TURING MACHINES TO WORD PROBLEMS

CHARLES F. MILLER III

Abstract. We trace the emergence of unsolvable problems in algebra and topology from the unsolvable halting problem for Turing machines.

§1. Introduction. Mathematicians have always been interested in being able to calculate with or about the things they study. For instance early developers of number theory and the calculus apparently did extensive calculations. By the early 1900s a number of problems were introduced asking for general algorithms to do certain calculations. In particular the tenth problem on Hilbert’s influential list asked for an algorithm to determine whether an integer polynomial in several variables has an integer solution.

The introduction by Poincaré of the fundamental group as an invariant of a topological space which can often be finitely described by generators and relations led to Dehn’s formulation of the word and isomorphism problem for groups. To make use of such group invariants we naturally want to calculate them and determine their properties. It turns out many of these problems do not have algorithmic solutions and we will trace the history and some of the ideas involved in showing these natural mathematical problems are unsolvable.

In the 1930s several definitions of computable functions emerged together with the formulation of the Church–Turing Thesis that these definitions captured intuitive notions of computability. Church and independently Turing showed that there is no algorithm to determine which formulas of first-order logic are valid, that is, the Entscheidungsproblem is unsolvable.

So it became meaningful to consider whether more standard mathematical questions have algorithmic solutions, and it seems Church encouraged several mathematicians including Emil Post to work on such problems. Indeed Post had also described an equivalent notion of a computing machine, though not in detail, and had anticipated many of these results.

In this article we describe how the unsolvable halting problem for Turing machines can be used to show the unsolvability of several important algorithmic problems in mathematics including the word and isomorphism problems for finitely presented groups and the homeomorphism problem for higher dimensional manifolds. Our intention is to provide some descriptions of what
was done following the historical trail and to make some additional comments about related results and about subsequent developments. Usually we give more modern versions of the constructions. We have included somewhat sketchy but essentially complete proofs for the unsolvability of the word and isomorphism problems based ultimately on Turing machines. But for many other results we give only a few comments on how they are proved. Full treatments of these advances are nowadays available in textbooks and we provide references to those sources for further details (see the last section below). Hopefully those readers with less technical interest will gloss over some of the detail and still gain an impression of how these results can be established.

Here is an outline of some milestones on the historical trail to these results.

– 1908 Tietze describes transformations of presentations of groups which realise all isomorphisms and formulates the isomorphism problem.
– 1911 Dehn poses the word, conjugacy and isomorphism problems for finitely presented groups.
– 1914 Thue poses what is essentially the word problem for finitely presented semigroups.
– 1930s Definitions of computable functions emerge and existence of unsolvable problems is established (Gödel, Church, Kleene, Turing, Post and others).
– 1947 Post and Markov independently construct finitely presented semigroups with unsolvable word problem. Post’s construction is built on Turing machines while Markov’s uses Post’s normal systems.
– 1950 Turing shows the word problem for cancellation semigroups is unsolvable by furthering Post’s construction.
– 1951 Markov shows the isomorphism problem and recognising most properties of semigroups are unsolvable.
– 1955 Novikov and independently 1957 Boone show the word problem for groups is unsolvable. Novikov uses Turing’s cancellation semigroup result while Boone uses Post’s semigroup construction. In 1958 Britton gives another proof and later, in 1963, he greatly simplifies Boone’s proof.
– 1956 Adian and independently 1958 Rabin show the isomorphism problem and recognising most properties of groups are unsolvable.
– 1958 Markov shows the homeomorphism problem for $n$-manifolds is unsolvable ($n \geq 4$).
– 1961 Higman characterises the subgroups of finitely presented groups thus connecting recursion theory with group theory in an unexpected way and giving a very different proof of the unsolvability of the word problem.

The mathematical logic “school” in Princeton during this era was actively involved in these developments. Early in his career Post had a post-doctoral fellowship at Princeton and worked on problems in logic. Alonzo Church
encouraged Post to work on these problems, and he was the doctoral supervisor of Turing, Boone, Rabin, Scott, and Collins (actually advised by Boone) who all contributed to the developments we describe. Also Gödel gave encouragement to Boone on several occasions.

It sometimes happens in mathematics that the same result is obtained quite independently by different researchers, particularly when answering a widely known problem. Our story has several instances of such independent discoveries. In my view it is desirable to give credit to all involved.

It has been my pleasure to know as friends at least four of the principals in our story: Bill Boone (who was my doctoral supervisor), John Britton, Graham Higman, and Sergei Adian. Likewise I will mention work of several other friends and collaborators. So my account is no doubt influenced by these various relationships.

§2. Presentations of Groups and Semigroups. A familiar example of a group is the dihedral group $D_3$ of order 6 consisting of the rigid motions of an equilateral triangle onto itself. It can be described by giving two generators which correspond to rotation through $120^\circ$ and flipping about a central axis. These operations have orders 3 and 2 respectively and the group $D_3$ is described by the presentation

$$D_3 = \langle a, b \mid a^2 = 1, b^3 = 1, a^{-1}ba = b^{-1} \rangle.$$ 

Here the symbols $a$ and $b$ are called generators and the equations they are subjected to are called defining relations.

Groups are often described as symmetry groups or permutation groups and defined by actions on some object. A more general and abstract approach emerged describing groups by generators and defining relations thought of as quotients of a free groups. So if $\{a_1, a_2, \ldots\}$ is a set of symbols representing generators of a group $G$, then a presentation of $G$ on these generators will have the form

$$G = \langle a_1, a_2, \cdots \mid u_1 = v_1, u_2 = v_2, \ldots \rangle.$$ 

Here the $u_i$ and $v_i$ are words on the generators $a_j$, that is, they are finite strings of the symbols $a_j$ and $a_j^{-1}$. Any such piece of notation describes a group as follows. Let $F$ be the free group with free basis $\{a_1, a_2, \ldots\}$ and let $N$ be the smallest normal subgroup of $F$ containing all the words $u_i v_i^{-1}$. The group presented by the above notation is then the quotient group $G = F/N$. Here, as usual, the inverse $v^{-1}$ of the word $v = v_1^\epsilon_1 \cdots v_k^\epsilon_k$ is the word $v^{-1} = v_k^{-\epsilon_k} \cdots v_1^{-\epsilon_1}$ where $\epsilon_j = \pm 1$. Note that by letting $r_i = u_i v_i^{-1}$ we can also describe $G$ as

$$G = \langle a_1, a_2, \cdots \mid r_1 = 1, r_2 = 1, \ldots \rangle$$

which is clearly equivalent and sometimes more convenient.
A presentation of a group $G$ as above is said to be finitely generated if the set of symbols $\{a_1, \ldots, a_n\}$ is finite. finitely related if the set of defining relations $\{u_1 = v_1, \ldots, u_m = v_m\}$ is finite and finitely presented if both are finite.

Given a finite presentation of the form $G = \langle a_1, \ldots, a_n | r_1 = 1, \ldots, r_m = 1 \rangle$ there is a standard recipe for constructing a finite 2-complex $Y$ with fundamental group $G \cong \pi_1(Y, o)$ as follows. Start with a single 0-cell $o$ and for each of the $n$ generators $a_i$ attach an oriented 1-cell by identifying its endpoints with $o$. This gives a 1-complex $Y^1$ with fundamental group which is free on the 1-cells corresponding to the $a_i$. Now for each relation of the form $r_i = 1$ we attach a 2-cell to $Y^1$ by subdividing the boundary of the 2-cell into 1-cells labeled according to the word $r_i$ and attaching (respecting orientation) the 2-cell to $Y^1$. The resulting 2-complex $Y$ is connected, has a single 0-cell, has $n$ 1-cells corresponding to the generators of the presentation for $G$ and has 2-cells corresponding to the defining relations in the presentation for $G$. Then one shows $G \cong \pi_1(Y, o)$ as desired. Conversely one can show that the fundamental group of any finite cell or simplicial complex is finitely presented. One can extend these observations to show that finitely presented groups are also the fundamental groups of closed, differentiable $n$-manifolds ($n \geq 4$). So groups with a finite presentation are of interest for geometric reasons and correspond to spaces with a certain geometric finiteness.

In 1908 Heinrich Tietze, in studying fundamental groups of manifolds and their topological invariance, introduced four operations on presentations which are now known as Tietze transformations. They can be described as: (1) adding a relation which is a consequence of the others; (2) removing a redundant relation; (3) defining a new symbol as a word in existing symbols; and (4) removing a symbol which can be defined in terms of the remaining symbols. Tietze then showed that if the groups described by two presentations are isomorphic, then there is a sequence of Tietze transformations leading from one of these presentations to the other. Actually this is a result of a very general nature applicable to algebraic systems describable by presentations.

Tietze noted that the problem of finding a method to determine whether two groups given by their presentations are isomorphic was unresolved. Thus he raised the algorithmic question known as the isomorphism problem for finitely presented groups.

Soon thereafter, in 1911, Max Dehn highlighted three basic algorithmic problems concerning groups defined by presentations which have played a central role:

word problem: Let $G$ be a group given by a finite presentation. Does there exist an algorithm to determine of an arbitrary word $w$ in the generators of $G$ whether or not $w =_G 1$?

conjugacy problem: Let $G$ be a group given by a finite presentation. Does there exist an algorithm to determine of an arbitrary pair of words $u$ and $v$ in the generators of $G$ whether or not $u$ and $v$ define conjugate elements of $G$?
isomorphism problem: Does there exist an algorithm to determine of an arbitrary pair of finite presentations whether or not the groups they present are isomorphic?

In geometric terms, the word problem for the fundamental group $\pi_1(Y, o)$ corresponds to the problem of determining whether or not a closed loop at $o$ in $Y$ is contractible. The conjugacy problem for $\pi_1(Y, o)$ corresponds to the problem of determining whether or not two closed loops are freely homotopic (intuitively whether one can be deformed into the other). Since homeomorphic spaces have isomorphic fundamental groups, the isomorphism problem is related to the homeomorphism problem, that is, determining whether given spaces are homeomorphic.

For the fundamental groups of closed 2-manifolds Dehn gave algorithms to solve each of these problems, though in general they turn out to be unsolvable. We will comment further on “Dehn’s algorithm” later when we consider the word problem in more detail.

Here are some observations about Dehn’s problems. If $u, v$ are any two words in the generators of a group $G$, then $u =_G v$ if and only if $uv^{-1} =_G 1$ so the problem of deciding equality of words is equivalent to the problem of deciding whether another word is equal to 1. If $w$ is a word of $G$ then $w =_G 1$ if and only if $w$ and 1 are conjugate in $G$, so the word problem for $G$ is always reducible to the conjugacy problem for $G$. 

Recall that if $F$ is a free group with free basis $\{a_1, \ldots, a_n\}$ then every element of $F$ is equal to a unique freely reduced word obtained by successively removing (deleting) inverse pairs, that is, removing subwords of the form $a_i a_i^{-1}$ or $a_i^{-1} a_i$.

If $G = F/N$ is given by a presentation as above, then the elements of $N$ are those words of $F$ which are freely equal to products of conjugates of the relators $r_j$ and their inverses. In particular one can systematically enumerate all such words, and so $\{w \in F \mid w =_G 1\}$ is a recursively enumerable set of words.

Suppose that $G = \langle a_1, \ldots, a_n \mid r_1 = 1, \ldots, r_m = 1 \rangle$ and $H = \langle b_1, \ldots, b_k \mid q_1 = 1, \ldots, q_p = 1 \rangle$ are two finite presentations, and suppose that $u_1, \ldots, u_k$ are words in the generators $a_i$ of $G$. Now the function $\phi$ defined by $b_i \mapsto u_i$ for $i = 1, \ldots, k$ can be formally extended to all words of $H$ of by setting

$$\phi(b_1^{e_1} \cdots b_k^{e_k}) = u_1^{e_1} \cdots u_k^{e_k}.$$ 

This $\phi$ defines a homomorphism from $H$ into $G$ if and only if each $\phi(q_j) =_G 1$ for $j = 1, \ldots, p$.

Now further suppose that the map $\phi$ above defines a monomorphism from $H$ into $G$ so that $\phi$ embeds $H$ in $G$. If $z$ is any word in the given generators of $H$, since we are given the images $\phi(b_i)$ of the generators, we can write down its image $\phi(z)$ as a word in the generators of $G$. Since we are assuming $\phi$ is a monomorphism we know $z =_H 1$ if and only if $\phi(z) =_G 1$. Since
the words in the generators of $G$ that are equal to 1 are recursively enumerable. It follows that $\{ z \in \{ a \} \mid z = H 1 \}$ is also recursively enumerable. Moreover, if we have an algorithm to solve the word problem for $G$ we can use it determine whether or not $z = H 1$. Note that these observations only require that $H$ be finitely generated (not necessarily finitely presented). These observations enable us to conclude the following:

**Proposition 1.** Let $G = \langle a_1, \ldots, a_n \mid r_1 = 1, \ldots, r_m = 1 \rangle$ be a finitely presented group given by generators and relations, and suppose $\{ u_1, \ldots, u_k \}$ is a finite set of words on the generators of $G$. Let $H$ denote the subgroup generated by $\{ u_1, \ldots, u_k \}$. Then:

1. $H$ can be presented on finitely many generators and a recursively enumerable set of defining relations.
2. If $G$ has a solvable word problem, then $H$ also has a solvable word problem. More generally the word problem for $H$ is reducible to the word problem for $G$.
3. The solvability of the word problem does not depend on the particular finite set of generators or finite presentation of the group $G$.

Sometimes it is necessary to distinguish carefully between a presentation which is a piece of notation and the group it describes as an algebraic object. So if $P$ is a presentation, the group it presents is written $gp(P)$. So in the above discussion about the presentation for $G$ we technically have $G = gp(P) = F/N$. One fact we need to have is the following.

**Proposition 2.** The set of pairs of presentations $(P_1, P_2)$ of groups, say on some fixed alphabet, such that $gp(P_1)$ and $gp(P_2)$ are isomorphic (as abstract groups) is recursively enumerable.

Here in outline are two ways to prove this using enumeration. (1) According to Tietze's Theorem if the groups presented are isomorphic there is a sequence to Tietze transformations leading from one presentation to the other. But the Tietze transformations applicable to a presentation are recursively enumerable. Hence by an elaborate enumeration we find a route from one of our presentations to the other. (2) Since the true equations among words in a finite presentation are always recursively enumerable, we can enumerate all pairs of homomorphisms $\phi : P_1 \to P_2$ and $\psi : P_2 \to P_1$. As such pairs are enumerated, compute the values of $\phi \circ \psi$ and $\psi \circ \phi$ on each of the respective generating sets. Then search by enumeration for those homomorphism which are inverse pairs and so fix the respective generating sets.

**Presentations of semigroups and string rewriting.** We recall that a semigroup with 1, also called a monoid, is a set $S$ with an associative multiplication $\cdot$ and a distinguished element denoted 1 so that $u \cdot 1 = u = 1 \cdot u$ for all $u \in S$. Suppose now that $A = \{ a_1, a_2, \ldots \}$ is a set of symbols and let $A^*$ denote the set of all finite strings of symbols from $A$, also called words on $A$. We allow the empty
word which is denoted by $1$. We define the product of two words $u, v \in A^*$ by concatenation, so $u \cdot v = uv$, which is clearly an associative multiplication. The resulting system which we denote by $\Phi(A)$ is the free semigroup with $1$ or free monoid with basis $A$ and it enjoys the usual universal mapping properties of free objects.

As in the case of groups one can impose relations on semigroup elements to obtain presentations. Thus a finitely presented semigroup $S$ is a semigroup described by a notation such as $S = sgp\langle a_1, \ldots, a_n \mid u_1 = v_1, \ldots, u_m = v_m \rangle$ where we have added the $sgp$ to distinguish from the group case. It is important to always keep in mind that there are in general no inverses for elements of a semigroup nor inverse symbols for the generating symbols. The semigroup presented by this notation is the quotient of the free semigroup $\Phi(A)$ by the smallest congruence determined by the defining relations \{ $u_1 = v_1, \ldots, u_m = v_m$ \}.

A finitely presented semigroup such as $S$ can also be described in terms of string rewriting. A rewrite rule is a pair of words $(u, v) \in A^* \times A^*$ which can be applied to strings in $A^*$ of the form $z_1 uz_2$ where $z_1, z_2 \in A^*$. The result of applying the rule $(u, v)$ to this string is $z_1 uvz_2$. The rule $(u, v)$ is usually written as $u \to v$ with the understanding that it enables one to replace any substring $u$ in a word by the substring $v$. A system consisting of $A^*$ together with a finite set $R$ of rewrite rules on $A^*$ is called a string rewriting system or semi-Thue system after Axel Thue who studied them systematically in 1914 [96]. If the rewrite rules $R$ are also symmetric, that is, if $v \to u$ is a rule whenever $u \to v$ is a rule, the semi-Thue system is sometimes called a Thue system and the rewrite rules are written as $u \leftrightarrow v$. One observes that a Thue system is exactly the same as a finitely presented semigroup as discussed above, but the emphasis is on string substitution rather than abstract algebra. Thue [96] raised the word problem for finitely presented semigroups which in the string rewriting context asks for an algorithm to determine whether, given two words $y$ and $z$ of a semigroup $S$ presented as above, there is a sequence of substring replacements of the form $u_i \leftrightarrow v_i$ leading from $y$ to $z$.

Below we will need some terminology for the direction of rewrite rules when a semigroup is given by a presentation $S = sgp\langle a_1, \ldots, a_n \mid u_1 = v_1, \ldots, u_m = v_m \rangle$. A string rewrite of the form $z_1 uz_2 \to z_1 vz_2$ is called a forward application of the defining relation $u_1 = v_1$, so the left-hand side is replaced by the right-hand side of the relation. Similarly a rewrite of the form $z_1 v_1 z_2 \to z_1 u_1 z_2$ is called a backward application where the right-hand side is replaced by the left-hand side. A proof in $S$ is a sequence of rewrites of the form $w_1 \to w_2 \to \cdots \to w_n$ where each $w_i \to w_{i+1}$ is a single application of a defining relation, so is either a forward or backward application (rewrite). Of course two words are equal in $S$ if and only if there is a proof beginning with one and ending with the other. A proof is said to be reversal-free or
without reversals if it does not contain two successive rewrites of the form
\[ z_1 u_i z_2 \rightarrow z_1 v_i z_2 \] or \[ z_1 u_i z_2 \rightarrow z_1 u_i z_2 \rightarrow z_1 v_i z_2, \] the second of which undoes the first.

