

# Computability of Bass-Serre structures

Christian Perfect

Andrew Duncan

January 27, 2012

## Introduction

Grzegorzcyk [Grz53] introduced a hierarchy of classes of recursive functions delineated by the number of times unbounded recursion is used in the construction of their members. Following Rabin [Rab60], Cannonito [Can66] defined a notion of computability for groups in the context of this hierarchy. Cannonito and Gatterdam [CG73] [Gat73] showed that, when taking a free product with amalgamation or an HNN extension of group(s) represented by functions computable at a particular level of the hierarchy, one creates a group with corresponding functions computable at most one level higher.

We describe the Grzegorzcyk hierarchy and present a clear notation for defining functions within it, along with constructions of many useful elementary functions.

Section 1 defines the notation used in this paper. Section 2 defines the Grzegorzcyk hierarchy of computable functions and gives definitions and positions on the hierarchy for many useful functions. Section 3 repeats Serre's [Ser77] formulation of graphs, and Section 4 gives conditions for a tree to be computable at a certain level of the Grzegorzcyk hierarchy.

In section 5 we restate [Can66]'s definition of an  $\mathcal{E}^n$ -computable group and  $\mathcal{E}^3$ -computable index of the free group on countably many generators.

Section 6 consists of a statement and proof of the main result of this paper – that the fundamental group of a graph of  $\mathcal{E}^n$  groups is  $\mathcal{E}^{n+1}$ -computable, generalising the results of [CG73]. Additionally, we show that the fundamental group of a graph of finitely-generated  $\mathcal{E}^n$ -computable groups is also  $\mathcal{E}^n$ -computable, and give an example of a graph of  $\mathcal{E}^3$ -computable groups whose fundamental group is at least  $\mathcal{E}^4$ -computable.

Finally, Section 7 defines computability for Stallings' pregroups [Sta71] in the context of the Grzegorzcyk hierarchy, generalising the results above.

# 1 Building blocks

## 1.1 Notation

Expressions consist of operators, function names, variables and punctuation. When defining expressions or new notation,  $E \equiv E' := D$  means we will write  $E$  or  $E'$  to mean  $D$ .

## 1.2 Constructing functions

Let  $\mathbb{Z}_{\geq 0}$  denote the non-negative integers. A function  $f$ , taking  $n$  arguments and resulting in a single value, is a map,

$$f : \mathbb{Z}_{\geq 0}^n \rightarrow \mathbb{Z}_{\geq 0}.$$

Usually we will write the expression  $f(x_1, \dots, x_n)$  to mean “the application of  $f$  to the arguments  $x_1, \dots, x_n$ .”

Functions in general can take any number of arguments, so it is clearly impractical to write out definitions of each operation for each possible arity. When the arity of a function  $f$  doesn't matter, we will write it as  $f(\mathbf{x})$ , where  $\mathbf{x}$  can represent an arbitrarily big vector. When a function has to take at least a specific number of arguments  $y_1, \dots, y_n$ , and potentially more, we will write something of the form  $f(\mathbf{x}, y_1, \dots, y_n)$ .

### 1.2.1 The primitive functions

We will begin by defining the *primitive functions*, which are functions of one of the following three types:

- *The zero function.* The zero function is a nullary function which returns the number 0. For brevity we will write the digit 0 to mean the zero function, instead of using the function notation with brackets defined above.
- *The successor function.* The successor function  $s(x) := x + 1$  is a unary function which adds 1 to its input.
- *The projection functions.* For every  $n > 0$  and every  $1 \leq i \leq n$  there is a projection function  $P_n^i(x_1, \dots, x_n) := x_i$ .  $P_n^i$  is an  $n$ -ary function which returns its  $i^{\text{th}}$  argument.

There are two operations we shall use to create new functions from these primitives.

### 1.2.2 Composition

Given an  $n$ -ary function  $f$  and  $m$ -ary functions  $g_1, \dots, g_n$ , the *composition* operator  $C$  constructs a new  $m$ -ary function  $h$ , where

$$h(\mathbf{x}) = f(g_1(\mathbf{x}), \dots, g_n(\mathbf{x})) := C(f, g_1, \dots, g_n)(\mathbf{x}).$$

If  $f$  is unary, a function  $h$  constructed by composition of  $f$  with another function  $g$  is denoted  $h = f \circ g$  where

$$f \circ g(\mathbf{x}) \equiv f(g(\mathbf{x})) := C(f, g)(\mathbf{x}).$$

We will omit the generalised arguments for most function definitions, like so:

$$\begin{aligned} f \circ g &:= C(f, g), \\ f(g_1, \dots, g_n) &:= C(f, g_1, \dots, g_n). \end{aligned}$$

The mixing of functions of different arities can be dealt with systematically. For every  $n$ -ary function  $f$ , and every  $m > n$ , we can define an  $m$ -ary function  $f_m$ ,

$$f_m := C(f, P_m^1, \dots, P_m^n),$$

such that, for all  $x_1, \dots, x_m$ ,  $f_m(x_1, \dots, x_n, \dots, x_m) \equiv f(x_1, \dots, x_n)$ .

Substitution of functions for parameters can also be dealt with systematically, by composing with the projection function  $P_i^i$  everywhere except at the position which will take the value of the substituted function.

Suppose we have two  $n$ -ary functions  $f(x_1, \dots, x_n)$  and  $g(x_1, \dots, x_n)$ . The function

$$f(x_1, \dots, x_{i-1}, g(x_1, \dots, x_n), x_{i+1}, \dots, x_n),$$

which computes  $f$  with its  $i^{\text{th}}$  parameter replaced by  $g(x_1, \dots, x_n)$ , is constructed like so:

$$f(x_1, \dots, x_{i-1}, g(x_1, \dots, x_n), x_{i+1}, \dots, x_n) := C(f, P_n^1, \dots, P_n^{i-1}, g, P_n^{i+1}, \dots, P_n^n).$$

### 1.2.3 Recursion

A function  $h$  is constructed by *recursion* from functions  $f$  and  $g$  if

$$\begin{aligned} h(\mathbf{x}, 0) &:= f(\mathbf{x}), \\ h(\mathbf{x}, n+1) &:= g(n, h(\mathbf{x}, n), \mathbf{x}), \end{aligned}$$

in which case we write  $h = R(f, g)$ .

Note that the argument of recursion  $n$  can be made to be in any position by composition with the projection functions. If we have a function  $h(x_1, \dots, x_m, n)$ , we can define a new function,

$$f'(x_1, \dots, x_k, n, x_{k+1}, \dots, x_m) := C(f, P_{m+1}^1, \dots, P_{m+1}^k, P_{m+1}^{k+2}, \dots, P_{m+1}^m, P_{m+1}^{k+1}),$$

so that

$$f'(x_1, \dots, x_k, n, x_{k+1}, \dots, x_m) := f(x_1, \dots, x_m, n).$$

### 1.3 Further notation

#### 1.3.1 Parameters

A *parameter* is a variable which is used in the definition of a function, not during evaluation. For instance, in the projection function  $P_n^i$ ,  $n$  and  $i$  are parameters – there is a different function for each pair of values  $(n, i)$ . Parameters do not count towards the arity of a function.

#### 1.3.2 Iteration

A function which evaluates  $n$  iterations of a function  $f$ , where  $n$  is a parameter, can be constructed by composing  $f$  with itself  $n - 1$  times.

$$f^{(n)}(\mathbf{x}) := \underbrace{C(f, C(f, C(f, \dots C(f, f)))}_{n-1 \text{ compositions}}(\mathbf{x}).$$

#### 1.3.3 Infix operators

When using infix notation, binary operators are written between their operands.

For a binary function  $f_*$ , representing the operation  $*$ , we will use  $x*y$  to denote  $f_*(x, y)$ , whenever the former notation is more conventional.

$$x * y := f_*(x, y).$$

#### 1.3.4 Sets

A set  $S \subseteq \mathbb{Z}_{\geq 0}^d$  is defined by its characteristic function  $\chi_S$ .

$$\chi_S(\mathbf{x}) := \begin{cases} 1 & \mathbf{x} \in S, \\ 0 & \mathbf{x} \notin S. \end{cases}$$

An *indexing* of a set  $X$  is an injective function

$$i : X \rightarrow \mathbb{Z}_{\geq 0}.$$

### 1.3.5 Relations

A  $d$ -ary relation on  $\mathbb{Z}_{\geq 0}$  is a subset  $R$  of  $\mathbb{Z}_{\geq 0}^d$ , so is defined by its characteristic function  $\chi_R$ . In addition, for binary relations we will write

$$xRy := \chi_R(x, y).$$

### 1.3.6 Logic

The concept of falsity will be represented by the number 0, and all other values will represent truth. With this in mind, predicates  $P(\mathbf{x})$  can be written out as ordinary functions once the logical operations have been defined.

## 2 The Grzegorzcyk hierarchy

A class  $\mathcal{C}$  is *closed under finite composition* if, whenever  $f, g_1, \dots, g_m$  all belong to  $\mathcal{C}$ , then  $C(f, g_1, \dots, g_m)$  also belongs to  $\mathcal{C}$ . In this case a function defined by finitely many compositions of elements of  $\mathcal{C}$  is also an element of  $\mathcal{C}$ .

A function  $f$  is said to be *bounded* by another function  $g$  if, for all  $\mathbf{x}$ ,  $f(\mathbf{x}) < g(\mathbf{x})$ .

A class  $\mathcal{C}$  is *closed under bounded recursion* if, whenever  $f, g, h$  belong to  $\mathcal{C}$  and  $R(f, g)(\mathbf{x})$  is bounded by  $h(\mathbf{x})$ , then  $R(f, g)$  also belongs to  $\mathcal{C}$ . If, for some  $f$  and  $g$ , there is no such  $h \in \mathcal{C}$  bounding  $R(f, g)$ , then  $R(f, g)$  is said to be obtained by an *unbounded recursion* on  $f$  and  $g$ .

