and h. Consequently, Y has a subword of S_k of the form

$$a_2 a_1^3 a_2 T_{kL\Omega+j} a_2 a_1^3 a_2 .$$

Now part (a) becomes obvious from the inequality $kL\Omega + h < (k+1)L\Omega$.

(b) Suppose E^5 is a subword of S_k and E^5 has a subword a_1^3. For $E^5 \ne a_1^3$, E^5 has a subword F^4 so that F starts with a_1^3. Then there are two different subwords of F^4 both equal to $a_1^3 T a_1^3$, where T has no subwords a_1^3. This, however, is a contradiction. Hence, a_1^3 does not occur in E^5. This means that E^5 is a subword of $a_1^2 a_2 T_{kL\Omega+j} a_2 a_1^2$ and the equality $E = 1$ follows from the 3-aperiodicity of $T_{kL\Omega+j}$. ∎

LEMMA 4. *Let* Δ *be a reduced diagram of rank* i, $\varphi(\partial\Delta)$ *be a cyclically reduced word, and* $\partial\Delta = s_1 t_1 s_2 t_2 \cdots s_l t_l$, *where* $l > 0$, s_j *is a subword of some* S_{k_j} *with* $|s_j| > (1 - 5h^{-1})|S_{k_j}|$, *and* $|t_j| \le \Omega$ *for all* $j = 1, \dots, l$. *Also suppose that for every* $j \pmod l$ *there are no* 0*-bonds between any* s_j *and* s_{j+1}, *and* $s_j t_j s_{j+1}$ *is geodesic in* Δ. *Then there exists a cell* Π *of rank* r *in* Δ *such that* $|A_r| > 0.2\beta^{-2}\Omega$ *and* $\partial\Pi$ *contains an arc* a *with the following properties*: $|a| > (\beta - 30n^{-1})|\partial\Pi|$ *and* $a^{-1} = u_1 v_1 \cdots u_{l'} v_{l'} u_{l'+1}$, *where* $l' > 0$, *each* u_j, $j = 1, \dots, l' + 1$, *is a subpath of* s_{d+j} [*indexes* $\pmod l$)] *for some* d *independent of* j *such that if* $s_{d+j} = s'_j u_j s''_j$, *then* $|s'_j|, |s''_j| < 2\Omega$, *and* $|v_j| < 4\Omega$ *for all* $j = 1, \dots, l'$. *In particular, every* $|u_j| > |s_{d+j}| - 4\Omega$.

Proof. First let us show that there are a cell Π in Δ and a contiguity subdiagram Γ of Π to $\partial\Delta$ such that $|\Gamma \wedge \Pi| > \beta|\partial\Pi|$ and $\Gamma \wedge \partial\Delta$ contains at least one of the sections $s_1, \dots, s_l$.

By Lemmas 5.7 and 6.2, there are a cell Π and a contiguity subdiagram Γ of Π to $\partial\Delta$ such that $|\Gamma \wedge \Pi| > \theta|\partial\Pi| > \beta|\partial\Pi|$. If $p = \Gamma \wedge \partial\Delta$ contains one of the sections $s_1, \dots, s_l$, we are done. Otherwise, p is obviously a subpath of $s_j t_j s_{j+1}$ for some j. Then, in view of Lemma 3.3, $s_j t_j s_{j+1}$ is not geodesic in Δ, contrary to the lemma's hypothesis.

Picking such Π, Γ with minimal $\tau(\Gamma)$, we can assume that Γ has no cell Π' and contiguity subdiagram Γ' with the foregoing properties.

Let $\partial\Gamma = bpcq$ be the standard contour of Γ, where $p = \Gamma \wedge \partial\Delta$, $q = \Gamma \wedge \Pi$. Assume there is a cell Π_0 in Γ that has a contiguity subdiagram Γ_0 to p with $|\Gamma_0 \wedge \Pi_0| > \beta|\partial\Pi_0|$.

LEMMA 4.A. (a) $\Gamma_0 \wedge p$ *is a subpath of* $s_j t_j s_{j+1}$ *for some* j *that has edges in common with* t_j.

(b) $|\Gamma_0 \wedge p| < 3.1\Omega$ *and* $|\partial\Pi_0| < 4\beta^{-1}\Omega$.

Proof. (a) This follows from the definition of a bond, choice of Π, Γ and Lemma 1.

(b) Let $\partial\Gamma_0 = b_0 p_0 c_0 q_0$ be the standard contour of Γ_0, where $p_0 = \Gamma_0 \wedge p$, $q_0 = \Gamma_0 \wedge \Pi_0$.

Keeping in mind part (a), we can write $p_0 = v_1 t v_2$, where v_1, t, v_2 are subpaths of s_j, t_j, s_{j+1}, respectively. (If p_0 is a subpath of, say, $s_j t_j$, then we assume $|v_2| = 0$ and so on.)

Proceeding by induction on $r(\Pi_0)$, let us first show that

$$|v_1|, |v_2| < 1.05\Omega. \tag{4}$$

Suppose E_1, E_2 are the bonds that define Γ_0.

Suppose first that $E_1 = E_2$. Then it follows from $|E_1 \wedge \Pi_0| > \beta n|\partial\Pi_0| > 1$ that E_1 is not a 0-bond and, in view of Lemma 3.2, we have

$r(\pi) < r(\Pi_0)$, where π is the principal cell of E_1. By the induction hypothesis applied to π and its contiguity subdiagram to p_0, part (b) is proven.

From now on assume $E_1 \neq E_2$. There are four cases to consider:

1. Both E_1 and E_2 are 0-bonds.
2. Precisely one of bonds E_1, E_2 is a 0-bond and for the other bond E the path $E \wedge p$ does not entirely contain t.
3. Precisely one of bonds E_1, E_2 is a 0-bond and for the other bond E the path $E \wedge p$ entirely contains t.
4. Neither E_1 nor E_2 is 0-bond.

Case 1. Consider a subdiagram Γ_0' of Γ_0 with $\partial\Gamma_0' = v_1' t v_2' q_0'$, where v_1', v_2' are some end, beginning of v_1, v_2, respectively, q_0' is a subpath of q_0, and the path $v_2' q_0' v_1'$ is reduced. It follows from hypotheses of Lemma 4 and Lemma 3(b) that Lemma 2 applies to Γ_0' and yields that

$$|v_1'|, |v_2'| \leq 1.04|t|, \qquad |q_0'| \leq 4|t|. \tag{5}$$

By Lemma 3(b),

$$|v_1| - |v_1'|, |v_2| - |v_2'| < 5|A_{j_0}|,$$

where $j_0 = r(\Pi_0)$. Then

$$|q_0| < |q_0'| + 10|A_{j_0}| < |q_0'| + 10(\beta n)^{-1}|q_0|.$$

Hence it follows from (5) that

$$\beta n|A_{j_0}| < |q_0| < 4\left(1 - 10(\beta n)^{-1}\right)^{-1}|t|$$

and so

$$5|A_{j_0}| < 20(\beta n)^{-1}\left(1 - 10(\beta n)^{-1}\right)^{-1}|t|.$$

Now we see from (5) that

$$|v_1|, |v_2| < 1.04|t| + 5|A_{j_0}| < 1.05|t| \leq 1.05\Omega,$$

as required in (4).

Case 2. For definiteness, let E_2 be a 0-bond. Denote the principal cell of E_1 by π_1. By induction hypothesis applied to π_1 and its contiguity subdiagram to p_0, we have $|v_1| < 1.05\Omega$.

Let $q_0 = q_{01} q_{02}$ be the factorization of q_0 defined by the vertex $(E_1 \wedge q_0)_-$. By Lemma 3.2, $|q_{02}| < (1+\varepsilon)|A_{j_0}|$, where $j_0 = r(\Pi_0)$. Hence

$$|q_{01}| > (\beta n - 1 - \varepsilon)|A_{j_0}|.$$

To estimate $|v_2|$, consider a subdiagram Γ_0' of Γ_0 such that $\partial\Gamma_0' = z v_2' q_{01}'$, where v_2' and q_{01}' are some beginning and end of v_2 and q_{01}, respectively, the path $v_2' q_{01}'$ is reduced, and $z = (v_2)_- - (q_{01}')_+$ is geodesic in Γ_0'. By Lemma 3(b),

$$\begin{aligned} |q_{01}'| &> |q_{01}| - 5|A_{j_0}| > (\beta n - 6 - \varepsilon)|A_{j_0}|, \\ |v_2'| &> |v_2| - 5|A_{j_0}|. \end{aligned} \tag{6}$$

Note it follows from Lemma 3.1 applied to E_1 that

$$|z| < 2\beta^{-1}|A_{j_0}| + |t| \le 2\beta^{-1}|A_{j_0}| + \Omega.$$

Hence, by Lemma 2 (in which $|p_2| = 0$) applied to Γ_0', we have

$$\begin{aligned} |q_{01}'| &\le 4|z| < 8\beta^{-1}|A_{j_0}| + 4\Omega, \\ |v_2'| &\le 1.04|z| < 2.1\beta^{-1}|A_{j_0}| + 1.04\Omega. \end{aligned} \tag{7}$$

It now follows from estimates (6) and (7) that

$$|A_{j_0}| < 4\big(\beta n - 8\beta^{-1} - 6 - \varepsilon\big)\Omega < 5\beta^{-1} n^{-1}\Omega,$$
$$|v_2| < \big(2.1\beta^{-1} + 5\big)|A_{j_0}| + 1.04\Omega.$$

Therefore,

$$|v_2| < \Big(1.04 + \big(2.1\beta^{-1} + 5\big)\cdot 5\beta^{-1}n^{-1}\Big)\Omega < 1.05\Omega,$$

as required.

Case 3. For definiteness, assume that E_2 is a 0-bond. By π_1 denote the principal cell of E_1. Let

$$p_0 = p_{01}p_{02}, \qquad q_0 = q_{01}q_{02}$$

be factorizations of p_0 and q_0 defined by the vertices $(E_1 \wedge p_0)_+$ and $(E_1 \wedge q_0)_-$, respectively.

Arguing as in Case 2, we can get estimates

$$|q_{02}| < (1+\varepsilon)|A_{j_0}|, \qquad |q_{01}| < 8\beta^{-1}|A_{j_0}| + 5|A_{j_0}|$$

following from Lemmas 3.2, 3.1, 3(b), and 2. Therefore,

$$|q_0| < \big(8\beta^{-1} + 6 + \varepsilon\big)|A_{j_0}| < \beta n|A_{j_0}|.$$

A contradiction to the inequality $|q_0| > \beta n|A_{j_0}|$ shows that this case is impossible.

Case 4. Let π_1 and π_2 denote the principal cells of bonds E_1 and E_2, respectively. According to part (a), both paths $E_1 \wedge p_0$ and $E_2 \wedge p_0$ have edges in common with t. Applying the induction hypothesis to π_1, π_2 and their contiguity subdiagrams to p_0, we have

$$|v_1|, |v_2| < 1.05\Omega,$$

as required.

Thus inequality (4) is proven. Now the inequality

$$|p_0| < 3.1\Omega$$

becomes obvious from $|t| \le \Omega$.

To estimate $|\partial\Pi_0|$ we apply Lemmas 3.1 and 6.1 to Γ_0 to get that

$$\rho\beta n|\partial\Pi_0| < \rho|q_0| < |p_0| + 2\gamma|\partial\Pi_0|.$$

Hence

$$|\partial\Pi_0| < (\rho\beta n - 2\gamma)^{-1} \cdot 3.1\Omega < 4\beta^{-1}\Omega.$$

Lemma 4.A is proven. ▮

LEMMA 4.B. *Suppose a section s_j of $\partial\Delta$ is a subpath of p. Then there is a 0-bond between s_j and q.*

Proof. Arguing on the contrary, consider a maximal system $\mathcal{C}$ of disjoint bonds $E_1, E_2, \dots, E_c$ in Γ between p and q (where E_1 and E_c are the bonds that define Γ; see Figs. 9.1 and 9.2 in [11]).

If $\mathcal{C}$ contains a bond E_k such that $E_k \wedge p$ is a subpath of s_j, then, by Lemma 4.A(a), E_k is a 0-bond. Hence, we may suppose that no $E_k \wedge p$ nor $E_k \in \mathcal{C}$ is a subpath of s_j. Then, keeping the notation of Lemma 9.1 relating to the system $\mathcal{C}$, we can consider a diagram Δ_t (sitting between E_t and E_{t+1} in Γ) such that, by Lemma 4.A(a), r_t is a subpath of $t_{j'-1}s_{j'} \cdots t_{j''-1}s_{j''}t_{j''}$, where $j' \le j \le j''$. It follows from Lemmas 3.1 and 4.A(b) that

$$|x_{t+1}|, |y_t| < (1 + 2\gamma) \cdot 4\beta^{-1}\Omega.$$

On the other hand, it follows from Lemma 4.A(b) that

$$|E_t \wedge p|, |E_{t+1} \wedge p| < 3.1\Omega,$$

whence

$$|r_t| \ge |s_j| - 6.2\Omega > (1 - 5h^{-1})|S_{k_j}| - 6.2\Omega > \big((1 - 5h^{-1})\beta^{-2} - 6.2\big)\Omega.$$

Note it follows from the choice of Γ, the geodesity of $s_{\bar{j}} t_{\bar{j}} s_{\bar{j}+1}$, and Lemma 3.3 that if a cell Π' of Γ has a contiguity subdiagram Γ' to p, then $|\Gamma' \wedge \Pi'| < \alpha|\partial\Pi'|$. This property of the section p of $\partial\Gamma$ ensures that if Δ' is a subdiagram of Γ with $\partial\Delta' = b'p'c'q'$, where p', q' are subpaths of p, q, respectively, then the analog of Lemma 6.5 applies to Δ' (for the geodesity or smoothness of either p', q' is used in proving Lemma 6.5 only to refer to Lemmas 3.3 and 3.4). By the analog of Lemma 6.5 applied to Δ_t and the maximality of C, we have $|r_t| < \mu(|x_{t+1}| + |y_t|)$. However,

$$\mu\big(|x_{x+1}| + |y_t|\big) < 2\mu(1 + 2\gamma) \cdot 4\beta^{-1}\Omega < \big((1 - 5h^{-1})\beta^{-2} - 6.2\big)\Omega.$$

A contradiction completes the proof of Lemma 4.B. ∎

As before, consider a maximal system C of disjoint bonds $E_1, E_2, \ldots, E_c$ between p and q in Γ.

Let s_{j_1+1} be the first section of $\partial\Delta$ that is entirely contained in p and let E_{k_1} be the bond in C such that E_{k_1} is the first 0-bond between s_{j_1+1} and q.

Put $p = p_1 p_2$ and $q = q_1 q_2$ are factorizations of p and q defined by the vertices $(E_{k_1} \wedge p)_-$ and $(E_{k_1} \wedge q)_+$, respectively. Let us estimate $|p_1|$. First notice that, by choice of s_{j_1+1} and Lemma 4.B, p_1 is a subpath of $s_{j_1} t_{j_1} s_{j_1+1}$. Therefore, if E_1 is a 0-bond, then the estimate

$$|p_1| < 3.08|t_{j_1}| + 5|A_r| < 5\big(\Omega + |A_r|\big),$$

where $r = r(\Pi)$, follows from Lemmas 2 and 3(b).

If E_1 is not 0-bond, then, by Lemma 4.A, $E_1 \wedge p$ has edges in common with t_{j_1} and so

$$|E_1 \wedge p| < 3.1\Omega, \qquad |\partial\pi_1| < 4\beta^{-1}\Omega, \tag{8}$$

where π_1 is the principal cell of the bond E_1.

