HIGMAN'S EMBEDDING THEOREM AND DECISION PROBLEMS

ALEX BURKA

Abstract. We exposit Higman's embedding theorem, which states the finitely generated and recursively presented subgroups of a finitely presented group are the same. To that end we present a few (semi)group-theoretic constructions—amalgamated free products, HNN extensions, and semigroups induced by Turing machines. After we prove Higman's theorem, our discussion culminates in some applications to decision problems in group theory and topology.

Contents

Introduction

Our story begins with Max Dehn's 1907 *Analysis situs*, which includes a detailed proof of the Classification of Surfaces. Its neat characterization of (closed, connected) 2-manifolds raises interest in the homeomorphism problem for higher dimensions: For which $n > 2$ can we classify the homeomorphism classes of manifolds in dimension n?

Dehn couldn't solve the homeomorphism problem for $n = 3$,[1] but introduced far-reaching techniques still relevant in geometric topology, e.g. Dehn surgeries. The Classification of Surfaces and Dehn surgeries involve a common group-theoretic apparatus: The former tells us all surface groups are finitely presented groups with a single relation. The latter bears a close relationship with Dehn's theory of tame knots, whose knot groups are finitely presentable via Wirtinger presentations. Indeed, Dehn recognized finitely presented groups as a salient feature of interesting

[1]Lickorish and Wallace did some forty years later.

topological questions and thus called for an investigation of them in their own right in *On Discontinuous Groups*, 1911. There, he posed three purely group-theoretic problems: the word, conjugacy, and isomorphism problems. We are interested in the first:

Problem 0.1. *"An element of a group is given as a product of generators. One is required to give a method whereby it may be decided in a finite number of steps whether this element is the identity or not."* [5]

Certain classes of groups were known to have soluble word problem, e.g. trefoil knot groups, braid groups, and 1-relator groups, among others. Of course, the only criterion for solubility is a concrete algorithm. Turing's work displays strong evidence for what is and is not computable. Post showed the word problem for f.p. semigroups wasn't computable. It is easy to construct f.g. groups whose word problem is not computable, but what about f.p. groups? As it turns out, there are finitely presented groups non-computable word problem. Higman's embedding theorem delivers an easy proof of this fact, as we shall see. Towards Higman's theorem, we develop some basic results about amalgamated free products and HNN extensions assuming a working knowledge of free groups and group presentations. We also describe how to represent a Turing machine with a semigroup, taking for granted some intuition for computability. Some topology is assumed in Section 3, §3.

1. Amalgamated Free Products and HNN Extensions

We begin by reviewing some basic facts about free groups and group presentations. Given a set X, the free group $\langle X\rangle$ on X satisfies the universal property that every function $X \to G$ factors through a unique group homomorphism $\langle X\rangle \to G$.[2] $\langle X\rangle$ and $\langle Y\rangle$ are isomorphic iff $|X| = |Y|$, so the rank of a free-group is well-defined.

The free group $\langle X\rangle$ exists for any set X, given by X-words modulo adjacent inverses under concatenation. That every group is a quotient of a free group is an easy corollary: Factor the identity (set) map id_G through a unique epimorphism $\phi : \langle G\rangle \to G$, so that $G \cong \langle G\rangle \ker\phi$. In general, if there exist sets S of generators and $R \subset S^\omega$ of relations and G is given by the quotient of $\langle S\rangle$ by the normal closure of $\langle R\rangle$ in $\langle S\rangle$, we say G has presentation $\langle S \mid R\rangle$. Abusing notation slightly, we often write $G = \langle S \mid R\rangle$. If $G = \langle S \mid R\rangle$, then G is finitely generated (f.g.) if S is finite and finitely presented (finitely presented) if f.g. and R is finite. Provided a presentation $\langle S' \mid R'\rangle$, we often write $\langle G, S' \mid R'\rangle$ instead of $\langle S, S' \mid R, R'\rangle$.

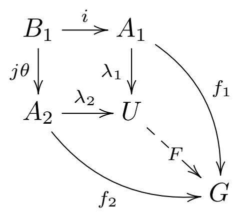
1.1. Amalgamated Free Products.

The free product $A_1 * A_2$ of (disjoint) groups A_1, A_2 satisfies the universal property that for every group G and homomorphisms $f_i : A_i \to G$, $i = 1, 2$, there is a unique homomorphism $\phi : A_1 * A_2 \to G$ through which f_1 and f_2 factor.[3] The free product of two groups always exists: if $A_1 = \langle S \mid R\rangle$ and $A_2 = \langle S' \mid R'\rangle$, then $A_1 * A_2$ has presentation $\langle S, S' \mid R, R'\rangle$. Amalgamated free products are a generalization of free products in which we identify isomorphic subgroups.

[2] $\langle X\rangle$ is unique up to isomorphism, if it exists. So we may speak of "the" free group on a set.

[3] The requirement $A_1 \cap A_2 = \emptyset$ is not a restriction, as we can always take disjoint isomorphic copies.

Definition 1.1. Let A_1, A_2 be groups and θ an isomorphism among subgroups $B_i \leq A_i$, $i = 1, 2$. Let i, j be the inclusion maps $B_i \to A_i$ for $i = 1, 2$, respectively. The **amalgamated free product** $A_1 *_\theta A_2$ along θ is the data of a group U and homomorphisms $\lambda_i : A_i \to U$ satisfying the following universal property: For every group G and homomorphisms $f_i : A_i \to G$ such that $f_1 i = f_2 j\theta$, there is a unique homomorphism $F : U \to G$ making the diagram commute:



Specifying the amalgamated free product by a universal property guarantees uniqueness up to isomorphism, if it exists. This is in fact the case:

Proposition 1.2. *Let A_i, B_i, θ be as above. Let X be the set $\{b\theta(b)^{-1} : b \in B_1\}$. Then:*

(a) $A_1 *_\theta A_2$ exists, given by the quotient of $A_1 * A_2$ by the normal closure of X along with maps $\lambda_i = \nu\mu_i$, where μ_i is the inclusion $A_i \to A_1 * A_2$ and ν is the canonical surjection $A_1 * A_2 \to (A_1 *_\theta A_2)/\langle X \rangle^{A_1 * A_2}$.

(b) If A_i has presentation $\langle S_i \mid R_i \rangle$, then $A_1 *_\theta A_2$ has presentation $\langle S_1, S_2 \mid R_1, R_2, X \rangle$.

The proof of Proposition 2.2a is routine, involving little more than constructing an appropriate homomorphism F and checking that the diagram commutes. 2.2b is an easy consequence. The reader may refer to [8] for details.

The presentation of an amalgamated free product as the quotient of $A_1 * A_2$ by the normal closure of X in $A_1 * A_2$ concretizes the idea of gluing A_1 and A_2 together along isomorphic subgroups $B_1 \cong B_2$. Indeed, under mild hypotheses the fundamental group of $X \cup Y$ is the amalgamated free product of $\pi_1(X)$ and $\pi_1(Y)$ along $\pi_1(X \cap Y)$.

Amalgamated free products admit nice normal form properties, as shown below. Recall a left transversal of a subgroup $K \leq G$ is a subset $T \subseteq G$ consisting of a chosen representative from the left cosets of K in G. [4] Choose left transversals T_i of $B_i \leq A_i$ so that the representative of B_i is the identity, and denote the representative of $a_k B_i$ by $\overline{a}_k$. Then for all $a_k \in G$, there is a unique $b_k \in B_i$ satisfying $a_k = \overline{a}_k b_k$.

Theorem 1.3. *(Normal Form for AFPs) In $A_1 *_\theta A_2$, every element is given by a unique normal form, i.e. a product $\overline{a}_1 \cdots \overline{a}_n b$ satisfying:*

(i) Each $\overline{a}_i$ is a member of T_1 or T_2.
(ii) $\overline{a}_j \in A_1$ iff $\overline{a}_{j+1} \in A_2$.
(iii) $b \in B_1$.

Proof of Existence. Let N the normal closure of X as in Proposition 2.2a. Every coset of N in $A_1 * A_2$ has a representative $p_1 q_1 \cdots p_n q_n$, where $p_i \in A_1$ and $q_i \in A_2$.

[4] If T is a left transversal of K in G, then G is the disjoint union of left cosets tK for $t \in T$, so that every $g \in G$ has a factorization $g = tk$ for a unique $t \in T$ and $k \in K$.

To prove the existence of normal forms, we induct on the length n of these coset representatives.

For $n = 1$, the specifications above furnish a unique $b_1 \in B_1$ for which $p_1 = \overline{a}_1 b_1$. But in $A_1 *_\theta A_2$, the members of B_1 are identified with their respective images under θ. We may thus write

$$p_1 q_1 = \overline{a}_1 b_1 q_1 = \overline{a_1}(\theta(b_1) q_1).$$

But $\theta(b_1)q_1 \in A_2$ gives us another expression $\theta(b_1)q_1 = \overline{a}_2 b_2$ for some unique $b_2 \in B_2$. Therefore $b_2' = \theta^{-1}(b_2)$ lives in B_1. It follows, $p_1 q_1 = \overline{a}_1 \overline{a}_2 b_2'$. This completes the base case.