Results analogous to most of those above for groups also apply to the semigroup case. Note that every finitely presented group can be viewed as a finitely presented semigroup by adding new symbols \( a_i^{-1} \) and rules \( a_i^{-1}a_i = 1 = a_i a_i^{-1} \). So a solution to the word problem for all semigroups would include a solution to the word problem for all groups. But as we shall describe there are semigroups and groups (more difficult to prove) for which such a solution is not possible.

§3. The word problem for semigroups and Post's construction. In 1947 Emil Post [83] and A. A. Markov [61] independently constructed finitely presented semigroups with unsolvable word problem. This was perhaps the first long standing mathematical algorithmic problem shown to be recursively unsolvable. Post's construction was based rather directly on a Turing machine with unsolvable halting problem while Markov's was based on Post's earlier unsolvability results on “canonical systems”.

Markov’s father was the mathematician A. A. Markov famous for his work in stochastic processes after whom Markov chains are named. The son, whose name was the same, made a number of important contributions to mathematical logic, in particular “unsolving” the word problem for semigroups, the isomorphism problem for semigroups and the homeomorphism problem for 4-manifolds.

The role of Emil Post in mathematical logic seems underappreciated. For much of his adult life he suffered recurring attacks of manic-depressive disorder (bipolar disorder) and many of his ideas were not published in sufficient detail or a timely way. It seems fair to say that he anticipated many of the important results of Gödel, Church and Turing. The appreciative article [92] on Post’s work by Stillwell and Martin Davis’s introduction to Post’s collected works [40] are recommended for trying to put Post’s work into perspective. In particular, in a short 1936 paper [79] Post describes a theory of computation which is easily seen to be equivalent to Turing machines, work done independent of Turing’s work. Also his classic 1944 paper [81] is the origin of the structure theory of recursively enumerable sets and degrees of unsolvability.

We are going to describe Post’s construction of a finitely presented semigroup with unsolvable word problem in some detail. The starting point is a Turing machine, or rather a slightly adapted form of a Turing machine whose tape is always finite but expandable. The instantaneous descriptions of the Turing machine correspond to certain words of the semigroup and basic moves of the machine to string rewrites in the semigroup. Since we have later
applications to the word problem for groups, we also do some tidying up after a successful computation.

Part of our intention here is to make clear the parametric nature of this and some later constructions and that the objects can be written down explicitly in terms of the underlying Turing machine. Much of this is available in textbooks, for instance in various editions of Rotman’s text [87].

**Turing machine refresher.** We now describe Post’s construction [83] of a finitely presented semigroup \((T)\) from a Turing machine \(T\) in some detail. The aim is to show that computations in the Turing machine correspond exactly to rewriting sequences for certain special words in the semigroup. The word problem for the semigroup is then reduced to the halting problem for the Turing machine, and so can be unsolvable.

For us a word is a finite string of symbols (all positive in the absence of groups). We want an informal model of a Turing machine computation in which the “tape” is potentially infinite and blank outside some finite segment, yet we want the “blank” to be a symbol in our alphabet. So we think of our tape as finite but expandable in the sense that when we reach either end, we add a blank to that end and proceed as usual. Here is the basic set up (which may be a variation on one the reader is familiar with) and the necessary background on the halting problem we require.

A (deterministic) Turing machine \(T\) consists of:

1. a finite set of symbols \(s_0, s_1, \ldots, s_M\), called the alphabet of \(T\), including a distinguished symbol \(s_0\) called blank;
2. a finite set of internal states \(q_0, q_1, \ldots, q_N\) with two distinguished states \(q_1\), called the start state, and \(q_0\), called the halting state; and
3. a non-empty finite set of quadruples, which are 4-tuples of one of the following three types: \(q_isjs_kq_m\) or \(q_isjRq_m\) or \(q_isjLq_m\), such that no two quadruples have the same first two letters.

An instantaneous description \(\alpha\) is a positive word on the letters \(s_j\) and exactly one \(q_i\) which is not at the right end. For example, \(s_2s_0s_5s_2\) is an instantaneous description which is to be interpreted as “the symbols on the tape are \(s_2s_0s_5s_2\) (with blanks elsewhere) and the machine is in state \(q_3\) scanning \(s_5\) (the symbol to the right of \(q_3\)).

If \(T\) is a Turing machine and \(\alpha, \beta\) are two instantaneous descriptions, then a basic move \(\alpha \to \beta\) of \(T\) is a string replacement of one of the following forms:

1. \(\alpha = Pq_isjQ \to Pq_ms_ks_0Q = \beta\) where \(q_isjks_0q_m \in T\);
2. \(\alpha = Pq_isjQ \to Pq-js_ks_0Q = \beta\) where \(q_isjRq_m \in T\);
3. \(\alpha = Pq_isjQ \to Pq_js_ks_0 = \beta\) where \(q_isjLq_m \in T\);
4. \(\alpha = Pq_isjQ \to Pq_ms_ks_0Q = \beta\) where \(q_isjLq_m \in T\);
5. \(\alpha = q_isjQ \to q_ms_0s_0Q = \beta\) where \(q_isjLq_m \in T\);

where \(P\) and \(Q\) are (possibly empty) words on the \(s_j\) letters.
The interpretation of the basic move of type (1) is “when in state \( q_i \) and scanning \( s_j \), replace \( s_j \) with \( s_k \) and enter state \( q_m \)”. The interpretation of type (2) is “when in state \( q_i \) and scanning \( s_j \), and \( s_k \) is the next symbol to the right, move right one space to scan that \( s_k \) and enter state \( q_m \)”. Similarly for types (4) and (5) with left in place of right.

An instantaneous description is terminal if there is no \( \alpha \to \beta \) with \( \alpha \to \beta \), that is, no further basic moves can be applied. A computation of a Turing machine is a finite sequence of basic moves

\[ \alpha_1 \to \alpha_2 \to \cdots \to \alpha_t \]

where \( \alpha_t \) is terminal.

Our requirement that \( T \) is deterministic, that is, no two quadruples have the same first two letters, implies that at most one basic move is applicable to any instantaneous description. So if \( \alpha \to \beta \) and \( \alpha \to \delta \) then \( \beta = \delta \).

If \( z \) is any word on the \( s_i \), we say that \( T(z) \) exists if there is a computation of \( T \) beginning with \( q_1z \) (recall that \( q_1 \) is our machines “start state”). That is, when started in state \( q_1 \) with tape containing exactly the word \( z \), the machine eventually halts in the sense that the computation reaches a terminal state. The algorithmic problem of determining whether \( T(z) \) exists is called the halting problem for \( T \).

We will assume henceforth that our Turing machines all have unique stopping state \( q_0 \), that is, an instantaneous description \( \alpha \) is terminal if and only if \( \alpha \) involves \( q_0 \). There is no loss in this assumption since, by renumbering and adding appropriate quadruples, one can easily modify any machine \( T \) to obtain a new machine \( T' \) on the same alphabet with unique stopping state \( q_0 \) such that \( T(z) \) exists if and only if \( T'(z) \) exists.

Some fundamental facts about Turing machines are summarised in the following (which may be theorems or definitions depending on your point of view):

**Theorem 3** (Turing, Post). Let \( Z \) be the set of all words on the alphabet \( \{s_1, \ldots, s_M\} \). Then:

1. A subset \( E \subseteq Z \) is recursively enumerable if and only if there is a Turing machine \( T \) with alphabet \( \{s_0, s_1, \ldots, s_M\} \) such that \( E = \{z \in Z \mid T(z) \text{ exists}\} \).
2. There is a Turing machine \( T \) such that \( E = \{z \in Z \mid T(z) \text{ exists}\} \) is not recursive.

Thus, assuming the Church–Turing thesis, there is a Turing machine \( T \) such that there is no algorithm to determine of an arbitrary \( z \in Z \) whether or not \( T(z) \) exists.
**Post’s semigroup construction.** We now describe Post’s construction [83] of a sort of faithful interpretation of the Turing machine $T$ inside a finitely presented semigroup. Some modifications to Post’s original construction are included for later application to the word problem for groups.

Let $\gamma(T)$ be the semigroup with presentation

$$\gamma(T) = \langle h, s_0, s_1, \ldots, s_M, q, q_0, q_1, \ldots, q_N \mid R(T) \rangle$$

where the relations $R(T)$ are

$$q_i s_j = q_l s_k \text{ if } q_i s_j s_k q_l \in T,$$

and for all $b = 0, 1, \ldots, M$:

$$q_i s_j s_b = s_j q_l s_b \text{ if } q_i s_j R q_l \in T,$$

$$q_i s_j h = s_j q_l s_0 h \text{ if } q_i s_j R q_l \in T,$$

$$s_b q_i s_j = q_l s_b s_j \text{ if } q_i s_j L q_l \in T,$$

$$h q_i s_j = h q_l s_0 s_j \text{ if } q_i s_j L q_l \in T,$$

$$q_0 s_b = q_0$$

$$s_b q_0 h = q_0 h$$

$$h q_0 h = q.$$

Intuitively the symbol $h$ marks the ends of the tape the Turing machine $T$ is reading. A word of $\gamma(T)$ is $h$-special if it has the form $huq_j v$ where $u$ and $v$ are $s$-words and $q_j \in \{q, q_0, q_1, \ldots, q_N\}$. The $h$-special words correspond exactly to the instantaneous descriptions of the Turing machine $T$ with $h$ symbols added at each end.

It is important to observe that the first five types of relations of $\gamma(T)$ that correspond to quadruples of $T$, when applied in the forward direction, exactly carry out a basic move of the $T$ on that description. So the operation of the Turing machine exactly corresponds to a sequence of forward moves in $\gamma(T)$.

The effect forward applications of the last three types of defining relations in the presentation of $\gamma(T)$ is to take any $h$-special word $huq_0 vh$, which corresponds to a stopped computation of $T$ and erase symbols as follows: first (1) erase all the $s_b$ symbols to the right of $q_0$ obtaining an the $h$-special word $huq_0 h$; then (2) erase all of the $s_b$ symbols to the left of $q_0 h$ to obtain the word $h q_0 h$; and finally (3) enter state $q$ no longer flanked by $h$ symbols. That is, erase the tape and then enter the (new) final state $q$.

A consequence of these observations is the following.

**Lemma 4.** Suppose $z$ is a word on $\{s_1, \ldots, s_M\}$ such that $T(z)$ exists, that is, the computation of $T$ beginning with $q_1 z$ halts in the stop state $q_0$. Then there is a proof in $\gamma(T)$ of the form

$$hq_1 z h \rightarrow \cdots \rightarrow h q_0 h \rightarrow q$$
in which each step is a forward application of a defining relation. Thus in \( \gamma(T) \) we have \( hq_1zh = q \).

To show the word problem for \( \gamma(T) \) is equivalent to the halting problem for \( T \) we need the converse of this lemma, namely if \( hq_1zh = q \) then \( T(z) \) exists. With a view to subsequent applications we follow the proof in [39], which is a variant of Post's original argument.

Since only the last relation \( hq_0h = q \) creates or destroys \( h \), we observe the following:

**Lemma 5.** If \( w_1 \) and \( w_2 \) are words of \( \gamma(T) \) and neither of them is the word \( q_0 \), and if \( w_1 \rightarrow w_2 \) is an application of a relation, then \( w_1 \) is \( h \)-special if and only if \( w_2 \) is \( h \)-special.

**Proof.** Clearly we may assume \( q_j \neq q_0 \). If \( q_j \neq q_0 \) the conclusion is immediate from the deterministic nature of the Turing machine \( T \). So suppose \( q_j \equiv q_0 \). If \( u \) and \( v \) are empty then there is a forward application of \( hq_0h = q \), but clearly this is the only possibility. If \( u \) is non-empty and \( v \) is empty, then just one of the relations \( s_0q_0h = q_0h \) has a forward application. Finally if \( v \) is non-empty, just one of the relations \( q_0s_0 = q_0 \) has a forward application. This completes the proof.

**Lemma 6.** Let \( w \equiv huq_jvh \) be an \( h \)-special word of \( \gamma(T) \). Then at most one of the relations of \( \gamma(T) \) has a forward application to \( w \).

**Proof.** Clearly we may assume \( q_j \neq q \). If \( q_j \neq q_0 \) the conclusion is immediate from the deterministic nature of the Turing machine \( T \). So suppose \( q_j \equiv q_0 \). If \( u \) and \( v \) are empty then there is a forward application of \( hq_0h = q \), but clearly this is the only possibility. If \( u \) is non-empty and \( v \) is empty, then just one of the relations \( s_0q_0h = q_0h \) has a forward application. Finally if \( v \) is non-empty, just one of the relations \( q_0s_0 = q_0 \) has a forward application. This completes the proof.

**Lemma 7.** Let \( w \equiv huq_jvh \) be an \( h \)-special word of \( \gamma(T) \). If \( w \equiv w_1 \rightarrow w_2 \rightarrow \cdots \rightarrow w_{n-1} \rightarrow q \) is a proof in \( \gamma(T) \), then either this proof contains a reversal or all the applications of relations are forward. In particular, the shortest such proof consists entirely of forward applications of relations.

**Proof.** Clearly the application \( w_{n-1} \rightarrow q \) is forward and \( w_{n-1} \equiv hqh \). Assume that not all applications of relations in the given proof are forward. Suppose that \( w_{j-1} \rightarrow w_j \) is the last backward application of a relation in the proof. Then \( w_j \rightarrow w_{j+1} \) and all subsequent applications are forward. But \( w_{j-1} \rightarrow w_j \) backward implies that \( w_j \rightarrow w_{j-1} \) is a forward application of the same relation.

Now \( w_{n-1} \) is \( h \)-special and none of \( w_j, \ldots, w_{n-1} \) could be \( q \) since only a backward application applies to \( q \). So each of \( w_j, \ldots, w_{n-1} \) must be \( h \)-special. Hence at most one relation has a forward application to \( w_j \) by the previous result. Thus \( w_{j+1} \equiv w_{j-1} \) and hence \( w_{j-1} \rightarrow w_j \rightarrow w_{j+1} \) is a reversal. This proves the lemma.

Since forward applications of relations correspond exactly to basic moves of \( T \) or to tape erasing moves at the end, we conclude the following.
Corollary 8. Suppose \( z \) is a word on \( \{s_1, \ldots, s_M\} \). Then \( hq_1zh = q \) in the semigroup \( \gamma(T) \) if and only if \( T(z) \) exists, that is, the Turing machine \( T \) halts when started in state \( q_1 \) with tape input exactly \( z \).

By choosing \( T \) to be a Turing machine with unsolvable halting problem, we obtain the main result.

Theorem 9 (Post–Markov). There is a finitely presented semigroup with unsolvable word problem.

The proof given above actually shows \( \gamma(T) \) has some additional properties which we record for later reference.

Corollary 10. Every non-trivial proof in \( \gamma(T) \) of the form

\[
q \rightarrow w_2 \rightarrow \cdots \rightarrow w_{n-1} \rightarrow q
\]

contains a reversal.

Proof. Clearly \( w_2 \equiv hq_0h \). So \( w_2 \rightarrow \cdots \rightarrow w_{n-1} \rightarrow q \) either contains a reversal (as claimed), or consists of all forward applications. Assume all the applications \( w_2 \rightarrow \cdots \rightarrow w_{n-1} \rightarrow q \) are forward. But the only forward application of a relation to \( w_2 \equiv hq_0h \) is \( hq_0h \rightarrow q \) so the given proof begins with a reversal. This proves the corollary.

An intuitive way to express the situation is that if we look the graph with vertices words of \( \gamma(T) \) which are equal to \( q \) and edges joining words when \( w_1 \rightarrow w_2 \) is a forward application of a relation, the resulting graph is a connected tree. This corresponds to the fact that the instantaneous descriptions (vertices) in terminating computations in our Turing machine with basic moves (edges) are a forest of trees.

For use in a later section, we modify the above by relabelling the alphabet to include \( h \) among the \( s_i \) and use the term \( s \)-word to describe any (positive) word on \( \{s_0, \ldots, s_M\} \). Then Post’s construction can be recast as yielding a finitely presented semigroup \( \Gamma(T) \) of the form

\[
\Gamma(T) = \langle q, q_0, \ldots, q_N, s_0, \ldots, s_M \mid F_i q_i G_i = H_i q_i K_i, i \in I \rangle
\]

where the \( F_i, G_i, H_i, K_i \) are (positive) \( s \)-words and \( q_i \in \{q, q_0, \ldots, q_N\} \). If \( T \) has an unsolvable halting problem, then the problem of deciding for arbitrary (positive) \( s \)-words \( X, Y \) whether or not \( Xq_1Y = q \) in \( \Gamma(T) \) is recursively unsolvable.

§4. Turing on the word problem for cancellation semigroups. Following on from Post’s construction, Turing [99] constructed a finitely presented cancellation semigroup with unsolvable word problem published in 1950. It is reported that he initially thought he had proved the desired result for groups but before presenting his work had realized that it only gave the cancellation semigroup case.
We will not look at Turing’s construction in any detail, partly because modern approaches achieve the desired unsolvable word problem for groups more simply and directly. But while going from semigroups to cancellation semigroups may seem but a step on the road to the goal of the result for groups, Turing’s paper played a very important role in developments. Novikov’s 1955 paper [76] containing the first published proof of the unsolvability of the word problem for groups is based on Turing’s result for cancellation semigroups. Moreover Boone’s independent 1957 proof of the result for groups, while based only on Post’s construction, used a new “phase change” idea which was suggested by Turing’s work.

Turing’s paper [99] suffers from a number of unfortunate misprints and other technical difficulties and is apparently difficult to follow. In his 1952 review of the paper for the Journal of Symbolic Logic [16] Boone outlines the proof and concludes:

The paper is rendered difficult to read by many minor errors of detail and arrangement, corrections of which must be omitted here for lack of space. In spite of these it is to be emphasized that Turing’s proof is in principle correct, as the reviewer has fully verified. Some years later in 1958 Boone published [18] his explication and analysis of Turing’s paper. While Boone’s 1952 review vouches for the correctness of Turing’s result, Novikov’s reliance on it was a potential concern. Then in 1958 Novikov and Adian published [77] a proof, based directly on the Post–Markov result for semigroups, that the word problem for one-sided cancellation semigroups is unsolvable. They then show how to modify Novikov’s proof for groups so that one-sided cancellation is all that is required.