Consider the subdiagram Δ_{k_1-1} of Γ "sitting" between E_{k_1-1} and E_{k_1}. By maximality of C, there are no bonds between p and q in Δ_{k_1-1}. Hence, applying Lemma 4.A to E_{k_1-1} to obtain the analogs of (8), we have from Lemmas 6.5 and 3.1 that the end p_{12} of p_1 given by

$$(p_{12})_- = \big(E_{k_1-1} \wedge p\big)_-$$

can be estimated as

$$|p_{12}| < \mu\big((1 + 2\gamma) \cdot 4\beta^{-1} + 1\big)\Omega + \Omega < 5.3\beta^{-1}\Omega.$$

Now it follows from Lemma 4.A and (8) that

$$|p_1| \le |E_1 \wedge p| + |t_{j_1}| + |p_{12}| < 3.1\Omega + \Omega + 5.3\beta^{-1}\Omega < 5.4\beta^{-1}\Omega.$$

Hence, in any case, it is proven that

$$|p_1| < \max\big(5(\Omega + |A_r|), 5.4\beta^{-1}\Omega\big). \tag{9}$$

It follows from Lemma 3.1 and inequalities (8) that

$$|b| < (1 + 2\gamma)|\partial\pi_1| < 4\beta^{-1}(1 + 2\gamma)\Omega.$$

Then, by Lemma 6.1 applied to a subdiagram of Γ whose contour is bp_1q_2, we have

$$|q_2| \le \rho^{-1}(|p_1| + |b|) < \rho^{-1}\max\big(5|A_r| + 5\beta^{-1}\Omega, 10\beta^{-1}\Omega\big). \tag{10}$$

Analogously, let s_{j_2+1} be the last section of $\partial\Delta$ that is contained in p and let E_{k_2} be the last 0-bond in $\mathcal{C}$ between s_{j_2+1} and q (perhaps, $j_1 = j_2$ and $k_1 = k_2$). Put $p = p_3p_4$ and $q = q_3q_4$ are factorizations of p and q defined by the vertices $(E_{k_2} \wedge p)_+$ and $(E_{k_2} \wedge q)_-$, respectively. Then, as before, we get

$$\begin{aligned} |p_4| &< \max\big(5(\Omega + |A_r|), 5.4\beta^{-1}\Omega\big), \\ |q_3| &< \rho^{-1}\max\big(5|A_r| + 5\beta^{-1}\Omega, 10\beta^{-1}\Omega\big). \end{aligned} \tag{11}$$

Now let us make a couple of remarks.

First, if E' and E'' are 0-bonds in Γ between some section s_j of p and q and Δ' is a subdiagram of Γ with $\partial\Delta' = p'q'$, where p', q' are subpaths of p, q so that $p'_- = (E' \wedge p)_+$ and $p'_+ = (E' \wedge p)_-$, then, by Lemmas 5.7, 3.4, and 1, Δ' has no cells, that is, $p' = (q')^{-1}$. In particular, $|q'| < 5|A_r|$ following from Lemma 3(b).

Second, if p entirely contains two consecutive sections s_j and s_{j+1} of $\partial\Delta$, $E_{k_3} \in \mathcal{C}$ is the last 0-bond between s_j and q, $E_{k_4} \in \mathcal{C}$ is the first 0-bond between s_{j+1} and q (E_{k_3}, E_{k_4} exist by Lemma 4.B), and Δ' given by $\partial\Delta' = p'q'$ is defined as before, then, by Lemma 2, $|p'|, |q'| \le 4\Omega$.

Assume that p entirely contains exactly l_0 sections $s_{j_1+1}, \dots, s_{j_1+l_0}$ of $\partial\Delta$, where $s_{j_1+l_0} = s_{j_2+1}$. Making use of the preceding remarks and inequalities (10) and (11), we can estimate $|q|$ as

$$|q| < 3\max\big(5|A_r| + 5\beta^{-1}\Omega\big), 10\beta^{-1}\Omega\big) + l_0(4\Omega + 5|A_r|). \tag{12}$$

If $l_0 \le 2$, we have

$$\beta n|A_r| < |q| < 5\max\big(5|A_r| + 10\beta^{-1}\Omega\big), 10\beta^{-1}\Omega\big).$$

Therefore,

$$|A_r| < 5\max\big(5\beta^{-1}(\beta n - 5)^{-1}\Omega, 10\beta^{-1}n^{-1}\Omega\big) < \Omega.$$

Now we see from (9), (11), and the foregoing remarks that s_{j_1+1} contains a subpath u such that u^{-1} is a subpath of p with

$$|u| > |s_{j_1}| - |p_1| - \max(|p_4|, 4\Omega) > (1 - 5h^{-1})|S_{k_{j_1}}| - 20\beta^{-1}\Omega > 5\Omega > 5|A_r|.$$

A contradiction to Lemma 3(b) shows that $l_0 > 2$. Thus, s_{j_1+2} is different from s_{j_1+1} and s_{j_2+2}, and so, by the foregoing remarks, s_{j_1+2} contains a subpath u such that u^{-1} is a subpath of p with

$$|u| > |s_{j_1+2}| - 8\Omega > (1 - 5h^{-1})|S_{k_{j_1+2}}| - 8\Omega > \beta^{-2}\Omega.$$

For $|u| < 5|A_r|$ we have

$$|A_r| > 0.2\beta^{-2}\Omega. \tag{13}$$

Hence, $\rho^{-1} \cdot 10\beta^{-1}\Omega < 60\beta|A_r|$ and so inequalities (10) and (11) imply that

$$|q_2|, |q_3| < 6|A_r|.$$

Let $E_{k_5} \in \mathsf{C}$ be the first 0-bond between s_{j_1+2} and q, and let $E_{k_6} \in \mathsf{C}$ be the last 0-bond between s_{j_2} and q. Consider the contiguity subdiagram Γ' between p and q that is defined by E_{k_5} and E_{k_6}. Let $\partial\Gamma' = p_2' q_2'$ and $q = q_1' q_2' q_3'$. Then, similar to (12), we have from (13) that

$$|q_1'| + |q_3'| < 3\max\big(5|A_r| + 10\beta^{-1}\Omega, 14\beta^{-1}\Omega\big) + 2\big(4\Omega + 5|A_r|\big) < 30|A_r|.$$

It is now clear that

$$|q_2'| > (\beta - 30n^{-1})|\partial\Pi|, \qquad (q_2')^{-1} = u_1 v_1 \cdots u_{l'} v_{l'} v_{l'+1}$$

with all of the properties of the lemma's statement. In view of inequality (13), Lemma 4 is completely proven. ∎

From now on we will be denoting the free Burnside group $B(m, n)$ of exponent n over the alphabet $\mathsf{A} = \{a_1, \ldots, a_m\}$ by $B(\mathsf{A}, n)$.

Let us introduce a countably infinite alphabet $\mathsf{X} = \{x_1, x_2, \ldots\}$ for use throughout the rest of the article. (We assume that A and X are disjoint.)

By $B(\mathsf{X}, n)$ denote the free Burnside group over X of exponent n.

As before, a word in $\mathsf{X}^{\pm 1}$ is also referred to as an X-*word*.

LEMMA 5. *Suppose* G *is a finite subgroup of* $B(\mathsf{A}, n)$. *Then there are* A*-words* $S_1, S_2, \dots$ *such that the subgroup* $\langle \mathsf{G}, S_1, S_2, \dots \rangle$ *of* $B(\mathsf{A}, n)$ *is isomorphic to the subgroup* $\langle \mathsf{G}, x_1, x_2, \dots \rangle$ *of* $B(\mathsf{A} \cup \mathsf{X}, n)$ *under* $\Psi(T) = T$ *if* $T \in \mathsf{G}$ *and* $\Psi(x_1) = S_1$, $\Psi(x_2) = S_2, \dots$.

Proof. Let G also stand for the set of all reduced in $B(\mathsf{A}, n)$ words that represent elements of G. Put

$$\Omega' = \max_{T \in \mathsf{G}} |T|.$$

Consider the words $S_1, S_2, \dots$ defined by formulas (3) for $\Omega = 4\Omega'$ and denote the set of all such $S_k^{\pm 1}$ by S.

By $W(\mathsf{G}, \mathsf{S})$ denote a word $\mathsf{G} \cup \mathsf{S}$ with no subwords of the form $G_1 G_2$, where $G_1, G_2 \in \mathsf{G}$ are nonempty. The image of $W(\mathsf{G}, \mathsf{S})$ under $a \to a$, $a \in \mathsf{A}$, $S_k \to x_k$ is denoted by $W(\mathsf{G}, \mathsf{X})$.

By $\tau_\Omega(\Delta)$ denote a truncated type of a diagram Δ over $B(\mathsf{A}, n)$: $\tau_\Omega(\Delta) = (l_{j_\Omega}, l_{j_\Omega + 1}, \dots)$, where j_Ω is the first rank for which $|A_{j_\Omega}| > 0.2\beta^{-2}\Omega$, and l_j is the number of cells of rank j in Δ.

By induction on $\tau_\Omega(\Delta)$, we will be proving that if Δ is a reduced diagram over $B(\mathsf{A}, n)$ such that $\varphi(\partial\Delta)$ is a word of the form $W(\mathsf{G}, \mathsf{S})$ and $W(\mathsf{G}, \mathsf{X})$ is freely reduced, then $W(\mathsf{G}, \mathsf{X}) = 1$ in $B(\mathsf{A} \cup \mathsf{X}, n)$.

Put $\partial\Delta = q_1 p_1 \cdots q_l p_l$, where $\varphi(q_j) \in \mathsf{S}$, $\varphi(p_j) \in \mathsf{G}$, $j = 1, \dots, l$. For each p_j pick a reduced diagram Δ_j over $B(\mathsf{A}, n)$ such that $\partial\Delta_j^{-1} = x_j p_j y_j t_j^{-1}$, where $\varphi(x_j)$ is an end of $\varphi(q_j)$, $\varphi(y_j)$ is a beginning of $\varphi(q_{j+1})$, the word $\varphi(t_j)$ is reduced in $B(\mathsf{A}, n)$, the word $\varphi(y_j t_j^{-1} x_j)$ is freely reduced, and Δ_j with preceding properties is maximal relative to $|x_j| + |y_j|$.

Suppose there is a 0-bond between x_j^{-1} and y_j^{-1} in Δ_j. Then we consider a maximal contiguity subdiagram Γ between x_j^{-1} and y_j^{-1} in Δ_j relative to $|\Gamma \wedge x_j^{-1}| + |\Gamma \wedge y_j^{-1}|$. It is easy to see from Lemmas 1 and 5.7 that Γ has no cells. Since t_j, p_j are geodesic in Δ_j, it follows from Lemmas 5.7, 3.3, and 1 that the entire Δ_j has no cells either. Then $|t_j| = 0$ for $y_j t_j^{-1} x_j$ is reduced. If, say, $|x_j| > 2h^{-1}|S_{k_j}| + |p_j|$, then $|\Gamma \wedge x_j^{-1}| > 2h^{-1}|S_{k_j}|$ and, by Lemma 3(a), we conclude that $\varphi(x_j) \equiv \varphi(y_{j+1})^{-1}$. However, then, by $\varphi(t_j) \equiv 1$, $W(\mathsf{G}, \mathsf{X})$ is not reduced, contrary to the hypothesis. Analogously,

$$|y_j| \le 2h^{-1}|S_{k_j}| + |p_j| < 2.1h^{-1}|S_{k_j}|.$$

Now suppose there is no 0-bond between x_j^{-1} and y_j^{-1} in Δ_j. Then Lemma 2 applied to Δ_j yields that

$$|x_j|, |y_j| \le 1.04|p_j|, \qquad |t_j| \le 4|p_j| \le \Omega.$$

Hence, by Lemma 6.2, if π is a cell in Δ_j, then

$$|\partial\pi| < (1 + 3.08/4)\rho^{-1}\Omega < 2\Omega$$

and so $r(\pi) < j_\Omega$.

Let us attach each Δ_j to Δ along $x_j p_j y_j$ producing a diagram Δ'. Let s_j be a subpath of q_j so that $q_j = y_{j-1}s_j x_j$ in Δ'. Clearly, $\partial\Delta' = s_1 t_1 s_2 t_2 \cdots s_l t_l$, where, by the preceding estimates, $|t_j| < \Omega$, $|s_j| < (1 - 5h^{-1})|S_{k_j}|$ and $\tau_\Omega(\Delta') = \tau_\Omega(\Delta)$. It is readily seen from the maximality of each Δ_j relative to $|x_j| + |y_j|$ that $\varphi(\partial\Delta')$ is reduced. Removing reducible pairs of cells in Δ' results in a reduced diagram Δ^0 which satisfies all of the assumptions of Lemma 4. [Note that $s_j t_j s_{j+1}$ is geodesic follows from the choice of Δ_j; also, $\tau_\Omega(\Delta^0) \le \tau_\Omega(\Delta)$.] Let Π be a cell of rank r in Δ^0 with an arc a, where $|a| > (\beta n - 30)|A_r|$ and $a^{-1} = u_1 v_1 \cdots u_{l'} v_{l'} u_{l'+1}$ as described in the conclusion of Lemma 4. Note it follows from Lemma 4 that we can consider a subdiagram Γ of Δ defined by $\partial\Gamma = ab$, where b is a subpath of $\partial\Delta$, and if π is a cell in Γ, then

$$|\partial\pi| < \rho^{-1}(4\Omega + 2\Omega + \Omega) < 8\Omega$$

following from Lemma 6.2. Consequently, we have from $1 - 5h^{-1} \gg 2h^{-1}$ and Lemma 3(a) that if A is a cyclic permutation of $A_r^{\pm 1}$ beginning with u_1 and $S_k^{\pm 1} \equiv C\varphi(u_1)D$ [we assume $\varphi(q_1) \equiv S_k^{\pm 1}$], then CAC^{-1} is equal in rank $< j_\Omega$ to a word $V(\mathsf{G}, \mathsf{S})$. Now it is clear that we can take Π and a part of Γ out of Δ^0 and, making use of cells of rank $< j_\Omega$, get a reduced diagram Δ^r with $\tau_\Omega(\Delta^r) < \tau_\Omega(\Delta^0) \le \tau_\Omega(\Delta^0)$ and $\varphi(\partial\Delta^r)$ is a word in $\mathsf{G} \cup \mathsf{S}$. Using diagrams of rank $< j_\Omega$, we can also assume that cyclic word $\varphi(\partial\Delta^r)$ has no subwords of the form $G_1 G_2$ with nonempty $G_1, G_2 \in \mathsf{G}$, that is, $\varphi(\partial\Delta^r) = W_r(\mathsf{G}, \mathsf{S})$ and $W_r(\mathsf{G}, \mathsf{X})$ is reduced. By construction, we have in the group $B(\mathsf{A} \cup \mathsf{X}, n)$ that $W_r(\mathsf{G}, \mathsf{X}) = W(\mathsf{G}, \mathsf{X})T^n$ with some word T. By the induction hypothesis, $W_r(\mathsf{G}, \mathsf{X}) = 1$ in $B(\mathsf{A} \cup \mathsf{X}, n)$ and so $W(\mathsf{G}, \mathsf{X}) = 1$ in $B(\mathsf{A} \cup \mathsf{X}, n)$ as well. ∎

LEMMA 6. *The subgroup* $\langle x_{2k-1}^{n/2} x_{2k}^{n/2} \mid k = 1, 2, \ldots \rangle$ *of* $B(\mathsf{X}, n)$ *is isomorphic to* $B(\mathsf{X}, n)$ *under the map* $x_{2k-1}^{n/2} x_{2k}^{n/2} \to x_k$, $k = 1, 2, \ldots$.

Proof. Denote the word $x_{2k-1}^{n/2} x_{2k}^{n/2}$ by Y_k and put $\mathsf{Y} = \{Y_1, Y_2, \ldots\}$. It suffices to show that if $W(\mathsf{Y}) = 1$ in $B(\mathsf{X}, n)$, then $W(\mathsf{X}) = 1$ in $B(\mathsf{X}, n)$ as well.

We will say that a nonempty cyclically reduced X-word U is an *almost* Y*-word* provided a finite sequence of replacements of the form $x_l^{\pm n/2} \to x_l^{\mp n/2}$ of some subwords $x_l^{\pm n/2}$ in U yields a reduced Y-word.

Assume $W(\mathsf{Y})$ is a nonempty cyclically reduced (as a Y-word) word with $|W(\mathsf{Y})| > 0$ and $W(\mathsf{Y}) = 1$ in $B(\mathsf{X}, n)$.