The *Grzegorzcyk hierarchy* is an infinite nested hierarchy of classes  $\mathcal{E}^n$  of functions which are each closed under finite composition and by bounded recursion.

$\mathcal{E}^0$  consists of just the primitive functions, and whatever else can be obtained by finitely many applications of the composition operator and bounded recursion.

$\mathcal{E}^{n+1}$ , for  $n \geq 0$ , is the smallest class which

- contains all of  $\mathcal{E}^n$ ;
- contains every function obtained by a single unbounded recursion on  $\mathcal{E}^n$  functions;
- is closed under finite composition and bounded recursion.

**Definition 2.1.** We say a function  $f$  is  $\mathcal{E}^n$ -computable iff  $f$  belongs to  $\mathcal{E}^n$ . Similarly, a set  $S$  is said to be  $\mathcal{E}^n$ -decidable iff its characteristic function  $\chi_S$  is  $\mathcal{E}^n$ -computable.

An indexing  $i : X \rightarrow Z_{\geq 0}$  is said to be  $\mathcal{E}^n$ -decidable iff its image is  $\mathcal{E}^n$ -decidable.

It can be difficult to determine whether a particular function belongs to a particular level of the hierarchy. A definition of the Grzegorzcyk hierarchy equivalent to the one given above uses a sequence of hierarchy functions which bound the classes  $\mathcal{E}^n$ . We repeat Rose's [Ros84] definitions and Lemmas on the subject here.

**Definition 2.2.** [Ros84] The sequence of primitive recursive functions  $E_n$  ( $n = 0, 1, 2, \dots$ ) is defined by

$$\begin{aligned} E_0(x, y) &= x + y, \\ E_1(x) &= x^2 + 2, \\ E_{n+2}(0) &= 2, \\ E_{n+2}(x + 1) &= E_{n+1}(E_{n+2}(x)). \end{aligned}$$

**Definition 2.3.** [Ros84] Paraphrasing Rose, for each  $n \geq 0$ ,  $\mathcal{E}^{n+1}$  is the smallest class which contains  $\mathcal{E}^n$  plus the function  $E_n$ , and is closed under composition and bounded recursion.

**Lemma 2.4.** [Ros84, Lemma 2.2.5] Every function  $\mathcal{E}^0$ -computable function  $f$  satisfies

$$f(x_1, \dots, x_n) \leq x_i + k,$$

for some integers  $k$  and  $i$  where  $1 \leq i \leq n$ .

**Lemma 2.5.** [Ros84, Lemma 2.2.6] If  $f$  is an  $n$ -ary function in  $\mathcal{E}^1$  there are integers  $a_0, a_1, \dots, a_n$  such that

$$f(x_1, \dots, x_n) \leq a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \dots + a_n \cdot x_n.$$

That is, every  $\mathcal{E}^1$ -computable function is bounded by a linear combination of its arguments.

**Lemma 2.6.** [Ros84, Lemma 2.2.7] If  $f$  is an  $n$ -ary function in  $\mathcal{E}^2$  there is a polynomial  $p$  in  $n$  variables with the property

$$f(x_1, \dots, x_n) \leq p(x_1, \dots, x_n).$$

**Lemma 2.7.** [Ros84, Lemma 2.2.9] If  $f$  is a  $k$ -ary function in  $\mathcal{E}^n$  and  $n \geq 2$  then there is an integer  $m$  such that

$$f(x_1, \dots, x_k) \leq E_{n-1}^{(m)}(\max(x_1, \dots, x_k)).$$

## 2.1 The stockpile of functions

Below we will give explicit definitions for the operations of elementary arithmetic, plus some other useful things, using the notation defined above. These functions all appear, though not always in the same notation, in Kleene's list of primitive recursive functions, Sections 44 and 45 of [Kle52]. Here we have tried to use definitions which, though not necessarily the most obvious, place functions as low down on the hierarchy as possible.

## 2.2 $\mathcal{E}^0$ functions

For a constant integer  $n$ , we can construct a function to add  $n$  to  $x$  by iterating the successor function  $n$  times on  $x$ :

$$+_n(x) \equiv x + n := s^{(n)}(x). \quad (1)$$

Note that  $n$  is a parameter, that is, there is a different function  $+_n$  for each  $n$ .

A nullary function returning any single constant can be constructed by composing  $+_n$  with the zero function

$$c_n() := 0 + n. \quad (2)$$

From now on we will just write the number  $n$  to mean the function  $c_n$ , when it is not likely to lead to confusion.

The rest of this section consists of descriptions of operations and their formal definitions.

Decrement by 1:

$$\begin{aligned} dec(0) &:= 0, \\ dec(n+1) &:= n. \end{aligned} \quad (3)$$

Proper subtraction:

$$\begin{aligned} x \dot{-} 0 &:= x, \\ x \dot{-} (y+1) &:= dec(x \dot{-} y). \end{aligned} \quad (4)$$

$x \dot{-} y$  is bounded by  $P_2^1 \equiv x$ , so  $\dot{-}$  belongs to  $\mathcal{E}^0$ .

Signature:

$$\begin{aligned} sg(0) &:= 0, \\ sg(x+1) &:= 1. \end{aligned} \quad (5)$$

Reverse signature:

$$\begin{aligned} \overline{sg}(0) &:= 1, \\ \overline{sg}(x+1) &:= 0. \end{aligned} \quad (6)$$

Logical AND:

$$\begin{aligned} x \wedge 0 &:= 0, \\ x \wedge (n+1) &:= x. \end{aligned} \quad (7)$$

$\wedge$  is bounded by  $P_2^1 \equiv x$ , so is in  $\mathcal{E}^0$ .

Logical OR:

$$\begin{aligned} x \vee 0 &:= x, \\ x \vee (n+1) &:= 1. \end{aligned} \tag{8}$$

$\vee$  is bounded by  $s \circ P_2^1 \equiv x+1$ , so is in  $\mathcal{E}^0$ .

Logical NOT:

$$\neg x := \overline{sg}(x). \tag{9}$$

Ordering:

$$\begin{aligned} x > y &:= x \dot{-} y, \\ x < y &:= y \dot{-} x, \\ x \geq y &:= \neg(y < x), \\ x \leq y &:= \neg(x > y). \end{aligned} \tag{10}$$

Equality:

$$x = y := (x \leq y) \wedge (y \leq x) \tag{11}$$

The smallest of two numbers:

$$\min(x, y) := x \dot{-} (x \dot{-} y). \tag{12}$$

The remainder when dividing  $x$  by  $y$ :

$$\begin{aligned} 0 \bmod y &:= 0, \\ (x+1) \bmod y &:= \text{rm}'(y \dot{-} ((x \bmod y) + 1), x \bmod y). \end{aligned} \tag{13}$$

$$\begin{aligned} \text{rm}'(0, y) &:= 0, \\ \text{rm}'(x+1, y) &:= y+1. \end{aligned}$$

$\text{rm}'$  is bounded by  $y+1$  so belongs to  $\mathcal{E}^0$ .  $x \bmod y$  is defined by recursion on  $\mathcal{E}^0$  functions and is bounded by  $y$ , so belongs to  $\mathcal{E}^0$ .

### 2.3 $\mathcal{E}^1$ functions

Adding two numbers together is achieved by a recursion on the successor function.

$$\begin{aligned} x + 0 &:= x, \\ x + (y+1) &:= (x + y) + 1. \end{aligned} \tag{14}$$

Because every function in  $\mathcal{E}^0$  is bounded by a function of the form  $x+n$ , where  $n$  is a parameter, addition must belong to  $\mathcal{E}^1$  and not  $\mathcal{E}^0$ .



We can define another version of  $+_n$  where  $n$  is unbound:

$$+_n(y) \equiv f_+(n, y) := n + y. \quad (15)$$

This version is in  $\mathcal{E}^1$ . This is really an abuse of notation, but it will be useful in a little while.

The biggest of two numbers:

$$\max(x, y) := x + (y \dot{-} x). \quad (16)$$

Note that  $\max(x, y)$  is bounded by one of  $P_2^1$  or  $P_2^2$  but, since it could be either one of them, neither projection will serve as an  $\mathcal{E}^0$  bound on  $\max$ , so  $\max(x, y)$  only belongs to  $\mathcal{E}^1$  and not  $\mathcal{E}^0$ .

The absolute difference of two numbers:

$$|x - y| := (x \dot{-} y) + (y \dot{-} x). \quad (17)$$

Any constant multiple  $n \cdot x$  is produced by  $n$  iterations of addition:

$$n \cdot x \equiv 0 + x + \dots + x := +_x^{(n)}(0). \quad (18)$$

So we can also construct any positive- and constant-coefficient linear polynomial  $a_0 + a_1x_1 + a_2x_2 + \dots a_nx_n$  by finite compositions of  $\mathcal{E}^1$  functions.

Remember that the functions we are considering have domain  $\mathbb{Z}_{\geq 0}^d$ , so computations involving subtraction can go wrong. The key point is that proper subtraction is not associative with addition. For example,  $(4 + 3) \dot{-} 5 \neq 4 + (3 \dot{-} 5)$ . For polynomials with negative coefficients, computing all the positive terms and then subtracting the negative terms using proper subtraction will give the closest match to the actual value.

