Submitted exclusively to the London Mathematical Society
DOI: 10.1112/S0000000000000000

GROUPS WITH CONTEXT-FREE CO-WORD PROBLEM

DEREK F. HOLT, SARAH REES, CLAAS E. RÖVER, RICHARD M. THOMAS

Abstract

We study the class of co-context-free groups. We define a co-context-free group to be one whose co-word problem (the complement of its word problem) is context-free. This class is larger than the subclass of context-free groups, being closed under the taking of finite direct products, restricted standard wreath products with context-free top groups, and passing to finitely generated subgroups and finite index overgroups. But we do not know of other examples of co-context-free groups. We prove that the only examples amongst polycyclic groups or the Baumslag-Solitar groups are virtually abelian. We do this by proving that languages with certain purely arithmetical properties cannot be context-free; this result may be of independent interest.

1. *Introduction*

Let G be a group with finite generating set X. The *word problem* of G with respect to X, denoted $W(G,X)$, is the set of all words in $(X \cup X^{-1})^*$ which represent the identity element of G. The *co-word problem* of G with respect to X, denoted $coW(G,X)$, is the complement of $W(G,X)$ in $(X \cup X^{-1})^*$, that is, the set of words which represent nontrivial elements of G.

In this paper we study groups whose co-word problem with respect to some finite generating set (and therefore, it turns out, with respect to any finite generating set) is a context-free language. For brevity we call such groups co-context-free ($co\mathcal{CF}$) groups. Notice that, since the class of regular languages is closed under complementation, groups with regular co-word problem are precisely the ones with regular word problem, or equivalently all finite groups [**1**].

Groups with context-free word problem, known as context-free ($\mathcal{CF}$) groups, were classified by Muller and Schupp in [**11**, **12**, **3**]; these are precisely the virtually free

---

2000 *Mathematics Subject Classification* 20F10, 68Q45 (primary), 03D40 (secondary).

This research was supported by EPSRC.

groups and their word problem is in fact deterministic context-free. Since the complement of a deterministic context-free language is also deterministic context-free, such groups are examples of $co\mathcal{CF}$-groups, but there are many other $co\mathcal{CF}$-groups besides these, whose co-word problems are nondeterministic context-free. The most obvious examples are finitely generated abelian groups; since any such group is a direct product of virtually cyclic groups, its co-word problem can be recognised by a machine which first chooses one component of that direct product, and then projects onto that component and uses the deterministic pushdown automaton which solves the co-word problem for that component.

In Section 2, in addition to proving that the property of being $co\mathcal{CF}$ is independent of the choice of finite generating set, we prove the technical results that this class of groups is invariant under passing to finitely generated subgroups, moving to finite index overgroups and taking finite direct products. We deduce the results from properties of the class of context-free languages, which also hold for many other classes of languages, so that our results extend to groups with co-word problem in various other language classes (and all the results in this section are stated and proved in this generality). The fact that finitely generated abelian groups are $co\mathcal{CF}$ follows immediately from this section. It may be worthy of comment that many of our results (though not in general the direct product result) are also well known to be true for the classes of groups whose word problems (rather than co-word problems) lie in particular formal language classes.

In Section 3 we specialise to the study of groups whose co-word problem is context-free. We note that the word problem of such a group is solvable in cubic time, but that the conjugacy problem and generalised word problem may be unsolvable. However, the order problem is shown to be solvable. We prove that the restricted standard wreath product of a $co\mathcal{CF}$-group and a $\mathcal{CF}$-group is a $co\mathcal{CF}$-group. Consequently there exist $co\mathcal{CF}$-groups which are not finitely presentable.

We do not know of any $co\mathcal{CF}$-groups which do not arise in the ways already described. We believe that the class of $co\mathcal{CF}$-groups is not closed under taking free products. Indeed, we conjecture that $\mathbb{Z}^2 * \mathbb{Z}$ is not a $co\mathcal{CF}$-group, but we have been unable to prove this. In Section 4 we obtain some negative results, proving that polycyclic groups and Baumslag-Solitar groups are only $co\mathcal{CF}$-groups when they are virtually abelian. We use a technique based on Parikh's Theorem [**13**], which gives conditions on the kinds of context-free languages which can arise as subsets of languages of the form $w_1^* w_2^* \dots w_n^*$ (commonly known as bounded languages). This is quite a general technique, which we might well apply to exclude other classes of groups, but it most definitely cannot work to exclude free products.

Note that the results which we prove in Section 4 (Propositions 14 and 11) about context-free languages may be of independent interest.

A further paper [**6**] by two of the authors of this one studies groups whose co-word problem is an indexed language (accepted by a nested stack automaton). In particular it is proved in that paper that for all currently known pairs of $co\mathcal{CF}$-groups G and H, the free product $G * H$ has indexed co-word problem.

2. *General properties of $co\mathcal{C}$- and $\mathcal{C}$-groups*

Let $\mathcal{C}$ be a class of languages. Following [**7**], $\mathcal{C}$ is closed under *inverse homomorphisms* if whenever $\phi : X^* \to Y^*$ is a monoid homomorphism and $L \subset Y^*$ is in $\mathcal{C}$, then $\phi^{-1}(L) \in \mathcal{C}$. (Note that here, and later, we do not demand that L is contained in $\phi(X^*)$; $\phi^{-1}(L)$ is defined to be $\{w \in X^* \mid \phi(w) \in L\}$.) The following result (at least as far as word problems are concerned) is well known.

Lemma 1. *Let $\mathcal{C}$ be a class of languages closed under inverse homomorphisms and let G be a finitely generated group. Then the following hold.*

(i) *$W(G, X) \in \mathcal{C}$ for some finite generating set X if and only if for every finite generating set Y, $W(G, Y) \in \mathcal{C}$.*

(ii) *$coW(G, X) \in \mathcal{C}$ for some finite generating set X if and only if for every finite generating set Y, $coW(G, Y) \in \mathcal{C}$.*

In this case we say that $\mathcal{C}$- or $co\mathcal{C}$-groups are *insensitive* to choice of generators. We shall write $X^{\pm *}$ as a short hand notation for $(X \cup X^{-1})^*$ and $X^{\pm 1}$ for $X \cup X^{-1}$.

Proof. Let X and Y be two finite generating sets for G. Define $h : Y^{\pm *} \to X^{\pm *}$ as the homomorphism induced by expressing each $y \in Y^{\pm 1}$ as a selected word

$h(y) \in X^{\pm *}$ representing the same element of G as y. Then the following diagram commutes.

$$\begin{tikzcd} Y^{\pm *} \arrow[dd, "h"'] \arrow[dr] & \\ & G \\ X^{\pm *} \arrow[ur] & \end{tikzcd}$$

So $W(G, Y) = h^{-1}(W(G, X))$ and $coW(G, Y) = h^{-1}(coW(G, X))$, and the proof is complete. $\square$

LEMMA 2. *Let $\mathcal{C}$ be a class of languages closed under inverse homomorphisms and intersection with regular sets. Then $\mathcal{C}$-groups, as well as co$\mathcal{C}$-groups, are closed under taking finitely generated subgroups.*

Proof. Let $H \subset G$ be finitely generated groups. We choose a finite generating set X for G which includes a generating set X_0 for H. Then, $W(H, X_0) = X_0^{\pm *} \cap W(G, X)$ and $coW(H, X_0) = X_0^{\pm *} \cap coW(G, X)$. Since $X_0^{\pm *}$ is a regular language in $X^{\pm *}$, the result follows, by Lemma 1 and the hypothesis that $\mathcal{C}$ is closed under intersection with regular languages. $\square$