By $\overline{W}(\mathsf{Y})$ denote any word obtained from $W(\mathsf{Y})$ by a finite sequence of replacements of the form $x_l^{\pm n/2} \to x_l^{\mp n/2}$, where $x_l^{\pm n/2}$ is a subword of

$W(Y)$. Clearly, any such $\overline{W}(Y)$ is also cyclically reduced and is an almost Y-word.

By $\tau_2(\Delta)$ denote a truncated type of a diagram Δ now defined as follows: $\tau_2(\Delta) = (l_k, l_{k+1}, \dots)$, where l_j is the number of cells in Δ of rank j and k is the first rank for which $|A_k| > 1$ (this means that we do not count cells π with $|\partial\pi| = n$).

Consider a reduced diagram Δ over $B(X, n)$ with $\varphi(\partial\Delta) = \overline{W}(Y)$ [this means that we pick Δ over such $B(i) = B(X', n)(i)$, where $X' \subset X$ is finite, that $\overline{W}(Y) = 1$ in $B(i)$] and Δ is minimal relative to $\tau(\Delta)$ over all such $\overline{W}(Y)$.

By induction on $\tau_2(\Delta)$, we will prove that $W(X) = 1$ in $B(X, n)$.

By Lemma 18.1, there is a cell Π of some rank j in Δ such that $\partial\Pi = uv$, $u > \beta n|\partial\Pi|$, and u^{-1} is a subpath of $\partial\Delta$. If $|A_j| = 1$, then we have $\varphi(u) = A_j^{\pm n/2}$ and a contradiction to choice of Δ is immediate. Hence $|A_j| > 1$. In this case, it is easy to see from $\beta n \gg 1$ that some cyclic permutation A of $A_j^{\pm 1}$ is an almost Y-word. Therefore, we may take Π and a part of u whose label is an almost Y-word out of Δ and get a diagram Δ' such that $\tau_2(\Delta') < \tau_2(\Delta)$ and $\varphi(\partial\Delta') \equiv \overline{W}'$ with $\overline{W}'$ an almost Y-word. It is easy to see that we can obtain a reduced almost Y-word $\overline{W}''$ from $\overline{W}'$ by making all possible cancellations in cyclic $\overline{W}'$ and using relations $x_k^n = 1$. Performing corresponding surgery on Δ' (and removing reducible pairs) results in a reduced diagram Δ'' such that $\tau_2(\Delta'') < \tau_2(\Delta)$ and $\varphi(\partial\Delta'') \equiv \overline{W}''$. Changing $\overline{W}''$ by another $\overline{\overline{W}}''$ as before if necessary, we may further assume that Δ'' is minimal relative to $\tau(\Delta'')$. Now the induction hypothesis applies to the pair Δ'', $\varphi(\partial\Delta'') \equiv \overline{W}''$ and, in view of our construction, yields the required equalities. ∎

LEMMA 7. *Suppose G is a finite 2-subgroup of $B(A, n)$. Then the centralizer $C_{B(A\cup X, n)}(G)$ of G in $B(A \cup X, n)$ contains a subgroup $H = \langle H_1, H_2, \dots\rangle$ such that H is isomorphic to the subgroup $\langle x_1^{n/2}, x_2^{n/2}, \dots\rangle$ of $B(X, n)$ under the map $H_k \to x_k^{n/2}$, $k = 1, 2, \dots$. In addition, $G \cap H = \{1\}$ and $C_{B(A\cup X, n)}(H) = G$.*

Proof. We will say that a word W in $(A \cup X)^{\pm 1}$ *strongly depends on* $x \in X$ if W is a word in $(A \cup \{x\})^{\pm 1}$, $W \neq 1$ in $B(A \cup X, n)$, and $\pi_A(W) = 1$, where $\pi_A\colon B(A \cup X, n) \to B(A, n)$ is the natural projection.

LEMMA 7.A. *For every $x_k \in X$ there is an A_{x_k}-word H_k, where $A_{x_k} = A \cup \{x_k\}$, such that H_k strongly depends on x_k, $H_k^2 = 1$ in $B(A \cup X, n)$, and $H_k \in C_{B(A\cup X, n)}(G)$.*

Proof. If G is trivial we put $H_k = x_k^{n/2}$. Proceeding by induction on $|G|$, denote by G_0 a normal subgroup in G of index 2 and put $G = \langle G_0, T\rangle$, where $T \in G$ is an A-word. By the induction hypothesis, there is a word H such that H has all of the properties stated in the lemma relative to G_0.

Consider the subgroup $G_1 = \langle G, H \rangle$. Clearly, G_0 is normal in G_1 and the images of H and T have order 2 in the quotient G_1/G_0.

It follows from $\pi_A(HT) = T$ and the assumptions that G is a 2-group and $T \notin G_0$ that the images of H, T generate a dihedral group modulo G_0 of some order $2d$, where d is even. Therefore, we can consider $H' = (HT)^{d/2}T^{-d/2}$. Clearly, the order of H' in G_1/G_0 is 2, H' commutes with T in G_1/G_0, and H' strongly depends on x_k. Hence, raising H' to a suitable power, we will get a word H_k such that H_k strongly depends on x_k, has order 2 in $B(A \cup X, n)$, and H_k normalizes G. However, then H_k also centralizers G following from $\pi_A(H_k) = 1$ and Lemma 7.A is proven. ∎

LEMMA 7.B. *Suppose*

$$H_{k_1} \cdots H_{k_l} = 1 \tag{14}$$

in $B(A \cup X, n)$. *Then* $x_{k_1}^{n/2} \cdots x_{k_l}^{n/2} = 1$ *in* $B(X, n)$.

Proof. Let Δ be a diagram over $B(A \cup X, n)$ (with the same meaning as in proving Lemma 6). By $\tau_X(\Delta)$ denote an X-*type* of Δ defined as follows: $\tau_X(\Delta) = (l_1, l_2, \dots)$, where l_j is the number of cells Π of rank j in Δ provided A_j^n is not A_x-word for any $x \in X$ (i.e., $A_{l_{k_t}}^n$ contains at least two distinct $x^{\pm 1}, y^{\pm 1} \in X$); otherwise, put $l_j = 0$.

Let us associate with Eq. (14) a reduced diagram Δ over $B(A \cup X, n)$ such that $\varphi(\partial\Delta)$ is a word of the form

$$B_1 C_1 B_2 C_2 \cdots B_l C_l, \tag{15}$$

where for every $t \pmod l$ the following are true: C_t is a word in $A_{x_{k_t}}^{\pm 1}$ and B_t is a word in $A^{\pm 1}$ such that there are words B_{t1}, B_{t2} in $A^{\pm 1}$ so that $B_t = B_{t1}B_{t2}$ in $B(A, n)$ and $B_{t2}C_tB_{t+11} = H_{k_t}$ in $B(A_{k_t}, n)$. Note it is possible that for every t the word B_t is empty, $C_t \equiv H_{k_t}$ and then $\varphi(\partial\Delta)$ turns into (14).

For such diagram Δ, by induction on $\tau_X(\Delta)$, we prove that $x_{k_1}^{n/2} \cdots x_{k_l}^{n/2} = 1$ in $B(X, n)$.

First notice that if C_t contains no occurrences of $x_{k_t}^{\pm 1}$, then H_{k_t} equals in $B(A \cup X, n)$ to an A-word. However, then $H_{k_t} = 1$ in $B(A \cup X, n)$ for H_{k_t} strongly depends on x_{k_t} [recall $\pi_A(H_{k_t}) = 1$ in $B(A, n)$]. Hence every C_t does contain occurrences of $x_{k_t}^{\pm 1}$.

Second, making use of relators of groups $B(A_{x_{k_t}}, n)$, $t = 1, \dots, l$, we can bring Δ with $\varphi(\partial\Delta)$ of the form (15) back to Δ' with $\varphi(\partial\Delta')$ of the form (14) and $\tau_X(\Delta') = \tau_X(\Delta)$.

Also, note that $H_t^2 = 1$ in $B(A_{x_{k_t}}, n)$ obviously follows from $H_t^2 = 1$ in $B(A \cup X, n)$.

These remarks, in particular, mean that we can assume that for every t (mod l), $k_t \neq k_{t+1}$ in (15).

Third, changing B_t, C_t if necessary, we can suppose that the word (15) is cyclically reduced and every C_t both begins and ends with $x_{k_t}^{\pm 1}$.

Fourth, we can assume that Δ contains no cell π such that $\varphi(\partial\pi)$ is an A-word and $\partial\pi$ has an edge e with $e^{-1} \in \partial\Delta$ and that Δ has no cell Π such that $\varphi(\partial\Pi)$ is an $\mathsf{A}_{x_{k_t}}$-word and $\partial\Pi$ has an edge e with $e^{-1} \in \partial\Delta$ and $\varphi(e) = x_{k_t}^{\pm 1}$ [for otherwise, we would take such π or Π along with e out of Δ changing B_t, C_t but not changing $\tau_{\mathsf{X}}(\Delta)$; note this is the main reason for replacing (14) by (15)].

To have the right to suppose that Δ (after all of the foregoing described transformations) is still reduced, we need to show that removal of reducible pairs in Δ does not increase the X-type $\tau_{\mathsf{X}}(\Delta)$. To do this, we first make a change in the definition of j-compatibility of sections q_1, q_2 (see [11, p. 13]): In addition to (A1) and (A2), we require that there is a contiguity subdiagram E between q_1 and q_2 such that E contains t,

$$|E \wedge q_1|, |E \wedge q_2| > N|A_j| = 484|A_j|,$$

(E) is a reduced diagram of rank $j-1$, and $E \wedge q_1$ and $|E \wedge q_2|$ are smooth sections of rank j of ∂E. In particular, $|q_1|, |q_2| > N|A_j|$ (otherwise, q_1, q_2 may not be compatible).

The definition of weak j-compatibility remains unchanged.

It is not difficult to check [see first Lemmas 18.5(c) and 19.1] that proofs of all lemmas [11] are retained. Taking advantage of the new definition of reducible pairs, we can prove the following lemma.

LEMMA 7.C. *It is possible to remove a reducible pair of cells in a diagram Δ over $B(\mathsf{A} \cup \mathsf{X}, n)$ so that the X-type $\tau_{\mathsf{X}}(\Delta)$ does not increase.*

Proof. Let Π_1, Π_2 be cells of rank j in Δ and let t be a path that makes Π_1, Π_2 be j-compatible. Assume that $\varphi(\partial\Pi_1)$, $\varphi(\partial\Pi_2)$ are A_x-words for some $x \in \mathsf{X}$. It follows from Lemma 18.4.3 and the definition of a bond that the contiguity subdiagram E_t between Π_1 and Π_2 that contains $\varphi(t)$ has the contour ∂E labelled by an A_x-word. Then, by a routine argument based on Lemma 18.1, for every cell π in E, $\varphi(\partial\pi)$ is an A_x-word. Then $\varphi(t)$ and so

$$\varphi(\partial\Pi_1)\,\varphi(t)\,\varphi(\partial\Pi_2)^{\pm 1}\,\varphi(t)^{-1}$$

are A_x-words. Therefore, a reduced diagram of rank j_1 for the equality

$$\varphi(\partial\Pi_1)\,\varphi(t)\,\varphi(\partial\Pi_2)^{\pm 1}\,\varphi(t)^{-1} \overset{j-1}{=} 1$$

contains only such cells π that $\varphi(\partial\pi)$ are A_x-words following from the same argument as before based on Lemma 18.1. This proves Lemma 7.C. ∎

Thus, by Lemma 7.C, we still can assume that Δ is a reduced diagram.

There is nothing to prove if $l = 0$. Let $l > 0$; that is, let the word (15) be nonempty. Since such Δ has cells, we can apply Lemma 18.1 to Δ and get a cell Π so that $\partial\Pi = uv$, where u^{-1} is a subpath of $\partial\Delta$ and $|u| > \beta|\partial\Delta|$. Denote $r(\Pi) = j$ and $\varphi(\partial\Pi) = A_j^{\pm n}$. Since $\beta n \gg 1$, it follows from choice of Δ that A_j contains at least two distinct $x^{\pm 1}, y^{\pm 1} \in X$. Therefore, there is a cyclic permutation A of $A_j^{\pm 1}$ such that A^s, where $s > [\beta n] - 2 > 3$, is a subword of cyclic word (15) such that

$$A^s \equiv \left(C_{t_1} B_{t_1+1} \cdots C_{t_1+t_2} B_{t_1+t_2+1}\right)^s, \tag{16}$$

meaning that $C_{t_1+t_2+1} \equiv C_{t_1}$, $B_{t_1+t_2+2} \equiv B_{t_1+1}$, and so forth.

Taking the subpath of $\partial\Delta \cap \partial\Pi$ labelled by A^{s-2} (located in the "middle" of the path labelled by A^s) and Π out of Δ, we will get a diagram Δ_1 such that $\varphi(\partial\Delta_1)$ has the form (15) and $\tau_X(\Delta_1) < \tau_X(\Delta)$. It remains to note that, by the induction hypothesis, the word (14), that corresponds to $\varphi(\partial\Delta_1)$ and is rewritten in $x_k^{n/2}$, equals 1 in $B(X, n)$ and differs from the word (14), that corresponds to $\varphi(\partial\Delta)$ and is rewritten in $x_k^{n/2}$, by a word conjugate in $B(X, n)$ to $(x_{k_{t_1}}^{n/2} \cdots x_{k_{t_1+t_2}}^{n/2})^n$. This implies that $x_{k_1}^{n/2} \cdots x_{k_l}^{n/2} = 1$ in $B(X, n)$ and Lemma 7.B is proven. ∎

Now the main claim of Lemma 7 becomes obvious from Lemma 7.B. Clearly, $G \cap H = \{1\}$ for each H_k strongly depends on x_k. Lemma 7 is proven. ∎

LEMMA 8. *Suppose K is a finite 2-subgroup of $B(A, n)$, G is a subgroup of index 2 in K, and D is an A-word so that $K = \langle G, D\rangle$. Then the centralizer $C_{B(A\cup X, n)}(G)$ of G in $B(A \cup X, n)$ contains a subgroup $H = \langle H_2, H_3, \dots\rangle$ such that the quotient group $\langle K, H\rangle/G$, where $\langle H, K\rangle \subseteq B(A \cup X, n)$, is isomorphic to the subgroup $\langle x_1^{n/2}, x_2^{n/2}, \dots\rangle$ of $B(X, n)$ under the map $D \to x_1^{n/2}$, $H_k \to x_k^{n/2}$, $k = 2, 3, \dots$. In addition, $G \cap H = \{1\}$.*

Proof. We will repeat the proof of Lemma 7 with necessary changes, keeping most of the notation and terminology introduced there.

Suppose W is a word in $(A \cup X)^{\pm 1}$. We will say that W is *G-regular* provided for every maximal A-subword V (relative to $|V|$) of the word W we have $V \in G$ in $B(A, n)$. We will also say that W is *cyclically G-regular* if $W = 1$ provided W is an A-word or, otherwise, for every maximal A-subword V (relative to $|V|$) of the cyclic word W we have $V \in G$ in $B(A,n)$.

LEMMA 8.A. *For every* $x_k \in \mathsf{X}$ *there is an* A_{x_k}*-word* H_k, *where* $\mathsf{A}_{x_k} = \mathsf{A} \cup \{x_k\}$, *such that* H_k *strongly depends on* x_k, $H_k^2 = 1$ *in* $B(\mathsf{A} \cup \mathsf{X}, n)$, $H_k \in C_{B(\mathsf{A} \cup \mathsf{X}, n)}(\mathsf{G})$, *and* H_k *is* G*-regular.*

Proof. Repeating the proof of Lemma 7.A, we note that the extra property of H_k is ensured by its explicit construction. ∎

To unify notation, put $H_1 = D$ and let $H_2, H_3, \ldots$ be unchanged.