The quotient:

$$\begin{aligned} \left\lfloor \frac{0}{y} \right\rfloor &:= 0, \\ \left\lfloor \frac{x+1}{y} \right\rfloor &:= \left\lfloor \frac{x}{y} \right\rfloor + \overline{sg}((x+1) \bmod y). \end{aligned} \quad (19)$$

$x$  divides  $y$ :

$$x \mid y := \overline{sg}(x \bmod y). \quad (20)$$

Bounded minimisation:

$$\begin{aligned}
\min_{i < 0} P(\mathbf{x}, i) &:= 0, \\
\min_{i < y+1} P(\mathbf{x}, i) &:= \left( \min_{i < y} P(\mathbf{x}, i) \right) + sg \left( \left( \min_{i < y} P(\mathbf{x}, i) \right) = y \wedge (P(\mathbf{x}, y) = 0) \right).
\end{aligned} \tag{21}$$

Given a predicate  $P(\mathbf{x}, i)$  and a bound  $y$ , this returns the least  $i < y$  such that  $P(\mathbf{x}, i)$  is true. If there is no such  $i$ ,  $y$  itself is returned.

Bounded maximisation:

$$\begin{aligned}
\max_{i < y} P(\mathbf{x}, i) &:= y - \min_{i < y} P'(\mathbf{x}, y, i), \\
P'(\mathbf{x}, y, i) &:= P(\mathbf{x}, y - i).
\end{aligned} \tag{22}$$

If  $P$  is an  $\mathcal{E}^1$  function then so are bounded minimisations and maximisations over  $P$ , because they are defined by a recursion over  $\mathcal{E}^1$  functions, bounded by  $y$ .

**Lemma 2.8.** Let  $A = \{a_0, a_1, \dots\} \subseteq \mathbb{Z}_{\geq 0}$  be a strictly monotonically increasing integer sequence. Define a function  $\sigma_A : \mathbb{Z}_{\geq 0} \rightarrow A$  such that  $\sigma_A(n) = a_n$ .

If  $\sigma_A$  is  $\mathcal{E}^n$ -computable,  $n \geq 1$ , then its inverse  $\sigma_A^{-1} : a_n \mapsto n$  is also  $\mathcal{E}^n$ -computable.

*Proof.* Since  $A$  is monotonically increasing,  $a_n \geq a_0 + n \geq n$ . So we can construct  $\sigma_A^{-1}$  through bounded minimisation on  $\sigma_A$ .

$$\sigma_A^{-1}(n) := \min_{i < n+1} (\sigma_A(i) = n)$$

Constructed this way,  $\sigma_A^{-1}$  is  $\mathcal{E}^n$ -computable since the function performing the bounded minimisation is  $\mathcal{E}^1$ -computable and the equality predicate is  $\mathcal{E}^0$ -computable.  $\square$

Note that  $\sigma_A^{-1}$  has as its domain all of  $\mathbb{Z}_{\geq 0}$  but, since it is not being used to decide membership of  $A$ , we don't care what value it takes on inputs not in  $A$ .

## 2.4 $\mathcal{E}^2$ functions

Multiplication is an unbounded recursion on the addition operation:

$$\begin{aligned}
x \cdot 0 &:= 0, \\
x \cdot (y + 1) &:= (x \cdot y) + x.
\end{aligned} \tag{23}$$

We can calculate  $x^n$ , when  $n$  is constant, by  $n$  compositions of multiplication

by  $x$ , and hence any (positive-coefficient) polynomial can be obtained by finite compositions of addition and multiplication. So polynomials are  $\mathcal{E}^2$ -computable.

Finite sum:

$$\begin{aligned} S_f(\mathbf{x}, 0) &\equiv \sum_{i < 0} f(\mathbf{x}, i) := 0, \\ S_f(\mathbf{x}, y + 1) &\equiv \sum_{i < y+1} f(\mathbf{x}, i) := f(\mathbf{x}, y) + \sum_{i < y} f(\mathbf{x}, i). \end{aligned} \tag{24}$$

Note that the level of the sum in the hierarchy really depends on the level of  $f(\mathbf{x}, i)$ . However, we can say that if  $f$  is in  $\mathcal{E}^{n+2}$  then  $S_f(\mathbf{x}, y)$  is also in  $\mathcal{E}^{n+2}$  because it entails  $y - 1$  additions. If  $y$  is a bound variable and  $f$  is in  $\mathcal{E}^1$ , then we could even place the sum in  $\mathcal{E}^1$  by constructing the sum through  $y$  compositions of addition.

Bounded existential quantifier:

$$\exists_{i < y} P(\mathbf{x}, i) := \sum_{i < y} P(\mathbf{x}, i). \tag{25}$$

Definition by cases: if  $P_i(\mathbf{x})$ ,  $0 \leq i \leq k$  is a finite collection of mutually exclusive, exhaustive predicates, then we can define

$$f(\mathbf{x}) := \begin{cases} g_0(\mathbf{x}) & \text{if } P_0(\mathbf{x}), \\ \vdots & \vdots \\ g_k(\mathbf{x}) & \text{if } P_k(\mathbf{x}). \end{cases}$$

In our language, this is:

$$f(\mathbf{x}) := \sum_{i < k+1} g_i(\mathbf{x}) \cdot sg(P_i(\mathbf{x})), \tag{26}$$

where  $k$  is a parameter. If all the  $P_i$  and  $g_i$  are in  $\mathcal{E}^n$ ,  $n \geq 2$ , then so is  $f$ , because we are just doing  $k + 1$  multiplications and additions.

The integer square root function  $\lfloor x^{\frac{1}{2}} \rfloor$  computes the greatest integer  $n$  such that  $n^2 < x$ .

$$\lfloor x^{\frac{1}{2}} \rfloor := \max_{n < x} (n \cdot n < x)$$

$J(x, y)$  is an  $\mathcal{E}^2$ -computable integer-pairing function  $J : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}_{\geq 0}$  defined by

$$J(x, y) := ((x + y)^2 + x)^2 + y.$$

The functions  $L$  and  $R$  compute the inverses of  $J$ , that is, they have the properties  $LJ(x, y) := x$  and  $RJ(x, y) := y$ :

$$R(z) := z - \left\lfloor z^{\frac{1}{2}} \right\rfloor, \quad (27)$$

$$L(z) := R\left(\left\lfloor z^{\frac{1}{2}} \right\rfloor\right) = \left\lfloor z^{\frac{1}{2}} \right\rfloor - \left\lfloor \left\lfloor z^{\frac{1}{2}} \right\rfloor^{\frac{1}{2}} \right\rfloor. \quad (28)$$

Both  $L$  and  $R$  are  $\mathcal{E}^2$ -computable.

Although functions in the Grzegorzcyk hierarchy only operate on  $\mathbb{Z}_{\geq 0}$ , we will need the ability to work with negative integers and to add them together. The function  $\rho$  is an injection of  $\mathbb{Z}$  into  $\mathbb{Z}_{\geq 0}$ . We can then define operations on  $\rho(\mathbb{Z})$  to simulate operations on  $\mathbb{Z}$ .

$$\rho(x) := \begin{cases} 2n & \text{if } n \geq 0, \\ -2n - 1 & \text{if } n < 0. \end{cases}$$

The operation  $+_{\mathbb{Z}}$  simulates addition on  $\mathbb{Z}$ , so that if  $x, y \in \mathbb{Z}$ , then  $\rho(x + y) = \rho(x) +_{\mathbb{Z}} \rho(y)$ :

$$x +_{\mathbb{Z}} y = \begin{cases} x + y + (x \bmod 2) & \text{if } x \equiv y \bmod 2, \\ \max(x, y) \dot{-} \min(x, y) \dot{-} (\min(x, y) \bmod 2) & \text{if } x \not\equiv y \bmod 2. \end{cases}$$

The operation  $-_{\mathbb{Z}}$  simulates negation on  $\mathbb{Z}$ , so that  $\rho(-x) = -_{\mathbb{Z}} \rho(x)$ :

$$-_{\mathbb{Z}} x = \begin{cases} x - 1 & \text{if } x \equiv 0 \bmod 2, \\ x + 1 & \text{if } x \equiv 1 \bmod 2. \end{cases}$$

## 2.5 $\mathcal{E}^3$ functions

$\mathcal{E}^3$  is where just about all the really useful operations are.

First of all, we can perform exponentiation by unbounded recursion on the multiplication operation:

$$\begin{aligned} x^0 &:= 1, \\ x^{y+1} &:= x \cdot x^y. \end{aligned} \quad (29)$$

Factorial:

$$\begin{aligned} 0! &:= 1, \\ (x+1)! &:= (x+1) \cdot x!. \end{aligned} \quad (30)$$

Finite product:

$$\begin{aligned} P_f(\mathbf{x}, 0) &\equiv \prod_{i < 0} f(\mathbf{x}, i) := 1, \\ P_f(\mathbf{x}, y + 1) &\equiv \prod_{i < y+1} f(\mathbf{x}, i) := f(\mathbf{x}, y) \cdot \prod_{i < y} f(\mathbf{x}, i). \end{aligned} \quad (31)$$

If  $f$  is in  $\mathcal{E}^3$  then so is the product  $P_f(\mathbf{x}, y)$ , as it consists of  $y - 1$  multiplications. If  $y$  is a parameter, then we can define the product using  $y$  compositions of multiplication instead of the recursion, so if  $f$  is in  $\mathcal{E}^2$  then the series product is too.

Bounded universal quantifier:

$$\forall_{i < y} P(\mathbf{x}, i) := \prod_{i < y} P(\mathbf{x}, i). \quad (32)$$

Primality:

$$x \text{ prime} := (x > 1) \wedge \forall_{z < x} ((z < 2) \vee \neg(z \mid x)). \quad (33)$$

The first prime after  $x$ :

$$\text{nextprime}(x) := \min_{i \leq x!+1} ((i \text{ prime}) \wedge (i > x)). \quad (34)$$

Note that this is an exceptionally inefficient way of finding primes, in terms of the number of computations performed.