The *shuffle* (cf. p. 290 in [**2**]) of a language $L_1 \subset \Sigma^*$ with a language $L_2 \subset \Delta^*$ is defined as

$$L_1 \leftrightarrow L_2 = \{x_1 y_1 \dots x_n y_n \mid x_1 x_2 \dots x_n \in L_1, y_1 y_2 \dots y_n \in L_2, x_i \in \Sigma^*, y_i \in \Delta^*\}.$$

LEMMA 3. *Let $\mathcal{C}$ be a class of languages closed under shuffle with regular languages and under union. Then the class of co$\mathcal{C}$-groups is closed under taking finite direct products.*

Proof. Let (A, X) and (B, Y) be groups with finite generating sets such that $coW(A, X) \in \mathcal{C}$ and $coW(B, Y) \in \mathcal{C}$. Then it is easy to see that

$$coW(A \times B, X \cup Y) = (coW(A, X) \leftrightarrow Y^{\pm *}) \cup (coW(B, Y) \leftrightarrow X^{\pm *}),$$

which is a language in $\mathcal{C}$ by the hypothesis. $\square$

A *generalised sequential machine* (gsm) is a deterministic finite state automaton with output capacity. A useful way of representing a gsm is as a finite graph with doubly labelled edges, where the vertices correspond to the internal states and the double label $x|y$ of an edge specifies, on the one hand, the input symbol x which allows movement along this edge and, on the other hand, the output string y which is to be appended to the output so far. These machines do not accept or recognise languages but generate so called *generalised sequential machine mappings*. However, we use accept states to confirm the validity of the input, and hence the output.

For example, the gsm given by the graph in Figure 1 maps a string w in $\{a, b\}^*$ to the string x^n where n is the number of a's in w. The accept state makes sure the input was in $\{a, b\}^*$. Let us agree that every input string is terminated by an end of input symbol EOF.

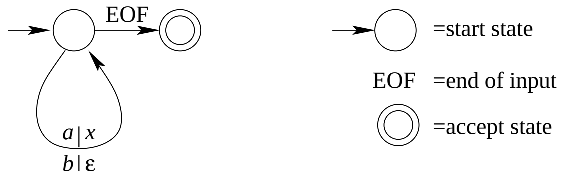

FIGURE 1

Generalised sequential machines are better known than, for instance the shuffle, which turns out to be the image of a gsm mapping provided that one of L_1, L_2 is a regular language. For suppose that M is a finite state automaton recognising the regular language $L_2 \subset \Delta^*$ and let Σ be a finite alphabet. View M as a labelled graph. Consider the gsm T obtained by modifying M as follows: replace each edge label x by the double label $\epsilon|x$, and for each state s and each $\sigma \in \Sigma$ introduce a new edge from s to itself with label $\sigma|\sigma$; here ϵ denotes the empty word. Then it is easy to check that the gsm mapping generated by T maps a language $L_1 \subset \Sigma^*$ to $L_1 \leftrightarrow L_2$, and we have the following direct consequence of Lemma 3.

COROLLARY 4. *Let $\mathcal{C}$ be a class of languages closed under union and gsm mappings. Then the class of co$\mathcal{C}$-groups is closed under taking finite direct products.*

When H is a subgroup of finite index in a group G, then we call G a *finite index overgroup* of H.

LEMMA 5. *Let $\mathcal{C}$ be a class of languages closed under union with regular sets and* *inverse gsm mappings. Then the classes of $\mathcal{C}$-groups and co$\mathcal{C}$-groups are closed under passing to finite index overgroups.*

Proof. Notice that a homomorphism is also a special case of a gsm mapping, whence we may use Lemma 1. Let H be a subgroup of finite index in G. Let T be a right transversal for H in G with $1 \in T$. Now every element g of G can be written in the form $g = ht$ with $h \in H$, $t \in T$. Let X be a finite generating set for H and put $Y = X \cup (T \setminus \{1\})$. Then Y is a (finite) generating set for G. For each $y \in Y^{\pm 1}$ and $t \in T$ fix a word $h_{ty} \in X^{\pm *}$ such that $ty =_G h_{ty}t'$ for some $t' \in T$. Now consider the gsm F given by the following graph (cf. Figure 2).

— state set $T \cup \{Q\}$ $(Q \notin T)$,

— $y|h_{ty}$ labelled edge from t to t', whenever $y \in Y^{\pm 1}$, $t, t' \in T$, and $ty = h_{ty}t'$,

— $EOF|t$ labelled edge from t to Q for all $t \in T \setminus \{1\}$,

— $EOF|\epsilon$ labelled edge from 1 to Q, (ϵ denotes the empty word), and

— $1 \in T$ is the start state and Q is the only accept state.

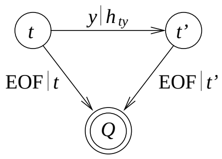

FIGURE 2

It is easy to see that the gsm mapping ϕ generated by F maps an arbitrary word $w \in Y^{\pm *}$ to a word $w't$ with $w' \in X^{\pm *}$, $t \in T$, and such that $w =_G w't$. We now have

$$coW(G, Y) = \phi^{-1}\left(coW(H, X) \cup X^{\pm *}(T \setminus \{1\})\right)$$

and

$$W(G, Y) = \phi^{-1}(W(H, X)),$$

and the proof is complete. $\square$

3. *Groups with context-free co-word problem*

Let $\mathcal{CF}$ denote the class of context-free languages. Using the fact that $\mathcal{CF}$ contains all regular languages and is closed under intersection with regular sets, union, gsm mappings, and inverse gsm mappings (eg. Theorem 1.7.2 and Chapter 3 in [**4**]), the results of the previous section immediately imply the following result.

PROPOSITION 6. *The class of co$\mathcal{CF}$-groups is insensitive to choice of generators and closed under passing to finitely generated subgroups, passing to finite index overgroups, and finite direct products.*

As noted in the introduction, the class of $\mathcal{CF}$-groups is precisely the class of virtually free groups which also coincides with the class of deterministic (co)$\mathcal{CF}$-groups; remember that deterministic context-free languages are closed under complementation. Since non-cyclic free abelian groups are direct products of virtually free groups but not virtually free, the proposition immediately implies that the class of $\mathcal{CF}$-groups is a proper subclass of the class of *co*$\mathcal{CF}$-groups.

Another useful property of *co*$\mathcal{CF}$-groups follows directly from the fact that every context-free language has membership problem solvable in cubic time (eg. pp. 139-140 in [**7**]).

PROPOSITION 7. *The word problem of every co$\mathcal{CF}$-group is solvable in cubic time in terms of the length of the input word.*

Since the direct product $F \times F$ of two copies of a free group F of rank at least two has a finitely generated subgroup G with unsolvable conjugacy problem and unsolvable generalised word problem in $F \times F$ (Theorem 23 in Chapter 3 of [**9**] or Theorem 4.6 in [**10**]), we have the following result contrasting Proposition 7.

PROPOSITION 8. *There exist co$\mathcal{CF}$-groups with unsolvable conjugacy problem and the generalised word problem is, in general, unsolvable for co$\mathcal{CF}$-groups.*

A group has *solvable order problem* if there is an algorithm which takes as input a word w in the generators and their inverses and decides whether the element represented by w has finite or infinite order.