For induction, assume we have a normal form

$$p_1 q_1 \cdots p_{n-1} q_{n-1} = \overline{a}_1 \cdots \overline{a}_k b.$$

Consider a word of the form $p_1 q_1 \cdots p_n q_n$. By hypothesis, there holds

$$p_1 q_1 \cdots p_n q_n = (\overline{a}_1 \cdots \overline{a}_k b) p_n q_n = (\overline{a}_1 \cdots \overline{a}_k)((b p_n) q_n).$$

But b lives in B_1, so bp_n lives in A_1. Therefore $(bp_n)q_n = (\overline{a}_{k+1} b') y_n$ for some $b' \in B_1$. In view of the identification $b\theta(b)^{-1}$ in $A_1 *_\theta A_2$, this latter term equates to $\overline{a}_{k+1}(\theta(b') y_n)$. Then $\theta(b') y_n = \overline{a}_{k+2} b''$ for some $b'' \in B_2$, so again we end up with an expression $\overline{a}_{k+2} b'' = \overline{a}_{k+2} \theta^{-1}(b'') =: \overline{a}_{k+2} b_{k+2}$. It follows,

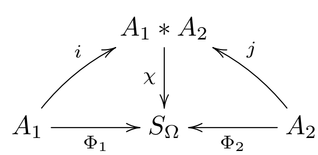
$$p_1 q_1 \cdots p_n q_n = \overline{a}_1 \cdots \overline{a}_{k+2} b_{k+2},$$

as desired. □

Proof of Uniqueness. Let Ω be the set of all normal forms. For each $a \in A_i$, let $\phi_a : \Omega \to \Omega$ be left-multiplication by a. Observe $\phi_1 = \mathrm{id}_\Omega$, $\phi_{a_1} \phi_{a_2} = \phi_{a_1 a_2}$, and $\phi_{a^{-1}} = \phi_a^{-1}$. Therefore the map $\Phi_i : A_i \to S_\Omega$ defined by $a \mapsto \phi_a$ is well-defined and in fact a homomorphism. The universal property of free products furnishes a unique homomorphism

$$\chi : A_1 * A_2 \to S_\Omega : \overline{a}_1 \cdots \overline{a}_n b \mapsto \phi_{\overline{a}_1} \cdots \phi_{\overline{a}_n} \phi_b$$

making the diagram commute:



Notice that for all $b \in B_1$ and normal forms w, there holds

$$\chi(b\theta(b)^{-1})(w) = \phi_{b\theta(b)^{-1}}(w) = w.$$

So $b\theta(b)^{-1}$ lives in $\ker \chi$ for all $b \in B_1$. Now, consider

$$\Psi : A_1 *_\theta A_2 \to S_\Omega : \overline{a}_1 \cdots \overline{a}_n bN \mapsto \chi(\overline{a}_1 \cdots \overline{a}_n b);$$

this is well-defined by the previous sentence. Two normal forms are equal iff they have the same spelling, and so $\Psi(w) \neq \Psi(v)$ in S_Ω when $w \neq v$ in Ω. Uniqueness of normal forms thus follows. □

The corollary below is quite useful and follows from the theorem above. Consult Lemma 1, [3] for an explicit proof:

Corollary 1.4. *Let $U = A_1 *_{A_1 \cap A_2} A_2$. If $x_1 \dots x_n$ live in $U \setminus A_1 \cap A_2$ and there holds $x_i \in A_1$ iff $x_{i+1} \in A_2$, then $x_1 \cdots x_n$ is not a member of $A_1 \cap A_2$. In particular, $x_1 \cdots x_n$ is nontrivial in U.*

1.2. **HNN Extensions.** An amalgamated free product is essentially a free product in which some specified pair of isomorphic subgroups are identified. By contrast, in an HNN extension, we treat isomorphic subgroups as conjugate with respect to a chosen isomorphism. In turn, the HNN extension induces this isomorphism as an inner automorphism among these subgroups. The first part of the theorem below makes sense of this idea, and originally motivated HNN extensions. The second part is a key ingredient of Britton's lemma.

Theorem 1.5. *Let G be a group.*

(a) If H_1, H_2 are isomorphic subgroups of G and $\phi : H_1 \cong H_2$, there is a $K \geq G$ in which ϕ is given by an inner automorphism.

(b) Suppose $\langle S \mid R \rangle$ is a presentation of G. Let P be a set disjoint from S indexed by I and A_j, B_J be S-words injexed by J. Further suppose

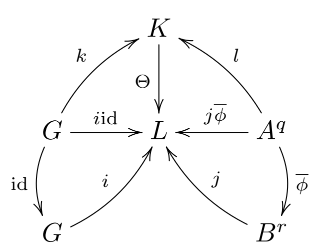
$$\langle S, P \mid R, p_{i,j}^{-1} A_j p_{i,j} = B_j : i, j \in I, j \in J \rangle$$

is a presentation of H. Let $J(i)$ be the set of j for which $p_{i,j}^{-1} A_j p_{i,j} = B_j$ as above. If there exist isomorphisms

$$\phi_i : \langle A_j : j \in J(i) \rangle \to \langle B_j : j \in J(i) \rangle : A_j \mapsto B_j$$

for all i, then $G \leq H$.

Proof. For (a), consider the subgroups K, L generated by $G \cup q^{-1} A q$ and $G \cup r^{-1} B r$ in $G * \langle q \rangle$ and $G * \langle r \rangle$, respectively. Then K, L are clearly free on the displayed generating sets. Define $\overline{\phi} : q^{-1} a q \mapsto r^{-1} \phi(a) r$ and consider the diagram



The universal property of free products gives rise to a unique map Θ making the diagram commute. In particular, $\Theta k = i\mathrm{id}$. Notice the generating sets of K, L have the same cardinality, whence Θ is an isomorphism. Moreover, we have

$$\Theta(q^{-1} a q) = j\overline{\phi}(q^{-1} a q) = r^{-1} \phi(a) r.$$

Put $H = (G * \langle q \rangle) *_{\Theta} (G * \langle r \rangle)$, so that H contains some isomorphic copy X of K as a subgroup by Theorem 2.3. Now, we have for all a in A, $q^{-1} a q = r^{-1} \phi(a) r$ in H by amalgamating along Θ. Setting $p = q r^{-1}$ thus completes the proof.

We can apply part (a) sufficiently many times to G to produce a group H' containing G with generators $S, p_i : i \in I$ such that $p_{i,j}^{-1} A_j p_{i,j} = B_j$ holds for all $i \in I$ and $j \in J(i)$. Therefore, there is an epimorphism $H \to H'$ under which the members of H are mapped to "themselves" in H subject to new relations furnished by the kernel of the map. In particular, any S-word W is the identity in H iff W is the identity in H' (under the homomorphism) iff W is the identity in G (as a subgroup of H'). □

In the discussion that follows, assume G, H satisfy the hypotheses of Theorem 2.5b.

Definition 1.6. H is an HNN extension of G with basis S and stable letters p_i if for all $i \in I$, there is an isomorphism

$$\phi_i : \langle A_j : j \in J(i)\rangle \to \langle B_j : j \in J(i)\rangle : A_j \mapsto B_j.$$

We now prove Britton's lemma, an important combinatorial characterization of trivial words in an HNN extension. Our proof follows Britton's argument in [3], but fills out the details and breaks it up into three stages. In Lemma 1.8, we consider one stable letter and assume $\phi = \mathrm{id}$. Lemma 1.9 expands 1.8 to nontrivial isomorphisms, and Theorem 1.10 expands 1.9 to any number of stable letters. For 1.8 and 1.9, let A, B be the subgroups of G generated by $\{A_j : j \in J\}$ and $\{B_j : j \in J\}$, respectively. Similarly for $A(i)$, $B(i)$ with $J(i)$ instead of J.

Definition 1.7. Let H be an HNN extension of G, $i \in I$, and $e = \pm 1$. We say $p_i^e C p_i^{-e}$ is a **pinch** in H if $e = -1$ implies C is an element of $A(i)$ in G and $e = 1$ implies C is an element of $B(i)$ in G. A word on the generators of H is p_i**-reduced** if it contains no pinches involving p_i.