§ 5. Further unsolvable problems for semigroups. In this section we describe a few further developments concerning decision problems for semigroups which have grown out of the above considerations.

Markov’s isomorphism and recognition results. We will call a property \( P \) of finitely presented semigroups abstract if it does not depend on the particular presentation chosen.

**Definition 11.** An abstract property \( P \) of finitely presented semigroups is a Markov property if there are finitely presented semigroups \( S_+ \) and \( S_- \) such that:

1. \( S_+ \in P \); and
2. if \( S_- \) is embedded in a finitely presented semigroup \( T \) then \( T \notin P \).

\( S_+ \) is called the positive witness and \( S_- \) is called the negative witness for \( P \).

For example the properties “being trivial” (that is, having only one element) is a Markov property and “being finite” is a Markov property of semigroups. There is of course an analogous definition of a “Markov property of groups” which will be considered below. Notice that “being a group” is a Markov
property of semigroups since there are semigroups not embeddable in a group. More examples will be given below for the case of groups. In 1951 Markov [62] proved the following:

**Theorem 12 (Markov).** Let \( \mathcal{P} \) be a Markov property of semigroups. There is no algorithm to determine of an arbitrary finite presentation \( \Pi \) of a semigroup whether or not the semigroup presented by \( \Pi \) has property \( \mathcal{P} \), that is symbolically, whether or not \( \text{sgp}(\Pi) \in \mathcal{P} \).

An immediate consequence of this theorem is that there is no algorithm to recognize the trivial semigroup and so the isomorphism problem is unsolvable.

**Corollary 13 (Markov).** There is no algorithm to determine of two finite presentations of semigroups whether or not the semigroups they present are isomorphic. \( \Box \)

**Proof.** We sketch a proof of Markov's theorem. We may suppose that we have three finite presentations of semigroups on disjoint sets of generators as follows:

1. \( S_0 = \text{sgp}(A_0 \mid R_0) \) which has unsolvable word problem;
2. \( S_+ = \text{sgp}(A_+ \mid R_+) \) which is a positive witness for property \( \mathcal{P} \); and
3. \( S_- = \text{sgp}(A_- \mid R_-) \) which is a negative witness for \( \mathcal{P} \).

For each pair of words \( u, v \) in the generators \( A_0 \) for \( S_0 \) we consider the presentation \( S_{u,v} \) having generators \( A_0 \cup A_- \cup \{ c, d \} \) where \( c, d \) are new symbols, and defining relations \( R_0 \cup R_- \) together with the relations \( cud = 1 \) and the relations \( cvd = cvd' \) where \( \beta \) ranges over all the generators of \( S_{u,v} \).

Clearly if \( u = v \) in \( S_0 \) then in \( S_{u,v} \) we have \( cud = cvd = 1 \); so by the last collection of relations all the generators of \( S_{u,v} \) are equal to 1 and \( S_{u,v} \) is trivial.

Suppose now that \( u \neq v \) in \( S_0 \). We claim that \( S_- \) is embedded in \( S_{u,v} \), that is, if \( y, z \) are two words on the generators of \( S_- \) then \( y = z \) in \( S_- \) if and only if \( y = z \) in \( S_{u,v} \). The necessity is clear. For sufficiency we argue by induction on the number \( n \) of applications of the relation \( 1 \rightarrow cud \) in the indicated direction thought of as a rewrite rule. In case \( n = 0 \), no \( c \) or \( d \) can appear in the proof and so the same proof is valid in the (ordinary) free product \( S_0 \ast S_- \) and so must give a proof in \( S_- \) which is embedded. So suppose \( n > 0 \) and consider the last application of \( 1 \rightarrow cud \). The only way to get rid of that \( c \) and \( d \) is by applying the rule \( 1 \leftarrow cud \) since \( u \neq v \) in \( S_0 \). So this last application must be reversed and can be omitted to give a shorter proof. This proves the claim.

It follow that the free product \( S_+ \ast S_{u,v} \in \mathcal{P} \) if and only if \( u = v \) in \( S_0 \). Since the word problem for \( S_0 \) is unsolvable, this proves the theorem. \( \Box \)

Further remarks on the word problem for semigroups. It turns out that groups defined by a single defining relation have a recursively solvable word problem. This was proved by Magnus using induction on the length of the relation (see below). But the word problem for semigroups with a single defining relation remains a seemingly difficult open problem.
Mathematicians seem to have some desire or impulse to write down quite explicitly a finite presentation of semigroup (or group) with an unsolvable word problem, that is, not depending on a parameter such as a Turing machine. The simplest examples (with respect to the total length of the presentation) of semigroups with unsolvable word problem were constructed independently in 1956 by G. S. Tseitin [98] and D. Scott [88]. Both presentations contain seven defining relations. A proof can be found in [5] for Tseitin’s example which we record and follows:

**Theorem 14.** The semigroup $TS$ presented by the five generators $a, b, c, d, e$ and the seven defining relations

$$ac = ca; \quad ad = da; \quad bc = cb; \quad bd = db; \quad ce = eca; \quad de = edb; \quad cca = ccae$$

has a recursively unsolvable word problem.

The idea here is that the semigroup $TS$ is in a sense universal for all semigroups $G$ that have a presentation of a given special form. So deductions of equations in any such $G$ have parallel deductions of appropriately encoded equations in this one fixed semigroup $TS$. See [5] for a full account.

Matiyasevich [67] constructed a semigroup with 3 defining relations and an unsolvable word problem. Two of his relations are short but the third, much longer relation identifies two words with several hundred symbols each. See the survey [5] by Adian and Durnev for more information on these and many other aspects of the word problem for semigroups.

§6. The word problem for groups. As we previously stated, Novikov in 1955 [76] and independently Boone in 1957 [15] showed the word problem for finitely presented groups is recursively unsolvable. Even in the many modern improved versions, this is still a hard result requiring about ten pages of technical argument and a substantial knowledge of combinatorial group theory.

Actually it was not certain for groups the word problem would turn out to be unsolvable. Indeed there are many large classes of groups for which the word problem can be solved. So before discussing how the finitely presented groups with unsolvable word problem were constructed, we indicate some solutions to the word problem for special classes including several known by the mid 1950s. We also discuss an interesting example (Higman’s non-hopfian group) that appears as a subgroup in both the Novikov and Boone groups and recall some relevant tools from combinatorial group theory.

6.1. Many, even most, groups have solvable word problem.

Dehn’s Algorithm: When Dehn introduce the word, conjugacy and isomorphism problems he provided algorithms for their solution for closed 2-manifolds. In fact the algorithms he described for the word and conjugacy problems work for a very large class of groups. In the case of the word problem *Dehn’s algorithm* for deciding whether a word $w$ is equal to 1 works by successively trying to apply the defining relations to find a way to shorten $w$. 
Definition 15. A finite presentation $G = \langle A \mid R \rangle$ is called a Dehn presentation if $R = \{u_1v_1^{-1}, \ldots, u_nv_n^{-1}\}$ where the lengths of the words satisfy $|v_i| < |u_i|$ and the following holds: if $w$ is a freely reduced word of $G$ with $w =_G 1$ then at least one of the $u_i$ is a subword of $w$.

Thus, if $w =_G 1$ and is freely reduced then for some $i$, we have $w$ is $w_1u_iw_2$ and $w =_G w_1v_iw_2$ where the right hand side is shorter than the left. So after at most $|w|$ steps of this kind or free reductions we must reach the identity $1$. This is known as Dehn’s algorithm for solving the word problem.

Dehn showed that the fundamental group of a closed orientable surface of genus $g \geq 2$ has a Dehn presentation. It is easy to see any finite group has a Dehn presentation by considering a multiplication table. However the free abelian group $\mathbb{Z}^n$ of rank $n \geq 2$ does not have a Dehn presentation (though it is easy to solve the word problem for free abelian groups). One can show that having a Dehn presentation does not depend on the chosen set of generators for the group.

If a group has a Dehn presentation, it is usually convenient to close the relations under inverses and cyclic permutations obtaining a symmetrized set of defining relations. Tartakovski [93], [94], [95] considered symmetrized presentations which have the property that any two non-inverse relations $r_i, r_j \in R$ cancel less than one-sixth their length in freely reducing their product $r_ir_j$ (as words in the free group on the generators. He showed groups satisfying such a cancellation condition have a Dehn presentation and so Dehn’s algorithm solves their word problem. Moreover, they have solvable conjugacy problem. These and related small cancellation groups were later studied geometrically by Lyndon and Schupp and a readable account of their theory can be found in the last chapter of [56].

Almost all finitely presented groups have solvable word problem: In the 1980s Gromov [46] introduced the study hyperbolic groups which have various definitions based on metric properties of their Cayley graphs. Among other characterisations of these groups, Gromov showed that a finitely presented group is hyperbolic if and only if it has a Dehn presentation. Intuitively the relators in a “random” finitely presentation of a group should not have long subwords in common. So they might be expected to satisfy a small cancellation condition and hence be hyperbolic. While this intuition is not exactly correct, it turns out that almost all groups are hyperbolic in a probabilistic sense made precise in Gromov [46] and in Olshanskii [78]. We record this informally as follows:

Theorem 16 (Gromov–Olshanskii). In a certain probabilistic sense, almost all finitely presented groups are hyperbolic and so have Dehn presentations and solvable word and conjugacy problems.
The geometry and algorithmic aspects of hyperbolic groups and more generally groups of non-positive curvature are topics of ongoing interest. A useful reference is the monograph [25] by Bridson and Haefliger.

**Generic-case complexity and the word problem:** Another interesting probabilistic aspect of the word problem arises from the observation that in many groups it is easy to see that certain plentiful types of words are not equal to 1. For example, suppose the finitely presented group $G = \langle a_1, \ldots, a_n \mid r_1 = 1, \ldots, r_m = 1 \rangle$ has an infinite cyclic quotient group, or equivalently that its abelianization $G/[G, G]$ is infinite. (The various groups with unsolvable word problem constructed by Novikov and by Boone all have this property.) Then we can easily adjust the presentation by making Tietze transformations so that at least one of the generators, say $a_1$, appears with exponent sum 0 in each of the relations. So if a word $w$ in the generators of $G$ has $w =_G 1$ then $w$ has exponent sum 0 on $a_1$, that is, $a_1$ and $a_1^{-1}$ each occur the same number of times in $w$. Otherwise the image of $w$ in $G/[G, G]$ would be different from 1.

Now a “random” word of $G$ almost never has exponent sum 0 on $a_1$. The set of words with non-zero exponent sum is generic in the sense that, probabilistically, almost all words have the property. Even though there may be no general algorithm for deciding whether or not $w =_G 1$, there is a partial algorithm which correctly answers “No” on a generic set of words, namely those with non-zero exponent sum on $a_1$.

Such observations have been made precise by Kapovich, Myasnikov, Schupp and Shpilrain in [54] where they introduce a notion of generic-case complexity $C$ where $C$ is a complexity class such as linear time or polynomial time. Here the complexity concerns the behavior of the partial algorithm only on the generic set, completely ignoring the behavior elsewhere. The following is one of their results.

**Theorem 17 ([54]).** Let $G$ be a finitely generated group. Suppose that $H$ is a finite index subgroup of $G$ that possesses an infinite quotient group $\overline{H}$ for which the word problem is solvable in the complexity class $C$. Then the word problem for $G$ has generic-case complexity in the class $C$.

The full worst-case word problem in groups covered by this theorem can be unsolvable, for instance in the groups of Novikov and of Boone. But the types of words which witness this unsolvability are of a very restricted form and lie outside a generic set.

Similar results are also obtained in [54] for the conjugacy problem and for the problem of deciding membership in a finitely generated subgroup. It does not seem to be known whether there is a finitely presented group whose word problem is not generically computable. But there are finitely presented groups which do not satisfy the hypothesis of the above theorem. By using a pair of effectively inseparable sets, the author has constructed (see [72]) a
Many interesting classes of groups have solvable word problem: In different direction, Wilhelm Magnus (who was a student of Dehn) had earlier studied groups given by a single defining relation and proved an important result called the Freiheitssatz [57]: a subset of the generators which omits a symbol in the relation freely generates a free subgroup. Magnus then used his methods to solve the word problem for one-relator groups [58].

As further large classes of groups having solvable word problem we mention finitely generated linear groups and finitely presented residually finite groups. A group is linear of degree $n$ if it has a faithful representation as a group of $n \times n$-matrices over a field, that is, it is isomorphic to a group of matrices. In an important paper in 1940, A. I. Malcev [60] used the compactness theorem from mathematical logic to prove that groups which are locally linear of degree $n$, that is, for which every finitely generated subgroup is linear of degree $n$, are (globally) linear of degree $n$.

We recall some definitions. A group $G$ is said to be residually finite if the intersection of the subgroups of finite index in $G$ is the trivial subgroup. Equivalently, $G$ is residually finite if for every non-trivial element $1 \neq g \in G$ there is a homomorphism $\theta: G \to H$ from $G$ to a finite group such that $\theta(g) \neq H\{1\}$ (where $\theta$ and $H$ depend on $g$). This the fact that $1 \neq g \in G$ is witnessed in the finite quotient $H$ of $G$. We further recall that H. Hopf (1932) had asked for for finitely generated groups $G$ whether $G \cong G/N$ implies $N = 1$. Nowadays we call a group $G$ hopfian if it satisfies the equivalent condition that every surjective map $\phi: G \to G$ is an isomorphism. In this 1940 paper [60] on linear groups, Malcev also proved the following results related to our story.

THEOREM 18 (Malcev).

1. Finitely generated linear groups of degree $n$ are residually finite.
2. Finitely generated residually finite groups are hopfian.
3. Hence, finitely generated linear groups are hopfian.

There are several ways these results are connected with the word problem. In 1943 J. C. C. McKinsey [68] gave a solution to the decision problem for “open sentences” in certain classes of algebras which he termed “finitely reducible”. In the case of groups this is essentially the finitely presented residually finite groups (as later observed by Verena Huber-Dyson [42]), though he only mentions finitely generated abelian groups explicitly.

THEOREM 19 (McKinsey). Finitely presented, residually finite groups have a solvable word problem.

PROOF. Let $G$ be finitely presented and residually finite. Since $G$ is finitely presented the $N$ set of words equal to 1 in $G$ is recursively enumerable. To
decide whether an arbitrary word \( w \in G \) is equal to 1, we start enumerating \( N \). At the same time we begin enumerating all homomorphisms from \( \phi : G \rightarrow H \) where \( H \) is a finite group, say given by a finite multiplication table. We can check whether a map on generators defines a homomorphism by checking whether the images of the relators are equal to 1 in \( H \) using its multiplication table. As such homomorphisms are enumerated, evaluate them on \( w \) and check to see whether \( \phi(w) = 1 \). If \( w \neq_G 1 \), eventually one of its images will be different from 1, so we have an answer. But if \( w =_G 1 \) then \( w \) will appear in the list of \( N \). Thus we have an algorithm to solve the word problem for \( G \).

While this style of argument is familiar in logic and recursion theory, many mathematicians find the argument disconcerting since there is no easy estimate in advance of how long one might have to wait for an answer.

Concerning finitely generated linear groups \( G \) of degree \( n \), they may not be finitely presented so McKinsey’s result may not apply. Nevertheless such a linear group over a given field is also linear over a larger field and so should be linear over a field in which one can actually do the field arithmetic computationally. Hence by multiplying the matrices corresponding to the group generators, one should be able to determine whether a product is equal to the identity matrix. This has all been made precise, in particular by Rabin [85] (see also Lipton and Zalcstein [55]).

**Theorem 20** (Rabin). A finitely generated linear group over a field has solvable word problem.

Linear groups are quite a large class of groups and many of the infinite groups mathematicians have studied historically are linear or nearly so, often for geometric reasons. So not only do “most” finitely presented groups have solvable word problem, so also do those arising in many traditionally studied contexts.

For a survey of results on decision problems for various classes of groups as of about 1990 see [72].

### 6.2. Some constructions in group theory and Higman’s non-hopfian group

In this subsection we discuss some examples relevant to our story and review some standard constructions from combinatorial group theory which will be used to “unsolve” the word problem below.

**The group \( BS(1,2) \).** Consider the group \( G = \langle x, s \mid xs = sx^2 \rangle \), which is often referred to as \( BS(1,2) \) for reasons we indicate below. This interesting group appears as a subgroup in important ways in the (very different) groups with unsolvable word problem constructed by Novikov and by Boone, and we want to examine it more closely. In fact this group \( G = BS(1,2) \) is a linear group and one can show the map

\[
x \mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad s \mapsto \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}
\]
is a homomorphism which embeds $G$ as a subgroup of $GL(2, \mathbb{Q})$. Hence $G$ is residually finite and hopfian and has solvable word problem.

But to better understand $G$ we describe an elementary solution to the word problem for $G$. The defining relation can be also written in a number of equivalent forms: $s^{-1}xs = x^2$ or $x = sx^2s^{-1}$ or $s^{-1}x = x^2s^{-1}$ or $x^{-1}s = sx^{-2}$ or $s^{-1}x^{-1} = x^{-2}s^{-1}$. We observe that, if one starts with any word $w$ on the generators $x$ and $s$, the relation can be applied to move the letter $s$ from right to left over $x^\pm 1$ symbols creating additional $x$’s. Similarly an $s^{-1}$ can be moved from left to right over $x^\pm 1$ symbols. So $w = _G s^i x^j s^{-k}$ where $i \geq 0$ and $k \geq 0$ and $j \in \mathbb{Z}$. In case $j = 2m$ is even and both $i > 0$ and $k > 0$ we can apply the relation $x = sx^2s^{-1}$ to deduce that $w = _G s^{i-1}x^ms^{-k}$ which has fewer $s$-symbols. This process is called pinching a pair of $s$-symbols, or an $s$-pinch. Continuing one obtains $w = _G s^i x^j s^{-k}$ where either $j$ is odd or at least one of $i$ or $k$ is 0—in either case no further pinches are possible. If the right hand side of this equation is not the trivial word, then one can show $w \neq _G 1$. So the method described solves the word problem for $G$.