Theorem 9. *Every co$\mathcal{CF}$-group has solvable order problem. Moreover, if w turns out to represent an element of finite order, then its order can be determined.*

Proof. Since w^* is a regular language, $L = \{\epsilon\} \cup (coW(G) \cap w^*)$ is a context-free language, where again ϵ denotes the empty word. Clearly, the element represented by w has infinite order in G if and only if $L = w^*$. Since w^* is a bounded language, the first part follows from Theorem 5.6.3(c) in [**4**]; it is decidable for arbitrary context-free languages M_1 and M_2, one of them bounded, whether $M_1 = M_2$. The second part is another application of the solvability of the membership problem for context-free languages; simply check whether $w^i \in L$ for $i = 1, 2, 3, \ldots$ until $w^j \notin L$. □

Recall that the class of context-free languages is the same as the class of languages accepted by nondeterministic pushdown automata. For later use and as an informal definition, let us describe such a machine P_H which accepts the word-problem of the free group H freely generated by the finite set Y. The input alphabet as well as the stack alphabet of P_H is $Y^{\pm 1}$ and P_H has one internal state q_0. Initially the stack of P_H is empty, that is the read-write head is scanning a special bottom of stack marker. When the read-write head is scanning the bottom of stack marker and the next input symbol is $y \in Y^{\pm 1}$, then y is pushed onto the stack (and strictly speaking the read head on the input tape is advanced, but we shall always gloss over this point). If, on the other hand, the read-write head is scanning some $y' \in Y^{\pm 1}$ and the next input symbol is $y \in Y^{\pm 1}$, then y is pushed onto the stack unless $y^{-1} = y'$ in which case y' is popped off the stack. Finally P_H accepts if and only if the read-write head is scanning the bottom of stack marker. Clearly, P_H is deterministic, and a pushdown automaton accepting the word problem of a virtually free group operates essentially in the same fashion, as can be seen from the proof of Lemma 5.

The following result is an interesting closure property for which we know only a machine-theoretic proof.

THEOREM 10. *Let G be a co$\mathcal{CF}$-group and let H be a $\mathcal{CF}$-group. Then the restricted standard wreath product, $G \wr H$, of G with H is a co$\mathcal{CF}$-group.*

Proof. Assume that G is generated by X and H is generated by Y. Then $W = G \wr H$ is generated by $X \cup Y$. By definition, the base group B of W is the direct product of copies G_h, $h \in H$, of G. We view B as the set of all functions b from H to G with $b(h)$ trivial for all but finitely many $h \in H$. Elements of B are multiplied component-wise. Now $h \in H$ acts on $b \in B$ by

$$b^h(h') = b(h'h^{-1}),$$

and W is the resulting semi-direct product. We identify G with the subgroup of the base group comprising those elements b with $b(h) = 1$ for all nontrivial $h \in H$. Below, by X-letters we mean elements of $X \cup X^{-1}$ and similarly for Y-letters.
Thus every element w of W is of the form bh, $b \in B$, $h \in H$, and w is nontrivial if and only if either h is nontrivial (in H) or b maps some element of H to a nontrivial element of G. The pushdown automaton P which accepts the co-word problem of W decides nondeterministically which of these possibilities it will try to confirm.
In order to see if h is nontrivial, P chooses to ignore all X-letters and simulates the deterministic pushdown automaton P_H described above, except that P accepts if and only if P_H rejects; note that B is the normal closure of G in W.
Let $w = w_1 w_2 \dots w_l \in (X^{\pm 1} \cup Y^{\pm 1})^*$, where $w_i \in X^{\pm 1} \cup Y^{\pm 1}$, and write $w_{(i)}$ for the prefix of length i of w. Furthermore, let $\overline{w}$ denote the word obtained from w by deleting all X-letters. Now suppose $w =_W bh$. Let $h' \in H$ and let I be the subset of all elements i of $\{1, 2, \dots, l\}$ such that $\overline{w_{(i)}} =_H h'^{-1}$ and $w_i \in X^{\pm 1}$. Then $\overline{w} =_H h$ and $b(h') =_G w_{i_1} w_{i_2} \dots w_{i_k}$, where $I = \{i_1, i_2, \dots, i_k\}$ and $i_j < i_{j+1}$ for $1 \leq j < k$. In other words, $b(h')$ is the subsequence of w consisting of all X-letters w_i for which $\overline{w_{(i)}}$ represents h'^{-1}.
Now we describe how P tries to verify that b maps some element of H to a nontrivial element of G. First P guesses a word v in $(Y^{\pm 1})^*$ and passes it to the deterministic pushdown automaton P_H. Let $h' \in H$ be the element represented by v. Now P shall investigate $b(h')$. By the previous paragraph, all P has to do now is pass Y-letters to P_H and ignore all X-letters unless P_H is in an accept configuration, in which case X-letters get passed to a pushdown automaton P_G accepting $coW(G, X)$. The

point is that, if P_H interprets the stack symbols of P_G as bottom of stack marker, then P_G and P_H can share the same stack because P_G only ever acts when the stack looks empty to P_H. It follows that this causes P to feed precisely the component $b(h')$ into P_G. We leave further details to the reader. $\square$

We conjecture that, if the standard restricted wreath product $G \wr H$ is a *co*$\mathcal{CF}$-group with nontrivial bottom group G, then the top group H has to be a $\mathcal{CF}$-group. Note that both G and H are *co*$\mathcal{CF}$-groups, by Proposition 6.

Using the same idea as in the proof of Theorem 10 one can show that the restricted standard wreath product of a finite group with a virtually cyclic top group has one counter co-word problem. One counter languages are those accepted by a pushdown automaton with only one stack symbol (apart from the bottom marker). Groups with one counter word problem are precisely the virtually cyclic groups (see [**5**]). This shows that there are groups whose co-word problems are among the simplest non regular languages and which are not finitely presentable.

It is well known that every group is the syntactic monoid of its word problem (see [**14**] for example). Since the syntactic monoid of L coincides with the syntactic monoid of $\Sigma^* \setminus L$, every group is also the syntactic monoid of its co-word problem. Thus every group with context-free co-word problem is the syntactic monoid of a context-free language. The examples and constructions given in this paper appear to extend the class of groups known to arise in this way.

4. *Groups whose co-word problem is not context-free*

In this section we prove that polycyclic groups and Baumslag-Solitar groups are *co*$\mathcal{CF}$-groups if and only if they are virtually abelian. We do this by means of semilinear sets which we define now.

Fix an integer $k > 0$, let $\mathbb{N}_0$ denote the non-negative integers including zero, and let $\mathbb{N}_0^k$ denote the set of all k-tuples of non-negative integers. Elements of $\mathbb{N}_0^k$ are added component-wise. A subset L_i of $\mathbb{N}_0^k$ is called *linear* if there exist $c_i \in \mathbb{N}_0^k$ and a finite subset $P_i = \{p_{i1}, \ldots, p_{ij_i}\}$ of $\mathbb{N}_0^k$ such that

$$L_i = \{\, c_i + \sum_{j=1}^{j_i} \alpha_{ij} p_{ij} \mid \alpha_{ij} \in \mathbb{N}_0 \,\}. \tag{4.1}$$

By definition, a subset L of $\mathbb{N}_0^k$ is *semilinear* if it is the union of finitely many linear sets.