Lemma 1.8. *Let $H = \langle S, p \mid R, p^{-1}A_j p = A_j : j \in J\rangle$ for some S-words A_j. If W is an S,p-word involving p such that W reduces to the identity, then W contains a pinch $p^e C p^{-e}$.*

Proof. We first show H is isomorphic to an amalgamated free product, and then induct on the number of p-separated S-words needed to describe W. To begin, let Q_1 be an isomorphic copy of A via an isomorphism ϕ, so that Q_1 has a presentation

$$\langle \phi(A_j) : j \in J \mid q_k = 1 : k \in K\rangle$$

for some $\phi(A_j)$-words q_k. Take the direct product $Q_2 = Q_1 \times \langle t\rangle$ for a letter t disjoint from Q_1. Proposition 2.2b shows $Q_3 = G *_\phi Q_2$ has presentation

$$\begin{aligned}\langle S, \phi(A_j) : j \in J, p \mid R, q_k = 1 : k \in K, p^{-1}\phi(A_j)p = \phi(A_j), \phi(A_j) = A_j : j \in J\rangle\\ = \langle S, p \mid R, p^{-1}A_j t = A_j : j \in J, r_k := \phi^{-1}(q_k) = 1 : k \in K\rangle.\end{aligned}$$

The latter presentation follows from the relations $\mathrm{A}_j\phi(A_j)^{-1}$. By construction, we have $q_k = 1$ in Q_3, but $q_k = 1$ always holds in Q_1. It follows, $r_k = \phi^{-1}(q_k) = 1$ in Q_0. Therefore $r_k = 1$ in G, since G contains Q_0. Consider the canonical homomorphism $G \to Q_3$ given by the composition of the inclusion map into the free product and quotient map. Then $r_k = 1$ in G implies $r_k = 1$ in Q_3. We may thus omit the relations r_k from the second presentation of Q_3, whence it becomes apparent that Q_3 and H are isomorphic.

Since $Q_3 \cong H$ and $G \leq Q_3$, G is isomorphic to a subgroup of H. In particular, an S-word w is trivial in G iff trivial in H. Already the generators and relations of H are contained in those of G. Altogether, we have $G \leq H$.

Let W be an unreduced S,p-word involving t. We may assume W has spelling

$$W_0 p^{e_1} W_1 \cdots p^{e_n} W_n$$

for nonzero integers e_j and S-words W_j such that $W_j = \emptyset$ implies $j = 0$ or $j = n$. There is nothing to prove if W contains a subword pp^{-1} or $p^{-1}p$. We induct on n: $n = 1$ is impossible, since $W = 1$ in H implies $W = 1$ in G and $W = 1$ in Q_2. Then $p^{e_1} = W_0^{-1}W_1^{-1}$ in $G \cap Q_2$. By Theorem 2.3, we have $G \cap Q_2 \cong A$. But recall A was generated by S-words A_j, hence prohibiting the membership of p^{e_1}.

Assume the result for $n-1$, and consider a word W spelled $W_0 p^{e_1} W_1 \cdots p^{e_n} W_n$. Since $W = 1$ in H, we have by Corollary 2.4 some j between 0 and n such that $W_j \in A$. We only need to consider the case when the signs of e_j and e_{j+1} agree, for there is nothing to prove when they don't. In this case, W contains as a subword $p^{e_j} W_j p^{e_{j+1}} W_{j+1}$. Because $W_j \in A$, it commutes with p. Hence W contains a subword $p^{e_j + e_{j+1}} W_j W_{j+1}$ in H. Having equated W to a term in H with $n-1$ factors with p and S-words, we have returned to the inductive hypothesis. □

Lemma 1.9. *Let $H = \langle S, p \mid R, p^{-1} A_j = B_j : j \in J\rangle$ for some S-words A_j, B_j. Suppose H is an HNN extension of G with basis S and stable letters p, so that we have an isomorphism $\phi : A \to B : A_j \mapsto B_j$. If W is an S,p-word involving p and W reduces to the identity in H, then W contains a pinch.*

Proof. Consider the following presentations:

$$H_1 = \langle S, q \mid R q^{-1} A_j q = B_j : j \in J\rangle,$$
$$H_2 = \langle S, q, r \mid R, q^{-1} A_j q = B_j, r^{-1} B_j r = B_j : j \in J\rangle,$$
$$H_3 = \langle S, q, r, p \mid R, q^{-1} A_j q = B_j, p = qr, p^{-1} A_j p (= r^{-1} q^{-1} A_j q r) = B_j : j \in J\rangle.$$

Notice $H_1 \cong H$ and $H_2 \cong H_3$. Moreover, we have $G \leq H$ by Theorem 2.5b. Take an S,p-word W as in the hypothesis. As in the proof above, we only need to consider when W is spelled

$$W_0 p^{e_1} W_1 \cdots p^{e_n} W_n.$$

Define V to be the word spelled

$$W_0 (qr)^{e_1} W_1 \cdots (qr)^{e_n} W_n.$$

Since the generating and relating sets of H_3 contain those of H, and since $W = 1$ holds in H, we get $V = 1$ in H_3 in view of the relation $p = qr$. V, H_1, and H_2 satisfy the hypotheses of Lemma 1.8, so that V must contain a pinch $r^{-1} C r$ or $r C r^{-1}$ where $C \in B$ in H_1. Analyzing the spelling of V, we have in the first case, C has spelling $q^{-1} W_i q$ for some $1 \leq i \leq n-1$. So $q^{-1} W_i q$ is in B, whence W_i is in $q B q^{-1}$ and thus in A. In the second case, C already has spelling W_i. □

Theorem 1.10. *(Brittn's Lemma) Let H be an HNN extension of G. If W contains a stable letter and $W = 1$ in H, then W contains a pinch.*

Proof. Suppose H has presentation

$$\langle G, p_i : i \in I \mid p_{i,j}^{-1} A_j p_{i,j} = B_j : j \in J(i)\rangle.$$

Order the elements of I (it is countable) and consider the following:

$$H_0 = \langle G\rangle, \quad H_k = \langle G, p_k \mid p_{k,j}^{-1} A_j p_{k,j} = B_j : j \in J(k)\rangle.$$

Since H is an HNN extension of G, there are isomorphisms $\phi_i : A(i) \to B(i)$ for all i. We thus have H_{k+1} is an HNN extension of H_k for all k with basis

$$S \cup \{p_i : 1 \leq i \leq k-1\}$$

and stable letter p_k, as the isomorphism condition is already fulfilled. We thus obtain a chain $H_k \leq H_{k+1}$ and $H_k \leq H$ for all k by Theorem 2.5b. Define $N = \max\{n \in \mathbb{N} : W \text{ involves } p_n\}$ and proceed by induction on N: Lemma 1.9 covers the case $N = 1$, so assume the result for $N-1$. Given the hypothesis $W = 1$ in H, the chain $H_0 \leq H_1 \leq \cdots$, and the maximality of N, it follows that W is trivial in H_N. By Lemma 1.9, W contains a subword $p_N^{-1}Cp_N$ or $p_NCp_N^{-1}$ where in the former case, C is a member of the subgroup $\langle B_j : j \in J(N)\rangle^{H_{N-1}}$ and in the latter, C is a member of the subgroup $\langle A_j : j \in J(N)\rangle^{H_{N-1}}$. The proof is complete if C is an S-word, for the aforementioned subgroups in H_{N-1} are subgroups in G as well. If not, C involves some $p_1, ..., p_{N-1}$ equating in E_{N-1} to a word D in $B(i)$ (in G). That is, $CD^{-1} = 1$ in E_{N-1}. By hypothesis, CD^{-1} contains a subword of the desired form. Since D does not contain any p_i, we know this subword occurs in C. Then, it also occurs in W. □

2. Turing Machines and Induced Semigroups

Having investigated amalgamated free products and HNN extensions, we shift our attention to the relevant notions from computability theory. Our treatment introduces one (of many) formalisms in view of constructing a semigroup that describes the behavior of a given Turing machine. Albeit self-contained, it omits motivation. For motivation, we refer the reader to [10].

Definition 2.1. Let S be a finite set of letters $s_0 \dots s_m$ and Q be a finite set of states $q_0 \dots q_n$ disjoint from S. s_0 is the blank symbol, and q_1 is the initial state. We denote the set of quadruples $Q\times S\times S\cup\{L, R\}\times Q$ by $\mathcal{Q}$, where $\{L, R\}$ is disjoint from Q and S. We further identify quadruples in $\mathcal{Q}$ with strings. A set $T \subseteq Q$ of quadruples is a deterministic Turing machine if for all $q_i^1 s_j^1 s_k^1 q_l^1$ and $q_i^2 s_j^2 s_k^2 q_l^2$ in $\mathcal{Q}$, $q_i^1 s_j^1 = q_i^2 s_j^2$ implies $s_k^1 q_l^1 = s_k^2 q_l^2$. As far as potentially confusing terminology goes, we call S the **alphabet** of T, Q the **states** of T, and $S \cup Q \cup \{L, R\}$ the **symbols** of T.

A description of a Turing machine T is a string

$$s_{i_1} \cdots s_{i_k} q_j s_{i_{k+1}} \cdots s_{i_l}$$

of letters and exactly one state not occuring at the right end. For descriptions X, Y of T, we have the basic move $X \to Y$ if one of the conditions below holds. Note P, Q are just strings of letters:

$$q_i s_j s_k q_l \in T \quad \wedge \quad (X = Pq_i s_j Q \quad, Y = Pq_l s_k Q);$$

$$q_i s_j R q_l \in T \quad \wedge \quad (X = Pq_i s_j s_t Q, Y = Ps_j q_l s_t Q) \quad \vee \quad (X = Pq_i s_j, Y = Ps_j q_l s_0);$$

$$q_i s_j L q_l \in T \quad \wedge \quad (X = Ps_j q_i s_t Q, Y = Pq_l s_j s_t Q) \quad \vee \quad (X = q_i s_j Q, Y = q_l s_0 s_j Q).$$

Stipulating that T is deterministic prohibits non-deterministic moves, wherein $X \to Y$ and $X \to Z$ but $Y \neq Z$.