Notice that the algorithm removes inverse pairs of $s$-symbols during the pinching operation, but inverse pairs of $s$-symbols were never inserted. Also observe that for the word $sxs^{-1}$ an $s$-pinch is not possible and this word is not equal in $G$ to any word with fewer $s$-symbols.

**Amalgamated free products.** We briefly recall the amalgamated free product construction. Suppose $H$ and $K$ are two groups with respective subgroups $A$ and $B$ which are isomorphic via an isomorphism $\phi: A \to B$. We can assume we have presentations $H = \langle X_H \mid R_H \rangle$ and $K = \langle X_K \mid R_K \rangle$ on disjoint alphabets. Then the free product of $H$ and $K$ with amalgamated subgroup $A = B$ is the group with presentation

$$H \star_{A = B} K = \langle X_H, X_K \mid R_H, R_K, a = \phi(a) \text{ for } a \in A \rangle.$$ 

There are two familiar (but not obvious) facts we need to know about this construction: (1) the inclusion maps on generators embed both $H$ and $K$ in $H \star_{A = B} K$; and (2) if $w = h_1k_1h_2k_2\ldots h_mk_m$ is an alternating product of elements of $H$ and $K$ (where possibly $h_1$ or $k_m$ does not appear) and if $w = 1$ in $H \star_{A = B} K$, then for some $i \in \{1, \ldots, m\}$ either $h_i \in A$ or $k_i \in B$ and so relations of the form $a = \phi(a)$ can be used to reduce the number of alternations.

**HNN extensions.** The group $BS(1, 2)$ above is an example of a very general construction, usually called an HNN extension, which is closely related to free products with amalgamated subgroups. It was introduced by Higman, Neumann and Neumann in their famous 1948 paper [50] and used to prove that any countable group $C$ can be embedded in a two generator group $H_C$ spending on $C$. As a consequence they showed that there are continuously many two generator groups.
The amalgamated free product construction identifies isomorphic subgroups of two different groups. The HNN extension or HNN construction takes two isomorphic subgroups of the same group and realises this isomorphism as a conjugation in a universal way. So suppose that the group $H = \langle X_H \mid R_H \rangle$ has a pair of isomorphic subgroups via the isomorphism $\phi : A \to B$. Then the HNN extension of $H$ with associated subgroups $A$ and $B$ and stable letter $s$ is the group with presentation

$$H \star_\phi = \langle X_H, s \mid R_H, s^{-1}as = \phi(a) \text{ for } a \in A \rangle$$

where $s$ is a new symbol not in $X_H$ called the stable letter. There are two important facts we need to know about this HNN construction (they are similar to the above statements about amalgamated free products):

**Lemma 21.** Let $H \star_\phi$ be an HNN extension as above. Then

1. (Higman–Neumann–Neumann) the inclusion map on generators embeds $H$ into $H \star_\phi$; and
2. (Britton) if $w$ is a word of $H \star_\phi$ in which $s$ or $s^{-1}$ appears and if $w = 1$ in $H \star_\phi$, then $w$ contains either a subword of the form $s^{-1}as$ or of the form $s\phi(a)s^{-1}$ and so the relations of the form $s^{-1}as = \phi(a)$ can be used to perform an $s$-pinch and reduce the number of $s$-symbols in $w$. \(\sqsubseteq\)

The first assertion of the lemma is a theorem Higman, Neumann and Neumann from [50]. The second assertion is usually known as Britton’s Lemma whose role in our story will be explained below. There are very similar statements in case one has several pairs of isomorphic subgroups and isomorphisms $\phi_i : A_i \cong B_i$ and corresponding new stable letters $s_i$, but for notational simplicity we here concentrate on the one stable letter case.

The amalgamated free product and HNN extension constructions are closely related and basic facts about one often can be deduced from the analogous facts about the other (see for instance [56]). They also play an important role in geometric topology, particularly the study of 3-manifolds. Suppose one has a nicely embedded two sided surface $Y$ in a 3-manifold $M$ (possibly with boundary) such that the inclusion embeds the fundamental group of $Y$ in the fundamental group of $M$. If one cuts $M$ along $Y$, the resulting manifold (with boundary) either (1) has 2 components so $Y$ separates and the fundamental group of $M$ is the corresponding amalgamated free product; or (2) has one component so $Y$ does not separate and the fundamental group of $M$ is the corresponding HNN extension. (The associated isomorphic subgroups come from the two copies of $Y$ in each case.)

As an example of an HNN extension we note that the group $G = \langle x, s \mid xs = sx^2 \rangle = \langle x, s \mid s^{-1}xs = x^2 \rangle$ introduced above is an HNN extension of the infinite cyclic group generated by $x$ with associated subgroups the cyclic
groups \( \langle x \rangle \) and \( \langle x^2 \rangle \) with the isomorphism \( \phi \) defined by \( \phi(x) = x^2 \). (Notice that we only need to include the relations \( a = \phi(a) \) for \( a \) ranging over a generating set for the subgroup \( A \)). One consequence of Britton’s Lemma for this group is that in the word \( sxsx^{-1} \) the subword \( x \) is not of the form \( \phi(x^i) \) and so no \( s \)-pinch is possible and \( sxsx^{-1} \) is not equal to a word with fewer \( s \)-symbols.

**Using Britton’s Lemma.** In our later discussion we will assume some familiarity with various ways of applying Britton’s Lemma, so we pause here to introduce some terminology and make a few observations.

Continuing with the above notation, suppose \( H \star \) is the HNN extension of \( H \) with stable letter \( s \) and associated subgroups \( A \) and \( B \) which are isomorphic via the map \( \phi: A \rightarrow B \). If \( w \) is any word of \( H \star \) (not necessarily = 1), then an \( s \)-pinch in \( w \) is a subword of the form \( s^{-1}us \) where \( u \in A \) or \( su^{-1} \) where \( u \in B \) (here we are assuming in either case that \( u \) does not contain \( s^{-1} \)). A word \( w \) is said to be \( s \)-reduced if it does not contain any \( s \)-pinch.

Changing our point of view slightly, suppose that \( H \) is given by a presentation \( H = \langle X_H \mid R_H \rangle \) and that the associated subgroups are given by sets of generators, say \( A = gp(a_1, a_2, \ldots) \) and \( B = gp(b_1, b_2, \ldots) \) so that \( \phi(a_i) = b_i \) where the \( a_i \) and \( b_i \) are certain words in the generators \( X_H \) for \( H \). Then the HNN extension can be presented as

\[
H \star = \langle X_H, s \mid R_H, s^{-1}a_1s = b_1, s^{-1}a_2s = b_2, \ldots \rangle
\]

\[
\cong \langle X_H, s \mid R_H, a_1s = sb_1, a_2s = sb_2, \ldots \rangle.
\]

Suppose the word \( w \) contains and \( s \)-pinch. For instance suppose it contains a subword of the form \( s^{-1}us \) with \( u \in A \). This means there is a word \( a_{i_1}^{e_1} \cdots a_{i_k}^{e_k} \) on the generators of \( A \) so that \( u = a_{i_1}^{e_1} \cdots a_{i_k}^{e_k} \) is a consequence of the defining relations \( R_H \) of \( H \). Then, applying the relations of \( H \star \) involving \( s \) in the form \( a_is = sb \) repeatedly moving \( s \) from right to left, we have the following calculation:

\[
w \equiv w_1s^{-1}usw_2 \equiv w_1s^{-1}a_{i_1}^{e_1} \cdots a_{i_k}^{e_k}sw_2
\]

\[
= w_1s^{-1}sb_{i_1}^{e_1} \cdots b_{i_k}^{e_k}w_2
\]

\[
= w_1b_{i_1}^{e_1} \cdots b_{i_k}^{e_k}w_2
\]

in which the result has fewer \( s \)-symbols. The analogous process transforms an \( s \)-pinch of the form \( sus^{-1} \) with \( u \in B \): first use the relations of \( H \) to express \( u \) as a word in the \( b_i \) and then move \( s \) from left to right over the \( b_i \)’s to give an equation

\[
w \equiv w_1sus^{-1}w_2 \equiv w_1sb_{i_1}^{e_1} \cdots b_{i_k}^{e_k}s^{-1}w_2
\]

\[
= w_1a_{i_1}^{e_1} \cdots a_{i_k}^{e_k}w_2
\]
with fewer $s$-symbols. Either process is called pinching out an $s$ or removing an $s$-pinch.

There is subtle but important difference here in the form we choose to express the defining relations. From a group theoretic point of view, the form $s^{-1}a_is = b_i$ emphasises that the subgroups $A$ and $B$ are conjugate realising the isomorphism $\phi(a_i) = b_i$. But thinking in terms of rewrite rules transforming words, to show say $s^{-1}a_1a_2s = b_1b_2$ we need to insert an inverse pair of $s$-symbols. Thus the calculation in detail would be $s^{-1}a_1a_2s = s^{-1}a_1ss^{-1}a_2s = s^{-1}a_1sb_2 = b_1b_2$. By contrast, using the relations in the form $a_is = sb_i$ as we have above, we have the detailed calculation $s^{-1}a_1a_2s = s^{-1}a_1sb_2s = s^{-1}sb_1b_2 = b_1b_2$ which does not involve the insertion of inverse pairs of $s$-symbols.

Clearly, by successively pinching out $s$-symbols we eventually arrive at an $s$-reduced word $\overline{w}$ such that $w =_{H*} \overline{w}$. It is fairly easy to show that a word $w$ of $H* \phi$ is $s$-reduced if and only if it is not equal to any word with fewer $s$-symbols.

For if a word $w$ is not equal to a word with fewer $s$-symbols it must clearly be $s$-reduced by the above. But if $w$ and $w'$ are two $s$-reduced words, then $1 = w^{-1}w'$ must reduce to an $s$-free word (= word without $s$-symbols) and in the reduction process each pinch must involve an $s^{\pm 1}$ from each. So $w$ and $w'$ are $s$-parallel in the sense that they have the same signed sequence of $s$ symbols.

Finally we note that (using the $a_is = sb_i$ form of the defining relations) the process of removing $s$-pinches never introduces an inverse pair of $s$ symbols. So if $w$ is a word of $H* \phi$ which is equal to an $s$-free word $w'$ it can be transformed into $w'$ without introducing any inverse pairs of $s$-symbols.

This fact is due to Novikov (his Principal Lemma). Novikov's Principal Lemma is essentially equivalent to Britton's Lemma as will be discussed below.

**Higman's non-hopfian group.** In 1951 Graham Higman [48] gave the first example of a finitely presented non-hopfian group. Higman's non-hopfian group can be presented as $H = \langle x,s_1,s_2 \mid s_1^{-1}xs_1 = x^2, s_2^{-1}xs_2 = x^2 \rangle$.

Now the map $\theta : H \to H$ defined by $s_i \mapsto s_i$ and $x \mapsto x^2$ is a surjective homomorphism from $H$ onto itself (easy check). Observe that the element $s_1xs_1^{-1}s_2x^{-1}s_2^{-1} \neq_H 1$ since, if we view it as an HNN extension with two stable letters $s_1$ and $s_2$, there is no possible $s_i$-pinch. But

$$\theta(s_1xs_1^{-1}s_2x^{-1}s_2^{-1}) = (s_1x^2s_1^{-1})(s_2x^{-2}s_2^{-1}) =_H xx^{-1} = 1$$

and so $\theta$ is not injective and $H$ is non-hopfian. Hence $H$ is not residually finite.

There are several alternative ways to view Higman's group $H$. For instance it is the amalgamated free product of two copies $G_i = \langle x, s_i \mid s_i^{-1}xs_i = x^2 \rangle$
amalgamated along the subgroup \( \langle x \rangle \), that is, \( H = G_1 \ast_{\langle x \rangle} G_2 \). Introducing a new symbol \( t \) a setting \( t = s_1s_2^{-1} \) one can apply a few Tietze transformations to obtain \( H \cong \langle x, t, s_1 \mid xt = tx, s_1^{-1}xs_1 = x^2 \rangle \) which is an HNN extension of the free abelian group on \( x, t \) with stable letter \( s_1 \).

It happens that Higman’s non-hopfian group also appears in important ways as subgroups of the (quite different) groups constructed by Novikov and by Boone. So they visibly contain a non-hopfian group and can not be residually finite (for reasons other than having an unsolvable word problem). We will see in Boone’s groups, for instance, the subgroup

\[
E_1 = \langle x, s_0, \ldots, s_M \mid s_b^{-1}xs_b = x^2 \text{ for } b = 0, \ldots, M \rangle
\]

which is an enlarged version of Higman’s group with more stable letters. A key technical property, proved in Lemma 29 below, is that certain words contain only positive occurrences of \( s_b \) symbols.

As further examples we mention the Baumslag–Solitar groups commonly denoted

\[
BS(n, m) = \langle x, s \mid s^{-1}xs = x^m \rangle
\]

studied by Baumslag and Solitar in [8]. Among other things they showed that the one-relator group \( BS(2, 3) \) is non-hopfian. This can be easily deduced from Britton’s Lemma. The map \( x \to x^2 \) and \( s \to s \) defines a homomorphism \( \phi \) from \( BS(2, 3) \) onto itself. Now the commutator \( [x, s^{-1}xs] = x^{-1}s^{-1}x^{-1}xs^{-1}xs \neq 1 \) by Britton’s Lemma since no \( s \)-pinch is possible because \( x^{\pm 1} \notin \langle x^2 \rangle \) and \( x^{\pm 1} \notin \langle x^3 \rangle \). But \( \phi([x, s^{-1}xs]) = [x^2, s^{-1}x^2s] = [x^2, x^3] = 1 \) and so \( BS(2, 3) \) is non-hopfian.

6.3. Novikov’s groups and principal lemma. In 1952 [74] Novikov announced the construction of a finitely presented group with unsolvable word problem, providing a sketch of the construction but no proofs. According to reviews by Hirsch [51] and by Boone [17] the construction is based on an unsolvable problem of Post concerning canonical systems and contains a number of misprints which make the exact meaning unclear.

In 1954 [75] Novikov published a construction of a finitely presented group with unsolvable conjugacy problem. While this follows trivially from the unsolvability of the word problem, Novikov provided a separate proof for the conjugacy problem which is considerably simpler and has many fewer generators and relations. Again this result is based on “Post systems”. See the reviews by Hirsch [52] and Britton [28].

Then in 1955 [76] Novikov published a rather lengthy paper proving the unsolvability of the word problem based on Turing’s cancellation semigroup with unsolvable word problem. The review by A. A. Markov [63] of this 1955 proof concludes as follows:
The exposition is rather inadequate, which makes for very difficult reading. There are numerous inaccuracies and some untrue assertions. In addition there are many misprints, of which only a part appear on the list of errata appended to the monograph. All these defects, however, do not affect the essence of the matter. After a careful study of the monograph the reviewer has come to the conclusion that the small errors in it can all be easily corrected, and hence that the result of this remarkable work is valid.

The review [27] for the Journal of Symbolic Logic by Britton (based on Hirsch’s translation of the monograph) notes that subsequent proofs have been given by Boone and by Britton himself and includes the comment

With a paper of this size and complexity, involving such profusion of detail, it is difficult to feel convinced that one has overlooked nothing. However, the reviewer now firmly believes that this proof is correct and contains no serious gaps.

The group constructed in this monograph differs significantly from the one outlined in the 1952 announcement and it is not so clear whether the earlier claimed proof held up. But it is clear that Novikov gave the first published proof of the unsolvability of the word problem for finitely presented groups. For some related historical comments, see the obituary of Boone [37].

We will not look at Novikov’s construction in any detail since the Post–Boone–Britton sequence of constructions provides a more straightforward approach than the Post–Turing–Novikov sequence of constructions. But we want to record an important tool in Novikov’s proof which he calls the “Principal Lemma”.

Novikov’s general approach to the word problem was combinatorial and more in the spirit of mathematical logic than group theory. As above for semigroups, relations are used to transform or rewrite words. But in addition to the defining relations specified in the presentation one has the inverse pair relations $xx^{-1} = 1$ and $x^{-1}x = 1$ for each generating symbol $x$. So two words are equal in the group defined by a presentation if and only if one can be obtained from the other by a sequence of elementary transformations of the form: (1) $z_1u_iz_2 \rightarrow z_1v_iz_2$ or $z_1v_iz_2 \rightarrow z_1u_iz_2$ where $u_i = v_i$ is a defining relation; or (2) inserting or deleting inverse pairs, as in $z_1z_2 \rightarrow z_1xx^{-1}z_2$ or $z_1xx^{-1}z_2 \rightarrow z_1z_2$.

Suppose the group $H = \langle X_H \mid R_H \rangle$ has a pair of isomorphic subgroups via the isomorphism $\phi: A \rightarrow B$. Let

$$H \star = \langle X_H, s \mid R_H, as = s\phi(a) \text{ for } a \in A \rangle$$

be the corresponding HNN extension with associated subgroups $A$ and $B$ and stable letter $s$. Note that if $A$ is generated by $\{a_1, a_2, \ldots, a_m\}$ we can use
instead the more efficient presentation
\[ H \star = (X_H, s \mid R_H, a_i s = s \phi(a_i) \text{ for } i = 1, \ldots, m). \]

In particular if \( H \) is finitely presented and \( A \) is finitely generated then \( H \star \) is finitely presented.

**Theorem 22 (Novikov’s Principal Lemma).** Let \( H \star \) be an HNN extension as above. Suppose that
1. \( w, z \) are two words of such that \( w = z \) in \( H \star \) and
2. the symbols \( s^{\pm 1} \) do not appear in \( z \).

Then there is an elementary sequence of transformations
\[ w \equiv w_1 \rightarrow w_2 \rightarrow \cdots \rightarrow w_n \equiv z \]
that contains no insertions of the letter \( s \).

We first observe that in the case \( w \) also does not contain the symbols \( s^{\pm 1} \), then the Principal Lemma implies \( w =_H z \) and so \( H \) is embedded in \( H \star \).

This fact was originally proved by Higman, Neumann and Neumann in their 1948 paper [50]. We further note that, in the case \( w \) does contain the symbols \( s^{\pm 1} \), the assertion of the Principal Lemma is easily seen to be equivalent to Britton’s Lemma. However, as we shall see below, the particular form of Britton’s Lemma makes it very useful for group-theoretic style arguments.