We shall use Parikh's Theorem: if $M \subset w_1^* w_2^* \ldots w_k^*$ is context-free, then $L = \{\, (n_1, n_2, \ldots, n_k) \in \mathbb{N}_0^k \mid w_1^{n_1} w_2^{n_2} \ldots w_k^{n_k} \in M \,\}$ is semilinear ([**13**] or Theorem 5.2.1 in [**4**]).

Our strategy to show that a group G is not a $co\mathcal{CF}$-group is to intersect its co-word problem with $w_1^* \ldots w_k^*$ for certain words w_i, and show that the corresponding subset of $\mathbb{N}_0^k$ is not semilinear. For then, by Parikh's Theorem, $w_1^* \ldots w_k^* \cap coW(G)$, and hence $coW(G)$, are not context-free, as context-free languages are closed under intersection with regular languages.

We deal with the last step in Propositions 11 and 14. Although Proposition 11 is a special case of Proposition 14, we include an independent proof of the former in the hope that it makes the fairly technical proof of Proposition 14 more accessible.

Proposition 11. *Let $L \subseteq \mathbb{N}_0^{r+1}$ for some $r \in \mathbb{N}$. Suppose that L has the following property. For every $k \in \mathbb{N}$ there exists $(a_1, \ldots, a_r) \in \mathbb{N}_0^r \setminus \{(0, \ldots, 0)\}$, such that*

(i) *there is a unique $b \in \mathbb{N}_0$ with $(a_1, \ldots, a_r, b) \notin L$; and*

(ii) *if $(a_1, \ldots, a_r, b) \notin L$ then $b \geq k \sum_{i=1}^r a_i$.*

Then L is not semilinear.

Proof. The proof is by contradiction. Suppose that L is the union of the linear sets $L_1, L_2, \ldots, L_n$, and let c_i and P_i be as in (4.1), for $1 \leq i \leq n$. We may clearly assume that the elements in each of the P_i are all nonzero.

Order the L_i such that, for $1 \leq i \leq m$, P_i contains an element $(0, \ldots, 0, n_i)$ (where, by assumption, $n_i > 0$), and for $m+1 \leq i \leq r$, P_i contains no such element. It is possible that $m = 0$ or $m = n$. Let N be the least common multiple of the n_i $(1 \leq i \leq m)$, where $N = 1$ if $m = 0$.

Fix some i with $m+1 \leq i \leq n$. If $p_{ij} = (x_1, \ldots, x_{r+1}) \in P_i$, then $x_1 + \ldots + x_r$ cannot be zero, and so there exists $t \in \mathbb{N}$ such that $x_{r+1} < t(x_1 + \ldots + x_r)$. We can clearly choose the same t for each $p_{ij} \in P_i$ and, since the sum of two elements $(x_1, \ldots, x_{r+1}) \in \mathbb{N}_0^{r+1}$ that satisfy the condition $x_{r+1} < t(x_1 + \ldots + x_r)$ also satisfies that condition, we have $x_{r+1} < t(x_1 + \ldots + x_r) + q$ for all $(x_1, \ldots, x_{r+1}) \in L_i$, where $q \in \mathbb{N}_0$ is a constant, which we could take to be the last component of c_i.

Now let $C \in \mathbb{N}$ be twice the maximum of all of the constants t, q that arise for all P_i with $m+1 \leq i \leq n$. Then $x_{r+1} < C(x_1 + \ldots + x_r)$ whenever $(x_1, \ldots, x_{r+1}) \in L_i$ with $m+1 \leq i \leq n$.

Let $k \geq 2\max\{N, C\}$ and let $(a_1, \ldots, a_r)$ satisfy (i) and (ii) of the proposition for this k, with $b \in \mathbb{N}_0$ say. The uniqueness of b implies that $(a_1, \ldots, a_r, b - N) \in L$ (note that $b > N$). In particular, $(a_1, \ldots, a_r, b - N) \in L_i$ for some i with $1 \leq i \leq n$. It follows that $i > m$, for otherwise we could add $(N/n_i)(0, \ldots, 0, n_i) \in \mathbb{N}P_i$ to $(a_1, \ldots, a_r, b - N)$ to get $(a_1, \ldots, a_r, b) \in L_i$ in contradiction to (i). So, by the previous paragraph, we have $b - N < C(a_1 + \ldots + a_r)$, or equivalently, $b < C(a_1 + \ldots + a_r) + N \leq 2\max\{C, N\}(a_1 + \ldots + a_r) \leq k(a_1 + \ldots + a_r)$, as $(a_1 + \ldots + a_r) \geq 1$. But this contradicts condition (ii) and the result is established. □

THEOREM 12. *A finitely generated nilpotent group has context-free co-word problem if and only if it is virtually abelian.*

Proof. By Proposition 6, every finitely generated virtually abelian group is a $co\mathcal{CF}$-group.

Now assume G is a finitely generated nilpotent but not virtually abelian group. It is well known that G has a torsion free nilpotent but not virtually abelian subgroup. Furthermore, every non abelian torsion free nilpotent group has a subgroup isomorphic to the Heisenberg group

$$H = \langle A, B, C \mid [A, B] = C, [A, C] = [B, C] = 1 \rangle,$$

where $[x, y]$ denotes the commutator $x^{-1}y^{-1}xy$. To see this let A be a non central element of the second term of the upper central series of G and let B be some element not commuting with A.

By Proposition 6, it suffices to show that H is not a $co\mathcal{CF}$-group. Since $[A^m, B^m] = C^{m^2}$ holds in H for all $m \geq 0$, the subset of $\mathbb{N}_0^5$ corresponding to the intersection of $(A^{-1})^*(B^{-1})^*A^*B^*(C^{-1})^*$ with $coW(H, \{A, B, C\})$ satisfies the condition of Proposition 11. (Given k, consider (m, m, m, m) with $m \geq 4k$.) This completes the proof. $\square$

Recall that a Baumslag-Solitar group is a group with a presentation of the form $\langle x, y \mid y^{-1}x^p y = x^q \rangle$, where $p, q \in \mathbb{Z} \setminus \{0\}$, eg. [**8**].

Theorem 13. *A Baumslag-Solitar group is a co$\mathcal{CF}$-group if and only if it is virtually abelian.*

Proof. Let G be given by the presentation $\langle x, y \mid y^{-1}x^p y = x^q \rangle$. It is well known that G is virtually abelian precisely when $p = \pm q$. We deal here with the case $0 < p < q$, the other cases being similar. Then the subset L of $\mathbb{N}_0^4$ corresponding to $coW(G) \cap (y^{-1})^*(x^{-1})^* y^* x^*$ does not contain (n, p^n, n, q^n) for any $n \in \mathbb{N}$ and q^n is the only value of x for which (n, p^n, n, x) is not in L. Moreover, since $q > p$, for every given k there exists n with $k(2n + p^n) \leq q^n$. But this simply means that L satisfies the hypothesis of Proposition 11 and the proof is complete. $\square$

Let us now turn to the generalisation of Proposition 11 which enables us to deal with polycyclic groups. We shall write elements $(a_1, \ldots, a_r, b_1, \ldots, b_s)$ of $\mathbb{N}_0^{r+s}$ as $(\mathbf{a}; \mathbf{b})$, where $\mathbf{a} = (a_1, \ldots, a_r) \in \mathbb{N}_0^r$ and $\mathbf{b} = (b_1, \ldots, b_s) \in \mathbb{N}_0^s$. We shall denote the zero vector in $\mathbb{N}_0^r$ or $\mathbb{N}_0^s$ by $\mathbf{0}$. For any vector $\mathbf{v}$, $\sigma(\mathbf{v})$ will denote the sum of its components: for example $\sigma(\mathbf{a}) = \sum_{i=1}^r a_i$.