The  $n^{\text{th}}$  prime, starting at  $n = 0$ :

$$\begin{aligned} p_0 &\equiv p(0) := 2, \\ p_{n+1} &\equiv p(n + 1) := \text{nextprime}(p_n). \end{aligned} \quad (35)$$

The exponent of  $x$  in the decomposition of  $y$ :

$$[y]_x := \max_{i < y} (x^i \mid y). \quad (36)$$

## 2.6 Lists

We will be working with  $n$ -tuples, or lists, extensively. All of the list operations can be defined as  $\mathcal{E}^3$  functions on numbers by way of the Gödel encoding.

The Gödel encoding of an  $n$ -tuple:

$$\mathbf{x} \equiv [x_0, \dots, x_{n-1}] := \prod_{i < n} p_i^{x_i+1}. \quad (37)$$

The  $+1$  in the exponent is so we can reliably determine the length of a list by counting distinct prime factors (the empty list should be encoded as 1: then it

will have length 0). The bold  $\mathbf{x}$  is shorthand so we don't have to write out the brackets formation. It shouldn't be confused with the use of  $\mathbf{x}$  previously to mean an arbitrary number of parameters in a function application. An encoded list  $\mathbf{x}$  is just a single number.

The value of the  $i^{\text{th}}$  position in an encoded list:

$$(\mathbf{x})_i := [\mathbf{x}]_{p_i} - 1. \quad (38)$$

The length of an encoded list:

$$|\mathbf{x}| := \min_{i < \mathbf{x}} ([\mathbf{x}]_{p_i} = 0). \quad (39)$$

(Thus the length of  $[x_0, \dots, x_{n-1}]$  is  $n$ .)

The index of the first occurrence of a given value in an encoded list:

$$\text{find}(\mathbf{x}, y) := \min_{i < |\mathbf{x}|} (\mathbf{x})_i = y. \quad (40)$$

Note that  $\text{find}(\mathbf{x}, y)$  returns  $|\mathbf{x}|$  if  $y$  does not occur in the list.

Whether an encoded list contains a given value:

$$y \in \mathbf{x} := \text{find}(\mathbf{x}, y) < |\mathbf{x}|. \quad (41)$$

The largest element of an encoded list:

$$\begin{aligned} \text{biggest}(\mathbf{x}) &:= \text{biggest}'(\mathbf{x}, |\mathbf{x}| - 1), \\ \text{biggest}'(\mathbf{x}, 0) &:= (\mathbf{x})_0, \\ \text{biggest}'(\mathbf{x}, n + 1) &:= \max((\mathbf{x})_{n+1}, \text{biggest}'(\mathbf{x}, n)). \end{aligned} \quad (42)$$

To obtain the sub-list of  $\mathbf{x}$  starting at position  $s$  and ending at position  $e$ :

$$\mathbf{x}[s \dots e] \equiv \text{slice}(\mathbf{x}, s, e) := \prod_{i=s'}^{e'} (p_{i-s'})^{(\mathbf{x})_i+1}, \quad (43)$$

$$\begin{aligned} e' &:= \max(0, \min(e, |\mathbf{x}| - 1)), \\ s' &:= \max(0, \min(s, e')). \end{aligned}$$

To append a single value to the end of an encoded list:

$$\text{splice}(\mathbf{x}, n) := \mathbf{x} \cdot p_{|\mathbf{x}|}^{n+1}. \quad (44)$$

To concatenate two encoded lists:

$$\mathbf{x} \mathbin{++} \mathbf{y} := \text{splice}(\mathbf{x}, (\mathbf{y})_0) \mathbin{++} \mathbf{y}[1 \dots |\mathbf{y}| - 1]. \quad (45)$$

$\mathbf{x} \dot{+} \mathbf{y}$  is bounded by  $p_{|\mathbf{x}|+|\mathbf{y}|}^{(|\mathbf{x}|+|\mathbf{y}|)(\text{biggest}(\mathbf{x})+\text{biggest}(\mathbf{y}))}$ , so is  $\mathcal{E}^3$ -computable.

### 3 Graphs

**Definition 3.1.** Following [Ser77], a graph is a set of vertices  $V$  and a set of directed edges  $E = E^+ \cup E^-$ , with two associated maps. The first,

$$e \mapsto (\iota(e), \tau(e)),$$

associates each edge with its initial and terminal vertices. The second,

$$e \mapsto \bar{e},$$

is an involution on the edges, which must satisfy the conditions that  $\bar{\bar{e}} = e$  and  $\iota(\bar{e}) = \tau(e)$ ,  $\forall e \in E$ . We will write  $e_{ij}$  to mean the edge with  $\iota(e_{ij}) = i$  and  $\tau(e_{ij}) = j$ . For any pair of opposing edges  $e$  and  $\bar{e}$ , one is ‘positive’ and belongs to  $E^+$ , and one is ‘negative’ and belongs to  $E^-$ .

**Definition 3.2.** A *path* in a graph is a sequence of edges  $(e_1, e_2, \dots, e_n)$  such that  $\tau(e_i) = \iota(e_{i+1})$ ,  $1 \leq i < n$ . A path is *reduced* if it does not contain an edge followed by its inverse, that is, if  $e_{i+1} \neq \bar{e}_i$ ,  $1 \leq i < n$ .

**Definition 3.3.** A *circuit* is a path  $(e_0, \dots, e_n)$  with  $\iota(e_1) = \tau(e_n)$ .

**Definition 3.4.** A graph is *connected* if there exists a path between each pair of its vertices.

### 4 Computable Trees

**Definition 4.1.** A tree  $T = (V, E)$  is a connected, non-empty graph with no reduced circuits.

Following [Jan08], a *rooted* tree has a particular vertex  $v_0$  said to be at the ‘top’ and called the *root*. The children of a vertex  $v$  are those vertices adjacent to  $v$  and further from the root than  $v$ . In an *ordered* tree, the set of children of each vertex is ordered, so there is a least, or *leftmost*, child, and the other children are greater, or on its right.

**Definition 4.2.** An *ordered recursive tree* is a rooted, ordered tree obtained by a recursive process which begins with the root node and adds vertices below existing ones, in effect enumerating the edges.

We can assign a labelling to an ordered recursive tree by giving vertices increasing integer labels corresponding to the order in which they were added to the tree, beginning with 0 for the root. This way, the sequence of labels encountered on a path heading downwards is always increasing, and the root can always be recognised.

Every edge  $e_{ij}$  can be given a unique label  $\max(i, j)$ . If positive edges are defined to be the ones pointing downwards (away from the root) then the label of a positive edge  $e_{i,j}$  is always  $j$ . For all edges  $e$ , both  $e$  and  $\bar{e}$  have the same label.

**Definition 4.3.** Suppose we have a tree  $T = (V, E)$  with root vertex  $v_0$ .  $T$  is  $\mathcal{E}^n$ -computable if there is an  $\mathcal{E}^n$ -decidable labelling

$$i : V \rightarrow \mathbb{Z}_{\geq 0},$$

(we can insist that  $i(v_0) = 0$ ) and an  $\mathcal{E}^n$ -computable ‘parent function’,

$$\phi_T : i(V) \rightarrow i(V),$$

which gives the label of the parent of a vertex. For completeness, we also require that  $\phi_T(v_0) = i(v_0) = 0$  and allow  $\phi_T(v_0)$  to take any value on the complement of  $i(V)$ .

**Definition 4.4.** A finite ordered recursive tree can be entirely described in a canonical way by a *depth-first walk*. Note that every node has a single parent and a finite ordered set of children. Begin by ‘visiting’  $v_0$ . The process on ‘visiting’ a vertex  $v_i$  goes as follows: for each child node  $v_j$ , travel along  $e_{ij}$  and ‘visit’  $v_j$ , then travel backwards along  $\bar{e}_{ij}$ . The sequence of labelled edges travelled along by beginning this process at  $v_0$  is called the *depth-first walk* of the tree.

Note that the root vertex’s label will not appear in this sequence because it has no positive edge leading into it.

Because an edge and its inverse have the same label, the first occurrence of a label in the sequence represents the positive edge, and the second occurrence represents the trip back up the negative edge.

The depth-first walk can be encoded as a Gödel number by the method described in Section 2.6. We will define a function  $\phi_{\mathbf{t}}(n)$  which takes an encoded depth-first walk  $\mathbf{t}$  of a tree and gives the label of the parent of the vertex with label  $n$ .

## 4.1 Some functions to work with encoded trees

Let  $\mathbf{t}$  be an encoded depth-first walk of a finite ordered tree.

To find the first and second occurrences of the edge with label  $n$  in the walk:

$$\begin{aligned} \text{fst}(\mathbf{t}, n) &:= \text{find}(\mathbf{t}, n), \\ \text{snd}(\mathbf{t}, n) &:= \text{find}(\mathbf{t}[\text{fst}(\mathbf{t}, n) + 1 \dots |\mathbf{t}| - 1], n). \end{aligned} \tag{46}$$

The function  $\text{find}$  is  $\mathcal{E}^3$ -computable, so the functions  $\text{fst}$  and  $\text{snd}$  are  $\mathcal{E}^3$ -computable.



To get the label of the  $n^{\text{th}}$  child of the root:

$$\begin{aligned}\text{child}(\mathbf{t}, 0) &:= (\mathbf{t})_0, \\ \text{child}(\mathbf{t}, n+1) &:= (\mathbf{t})_{\text{snd}(\mathbf{t}, \text{child}(\mathbf{t}, n)) + 1}.\end{aligned}\tag{47}$$

The function  $\text{child}$  is constructed from  $\mathcal{E}^3$  functions and is bounded by the supremum of the labels of the vertices in the tree. As the tree is finite,  $\text{child}$  is  $\mathcal{E}^3$ -computable.

Note that the depth-first walk of the subtree descended from a vertex  $v_n$  is the subsequence found between the two occurrences of  $n$  in  $\mathbf{t}$ .