LEMMA 8.B. *Suppose*

$$H_{k_1} \cdots H_{k_l} = G_0 \tag{17}$$

in $B(\mathsf{A} \cup \mathsf{X}, n)$, *where* $G_0 \in \mathsf{G}$ *is an* A*-word. Then* $x_{k_1}^{n/2} \cdots x_{k_l}^{n/2} = 1$ *in* $B(\mathsf{X}, n)$.

Proof. Let us modify the definition of X-type $\tau_{\mathsf{X}}(\Delta)$ of a diagram Δ over $B(\mathsf{A} \cup \mathsf{X}, n)$ as follows: $\tau_{\mathsf{X}}(\Delta) = (l_1, l_2, \ldots)$, where l_j is the number of cells Π of rank j in Δ provided either A_j^n is not A_x-word for any $x \in \mathsf{X}$ or A_j^n is an A_x-word for some $x \in \mathsf{X}$ such that A_j^n is not cyclically G-regular; otherwise, put $l_j = 0$.

As before, associate with Eq. (17) a reduced diagram Δ over $B(\mathsf{A} \cup \mathsf{X}, n)$ such that $\varphi(\partial\Delta)$ is a word of the form

$$B_1 C_1 B_2 C_2 \cdots B_l C_l, \tag{18}$$

where for every $t \pmod l$ the following are true:

(1) C_t is a G-regular word in $\mathsf{A}_{x_{k_t}}^{\pm 1}$ provided $k_t > 1$ and C_t is an A-word provided $k_t = 1$.

(2) B_t is a word in $\mathsf{A}^{\pm 1}$ such that there are A-words B_{t1}, B_{t3}, B_{t2} so that $B_t = B_{t1} B_{t3} B_{t2}$ in $B(\mathsf{A}, n)$ and either $B_{t2} C_t B_{t+11} = H_{k_t}$ modulo all cyclically G-regular relators of $B(\mathsf{A}_{k_t}, n)$, provided $k_t > 1$, or $B_{t2} C_t B_{t+11} = H_{k_t}$ in $B(\mathsf{A}, n)$, provided $k_t = 1$.

(3) The words B_{t1}, B_{t3}, B_{t2} introduced in part (2) have the following additional properties: For every t, B_{t1}, B_{t2} are G-regular; $B_{t3} = G_0^{-1}$ in $B(\mathsf{A}, n)$ if $t = 1$ and $B_{t3} \equiv 1$ if $t \neq 1$.

For such diagram Δ, by induction on $\tau_{\mathsf{X}}(\Delta)$, we prove that $x_{k_1}^{n/2} \cdots x_{k_l}^{n/2} = 1$ in $B(\mathsf{X}, n)$.

As before, every C_t does contain occurrences of $x_{k_t}^{\pm 1}$, provided $t > 1$, following from strong dependence of C_t on x_{k_t}.

Suppose there is a $t \pmod l$ such that $k_t = k_{t+1}$ in (18). If $k_t = 1$, then using only cells that correspond to relators of $B(\mathsf{A}, n)$, we can bring Δ to Δ' with $\varphi(\partial\Delta')$ having form (18) with $l' = l - 2$. Clearly, $\tau_{\mathsf{X}}(\Delta') = \tau_{\mathsf{X}}(\Delta)$. To consider the case $k_t > 1$ we need

LEMMA 8.C. *Suppose W is a cyclically G-regular word in $(\mathsf{A} \cup \mathsf{X})^{\pm 1}$, $W = 1$ in $B(\mathsf{A} \cup \mathsf{X}, n)$ and Δ is a reduced diagram over $B(\mathsf{A} \cup \mathsf{X}, n)$ for this equality. Then for every cell Π in Δ the word $\varphi(\partial\Pi)$ is cyclically G-regular.*

Proof. Without loss of generality, we may assume W to be cyclically reduced. However, then the claim can be obtained by a routine argument based on Lemma 18.1. ∎

Coming back to the case $k_t > 1$, we see from Lemma 8.C that it suffices to use only the cells that correspond to cyclically G-regular relators of $B(\mathsf{A}_{x_{k_t}}, n)$ to bring Δ to Δ' with $\varphi(\partial\Delta')$ of the form (18) where $l' = l - 2$. Again, $\tau_{\mathsf{X}}(\Delta') = \tau_{\mathsf{X}}(\Delta)$.

Therefore, we can assume that for every $t \pmod l$, $k_t \neq k_{t+1}$ in (18).

As before, by changing B_t, C_t if necessary, we can assume that the word (18) is cyclically reduced and every C_t both begins and ends with $x_{k_t}^{\pm 1}$ provided $k_t > 1$.

We can also suppose that Δ contains no cell π such that $\varphi(\partial\pi)$ is an A-word and $\partial\pi$ has an edge e with $e^{-1} \in \partial\Delta$; neither does Δ have cell Π such that $\varphi(\partial\Pi)$ is a G-regular $\mathsf{A}_{x_{k_t}}$-word and $\partial\Pi$ has an edge e with $e^{-1} \in \partial\Pi$ and $\varphi(e) = x_{k_t}^{\pm 1}$ [otherwise, we could take such π or Π along with e out of Δ and suitably change B_t, C_t without changing $\tau_{\mathsf{X}}(\Delta)$].

At last, we change the definition of j-compatibility as was done in the proof of Lemma 7 to be able to prove

LEMMA 8.D. *It is possible to remove a reducible pair of cells in a diagram Δ over $B(\mathsf{A} \cup \mathsf{X}, n)$ so that the X-type $\tau_{\mathsf{X}}(\Delta)$ does not increase.*

Proof. First let us prove three auxiliary lemmas (8.D.1–8.D.3).

LEMMA 8.D.1. *Suppose Δ is a reduced diagram over $B(\mathsf{A} \cup \mathsf{X}, n)$ such that $\partial\Delta = qt$, where q is either smooth of some rank j or q is geodesic in Δ and $\varphi(t)$ is G-regular. Then $\varphi(q)$ is also G-regular.*

Proof. Without loss of generality we can assume that t is a reduced path. Arguing as in the proof of Lemma 18.1, we can get from Lemmas 5.7, 3.3, and 3.4 that there is a cell Π in Δ such that $\partial\Pi = uv$, u^{-1} is a subpath of t, and $|u| > \beta n |\partial\Pi|$. It follows from the inequality $\beta n > 2$ that $\varphi(\partial\Pi)$ is cyclically G-regular. It remains to take Π and u out of Δ and apply the induction hypothesis. ∎

LEMMA 8.D.2. *Suppose Δ is a reduced diagram over $B(\mathsf{A} \cup \mathsf{X}, n)$ such that $\partial\Delta = bpcq$, where b, c are geodesic in Δ and if $U \in \{\varphi(p), \varphi(q)\}$, then U has the following properties: U is G-regular; if $U = V$ in $B(\mathsf{A} \cup \mathsf{X}, n)$, then V is not A-word and $|V| > 3\max(|b|, |c|)$. Then for every cell Π in Δ the word $\varphi(\partial\Pi)$ is cyclically G-regular.*

Proof. Proceeding by induction on $\tau(\Delta)$, first note that, without loss of generality, we may assume $\varphi(p)$, $\varphi(q)$ to be reduced words. If Δ contains no cells, i.e., $\tau(\Delta) = 0$, then the claim follows from the assumption that $\varphi(p)$, $\varphi(q)$ are not A-words.

Suppose Δ contains cells. By Lemmas 5.7 and 9.2, there is a θ-cell in Δ. Denote the contiguity subdiagrams of Π to b, p, c, q by $\Gamma_b, \Gamma_p, \Gamma_c, \Gamma_q$, respectively, and their contiguity degrees by ψ_b, ψ_p, ψ_c, ψ_q.

Suppose $\psi_p > \beta$ (the case $\psi_q > \beta$ is analogous). Then it follows from the definition of contiguity subdiagrams that there is a cell π in Γ_p such that $\partial\pi = uv$ and u^{-1} is a subpath of p with $|u| > \beta|\partial\pi|$. It is clear from $\beta n > 2$ that $\varphi(\partial\pi)$ is cyclically G-regular. It remains to take π along with u out of Δ and apply the induction hypothesis.

Therefore, we can assume that $\psi_p, \psi_q \le \beta$. Then $\psi_p + \psi_q > \theta - 2\beta$. Applying Lemmas 3.1, 3.3, and 6.1 to Γ_p, Γ_c, we have

$$(\rho(\theta - 2\beta) - 4\gamma)|\partial\Pi| < |b| + |c|,$$

whence

$$|\partial\Pi| < (\rho(\theta - 2\beta) - 4\gamma)^{-1}(|b| + |c|) < 1.2(|b| + |c|).$$

By Lemma 3.3, $\psi_b, \psi_c < \alpha$. Hence $\psi_b, \psi_c > \theta - 2\beta - \alpha$ and we can consider a bond E between b and c consisting of Π, Γ_b, Γ_c. Then, by definitions and Lemma 3.1, there is a path s in Δ homotopic to p such that

$$|s| < |b| + |c| + (1 - \theta + \beta)|\partial\Pi| < 1.02(|b| + |c|).$$

Since $\varphi(s) = \varphi(p)$ in $B(\mathsf{A} \cup \mathsf{X}, n)$, it follows from the lemma's hypothesis that $|s| > 3\max(|b|, |c|)$. A contradiction completes the proof of Lemma 8.D.2. ∎

LEMMA 8.D.3. *Suppose Δ is a reduced diagram over $B(\mathsf{A} \cup \mathsf{X}, n)$ with $\partial\Delta = bpcq$, p, q are smooth A-periodic sections, where A is a period of some rank i, so that $|p|, |q| > N|A|$, A^n is cyclically G-regular and not A-word, Δ itself is a contiguity subdiagram between p and q, and $o_1 = (e_p)_\pm$ and $o_2 = (e_q)_\pm$, where $e_p \in p$ and $e_q \in q$ are some edges so that $\varphi(e_p), \varphi(e_q) \in \mathsf{X}^{\pm 1}$. Then there is a path $r = o_1 - o_2$ in Δ such that $\varphi(r)$ is G-regular.*

Proof. Applying Lemmas 9.1 and 5.4 to Δ if necessary and keeping their notation, we can assume that $\Delta = \Delta(1, k)$ is rigid and, by Lemma 9.5,

$$|q(k,1)| > (N - 4.4)|A|,$$

$$|q(k,1)| < 2N|A|, \tag{19}$$

$$|b|, |c| < 0.003|A|, \qquad |x_{t'}|, |y_{t''}| < \xi|A|, \tag{20}$$

where t' and t'' satisfy $1 < t' \le k$ and $1 \le t'' < k$, and $\partial E_t = x_t p_t y_t q_t$, $t = 1, \dots, k$.

Suppose $\bar{b}$ is a geodesic in Δ that is homotopic to b. Consider a subdiagram Δ_b of Δ given by $\partial\Delta_b = (\bar{b})^{-1} b$. Assume that Δ_b contains a vertex of a bond $E_l \in \mathbf{C} = \mathbf{C}(\Delta)$ with $l > 1$. Then we have the following estimates from Lemma 6.1 and the inequality $|\bar{b}| \le |b|$:

$$\operatorname{dis}_q\big((E_l \wedge q)_+, q_+\big) \le \rho^{-1}\big(|b| + |x_l|\big) < 0.004|A|.$$

In view of inequalities (19) and Lemmas 9.3 and 9.5, there is a subdiagram of $\Delta = \Delta(1, k)$ of the form $\Delta(m_1, k)$ so that

$$|q(k, 1)| - |q(k, m_1)| < (1 + \varepsilon + 3\xi + 0.004)|A| < 1.1|A|$$

and $\Delta(m_1, k)$ has no vertices in common with Δ_b.

Quite analogously, consider a geodesic path $\bar{c}$ homotopic to c and a subdiagram Δ_c of Δ given by $\partial\Delta_c = c(\bar{c})^{-1}$. As before, there is a subdiagram $\Delta(1, m_2)$ of $\Delta = \Delta(1, k)$ such that

$$|q(k, 1)| - |q(m_2, 1)| < 1.1|A|$$

and $\Delta(1, m_2)$ has no vertices in common with Δ_c.

Hence the diagram $\Delta(m_1, m_2)$ is contained in the diagram $\bar{\Delta}$ given by $\partial\bar{\Delta} = \bar{b}p\bar{c}q$ and

$$\begin{aligned} &|q(m_2, m_1)| > |q(k, 1)| - 2.2|A| > (N - 6.6)|A|, \\ &|q(m_2, m_1)| < 2N|A|. \end{aligned} \tag{21}$$

It follows from inequalities (19) and Lemmas 6.1 and 18.4.1 that Lemma 8.D.2 applies to the diagram $\Delta'(m_1, m_2)$, where $\Delta'(m_1, m_2)$ is obtained from $\Delta(m_1, m_2)$ by $A^{\pm 1}$-extensions of $p(m_1, m_2)$ and $q(m_2, m_1)$ on both sides of length $< |A|$ (and corresponding lengthening of x_{m_1} and y_{m_2}) in order to get

$$\partial\Delta'(m_1, m_2) = x'_{m_1} p'(m_1, m_2) y'_{m_2} q'(m_2, m_1),$$

where $\varphi(p'(m_1, m_2))$ and $\varphi(q'(m_2, m_1))$ are $\mathbf{G}$-regular words.

By Lemma 8.D.2, for every cell π in $\bar{\Delta}$ and hence in $\Delta'(m_1, m_2)$ we have that the word $\varphi(\partial\pi)$ is cyclically $\mathbf{G}$-regular.

To simplify notation, rename $\Delta'(m_1, m_2)$ by Δ' and the sections $x'_{m_1}, p'(m_1, m_2), y'_{m_2}, q'(m_2, m_1)$ of $\partial\Delta'(m_1, m_2)$ by b', p', c', q', respectively. Thus $\partial\Delta' = b'p'c'q'$.

Note that if $\partial\Delta'$ has an edge e_0 so that $\varphi(e_0) \in X^{\pm 1}$ and $e_0 \in p'$, $e_0^{-1} \in q'$, then Lemma 18.D.3 is true following from G-regularity of $\varphi(p')$, $\varphi(q')$. Assume there are a sequence of cells $\pi_1, \pi_2, \dots, \pi_l$ in Δ' and a sequence of edges $e_j, f_j \in \partial\pi_j$, $j = 1, \dots, l$, so that $\varphi(e_j), \varphi(f_j) \in X^{\pm 1}$, $e_1^{-1} \in q'$, $e_2 = f_1^{-1}, \dots, e_l = f_{l-1}^{-1}$, $f_l^{-1} \in p'$. Then it is easy to see that Lemma 18.D.3 is again true following from cyclic G-regularity of words $\varphi(\partial\pi_1), \dots, \varphi(\partial\pi_l)$ and G-regularity of $\varphi(p')$, $\varphi(q')$.

Therefore, we can assume that there are no such edges e_0 and sequences of cells π_j and edges e_j, f_j in Δ'.

Consider an edge $e \in \partial\Delta'$ with $\varphi(e) \in X^{\pm 1}$. By $\Delta(e, X)$ denote a subdiagram of Δ' such that $\Delta(e, X)$ has the following properties [(1P) and (2P)] and $\Delta(e, X)$ is maximal relative to $\tau(\Delta(e, X))$:

(1P) If π is a cell in $\Delta(e, X)$, then there are sequences of cells $\pi_1, \pi_2, \dots, \pi_l$ in $\Delta(e, X)$ and edges $e_1, f_1, \dots, e_l, f_l$ such that $e_j, f_j \in \partial\pi_j$, $\varphi(e_j), \varphi(f_j) \in X^{\pm 1}$, and $e_1^{-1} = e$, $e_2 = f_1^{-1}, \dots, e_l = f_{l-1}^{-1}$, $f_l^{-1} \in \partial\pi$.

(2P) The contours of $\Delta(e, X)$ are cyclically reduced paths.