Proposition 14. *Let $L \subseteq \mathbb{N}_0^{r+s}$ for some $r, s \in \mathbb{N}$. Suppose that L has the following property. For every $k \in \mathbb{N}$, there exists $\mathbf{a} \in \mathbb{N}_0^r \setminus \{\mathbf{0}\}$, such that*

(i) *there is a unique $\mathbf{b} \in \mathbb{N}^s$ with $(\mathbf{a}; \mathbf{b}) \notin L$; and*
(ii) *if $(\mathbf{a}; \mathbf{b}) \notin L$ then $b_j \geq k\sigma(\mathbf{a})$ for $1 \leq j \leq s$.*

Then L is not semilinear.

Proof. Again we argue by contradiction. Suppose that L is semilinear and is the union of linear sets $L_1, \ldots, L_n$, where

$$L_i = \{ \mathbf{u}_i + \sum_{j=1}^{j_i} \alpha_{ij} \mathbf{v}_{ij} \mid \alpha_{ij} \in \mathbb{N}_0 \},$$

$P_i = \{\mathbf{v}_{i1}, \ldots, \mathbf{v}_{ij_i}\}$ is the set of periods of L_i, and $\mathbf{u}_i$, $\mathbf{v}_{ij}$ all lie in $\mathbb{N}_0^{r+s}$. We may assume that all periods are nonzero.

We order the L_i such that, for $1 \le i \le m$, P_i contains an element $(\mathbf{0};\mathbf{b})$ (where, by assumption, $\mathbf{b} \ne \mathbf{0}$), and for $m+1 \le i \le n$, P_i contains no such element. It is possible that $m = 0$ or $m = n$.

Our strategy is similar to that used in the proof of Proposition 11, in that we aim to find a vector $(\mathbf{a};\mathbf{b})$ that is not in L, where $\mathbf{b}$ is 'large' compared with $\mathbf{a}$, and to consider a suitable $(\mathbf{a};\mathbf{b}-\mathbf{v})$ which is in L. The largeness of $\mathbf{b}$ will enable us to conclude, as in the earlier proof, that $(\mathbf{a};\mathbf{b}-\mathbf{v})$ is not in L_i for $i > m$. The argument for $i \le m$ is more complicated here than in Proposition 11. We shall identify a subset $\mathcal{L}$ of $\{L_1, L_2, \ldots, L_m\}$ consisting of those L_i that are in some sense close to $(\mathbf{a};\mathbf{b})$, and show that $(\mathbf{a};\mathbf{b}-\mathbf{v})$ has to lie in an L_i in $\mathcal{L}$. The vector $\mathbf{v}$ will be chosen such that $(\mathbf{0};\mathbf{v})$ is in the space spanned by the periods of L_i for all $L_i \in \mathcal{L}$, and so can be added to $(\mathbf{a};\mathbf{b}-\mathbf{v})$ to obtain a vector in L, thereby yielding a contradiction. Before we can choose $(\mathbf{a};\mathbf{b})$ and $\mathbf{v}$, we need to define a number of constants.

As in the proof of Proposition 11, we can find a constant C_1 such that, for any $(\mathbf{a};\mathbf{b}) \in L_i$ with $m+1 \le i \le n$, we have $b_j < C_1\sigma(\mathbf{a})$ for $1 \le j \le s$.

For $1 \le i \le m$, we assume that the periods in P_i are ordered such that $\mathbf{v}_{ij} = (\mathbf{0};\mathbf{b}_{ij})$ for $1 \le j \le k_i$, and $\mathbf{v}_{ij} = (\mathbf{a};\mathbf{b})$ with $\mathbf{a} \ne \mathbf{0}$ for $k_i < j \le j_i$. By definition of m, we have $k_i > 0$ for $1 \le i \le m$.

For $1 \le i \le m$, let R_i, Q_i, Z_i be respectively the real, rational, and integral submodules of $\mathbb{R}^s, \mathbb{Q}^s, \mathbb{Z}^s$ spanned by $\{\mathbf{b}_{ij} \mid 1 \le j \le k_i\}$. Let

$$R_i^+ = \{\sum_{j=1}^{k_i} \lambda_j \mathbf{b}_{ij} \mid \lambda_j \in \mathbb{R}, \lambda_j \ge 0\},$$

and define Q_i^+ and Z_i^+ correspondingly.

Let $(\mathbf{a};\mathbf{b}) \in L_i$, where $1 \le i \le m$ and $\mathbf{a} \ne \mathbf{0}$. We shall show now that $\mathbf{b}$ is close to an element of Z_i^+. Note that $\mathbf{b}$ is the sum of a constant vector coming from $\mathbf{u}_i$, an element of Z_i^+, and at most $\sigma(\mathbf{a})$ vectors coming from periods $\mathbf{v}_{ij}$ with $j > k_i$. It follows that there is a constant C_2 such that for any i with $1 \le i \le m$, and any such $(\mathbf{a};\mathbf{b})$ with $\mathbf{a} \ne \mathbf{0}$, we have $d(\mathbf{b}, Z_i^+) \le C_2\sigma(\mathbf{a})$, and hence $d(\mathbf{b}, R_i^+) \le C_2\sigma(\mathbf{a})$, where d is the standard Euclidean metric on $\mathbb{R}^s$.

As mentioned above, we shall eventually be choosing a specific subset $\mathcal{L}$ of $\{L_1, L_2, \ldots, L_m\}$. We cannot do this yet, because we are not ready to choose our vector $(\mathbf{a};\mathbf{b})$; we need first to establish some general properties of such subsets $\mathcal{L}$. Define $\mathcal{I}$ to be the set of all subsets I of $\{1, 2, \ldots, m\}$ such that the intersection of the R_i^+ with $i \in I$ is not $\{\mathbf{0}\}$, that is

$$I \in \mathcal{I} \iff \cap_{i \in I} R_i^+ \neq \{\mathbf{0}\}$$

Suppose first that $I \in \mathcal{I}$. It follows from Lemma 15 below that the intersection of the Q_i^+ contains a nonzero vector in $\mathbb{Q}^s$, and hence by multiplying by the least common multiple of the denominators of the coefficients, the intersection of the Z_i^+ for $i \in I$ contains a nonzero vector in $\mathbb{Z}^s$ and hence in $\mathbb{N}_0^s$. For each $I \in \mathcal{I}$ choose one such vector $\mathbf{v}_I$ and let l be the largest Euclidean length of any of these $\mathbf{v}_I$, $I \in \mathcal{I}$. One of these vectors $\mathbf{v}_I$ will later serve as the vector $\mathbf{v}$ in $(\mathbf{a}; \mathbf{b} - \mathbf{v})$ as discussed above.