Given a string $\sigma \in S^{<\omega}$ whose first character is not blank, the description $q_1\sigma$ is called an input to T. A terminating computation $T(\sigma)$ with input $q_1\sigma$ is a finite collection of descriptions $X_0 \dots X_p$ ordered such that

$$q_1\sigma = X_0 \to \cdots \to X_p$$

and for no $i < p$ does there hold $X_p \to X_i$. We write $T(\sigma)\downarrow$ if there is a terminating computation with input $q_1\sigma$, and $T(\sigma)\uparrow$ otherwise. The halting set $E(T)$ of T is

$$\{\sigma \in S^{<\omega} : T(\sigma)\downarrow\},$$

which T is said to enumerate. q_0 is a stopping state of T if for every $\sigma \in E(T)$, the terminal description of the computation $T(\sigma)$ involves the state q_0. Note we may assume any Turing machine has stopping state q_0 without affecting its halting set.

A subset $E \subset S^*$ is computably enumerable (c.e.) if $E = E(T)$ for some Turing machine T. E is computable if E and E^c are c.e.

Is there a c.e. but not computable set? The affirmative answer to this question underlies the classical theory of computability. Let $S = \{0, 1\}$ and Q be any finite set of states. We can encode natural numbers n in S by consecutive sequences of $n+1$ 1s. Furthermore, observe that the number of Turing machines on any finite alphabet and finite state set is countable. It is further assumed that an effective enumeration $T_0, T_1, \dots$ of all Turing machines on S and Q exists, given by the so-called universal Turing machine U. See [10] or [4] for the reasons behind this.

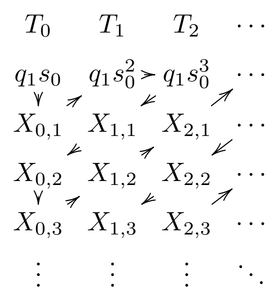
Definition 2.2. The **halting problem** is the set of natural numbers n on which the n-th Turing machine halts:

$$\mathbf{0}' = \{n \in \mathbb{N} : T_n(n)\downarrow\}.$$

Proposition 2.3. $\mathbf{0}'$ *is c.e. but not computable.*

Proof. Assume the complement of $\mathbf{0}'$ were c.e., so that $[\mathbf{0}']^c = E(T_N)$ for some N. But then $N \in \mathbf{0}'$ iff $N \in [\mathbf{0}']^c$: $N \in \mathbf{0}'$ iff $T_N(N)$ halts, iff $N \in [\mathbf{0}']^c$, for $[\mathbf{0}']^c = E(T_N)$.

But $\mathbf{0}'$ is c.e. To each $n \in \mathbb{N}$, there is a (possibly nonterminal) computation $q_1 s_1^{n+1} = X_{0,n} \to \cdots \to X_{k,n} \to \dots$ in T_n. Beginning with input $q_0 s_1^{n+1}$, consider the Turing machine T corresponding to the following computation:



If $X_{k,n} \to X_{k+1,n}$ is defined, continue along the path; otherwise, stop. If $n \in \mathbf{0}'$, then there are only finitely many descriptions $\{X_{k,n} : 1 \leq k \leq K(n)\}$. By construction, $T(q_1 s_1^{n+1})$ will eventually reach $X_{K(n),n}$, at which point it halts. So $n \in E(T)$, as desired. ☐

Our formulation (2.1) helps us easily construct a semigroup that simulates a given Turing machine T. Indeed, the quadruples in T correspond to the relations we wish to impose on the semigroup freely generated by alphabet and states of T. Additional symbols h and q are included to represent the ends of the tape and a sort of terminal state, respectively. The details of our construction allow us to establish Lemma 2.5, which asserts that the halting set $E(T)$ corresponds to an algebraic condition in the induced semigroup.

First, a notational remark: Fix a generating set X and consider the semigroup Γ generated by X subject to some relations $A_j = B_j$ for some X-words A_j, B_j indexed by $j \in J$. For X-words W, V and P, Q, we write $W \to V$ if W and V are spelled PA_jQ and PB_jQ for some j, or vice-versa. Observe W and V are equal in Γ iff there are finitely many $U_1 \dots U_n$ such that

$$W \equiv U_1 \to \cdots \to U_n \equiv V.$$

Definition 2.4. Let T be a Turing machine with stopping state q_0 and symbols $\{L, R, q_0 \ldots q_N, s_0 \ldots s_M\}$. Its induced semigroup $\Gamma(T)$ is defined by the presentation

Generators $G(T)$	Relations $R(T)$
q, h	$q_i s_j = q_l s_k$ if $q_i s_j s_k q_l \in T$
$q_0, ..., q_N$	$q_i s_j s_b = s_j q_l s_b$ if $q_i s_j R q_l \in T$
$s_0, ..., s_M$	$q_i s_j h = s_j q_l s_0 h$ if $q_i s_j R q_l \in T$
	$s_b q_i s_j = q_l s_b s_j$ if $q_i s_j L q_l \in T$
	$h q_i s_j = h q_l s_0 s_j$ if $q_i s_j L q_l \in T$
	$q_0 s_b = q_0, s_b q_0 h = q_0 h, h q_0 h = q$

Lemma 2.5. *Let T be a Turing machine with stopping state q_0, W be the set of positive words on the alphabet S of T, and $E = E(T)$ be the halting set of T. For any $w \in W$, we have $w \in E$ iff $hq_1wh = q$ in $\Gamma(T)$.*

Proof. Suppose $w \in W$ and $hq_1wh = q$ in $\Gamma(T)$. By the note above, there are elementary moves $U_0, ..., U_n$ in $\Gamma(T)$ so that there holds

$$hq_1wh \equiv U_0 \to \cdots \to U_n \equiv hq_0h \to q.$$

Notice that if U, V are words in $\Gamma(T)$ with $U \to V$, and neither U nor V is spelled q, then we have $U \equiv hXh$ iff $V \equiv hYh$ for some descriptions X, Y of T; for the only operation in $\Gamma(T)$ killing h is $hq_0h = q$, thus the observation follows from the hypothesis $U, V \not\equiv q$. In particular, each U_k is of the form hX_kh for a description X_k. Let m be the first k such that U_k involves q_0; we proceed by induction on m. Clearly $m \geq 1$, since X_0 involves q_1 and no other q_i. Observe that if we have $U \equiv hXh$ for a description X, $V \not\equiv q$, and $U \to V$ in $\Gamma(T)$ describes one of the first five relations, then $V \equiv hYh$ for a description Y such that $X \to Y$ or $Y \to X$ in T. Thus for all i, we have $X_i \to X_{i+1}$ or $X_{i+1} \to X_i$. But X_n involves stopping state q_0, hence $X_{n-1} \to X_n$. For $m = 1$, we recover $q_1 w \equiv X_0 \to X_1$ as a computation in T, verifying $T(w)\downarrow$. Assume the result for all $m < M$ for $M > 1$. If we have

$$X_0 \to X_1 \to \cdots \to X_{M-1} \to X_M,$$

we are done. Otherwise, there exists k such that $X_k \to X_{k-1}$ and $X_k \to X_{k+1}$. But since T is deterministic, this implies $X_{k-1} \equiv X_{k+1}$, hence reducing to the case $M - 1$. Relabelling the descriptions $X_1, ..., \hat{X}_k, ...X_n$ to $X_0, ..., X_{M-1}$, the computation $T(w)$ must halt by the inductive hypothesis.

The forward implication is easier. If $w \in E$, there is a computation

$$T(w) \equiv X_0 \to \cdots \to X_n.$$

Notice how the first five relations in $\Gamma(T)$ arise from the basic moves of the machine outlined in Definition 3.1. Repeated application gives us

$$hq_1wh (= hX_0h) = hX_1h = \cdots = hX_nh.$$

Since T has stopping state q_0, X_n must involve q_0—that is, there holds $X_n \equiv Fq_0G$ for S-words F, G. But for all $s_k \in S$, we have $q_0 s_k = q_0$ and so there holds in $\Gamma(T)$

$$hX_nh = hFq_0Gh = hFq_0h.$$

Similarly, the relations $s_k q_0 h$ give us in $\Gamma(T)$

$$hFq_0h = hq_0h = q.$$

That is, $hq_1wh = q$ holds in $\Gamma(T)$, as desired. □

3. Higman's Embedding Theorem

Here, we prove the main result of this paper. Our proof does not follow Higman's 1961 paper, in which he developed his theory of benign subgroups to reach the result. Instead, we follow Aanderaa [1], which enlists the techniques developed in the last two sections.

How can we characterize the f.g. subgroups of an f.p. group computability-theoretically? Higman's theorem shows they are exactly the recursively presented subgroups:

Definition 3.1. An f.g. group is **recursively presented** if it has a presentation $\langle s_1 \dots s_M \mid w = 1 : w \in R\rangle$ where R is a c.e. set of positive words on the generators. It is **recursively presentable** if it has some recursive presentation.