As we have already noted, following Novikov’s 1955 proof of the unsolvability of the word problem, Boone gave a rather different proof in 1957 based directly on Post’s semigroup. In 1958, Britton [29] gave a somewhat different proof based on Turing’s cancellation semigroups result and some of his own earlier work on solutions to the word problem. In this paper a version of what is called Britton’s Lemma appears but under the unnecessarily restrictive assumption that the associated subgroups intersect trivially. As his review above indicates, Britton would have been familiar with Novikov’s work and Principal Lemma, but he establishes it using the earlier paper of Higman, Neumann and Neumann [50].


In later years Boone, who seemingly never tried to read Novikov’s work in detail, asked the writer if there might be a more combinatorial style proof of Britton’s Lemma. He seemed surprised when I explained to him that Novikov had proved (in a combinatorial manner) his Principal Lemma which is easily shown to be equivalent to Britton’s Lemma.

**6.4. Boone’s groups.** Boone’s initial construction of a group with unsolvable word problem was developed in a six part series of papers over 5 years.
In the first four articles it is shown the unsolvability of the “quasi-magnus problem”, that is, whether or not an arbitrary word is expressible as a product of positive powers of a certain proper subset of the generators. The unsolvability of the word problem itself occupies the last two parts which are in principle independent of the first four, but the exposition relies on the earlier parts.

We now describe a variant of Boone’s construction of a finitely presented group with unsolvable word problem. The construction was modified a number of times: first by Boone himself [19] and then by Britton [30] who gave a much more group theoretic proof. Further variations and simplifications were then contributed by Boone, Collins and myself, and were presented in various seminars. Eventually these were included in later editions of Rotman’s group theory text [87]. We also borrow from the discussion in [39] which provides for additional applications to asphericity. The theme is still the same so we continue to refer to them as Boone’s groups. An elegant further simplification of Boone’s approach was introduced by Borisov [24] and is incorporated in the exposition in [5], but we have chosen to follow the earlier treatments since the involvement of Turing machines is more apparent.

A very attractive aspect of Boone’s group is that it encodes Post’s semigroup construction in a very direct way so that the underlying Turing machine and its computations are still visible in the group. Actually, as Boone liked to describe the situation, the words for which the word problem is shown to be difficult encode four copies of the Turing machine operating simultaneously. In addition it is exhibited as a tower of successive HNN extensions which enables one to use techniques of combinatorial group theory.

So the construction of Boone’s group begins with a Turing machine $T$ having an unsolvable halting problem. The semigroup construction of Post is then applied to obtain a finitely presented semigroup $\Gamma(T)$ of the form

$$\Gamma(T) = \langle q, q_0, \ldots, q_N, s_0, \ldots, s_M | F_i q_i G_i = H_i q_i K_i \ (i \in I) \rangle$$

where the $F_i, G_i, H_i, K_i$ are positive $s$-words and $q_i \in \{q, q_0, \ldots, q_N\}$. In the present context, an $s$-word is a word on the symbols $s_0, \ldots, s_M$ and their inverses. An $s$-word is positive if it contains no $s^{-1}$ symbols. By Post’s results:

**Lemma 23.** The problem of deciding for an arbitrary pair of positive $s$-words $X, Y$ whether or not $Xq_i Y = q$ in $\Gamma(T)$ is recursively unsolvable.

As before we use $X \equiv Y$ to mean the words $X$ and $Y$ are identical (letter by letter). If $X \equiv s_{b_1}^{e_1} \cdots s_{b_m}^{e_m}$ is an $s$-word, we define $X^\# \equiv s_{b_1}^{-e_1} \cdots s_{b_m}^{-e_m}$. Note that $X^\#$ is not the same as $X^{-1}$. Also, if $X$ and $Y$ are $s$-words, then $(X^\#)^\# \equiv X$ and $(XY)^\# = X^\# Y^\#$.

**Boone’s group** $B = B(T)$ is then the finitely presented group depending on $\Gamma(T)$ described as follows:

- **generators:** $q, q_0, \ldots, q_N, s_0, \ldots, s_M, r_i \ (i \in I), x, t, k$;
relations: for all $i \in I$ and all $b = 0, \ldots, M$,
\[
\begin{align*}
x_i s_b &= s_b x^2 \\
r_i s_b &= s_b x r_i x \\
r_i^{-1} F_i^b q_i G_r = H_i^b q_i K_i \\
t r_i &= r_i t \\
t x &= x t \\
k r_i &= r_i k \\
k x &= x k \\
k(q^{-1} t q) &= (q^{-1} t q) k.
\end{align*}
\]

The subsets $\Delta_1 \subset \Delta_2 \subset \Delta_3$ of the relations each define a presentation of a group $B_i$ generated by the symbols appearing in the $\Delta_i$.

**Definition 24.** A word $\Sigma$ is special if $\Sigma \equiv X^q_j Y$ where $X$ and $Y$ are positive s-words and $q_j \in \{q, q_0, \ldots, q_N\}$.

The main technical result linking the word problem in $B(T)$ as presented above to the word problem in Post’s semigroup $\Gamma(T)$ is the following:

**Lemma 25 (Boone’s Lemma).** If $\Sigma \equiv X^q_j Y$ is a special word in $B(T)$, then
\[
k(\Sigma^{-1} t \Sigma) = (\Sigma^{-1} t \Sigma) k \text{ in } B(T)
\]
if and only if $X q_j Y = q$ in $\Gamma(T)$.

Because of the properties (Lemma 23) of Post’s semigroup, Boone’s Lemma immediately implies the finitely presented group $B(T)$ has an unsolvable word problem.

**Theorem 26 (Novikov–Boone).** There exists a finitely presented group $B$ with recursively unsolvable word problem.

We here give a sketchy but essentially complete proof of Boone’s Lemma and make some other observations so as to give some feeling for why the construction works. For additional details the reader is referred to the text book by Rotman [87].

From the very form of the various defining relations for $B = B(T)$ above, we suspect that the group $B$ is built up by successively forming HNN extensions. To make this precise we let $B_0$ denote the infinite cyclic group generated by $x$, and let $Q$ denote the free group with basis $\{q, q_0, \ldots, q_N\}$. Using some standard facts from combinatorial group theory, one can easily show the following (this is Lemma 12.11 in [87]).

**Lemma 27.** In the chain
\[
B_0 \leq B_1 \leq B_1 \ast Q \leq B_2 \leq B_3 \leq B
\]
each group is an HNN-extension of its predecessor; moreover, the free product $B_1 \ast Q$ is an HNN-extension of $B_0$. 
One consequence of this is that copies of the group $BS(1, 2) = \langle x, s \mid xs = sx^2 \rangle$ and also Higman’s non-hopf group are embedded as subgroups of $\mathcal{B}$. Indeed we observe that $\mathcal{B}_1$ is an enlarged version of Higman’s group obtained by adding even more stable letters $s_b$, and it is non-hopfian by the same argument. As we have noted before these groups also appear as subgroups of Novikov’s groups and play an important role there as well.

We previously observed in our discussion of $BS(1, 2)$ that we can move $s_b$ symbols from right to left over $x^{\pm 1}$. So if $R_1$ is any word on the $r_i$ and $x$, then $R_is_b = s_bR_2$ where $R_2 = s_b^{-1}R_is_b$ is another word on $r_i$ and $x$.

Similarly, if $L_1$ is a word on $r_i$ and $x$, then $s_b^{-1}$ can be moved from left to right over $L_1$. So $s_b^{-1}L_1 = L_2s_b^{-1}$ where $L_2 = s_b^{-1}L_1s_b$ is again another word on $r_i$ and $x$.

More generally these considerations apply to moving any positive $s$-word $V$ from right to left and $V^\#$ from left to right across a word in $r_i$ and $x$. Also observe that only the relations of $\mathcal{B}_2$ have been used in these calculations. We state this as follows:

**Lemma 28.** Suppose that $V$ is a positive $s$-word and that $R_1$ and $L_1$ are words on $r_i$ and $x$. Then there are words $R_2$ and $L_2$ on $r_i$ and $x$ so that $R_1V = L_2VR_2$ and $V^\#L_2 = L_2V^\#$.

Next consider one of the defining relations of $\mathcal{B}_2$ of the form $r_i^{-1}F_i^\#q_iG_i = H^\#K_i$. Recall that this relation comes from Post’s construction so it corresponds a basic step in a Turing machine. Suppose that $U, V$ are two positive $s$-words. Then, by the relation and the previous Lemma,

$$U^\#(H^\#K_i) = U^\#(r_i^{-1}F_i^\#q_iG_i) = L_2VR_2$$

for certain words $L, R$ on $r_i$ and $X$. Similarly

$$U^\#(r_iH^\#q_iK_i) = U^\#(F_i^\#q_iG_i) = L_2VR_2$$

for certain words $L, R$ on $r_i$ and $X$. We use these observations in the following:

**Proof of sufficiency in Boone’s Lemma.** Suppose $X, Y$ are two positive $s$-words such that $Xq_jY = q \in \Gamma(T)$. Then there is a sequence of applications of the defining relations of $\Gamma(T)$, say

$$Xq_jY \equiv w_1 \rightarrow w_2 \rightarrow \cdots \rightarrow w_n \equiv q$$

so that for each $i$, one of the words $w_i$ and $w_{i+1}$ has the form $U^\#F_i^\#q_iG_iV$ and the other has the form $U^\#H^\#q_iK_iV$.

Now if we start with a special word $\Sigma \equiv X^\#q_1Y$ and apply the corresponding relations in $B(T)$ (in the same direction as the above sequence $w_j \rightarrow w_{j+1}$) and move the $s$-words towards the $q_i$ as in the above discussion we get a corresponding sequence of equalities in $B(T)$ of the form $w^*_i = L_1w^*_{i+1}R_i$ where $w^*_i$ is the special word of $B$ corresponding to $w_i$ and $L_1, R_i$ are words
on the \( r_j \) and \( x \). Setting \( L = L_1 \cdots L_{n-1} \) and \( R = R_{n-1} \cdots R_1 \), it follows that in \( B(T) \) we have \( \Sigma = LqR \) where \( L, R \) are words in \( r_j \) and \( x \). Note that in arriving at this equation only the relations of \( B_2 \) are used.

But the generators \( k \) and \( t \) commute with the \( r_j \) and with \( x \) so they commute with \( L \) and \( R \). Therefore

\[
 k(\Sigma^{-1}t\Sigma) = k(R^{-1}q^{-1}L^{-1}tLqR) = k(R^{-1}q^{-1}tqR) = R^{-1}k(q^{-1}tq)R \\
= R^{-1}(q^{-1}tq)kR \quad \text{(since } k \text{ commutes with } q^{-1}tq) \\
= (R^{-1}q^{-1}tqR)k = (R^{-1}q^{-1}L^{-1}tLqR)k = (\Sigma^{-1}t\Sigma)k
\]

as claimed. This proves sufficiency in Boone’s Lemma.

**Discussion.** Conceptually, the proof of sufficiency establishes two connected implications between equations in various systems which we can write as

\[
 Xq_jY = q \implies \Sigma \equiv X^nq_jY \equiv B_3(T) LqR \implies k(\Sigma^{-1}t\Sigma) = B(T) (\Sigma^{-1}t\Sigma)k
\]

where \( X, Y \) are positive words on the \( s_b \) and \( L, R \) are words on the \( r_j \) and \( x \). In the proof of necessity, the converse of each of these implications is established. The converse of the second implication is fairly straightforward using Britton’s Lemma. The converse of the first implication is more delicate because one needs to show certain words on the \( s_b \) are positive and hence the semigroup rewrites can be carried out.

**Proof of necessity in Boone’s Lemma.** Suppose \( \Sigma \equiv X^nq_jY \) is a special word in \( B = B(T) \) and that \( k(\Sigma^{-1}t\Sigma) = B(\Sigma^{-1}t\Sigma)k \). Then

\[
 k(\Sigma^{-1}t\Sigma)k^{-1}(\Sigma^{-1}t\Sigma) = B_1 1
\]

and since \( B \) is an HNN extension of \( B_3 \) with stable letter \( k \), there must be a \( k \)-pinch. There are exactly two \( k \equiv 1 \) which must pinch over the word between them and hence \( \Sigma^{-1}t\Sigma \) must be equal in \( B_3 \) to a word \( W \) on \( \{r_i(i \in I), x, q^{-1}tq\} \). Now \( B_3 \) is an HNN extension of \( B_2 \) with stable letter \( t \) which commutes with \( \{r_i(i \in I), x\} \). So pinching as many \( t \)-symbols from \( W \) as possible, we obtain a \( t \)-reduced word \( W_1 = B_3 W \) and still on \( \{r_i(i \in I), x, q^{-1}tq\} \). Thus \( \Sigma^{-1}t\Sigma = B_3 W_1 \) and both sides are \( t \)-reduced and hence \( W_1 \) has a single \( q^{-1}tq \). Thus we may suppose \( W_1 \equiv R^{-1}q^{-1}tqR_0 \) where \( R, R_0 \) are words on \( \{r_i(i \in I), x\} \).

Thus \( \Sigma^{-1}t\Sigma R^{-1}q^{-1}tqR_0 = B_3 1 \) and by Britton’s Lemma there must be a \( t \)-pinch and

\[
 \Sigma R^{-1}q^{-1} \equiv X^nqY R^{-1}q^{-1} = B_3 L
\]

where \( L \) is a word on \( \{r_i(i \in I), x\} \). Thus in \( B_2 \) we have the equation

\[
 \Sigma \equiv X^nqY = B_3 LqR
\]

where \( L, R \) are words on \( \{r_i(i \in I), x\} \).
Remark. The argument so far establishes the converse of the second implication in the “Discussion” above. So in \( B \) we know that \( k(\Sigma^{-1}Tq) = (\Sigma^{-1}Tq)k \) if and only if \( \Sigma \equiv X^q_1Y = LqR \). The remainder of the proof establishes the converse of the first implication, namely that if \( \Sigma \equiv X^q_1Y = LqR \) in \( B = B(T) \), then \( Xq_1T = q \) in \( \Gamma(T) \).

Now \( B_2 \) is an HNN extension of \( B_1 \ast Q \) with stable letters the \( \{r_i(i \in I)\} \) and it is easy to show that the elements \( \{r_i(i \in I), x\} \) freely generate a free subgroup of \( B_2 \) (see Lemma 31 below). So we may assume that \( L \) and \( R \) are freely reduced and hence \( r_i \)-reduced for \( i \in I \). But by Britton’s Lemma we must be able to successively transform \( LqR \) by carrying out \( r_i \)-pinches across what is between them to obtain a word which no longer involves \( r_i \) and is in fact equal in \( B_1 \ast Q \) to \( \Sigma \). In particular \( L^{-1} \) and \( R \) must be \( r_i \)-parallel in the sense that the sequences of \( r_i \) that appear must be the same.

Rewriting the equation \( \Sigma = LqR \) instead as \( L^{-1}\Sigma R^{-1} = q \), the sequence of successive \( r_i \)-pinches must act on \( \Sigma \equiv X^q_1Y \) eventually leading to the halting state \( q \). Intuitively this sequence of \( r_i \)-pinches just encodes a proof in Post’s semigroup \( \Gamma(T) \) that \( Xq_1Y =_{\Gamma(T)} q \). That is exactly what the remainder of our proof is devoted to showing.

So let us examine the first such \( r_i \)-pinch. We can write \( L^{-1} \equiv L_1r_i^{-\epsilon}x^{\alpha} \) and \( R^{-1} \equiv x^\beta r_i^x R_1 \) where \( \epsilon \) is \( \pm 1 \) and \( \alpha, \beta \) are integers and \( L_1, R_1 \) are the remainder of \( L^{-1} \) and \( R^{-1} \). We know here that \( r_i^{-\epsilon}x^{\alpha}X^q_1Yx^\beta r_i^x \) is a pinch and thus \( x^{\alpha}X^q_1Yx^\beta \) is equal in \( B_1 \ast Q \) to a word in the appropriate subgroup associated to \( r_i \) depending on \( \epsilon \).

Consider the case \( \epsilon = +1 \). We then have

\[
x^{\alpha}X^q_1Yx^\beta = (x^{\alpha}X^q_1F_i^{-1}F_i^x)q_0(G_i^{-1}Yx^\beta)
\]

where \( q_0 \) is \( q_1 \) and \( (G_i^{-1}Yx^\beta) \) and \( (x^{\alpha}X^q_1F_i^{-1}) \) are equal in \( B_1 \) to words in the elements \( s_{\alpha,x} \) corresponding to the (free) generators of the left hand associated subgroup in the relations \( r_i^{-1}s_{\alpha,x}r_i = s_{\alpha,x}^{-1} \).

We now prove an elementary but important lemma about certain words in \( B_1 \) which explains why the group \( B_1 \) plays a crucial role in the construction.

Lemma 29. Suppose that \( U \) and \( V \) are positive word in the \( s_{\alpha,x} \)-symbols and that \( U^{-1}V \) is freely reduced as written, that is, the last symbol of \( U^{-1} \) is not the inverse of the first symbol of \( V \). If the word \( U^{-1}Vx^\beta \) is equal in \( B_1 \) to a word in the elements \( s_{\alpha,x} \), then \( U \) must be empty. Similarly, if \( U^{-1}Vx^\beta \) is equal to a word in the elements \( s_{\alpha,x}^{-1} \), then \( U \) must be empty.

Proof. Note that the \( s_{\alpha,x} \)-symbols freely generate a free subgroup which is a retract of \( B_1 \). Suppose that \( U \) is not the empty word. If we write out \( U^{-1}Vx^\beta \) in detail it has the form

\[
U^{-1}Vx^\beta \equiv s_{\alpha_1}^{-1}s_{\alpha_2}^{-1}\cdots s_{\alpha_n}^{-1}s_{\beta_1}\cdots s_{\beta_m}x^\beta.
\]
Assume this is equal in \( B_i \) to a word in the \( s_b x \) which must have the same retraction onto the free group on \( s_b \). So we must have

\[
U^{-1} V x^\beta \equiv s_b x^{-1} \cdots s_b^{-1} s_{c_1} \cdots s_{c_p} x^\beta = B_i x^{-1} s_b \cdots x^{-1} s_b^{-1} s_{c_1} x \cdots s_{c_p} x.
\]

Equivalently this can be expressed as

\[
x^{-\beta} s_{c_p}^{-1} \cdots s_{c_1}^{-1} s_b \cdots s_b^{-1} x^{-1} s_b^{-1} \cdots s_{c_1} x s_{c_p} x = B_i 1.
\]

Now \( B_i \) is an HNN extension with stable letters the \( s_b \)-symbols so by Britton’s Lemma there must be an \( s_b \)-pinch. But by the assumptions on free reductions, the only place such a pinch could occur is at \( s_b x^{-1} s_b^{-1} \). But this is not a pinch since the relevant relation is \( x = s_b x^2 s_b^{-1} \) and \( x^{-1} \) does not lie in the subgroup generated by \( x^2 \). So we have a contradiction, proving the claim. The proof for equality to words in \( s_b x^{-1} \) is very similar with \( x^{-1} \) in place of \( x \).