Suppose, on the other hand, that I is a subset of $\{1, 2, \ldots, m\}$ but not in $\mathcal{I}$. In this case, we shall show that a 'large' vector in $\mathbb{R}^s_{\geq 0}$ is a long way from R_i^+ for at least one $i \in I$. For $c \in \mathbb{R}, c > 0$, let $H_c = \{\mathbf{v} \in \mathbb{R}^s_{\geq 0} \mid \sigma(\mathbf{v}) = c\}$. Let $\mathbf{v} \in H_c$. Since $\mathbf{v}$ does not lie in all of the R_i^+ with $i \in I$, and each R_i^+ is closed, we have $\max\{d(\mathbf{v}, R_i^+) \mid i \in I\} > 0$. Clearly $\max\{d(\mathbf{v}, R_i^+) \mid i \in I\}$ is a continuous function of $\mathbf{v}$, and since H_c is compact, and the image of a continuous function on a compact set is compact, there exists $D_c \in \mathbb{R}$, $D_c > 0$ such that $\max\{d(\mathbf{v}, R_i^+) \mid i \in I\} \geq D_c$ for all $\mathbf{v} \in H_c$. Now to every $\mathbf{v} \in H_1$ and $\mathbf{w} \in R_i^+$ there are corresponding points $c\mathbf{v} \in H_c$ and $c\mathbf{w} \in R_i^+$ with $d(c\mathbf{v}, c\mathbf{w}) = cd(\mathbf{v}, \mathbf{w})$, so we can take $D_c = cD_1$ for all $c > 0$.

We are finally ready to choose our vector $(\mathbf{a};\mathbf{b})$ and to define our subset $\mathcal{L}$ of $\{L_1, L_2, \ldots, L_m\}$. We apply the hypothesis of the theorem with k chosen such that $D_1 k > C_2 + l$ and $k > C_1 + l$. Then there exist corresponding $\mathbf{a}$ and $\mathbf{b}$ for this k. Let $I = \{i \in \{1, 2, \ldots, m\} \mid d(\mathbf{b}, R_i^+) \leq (C_2 + l)\sigma(\mathbf{a})\}$. (Then our $\mathcal{L}$ is just $\{L_i \mid i \in I\}$.)

Suppose first that $I \notin \mathcal{I}$. Then, as we saw above, if $c = \sigma(\mathbf{b})$, then there is an $i \in I$ with $d(\mathbf{b}, R_i^+) \geq D_c = cD_1$. By hypothesis, each component b_j of $\mathbf{b}$ is at least $k\sigma(\mathbf{a})$, and so certainly $c \geq k\sigma(\mathbf{a})$, and hence $d(\mathbf{b}, R_i^+) \geq D_1 k\sigma(\mathbf{a}) > (C_2 + l)\sigma(\mathbf{a})$, contrary to the definition of I.

Hence $I \in \mathcal{I}$ and, as we saw above, the intersection of the Z_i^+ for $i \in I$ contains the nonzero vector $\mathbf{v}_I$; in particular $|\mathbf{v}_I| \le l$. By the hypothesis of the theorem, $\mathbf{b}$ is unique with $(\mathbf{a};\mathbf{b}) \notin L$, and so $(\mathbf{a};\mathbf{b}-\mathbf{v}_I) \in L$.

Since each component b_j of $\mathbf{b}$ satisfies $b_j \ge k\sigma(\mathbf{a}) > (C_1 + l)\sigma(\mathbf{a})$, and the components of $\mathbf{v}_I$ are at most l, the components of $\mathbf{b}-\mathbf{v}_I$ are at least $C_1\sigma(\mathbf{a})$, and so, by definition of C_1, we cannot have $(\mathbf{a};\mathbf{b}-\mathbf{v}_I) \in L_i$ for $i > m$.

If $i \in \{1,2,\ldots,m\}$ with $i \notin I$, then by choice of I we have $d(\mathbf{b}, R_i^+) > (C_2+l)\sigma(\mathbf{a})$, and since $|\mathbf{v}_I| \le l$, this implies that $d(\mathbf{b}-\mathbf{v}_I, R_i^+) > C_2\sigma(\mathbf{a})$. But then, by definition of C_2, we cannot have $(\mathbf{a};\mathbf{b}-\mathbf{v}_I) \in L_i$.

Hence we must have $(\mathbf{a};\mathbf{b}-\mathbf{v}_I) \in L_i$ for some $i \in I$. But then $\mathbf{v}_I \in Z_i^+$, and so $(\mathbf{a};\mathbf{b}-\mathbf{v}_I) + (\mathbf{0};\mathbf{v}_I) = (\mathbf{a};\mathbf{b}) \in L_i \subseteq L$, contrary to assumption. □

After this final contradiction, the proof of Proposition 14 is complete once we establish the following result which we used above.

LEMMA 15. *Let $s \in \mathbb{N}$, let $P_1, P_2, \ldots, P_m$ be finite subsets of $\mathbb{Q}^s$, and let $Q_1, \ldots, Q_m$ and $R_1, \ldots, R_m$ be respectively the rational and real subspaces of $\mathbb{Q}^s$ and $\mathbb{R}^s$ spanned by $P_1, \ldots, P_m$. Then the following hold.*

(i) *The intersection Q_I of the Q_i has the same dimension as the intersection R_I of the R_i.*

(ii) *Suppose that, for some i with $1 \le i \le m$, we have $\lambda_1\mathbf{v}_1 + \ldots \lambda_r\mathbf{v}_r \in R_I$ where, for $1 \le j \le r$, $\mathbf{v}_j \in P_i$, and $\lambda_j \in \mathbb{R}$ with $\lambda_j > 0$. Then there exist $\mu_1, \ldots, \mu_r \in \mathbb{Q}$ with each $\mu_j > 0$ and $\mu_1\mathbf{v}_1 + \ldots \mu_r\mathbf{v}_r \in Q_I$.*

Proof. We prove (i) by induction on m. For $m = 1$, $Q_I = Q_1$, $R_I = R_1$, and the dimensions of Q_I and R_I are equal to the $\mathbb{Q}$-rank and $\mathbb{R}$-rank of the matrix A of which the rows are the vectors in P_1. But these ranks are both equal to the degree of the largest square submatrix of A having nonzero determinant, so they are equal.

When $m > 1$, we apply the formula $\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V)$ for subspaces U and V of a vector space, with U equal to the intersection of $Q_1, \ldots, Q_{m-1}$ (respectively $R_1, \ldots, R_{m-1}$), and $V = Q_m$ (respectively R_m), and the result follows by induction on m.

By (i), any basis of Q_I is a basis of R_I. Using this basis and the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$, it follows that Q_I is dense in R_I. Now let $\mathbf{v} := \lambda_1\mathbf{v}_1 + \ldots \lambda_r\mathbf{v}_r \in R_I$ be as in (ii), with each $\lambda_j \in \mathbb{R}$, $\lambda_j > 0$. We may assume that $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is the whole of P_i, since otherwise we can prove the whole lemma with P_i replaced by that subset. Then, for any $\epsilon > 0$, we can find $\mathbf{v}' \in Q_I$ with $|\mathbf{v}' - \mathbf{v}| < \epsilon$. By definition of Q_I, we have $\mathbf{v}' = \nu_1\mathbf{v}_1 + \ldots \nu_r\mathbf{v}_r$ for some $\nu_j \in \mathbb{Q}$. We would like each $|\nu_j - \lambda_j|$ to be small, in order to force ν_j to be positive, but since P_i is not necessarily a subset of Q_I, we cannot achieve this simply by choosing the ν_j to be rational numbers close to the λ_j.