Requiring that R consist of positive words on the generators is not a restrictin. Given a group presentation $\langle S \mid R\rangle$, we can adjoin new generators t_s to S and relations $t_s s = 1$ whenever an instance of s^{-1} occurs in R, for some $s \in S$. Evidently, this new presentation determines the same group.

Theorem 3.2. *(Higman, 1961) An f.g. group G is recursively presentable iff it embeds into some f.p. group.*

In what follows, we assume the following notational conventions:

Notation 3.3. Suppose $\langle u_1 \dots u_m \mid w = 1 : w \in R\rangle$ is a recursive presentation of a recursively presentable group G. By definition, some Turing machine T, with alphabet $\{s_0 \dots s_M\} \supseteq \{u_1 \dots u_m\}$, states $\{q_0 \dots q_N\}$, and stopping state q_0, enumerates R.[5] For convenience, we often abbreviate the relating set of the induced semigoup $\Gamma(T)$ by expressions $F_i q_{i_1} G_i = H_i q_{i_2} K_i$ for some positive words on the alphabet $\{s_1 \dots s_M, h\}$ and $q_{i_1}, q_{i_2} \in \{q, q_0 \dots q_N\}$ indexed by a set I. For a word $X \equiv x_{j_1} \dots x_{j_n}$, we write $\overline{X}$ for the word $x_{j_1}^{-1} \dots x_{j_n}^{-1}$.

[5] We switch notation from our usual set $\{s_1 \dots s_M\}$ of generators so that we don't clash with notation in the induced semigroup $\Gamma(T)$.

Construction 3.4.

$$
\begin{aligned}
G =&\langle u_1, ..., u_m \mid w = 1 : w \in R\rangle,\\
Q_0 =&\langle x\rangle,\\
Q_1 =&\langle Q_0, h, s_k : 1 \leq k \leq M \mid xs_k = s_k x^2; xh = hx^2\rangle \qquad ,\\
Q_1' =&Q_1 * \langle q, q_1, ..., q_N\rangle,\\
Q_2 =&\langle Q_1, r_i : i \in I, q, q_l : 1 \leq l \leq N \mid r_i s_k = s_k x r_i x; r_i h = hxr_i x; r_i^{-1}\overline{F}_i q_{i_1} G_i r_i = \overline{H}_i q_{i_2} K_i\rangle,\\
Q_3 =&\langle Q_2, t \mid tr_i = r_i t, tx = xt\rangle,\\
Q_3' =&\langle Q_2, t_0 \mid t_0^{-1}(q_1^{-1}hr_ih^{-1}q_1)t_0 = q_1^{-1}hr_ih^{-1}q_1; t_0^{-1}(q_1^{-1}hxh^{-1}q_1)t_0 = q_1^{-1}hxh^{-1}q_1\rangle,\\
Q_* =&\langle Q_3, k \mid kr_i = r_i k; kx = xk; k(q^{-1}tq) = (q^{-1}tq)k\rangle,\\
Q_*' =&\langle Q_3', k_0 \mid k_0^{-1}(hr_ih^{-1})k_0 = hr_ih^{-1}, k_0^{-1}(hxh^{-1})k_0 = hxh^{-1};\\
&\qquad k_0^{-1}(hq^{-1}h^{-1}q_1t_0q_1^{-1}hgh^{-1})k_0 = hq^{-1}h^{-1}q_1t_0q_1^{-1}hqh^{-1}\rangle,\\
Q_4 =&Q_*' * G,\\
Q_5 =&\langle Q_4, b_s : 1 \leq s \leq m \mid b_s^{-1}u_jb_i = u_j; b_s^{-1}a_jb_s = a_j; b_s^{-1}k_0b_s = k_0u_s^{-1} : 1 \leq s, j \leq m\rangle,\\
Q_6 =&\langle Q_5, d \mid d^{-1}k_0d = k_0, d^{-1}a_sb_sd = a_s : 1 \leq s \leq m\rangle,\\
Q_7 =&\langle Q_6, e \mid e^{-1}t_0e = t_0d; e^{-1}k_0e = k_0, e^{-1}a_se = a_s : 1 \leq s \leq m\rangle.
\end{aligned}
$$

Lemma 3.5. *Q_1 is an HNN extension of Q_0. Q_2 is an HNN extension of Q_1'. Q_3 and Q_3' are isomorphic and are HNN extensions of Q_2. Q_* and Q_*' are isomorphic and are HNN extensions of Q_3 and Q_3', respectively. Q_5 is an HNN extension of Q_4, Q_6 is an HNN extension of Q_5, and Q_7 is an HNN extension of Q_6.*

Proof. See [1] and [8]. Albeit arid, these proofs aren't all trivial. In particular, showing Q_7 is an HNN extension of Q_6 isn't easy. □

Lemma 3.6. *Suppose H is an HNN extension of G with stable letters $\{r_i : i \in I\}$, and suppose W, V are r_i-reduced words for all i, of the form $W_0 r_i^{e_1} W_1 \cdots r_i^{e_m} W_m$ and $V_0 r_i^{f_1} \cdots r_i^{f_n} V_n$ where W_j, V_j do not involve any stable letters, $1 \leq j \leq m, n$. If $W = V$ in H, then $m = n$, $e_j = f_j$ for all j, and $r_i^{e_m} W_m V_n^{-1} r_i^{-f_n}$ is a pinch.*

Proof. We induct on $\max\{m, n\}$. By assumption and Britton's lemma, WV^{-1} contains a pinch. But W and V are r_i-reduced for all i, whence the pinch must occur at the interfact $r_i^{e_m} W_m V_n^{-1} r_i^{-f_n}$, as claimed. Therefore $e_m = f_n$, because fm and $-f_n$ must differ in sign. Now, clearly the lemma must hold for $\max\{m, n\} = 1$. Assume the result for $\max\{m, n\} = N - 1$ and consider when $\max\{m, n\} = N$. By definition of a pinch, if $e_m = f_n = -1$, then Britton's lemma guarantees $W_m V_n^{-1} = A_1^{g_1} \cdots A_k^{g_m}$ where $A_j \in A(r_i)$ and $g_j = \pm 1$ for all $1 \leq j \leq k$. Then our original pinch $r_i^{-1} W_m V_n^{-1} r_i$ equates in H to

$$r_i^{-1} A_1^{g_1} \cdots A_k^{g_m} r_i = r_i^{-1} A_1^{g_1} r_i r_i^{-1} \cdots r_i^{-1} A_k^{g_m} r_i,$$

which in turn equates in H to $B_1^{g_1} \cdots B_k^{g_k}$ for some $B_j \in B(r_i)$. Having eliminated one r_i from the pinch $r_i^{-1} W_m V_n^{-1} r_i$, we've reduced to the inductive hypothesis. A similar argument works for the case $e_m = f_n = 1$. □

Lemma 3.7. *Let w be a positive word on $\{s_0 \dots s_M\}$, and let $\tau = h^{-1} q_1 wh$ and $\sigma = hq_1wh$. Then $\sigma = q$ in $\Gamma(T)$ iff $\tau^{-1} t \tau$ commutes with k in Q_*.*

Proof of Necessity in Lemma 3.7. Let $\Lambda := k^{-1}(\tau^{-1}t\tau)k(\tau^{-1}t^{-1}\tau)$, so that $\Lambda = 1$ in Q_* and involves a stable letter of Q_* by assumption. It follows by Lemma 1.8, Λ contains a pinch $k^{-1}Ck$ where C is a word on some r_i, x, and $q^{-1}tq$. Inspecting the form of Λ, we observe $C = \tau^{-1}t\tau$ in Q_3, whence $\tau^{-1}t\tau = R_n^{-1}(q^{-1}tq)^{e_n}\cdots R_1^{-1}(q^{-1}tq)^{e_1}R_0$ for some x, r_i-words R_j and $e_j = \pm 1$. Let n be such that the preceding equation contains the minimal number of terms, and let M be the word with spelling

$$\tau^{-1}t\tau R_0(q^{-1}t^{e_1})R_1(q^{-1}t^{e_2}q)R_2\cdots(q^{-1}t^{e_n}q)R_n$$

so that we have $M = 1$ in Q_3. By Lemma 1.8 again, M contains a pinch tCt^{-1} or $t^{-1}Ct$ where $C = R_*$ in Q_2, for some word R_* on r_i and x. Assume the initial t or t^{-1} in tCt^{-1} or $t^{-1}Ct$, respectively, occurs in t^{e_j} for some $j \geq 1$. Since C doesn't involve t, we have $t^eCt^{-e} \equiv t^{e_j}qR_jq^{-1}t^{e_{j+1}}$. In Q_3, the equation

$$q^{-1}t^{e_j}qR_jq^{-1}t^{e_{j+1}}q \equiv q^{-1}t^eCt^{-e}q = q^{-1}t^eRt^{-e}q = q^{-1}Rq = q^{-1}(qR_jq^{-1})q = R_j.$$

follows in view of the commutation relations $tx = xt$ and $tr_i = r_it$. We have thus contradicted the minimality of n, forbidding this case. If, in fact, the initial t or t^{-1} occurs in the first possibel place in W, then we have $tCt^{-1} \equiv t\tau R_0q^{-1}t^{e_1}$. so that $R^{-1}\tau R_0 = q$ in Q_2. (R, R_0 are words on r_i and x.) Expanding and rearranging, we have $R^{-1}h^{-1}q_1 = qR_0^{-1}h^{-1}w^{-1}$ in Q_2. This final equation still holds when R and R_0 are freely reduced, wherefore $R^{-1}h^{-1}q_1$ and $qR_0^{-1}h^{-1}w^{-1}$ are r_i-reduced for all i, since w is a positive word on $\{s_0, \ldots, s_M\}$ and R, R_0 are freely reduced x, r_i-words.