It is immediate from the definition of $\Delta(e, X)$ that the following are true: $\Delta(e, X)$ is connected (but need not be simply connected); if s is a contour of $\Delta(e, X)$ that bounds a bounded hole in $\Delta(e, X)$, then $\varphi(s)$ is an A-word; if s is the contour of $\Delta(e, X)$ that does not bound a bounded hole in $\Delta(e, X)$, then for every edge $e \in s$ with $\varphi(e) \in X^{\pm 1}$ we have $e \in \partial\Delta'$.

Let q_0' be a subpath of q' such that $|q_0'| = |A|$ and q_0' is a "middle" of q', that is, if $q' = q_1' q_0' q_2'$, then

$$|q_1'|, |q_2'| \geq (|q'| - |A|)/2 - 1 > \frac{N - 10}{2}|A| \tag{22}$$

following from (21) and construction of Δ'.

Let e' be an edge of q_0' with $\varphi(e') \in X^{\pm 1}$ (recall A is not A-word).

Consider a diagram $\Delta(e', X)$ and denote its contour that bounds the unbounded component of the complement to $\Delta(e', X)$ by s'. Note that if $(e')^{-1} \in b'$ or $(e')^{-1} \in c'$, then we can get an easy contradiction to inequalities (22) by making use of Lemma 6.1 and estimates

$$|b'|, |c'| < 2.003|A|$$

that follow from (20) and construction of Δ'.

Hence it follows from the foregoing assumption (about some edges and sequences of cells and edges) that $\Delta(e', X)$ contains cells and s' has no edge e such that $\varphi(e) \in X^{\pm 1}$ and $e \in p'$.

Suppose s' has no edge e with $\varphi(e) \in X^{\pm 1}$ and $e \in b'$ or $e \in c'$. Then there is a disk subdiagram Γ in Δ' such that $\partial\Gamma = q_3' t$, where q_3' is a subpath of q' containing e' and $\varphi(t)$ is an A-word. Since q_3' is smooth in Γ and the word $\varphi(q_3')$ contains $\varphi(e') \in X^{\pm 1}$, we have a contradiction to Lemma 18.4.2.

Therefore, we may assume that there is an edge $e_b \in s'$ such that $\varphi(e_b) \in X^{\pm 1}$ and $e_b \in b'$. First suppose that there is no $e_c \in s'$ such that $\varphi(e_c) \in X^{\pm 1}$ and $e_c \in c'$. (The case when such e_c exists and e_b does not is analogous.) Then there is a disk subdiagram Γ in $\Delta(m_1, m_2)$ given by

$$\partial\Gamma = q_2(m_2, m_1)b_1 t,$$

where $q(m_2, m_1) = q_1(m_2, m_1)q_2(m_2, m_1)$, the path $q_2(m_2, m_1)$ contains the edge e', $x_{m_1} = b_1 b_2$, and t is a reduced path in $\Delta(m_1, m_2)$ with $\varphi(t)$ an A-word. It follows from (19), (22), and construction of Δ' from $\Delta(m_1, m_2)$ that

$$|q_2(m_2, m_1)| > \frac{N-14}{2}|A|, \qquad |b'| \le |x_{m_1}| < 0.003|A|. \tag{23}$$

Picking such Γ with minimal $\tau(\Gamma)$, we have that if π is a cell in Γ with $\varphi(\partial\pi)$ an A-word, then $\partial\pi^{-1}$ has no edges in common with t [for otherwise, we could decrease $\tau(\Gamma)$ by changing t]. Note it follows from Lemma 18.4.3 that if Π' is a cell in Γ and Γ' is a contiguity subdiagram of Π' to t with $|\Gamma' \wedge \Pi'| > \beta|\partial\Pi'|$, then $\varphi(\partial\Gamma'), \varphi(\partial\Pi')$ are A-words. Then, by a routine argument based on Lemma 18.1, Γ' consists of cells π with $\varphi(\partial\pi)$ an A-word. A contradiction to choice of Γ shows that there are no such Π', Γ' in Γ. Then, as in proving Lemma 1(b), it follows from Lemma 5.7 that t is geodesic in Γ. By Lemma 6.5 applied to Γ, we have from (23) the existence of a bond E between t and $q_2(m_2, m_1)$. Since any bond between t and $q_2(m_2, m_1)$ is a 0-bond, we can assume that

$$\mathrm{dis}_{q_2(m_2, m_1)}\big((E \wedge q_2(m_2, m_1))_+, q_2(m_2, m_1)_+\big) < \mu \cdot 0.003|A| < 0.1|A|$$

following from the second inequality in (23). Now we see from the first inequality in (23) that there is a subdiagram Γ_0 in Γ such that $\partial\Gamma_0 = q_3' t'$, where q_3' is a subpath of $q_2(m_2, m_1)$ containing e' and $\varphi(t')$ an A-word. A contradiction to Lemma 18.4.2 completes this case.

At last, suppose that there are edges $e_b, e_c \in s'$ such that $\varphi(e_b), \varphi(e_c) \in X^{\pm 1}$ and $e_b \in b'$, $e_c \in c'$. Then we can consider a disk subdiagram Γ in $\Delta(m_1, m_2)$ given by

$$\partial\Gamma = b_1 t c_2 q(m_2, m_1),$$

where $x_{m_1} = b_1 b_2$, $y_{m_2} = c_1 c_2$, and t is a reduced path in $\Delta(m_1, m_2)$ with $\varphi(t)$ an A-word. Picking such Γ minimal relative to $\tau(\Gamma)$, we can repeat the preceding argument to prove the existence of two distinct 0-bonds in Γ between t and $q(m_2, m_1)$ such that the contiguity subdiagram defined by the bonds contains e'. A contradiction to Lemma 18.4.2 completes the proof of Lemma 8.D.3. ∎

Let us prove Lemma 8.D. Suppose some cells Π_1, Π_2 form a reducible j-pair in a diagram Δ, E is a contiguity subdiagram between Π_1 and Π_2 that contains the path t with $|t| < \delta|A_j|$ which makes Π_1, Π_2 be j-compatible.

If $\varphi(\partial \Pi_1)$ is A-word, then $\varphi(\partial E)$ is also A-word and so, by the same argument as in proving Lemma 8.C, $\varphi(t)$ is A-word. Now it remains to refer to Lemmas 18.4.2 and 8.C to draw the required conclusion.

Assume $\varphi(\partial \Pi_1)$ is cyclically G-regular but not A-word. Cyclically permutting A_j if necessary, suppose that A_j begins with a letter $x \in \mathsf{X}^{\pm 1}$. Then, by Lemma 8.D.3, we can replace t by a homotopic to t path t' (for t_-, t_+ are phase vertices of $\partial \Pi_1, \partial \Pi_2$, respectively) such that $\varphi(t')$ is G-regular. Now Lemma 8.D follows from Lemma 8.C. ∎

Coming back to proving Lemma 8, we still can assume Δ to be reduced following from Lemma 8.D.

Assuming $l > 0$ (otherwise the claim is obvious), we apply Lemma 18.1 to Δ to get a cell Π with $\partial \Pi = uv$, where u^{-1} is a subpath of $\partial\Delta$ and $|u| > \beta|\partial\Delta|$. Put $r(\Pi) = j$ and $\varphi(\partial\Delta) = A_j^{\pm n}$. Since $\beta n > 2$, it follows from choice of Δ that A_j contains at least one $x^{\pm 1} \in \mathsf{X}$ and, if A_j contains occurrences of precisely one $x^{\pm 1} \in \mathsf{X}$, A_j^n is not G-regular. Then it is not difficult to see that there is a cyclic permutation A of $A_j^{\pm 1}$ such that A^s, where $s > [\beta n] - 2 > 3$, is a subword of cyclic word (18) such that A begins with some $x^{\pm 1} \in \mathsf{X}$ and

$$A^s \equiv \left(C_{t_1} B_{t_1+1} \cdots C_{t_1+t_2} B_{t_1+t_2+1}\right)^s,$$

meaning that $C_{t_1+t_2+1} \equiv C_{t_1}$, $B_{t_1+t_2+2} \equiv B_{t_1+1}$, and so forth. Then it is clear that A begins with $x_{k_{t_1}}^{\pm 1}$, $t_1 > 1$, and A^n is not G-regular by the properties of (18).

If the subword $(C_{t_1} B_{t_1+1} \cdots C_{t_1+t_2} B_{t_1+t_2+1})^s$ of (18) contains no B_1, then we can repeat the corresponding argument in proving Lemma 7.B to get the required conclusion from the induction hypothesis [note $\tau_{\mathsf{X}}(\Delta_1) < \tau_{\mathsf{X}}(\Delta)$ now follows from the fact that A_j^n is not G-regular]. Otherwise, we put $A^s \equiv A^{s_1} A^{s_2}$, where $|s_1 - s_2| \le 1$ and pick that A^{s_i} which contains no B_1. Since $[[\beta n] - 2]/2 > 3$, it is possible to argue as previously to finish the proof of Lemma 8.B. ∎

The first claim of Lemma 8 becomes obvious from Lemma 8.B. Clearly, $\mathsf{K} \cap \mathsf{H} = \{1\}$ for each H_k, $k \ge 2$, strongly depends on x_k. Lemma 8 is proven. ∎

LEMMA 9. *Let $X = X_1 \cup X_2$ be a disjoint union of nonempty subsets X_1, X_2. Suppose X is an X_1-word, $X \neq 1$ in $B(X_1, n)$, and Y, Z are X_2-words such that $Y \neq 1$, $Z \neq 1$, and $Y \neq Z$ in $B(X_2, n)$. Then the subgroup $\langle X, Y, Z\rangle$ of $B(X, n)$ is infinite.*

Proof. By Lemma 1.7, for every $j > 0$ we can consider a 3-aperiodic word $D_j(a_1, a_2)$ in the alphabet $\{a_1, a_2\}$ whose length is j. By $D_j(a_1, a_2, a_3)$ denote a word in $\{a_1, a_2, a_3\}$ of length $2j+1$ such that if $D_j(a_1, a_2) \equiv a_{k_1}a_{k_2} \cdots a_{k_j}$ then

$$D_j(a_1, a_2, a_3) \equiv a_3 a_{k_1} a_3 a_{k_2} \cdots a_3 a_{k_j} a_3.$$

Consider the word $D_j(Y, Z, X)$. Assume that

$$D_{j_1}(Y, Z, X)^{-1} D_{j_2}(Y, Z, X) = 1 \tag{24}$$

in $B(X, n)$ with $j_1 < j_2$.

Assuming that $D_{j_1}(a_1, a_2)$ is a beginning of $D_{j_2}(a_1, a_2)$ let us make all possible cancellations in the cyclic word $D_{j_1}(Y, Z, X)^{-1} D_{j_2}(Y, Z, X)$ regarding it as a word in $\{Y, Z, X\}^{\pm 1}$ and denote thus obtained cyclically reduced word by

$$U_1 \cdots U_k, \tag{25}$$

where $U_i \in \{Y, Z, X\}$. It follows from definitions that $U_1 \in \{Y, Z\}$ and U_a is an $X_{1+a \pmod 2}$-word.

Consider a reduced diagram Δ over $B(X, n)$ such that $\varphi(\partial\Delta)$ is the word (25). If Π is a cell in Δ such that $\varphi(\partial\Pi)$ is X_1- or X_2-word and $\partial\Pi$ has an edge e so that $e^{-1} \in \partial\Delta$, then we take Π and e out of Δ, thereby decreasing $\tau(\Delta)$.

Therefore, we can assume that

$$\varphi(\partial\Delta) = U_1' \cdots U_k',$$

where $U_i' = U_i$ in $B(X_{i+1 \pmod 2}, n)$ and Δ has no cell Π such that $\varphi(\partial\Pi)$ is X_1- or X_2-word and $\partial\Pi$ has an edge e with $e^{-1} \in \partial\Delta$. Clearly, we may suppose the words U_i', V_j' to be reduced.

Arguing as in the proof of Lemma 18.1, we can find a cell Π of rank j in Δ such that $\partial\Pi = uv$, where u^{-1} is a subpath of $\partial\Delta$ of length $> \beta n/2$ (the coefficient $1/2$ shows up because of possible cancellations in $U_k' U_1'$). In addition, $\varphi(u)^{-1}$ is a subword of the (noncyclic) word $U_1' \cdots U_k'$.

Since $\varphi(\partial\Pi)$ is neither X_1- nor X_2-word and $\beta n/2 \gg 1$, we can find a cyclic permutation A of $A_j^{\pm 1}$ such that A begins with a letter in $X_1^{\pm 1}$, ends with a letter in $X_2^{\pm 1}$, and A^s with $s > \beta n/2 - 3 > 5$ is a subword of $U_1' \cdots U_k'$. It follows from these properties of A^s that the "middle" subword A^{s-2} of $A^s \equiv AA^{s-2}A$ regarded as a subword of $U_1' \cdots U_k'$ has the form $(U_t' \cdots U_{t+d}')^{s-2}$, meaning that $U_{t+d+1}' \equiv U_t'$, $U_{t+d+2}' \equiv U_{t+1}'$, and so forth. However, then the word $D_{j_1}(a_1, a_2)$ obviously contains a nonempty subword E^{s-2} with $s - 2 > 3$. This contradiction proves that equality (24) is impossible and therefore the subgroup $\langle Y, Z, X\rangle$ is infinite. ∎

LEMMA 10. *In conditions of Lemma* 8, *the subgroup* $\mathsf{H}_0 = \langle H_{2k}H_{2k+1} \mid k = 1, 2, \dots\rangle$ *of* $C_{B(\mathsf{A}\cup\mathsf{X}, n)}(\mathsf{G})$ *is isomorphic to* $B(\mathsf{X}, n)$ *under the map* $H_{2k}H_{2k+1} \to x_k$ *and the following is true*: *Suppose* $E_1, E_2 \in \mathsf{H}_0$ *and* $E_1 \neq 1$, $E_2 \neq 1$, $E_1 \neq E_2$ *in* $B(\mathsf{A}\cup\mathsf{X}, n)$. *Then the subgroup* $\langle D, E_1, E_2\rangle$ *of* $B(\mathsf{A}\cup \mathsf{X}, n)$ *is infinite*.

Proof. The first claim is immediate from Lemmas 6 and 8. Furthermore, in view of Lemmas 6 and 8, it suffices to show that if $X = x_1^{n/2}$ and Y, Z are X-words that are nontrivial and distinct in $B(\mathsf{X}, n)$ and have no occurrences of $x_1^{\pm 1}$, then the subgroup $\langle X, Y, Z\rangle$ of $B(\mathsf{X}, n)$ is infinite. This, however, follows from Lemma 9. ∎

LEMMA 11. *Suppose* G *is a finite subgroup of* $B(\mathsf{A}, n)$ *and* G *contains a nontrivial element of odd order*. *Then*

(a) G *is contained in a unique maximal finite subgroup* $F(\mathsf{G})$, *and* $F(\mathsf{G})$ *contains* $C_{B(\mathsf{A}, n)}(\mathsf{G})$.

(b) *Either* $C_{B(\mathsf{A}, n)}(\mathsf{G})$ *is a* 2-*group or*, *otherwise*, $C_{B(\mathsf{A}, n)}(C_{B(\mathsf{A}, n)}(\mathsf{G}))$ *is contained in* $F(\mathsf{G})$.

Proof. (a) Suppose G has height j. By Lemma 15.2, we can assume that G is a subgroup of $\mathsf{K}(A_j)$ and contains $A_j^{n_2 l}$ with some $l \not\equiv 0 \pmod{n_1}$. Suppose $\mathsf{K}(A_j)$ is contained in a finite subgroup G_1 of $B(\mathsf{A}, n)$. By Lemma 15.2, G_1 is conjugate in $B(\mathsf{A}, n)$ to a subgroup of $\mathsf{K}(A_{j_1})$, where $j_1 = h(\mathsf{G}_1)$. Since the word $A_j^{n_2 l} \in \mathsf{G}$ has odd order > 1 by Lemma 10.3(a), we have from Lemma 15.2 (applied to G_1) that $j_1 = j$ and so $\mathsf{G}_1 = \mathsf{K}(A_j)$, that is, $\mathsf{K}(A_j)$ is a maximal finite subgroup of $B(\mathsf{A}, n)$, as required.