Let $\phi : \mathbb{R}^r \to \mathbb{R}^s$ be the linear map which maps $(\alpha_1, \alpha_2, \ldots, \alpha_r) \in \mathbb{R}^r$ to $\alpha_1\mathbf{v}_1 + \ldots + \alpha_r\mathbf{v}_r$. Then $\phi(\mathbf{u}) = \mathbf{v}$ and $\phi(\mathbf{u}') = \mathbf{v}'$, where $\mathbf{u} = (\lambda_1, \ldots, \lambda_r)$ and $\mathbf{u}' = (\nu_1, \ldots, \nu_r)$. We are looking to express $\mathbf{v}'$ as a sum $\mu_1\mathbf{v}_1 + \ldots \mu_r\mathbf{v}_r$ with each $\mu_j \in \mathbb{Q}$ and $|\mu_j - \lambda_j|$ small. In other words, we are looking for $\mathbf{u}'' = (\mu_1, \ldots, \mu_r) \in \mathbb{Q}^r$ with $\phi(\mathbf{u}'') = \phi(\mathbf{u}')$ and $|\mathbf{u}'' - \mathbf{u}|$ small. Let $N_\mathbb{R}$ be the nullspace of ϕ, and let W be a complementary subspace to $N_\mathbb{R}$ in $\mathbb{R}^r$. Then ϕ maps W isomorphically onto $\mathrm{im}(\phi)$. Let $\mathbf{x} \in W$ with $\phi(\mathbf{x}) = \mathbf{v}' - \mathbf{v}$. Since $|\mathbf{v}' - \mathbf{v}| < \epsilon$, we have $|\mathbf{x}| < K\epsilon$, where K is the norm of the inverse map of $\phi_W : W \to \mathrm{im}(\phi)$. We have $\phi(\mathbf{x} + \mathbf{u}) = \mathbf{v}' - \mathbf{v} + \mathbf{v} = \mathbf{v}'$, and $|(\mathbf{x} + \mathbf{u}) - \mathbf{u}| = |\mathbf{x}|$ is small, but $\mathbf{x} + \mathbf{u} \notin \mathbb{Q}^r$, so we cannot just take $\mathbf{u}'' = \mathbf{x} + \mathbf{u}$. We have $\mathbf{x} + \mathbf{u} - \mathbf{u}' \in N_\mathbb{R}$, and by a similar argument to that used in the proof of (i), the dimension of $N_\mathbb{R}$ is equal to the dimension of the rational space $N_\mathbb{Q} := N_\mathbb{R} \cap \mathbb{Q}^r$, so $N_\mathbb{Q}$ is dense in $N_\mathbb{R}$, and we can find $\mathbf{y} \in \mathbb{N}_Q$ with $|\mathbf{y} - (\mathbf{x} + \mathbf{u} - \mathbf{u}')| < \epsilon$. So $\mathbf{u}' + \mathbf{y} \in \mathbb{Q}^r$ with

$$|(\mathbf{u}' + \mathbf{y}) - \mathbf{u}| \leq |\mathbf{y} - \mathbf{x} - \mathbf{u} + \mathbf{u}'| + |\mathbf{x}| < (K + 1)\epsilon,$$

and so, with a suitable small choice of ϵ, $\mathbf{u}'' := \mathbf{u}' + \mathbf{y}$ will have the required properties. That is, $\mu_1\mathbf{v}_1 + \ldots + \mu_r\mathbf{v}_r = \mathbf{v}' \in Q_I$ with each $\mu_j \in \mathbb{Q}$ and $\mu_j > 0$, where $\mathbf{u}'' = (\mu_1, \ldots, \mu_r)$. Thus the lemma holds. ☐

Theorem 16. *A polycyclic group has context free co-word problem if and only if it is virtually abelian.*

Proof. Let G be a counterexample. Since all subgroups of a polycyclic group are finitely generated, we can, using Proposition 6, at any time replace G by any subgroup of G that is not virtually abelian.

Clearly G is infinite, so by 5.4.15 of [**15**] G has a nontrivial free abelian normal subgroup N. Choose such a subgroup of largest possible rank. Since G is not virtually abelian, G/N is also infinite, and so has a nontrivial free abelian normal subgroup M/N. If M were virtually abelian, then it would have a free abelian subgroup L of finite index t, say, and then M^t would be free abelian of finite index in M and normal in G. But then the rank of M^t would be greater than that of N, contrary to the choice of N, so M is not virtually abelian, and we can therefore assume that $M = G$. So G/N is free abelian.

Let $g \in G \setminus N$. If $\langle g, N\rangle$ were virtually abelian, then some power g^t of g with $t > 0$ would centralise N, and then $\langle g^t, N\rangle$ would be a normal free abelian subgroup of G of rank greater than that of N. So $\langle g, N\rangle$ is not virtually abelian, and we can assume that $G = \langle g, N\rangle$.

In general, the minimal polynomial over $\mathbb{Q}$ of a matrix with entries in $\mathbb{Z}$ divides the characteristic polynomial, and hence, by Gauss' Lemma, is a monic polynomial in $\mathbb{Z}[x]$. If the inverse of the matrix also has integral entries then its determinant is ± 1, and hence so also is the constant term of the minimal polynomial.

This applies, in particular, to the minimal polynomial $p(x)$ of the action of g by conjugation on N. If each of the irreducible factors of p are cyclotomic polynomials, then the action of some power g^t $(t > 0)$ of g on N satisfies a polynomial of the form $(x-1)^m$ for some $m > 0$, and then $\langle g^t, N\rangle$ is nilpotent, G is virtually nilpotent, and the result follows from Theorem 12.

Hence we can assume that $p(x) = q(x)r(x)$, where $q \in \mathbb{Z}[x]$ is irreducible and is not cyclotomic. We can regard N as a $\mathbb{Z}[x]$-module, where the action of x is the conjugation action of g, and then $R := r(x)N$ is a submodule on which g acts with minimal polynomial $q(x)$. We can now assume that $G = \langle g, R\rangle$, and then $N = R$ and $p(x) = q(x)$ is irreducible. Similarly, by replacing N by a submodule if necessary, we can assume that N is generated as a $\mathbb{Z}[x]$-module by a single element v, and hence the rank n of N is equal to the degree of p. We can also assume that, for any $t > 0$, the minimal polynomial $p_t(x)$ of the action of g^t on N is irreducible and has degree n, for otherwise we can replace G by $\langle g^t, N\rangle$.

Let

$$p_t(x) = x^n + a_{t,n-1}x^{n-1} + \ldots + a_{t1}x + a_{t0}.$$

By Lemma 11.6 of [**16**], if all of the complex roots $\lambda_1, \ldots, \lambda_n$ of $p(x)$ had absolute value 1, then they would all be roots of unity, but then $p(x)$ would be cyclotomic, which we are assuming is not the case. Since the product of these roots is the constant term $a_{1,0}$ of $p(x)$, which is 1 or -1, there must be some root which has absolute value greater than 1. Let us order the λ_i such that $\lambda_1, \ldots, \lambda_m$ are the roots with largest absolute value $\lambda = |\lambda_1| = \ldots = |\lambda_m|$. For any $t > 0$, the roots of $p_t(x)$ are λ_i^t. Since the coefficient $a_{t,n-m}$ of x^{n-m} in $p_t(x)$ is plus or minus the sum of the products of the λ_i^t taken m at a time, and $\lambda_1^t \lambda_2^t \ldots \lambda_m^t$ is the unique such product with largest absolute value λ^{tm}, there is a constant $c > 0$ with the property that, for large enough t, we have $|a_{t,n-m}| > c\lambda^t$.