To complete this proof, we need to show $hq_1wh = q$ in $\Gamma(T)$ provided $R^{-1}h^{-1}q_1$ and $qR_0^{-1}h^{-1}w^{-1}$ are r_i-reduced for all i. We induct on the number N of instances of r_i's in R^{-1}, which coincides with the number of instances of r_i's in R_0^{-1} by Lemma 3.6. The case $N = 0$ is not possible[6]; the base case $N = 1$ follows a similar argument to the inductive step, so we omit a complete description here. Assume the result for all $T < N$, and assume there are N r_i's that occur in R^{-1} and R_0. By Lemma 3.6, we may write $R^{-1}\tau R_0 \equiv R_1[r_i^ex^m\tau x^nr_i^{-e}]R_2 = q$ in Q_2, where there holds $R^{-1} = R_1r_i^e$ and $R_0 = r_i^{-e}R_2$. Observe the bracketed term is a pinch. By Britton's lemma, the term $x^m\tau x^n$ lies in $A(i) = \langle \overline{F}_iq_{i_1}G_i, s_1x_1, \ldots, s_Mx\rangle^{Q_1'}$ when $e = -1$ and $B(i) = \langle \overline{H}_iq_{i_2}K_i, s_1x^{-1}, \ldots, s_Mx^{-1}\rangle$ when $e = 1$. We will only treat the former case, the latter being similar.

For the case $e = -1$, we observe $j = i_1$ and $q_j = q_{i_1}$. Because $x^m\tau x^n$ is a member of $A(i)$, we have

$$\Lambda' :\equiv x^m\tau x^nW_0(\overline{F}_iq_{i_1}G_i)^{f_1}W_1\cdots(\overline{F}_iq_{i_1}G_i)^{f_n}W_n = 1$$

in Q_1' where $f_j = \pm 1$ and W_j are words on the free generators $\{s_1x, \ldots, s_Mx\}$. Hence, we may assume the W_j's are freely reduced without affecting the equality just stated. Assume Λ' is splled with the minimal number of terms n. Since Q_1' is an HNN extension of Q_1 and Λ' involves a stable letter, Britton's lemma guarantees a pinch in Λ'. If this pinch occurs at the first occurence of q_{i_1}, then $f_1 = -1$ and $whx^nW_0G_i^{-1} = 1$ in Q_1. If it occurs anywhere else, say the j-th q_{i_1}, it follows,

[6] If we have $N = 0$, then it follows $q = R^{-1}\tau R_0 = x^m\tau x^n$ in Q_2 for some integers m, n. Since $Q_1' \leq Q_2$ holds and no r_i's occur by hypothesis, $x^m\tau x^n = q$ in Q_1'. This equation holds in the free product Q_1' iff $q_1 = q$ and $x^mh^{-1}whx^n = 1$ in Q_1, implying $m = n = 0$ and in turn implying τ is empty, a contradiction.

$u_j = 1$, a contradiction lest we violate the minimality of n. Thus the equations $f_1 = -1$ and

$$M \equiv x^m h^{-1} q_1 w h x^n W_0 G_i^{-1} q_{i_1}^{-1} \overline{F}_i^{-1} W_1 = 1$$

in Q_1' follow. We already showed $whx^n W_0 G_i^{-1} = 1$ in Q_1. But Q_1' is a free product, whence $\overline{F}_i^{-1} W_1 = 1$ in Q_1. Conjugating, we conclude $x^n W_0 G_i^{-1} wh = 1 = h^{-1} \overline{F}_i^{-1} u_1 x^m$ in Q_1.

The words W_j are, without a loss of generality, freely reduced. In particular, they contain no subwords of the frm $s_k s_k^{-1}$ or $s_k^{-1} s_k$ for $1 \leq k \leq M$. Cancelling any remaining subwords of these forms does not affect any of the equalities above. Following this reduction, the first surviving letter in $G_i^{-1} wh$ is positive. As G_i is a positive word on $\{s_0 \dots s_m\}$, G_i^{-1} necessarily vanishes upon free reduction. If not, $G_i^{-1} wh$ begins with s_k^{-1} for some k. But then $x^n W_0 G_i^{-1} wh$ (which is trivial in Q_1) involves s_k, so that Britton's lemma guarantees a pinch $s_k^e C s_k^{-e}$. Since W_0 is freely reduced, this pinch cannot occur within a subword of $x^n W_0$. Thus, the last letter of the pinch is the first surviving letter of G_i^{-1}. If, as we assumed, this letter is not positive, then we have $e = 1$ and $s_k^e C s_k^{-e} \equiv s_k x s_k^{-1}$ since C is an x-word. But Britton's lemma again guarantees $x \in \langle x^2 \rangle$, a contradiction. By a similar argument, $\overline{F}_i^{-1}$ vanishes in $h^{-1} \overline{F}_i^{-1}$.

Define $V_0 :\equiv r_i^{-1} W_0^{-1} r_i$. Then, V_0^{-1} is a word on $s_1 x^{-1}, \dots, s_M x^{-1}$, in view of the relations $r_i^{-1} s_k x r_i = s_k x^{-1}$ for all $1 \leq k \leq M$. Let ψ be the automorphism of Q_1 defined by $x \mapsto x^{-1}, s_k \mapsto s_k^{-1}$, so that $\psi : W_0^{-1} \mapsto V_0^{-1}$. Then, we have $V_0^{-1} = X_0 x^{-n}$ in Q_1, where X_0 is a positive word on $s_0 \dots s_m$ with spelling $G_i^{-1} wh$. Letting $V_1 :\equiv r_i^{-1} W_1 r_i$, we have by a similar argument $V_1^{-1} = x^{-m} X_1$, where X_1 is a negative word on $\{s_0 \dots s_m\}$ with spelling $h^{-1} \overline{F}_i^{-1}$. This leads us to the equation $q = R_1 x^{-m} X_1 \overline{H}_i q_{i_2} K_i X_0 x^{-m} R_2$ in Q_2. But $R_1 x^{-m}$ and $x^{-n} R_2$ are x, r_i-reduced words exhibiting at most $N-1$ occurrences of various r_i. One can easily check $X_1 \overline{H}_i$ and $K_i Y_i$ are freely reduced. By the inductive hypothesis, $\overline{X}_1 H_i$ and $K_i X_0$ are positive words satisfying $h^{-1} \equiv \overline{X}_1 F_1$ and $wh \equiv K_i X_0$. For the same reason, $X_1 H_i q_{i_2} K_i X_0 = q$ holds in $\Gamma(T)$. It follows, $hq_1 wh \overline{X}_1 F_i q_i G_i X_1 = X_0 H_i q_{i_2} K_i Y_1$ in $\Gamma(T)$. So we've finally obtained the desired equation $hq_1 wh = q$ in $\Gamma(T)$. □

Proof of Sufficiency in Lemma 3.7. If $\sigma = q$ in $\Gamma(T)$, there are basic moves

$$\sigma \equiv U_0 \to \cdots \to U_n \equiv hq_0 h \to q$$

for some words $U_0 \dots U_n$ in $\Gamma(T)$ of the form $U_k \equiv X F_i q_{i_1} G_i Y$, $U_{k+1} \equiv X H_i q_{i_2} K_i Y$, where X, Y are positive S-words. In Q_*, the relations inherited from Q_2 deliver

$$\overline{X}(\overline{H}_i q_{i_2} K_i) Y = \overline{X}(r_i^{-1} \overline{F}_i q_{i_1} G_i r_i) Y = L \overline{X}(\overline{F}_i q_{i_1} G_i) Y R$$

for some words L, R on various r_i and x. Now, we have $U_k = U_{k+1}$ in $\Gamma(T)$, whence $U_k' = U_{k+1}'$ in Q_* where $U_k' = \overline{X}(\overline{F}_i q_{i_1} G_i) Y$ and $U_{k+1}' = \overline{X}(\overline{H}_i q_{i_2} K_i) Y$ as above. For each k, reassign the specific instances of L's and R's above to L_k and R_k accordingly, Now, let $L = L_1 \cdots L_{n-1}$ and $R = R_{n-1} \cdots R_1$. In Q_*, we have $U_1' = L U_n' R$. The equation $\tau = LqR$ in Q_* now readily follows from the correspondence $U_1 = \sigma$ iff $U_1' = \tau$ and $U_n' = q$. Inspecting the relations in Q_* inherited from Q_3 and those involving k, we observe t and k commut with x and r_i (for all i), whence t and k commute with L and R. Therefore the following holds in Q_*:

$$\begin{aligned}(\tau^{-1}t\tau)k =& R^{-1}q^{-1}(L^{-1}tL)q(Rk) = R^{-1}((q^{-1}tq)k)R \\ =& (R^{-1}k)q^{-1}tqR = kR^{-1}q^{-1}L^{-1}tLqR = k(\tau^{-1}t\tau),\end{aligned}$$

as desired. □

Proof of Necessity in Theorem 3.2. Identify G with its image under the embedding into a finitely presented group H. If H is f.p. then, *a fortiori*, H is f.g. and recursively presented. That is, we have a presentation $\langle S \mid R\rangle$ for a finite set S and a c.e. set R of positive S-words. By definition, $H \cong \langle S\rangle/\langle R\rangle^{\langle S\rangle}$. Since R is c.e. and $\langle R\rangle^{\langle S\rangle}$ consists of conjugates of elements of R, $\langle R\rangle^{\langle S\rangle}$ is also c.e. If $G \leq H$, then $G \cong K/(K \cap \langle R\rangle^{\langle S\rangle})$, where K is an f.g. subgroup of $\langle S\rangle$. K must be c.e., whence $K \cap \langle R\rangle^{\langle S\rangle}$ must be c.e. Indeed, a finite intersection of c.e. sets is c.e. Given $\langle K \cap \langle R\rangle^{\langle S\rangle K}\rangle = K \cap \langle R\rangle^{\langle S\rangle}$, we have a presentation $\langle K \mid K \cap \langle R\rangle^{\langle S\rangle}\rangle$ of G. If $K \cap \langle R\rangle^{\langle S\rangle}$ involves any negative letters, produce an equivalent presentation omitting them. In any event, G exhibits a recursive presentation. □

Proof of Sufficiency in Theorem 3.2. In Lemma 3.7, we showed $\sigma = q$ in $\Gamma(T)$ iff $(\tau^{-1}t\tau)k = k(\tau^{-1}t\tau)$ in Q_*, where $= hq_1wh$ and $\tau = h^{-1}q_1wh$. In Lemma 2.5, we showed $\sigma = q$ in $\Gamma(T)$ iff $w \in R$ (since T enumerates R). Recall that the relations $t_0 = q_1^{-1}hth^{-1}q_1$ and $k_0 = hkh^{-1}$ hold in Q_3' and Q_*', respectively. Also, recall that we have isomorphisms $Q_3' \cong Q_3$ and $Q_*' \cong Q_*$. If w is a positive word on $s_0 \dots s_M$, then $w \in R$ iff in Q_*',

$$(*) \quad k_0(w^{-1}t_0w) = (w^{-1}t_0w)k_0.$$

If w is a word on $a_1 \dots a_m$, let $w^{[b]}$ and $w^{[u]}$ be the words obtained by replacing instances of a_i with b_i and u_i, respectively.

By Theorem 2.5b and Lemma 4.5, we have

$$Q_0 \leq Q_1 \hookrightarrow Q_1' \leq Q_2 \leq Q_3' \leq Q_* \hookrightarrow Q_4 \leq Q_5 \leq Q_6 \leq Q_7.$$

Recall $\langle S \mid R\rangle \leq \langle S' \mid R'\rangle$, where $S \subset S'$ and $R \subset R'$ iff the homomorphism sending an S-word to "itself" in S' is an embedding. Now, G embeds into $Q_4 := Q_* * G$ via the canonical maps, whence G embeds into Q_7 by iterated applications of our remarks above. Showing Q_7 exhibits a finite presentation thus completes the proof of sufficiency for Higman's theorem.

To show Q_7 exhibits a finite presentation, consider Q_7', which omits relations of the form $w^{[u]} = 1$ for $w \in R$. It suffices to demonstrate that these omitted relations may be derived from the remaining relations in Q_7', of which there are only finitely many. To that end, take any $w \in R$. Then:

(0) $k_0^{-1}w^{-1}t_0wk_0k_0 = w^{-1}t_0w$.

(1) $e^{-1}(k_0^{-1}w^{-1}t_0wk_0)e = e^{-1}w^{-1}t_0we$.

(2) $k_0^{-1}w^{-1}e^{-1}t_0ewk_0 = w^{-1}e^{-1}t_0ew$.

(3) $k_0^{-1}w^{-1}t_0dwk_0 = w^{-1}t_0dw$.

(4) $(k_0w^{-1}t_0wk_0)k_0^{-1}w^{-1}dwk_0 = (w^{-1}t_0w)w^{-1}dw$.

(5) $k_0^{-1}w^{-1}dwk_0 = w^{-1}dw$.

(6) $dwd^{-1} = ww^{[b]}$.

(7) $w^{-1}dw = w^{-1}dd^{-1}ww^{[b]}d = w^{[b]}d$.

(8) $k_0^{-1}w^{[b]}dk_0 = w^{[b]}d$.

(9) $k_0^{-1}w^{[b]}k_0 = w^{[b]}, \;\; k_0^{-1}w^{[b]}k_0 = w^{[b]}w^{[u]}$.

(10)$w^{[u]} = 1$.

(0) follows from (∗), and (1) follows immediately. Observing that e commutes with k_0 and all a_i in the presentation of Q_7, we derive (2). Then (3) follows from the relation $e^{-1}t_0e = t_0d$ in Q_7. In (4) we insert the identity $wk_0k_0^{-1}w^{-1}$ on the left and ww^{-} on the right, and then rearrange. By rearranging, (5) readily follows by (∗). The relations $d^{-1}a_ib_id = a_i$ and $a_ib_j = b_ja_i$ give rise to (6): Since w is a word $a_{i_1}\cdots a_{i_n}$ on a_i in Q_7, we have

$$da_{i_1}\cdots a_{i_n}d^{-1} = dd^{-1}a_{i_1}b_{i_1}d\cdots d^{-1}a_{i_n}b_{i_n}dd^{-1} = a_{i_1}b_{i_1}\cdots a_{i_n}b_{i_n} = ww^{[b]}.$$

Then (7) readily follows from (6), and (8) follows by combining (5) and (7). The first part of (9) expresses the commutativity relation on k_0 and d in Q_6. The relations $k_0b_ik_0 = b_iu_i$ and $b_iu_j = u_jb_i$ derive the second part, since if $w^{[b]} = b_{i_1}\cdots b_{i_n}$ and $w^{[u]} = u_{i_1}\cdots u_{i_n}$, then:

$$k_0^{-1}b_{i_1}\cdots b_{i_n}k_0 = k_0^{-1}b_{i_1}k_0k_0^{-1}b_{i_2}k_0\cdots k_0^{-1}b_{i_k}k_0 = b_{i_1}u_{i_1}\cdots b_{i_n}u_{i_n} = w^{[b]}w^{[u]}.$$

Combining the two parts of (9) proves $w^{[b]}w^{[u]} = w^{[b]}$, from which the desired result (10) follows. This completes the proof. □

Higman's theorem holds for groups, but what about for other algebraic objects?

Question 3.8. *(Boone, Lawvere) How general is Higman's theorem? Which algebraic categories exhibit an analog of Higman's theorem? Does such an analog exist for every single-sorted algebraic theory?*

4. Decision Problems

We conclude the present paper with some applications to decision problems. Our first application is the Novikov–Boone theorem, which readily follows from Higman's theorem. With the existence of an f.p. group with unsolvable word problem secured by Novikov–Boone, we may prove a much more general undecidability result known as the Adian–Rabin theorem. A special case of this theorem shows the homeomorphism problem for manifolds of dimension > 3 is not computable, bringing us full-circle.

4.1. **The Word Problem.** Baumslag described Novikov and Boone's independent proofs of Novikov–Boone as a "combinatorial tour-de-force" [2]. In contrast to Novikov and Boone's dense combinatorial arguments, a strikingly simple proof follows immediately from Higman's theorem:

Definition 4.1. Let G be an f.g. group with generators $\{x_1 \dots x_n\}$, and let W be the set of positive words on these generators. The **word problem** for G is decidable$if the set${w $\in W : w =_G 1\}$ is computable.

Theorem 4.2. *There exists an f.p. group with undecidable word problem.*

Proof. Let $\langle a, b\rangle$ and $\langle c, d\rangle$ be disjoint, isomorphic copies of the free group of rank 2, and let $G = \langle a, b\rangle *_\phi \langle c, d\rangle$ where ϕ is an isomoprhism $\langle a^{-n}ba^n : n \in \mathbf{0}'\rangle \to \langle c^{-n}dc^n : n \in \mathbf{0}'\rangle$ defined on the generators by $a^{-n}ba^n \mapsto c^{-n}dc^n$ for all $n \in \mathbf{0}'$. Then G exhibits a presentation $\langle a, b, c, d \mid a^{-n}ba^nc^{-n}d^{-1}c^n : n \in \mathbf{0}'\rangle$. Inpsecting this presentation, we note that for any word $w \equiv a^{-n}ba^nc^{-n}d^{-1}c^n$ with $n < \omega$, we have $w = 1$ in G iff $n \in \mathbf{0}'$. Therefore, G must have undecidable word problem, lest we compute $\mathbf{0}'$.