Let us show that

$$C_{B(\mathsf{A}, n)}(\mathsf{G}) \subseteq \mathsf{K}(A_j).$$

Suppose $X \in C_{B(\mathsf{A}, n)}(\mathsf{G})$ and Δ is a disk diagram of rank $i > j$ for the equality

$$XA_j^{n_2 l}X^{-1}A_j^{-n_2 l} = 1$$

so that $\partial\Delta = xq_1yq_2$, where

$$\varphi(x) = \varphi(y)^{-1} = X, \qquad \varphi(q_1) = \varphi(q_2)^{-1} = A_j^{n_2 l}.$$

Let Δ_0 be an annular diagram obtained from Δ by gluing x and y^{-1}. By Lemma 10.3(a), $\varphi(q_1) = A_j^{n_2 l}$ has odd order > 1 in $B(\mathsf{A}, n)$. On the other hand, by Lemma 10.4(b), any word $A_{j'}^{k'}U'$, where U' is an $\mathsf{F}(A_{j'})$-involution, has order $2^k > 1$ in $B(\mathsf{A}, n)$. Consequently, we have from the definition of a reducible cell that Δ_0 may not contain reducible cells and that any diagram obtained from Δ_0 by removal of reducible pairs may not contain reducible cells either. It now follows from Lemma 10.8(a) that $X \in \mathsf{K}(A_j)$, as required.

(b) Keeping the notation introduced previously, first assume that there is an $\mathsf{F}(A_j)$-involution U in G. Let us show that in this case $C_{B(\mathsf{A}, n)}(\mathsf{G})$ is a 2-group. Arguing on the contrary, note that $A_j^{n/2} \in C_{B(\mathsf{A}, n)}(\mathsf{G})$ following from Lemma 15.10. Hence, $h(C_{B(\mathsf{A}, n)}(\mathsf{G})) = j$ and, by part (a) and Lemma 15.2, $C_{B(\mathsf{A}, n)}(\mathsf{G})$ contains $A_j^{n_2 l'}$ with some $l' \not\equiv 0 \pmod{n_1}$. Then

$$UA_j^{n_2 l'}U^{-1} = A_j^{-n_2 l'}F$$

in $B(\mathsf{A}, n)$ with some $F \in \mathsf{F}(A_j)$. However, $A_j^{-n_2 l'}F \neq A_j^{n_2 l'}$ by Lemma 10.1(a). This contradiction proves that $C_{B(\mathsf{A}, n)}(\mathsf{G})$ is a finite 2-group.

Suppose there are no $\mathsf{F}(A_j)$-involutions in G, that is, $\mathsf{G} \subseteq \langle A_j, F(A_j)\rangle$. It follows from Lemmas 15.7 and 15.8 that $A_j^{n_2} \in C_{B(\mathsf{A}, n)}(\mathsf{G})$, whence

$$A_j^{n_2} \in C_{B(\mathsf{A}, n)}\big(\langle A_j, \mathsf{F}(A_j)\rangle\big),$$

$$A_j^{n_2} \in C_{B(\mathsf{A}, n)}(\mathsf{G}).$$

Now we can apply the argument of proving part (a) to $C_{B(\mathsf{A}, n)}(\mathsf{G})$ and obtain that $C_{B(\mathsf{A}, n)}(C_{B(\mathsf{A}, n)}(\mathsf{G}))$ is also a subgroup of $\mathsf{K}(A_j)$, as desired. ▮

LEMMA 12. *Suppose S is not a locally finite subgroup of $B(\mathsf{A}, n)$. Then its centralizer $C_{B(\mathsf{A}, n)}(\mathsf{S})$ is finite.*

Proof. Arguing on the contrary, assume $C_{B(\mathsf{A}, n)}(\mathsf{S})$ is infinite and so contains distinct and reduced in $B(\mathsf{A}, n)$ words W_k, $k = 1, 2, \dots$. By $X_1, \dots, X_{l_0} \in \mathsf{S}$ denote some words such that the subgroup $\langle X_1, \dots, X_{l_0}\rangle$ of $B(\mathsf{A}, n)$ is infinite. Let Δ_{lk} be a reduced diagram over $B(\mathsf{A}, n)$ so that

$$\varphi(\partial\Delta_{lk}) = X_l W_k X_l^{-1} W_k^{-1},$$

$l = 1, \dots, l_0$, $k = 1, 2, \dots$.

Denote the strict rank $r(\Delta_{lk})$ by r_{lk}. First assume that the set $\{r_{lk} \mid l = 1, \dots, l_0,\ k = 1, 2, \dots\}$ is unbounded. Put

$$s = \max\big\{h(X_{l_1}), h(X_{l_1}X_{l_2}) \mid 1 \le l_1, l_2 \le l_0\big\}$$

and j_1 to be the first rank for which the following inequalities hold:

$$|A_{j_1}| > \tfrac{1}{7}(n/2 + 2\delta)|A_2|,$$
$$|A_{j_1}| > \tfrac{1}{7} \max_{1 \le l \le l_0} |X_l|.$$

Let $\Delta_{l'k'}$ be a diagram such that $|A_{r_{l'k'}}| \ge |A_{j_1}|$ and

$$j_0 = \max_{1 \le l \le l_0} r_{lk'},$$
$$j_0 = r_{\bar{l}k'}.$$

By definitions,

$$|A_{j_0}| > \tfrac{1}{7}(n/2 + 2\delta)|A_s|,$$
$$|A_{j_0} > \tfrac{1}{7} \max_{1 \le l \le l_0} |X_l|. \tag{26}$$

Taking the diagram $\Delta_{\bar{l}k'}$ as the diagram $\Delta(1)$ in the hypothesis of Lemma 13.2 and any $\Delta_{lk'}$ as $\Delta(2)$, in view of inequalities (26), we have, by Lemma 13.2, that the subgroup $\langle X_1, \dots, X_{l_0} \rangle$ is conjugate [by $\varphi(x)^{-1}$ in notation of Lemma 13.2] to a subgroup of $\mathsf{K}(A_j)$. However, $\mathsf{K}(A_j)$ is finite, contrary to the infiniteness of $\langle X_1, \dots, X_{l_0} \rangle$.

Hence, it is proven that for all k, l we have $r(\Delta_{lk}) < i$ for some i. Then the centralizer $C_{B(i)}(\mathsf{S}')$, where $B(i) = B(\mathsf{A}, n)(i)$ and $\mathsf{S}' = \langle X_1, \dots, X_{l_0} \rangle$, contains the words W_k, $k = 1, 2, \dots$.

Consider the subgroup $\mathsf{W}' = \langle W_k \mid k = 1, 2, \dots \rangle$ of $B(i)$. Assume that W' has no elements of infinite order. Then W' is an infinite periodic subgroup of $B(i)$ following from the infiniteness of the subset $\{W_k \mid k = 1, 2, \dots\} \subseteq B(\mathsf{A}, n)$. On the other hand, $B(i)$ is hyperbolic by Lemma 21.1 and it is known (see [6, 4] or [12, Lemma 17 and the remark following Lemma 18]) that a hyperbolic group contains no infinite periodic subgroups. This contradiction proves that W' has an element, say V, of infinite order. Consider the centralizer $\mathsf{C}' = C_{B(i)}(V)$. Clearly, $\mathsf{S}' \subseteq \mathsf{C}'$. By [4, Theorem 38], the index $[\mathsf{C}' : \langle V \rangle]$ is finite. Hence, the quotient $\mathsf{C}'/(\mathsf{C}')^n$ is also finite. Note the subgroup S of $B(i)/B(i)^n = B(\mathsf{A}, n)$ embeds in $\mathsf{C}'/(\mathsf{C}')^n$ and so is also finite. A contradiction completes the proof of Lemma 12. ∎

LEMMA 13. *Suppose G is a finite subgroup of $B(\mathsf{A}, n)$ and $X_1, \dots, X_k \in \mathsf{G}^2$. Then*

$$h(\langle X_1, \dots, X_k \rangle) = \max\big(h(X_1), \dots, h(X_k)\big).$$

Proof. This is deduced from Lemmas 15.7 and 15.8 by induction on $h(\mathsf{G})$ in the same manner as Lemma 16.1. ∎

LEMMA 14. *Suppose G is a finite 2-subgroup of $B(\mathsf{A}, n)$ and $X \in \mathsf{G}^2$ has positive height j. Then there is a homomorphism*

$$\sigma_j \colon \mathsf{G} \to D(2n_2)$$

such that $\sigma_j(Y) \neq 1$ for every $Y \in \mathsf{G}^2$ with $h(Y) = j$ and $\sigma_j(Z) = 1$ for every $Z \in \mathsf{G}^2$ with $h(Z) < j$.

Proof. First suppose $h(\mathsf{G}) = j$. By Lemma 15.2, we may assume that G is a subgroup of $\mathsf{K}(A_j)$. It follows from the definitions of subgroup $\mathsf{F}(A_j)$ and $\mathsf{F}(A_j)$-involutions that every $Z \in \mathsf{G}^2$ with $h(Z) < j$ belongs to $\mathsf{F}(A_j)$. Hence, applying the homomorphism κ_1 of Lemmas 15.7 and 15.8 to $\mathsf{K}(A_j)$, we see that the restriction of κ_1 on G is a desired σ_j.

Suppose $h(\mathsf{G}) = i > j$. By Lemma 15.2, we may assume that $\mathsf{G} \subseteq \mathsf{K}(A_i)$. As before, by definitions of $\mathsf{F}(A_i)$ and $\mathsf{F}(A_i)$-involutions, for every $Z \in \mathsf{G}^2$ with $h(Z) < i$ we have $Z \in \mathsf{F}(A_i)$. Applying the homomorphism κ_2 to $\mathsf{K}(A_i)$ whose restriction on $\mathsf{F}(A_i)$ is identical, we get that

$$\kappa_2(\mathsf{G}) \subseteq \kappa_2(\mathsf{K}(A_i)),$$

the subgroup $\kappa_2(\mathsf{G})$ has height $i' < i$, and $\kappa_2(Z) = Z$ for every $Z \in \mathsf{G}^2$ with $h(Z) \leq j$. Since $\kappa_2(\mathsf{G}^2) \subseteq \kappa_2(\mathsf{G})^2$, it remains to refer to the induction hypothesis. ∎

2. PROOFS OF THEOREMS AND COROLLARIES

Proof of Theorem 1. (a) This immediately follows from Lemmas 5, 7, and 6.

(b) First suppose that G is a locally finite 2-group. If G is finite, then the infiniteness of its centralizer $C_{B(\mathsf{A}, n)}(\mathsf{G})$ follows from part (a). Assume that G is infinite. Then $\mathsf{G} = \bigcup_{t=1}^{\infty} \mathsf{G}_t$, where G_t are finite 2-subgroups of $B(\mathsf{A}, n)$, $t = 1, 2, \dots$. By [11, Theorem A(c)], every G_t is isomorphic to a subgroup of $D(2n_2)^{l_t}$ for some l_t. Hence, G belongs to the quasivariety $\operatorname{qvar} D(2n_2)$ of groups generated by $D(2n_2)$. By Mal'cev theorem (see [15, Corollary 9, Chap. V.11]), a group of $\operatorname{qvar} D(2n_2)$ embeds in a Cartesian product of copies of $D(2n_2)$. (Another, independent from this result from model theory, proof of such an embedding for G will be given in the proof of Theorem 2.) This implies that G is residually finite. Suppose the center $C(\mathsf{G})$ of G is finite and $Z_1, \dots, Z_k$ are all distinct nontrivial elements in $C(\mathsf{G})$. Since G is residually finite, there are normal subgroups $N_1, \dots, N_k$ in G of finite index such that $Z_1 \notin N_1, \dots, Z_k \notin N_k$. Then the intersection $N = N_1 \cap \cdots \cap N_k$ trivially intersects with $C(G)$ and has finite index in G, in particular, $N \neq \{1\}$. However, G is nilpotent [for $D(2n_2)$ is a finite 2-group] and so every nontrivial normal subgroup of G must have nontrivial intersection with $C(G)$ (e.g., see [21, Theorem 5.41]). A contradiction proves the following:

CLAIM 1. *The center $C(G)$ of an infinite locally finite 2-subgroup G of $B(A, n)$ is infinite.*

Now it becomes obvious that the centralizer $C_{B(A,n)}(G)$ is infinite as required.

Conversely, suppose that the centralizer $C_{B(A,n)}(G)$ is infinite. Then it follows from Lemma 12 that G is locally finite. That G is a 2-group is obvious from Lemma 11(a).

(c) By parts (b) and (a), the subgroup $C_{B(A,n)}(C_{B(A,n)}(G))$ is a finite group. Since

$$C_{B(A,n)}\Big(C_{B(A,n)}\big(C_{B(A,n)}(G)\big)\Big)$$

contains $C_{B(A,n)}(G)$ and so is not locally finite, we have from Lemma 11(a) that

$$C_{B(A,n)}\big(C_{B(A,n)}(G)\big)$$

is a finite 2-group.

Assume $G \neq C_{B(A,n)}(C_{B(A,n)}(G))$. Since $G \subset C_{B(A,n)}(C_{B(A,n)}(G))$, there is a finite subgroup K in $C_{B(A,n)}(G)$ that contains G as a subgroup of index 2. By definition, K centralizes $C_{B(A,n)}(G)$. However, the existence of such K obviously contradicts Lemmas 5 and 8. ∎

Proof of Corollary 1. By [11, Theorem A(c)], every finite subgroup G of $B(A, n)$ is isomorphic to a subgroup of $D(2n_1) \times D(2n_2)^l$ for some l.

Note that, by Lemma 6, the subgroup $\langle a_1^{n/2}, a_2^{n/2} \rangle$ of $B(A, n)$ is isomorphic to $D(2n)$ and so $B(A, n)$ contains subgroups isomorphic to $D(2n_1)$ and $D(2n_2)$. Now the converse obviously follows from Theorem 1(a). ∎

Proof of Corollary 2. This easily follows from the Mal'cev theorem cited in the proof of Theorem 1(b) and Lemma 11. ∎

Proof of Corollary 3. A theorem of Gorchakov (see [5, Theorem 2.6] or [22, Theorem 2.13]) claims that an *FC*-subgroup of a Cartesian power of a finite group F embeds in a direct product of finite groups of bounded exponents. However, it seems to have gone unnoticed that Gorchakov's proof (which is rather involved) actually yields more: Any *FC*-subgroup of a Cartesian power of a finite group F embeds in a direct power of F. To see this, it suffices just to follow proofs in [22] (or [5]): In [22, Lemmas 2.17 and 2.18] the groups $G(J)$ and G are embeddable in a direct power of F because they embed in the direct product of finitely many copies of F and an Abelian group whose exponent divides that of the center of F. In the definition of a reducing subset J, require that $G(J)$ is embeddable in a Cartesian product of finite subgroups each of which is isomorphic to a proper subgroup of F. Then the proof of the central Lemma 2.19 [22] is retained.

Putting together this version of Gorchakov's theorem and the Mal'cev theorem cited in the proof of Theorem 1(b), we have the following rather interesting proposition:

PROPOSITION. *Any FC-group of the quasivariety* qvar F, *where F is a finite group, embeds in a direct power of F.*

Now Corollary 3 becomes obvious from Theorem 1(a). ∎

Proof of Theorem 2. Suppose L is an infinite locally finite subgroup of $B(\mathsf{A}, n)$. It follows from Lemma 11(a) that L contains no nontrivial elements of odd order, that is, L is a 2-group.

Recall by D we denote the Cartesian product of groups D_i, $i = 1, 2, \ldots,$ isomorphic to $D(2n_2)$. Denote the projection $\mathsf{D} \to D_i$ by π_i. Recall if $x \in \mathsf{D}$, then $\operatorname{supp} x = \{i \mid \pi_i(x) \neq 1\}$.