Let $I = \{0, 1, \ldots, n-1\}$, and choose a maximal subset J of I with the property that, for any constant $k \in \mathbb{N}$, there exists $t > 0$ such that $|a_{tj}| > kt$ for all $j \in J$. The statement at the end of the preceding paragraph shows that $\{n-m\}$ has this property, so J is nonempty. Now there must exist a constant $C > 0$ such that whenever $|a_{tj}| > Ct$ for all $j \in J$, we have $|a_{ti}| \leq Ct$ for all $i \in I \setminus J$, since otherwise J would not be maximal. [Suppose not. Then for all $C > 0$, there exists $t > 0$ such that $|a_{tj}| > Ct$ for all $j \in J$, and also $|a_{ti}| > Ct$ for some $i \in I \setminus J$. Then, since I is finite, there must be some $i \in I \setminus J$ that occurs in this way for infinitely many $C \in \mathbb{N}$, and then $J \cup \{i\}$ has the property for which J was supposed to be maximal.]

For each $k \in \mathbb{N}$, let t_k be a specific value of t for which the property in the preceding paragraph holds for the set J, and let $(s_{k,n-1}, s_{k,n-2}, \ldots, s_{k1}, s_{k0})$ be the sequence of signs of the coefficients $(a_{t_k,n-1}, a_{t_k,n-2}, \ldots, a_{t_k,1}, a_{t_k,0})$ in $p_{t_k}(x)$, where each s_{kj} is 1 or -1. There must be some such sequence $(s_{n-1}, s_{n-2}, \ldots, s_1, s_0)$ of signs that occurs for infinitely many k, and then we can alter our choice of t_k if necessary to ensure that this same sequence occurs for all $k \in \mathbb{N}$.

Let v be a nontrivial element of N. Then, by definition of $p_t(x)$, we have

$$(g^{-nt} v g^{nt})(g^{-(n-1)t} v^{a_{t,n-1}} g^{(n-1)t}) \ldots (g^{-t} v^{a_{t1}} g^t)(v^{a_{t0}}) =$$
$$g^{-nt} v g^t v^{a_{t,n-1}} g^t \ldots g^t v^{a_{t1}} g^t v^{a_{t0}} = 1.$$

and furthermore, $a_{t,n-1}, a_{t,n-2}, \ldots, a_{t1}, a_{t0}$ are the only values of $b_{n-1}, b_{n-2}, \ldots, b_1, b_0$, respectively, for which

$$g^{-nt} v g^t v^{b_{n-1}} g^t v^{b_{n-2}} \ldots g^t v^{b_1} g^t v^{b_0} = 1,$$

for otherwise g^t would satisfy a polynomial of degree less than n in its action on the $\mathbb{Z}[x^t]$-submodule of N generated by v, and then $p_t(x)$ would not be irreducible, contrary to what we assumed above.

Now let L be the subset of $\mathbb{N}_0^{2n+1}$ defined by $(a_1, a_2, \ldots, a_{2n+1}) \in L$ if and only if

$$\overline{g}^{a_1} v g^{a_2} v_{n-1}^{a_3} g^{a_4} v_{n-2}^{a_5} g^{a_6} \ldots g^{a_{2n-2}} v_1^{a_{2n-1}} g^{a_{2n}} v_0^{a_{2n+1}} \neq 1,$$

where $\overline{g} = g^{-1}$ and, for $0 \leq j \leq n-1$, $v_i = v$ if $s_i = 1$ and $v_i = v^{-1}$ if $s_i = -1$. Then, for any $k \in \mathbb{N}$, there exists $t = t_k \in \mathbb{N}$, such that

$$(nt, t, |a_{t,n-1}|, t, |a_{t,n-2}|, t, \ldots, t, |a_{t1}|, t, |a_{t0}|)$$

is the unique element of $\mathbb{N}_0^{2n+1}$ of the form $(nt, t, a_3, t, a_5, t, \ldots, t, a_{2n-1}, t, a_{2n+1})$ that is not in L. Furthermore, provided that $k > C$, we have $|a_{tj}| \geq kt$ for $j \in J$ and $|a_{tj}| \leq Ct$ for $j \in I \setminus J$. Since the semilinearity of a subset of $\mathbb{N}_0^{2n+1}$ does not depend on the order of its components, we can now apply Proposition 14 to L to deduce that L is not semilinear. In this application, the final s components of the vectors in Proposition 14 correspond to the components of the vectors in $\mathbb{N}_0^{2n+1}$ containing the entries $|a_{t,j}|$ with $j \in J$ in the displayed vector above; so $s = |J|$. The first r components of the vectors in Proposition 14 correspond to all other components of the vectors in $\mathbb{N}_0^{2n+1}$.

Since L is the intersection of the co-word problem of G with a regular set, this completes the proof of the theorem. $\square$

Acknowledgements. The fourth author would like to thank Hilary Craig for all her help and encouragement.

References

1. V. A. Anisimov, 'The group languages', *Kibernetika* 4 (1971) 18–24.

2. R. Beigel and R. W. Floyd, *The language of machines: an introduction to computability and formal languages* (Computer Science Press, New York, 1994).

3. M. Dunwoody, 'The accessibility of finitely presented groups', *Invent. Math.* 81 (1985) 449–457.

4. S. Ginsburg, *The mathematical theory of context-free languages* (McGraw-Hill, New York, 1966).

5. T. Herbst, 'On a subclass of context-free groups', *RAIRO Inform. Théoret. Appl.* 25 (1991) 255–272.

6. D. F. Holt and C. E. Röver, 'Groups with indexed co-word problem', in preparation.

7. J. E. Hopcroft and J. D. Ullman, *Introduction to automata theory, languages, and computation* (Addison-Wesley Publishing Co., Reading, Mass., 1979).

8. A. Karrass, W. Magnus and D. Solitar, *Combinatorial Group Theory* (Dover Publications Inc., New York, 1976).

9. C. F. Miller III, *On group-theoretic decision problems and their classification*, no. 68 in Annals of Mathematics Studies (Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1971).

10. C. F. Miller III, 'Decision problems for groups — survey and reflection', *Algorithms and Classification in Combinatorial Group Theory* (eds. G. Baumslag and C. F. Miller III, Springer-Verlag, 1992), pp. 1–60.

11. D. E. Muller and P. E. Schupp, 'Groups, the theory of ends, and context-free languages', *J. Comp. System Sci.* 26 (1983) 295–310.

12. D. E. Muller and P. E. Schupp, 'The theory of ends, pushdown automata, and second-order logic', *Theoret. Comput. Sci.* 37 (1985) 51–75.

13. R. J. Parikh, 'Language generating devices', *M.I.T. Res. Lab. Electron. Quart. Prog. Rept.* 60 (1961) 199–212.

14. D. W. Parkes and R. M. Thomas, 'Syntactic monoids and word problems', *Arab. J. Sci. Eng. Sect. C Theme Issues* 25 (2000) 81–94.

15. D. J. S. Robinson, *A Course in the Theory of Groups* (Springer-Verlag, New York, Heidelberg, Berlin, 1993).

16. I. N. Stewart and D. O. Tall, *Algebraic Number Theory* (Second Edition) (Chapman & Hall, London, 1987).

D. F. Holt
Mathematics Institute
University of Warwick
Coventry CV4 7AL

dfh@maths.warwick.ac.uk

S. Rees and C. E. Röver
University of Newcastle-upon-Tyne
School of Mathematics and Statistics
Merz Court
Newcastle upon Tyne NE1 7RU

sarah.rees@newcastle.ac.uk
c.h.e.w.roever@newcastle.ac.uk

R. M. Thomas
Department of Mathematics and
Computer Science
University of Leicester
University Road
Leicester LE1 7RH

rmt@mcs.le.ac.uk