Returning now to the proof of necessity in Boone’s Lemma, we have shown that \( (G_i^{-1} Y x^\beta) \) is equal in \( B_i \) to a word in the elements \( s_b x \). So by Lemma 29 if we freely reduce \( G_i^{-1} Y \) there are no negative symbols remaining. Hence \( G \) begins with the word \( G_i \) and \( Y \equiv G_i Y_1 \) where \( Y_1 \) is the remainder of \( Y \) and is a positive \( s \)-word. Similar considerations apply to \( (x^\alpha X^\# F_i^{\#} r_\epsilon^{-1})^{-1} = F_i^{\#} (X^\#)^{-1} x^{-\alpha} \) and we conclude that \( X^\# \equiv X^\# F_i^{\#} \) where \( X_1 \) is a positive \( s \)-word. That is, the \( r_i \)-pinch we are examining sends

\[
x^\alpha X^\# q_j Y x^\beta \equiv x^\alpha X^\# F_i^{\#} q_j G_i Y_1 x^\beta
\]

to

\[
x^{-\alpha} X^\# H_i^{\#} q_j K_i Y_1 x^{-\beta}
\]

which corresponds exactly to the rewrite \( X_1 F_i q_j G_i Y_1 \rightarrow X_1 H_i q_j K_i Y_1 \) in the semigroup \( \Gamma (T) \).

The case \( \epsilon = -1 \) is entirely analogous except that one uses Lemma 29 for words in \( s_b x^{-1} \) because the associated subgroup relations are \( s_b x = r_i s_b x^{-1} r_i^{-1} \).

It now follows inductively that the sequence of pinches is exactly parallel to the corresponding sequence of rewrites (including direction) in \( \Gamma (T) \) and so in particular \( X q_j Y = r_i (T) q \) as claimed. This completes the proof of necessity in Boone’s Lemma.

Given the recursive unsolvability of the halting problem for a suitable Turing machine \( T \) we have now constructed explicitly in terms of \( T \) a finitely presented semigroup \( \Gamma (T) \) with unsolvable word problem (following Post) and incorporated that into an explicit construction of a finitely presented group \( B(T) \) with unsolvable word problem (following Boone et al.).

It is sometimes useful that we know quite a lot about the structure of \( B(T) \) via the combinatorial group theory of HNN extensions. For instance it is quite easy to prove the \( B(T) \) is torsion free. We pause here to record another
property of $B(T)$ observed by Collins and Miller [39] which has interesting applications for decision problems in topology.

We first observe there is a somewhat different way to view $B_3$ as a tower of HNN extensions. Let $A$ be the group with presentation

$$A = \langle x, s_0, \ldots, s_M, q, q_0, \ldots, q_N, t \mid t^{-1}xt = x, s_0^{-1}xs_0 = x^2, b = 0, \ldots, M \rangle.$$ 

Then $A$ is an HNN-extension of $B_0 = \langle x \rangle$ with all the listed generators other than $x$ as stable letters. In particular, the associated subgroups are either cyclic or trivial and hence are finitely generated free groups. We now have the following easy fact.

**Lemma 30.** In the chain

$$B_0 \leq A \leq B_3,$$

each group is an HNN-extension of its predecessor; moreover, the associated subgroups are finitely generated free groups.

**Proof.** Let $F$ denote the free group on the stable letters of $A$ and $\phi : A \to F$ the retraction sending stable letters to themselves and $x$ to 1. One of the associated subgroups for $r_i$ is generated by the $M + 3$ elements \{ $F^n q_i, t, s_0x, \ldots, s_Mx$ \}. The image of this subgroup under $\phi$ is easily seen to be the (free) subgroup of $F$ generated by \{ $q_i, t, s_0, \ldots, s_M$ \} which has rank $M + 3$. Hence the associated subgroup is free on the given generators. Similar considerations show the other associated subgroup for $r_i$ is free. This completes the proof. \(\Box\)

We have previously made use of the following easy result.

**Lemma 31.** The elements \{ $x, r_i, i \in I$ \} freely generate a free subgroup of $B_3$.

**Proof.** Consider $B_3$ as an HNN-extension of $A$ and adopt the notation in the previous proof. If $w$ is a (non-empty) freely reduced word in $x$ and the $r_i$ and if $w = 1$ in $B_3$ then Britton’s Lemma implies $w$ contains a subword of the form $r_i^{-e}x^n r_i^e$ and $x^n$ belongs to one of the associated subgroups of $r_i$ depending on the sign of $e$. But $x \in \ker \phi$ while $\ker \phi$ intersects the associated subgroups in the identity. Thus $n = 0$ and $w$ is not freely reduced which is a contradiction. \(\Box\)

The result we want to state is the following improvement of the previous lemma.

**Lemma 32** (Collins–Miller). The elements \{ $q_i^{-1}tg, x, r_i, i \in I$ \} freely generate a free subgroup of $B_3$. Hence in the chain

$$B_0 \leq A \leq B_3 \leq B,$$

each group is an HNN-extension of its predecessor; moreover, the associated subgroups are finitely generated free groups. \(\Box\)

The proof of this is more difficult and uses some details of our arguments above as well as the determinism of Turing machines via our earlier Corollary 10. Intuitively, just as halting computations form a tree, so do semigroup
proofs of $Xq, Y = q$ and group theory proofs of $\Sigma = LqR$. This tree-like behaviour implies a uniqueness showing the subgroup is free. The reader is referred to [39] for more details. We record one consequence of this lemma as follows.

**Corollary 33** (Collins–Miller). The finitely presented group $B = B(T)$ having unsolvable word problem is of cohomological dimension 2. 

Later we explain how Lemma 32 combines with other results to show one cannot recognise asphericity for finite 2-complexes.

In the 1960s, L. A. Bokut’ introduced a more systematic study of towers of HNN extensions of the type found in the constructions of Boone and Novikov. He described certain normal forms in such groups, calling them groups with standard normal form. His methods are very useful in studying the word and conjugacy problems in such groups, and in obtaining recursively enumerable degree results. See the article by Bokut’ [13] and the monograph by Bokut’ and Kukin [14] for expositions of these methods and for versions of the Novikov and Boone constructions.

6.5. An explicit presentation for a group with unsolvable word problem. As for the case of semigroups, sometimes there is interest in writing down explicitly a finite presentation of a group having unsolvable word problem. Here we describe a reasonably small presentation provided by Collins [38] based on results of Tseitin and Borisov. The starting point is another small presentation of a semigroup with unsolvable word problem due to Tseitin [98] that is a variant of the one described above in Theorem 14. We call this presentation $C$ and it is given by:

Generators:

$$a, b, c, d, e.$$ 

Relations:

$$ac = ca, ad = da, bc = cb, bd = db.$$ 
$$ce = eca, de = edb, edca = cdcab.$$ 
$$eaaa = aaaa, daaa = aaaa.$$ 

Tseitin proved that the problem of deciding if $w \models_C aaa$, for arbitrary $w$, is unsolvable. The transition from this particular word problem for a semigroup presentation to the word problem for a group presentation is based on Boone’s construction described above but incorporating an elegant simplification of Boone’s approach by Borisov [24].

Applied to Tseitin’s presentation $C$. Borisov’s method yields the following presentation $B$ of a group:

Generators:

$$a, b, c, d, e, p, q, r, t, k.$$
Relations:
\[
p^{10}a = ap, \quad p^{10}b = bp, \quad p^{10}c = cp, \quad p^{10}d = dp, \quad p^{10}e = ep,
\]
\[
qa = aq^{10}, \quad qb = bq^{10}, \quad qc = cq^{10}, \quad qd = dq^{10}, \quad qe = eq^{10},
\]
\[
ra = at, \quad rb = br, \quad rc = cr, \quad rd = dr, \quad re = er.
\]
\[
pacr = rpcaq, \quad p^2adq^2r = rp^2daq^2,
\]
\[
p^3bcq^3r = rp^3cbq^3, \quad p^4bdq^4r = rp^4dbq^4,
\]
\[
p^5ceq^5r = rp^5ecaq^5, \quad p^6deq^6r = rp^6ed bq^6, \quad p^7cdeq^7r = rp^7cdceq^7,
\]
\[
p^8caaaq^8r = rp^8aaaaq^9, \quad p^9daaaq^9r = rp^9aaaaq^9.
\]
\[
p = tp, \quad qt = tq.
\]
\[
k(aaa)^{-1}t(aaa) = (aaa)^{-1}t(aaa)k.
\]

This presentation \( B \) has 27 relations among 10 generators which require 421 occurrences of a generator. Borisov proved that for any semigroup word \( w = w(a, b, c, d, e) \)
\[
k(w^{-1}tw) =_B (w^{-1}tw)k \text{ if and only if } w =_C aaa
\]
and the unsolvability of the word problem for \( B \) follows from Tseitin’s result for the semigroup \( C \).

§7. The isomorphism problem for groups and recognizing properties. We now apply the existence of a finitely presented group with unsolvable word problem to show that a wide variety of group theoretic questions cannot be answered algorithmically. First we prove the Theorem of Adian [2], [4], [3] and Rabin [84] that Markov properties of groups are not recursively recognisable. In particular, the problem of deciding of a finite presentation of a group whether or not the group is trivial is recursively unsolvable. Hence, the isomorphism problem for groups is recursively unsolvable.

Using a finitely presented group with unsolvable word problem, one can produce a finitely generated subgroup \( H \) of the direct product \( F \times F \) of two non-abelian free groups such that the problem of deciding membership in \( H \) is recursively unsolvable. Combining this with the Adian–Rabin Theorem, we show the problem of whether a finite subset of \( F \times F \) generates the whole group is recursively unsolvable. We also show that recognising certain standard properties of elements, such as being a commutator \([x, y]\), can be unsolvable in certain finitely presented groups.

7.1. The Adian–Rabin Theorem. By analogy to Markov’s results for semigroups we look more generally at abstract properties not depending on the presentation. The notion of a Markov property for groups is entirely analogous to the semigroup case.
Definition 34. An abstract property $P$ of finitely presented groups is said to be a Markov property if there are two finitely presented groups $G_+$ and $G_-$ such that:

1. $G_+$ has the property $P$; and
2. if $G_-$ is embedded in a finitely presented group $H$ then $H$ does not have property $P$.

These groups $G_+$ and $G_-$ will be called the positive and negative witnesses for the Markov property $P$ respectively.

It should be emphasized that if $P$ is a Markov property then the negative witness does not have the property $P$, nor is it embedded in any finitely presented group with property $P$.

For example the property of “being finite” is a Markov property. For $G_+$ one can take $\langle a \mid a^2 = 1 \rangle$ which is a finite group. For $G_-$ one can take the group $\langle b, c \mid b^{-1}cb = c^2 \rangle$ which is an infinite group and therefore not embedded in any finite group. The property “being infinite” is not a Markov property, but it is a co-Markov property in the sense that its complement is Markov.

An example of a property which is neither a Markov property nor a co-Markov property is the property of being perfect, that is $G/\langle [G, G] \rangle \cong 1$.

An abstract property $P$ of finitely presented groups is hereditary if $H$ is finitely presented and embedded in $G$ and $G \in P$ imply that $H \in P$, that is, the property $P$ is inherited by finitely presented subgroups. A property of finitely presented groups $P$ is non-trivial if it is neither the empty property nor is it enjoyed by all finitely presented groups. Suppose $P$ is a non-trivial, hereditary property of finitely presented groups. Then, since $P$ is non-trivial, there are groups $G_+ \in P$ and $G_- \notin P$. But if $G_-$ is embedded in a finitely presented group $H$, then $H \notin P$ because $P$ is hereditary. Thus $P$ is a Markov property with witnesses $G_+$ and $G_-$. This shows the following:

Lemma 35. If $P$ is a non-trivial hereditary property of finitely presented groups, then $P$ is a Markov property.

As discussed below, Graham Higman has constructed a universal finitely presented group, say $U$. If $P$ is a Markov property with positive and negative witnesses $G_+$ and $G_-$, then $G_-$ is embedded in $U$ so $U \notin P$. Moreover, if $U$ is embedded in a finitely presented group $H$ then so is $G_-$ and hence $H \notin P$. Thus $P$ is a Markov property with positive and negative witnesses $G_+$ and $U$. Hence $U$ is a negative witness for every Markov property.

An abstract property $P$ of finitely presented groups is recursively recognizable if there is an recursive method which when applied to an arbitrary finite presentation $\pi$ determines whether or not $gp(\pi)$ has the property $P$. That is, if $\{ \pi \mid gp(\pi) \in P \}$ is a recursive set of finite presentations (where all presentations are assumed to be on some fixed alphabet).
The main unsolvability result concerning the recognition of properties of finitely presented groups is the following:

**Theorem 36 (Adian–Rabin).** If \( P \) is a Markov property of finitely presented groups, then \( P \) is not recursively recognizable.

Before indicating a proof of this result, we note the following easy corollaries:

**Corollary 37.** The following properties of finitely presented groups are not recursively recognizable:

1. being the trivial group;
2. being finite;
3. being abelian;
4. being nilpotent;
5. being solvable;
6. being free;
7. being torsion-free;
8. being residually finite;
9. having a solvable word problem;
10. being simple; and
11. being automatic.

Each of (1) through (9) is a non-trivial, hereditary property and hence is a Markov property. For (10), it is known that finitely presented, simple groups have solvable word problem and hence a group with unsolvable word problem is a negative witness. Similarly for (11), automatic groups have solvable word problem and so being automatic is a Markov property.

**Corollary 38.** The isomorphism problem for finitely presented groups is recursively unsolvable.

For by (1) in the previous corollary there is no algorithm to determine of an arbitrary presentation \( \pi \) whether or not \( gp(\pi) \cong 1 \).

**Proof of the Adian–Rabin Theorem.** We are going to give a simple proof of the Adian–Rabin Theorem which is our modification [72] of one given by Gordon. The construction is quite straightforward and variations on the details can be applied to obtain further results. So suppose that \( P \) is a Markov property and that \( G_+ \) and \( G_- \) are witnesses for \( P \). We also have available a finitely presented group \( U \) having unsolvable word problem.

Using these three items of initial data, we construct a recursive family of finite presentations \( \{ \pi_w \mid w \in U \} \) indexed by the words of \( U \) so that if \( w =_U 1 \) then \( gp(\pi_w) \cong G_+ \) while if \( w \neq_U 1 \) then \( G_- \) is embedded in \( U \). Thus \( gp(\pi_w) \in P \) if and only if \( w =_U 1 \). Since \( U \) has unsolvable word problem, it follows that \( P \) is not recursively recognizable.

The family \( \{ \pi_w \mid w \in U \} \) is rather like a collection of buildings constructed from playing cards standing on edge. Such a building can be rather unstable...
so that if an essential card is removed (corresponding to \( w = v_1 \)) then the entire structure will collapse. The technical result we need provides such a construction and is somewhat independent of the present context.

**Lemma 39 (Main Technical Lemma).** Let \( K \) be a group given by a presentation on a finite or countably infinite set of generators, say

\[
K = \langle x_1, x_2, \ldots | r_1 = 1, r_2 = 1, \ldots \rangle.
\]

For any word \( w \) in the given generators of \( K \), let \( L_w \) be the group with presentation obtained from the given one for \( K \) by adding three new generators \( a, b, c \) together with defining relations

\[
\begin{align*}
a^{-1}ba &= c^{-1}b^{-1}cbc, \\
a^{-2}b^{-1}aba^2 &= c^{-2}b^{-1}cbc^2, \\
a^{-3}[w, b]a^3 &= c^{-3}bc^3, \\
a^{-(3+i)x_i}b(a^{3+i}) &= c^{-(3+i)}b(c^{3+i}), \quad i = 1, 2, \ldots
\end{align*}
\]

where \([w, b]\) is the commutator of \( w \) and \( b \). Then:

1. if \( w \neq K \) then \( K \) is embedded in \( L_w \) by the inclusion map on generators;
2. the normal closure of \( w \) in \( L_w \) is all of \( L_w \); in particular, if \( w = K \) then \( L_w \cong 1 \), the trivial group; and
3. \( L_w \) is generated by the two elements \( b \) and \( ca^{-1} \).

If the given presentation of \( K \) is finite, then the specified presentation of \( L_w \) is also finite.

**Proof.** Suppose first that \( w \neq K \). In the free group \( \langle b, c \rangle \) on generators \( b \) and \( c \) consider the subgroup \( C \) generated by \( b \) together with the right hand sides of the equations (1) through (4). It is easy to check that the indicated elements are a set of free generators for \( C \) since in forming the product of two powers of these elements or their inverses some of the conjugating symbols will remain uncancelled and the middle portions will be unaffected.

Similarly, in the ordinary free product \( K \ast \langle a, b \rangle \) of \( K \) with the free group on generators \( a \) and \( b \) consider the subgroup \( A \) generated by \( b \) together with the left hand sides of the equations (1) through (4). Using the assumption that \( w \neq K \) it is again easy to check that the indicated elements are a set of free generators for \( A \).

Thus assuming \( w \neq K \), the indicated presentation for \( L_w \) together with the equation identifying the symbol \( b \) in each the two factors is the natural presentation for the free product with amalgamation

\[
(K \ast \langle a, b \rangle) \ast \langle b, c \rangle, \quad A = C.
\]

So if \( w \neq K \), then \( K \) is embedded in \( L_w \) establishing the first claim.
Now let $N_w$ denote the normal closure of $w$ in $L_w$. Clearly $[w, b] \in N_w$ so by equation (3), $b \in N_w$. But equations (1) and (2) ensure that $a, b, c$ are all conjugate and so $a, b, c$ all belong to $N_w$. Finally, since each of the system of equations (4) can be solved to express $x_i$ in terms of $a, b, c$, it follows that $x_i \in N_w$ for $i = 1, 2, \ldots$. Thus each of the generators of $L_w$ belongs to $N_w$ and so $L_w = N_w$. This verifies the second assertion.