CLAIM 2. *There is a monomorphism* $\Omega\colon \mathsf{L} \to \mathsf{D}$ *such that* supp *z is finite for every $z \in \mathsf{L}^2$.*

Proof. Let L_j denote the subgroup of L generated by all elements of L^2 whose heights do not exceed j. It follows from Lemma 13 that $h(\mathsf{L}_j) \le j$. Hence, by Lemma 15.2, L_j is finite. Let $j_1, j_2, \ldots$ be the subsequence of $1, 2, \ldots$ of all js with $h(\mathsf{L}_j) = j$. Clearly,

$$\mathsf{L}^2 = \bigcup_{t=1}^{\infty} \mathsf{L}_{j_t}.$$

Consider some finite subgroups G_t, $t = 1, 2, \ldots,$ of L so that $\mathsf{L}_{j_t} \subset \mathsf{G}_t^2$, $\mathsf{G}_t \subseteq \mathsf{G}_{t+1}$, and $\bigcup_{t=1}^{\infty} \mathsf{G}_t = \mathsf{L}$. Fixing $j_t \ge 1$, we have from Lemma 14 that the set $\Sigma_s(j_t)$ of all homomorphisms $\sigma\colon \mathsf{G}_s \to D(2n_2)$, where $s \ge j_t$, such that $\sigma(X) \neq 1$ for every $X \in \mathsf{G}_s^2$ with $h(X) = j_t$ and $\sigma(Y) = 1$ for every $Y \in \mathsf{G}_s^2$ with $h(Y) < j_t$ is nonempty. Since the restriction of any $\sigma \in \Sigma_{s+1}(j_t)$ on G_s belongs to $\Sigma_s(j_t)$, we have, by standard "compactness" argument, the existence of a sequence $\sigma_t, \sigma_{t+1}, \ldots,$ where $\sigma_t \in \Sigma_t(j_t)$, $\sigma_{t+1} \in \Sigma_{t+1}(j_t), \ldots,$ such that the restriction of σ_{t+k+1} on G_{t+k} is σ_{t+k}. It follows from definitions that such a sequence defines a homomorphism $\sigma_{j_t}\colon \mathsf{L} \to D(2n_2)$ such that $\sigma_{j_t}(X) \neq 1$ for every $X \in \mathsf{L}^2$ with $h(X) = j_t$ and $\sigma_{j_t}(Y) = 1$ for every $Y \in \mathsf{L}^2$ with $h(Y) < j_t$. It is clear that for given $Z \in \mathsf{L}^2$ almost all $\sigma_{j_t}(Z)$, $t = 1, 2, \dots,$ are 1s.

Suppose $U \notin \mathsf{L}^2$. Then the existence of a homomorphism $\varepsilon_U\colon \mathsf{L} \to D(2n_2)$ such that $\varepsilon_U(U) \neq 1$ and $\varepsilon_U(\mathsf{L}^2) = 1$ is obvious.

Now we see that the homomorphisms σ_{j_t}, $t = 1, 2, \dots,$ and ε_U, $U \notin \mathsf{L}^2$, define a system of residualizing homomorphisms $\mathsf{L} \to D(2n_2)$ so that for every $V \in \mathsf{L}$ almost all $\sigma_{j_t}(V^2)$, $\varepsilon_U(V^2)$ are 1s. This proves Claim 2. ▮

Recall

$$D_i = \langle b_i, c_i \parallel b_i^2 = c_i^{n_2} = 1, b_i c_i b_i = c_i^{-1} \rangle,$$

$B_i = \langle b_i \rangle$, and $C_i = \langle c_i \rangle$, where $i = 1, 2, \dots,$ B and D are the Cartesian products of groups B_i and D_i, respectively, C is the direct product of groups C_i, the groups B and C are naturally embedded in D, and $\mathsf{E} = \mathsf{B}\mathsf{C}$.

In view of Claim 2, it suffices to prove that a countable subgroup H of D such that $\operatorname{supp} x$ is finite for every $x \in H^2$ is isomorphic to a subgroup of E.

Without loss of generality, we may assume that $\mathsf{C} \subseteq H$. Let $h_1, h_2, \dots$ be a system of generators for H modulo C. Since $\operatorname{supp} h_t^2$ is finite, every h_t has only finitely many projections $\pi_i(h_t)$ such that $\pi_i(h_t) \in C_i$ and $\pi_i(h_t)$ is not central in D_i. Hence, multiplying h_t by an element of C, we can eliminate all such projections of h_t. In particular, we can assume $h_t^2 = 1$.

Suppose $[h_1, h_t] \neq 1$. It follows from $[h_1, h_t] \in H^2$ that $\operatorname{supp}[h_1, h_t]$ is finite. Note that if $x, y \in D_i$ are such that $x^2 = y^2 = 1$ and $[x, y] \neq 1$, then, by $x \in yC_i$, there is $z \in C_i$ so that $yz = x$. This remark enables us to multiply h_t, $t \geq 2$, by an element of C so that for the resulting element h'_t we get $(h'_t)^2 = 1$ and $[h_1, h'_t] = 1$.

Rename $h_1, h'_2, h'_3, \dots$ back to $h_1, h_2, h_3, \dots$. Suppose $[h_2, h_t] \neq 1$. Then for every $i \in \operatorname{supp}[h_2, h_t]$, the element $\pi_i(h_1)$ is central in D_i. Hence, as before, multiplying $h_3, h_4, \dots$ by elements of C, we can achieve that $h_t^2 = [h_1, h_t] = [h_2, h_t] = 1$ for all t. If we keep on doing this, we will have a generating set $h_1, h_2, \dots$ for H modulo C so that $h_t^2 = [h_t, h_s] = 1$ for all t, s.

Since the intersection $\langle h_1, h_2, \dots \rangle \cap \mathsf{C}$ is a direct factor of $\langle h_1, h_2, \dots \rangle$, we may change $h_1, h_2, \dots$ if necessary and assume that

$$\langle h_1, h_2, \dots \rangle \cap \mathsf{C} = \{1\}. \tag{27}$$

Let us further change $h_1, h_2, \dots$ so that $\mathbb{N} = \{1, 2, \dots\} = T_0 \cup T_1$,

$$\langle h_t \mid t \in T_0 \rangle = \langle h_1, h_2, \dots \rangle \cap \hat{\mathsf{C}},$$

where $\hat{C}$ is the Cartesian product of groups C_i and $\langle h_1, h_2, \dots \rangle$ is the direct product of $\langle h_t \mid t \in T_0 \rangle$ and $\langle h_t \mid t \in T_1 \rangle$. It is clear that no nontrivial element of $\langle h_t \mid t \in T_1 \rangle$ centralizes C and that every element of $\langle h_t \mid t \in T_0 \rangle$ centralizes C.

Since $\langle h_t \mid t \in T_0 \rangle$ is a direct factor of H, it follows from (27) that we can embed H in $H' \subseteq D'$ so that the index set for D' is $T' = \mathbb{N} \cup T_0'$ ($\mathbb{N}$ and T_0' are disjoint), the subgroup H' of D' contains C', H' is generated modulo C' by h_t, $t \in T_1$, and no nontrivial element of $\langle h_t \mid t \in T_1 \rangle$ centralizes C'.

Finally, rename everything back (note T_0 will be empty) and consider the elements $\bar{h}_t$ such that $\pi_i(\bar{h}_t) = b_i(h_t)$ if $\pi_i(h_t)$ is not central in D_i and $\pi_i(\bar{h}_t) = 1$ otherwise. It follows from the preceding properties of $h_1, h_2, \dots$ that the subgroup $\bar{H}$ of D generated by $\bar{h}_1, \bar{h}_2, \dots$ and C is isomorphic to H under the map which is identical on C and sends $\bar{h}_t$ to h_t. Clearly, $\bar{H}$ is a subgroup of E, as required.

Let us show the converse, that is, every countable subgroup H of $E = BC$ is embeddable in $B(A, n)$. Without loss of generality, we may assume that $C \subseteq H$, H is generated modulo C by $h_1, h_2, \dots \in B$, and $b_t \in B$ for all t.

It is not difficult to see that the standard process of bringing a matrix over a field to a row echelon form also applies to ω-infinite matrices. (To see this it suffices to mark the rows of the original matrix A that will contain leading elements in the process and note that each row of A will be involved in no more than $N + 1$ elementary transformations, where N is the number of the marked rows located above that row.) Hence, we may also assume that the elements $h_1, h_2, \dots$ are in row echelon form, that is, the first element, j_t, in $\operatorname{supp} h_t$ is t (recall all $b_t \in H$).

By H_i denote the subgroup of H generated by $C_1, \dots, C_i$ (which are identified with the factors of $C \subset D$) and $h_1, \dots, h_i$. By induction on $i \geq 1$, we will construct monomorphisms

$$\Psi_{i+1}\colon H_{i+1} \to B(A, n)$$

such that the restriction $\Psi_{i+1}|_{H_i}$ equals Ψ_i ($H_0 = \{1\}$). Since H_1 is the dihedral group D_1, the existence of Ψ_1 is obvious.

Assume that $\Psi_i\colon H_i \to B(A, n)$, $i \geq 1$, with $\Psi_i|_{H_{i-1}} = \Psi_{i-1}$, already constructed. By Lemmas 5 and 7, there is a word $X \in C_{B(A,n)}(\Psi_i(H_i))$ of order 2 which is not in $\Psi_i(H_i)$. Put

$$\Psi_{i+1}(h_{i+1}) = X.$$

Consider a subgroup G in $K = \langle \Psi_i(H_i), X \rangle$ of index 2 such that $\Psi_i(h_t) \in G$ if and only if h_t commutes with c_{i+1} [i.e., $\pi_{i+1}(h_t) = 1$].

Clearly, $X \notin G$. Hence it follows from Lemmas 5, 8, and 6 that $C_{B(A,n)}(G)$ contains a word Y of order 2 such that XY has order n. To define Ψ_{i+1} put

$$\Psi_{i+1}(c_{i+1}) = (XY)^{n_1} \quad \text{and} \quad \Psi_{i+1}|_{H_i} = \Psi_i.$$

It follows from definitions that for every $h_t \in H_{i+1}$ one has

$$\Psi_{i+1}(h_t)\Psi_{i+1}(c_{i+1})\Psi_{i+1}(h_t)^{-1} = \Psi_{i+1}(c_{i+1})^{\varepsilon(t)},$$

where $\varepsilon(t) = 1$ if $\pi_{i+1}(h_t) = 1$ for $\Psi_{i+1}(h_t) \in G$ and $\varepsilon(t) = -1$ if $\pi_{i+1}(h_t) = b_t$ for $\Psi_{i+1}(h_t) \notin G$.

Therefore, the definition of Ψ_{i+1} is correct.

To complete the proof it remains to note that $H = \bigcup_{i=1}^{\infty} H_i$ and so the map $\Psi: H \to B(A,n)$ given by $\Psi|_{H_i} = \Psi_i$ is an embedding. Theorem 2 is proven. ∎

Proof of Corollaries 4 *and* 5. They are immediate from Theorem 2. ∎

Proof of Theorem 3. (a) It suffices to show that the intersection of two distinct maximal locally finite subgroups K and L of $B(A,n)$ is finite. Arguing on the contrary, assume that K and L have an infinite intersection. Then, by Lemma 11, K and L are 2-groups. First let us prove:

CLAIM 3. *The intersection $U = K \cap L$ is an FC-group.*

Proof. To prove this we will make use of some of quasi-identities that hold in $D(2n_2)$ and the fact that any locally finite 2-subgroup of $B(A,n)$ belongs to qvar $D(2n_2)$ (see Corollary 2).

Note the proposition "Every element of a group G that is a product of squares and has order ≤ 2 is in the center of G" can be written as an infinite system of quasi-identities that hold in $D(2n_2)$ and, therefore, holds in U, K, L.

Suppose $Z \in U$ has an infinite conjugacy class. Then the subgroup $V = \langle [Z,T] \mid T \in U \rangle$ is infinite. Since $V \subseteq U^2$, the subgroup U^2 is also infinite. By Claim 1, the center $C(U^2)$ of U^2 is infinite. Then, by Prüffer's theorem (e.g., see [21, Corollary 10.37]), the maximal subgroup $T(U^2)$ of $C(U^2)$ of exponent 2 is also infinite. By the preceding proposition applied to K and L, we have that $T(U^2)$ centralizes both K and L. However, then the subgroup $\langle K, L \rangle$ has an infinite centralizer in $B(A,n)$ and, by Lemma 12, is locally finite. This contradiction to the maximality of K, L proves Claim 3. ∎

Consider a central series

$$K = K_0 \supset K_1 \supset \cdots \supset K_p = \{1\}$$

in K whose factors have exponent 2 and K_{p-1} is infinite. Such a series exists by Claim 1. Let q be the minimal index so that $K_q \subseteq L$. Without loss of generality we may assume that there is no maximal locally finite subgroup L' in $B(A, n)$ such that $L' \neq K$ and $K_{q-1} \subseteq L'$ (otherwise, we would consider L' instead of L; note $K \cap L'$ would still be infinite).

Pick $X \in K_{q-1}$ so that $X \notin U$. Also, pick $Y \in L$ so that Y normalizes U, $Y^2 \in U$, and $Y \notin U$ (the existence of such Y follows from nilpotency of L; see [21, Theorem 5.41]).

Assume that the set $\{ZXZ^{-1} \mid Z \in U\}$ is infinite. Note that $\langle X, Y \rangle$ generate a dihedral subgroup modulo U. Hence the subgroup $\langle X, Y, U \rangle$ is locally finite. Consider a maximal locally finite subgroup P of $B(A, n)$ that contains $\langle X, Y, U \rangle$. Then $P \neq K$ for $Y \in P$ and the intersection $K \cap P$, containing $\langle X, U \rangle$, is not an FC-group. A contradiction to Claim 3 proves that the set $\{ZXZ^{-1} \mid Z \in U\}$ is finite.

Quite analogously, the set $\{ZYZ^{-1} \mid Z \in U\}$ is also finite. Consequently, the centralizer $C_U(X)$ of any $X \in K_{q-1} \setminus U$ in U has finite index in U and the index of $C_U(Y)$ in U is also finite.

Let $X_1, \dots, X_l \in K_{q-1}$. Then the intersection

$$U_0 = C_U(X_1) \cap \cdots \cap C_U(X_l) \cap C_U(Y)$$

also has finite index in U. Therefore, the intersection $U_0 \cap C(U)$ is infinite by Claim 1. This implies that the subgroup $\langle X_1, \dots, X_l, Y, U \rangle$ has an infinite centralizer in $B(A, n)$ and hence, by Lemma 12, is locally finite. Consequently, the subgroup $\langle K_{q-1}, Y \rangle$ is locally finite. Consider a maximal locally finite subgroup L_1 of $B(A, n)$ that contains $\langle K_{q-1}, Y \rangle$. It is clear that $L_1 \neq K$ for $Y \notin K$ and that $K \cap L_1$ contains K_{q-1}. This contradiction to the choice of q completes the proof of part (a).

(b) Let G be a finite 2-subgroup of $B(A, n)$, $\mathcal{G}$ be a group isomorphic to G, and $\Phi_0 \colon \mathcal{G} \to G$ an isomorphism. Suppose $M = M_1 \times M_2 \times \cdots$ is the direct product of finite nontrivial groups $M_1, M_2, \dots$ and M_k is isomorphic to a subgroup of $D_{1k} \times \cdots \times D_{l_k k}$, where $D_{1k}, \dots, D_{l_k k}$ are dihedral groups of order $2n_2$. First let us show that there is a monomorphism

$$\Phi \colon \mathcal{G} \times M \to B(A, n)$$

such that the restriction of Φ on $\mathcal{G}$ is Φ_0 and $\Phi(\mathcal{G} \times M)$ is a maximal locally finite subgroup of $B(A, n)$. Without loss of generality, we may assume that $|M_k| > 2$, $k = 1, 2, \dots$. Identifying M_k with the corresponding factor of M, denote

$$P_i = \mathcal{G} \times M_1 \times \cdots \times M_i \subset \mathcal{G} \times M,$$

where $P_0 = \mathcal{G}$.