But $\mathbf{0}'$ is c.e., whence the presentation above is recursive. By Higman's theorem, there is an f.p. group H into which G embeds. Thus, H is an f.p. group with undecidable word problem. □

4.2. **Markov Properties and a Universal F.P. Group.** The existence of an f.p. group with undecidable word problem lays the foundation for a more general undecidability result. This theorem, due to S. Adian and M.O. Rabin, roughly asserts that Markov properties are not decidable. As it turns out, most reasonable properties of a group are Markov. Example 4.4 lists some of them, but many more exist.

Once we establish Adian–Rabin, we introduce the notion of a universal f.p. group, i.e. one into which every f.p. group embeds.

Definition 4.3. Let P be a property of a group. P is a **Markov property** if P is preserved under isomorphism, there exists an f.p. group satisfying P, and there exists an f.p. group not embeddable into any f.p. group satisfying P.

Example 4.4. The following are Markov properties:

(1) *Trivial:* Only the trivial group embeds into the trivial group.
(2) *Finite:* $\mathbb{Z}$ does not embed into any finite group.
(3) *F.G. and Free:* If a group embeds into a free group, its image is free (Nielsen-Schreier). Moreover, Markov properties are preserved under isomorphism. Thus any non-free, f.p. group like $\mathbb{Z}^2$ will do.
(4) *F.G. with Decidable Word Problem:* Finite groups, free groups, and residually finite groups are all f.p. groups with solvable word problem. Novikov–Boone furnishes an f.p. group with unsolvable word problem, and clearly no such group embeds in a group with solvable word problem.

Theorem 4.5. *(Adian–Rabin) Fix a countable alphabet S. For any Markov property P, the set of f.p. groups generated by S satisfying P is not computable.*

Proof. By definition, there exist f.p. groups A satisfying P and B not embeddable into any f.p. group satisfying P. Let C be an f.p. group with undecidable word problem. Then $C * B$ has undecidable word problem, and exhibits a finite presentation $\langle x_1 \dots x_n \mid R\rangle$. Let w be a word on the generators of C, and consider the following group presentations:

$$\begin{aligned} D &= (C * B) * \langle p \rangle = \langle q_0 := p, q_1 := px_1, ...q_n := px_n \mid R' \rangle, \\ E &= \langle D, r_0, ..., r_n \mid r_i q_j^2 r_i^{-1} = q_j : 1 \leq i, j \leq jn \rangle, \\ F &= \langle E, s \mid s r_i s^{-1} = r_i^2 : 1 \leq i \leq n \rangle, \\ G &= \langle a, b \mid bab^{-1} = a^2 \rangle, \\ H &= \langle G, c \mid cbc^{-1} = b^2 \rangle, \\ J &= \langle E * H \mid sa^{-1}, wq_0 w^{-1} q_0^{-1} c^{-1} \rangle, \\ K &= J * A. \end{aligned}$$

R' swaps instances of x_i with $p^{-1}q_i$. E is an HNN extension of D with stable letters r_i. The subgroups of E generated by $r_0 \dots r_n$ and $r_0^2 \dots r_n^2$ are freely generated (contrapose Theorem 1.10), whence isomorphism. Therefore F is an HNN extension of E with stable letter s. H and G are HNN extensions of G and F, respectively. If the word w is trivial in C, then the relations in J shows us all its generators are trivial. That is, J is trivial. So $K = A$, whence K satisfies P. If w is not trivial in C, then a, c and $s, wq_0 w^{-1} q_0^{-1}$ freely generate the subgroups they determine in H and F, since $wq_0 w^{-1} q_0^{-1}$ has infinite order in E and c is stable. Therefore, $\langle a, c \rangle^H$ and $\langle s, wq_0 q^{-1} q_0^{-1} \rangle$ are isomorphic via $\rho : a \mapsto s, c \mapsto wq_0 w^{-1} q_0^{-1}$. As suspected, there is an isomorphism $J \cong F *_\rho H$ after all. So F embeds into J, and J embeds into K. But B embeds into F by construction, hence into J. Therefore, J cannot satisfy P by definition of B. For the same reason, neither can K.

We've thus shown $w = 1$ in C iff K satisfies P. Since C has undecidable word problem, we cannot decide whether K satisfies P. □

Corollary 4.6. *The isomorphism problem is not computable.*

Proof. Triviality is Markov. □

Theorem 4.7. *There is a universal f.p. group, i.e. one containing an isomorphic copy of every f.p. group.*

Proof. Up to isomorphism, there are countably many f.p. groups. Then their free product A has countably many generators $a_0, a_1, ..., a_n, ...$ where $a_0 = 1$. Let $B = A * \langle x, y \rangle$, and consider the subgroups $C = \langle x, a_n y^{-n} x y^n : n \in \mathbb{N} \rangle^H$ and $D = \langle y, x^{-n} y x^n : n \in \mathbb{N} \rangle^H$. Observe that D is a free subgroup of $\langle x, y \rangle$, hence of B. It follows, C is a free subgroup of B and there is an isomorphism $C \cong D$ via $\phi : x \mapsto y$ and $a_n y^{-n} x y^n \mapsto x^{-n} y x^n$ for all $n \geq 1$. Theorem 2.5a furnishes a group E containing B and some new letter t in E such that t induces ϕ by an inner automorphism, that is $\phi(c) = t^{-1} c t$ for all $c \in C$. Let F be the subgroup of E generated by x, t. We claim F contains A as a subgroup. For all $n \geq 1$, notice $t^{-1}(a_n y^{-n} x y^n) t = \phi(a_n y^{-n} x y^n) = x^{-n} y x^n$, so that $a_n = t x^{-n} y x^n y^{-n} x^{-1} y^n$. Since $y = \phi(x) = t^{-1} x t$, it follows that $a_n \in F$ for all $n \geq 1$. (There is nothing to check for a_0.) We have thus shown our free product A of every f.p. group embeds into a group F with 2 generators. It is a tedious but relatively straightforward exercise to show F does, in fact, have a recursive presentation. By Theorem 4.2, F embeds into an f.p. group U. Now, given any f.p. group G, we have $G \hookrightarrow A \hookrightarrow F \hookrightarrow U$. □

Corollary 4.8. *An f.p. group is not universal iff it satisfies a Markov property.*

Proof. Fix a Markov property P, so that there exists an f.p. group B not embeddable into any f.p. group satisfying P. If G is a universal f.p. group, then B embeds into G. Thus G satisfies no Markov property, lest we contradict our first assertion. Of course, not being universal holds up to isomorphism, there exist non-universal f.p. groups, and the universal f.p. group cannot embed into any non-universal one. That is, not being universal is Markov. □

4.3. **The Homeomorphism Problem Revisited.** Here, we return to what originally motivated these group-theoretic decision problems and consider the general problem of classifying manifolds up to homeomorphism. For $n \geq 4$, is the set of all manifolds M with $M \approx N$ computable for all closed, connected n-manifolds N?

Lemma 4.9. *For all $n \geq 4$, and for every f.p. group H, there is a closed, connected n-manifold M whose fundamental group is isomorphic to H.*

Proof. See [6]. □

Theorem 4.10. *For all $n \geq 4$, the set of simply connected n-manifolds is not computable. Thus the homeomorphism problem is not decidable for $n \geq 4$.*

Proof. The first claim follows immediately from Adian–Rabin, Lemma 4.9, and the fact that triviality is Markov. The second follows from the first, and independently from Corollary 4.8 and Lemma 4.9. □

Acknowledgements

Thanks to Isabella Scott for guidance throughout the program and perceptive feedback on my paper. Thanks to Shmuel Weinberger to clarify some questions and provide me further directions of study. Thanks to Denis Hirshfeldt for his reverse math lectures. Finally, thanks to Peter May, whose efforts make this wonderful program possible.

References

[1] Stal Aanderaa. A Proof of Higman's Embedding Theorem, Using Britton Extensions of Groups. North-Holland Publishing Company. 1973.
[2] Gilbert Baumslag. Topics in Combinatorial Group Theory. Springer Basel AG. 1993.
[3] John L. Britton. The Word Problem. Annals of Mathematics, Second Series, Vol. 77. 1963.
[4] S. Barry Cooper. Computability Theory. Chapman and Hall/CRC. 2004.
[5] M. Dehn. Uber unendliche diskontinuierliche gruppen. Annals of Mathematics, First Series, Vol. 71. 1911.
[6] Mathematical Notes. "Amalgamated Products and HNN Extensions." Link.
[7] David Peifer. Max Dehn and the Origins of Topology and Infinite Group Theory. The Mathematical Association of America. 2015.
[8] Joseph Rotman. An Introduction to the Theory of Groups. Springer-Verlag New York, Inc. 1995.
[9] John Stillwell. The Word Problem and the Isomorphism Problem for Groups. Bulletin of the American Mathematical Society, Volume 6, Number 1. 1982.
[10] A. M. Turing. On Computable Numbers, with an Application to the *Entscheidungsproblem*. Proceedings of the London Mathematical Society. 1936.
[11] Shmuel Weinberger. Computers, Rigidity, and Moduli. Princeton University Press. 2005.