To find the subtree descended from the  $n^{\text{th}}$  child node of the root:

$$\text{subtree}(\mathbf{t}, n) := \mathbf{t}[\text{fst}(\text{child}(\mathbf{t}, n)) + 1 \dots \text{snd}(\text{child}(\mathbf{t}, n)) - 1].\tag{48}$$

The function  $\text{subtree}$  is constructed from  $\mathcal{E}^3$  operations so is  $\mathcal{E}^3$ -computable.

To obtain the number of children of the root node:

$$\text{kids}(\mathbf{t}) := \min_{i < \mathbf{t}} (\text{child}(\mathbf{t}, i) = 0).\tag{49}$$

The function  $\text{kids}$  is constructed by bounded minimisation on  $\mathcal{E}^3$  operations so is  $\mathcal{E}^3$ -computable.

To decide whether the root has a child with label  $n$ :

$$\text{haschild}(\mathbf{t}, n) := \left( \min_{i < \text{kids}(\mathbf{t})} (\text{child}(\mathbf{t}, i) = n) \right) < \text{kids}(\mathbf{t}).\tag{50}$$

Finally, we present the parent function  $\phi_{\mathbf{t}}$ . The idea is basically to do the depth-first walk, remembering which vertex you just came from. If you reach a subtree which doesn't contain  $n$  then turn back, or if you reach the required vertex then return the label of the vertex you just came from.

$$\begin{aligned}\phi_{\mathbf{t}}(0) &:= 0, \\ \phi_{\mathbf{t}}(n+1) &:= \mathbf{p}'(\mathbf{t}, n+1, 0).\end{aligned}\tag{51}$$

$$\mathbf{p}'(\mathbf{t}, n, p) := \begin{cases} 0 & n \notin \mathbf{t}, \\ p & \text{haschild}(\mathbf{t}, n), \\ \mathbf{p}'(\text{subtree}(\mathbf{t}, z), n, \text{child}(\mathbf{t}, z)) & \text{otherwise.} \end{cases}$$

$$z = z(\mathbf{t}, n) := \min_{i < \text{kids}(\mathbf{t})} (n \in \text{subtree}(\mathbf{t}, i)).$$

The function  $\mathbf{p}'$  is  $\mathcal{E}^3$ -computable because it is bounded by  $\mathbf{t}$ .

We will also need to know if the tree contains an edge  $e_{ij}$ :

$$e_{ij} \in \mathbf{t} \Leftrightarrow (\phi_{\mathbf{t}}(i) = j \vee \phi_{\mathbf{t}}(j) = i) \wedge i \neq j. \quad (52)$$

**Lemma 4.5.** A finite ordered tree is  $\mathcal{E}^3$ -computable.

*Proof.* As the tree is finite, of course its labelling is an  $\mathcal{E}^3$ -decidable subset of  $\mathbb{Z}_{\geq 0}$ , and  $\phi_{\mathbf{t}}$  as defined above is  $\mathcal{E}^3$ -computable.  $\square$

## 5 Computable Groups

AJD To make the relationship between the word problem and the level of computability clear some notion like Cannonito and Gatterdam's standard index is required. This section needs to incorporate such an idea. It would help to make the definition of standard index clearer than in C&G. Following [Can66], we say a group  $G$  is  $\mathcal{E}^n$ -computable if it has a triple of  $\mathcal{E}^n$ -computable functions  $(i, m, j)$ , such that

- $i : G \rightarrow \mathbb{Z}_{\geq 0}$  is an injection of  $G$  onto an  $\mathcal{E}^n$ -decidable subset of  $\mathbb{Z}_{\geq 0}$ ;
- $m : i(G) \times i(G) \rightarrow i(G)$  computes the product of two group elements, i.e.

$$m(i(g_1), i(g_2)) = i(g_1 g_2), \forall g_1, g_2 \in G;$$

- $j : i(G) \rightarrow i(G)$  computes the inverse of any element of  $G$ .

Note that  $i$  itself is not  $\mathcal{E}^n$ -computable but its image has  $\mathcal{E}^n$ -computable characteristic function.

The concept of a *standard index* for a group presentation is useful as a tool for restricting one's attention to group indices which "behave well". We found the definitions of [Can66] unclear so we reproduce the relevant portions of that paper below, rephrased to reduce ambiguity.

A standard index for a group  $G$ , with respect to a presentation  $\langle X \mid R \rangle$ , is defined using the standard index for the free group on  $X$ ,

**Definition 5.1.** Here we will define the *standard index* of the free group  $F(X)$ , where  $X = \{a_0, a_1, \dots\}$  is a countable set of generators.

The elements of  $F(X)$  are freely-reduced words of the form  $a_{i_0}^{\alpha_0} \dots a_{i_r}^{\alpha_r}$ , with all the  $\alpha_i$  positive or negative (but non-zero) integers.

A word can be represented as a list of pairs of generators and exponents, i.e. as  $\mathbf{w} = [(i_0, \alpha_0), \dots, (i_r, \alpha_r)]$ . By using the injection  $\rho$  of  $\mathbb{Z}$  into  $\mathbb{Z}_{\geq 0}$  and the pairing function  $J$ , this yields an  $\mathcal{E}^3$ -computable index of  $F(X)$ :

$$i(w) := [J(i_0, \rho(\alpha_0)), \dots, J(i_r, \rho(\alpha_r))].$$

Naturally, the empty word  $\epsilon$  is equivalent to the empty list, so  $i(\epsilon) := 1$ . From now on, we will identify the list  $\mathbf{w}$  with  $i(w)$ .

One of the conditions of computability is that the image of  $i$  should be decidable. Assuming that  $X$  is infinite, the only case in which a number does not represent a group element is when one of the exponents in the number's prime decomposition is not in the image of the pairing function  $J$ . This condition can be written down fairly straightforwardly:

$$x \in i(F(X)) \Leftrightarrow \forall_{i \leq |x|} (J(L((x)_i), R((x)_i)) = (x)_i)$$

If  $X$  is finite, another condition can easily be added to check that the left-hand sides of the generator-exponent pairs are all valid.

**CP** Hopefully it's clear how the above condition works.

The multiplication function is defined using list operations. For words  $w_1 = a_{i_0}^{\alpha_0} \dots a_{i_r}^{\alpha_r}$  and  $\mathbf{w}_2 = a_{j_0}^{\beta_0} \dots a_{j_s}^{\beta_s}$ ,

$$m(\mathbf{w}_1, \mathbf{w}_2) = \begin{cases} \mathbf{w}_1 \mathbin{++} \mathbf{w}_2 & \text{if } i_r \neq j_0, \\ \mathbf{w}_1[0 \dots r-1] \mathbin{++} [J(i_0, \alpha_r + \beta_0)] \mathbin{++} \mathbf{w}_1[1 \dots s] & \text{if } i_r = j_0 \wedge \alpha_r + \beta_0 \neq 0, \\ m(\mathbf{w}_1[0 \dots r-1], \mathbf{w}_2[1 \dots s]) & \text{if } i_r = j_0 \wedge \alpha_r + \beta_0 = 0. \end{cases}$$

Finally, we must define the inverse function,  $j$ .

$$j(\mathbf{w}) := \text{reverse}(\mathbf{w}), \tag{53}$$

$$\text{reverse}(1) := 1, \tag{54}$$

$$\text{reverse}(\mathbf{w}) := \text{reverse}(\mathbf{w}[1 \dots |\mathbf{w}|]) \mathbin{++} [J(L((\mathbf{w})_0), -\beta_0 R((\mathbf{w})_0))]. \tag{55}$$

All of  $i$ ,  $m$  and  $j$  are  $\mathcal{E}^3$ -computable, so every free group  $F(X)$  is  $\mathcal{E}^3$ -computable.

**Definition 5.2.** For a countably-generated presentation  $P = \langle X \mid R \rangle$  of a group  $G$  with  $\mathcal{E}^n$ -decidable word problem,  $n \geq 3$ , there is an associated  $\mathcal{E}^n$ -computable *standard index*  $(i_P, m_P, j_P)$ , defined as follows.

There is a surjection  $F(X) \xrightarrow{\pi} G$ .

Let  $K = \ker(\pi)$ , the set of words on  $X^*$  trivial in  $G$ .

Let  $(i, m, j)$  be the standard index on  $F = F(X)$ . As stated before, let  $i(K)$  be  $\mathcal{E}^n$ -decidable,  $n \geq 3$ .

First we define the predicate  $E(x, y)$  which decides whether two words  $x, y \in X^*$

are equivalent in  $G$ .

$$E(x, y) := x \in i(F) \wedge y \in i(F) \wedge m(j(x), y) \in i(K).$$

$E$ -equivalence classes of words in  $X^*$  correspond to the cosets of  $K$  in  $F(X)$ , and hence to elements of  $G$ . Given any word  $x \in X^*$ , we can find the least word  $y$  in its  $E$ -equivalence class by an  $\mathcal{E}^n$ -computable bounded minimisation operation  $r$ .

$$r(x) := \min_{y \leq x} (E(x, y)). \quad (56)$$

We will use this number as the index of the corresponding element of  $G$ .

Now the multiplication and inversion operations can be done on the free group index  $i(F(X))$ , and the minimisation  $r$  finds the corresponding element of  $i_P(G)$ .

$$i_P(x) := r \circ i(x), \quad (57)$$

$$m_P(x, y) := r(m(x, y)), \quad (58)$$

$$j_P(x) := r \circ j(x). \quad (59)$$

$(i_P, m_P, j_P)$  is the *standard index* of  $G$  with respect to  $\langle X \mid R \rangle$ , and is  $\mathcal{E}^n$ -computable.

Cannonito and Gatterdam [CG73] introduced the *relativised Grzegorzcyk hierarchy* relative to a set  $A$ , which includes the characteristic function of  $A$  as a primitive function.