To embed $G \times M$ in $B(\mathcal{A}, n)$ we will construct monomorphisms $\Phi_i \colon P_i \to B(\mathcal{A}, n)$, $i \geq 0$, so that the restriction $\Phi_i|_{P_{i-1}}$ of Φ_i on P_{i-1} is Φ_{i-1}, where $i \geq 1$.

Let $\mathcal{B} = (B_1, B_2, \dots)$ be a list of $\mathcal{A}$-words that represent all nontrivial elements of $B(\mathcal{A}, n)$. Recall $\Phi_0(G) = \Phi_0(P_0) = \mathcal{G}$.

Proceeding by induction on $i \geq 0$, suppose a monomorphism

$$\Phi_i \colon P_i \to B(\mathcal{A}, n)$$

so that $\Phi_i|_{P_{i-1}} = \Phi_{i-1}$, where $i \geq 1$, is already constructed for $i \geq 0$. Let B_{k_i} be the first word in $\mathcal{B}$ such that $k_i > k_{i-1}$ ($k_0 = 0$ by definition), $B_{k_i} \notin \Phi_i(P_i)$, and the subgroup

$$\mathcal{K}_i = \langle B_{k_i}, \Phi_i(P_i) \rangle$$

is a finite 2-subgroup of $B(\mathcal{A}, n)$. By Theorem 1(a), such B_{k_i} does exist. By $\mathcal{G}_i$ denote a subgroup of index 2 in $\mathcal{K}_i$ that contains $\Phi_i(P_i)$. It follows from Lemmas 5, 8, 6, and 10 that $C_{B(\mathcal{A}, n)}(\mathcal{G}_i)$ contains a subgroup $\mathcal{D}_{i+1}$ isomorphic to $D_{1 i+1} \times \cdots \times D_{l_{i+1} i+1}$ such that

$$\langle \Phi_i(P_i), \mathcal{D}_{i+1} \rangle = \Phi_i(P_i) \times \mathcal{D}_{i+1}$$

and if $X, Y \in \mathcal{D}_{i+1}$ are distinct and nontrivial elements in $B(\mathcal{A}, n)$, then $\langle X, Y, B_{k_i} \rangle$ is an infinite subgroup of $B(\mathcal{A}, n)$. Now we are able to define Φ_{i+1}: Put $\Phi_{i+1}|_{P_i} = \Phi_i$ and $\Phi_{i+1}|_{M_{i+1}}$ is an isomorphism between M_{i+1} and a subgroup of $\mathcal{D}_{i+1}$.

The inductive construction of monomorphisms Φ_i for all $i = 1, 2, \dots$ is now complete and hence a monomorphism $\Phi \colon G \times M \to B(\mathcal{A}, n)$ given by $\Phi|_{P_i} = \Phi_i$ is also constructed.

Let us prove that $\Phi(G \times M)$ is a maximal locally finite subgroup of $B(\mathcal{A}, n)$. Arguing on the contrary, assume the existence of a word $B \notin \Phi(G \times M)$ such that the subgroup $\langle B, \Phi(G \times M) \rangle$ of $B(\mathcal{A}, n)$ is locally finite. Then, by Theorem 2, $\langle B, \Phi(P_j) \rangle$ are finite 2-subgroups for all $j \geq 1$, whence $B = B_{k_i}$ for some $i \geq 1$. However, in view of $|M_{i+1}| > 2$, we have that $\langle B, \Phi(M_{i+1}) \rangle$ is infinite. A contradiction completes the proof of the maximality of $\Phi(G \times M)$.

It remains to show that there are continuously many nonisomorphic groups of the form $G \times M_1 \times M_2 \times \cdots$. It is known due to Hirshon [9] that an isomorphism of the direct products $H \times A$ and $H \times B$, where H is finite and A, B are arbitrary groups, implies that A and B are isomorphic. Hence it suffices to prove that there are continuously many nonisomorphic groups of the form $M_1 \times M_2 \times \cdots$.

Let

$$G_j = \left\langle a_{1j}, \dots, a_{jj}, b_j \,\|\, b_j^2 = a_{kj}^4 = [a_{kj}, a_{lj}] = 1, \right.$$
$$\left. b_j a_{kj} b_j^{-1} = a_{kj}^{-1}, 1 \le k, l \le j \right\rangle,$$

where $j > 1$. It is easy to see that G_j is isomorphic to a subgroup of $D(2n_2)^j$.

Suppose $J = (j_1, j_2, \dots)$ is a strictly increasing infinite sequence of positive integers greater than 1 and $G(J) = G_{j_1} \times G_{j_2} \times \cdots$ is the direct product of groups $G_{j_1}, G_{j_2}, \dots$. By π_k denote the projection $G(J) \to G_{j_k}$.

Suppose $g \in G(J)$ and there is a k such that $\pi_k(g) \notin N_{j_k}$, where $N_{j_k} = \langle a_{1j_k}, \dots, a_{j_k j_k} \rangle$. Then the elments $a_{1j_k}^{\varepsilon_1} \cdots a_{j_k j_k}^{\varepsilon_k}$, where $\varepsilon_1, \dots, \varepsilon_k \in \{0, 1\}$, belong to distinct cosets of $G(J)$ by $C_{G(J)}(g)$. Hence, $[G(J){:}C_{G(J)}(g)] \ge 2^{j_k}$. Quite analogously, if $\pi_{k_t}(g) \notin N_{j_{k_t}}$ for all k_t, $t = 1, \dots, s$, then

$$\left[G(J){:}\, C_{G(J)}(g)\right] \ge 2^{j_{k_1} + \cdots + j_{k_s}}. \tag{28}$$

Assume $J \ne J'$, $J' = (j'_1, j'_2, \dots)$, and $\Psi{:}\, G(J) \to G(J')$ is an isomorphism. Note

$$\left[G(J){:}\, C_{G(J)}\left(b_{j_k}\right)\right] = 2^{j_k}. \tag{29}$$

By the Hirshon result cited previously, we can assume $j_1 < j'_1$. Consider $\Psi(b_1)$. It follows from (29) and (28) that $\pi_l(\Psi(b_1)) \in N_{j'_l}$ for every l. Since $b_1^2 = 1$, we have that $\Psi(b_1)$ is a central element of $G(J')$. A contradiction proves that $G(J)$ and $G(J'_v)$ are not isomorphic when $J \ne J'$. Since the set of such Js is continuous, part (b) is proven.

(c) This is shown in Lemma 11. ∎

Proof of Theorem 4. Let us begin with proving:

CLAIM 4. *Suppose L is an infinite locally finite subgroup of $B(\mathsf{A}, n)$, G is a finite subgroup of L, and B is a word, $B \notin \mathsf{L}$. Then there is an $X \in C_{B(\mathsf{A},n)}(\mathsf{G})$ such that the subgroup $\langle XBX^{-1}, \mathsf{L} \rangle$ is not locally finite.*

Proof. Arguing on the contrary, assume that $\langle XBX^{-1}, \mathsf{L} \rangle$ is locally finite for every $X \in C_{B(\mathsf{A},n)}(\mathsf{G})$.

By Theorem 3(a), there is a unique maximal locally finite subgroup $M(\mathsf{L})$ that contains L and, therefore, contains all XBX^{-1}, where $X \in C_{B(\mathsf{A},n)}(\mathsf{G})$.

Consider the subgroup

$$\mathsf{S} = \left\langle \mathsf{G}, XBX^{-1} \,\middle|\, X \in C_{B(\mathsf{A},n)}(\mathsf{G}) \right\rangle \subseteq M(\mathsf{L}).$$

Note $C_{B(A,n)}(G)$ normalizes S. Hence, picking any finite subgroup F in $C_{B(A,n)}(G)$ of order ≥ 4, we have that $\langle S, F \rangle$ is a locally finite subgroup. On the other hand, we can find a finite subgroup $K \subseteq S$ that contains G as a subgroup of index 2. Let $K = \langle D, G \rangle$, where D is a word. Then it is immediate from Lemmas 5, 8, and 10 that one can pick F so that $\langle D, F \rangle$ is not locally finite. Since $D \in S$ and $\langle S, F \rangle$ is locally finite, we have a contradiction that proves Claim 4. ∎

Let $L = \bigcup_{k=1}^{\infty} L_k$ be an infinite locally finite subgroup of $B(A, n)$, L_k, $k = 1, 2, \ldots,$ be finite subgroups so that $L_k \subseteq L_{k+1}$, and $B = \{B_1, B_2, \ldots\}$ be a list of all elements of $B(A, n)$.

Proceedings by induction on $i \geq 1$, we will construct a sequence of words X_i and a sequence of indexes k_i, $i = 1, 2, \ldots,$ to obtain a maximal locally finite subgroup M of $B(A, n)$ that is locally conjugate to L.

Rename $L = L(0)$, $L_k = L_k(0)$ and, to unify notation, put that $k_0 = 0$, $X_0 = 1$, and $L_{k_0}(0) = \{1\}$.

Suppose $i \geq 1$ and assume that the words $X_0, \ldots, X_{i-1}$, indexes $k_0, \ldots, k_{i-1}$, and subgroups $L(i-1)$, $L_k(i-1)$, where

$$L(i-1) = \bigcup_{k=1}^{\infty} L_k(i-1), \qquad L_k(i-1) \subseteq L_{k+1}(i-1)$$

are already defined.

First assume $B_i \in L(i-1) = \bigcup_{k=1}^{\infty} L_k(i-1)$. Then there is a t such that $t > k_{i-1}$ and $B_i \in L_t(i-1)$. In this case, define $k_i = t$ and $X_i = 1$.

Assume $B_i \notin L(i-1)$. It follows from Claim 4 that there is a word

$$Y \in C_{B(A,n)}\big(L_{k_{i-1}}(i-1)\big)$$

[recall $L_{k_0}(0) = \{1\}$] such that the subgroup $\langle YB_iY^{-1}, L(i-1)\rangle$ is not locally finite. Then there is an $s > k_{i-1}$ so that the subgroup $\langle YB_iY^{-1}, L_s(i-1)\rangle$ is not locally finite either. Define $k_i = s$ and $X_i = Y^{-1}$. Note that, by definitions, the subgroup

$$\langle B_i, X_i L_{k_i}(i-1) X_i^{-1} \rangle \tag{30}$$

is not locally finite [provided $B_i \notin L(i-1)$].

Thus, in either case, X_i and k_i are defined. At last, put

$$L(i) = X_i L(i-1) X_i^{-1}, \qquad L_k(i) = X_i L_k(i-1) X_i^{-1}.$$

Clearly,

$$L(i) = \bigcup_{k=1}^{\infty} L_k(i-1), \qquad L_k(i) \subseteq L_{k+1}(i).$$

The inductive definition is now complete and so we have sequences of words $X_1, X_2, \ldots$ and indexes $k_1 < k_2 < \cdots$ such that

$$L(i) = X_i X_{i-1} \cdots X_1 L(0) X_1^{-1} \cdots X_{i-1}^{-1} X_i^{-1},$$

$$L_k(i) = X_i X_{i-1} \cdots X_1 L_k(0) X_1^{-1} \cdots X_{i-1}^{-1} X_i^{-1},$$

and $X_i \in C_{B(A,n)}(L_{k_{i-1}}(i-1))$.

Let us show that the subgroup

$$M = \bigcup_{i=1}^{\infty} L_{k_i}(i)$$

is locally conjugate to L in $B(A,n)$ and M is a maximal locally finite subgroup of $B(A,n)$.

Note that the map Φ_i: $L_{k_i}(0) = L_{k_i} \to L_{k_i}(i)$ given by

$$\Phi_i(Z) = X_i X_{i-1} \cdots X_1 Z X_1^{-1} \cdots X_{i-1}^{-1} X_i^{-1}$$

is an isomorphism and the restriction $\Phi_i|_{L_{k_{i-1}}}$ of Φ_i on $L_{k_{i-1}}$ is Φ_{i-1} following from $X_i \in C_{B(A,n)}(L_{k_{i-1}}(i-1))$. Hence the map Φ: $L \to M$ given by $\Phi|_{L_{k_i}} = \Phi_i$ is an isomorphism that makes L be locally conjugate to M.

Suppose M is not maximal. Then there is a $B_i \in B$ such that $B_i \notin M$ and the subgroup $\langle B_i, M \rangle$ is still locally finite. Since $L_{k_i}(i) \subset M$, the subgroup $\langle B_i, L_{k_i}(i) \rangle$ is also locally finite. Hence it follows from $L_{k_i}(i) = X_i L_{k_{i-1}}(i-1) X_i^{-1}$ and defintions of X_i and k_i [see also (30)] that $B_i \in L_{k_i}(i) \subset M$. A contradiction completes the proof of Theorem 4. ∎

Proof of Corollary 6. This follows from Theorems 2 and 4. ∎

ACKNOWLEDGMENTS

The first author was supported in part by an Alfred P. Sloan Research Fellowship, a Beckman Fellowship, and NSF grant DMS-95-01056. The second author was supported in part by Russian Fund of Fundamental Research grant 96-01-0420.

REFERENCES

1. S. I. Adian, ‘‘The Burnside Problem and Identities in Groups,’’ Nauka, Moscow, 1975; English transl.: Springer-Verlag, New York, 1979.
2. G. Baumslag, ‘‘Topics in Combinatorial Group Theory,’’ Birkhaüser, Boston, 1993.

3. W. Burnside, On unsettled question in the theory of discontinuous groups, *Quart. J. Pure Appl. Math.* **33** (1902), 230–238.
4. E. Ghys and P. de la Harpe (Eds.), ''Sur les Groupes Hyperboliques d'Aprés Mikhael Gromov,'' Birkhaüser, Boston, 1990.
5. Yu. M. Gorchakov, ''Groups with Finite Conjugacy Classes,'' Nauka, Moscow, 1978.
6. M. Gromov, Hyperbolic groups, *in* ''Essays in Group Theory,'' S. M. Gersten, Ed., Springer-Verlag, New York, 1987, pp. 75–263.
7. N. Gupta, On groups in which every element has finite order, *Amer. Math. Monthly* **96** (1989), 297–308.
8. D. Held, On abelian subgroups of an infinite 2-group, *Acta Sci. Math. (Szeged)* **27** (1966), 97–98.
9. R. Hirshon, On cancellation in groups, *Amer. Math. Monthly* **76** (1969), 1037–1039.
10. S. V. Ivanov, On the Burnside problem on periodic groups, *Bull. Amer. Math. Soc.* **27**, No. 2 (1992), 257–260.
11. S. V. Ivanov, The free Burnside groups of sufficiently large exponents, *Int. J. Algebra Comput.* **4**, No. 1–2 (1994), 1–308.
12. S. V. Ivanov and A. Yu. Ol'shanskii, Hyperbolic groups and their quotients of bounded exponents, *Trans. Amer. Math. Soc.* **348** (1996), 2091–2138.
13. A. G. Kurosh, ''The Theory of Groups,'' Vols. I, II, 2nd ed., Chelsea, New York, 1960.
14. W. Magnus, I. Karrass, and D. Solitar, ''Combinatorial Group Theory,'' Interscience, New York, 1966.
15. A. I. Mal'cev, ''Algebraic Systems,'' Nauka, Moscow, 1970; English transl.: Springer-Verlag, New York, 1973.
16. H. Neumann, ''Varieties of Groups,'' Springer-Verlag, New York, 1967.
17. P. S. Novikov and S. I. Adian, On infinite periodic groups, I, II, III, *Math. USSR-Izv.* **32** (1968), 212–244, 251–524, 709–731.
18. A. Yu. Ol'shanskii, On the Novikov–Adian theorem, *Math. USSR-Sb.* **118** (1982), 203–235.
19. A. Yu. Ol'shanskii, ''Geometry of Defining Relations in Groups,'' Nauka, Moscow, 1989; English transl.: *Math. Appl. (Soviet Ser.)* **70** (1991).
20. A. Yu. Ol'shanskii, Quasiidentities of finite groups, *Siberian Math. J.* **15** (1974), 1409–1413.
21. J. J. Rotman, ''An Introduction to the Theory of Groups,'' 4th ed., Springer-Verlag, New York, 1995.
22. M. J. Tomkinson, ''FC-Groups,'' Pitman, London, 1984.