Finally, let $M$ be the subgroup of $L_w$ generated by $b$ and $ca^{-1}$. Equation (1) can be rewritten as $b(ca^{-1})b(ca^{-1})^{-1}b^{-1} = c$ so that $c \in M$. But then from $ca^{-1} \in M$ it follows that $a \in M$. Finally from the system of equations (4) which can be solved for the $x_i$ in terms of $a, b, c$ it follows that $x_i \in M$ for $i = 1, 2, \ldots$ and so $M = L_w$. (For later use we note that neither equation (2) nor equation (3) was used in the proof of the final assertion). This completes the proof of the lemma.

Using this technical lemma it is easy to complete the proof of the Adian–Rabin Theorem. We are given the three finitely presented groups $U$, $G_+$ and $G_-$ which can be assumed presented on disjoint alphabets as follows:

$$U = \langle y_1, \ldots, y_k \mid r_1 = 1, \ldots, r_p = 1 \rangle,$$
$$G_- = \langle s_1, \ldots, s_m \mid u_1 = 1, \ldots, u_s = 1 \rangle, \text{and}$$
$$G_+ = \langle t_1, \ldots, t_n \mid v_1 = 1, \ldots, v_t = 1 \rangle.$$

Let $K = U * G_-$ the ordinary free product of $U$ and $G_-$ presented as the union of the presentations of its factors. Since $U$ has unsolvable word problem, $K$ also has unsolvable word problem. Also both $U$ and $G_-$ are embedded in $K$ by the inclusion map on generators. For any word $w$ in the generators of $U$ (these are also generators of $K$) form the presentation $L_w$ as in the Main Technical Lemma. Finally we form the ordinary free product $L_w * G_+$.

A presentation $\pi_w$ for these groups $L_w * G_+$ can be obtained by simply writing down all of the above generators together with all of the above defining equations. Such a presentation is defined for any word $w$ in $U$ whether or not $w \neq U 1$. But it follows from the lemma that if $w \neq U 1$ then the group $G_-$ is embedded in $gp(\pi_w) = L_w * G_+$ and so $gp(\pi_w) \notin P$ by the definition of a Markov property. On the other hand, if $w = U 1$ then by the lemma $L_w \cong 1$ and so $gp(\pi_w) \cong G_+$ and hence $gp(\pi_w) \in P$.

Thus we have shown that the recursive collection of presentations

$$\{ \pi_w \mid w \text{ a word in } U \}$$

has the property that $gp(\pi_w) \in P$ if and only if $w = U 1$. Since $U$ has unsolvable word problem, it follows that $P$ is not recursively recognizable. This completes the proof of the Adian–Rabin Theorem.

We remark that among other consequences of the above technical lemma (Lemma 39) is the theorem of Higman. Neumann and Neumann [50] that any countable group $K$ can be embedded in a two generator group.
7.2. Membership and generation problems. Suppose $G$ is a finitely presented group and $H$ is a finitely generated subgroup given by a finite set of generating words. The membership problem or generalised word problem for $H$ in $G$ is to determine if an arbitrary word $w$ in the generators of $G$ represents an element of $H$, that is, whether or not $w$ is equal to some word in the given generators of the subgroup $H$. Notice that the word problem is just the membership problem for the trivial subgroup generated by $1 \in G$.

Using a group with unsolvable word problem, Mikhailova [69] shows the membership problem is unsolvable in what at first seem rather elementary groups.

**Theorem 40 (Mikhailova).** Let $F$ be a non-abelian free group of finite rank. Then the direct product $F \times F$ has a finitely generated subgroup $H$ for which membership is unsolvable.

**Proof.** Let $F = \langle a_1, \ldots, a_n \mid \rangle$ be a non-abelian free group and let $U = F = \langle a_1, \ldots, a_n \mid r_1 = 1, \ldots, r_m = 1 \rangle$ be a finite presentation of a group with unsolvable word problem (on the same set of generating symbols). We let $\phi : F \to U$ denote the quotient homomorphism from $F$ onto $U$. Now form the direct product of two copies of $F \times F$ for which one can easily write down a finite presentation.

Let $H \subseteq F \times F$ be the pullback or fibre product of two copies of $\phi$. Then

$$H_U = \{ (x, y) \in F \times F \mid \phi(x) = \phi(y) \} = \{ (x, y) \in F \times F \mid x =_U y \}.$$ 

It is a fairly straightforward exercise to see that a set of generators for $H_U$ is

$$S = \{ (a_1, a_1), \ldots, (a_n, a_n), (1, r_1), \ldots, (1, r_m) \}.$$ 

These pairs clearly belong to $H$. If $K$ is the subgroup they generate then the intersection of the normal closure in $K$ of the $(1, r_i)$ with the second factor is just ker $\phi$. And since the diagonal belongs to $K$ the intersection with the first factor is again ker $\phi$. So $K = H_U$. Now $(x, y) \in H_U$ if and only if $x =_U y$ and so the membership problem for $H_U$ is unsolvable. \qed

In [70] (see also [72]) the author combined this with the Adian–Rabin construction to show the following result on generation of groups.

**Theorem 41 (Miller).** Let $F$ be a non-abelian free group of finite rank. Then there is no algorithm to determine if an arbitrary finite set of elements in $F \times F$ whether or not they generate $F \times F$.

**Proof.** Let $F = \langle a_1, \ldots, a_n \mid \rangle$ be a non-abelian free group. Suppose the finitely presented group $U$ has unsolvable word problem. Let $L_w$ for $w \in U$ be as in the proof if Adian–Rabin for the Markov property “being trivial”. So $L_w \cong 1$ if and only if $w =_U 1$. Now each $L_w$ can be written as a presentation on two generators $a_1, a_2$ and we add more generators $a_3 = 1, \ldots, a_n = 1$ if $n > 2$ so that $\phi_w : F \to L_w$ is a surjective homomorphism. As in the previous proof, let $S_w$ be the specified finite generating set for the pullback $H_{L_w}$ of two
copies of $\phi_w$. Then $S_w$ generates $F \times F$ if and only if $H_{L_w} = F \times F$ if and only if $w = U$. But this last is a recursively unsolvable problem.

We remark that in case $w \neq U$ the subgroup $H_{L_w}$ is finitely generated but not finitely presented [47], [7] and has an unsolvable conjugacy problem [70]. In particular it cannot be isomorphic to $F \times F$ even as an abstract group. From this we conclude the following:

**Corollary 42.** There is no algorithm to determine of two finite sets of elements in $F \times F$ whether or not the subgroups they generate are isomorphic.

Recall that a free group $F$ has a faithful representation as a group of two-by-two integer matrices. So $F \times F$ is a subgroup of $SL(4, \mathbb{Z})$. Hence these results carry over to analogous statements for certain matrix groups and subgroups generated by finite set of matrices.

### 7.3. Some other decision problems about elements

We next briefly consider a few other local decision problems concerning elements in a group. The observations here are from the paper [6] by Baumslag, Boone and Neumann.

The structure of finitely generated abelian groups can be completely determined from a finite presentation of such a group, and in particular one can solve the word problem for such groups. Consequently, if $G$ is an arbitrary finitely presented group one can effectively determine the structure of its abelianization $G/[G, G]$. So for instance, there is an algorithm to decide whether $G$ is perfect, that is $G = [G, G]$. Moreover, since one can solve the word problem for $G/[G, G]$ it follows that one can decide of a arbitrary word $w$ of $G$ whether or not $w \in [G, G]$.

However, it seems that almost any property of elements of a finitely presented group which is not determined by the abelianization $G/[G, G]$ will be recursively unrecognizable. The following result shows a few common properties of elements are not recognizable.

**Theorem 43 (Baumslag, Boone and Neumann).** There is a finitely presented group $G$ such that there is no algorithm to determine whether or not a word in the given generators represents:

1. an element of the center of $G$;
2. an element which commutes with a given element of $G$;
3. an $n$-th power, where $n > 1$ is a fixed integer;
4. an element whose class of conjugates is finite;
5. a commutator; and
6. an element of finite order $> 1$.

**Proof.** Fix a finitely presented group $U$ having unsolvable word problem. Define $G$ to be the ordinary free product of $U$ with a cyclic group of order 3 and an infinite cyclic group, that is,

$$G = U \ast \langle s \mid t^3 = 1 \rangle.$$
We use the commutator notation \([x, y] = x^{-1}y^{-1}xy\). In the following, \(w\) is a variable for an arbitrary word in the generators of \(U\).

The center of \(G\) is trivial so \(w\) lies in the center of \(G\) if and only if \(w = U^{-1}\). So there is no algorithm to determine whether an arbitrary word of \(G\) lies in the center. This gives the first assertion. Similarly, \(w\) commutes with \(s\) if and only if \(w = U^{-1}\) which establishes the second assertion. The element \(s^n[w, w] = s^n\) is an \(n\)-th power if and only if \(w = U^{-1}\) since if \(w \neq U^{-1}\) the conjugates \(s^{-i}ws^i\) would all be distinct. This gives the fourth assertion. For the fifth assertion, note that \([s, t, w] = [s, t]w\) is a commutator if and only if \(w = U^{-1}\). Finally for the sixth assertion, observe that \(tw\) has infinite order if and only if \(w \neq U^{-1}\), while if \(w = U^{-1}\) then \(tw\) has order 3. This completes the proof.

\(\square\)

\section*{§8. The Higman Embedding Theorem.} Recall that a finitely generated, recursively presented group is one defined by a presentation \(H = \langle a_1, \ldots, a_n \mid R \rangle\) where \(R\) is a recursively enumerable set of defining relations. In contrast to the difficulties encountered for finitely presented groups, it is easy to give examples of finitely generated, recursively presented groups with unsolvable word problem. For example, let \(S \subset \mathbb{N}\) be a recursively enumerable set of natural numbers which is not recursive. Define the recursively presented group

\[ H_S = \langle a, b, t \mid t^{-1}a^{-i}ba^i = a^{-i}ba^i \forall i \in S \rangle. \]

Now \(H_S\) can be described as the HNN extension of the free group \(\langle a, b \rangle\) with associated subgroup (freely) generated by the \(a^{-i}ba^i, i \in S\). It follows from Britton’s lemma that the commutator \([t, a^{-i}ba^i] = H_S^{-1}\) if and only if \(i \in S\). Thus the word problem for \(H_S\) is recursively unsolvable.

Graham Higman observed that if one could embed this group \(H_S\) in a suitable finitely presented group, say \(G_S\) then, by Proposition 1, the group \(G_S\) would also have an unsolvable word problem. So finding such and embedding would give an alternative proof of the Novikov–Boone Theorem. In [49] Higman succeeded in showing such an embedding exists. Moreover, Higman gave a complete characterisation of the finitely generated subgroups of finitely presented groups—namely they are the recursively presented groups.

\textbf{Theorem 44 (Higman Embedding Theorem).} A finitely generated group \(H\) can be embedded in a finitely presented group if and only if \(H\) is recursively presented.

This is an important and, at the time, totally unanticipated result connecting group theory and recursive function theory.

That finitely generated subgroups of finitely presented groups are recursively presented is contained in our Proposition 1 above. The difficult part of this theorem is to show that a recursively presented group can be embedded in a finitely presented group.
As already pointed out, the Novikov–Boone Theorem is an easy corollary of Higman’s Embedding Theorem. Another consequence is the existence of universal finitely presented groups.

**Corollary 45 (Higman).** There exists a universal finitely presented group; that is, there exists a finitely presented group $G$ which contains an isomorphic copy of every finitely presented group.

To prove this one systematically enumerates all finite presentations on a fixed countable alphabet. Using the theorem of Higman, Neumann and Neumann, the free product of all of these can be embedded in a two generator group which will be recursively presented. This group can then be embedded in a finitely presented group which is the desired universal group.

Higman’s original proof of the Embedding Theorem has three stages. First he recasts the notion of an recursively enumerable set in more algebraic terms. He then introduces the following definition:

**Definition 46.** A subgroup $H$ of a finitely generated group $G$ is called benign in $G$ if the group $G \ast_H = \langle G, t \mid t^{-1}ht = h, h \in H \rangle$ is embeddable in a finitely presented group.

Of course $G \ast_H$ is the HNN extension in which the stable letter $t$ just commutes with the subgroup. Notice that if $H \subseteq G \subseteq K$ and if $H$ is benign in $K$, then $H$ is also benign in $G$.

Observe that if $H$ is benign in $G$, then the set of words in the generators of $G$ and their inverses that represent elements of $H$ is recursively enumerable. For if $H$ is benign in $G$ then $G \ast_H$ is embeddable in a finitely presented group and so we can enumerate the set of words $u$ in the generators of $G$ such that $[t, u] = 1$ which are just the elements of $H$. The main task in Higman’s proof is showing the following converse for the case of free groups.

**Lemma 47 (Higman).** A subgroup of a finitely generated free group is benign if and only if it is recursively enumerable.

The hard work then is to show that if $H$ is a recursively enumerable subgroup of a free group $F$, then the HNN extension $F \ast_H$ can be embedded in a finitely presented group. As previously noted this already implies the unsolvability of the word problem (Novikov–Boone Theorem) which was Higman’s starting point.

But having proved this lemma, he then gives a very brief but beautiful argument—dubbed the “Higman Rope Trick” in the monograph [56] by Lyndon and Schupp—which completes the full embedding result. I once asked Higman why he had placed this at the end of the paper since it seemed to explain why he did certain earlier steps. He responded that the paper was in the order he found the results, so he had proved the unsolvability of the word
problem before he realised this beautiful additional argument would prove the embedding and characterisation results. We include this argument here.

**Lemma 48 (The Higman Rope Trick).** *If $R$ is a benign normal subgroup of the finitely generated free group $F$, then $F/R$ is embeddable in a finitely presented group.*

**Proof.** By hypothesis the group $F*R = \langle F, t \mid t^{-1}rt = r, r \in R \rangle$ is embeddable in a finitely presented group $H$. Suppose the given generators of $F$ are $x_1, \ldots, x_n$. Applying Tietze transformations if necessary we can assume the $x_i$ are among the generating symbols of the given finite presentation of $H$. Let $\overline{F}$ be an isomorphic copy of $F$ with generators $\overline{x}_1, \ldots, \overline{x}_n$. If $w$ is a word in the $x_i$, we denote by $\overline{w}$ the corresponding word in the $\overline{x}_i$.

In $F*R$ the subgroup $L$ generated by $F$ and $t^{-1}Ft$ is isomorphic to their amalgamated free product with the subgroup $R = t^{-1}Rt$ amalgamated. Define a map $\phi : L \to \overline{F}/R$ by $\phi(w) = \overline{w}$ and $\phi(t^{-1}wt) = 1$. Since the definitions of $\phi$ agree on the amalgamated subgroup, $\phi$ defines a homomorphism.

Consider the direct product $H \times \overline{F}/R$ for which we will use ordered pair notation. Viewing $L$ as a subgroup of $H$, we can define a map $\psi : L \to L \times \overline{F}/R$ by $\psi(z) = (z, \phi(z))$ for $z \in L$. Clearly $\psi$ is an injective homomorphism. Hence we can form the HNN extension

$$K = \langle H \times \overline{F}/R, s \mid s^{-1}(z, 1)s = (z, \phi(z)) \text{ for } z \in L \rangle.$$

This is the required finitely presented group containing $\overline{F}/R$. It is visibly finitely generated. As defining relations we may take the relations of $H$, the relations of $\overline{F}/R$ (that is, $\overline{r} = 1$ for $\overline{r} \in \overline{R}$), the relations which say the generators of $H$ commute with the generators of $\overline{F}/R$, and the relations $s^{-1}(z, 1)s = (z, \phi(z))$ for $z$ in a set of generators of $L$. Each of these collections of relations is finite except for the relations of $\overline{F}/R$, so that it is sufficient to show these are redundant. For any word $w$ on the generators of $F$, by the relations for conjugation by $s$ and commutativity of $H$ and $\overline{F}/R$ we have

$$s^{-1}(w, 1)s = (w, \overline{w}) \text{ and } s^{-1}(t^{-1}wt, 1)s = (t^{-1}wt, 1).$$

But if $r \in R$, the relations of $H$ imply that $r = t^{-1}rt$ and so applying the previous equations we have

$$(r, 1) = s^{-1}(r, 1)s = s^{-1}(t^{-1}rt, 1)s = (t^{-1}rt, 1) = (r, 1).$$

Hence $\overline{r} = 1$. That is the relations of $\overline{F}/R$ are consequences of the remaining relations and so are redundant, as claimed. This completes the proof. \(\square\)

In 1970 Aanderaa [1] found another proof of Higman’s Embedding Theorem based upon using suitable HNN extensions of Boone’s groups and their built-in Turing machines. Aanderaa’s proof can also be found in Rotman’s textbook [87]. Other proofs are to be found in Lyndon–Schupp [56] and the appendix to Shoenfield’s logic text [90].
§9. Decision problems in geometric topology. In 1958 Markov [66], [64] announced the unsolvability of the homeomorphism problem for \( n \)-manifolds for \( n \geq 4 \) and provided a brief sketch.

In the classic textbook by Seifert and Threlfall [89] there is a construction from an arbitrary finite presentation \( G = \langle X \mid R \rangle \) of an \( n \)-manifold \( M^n \) for \( n \geq 4 \) with \( \pi_1(M^n) \cong G \) roughly as follows. First construct in the manner we described before a finite 2-complex \( Y \) using the presentation with \( \langle X \mid R \rangle \) with a single 0-cell, 1-cells corresponding to the generators and 2-cells corresponding to the relators so that fundamental group of \( Y \) is \( G \). Then \( Y \) can be embedded in Euclidean space \( \mathbb{R}^{n+1} \) and we take as \( M^n(X; R) \) the boundary of a suitable neighbourhood of the embedded image of \( Y \). Then \( M^n(X; R) \) is a closed (i.e., compact, no boundary) \( n \)-manifold and one can check that the fundamental group of \( M^n(X; R) \) is isomorphic to \( G \).