**Definition 5.3.** [CG73, Definition 2.3] Let  $A \subset \mathbb{Z}_{\geq 0}$ . Then the class of functions  $\mathcal{E}^n(A)$  for  $n \geq 2$  (with domain  $\mathbb{Z}_{\geq 0}^k$  for arbitrary  $k$  and range  $\mathbb{Z}_{\geq 0}$ ) is the smallest class of functions, closed under composition and bounded recursion, containing all of  $\mathcal{E}^n$  in addition to

$$\chi_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \quad (60)$$

**Lemma 5.4.** Let  $A \subset \mathbb{Z}_{\geq 0}$ . If  $c_A$  is  $\mathcal{E}^m$ -computable, then any  $\mathcal{E}^n(A)$ -computable function,  $n \leq m$ , is  $\mathcal{E}^m$ -computable.

**Definition 5.5.** The *word problem*  $WP(G)$ , associated with a group  $G$  with generating set  $X$ , is the set of all words  $w \in X^*$  such that  $w =_G 1$ . That is, it is the set of all words representing the trivial element of  $G$ .

**Theorem 5.6.** [CG73, Theorem 3.2] For every countable group  $G$  there exists  $A \subset \mathbb{Z}_{\geq 0}$  such that  $G$  is  $\mathcal{E}^3(A)$ -computable. In particular, the word problem of  $G$  is one such  $A$ .

**Corollary 5.7.** Let  $G$  be a countable group with generating set  $X$ . Let  $i_F$  be the index of  $F(X)$  defined in 5.1.

$i_F(WP(G))$  is  $\mathcal{E}^n$ -decidable,  $n \geq 3$ , if and only if  $G$  is  $\mathcal{E}^n$ -computable.

**Definition 5.8.** Let  $H$  be a subgroup of an  $\mathcal{E}^n$ -computable group  $G$ .  $H$  is  $\mathcal{E}^m$ -decidable, with  $m$  necessarily larger than or equal to  $n$ , if the characteristic function of  $i(H) \subseteq i(G)$  is  $\mathcal{E}^m$ -computable.

**Definition 5.9.** Let  $G_1$  and  $G_2$  be  $\mathcal{E}^n$ -computable groups with  $\mathcal{E}^n$ -computable index functions  $i_1$  and  $i_2$ , respectively. A group homomorphism  $\phi : G_1 \rightarrow G_2$  is  $\mathcal{E}^n$ -computable with respect to  $i_1, i_2$  if there exists an  $\mathcal{E}^n$ -computable function

$$\hat{\phi} : i_1(G_1) \rightarrow i_2(G_2)$$

such that

$$\hat{\phi}(i_1(g)) = i_2(\phi(g)), \forall g \in G_1.$$

## 6 Computability of Bass-Serre groups

Let  $\Gamma = (V, E)$  be a connected graph.

**Definition 6.1.** Associate with each vertex  $v$  a vertex group  $G_v = \langle X_v \mid R_v \rangle$ , with the  $X_v$  all pairwise disjoint. Call the set of all vertex groups  $\mathbf{G}$ .

To each edge  $e \in E$  associate two isomorphic *edge groups*  $A_e \leq G_{i(e)}$  and  $B_e \leq G_{\tau(e)}$ , with  $A_e = B_{\bar{e}}$ , and an isomorphism  $\phi_e : A_e \rightarrow B_e$  satisfying  $\phi_e^{-1} = \phi_{\bar{e}}$ .

$(\mathbf{G}, \Gamma)$  is called a *graph of groups*.

**Definition 6.2.** Let  $T$  be a spanning tree of  $\Gamma$  (a subtree of  $\Gamma$  containing every vertex), with a root vertex  $v_0$ . Then the *fundamental group* [Ser77] of  $(\mathbf{G}, \Gamma)$ , with respect to  $T$  and  $v_0$ , is the group  $G = \pi_1(\mathbf{G}, \Gamma, T, v_0)$ , defined as follows.

The generators of  $G$  are

$$X = \left( \bigcup_{v \in V} X_v \right) \cup \left( \bigcup_{e \in E} \{e\} \right),$$

and the relations are

$$R = \left( \bigcup_{v \in V} R_v \right) \cup \{e^{-1}ae = \phi_e(a), \forall a \in A_e, \forall e \in E\} \cup \{e^{-1} = \bar{e}, \forall e \in E\} \cup \{e = 1, \forall e \in T\}.$$

$G$  can also be defined as the free product of all the vertex groups and the free group on the graph's edges minus the spanning tree, with the appropriate edge groups amalgamated.

$$G = \pi_1(\mathbf{G}, \Gamma, T, v_0) = (*_{v \in V} G_v) \underset{\phi_e}{*} F(E(\Gamma \setminus T)) \quad (61)$$

**Definition 6.3.** A word  $w \in X^*$  is *admissible* if it can be factored in the form

$$w = a_0 e_1 a_1 e_2 \dots a_{n-1} e_n a_n$$

where  $(e_1, \dots, e_n)$  is a circuit in  $\Gamma$  with vertex sequence  $(v_0, \dots, v_n)$ , such that  $v_0 = v_n$ ,  $v_i = \iota(e_{i+1}) = \tau(e_i)$ ,  $0 < i < n$ , and  $a_i \in G_{v_i}$ ,  $i = 0, \dots, n$ . The sequence  $(a_0, e_1, \dots, a_{n-1}, e_n, a_n)$  is an *admissible sequence* for  $w$ .

It will be more convenient from now on to refer to  $G_{v_i}$  by  $G_i$ .

**Lemma 6.4.** [KWM05, Proposition 2.4] Let  $w \in X^*$  be an admissible word, and suppose  $w =_G 1$ . Then either:

- $n = 0$  and  $a_0 = 1$  in  $G_{v_0}$ , or
- $n > 0$  and there exists  $0 < i < n$  such that  $e_{i+1} = \bar{e}_i$  and  $a_i \in B_{e_i}$ .

**Definition 6.5.** Let  $\phi_T(n)$  be the parent function for  $T$ . The path in  $T$  from  $v_i$  to  $v_j$  can be found by freely reducing the path  $\psi(i, j)$ , defined by:

$$\begin{aligned} \psi(0, 0) &:= \epsilon, \\ \psi(i, 0) &:= (e_{i, \phi_T(i)}) \cdot \psi(\phi_T(i), 0), \\ \psi(0, i) &:= \psi(i, 0)^{-1}, \\ \psi(i, j) &:= \psi(i, 0)\psi(0, j). \end{aligned} \tag{62}$$

The free reduction of  $\psi(i, j)$  will be implicit from now on.

What we really want is a function  $\pi : G \rightarrow X^*$  such that  $\pi(g)$  is an admissible word which represents  $g \in G$  and whose circuit begins and ends at  $v_0$ .

If  $g$  belongs to some  $G_{v_i}$ , then  $\pi(g) = \psi(0, i) \cdot g \cdot \psi(i, 0)$ . Otherwise, if  $g = e_{ij}$ , some edge letter, then  $\pi(e_{ij}) = \psi(0, i) \cdot e_{ij} \cdot \psi(j, 0)$ .

**Lemma 6.6.** Let  $w = x_1 \dots x_n$  be a word in  $X^*$ , representing the group element  $g \in G$ . Then by replacing each generator  $x_i$  by  $\pi(x_i)$  and freely reducing the resulting word, we obtain an admissible word starting at  $v_0$  which is equivalent to  $w$ . Call this word  $\pi(w)$ .

*Proof.* The new word is equivalent to  $w$  because the only letters we are adding are edge letters from the spanning tree, which are all trivial in the presentation of  $G$ . Also, it is admissible since each  $x_i$  in  $w$  is replaced by an admissible word with circuit starting and ending at  $v_0$ .  $\square$

**Theorem 6.7.** Suppose we have  $G = \pi_1(\mathbf{G}, \Gamma, T, v_0)$ , where  $\Gamma$  has  $\nu$  vertices. Suppose  $n \geq 3$ . Suppose all the  $G_i$  are finitely groups with  $\mathcal{E}^n$ -computable standard indices with respect to presentations  $\langle X_i \mid R_i \rangle$ . Assume all the edge groups are  $\mathcal{E}^n$ -decidable, and all the isomorphisms  $\phi_e$  are  $\mathcal{E}^n$ -computable. Then  $G$  is  $\mathcal{E}^{n+1}$ -computable.

*Proof.* We will show that the word problem of  $G$  is  $\mathcal{E}^{n+1}$ -decidable. We can assume that the generating sets of all the vertex groups  $G_v$  are disjoint, and that there is an  $\mathcal{E}^n$ -computable structure  $(i_v, m_v, j_v)$  associated with each  $G_v$ .

Encode the depth-first walk of the spanning tree  $T$  as a Gödel number  $\mathbf{t}$  per Section 4.1. Then the vertex-parent function  $\phi_{\mathbf{t}}$  is  $\mathcal{E}^3$ -computable. From now on each vertex group can be referred to by the integer label of the vertex it belongs to.

We will define  $i(F(X))$ , an index of the free group on the generators of  $X$ , and use that to decide the word problem of  $G$ . First, we must assign numbers to the generators of  $G$ . Let  $i_X(g)$  denote the number corresponding to the generator  $g$ . We will begin by using the first  $\nu^2$  numbers to represent potential edges. For  $0 \leq a, b \leq \nu - 1$  define  $i_X(e_{a,b}) := a \cdot \nu + b$ .

Each vertex group  $G_i$  has a generating set  $X_i$ . For each  $x_{i,j} \in X_i$ , define  $i_X(x_{i,j}) := \nu^2 + J(i, j)$ .

Now that every generator has been given an integer label,  $i(F(X))$  works the same way as the “standard” free group index described in [CG73, Lemma 3.1] – elements of  $F(X)$  are freely-reduced words, encoded as Gödel lists.

Given a word  $w$ , we wish to decide whether  $w =_G 1$ . First, compute  $w' = \pi(w)$ . We now need to split  $w'$  into an admissible sequence  $(a_0, e_1, \dots, e_n, a_n)$ .

AJD The argument needs tidying. You refer to  $w'$  here, but later you replace every letter  $w_i$  of  $w$  with  $\pi(w_i)$  – in fact if the previous sentence is deleted it’s easier to follow what’s going on.

To do this, we must decide for each letter in  $w'$  either which vertex group it is from or if it is an edge letter. AJD You may as well describe the equivalence class on  $X$ . Then you don’t have to bother with  $w'$  at this point, and the reintroduction of  $w$  below will cause no problem. Assign to each letter  $w_i$  of  $w'$  a code as follows: If  $w_i$  belongs to  $G_i$ , then its code is  $i$ . If it is an edge letter, then its code is  $\text{biggest}(\mathbf{t}) + 1$ . We can then define an  $\mathcal{E}^3$ -decidable equivalence relation  $\approx$  on the indices of the elements of  $G$ , whose equivalence classes correspond to the vertex groups plus one more for each edge letter. AJD One for the set of edges?

As  $\approx$  has finitely many equivalence classes, and they are all  $\mathcal{E}^3$ -decidable,  $\approx$  is  $\mathcal{E}^3$ -decidable.

Each letter  $w_i$  in  $w$  can then be replaced by  $\pi(w_i)$ . When working on the encoded version of  $w$ , this process is  $\mathcal{E}^3$ -computable, since  $\phi_{\mathbf{t}}$  is  $\mathcal{E}^3$ -computable, and we only want to go as far as producing a list of generators, not a single element of  $i(G)$ . By concatenating all of the  $\pi(w_i)$ , we will create an admissible word equivalent to  $w$ . AJD It looks from what follows as though you mean to freely reduce the product at this point.

The next task is to split  $w'$  into ‘syllables’, or contiguous subwords. A syllable is either a single edge letter or a word from one of the vertex groups.

From now on, let  $\mathbf{w}$  be an encoded admissible word. We will explain how to construct an encoded admissible sequence for  $\mathbf{w}$ .

Define a function which gives the position of the start of the syllable to which

$(\mathbf{w})_i$  belongs:

$$\begin{aligned} \text{backtrack}(\mathbf{w}, 0) &:= 0, \\ \text{backtrack}(\mathbf{w}, i+1) &:= \begin{cases} i+1 & (\mathbf{w})_{i+1} < \nu^2, \\ i+1 & (\mathbf{w})_i \not\approx (\mathbf{w})_{i+1}, \\ \text{backtrack}(\mathbf{w}, i) & (\mathbf{w})_i \approx (\mathbf{w})_{i+1}. \end{cases} \end{aligned} \quad (63)$$

The function  $\text{backtrack}$  is constructed from  $\mathcal{E}^3$ -computable operations and is bounded by  $|\mathbf{w}|$ , so is itself  $\mathcal{E}^3$ -computable.

Now, the start of the  $n^{\text{th}}$  syllable is given by:

$$\begin{aligned} \text{start}(\mathbf{w}, 0) &:= 0, \\ \text{start}(\mathbf{w}, n+1) &:= \min_{\text{start}(\mathbf{w}, n)+1 \leq i < |\mathbf{w}|} (\text{backtrack}(\mathbf{w}, i) \neq \text{start}(\mathbf{w}, n)). \end{aligned} \quad (64)$$

The number of syllables can be computed like so:

$$\text{numsyllables}(\mathbf{w}) = \min_{i < |\mathbf{w}|} (\text{start}(\mathbf{w}, i) = |\mathbf{w}|). \quad (65)$$

And now the  $n^{\text{th}}$  syllable itself can be found:

$$\text{syllable}(\mathbf{w}, n) = \mathbf{w}[\text{start}(\mathbf{w}, n) \dots \text{start}(\mathbf{w}, n+1) - 1]. \quad (66)$$

The function  $\text{syllable}$  is constructed from  $\mathcal{E}^3$ -computable functions and is bounded by  $\mathbf{w}$  so is  $\mathcal{E}^3$ -computable.

We can then use the conditions of Lemma 6.4 to determine if  $\mathbf{w}$ , encoded as  $\mathbf{w}' = (a_0, e_1, \dots, e_n, a_n)$ , is trivial.

If  $n = 0$ , then  $w' = 1 \Leftrightarrow a_0 = 1$ , which is an  $\mathcal{E}^n$ -decidable question.

If  $n > 0$ , then we need to find a sequence of the form  $e^{-1}A_e e$  or  $eB_e e^{-1}$ . The first of these is given by

$$\begin{aligned} \min_{i < \text{numsyllables}(\mathbf{w}') - 2} & ((\text{syllable}(\mathbf{w}', i) = e \in E) \wedge \\ & (\text{syllable}(\mathbf{w}', i+1) \in B_e) \wedge \\ & (\text{syllable}(\mathbf{w}', i+2) = \text{syllable}(\mathbf{w}', i)^{-1})) \end{aligned} \quad (67)$$

(and the same the other way round for  $A_e$ .) Note that since the first syllable must be an edge letter, we can say the last syllable is the inverse of the first by checking its length is 1, then computing its inverse, which is a matter of applying the inverses  $L$  and  $R$  of the pairing function  $J$ . We don't need to know how to compute the whole multiplication table to do this.

To decide if the middle syllable, which is a word on the encoded generators of some vertex group  $G_v$ , belongs to the appropriate edge group, it must be



rewritten using the original index and multiplication function of  $G_v$  provided in the setup, which is an  $\mathcal{E}^n$  operation. Once the word is in this form, membership of the edge group can be decided via an  $\mathcal{E}^n$  operation, by the assumptions of the theorem statement. **AJD** The middle syllable is a sequence of generators of  $G_v$ . It is given that  $B_e$  is  $\mathcal{E}^n$ -decidable, so each generator can be tested for membership of  $B_e$ . However, to test whether or not the product of the generators is in  $B_e$  needs a recursion using the multiplication function  $m_v$  for  $G_v$ . This recursion may not be bounded by a  $\mathcal{E}^n$  function.

If the result of that calculation is  $\text{numsyllables}(\mathbf{w}') - 2$  then  $w$  is not trivial. **AJD** Do you mean that if the result of applying (67) above is  $\text{numsyllables}(\mathbf{w}') - 2$  then  $w$  is not trivial? Otherwise, we can replace the found sequence  $eB_e e^{-1}$  with  $\phi_e(\text{syllable}(\mathbf{w}', i + 1))$ , and try again. The new word is still admissible and has fewer syllables than the original one, so repeated applications of this process will eventually lead to a word of one syllable or a negative answer. Because  $\phi_e$  might increase the index of the word, this recursion means the process is only computable on the next level of the Grzegorzczuk hierarchy.

So the word problem  $WP(G)$  is  $\mathcal{E}^{n+1}$ -decidable, and hence  $G$  is  $\mathcal{E}^{n+1}$ -computable by Corollary 5.7.  $\square$

Theorems [CG73, 4.6] and [CG73, 5.3] follow as corollaries of the above result, as free products with amalgamation and HNN extensions can be considered as the fundamental groups of graphs with one edge connecting two vertices and one vertex, respectively.

**Lemma 6.8.** The  $n$ th prime  $p_n \approx n \log n$ .

**Lemma 6.9.**  $\exp(x, y) := x^y \in \mathcal{E}^3$ .

**Lemma 6.10.** The Gödel encoding of any word of length  $n$  on a finite alphabet is bounded above by  $p_n^{J(n,n)}$ , an  $\mathcal{E}^3$  function.

**Corollary 6.11.** *In addition to the assumptions of Theorem 6.7, also assume that the edge groups are finitely generated: so each edge homomorphism  $\phi_{e_{ij}}$  sends a finitely generated subgroup  $E_{ij} \leq G_i$  to another finitely generated subgroup  $E_{ji} \leq G_j$ .*

*Then  $G = \pi_1(\mathbf{G}, \Gamma, T, v_0)$  is  $\mathcal{E}^n$ -computable.*

**AJD** To make this work the computation of the product of the elements of syllables above needs to be shown to be in  $\mathcal{E}^n$ . Probably the use of standard indices for the vertex groups would do the trick.

*Proof.* In the algorithm from Theorem 6.7, the potential for unbounded recursion comes from applying the edge homomorphisms repeatedly. We will show that in this case the result of applying the edge homomorphisms repeatedly is bounded by an  $\mathcal{E}^3$ -computable function.

Let  $w$  be a word on  $G$  of length  $n$ . An upper bound on the number of syllables in  $w$  is  $n$ , so let's say there are  $n$  applications of edge homomorphisms.

A homomorphism  $\phi_{e_{ij}}$  rewrites generators  $x \in X_i$  as words  $\phi_{e_{ij}}(x) \in X_j^*$ . As  $G_i$  is finitely generated, there is a word  $\phi_{e_{ij}}(x)$  of maximum, finite, length  $k_{ij}$ . So a word  $w_i \in E_{ij}$  of length  $m$  is rewritten to a word  $w_j \in E_j$  of length at most  $k_{ij}m$ .

Because the graph is finite, we can find a greatest  $k \in \{k_{ij}\}$ . So after reducing all the syllables (assuming that can be done for the input word), the resulting word in  $G_0$ , the group attached to the graph's root vertex, has length at most  $k^n n = L$ .