By the proof of the Adian–Rabin Theorem given above we know there is a recursive collection of finite presentations \( \langle X_i \mid R_i \rangle \) all having the same set of generators and the same number of defining relations—indeed all except one defining relation are the same for all the presentations since they have the form

\[
P(w) = \langle x_1, \ldots, x_n \mid r_1(w) = 1, r_2 = 1, \ldots, r_m = 1 \rangle
\]

where \( w \) is a word which is the only parameter. (What is important here is that the number of generators and the number of relations are constant.) Recall that there is an auxiliary group \( U \) having unsolvable word problem and \( w \) is also word of \( U \) and \( gp(P_w) \cong 1 \) if and only if \( w =_U 1 \). So the triviality problem for \( P(w) \) is unsolvable. So “being simply connected”, that is, having trivial fundamental group, is not recursively recognisable for 2-complexes or \( n \)-manifolds.

**Theorem 49.** The problem of deciding whether or not a finite 2-complex is simply connected is recursively unsolvable. Similarly the problem of deciding whether or not a closed \( n \)-manifold \( n \geq 4 \) is simply connected is recursively unsolvable.

Markov modifies the above construction by adding a fixed number \( t \) of copies of the trivial relation \( 1 = 1 \) to each of the presentations so the presentations become

\[
P_{n,m,t}(w) = \langle x_1, \ldots, x_n \mid r_1(w) = 1, r_2 = 1, \ldots, r_m = 1, 1 = 1 \ (t \text{ times}) \rangle.
\]

The point adding these trivial relations is that if \( P_{n,t} \) and \( P_{m,t} \) are isomorphic then there is a sequence of “restricted” Tietze transformations on the presentations which never changes the quantity \( m + t - n \). Here the Tietze transformations are restricted to certain operations which (1) operate on relations by multiplying one by another. taking inverses, taking cyclic permutations and inserting or deleting inverse pairs but do not change the number of relations; and (2) add a new generator and a new relator setting the new generator equal.
to a word in the existing generators, or the inverse of this operation. The copies of the trivial relation provide room for the necessary adding of consequences or relations to transform $P_{n,m,t}(w)$ to another presentation $P'_{n',m',t'}(w')$ of the same group provided $m' + t' - n' = m + t - n$ and $t' \geq m + n'$ and $t' \geq m' + n$ (see [21] for more details).

Then Markov shows that for the manifolds $M^n(P_{n,m,t}(w))$ constructed from the presentations $P_{n,m,t}(w)$, if the groups are isomorphic, then one manifold can be obtained from the other by a corresponding sequence of homeomorphisms involving sliding of handles. Of course if the fundamental groups are not isomorphic, the manifolds could not be homeomorphic. In light of the Adian–Rabin Theorem this proves the following:

**Theorem 50 (Markov).** The homeomorphism problem for closed $n$-manifolds with $n \geq 4$ is recursively unsolvable.

The thoughtful reader may be wondering whether this makes sense since there are certainly continuously many closed $n$-manifolds and in what manner are the ones we are interested in described. Also what is the situation for diffeomorphic as opposed to homeomorphic? Markov’s two announcements [66], [64] on the result provide a reasonable outline of his proofs. A third announcement [65] describes the analogous result for combinatorial equivalence.

In a 1968 paper [21], Boone, Haken and Poénaru dealt with these questions in considerable detail. They clarified what might be meant by presentation of a manifold and a differentiable structure. They then showed that Markov’s result holds for several different kinds of equivalence. We state their result as follows.

**Theorem 51 (Markov–Boone–Haken–Poénaru).** For each dimension $n \geq 4$, there is a recursive class $C$ of finite presentations of $n$-manifolds, endowed with a differentiable and a compatible combinatorial structure so that the problem of deciding for an arbitrary pair $M_1, M_2 \in C$ whether or not $M_1 \sim M_2$ is recursively unsolvable where $\sim$ is any of diffeomorphic, homeomorphic, combinatorially equivalent, or homotopy equivalent.

Their paper [21] is mainly concerned with recursively enumerable degree analogs of Markov’s unsolvability results, but it also provides details and foundational material on these questions.

One question left unanswered by Markov’s Theorem and construction is whether one can determine whether or not a closed $n$-manifold is homeomorphic to the $n$-sphere $S^n$. The manifolds given by the construction outlined above are never spheres since they have the wrong homology groups because the number of relations is much larger than the number of generators. Using universal central extensions and additional topological techniques S. P. Novikov (who is the son of P. S. Novikov) in 1962 proved the sphere recognition problem is unsolvable.
Theorem 52 (S. P. Novikov). The problem of determining whether or not a closed \( n \)-manifold with \( n \geq 5 \) is homeomorphic to the \( n \)-sphere \( S^n \) is recursively unsolvable.

A more recent exposition of the topological unsolvability results of Markov and S. P. Novikov can be found in the paper by Chernavsky and Leksine [31].

Recall that a finite complex \( Y \) is aspherical if its higher dimensional homotopy groups vanish, that is, \( \pi_n(Y) = 0 \) for \( n \geq 2 \), or equivalently if its universal covering space \( \tilde{Y} \) is contractible. As we mentioned earlier, by combining the stronger properties of Boone's groups and the above version of the Adian–Rabin construction together with results of Chiswell, Collins and Huebschmann [32], one can show [39] that asphericity of 2-complexes is not recognisable.

Theorem 53 (Collins–Miller). There is a recursive class of finite 2-complexes \( C \) such that the problem of determining whether or not a complex \( Y \in C \) is aspherical is recursively unsolvable.

This result follows from the constructions we have described and the fact that asphericity of presentations is preserved under the amalgamated free product construction and HNN extensions provided the associated subgroups are finitely generated free groups.

§10. Some connections with recursive function theory. As the Higman Embedding Theorem suggests, there are numerous connections between the word problem for groups and the theory of recursively enumerable sets. In his classic 1944 paper [81] Post introduced several notions of reducibility and in particular Turing reducibility: if \( A \) and \( B \) are two sets of objects, then \( A \leq_T B \) means that an (hypothetical) algorithm to answer questions about membership in \( B \) would yield an algorithm to answer questions about membership in \( A \). This can be made precise for subsets of the natural numbers \( \mathbb{N} \) in the following manner. The recursive functions can be defined by declaring that a certain collection of basic functions (addition, multiplication, projection) are recursive and then closing under the usual operations of composition, minimalization and recursion. A function is said to be \( B \)-recursive if it is among the functions obtained from the base functions together with the characteristic function of \( B \) by closing under the usual operations. Then \( A \leq_T B \) or \( A \) is Turing reducible to \( B \) is defined to mean the the characteristic function of \( A \) is \( B \)-recursive. Of course if \( B \) is already recursive (that is, membership in \( B \) is decidable) and \( A \leq_T B \), then \( A \) is also recursive.

Two sets (of natural numbers) are Turing equivalent \( A \equiv_T B \) if both \( A \leq_T B \) and \( B \leq_T A \). One can check that \( \equiv_T \) is an equivalence relation on the natural numbers \( \mathbb{N} \) and hence it defines a partition of the power set of \( \mathbb{N} \) into disjoint collections of subsets called (Turing) degrees of unsolvability. Those degrees of unsolvability which contain a recursively enumerable set are called r.e. degrees.
of unsolvability and are of particular interest. The r.e. degrees are partially ordered by $\leq_T$. There is a smallest r.e. degree denoted by $0$ which consists of the recursive sets and a largest r.e. degree denoted $0'$ which is the degree of the general halting problem for all Turing machines. In his 1944 paper [81] Post asked whether there are any more r.e. degrees, which is known as Post's problem. In the late 1950s this problem was solved in the affirmative independently by Friedberg and Muchnik (see the textbooks [86] or [91]) who introduced the priority method to show there is a pair of mutually incomparable r.e. degrees strictly between $0$ and $0'$. It is now known the r.e. degrees have a rich structure: for example, they are dense with respect to the partial order $\leq_T$.

These concepts can, as usual, be carried over to the realm of finitely presented groups by using a suitable coding device or Gödel numbering. From our previous discussion we know that if $\mathcal{P}$ is a finite presentation of a group, then the set of words on the generators which are equal to 1 is recursively enumerable. Quite generally one can show that if $\mathcal{P}_1$ and $\mathcal{P}_2$ are two presentations on finite generating sets for the same group $G$, then set of words equal to 1 in $\mathcal{P}_1$ is Turing equivalent to the set of words equal to 1 in $\mathcal{P}_2$. So if we use the notation $WP(G)$ for the word problem for $G$, that is, the algorithmic problem of deciding which words are equal to 1, then the degree of unsolvability of the word problem for $G$ does not depend on the presentation. Similar considerations apply to the degree of the conjugacy problem for $G$, thought of as a set of pairs of conjugate elements in $G$, which we denote by $CP(G)$.

We observe that if $G$ is finitely presented then both $WP(G)$ and $CP(G)$ are r.e. degrees. From our previous discussion we know $WP(G) \leq_T CP(G)$. Moreover, if $H$ is a finitely generated subgroup of $G$, then $WP(H) \leq_T WP(G)$.

The solution of Post’s problem occurred around the time that Novikov and Boone proved there are finitely presented groups with unsolvable word problem. So it was natural to ask whether the word problem for groups could have any of the many intermediate r.e. degrees between $0$ and $0'$. Not so surprisingly the answer is affirmative, but there are considerable technical difficulties to overcome since one needs to understand the word problem in the groups in question rather thoroughly. There were a number of independent proofs of the result, all at about the same time in the mid 1960s.

**Theorem 54** (Bokut’ [9, 10], Boone [20], Clapham [33], Fridman [44, 45]). Let $D$ be an r.e. degree of unsolvability. Then there is a finitely presented group $G$ with word problem of degree $D$. In more detail, there is an explicit construction which, when applied to a Turing machine $T$ with halting problem of degree $D$, yields a finitely presented group $G(T)$ such that the $WP(G(T))$ is Turing equivalent to the halting problem for $T$.

The proofs of Bokut’. Boone and Fridman are all based on variants of the Post–Boone construction that we have discussed above. Clapham’s proof is
based on the proof of Higman’s Embedding Theorem, and in a subsequent paper he shows the embedding can be done in a “degree preserving” way.

**Theorem 55** (Clapham [34]). *If* \( H \) *is a finitely generated, recursively presented group, then* \( H \) *can be embedded in a finitely presented group* \( G \) *such that* \( WP(H) \cong_T WP(G) \).

Another natural question is whether the word and conjugacy problems must have the same degree, or more generally what r.e. degrees of unsolvability can they be? As we mentioned earlier, in [75] Novikov gave a construction of a finitely presented group with an unsolvable conjugacy problem which was different from and considerably simpler than his construction [76] for the word problem. In 1960 Fridman [43] showed that these groups of Novikov with unsolvable conjugacy problem have solvable word problem. So indeed the two problems can be different. Then in the mid 1960s Bokut’ and Collins independently showed there are finitely presented groups with solvable word problem but conjugacy problem of arbitrary degree.

**Theorem 56** (Bokut’ [11, 12], Collins [35], Fridman [43]). *Let* \( D \) *be an r.e. degree of unsolvability. Then there is a finitely presented group* \( G \) *with solvable word problem but conjugacy problem* \( D \).

Finally Collins showed that, in a suitable variant, the Post–Boone–Britton construction gives groups whose word and conjugacy problem have the same degree. So taking a free product with the groups of the previous theorem, one gets a complete picture.

**Theorem 57** (Collins [35]). *Let* \( D_1 \) *and* \( D_2 \) *be two r.e. degrees of unsolvability and assume* \( D_1 \leq_T D_2 \). *Then there is a finitely presented group* \( G \) *with word problem of degree* \( D_1 \) *and conjugacy problem of degree* \( D_2 \).

The explicit nature of the construction of groups with word problem equivalent to the halting problem for a Turing machine in Theorem 54 has useful consequences. Since “having a solvable word problem” is a hereditary property, we know from the Adian–Rabin Theorem that it is not recursively recognisable. As we indicated earlier, there are many interesting classes of groups for which the word problem can be solved. So one might hope there is some sort of unified approach to solving the word problem. But Boone and Rogers [23] combined a bit of recursive function theory with the explicit nature of Theorem 54 to observe the following.

**Theorem 58** (Boone–Rogers). *The set of all (Gödel numbers of) finitely presented groups with solvable word problem is \( \Sigma^0_3 \)-complete in the arithmetic hierarchy.*

As a consequence of this one can easily show the following statements which (negatively) answer questions raised by Whitehead, Church and Higman respectively.
**Corollary 59.**

1. There is no recursive enumeration of all finite presentations of groups having solvable word problem.
2. There is no single partial algorithm which solves the word problem for all finite presentations of groups having solvable word problem.
3. There is no universal solvable word problem group, that is, there is no group with solvable word problem which contains an isomorphic copy of every finitely presented group with solvable word problem.

Despite all these negative results about the word problem, Boone and Higman managed to give an algebraic characterisation of groups with solvable word problem. We have previously mentioned that finitely presented simple groups have solvable word problem. In fact somewhat more is true as we now sketch. Suppose that we have three groups $H \subseteq S \subseteq G$ where $G$ is finitely presented, $S$ is simple and $H$ is finitely generated. Notice there are no finiteness or recursion assumptions about $S$. We claim that $H$ has a solvable word problem. Fix a non-trivial element $1 \neq s_0 \in S$. For any word $w$ in the generators of $H$ we let $G_w$ be the presentation obtained from the given presentation of $G$ by adding the relation $w = 1$. We now enumerate the list $L$ of words equal to 1 in $G$ and the list $L_w$ of all words equal to one in $G_w$. Now if $w =_H 1$, then $w$ will appear in $L$. But if $w \neq_H 1$, then $s_0 \in L_w$ since $1 \neq w \in S$ and $S$ is simple. Exactly one of these must eventually occur enabling us to decide whether or not $w =_H 1$. This proves the claim which is half of the following Boone–Higman result.

**Theorem 60 (Boone–Higman [22]).** A finitely generated group has a solvable word problem if and only if it can be embedded in a simple subgroup of a finitely presented group.

Sometimes technical results and methods from recursive function theory find application to decision problems about groups. As one example, the author obtained the following which has a number of applications, for instance to the study of algebraically closed groups.

**Theorem 61 (Miller [71, 72]).** There is a finitely presented group $G$ with unsolvable word problem such that every non-trivial quotient group of $G$ has unsolvable word problem.

The proof uses a pair of disjoint recursively enumerable sets $A$ and $B$ which are recursively inseparable (from recursive function theory). The construction applies the Higman Embedding Theorem together with techniques from the proof of the Adian–Rabin Theorem.

A remarkable instance of unsolvability has recently been obtained by Myasnikov and Osin who were partly motivated by questions about the existence of groups with “generically hard” word problem. They define a finitely generated group $G$ to be algorithmically finite if there is no algorithm which
enumerates an infinite set of pairwise distinct elements of $G$. An algorithmically finite group $G$ which is infinite must have an unsolvable word problem (though it may not even be recursively presented).

Recall that a section of a group $G$ is a quotient group of a subgroup of $G$. It is easy to see that a finitely generated section of an algorithmically finite group is either finite or has an unsolvable word problem. So such a group must be a torsion group. (Of course the powers of an element of infinite order would be a recursively enumerable set of distinct elements.)

It is not known whether infinite, finitely presented torsion groups exist, so it is also unknown whether infinite, finitely presented algebraically finite groups exist. But Myasnikov and Osin show the following striking result.

**Theorem 62** (Myasnikov–Osin [73]). *There is an infinite, finitely generated recursively presented group $G$ which is algorithmically finite.*

They call such a group a “Dehn monster”. The proof of the theorem uses the construction of Golod and Shafarevich of an infinite torsion group combined with an analog of Post’s [81] simple set construction.

Finally we mention that generic case complexity, which we informally described in an earlier section, has been widely studied in connection with decision problems for groups. A partial motivation for these investigations has been the potential for applications to cryptography. But quite recently these notions have been studied in the context of recursive function theory by Jockusch, Schupp and Downey (see in [53]). This is an area of ongoing research.

§11. Hints about further directions and references. In this section we make some comments and suggestions for further reading and references concerning the material covered above. A particularly relevant group theory reference is the textbook by Rotman [87] which includes a lot of the material with more details and proofs in its final chapter.

A standard reference for combinatorial group theory is the monograph [59] by Magnus, Karrass and Solitar but it doesn’t deal with unsolvability results. Of more direct relevance for the present article is the classic monograph [56] by Lyndon and Schupp which contains a great deal of useful material as well as treatments of many of the results described above, but usually from a different point of view. Combinatorial group theory has in many ways evolved towards geometric group theory which was given great impetus by Gromov’s work [46]. The monograph [25] by Bridson and Haefliger has a wealth of material concerning geometric group theory up to about 1999. Algorithmic questions play a substantial role in geometric group theory which continues to be a very active area.

In about 1990, I wrote the survey article [72] which tries to describe the status of decision problems in group theory at the time. While it has some
unfortunate omissions (such as neglecting Bokut’s work), it hopefully gives a fairly broad picture of developments.

Of more specific relevance to the present article is the survey [5] by Adian and Durnev which contains a great deal of material concerning decision problems for both groups and semigroups, including a lot of proofs. It is recommended for further mathematical details about various topics, particularly concerning semigroups. Another useful work is the monograph [14] by Bokut’ and Kukin which discusses algorithmic questions in a wider variety systems including associative algebra, lie algebras and varieties. They also discuss the Novikov and Boone constructions and subsequent related results including Bokut’s standard normal form approach.

Finally I would like to thank Gilbert Baumslag, Don Collins, Rod Downey, Paul Schupp and Hamish Short for reading various drafts of this article and providing many useful comments, suggestions and corrections.

REFERENCES


TURING MACHINES TO WORD PROBLEMS


[51] K. A. Hirsch, Review of [74], *Mathematical Reviews* (1953), MR0052436 (14,618h) 20.0X.

[52] Review of [75], *Mathematical Reviews* (1956), MR0075196 (17,706a) 20.0X.


[63] Review of [76], *Mathematical Reviews* (1956), MR0075197 (17.706b) 20.0X.

384 CHARLES F. MILLER III