Hence the index of the resulting word is bounded by

$$p_{k^n n}^{J(k^n n, k^n n)},$$

an  $\mathcal{E}^3$  function, by Lemma 6.10. So recursion on the edge homomorphisms is bounded by an  $\mathcal{E}^n$  function, and so  $G$  is  $\mathcal{E}^n$ -computable. □

**Corollary 6.12.** *There exists a graph of  $\mathcal{E}^3$ -computable groups whose fundamental group is  $\mathcal{E}^4$ -computable.*

*Proof.* Let  $G_0 = \langle x, y \rangle$  and  $G_1 = \langle a, b \rangle$  be two copies of the free group on two generators. Let  $(\mathbf{G}, \Gamma)$  be the graph of groups with  $\mathbf{G} = \{G_0, G_1\}$  and  $\Gamma$  the graph with a single edge  $e$  from  $v_0$  to  $v_1$ ; so  $\Gamma = T = (\{v_0, v_1\}, \{e, \bar{e}\})$ .

Let  $E_{01} = \langle \{x^{-n}yx^n, n \in \mathbb{N}\} \rangle$ ,  $E_{10} = \langle \{a^{-n}ba^n, n \in \mathbb{N}\} \rangle$  be the (infinitely generated) identified subgroups of  $G_0$  and  $G_1$ , respectively. In fact the given generators of the edge subgroups freely generate these subgroups. This means that a homomorphism can be defined merely by defining the images of the generators.

We begin by defining the homomorphism  $\phi_e$  for the generators of  $G_0$  with even powers of  $x$ :

$$\phi_e(x^{-2n}yx^{2n}) := a^{-(n!)}ba^{n!}, \forall n \in \mathbb{N}.$$

Next we define a function to decide membership of the set of factorial numbers,  $\{d!, d \in \mathbb{Z}_{\geq 0}\}$ .

$$\begin{aligned} \text{fi}(n) &:= \text{fi}'(n, n). \\ \text{fi}'(n, 0) &:= 0, \\ \text{fi}'(n, d+1) &:= \begin{cases} d+1, & (d+1)! = n, \\ \text{fi}'(n, d), & (d+1)! \neq n. \end{cases} \end{aligned} \tag{68}$$

Helpfully,  $\text{fi}$  also acts as the inverse of the factorial operation, since it returns the integer  $d$  such that  $n = d!$ , when such a  $d$  exists.

The function  $\text{fi}(n)$  is  $\mathcal{E}^3$ -computable since it is bounded by  $n$  and is constructed from  $\mathcal{E}^3$ -computable functions.

In order for  $\phi_e$  to be an isomorphism, we need to identify the remaining generators of  $G_0$  (those with odd powers of  $x$ ) with the remaining generators of  $G_1$  (those with non-factorial powers of  $x$ ).

To do that, we need a bijection between  $\mathbb{N}$  and the non-factorial numbers.

Define

$$\begin{aligned} \text{nf}(1) &:= 3, \\ \text{nf}(n+1) &:= \begin{cases} \text{nf}(n) + 1, & \neg(\text{fi}(\text{nf}(n) + 1)), \\ \text{nf}(n) + 2, & \text{fi}(\text{nf}(n) + 1). \end{cases} \end{aligned} \quad (69)$$

Again,  $\text{nf}(n)$  is  $\mathcal{E}^3$ -computable. Its inverse is also  $\mathcal{E}^3$ -computable since  $\text{nf}^{-1}(n) = \min_{i < n} (\text{nf}(i) = n)$ .

Define

$$\phi_e(x^{-(2n-1)}yx^{2n-1}) := a^{-\text{nf}(n)}ba^{\text{nf}(n)}, \forall n \in \mathbb{N}.$$

Now we can define  $\phi_{\bar{e}}$  to be the inverse of  $\phi_e$ .

$$\phi_{\bar{e}}(a^{-n}ba^n) := \begin{cases} x^{-2 \cdot \text{fi}(n)}yx^{2 \cdot \text{fi}(n)}, & \text{fi}(n), \\ x^{-\text{nf}^{-1}(n)}yx^{\text{nf}^{-1}(n)}, & \neg \text{fi}(n). \end{cases}$$

Note that since  $G_0$  and  $G_1$  are  $\mathcal{E}^3$ -computable, the edge isomorphism is  $\mathcal{E}^3$ -computable in both directions.

Construct the index for  $G = G_0 *_{\phi_e} G_1$  by following the method in 6.7.

Consider a word  $w$  of the form

$$w = ea^{-1}e^{-1} \dots x^{-1}ea^{-1}e^{-1}x^{-2n}yx^{2n}eae^{-1}x \dots eae^{-1}.$$

In order to decide if this word is trivial, it must be simplified to something of the form  $x^{-m}yx^m$ . Start with the central subword  $x^{-2n}yx^{2n}$ . Each conjugation of  $x^{-2n}yx^{2n}$  by  $eae^{-1}x$  causes the number  $2n$  to be raised to (roughly) the factorial of itself. Since there can be arbitrarily many such conjugations, the final value of  $m$  will not be bounded by any  $\mathcal{E}^3$  function.

As  $m$  depends on the input, the length of the simplified word can only be computed by unbounded recursion on an  $\mathcal{E}^3$  function, so is  $\mathcal{E}^4$ -computable and not  $\mathcal{E}^3$ -computable.

Hence, the word problem of  $G$  is  $\mathcal{E}^4$ -computable, and by Corollary 5.7,  $G$  is  $\mathcal{E}^4$ -computable.  $\square$

## 7 Stallings' pregroups and their universal groups

We now turn to the notion of pregroups in the sense of Stallings, [Sta71], [Sta87]. A *pregroup*  $P = (P, D, m, \varepsilon, {}^{-1})$  consists of a set  $P$ , a distinguished element  $\varepsilon$ , a subset  $D \subset P \times P$ , equipped with a partial multiplication  $m : D \rightarrow P$ ,  $(a, b) \mapsto ab$ , and an involution (or *inversion*)  $P \rightarrow P$ ,  $a \mapsto a^{-1}$ , satisfying the following axioms for all  $a, b, c, d \in P$ . (By “ $ab$  is defined” we mean to say that  $(a, b) \in D$  and  $m(a, b) = ab$ .)

(P1)  $a\varepsilon$  and  $\varepsilon a$  are defined and  $a\varepsilon = \varepsilon a = a$ ;

(P2)  $a^{-1}a$  and  $aa^{-1}$  are defined and  $a^{-1}a = aa^{-1} = \varepsilon$ ;

(P3) if  $ab$  is defined, then so is  $b^{-1}a^{-1}$ , and  $(ab)^{-1} = b^{-1}a^{-1}$ ;

(P4) if  $ab$  and  $bc$  are defined, then  $(ab)c$  is defined if and only if  $a(bc)$  is defined, in which case

$$(ab)c = a(bc);$$

(P5) if  $ab, bc$ , and  $cd$  are all defined then either  $abc$  or  $bcd$  is defined.

It is shown in [Hoa88] that (P3) follows from (P1), (P2), and (P4), hence can be omitted. From now on, when we write  $ab = c \in P$ , we mean that  $a, b$  and  $c$  are in  $P$ ,  $(a, b) \in D$  and  $m(a, b) = c$ .

The *universal group*  $\mathbf{U}(P)$  of the pregroup  $P$  can be defined as the quotient monoid

$$\mathbf{U}(P) = P^* / \{ ab = c \mid m(a, b) = c \},$$

where  $P^*$  denotes the free monoid on  $P$ . Note that in  $\mathbf{U}(P)$  the identity element  $\varepsilon \in P$  is identified with the empty word  $1 \in P^*$ . The elements of  $\mathbf{U}(P)$  may therefore be represented by finite sequences  $(a_1, \dots, a_n)$  of elements of  $P$  such that  $a_i a_{i+1}$  is not defined in  $P$ , for  $1 \leq i < n$ , and  $a_i \neq \varepsilon$ , for all  $i$ . (The last condition is necessary only to exclude the sequence  $(\varepsilon)$ .) Such sequences are called  *$P$ -reduced* sequences. Since every element in  $\mathbf{U}(P)$  has an inverse, it is clear that  $\mathbf{U}(P)$  forms a group.

If  $\Sigma$  is any set, then the disjoint union  $P = \{\varepsilon\} \cup \Sigma \cup \bar{\Sigma}$ , where  $\bar{\Sigma} = \{\bar{a} : a \in \Sigma\}$  is a copy of  $\Sigma$ , yields a pregroup with involution given by  $\bar{\varepsilon} = \varepsilon$ ,  $\bar{\bar{a}} = a$ , for all  $a \in \Sigma$ , such that  $p\bar{p} = \varepsilon$ , for all  $p \in P$ . In this case the universal group  $\mathbf{U}(P)$  is nothing but the free group  $F(\Sigma)$ .

The universal property of  $\mathbf{U}(P)$  holds trivially, namely the canonical morphism of pregroups  $P \rightarrow \mathbf{U}(P)$  defines the left-adjoint functor to the forgetful functor from groups to pregroups.

### 7.1 Stallings' Theorem

Stallings [Sta71] showed that composition of the inclusion map  $P \rightarrow P^*$  with the standard quotient map  $P^* \rightarrow \mathbf{U}(P)$  is injective, where  $P^*$  is the free monoid

on  $P$ . His construction involves defining an equivalence relation, which we shall now describe, on the set of  $P$ -reduced sequences. First define the binary relation  $\sim$  on the set  $P$ -reduced sequences of elements of  $P$  by

$$(a_1, \dots, a_i, a_{i+1}, \dots, a_n) \sim (a_1, \dots, a_i c, c^{-1} a_{i+1}, \dots, a_n),$$

provided  $(a_i, c), (c^{-1}, a_{i+1}) \in D$ . Then Stallings' equivalence relation  $\approx$  is the transitive closure of  $\sim$ .

Canonical examples of pregroups arise from free products of groups with amalgamation and HNN-extensions (see [DDM10] for details).

The following is the principal result on the universal groups of pregroups.

**Theorem 7.1** (Stallings [Sta71]). *Let  $P$  be a pregroup. Then:*

- 1) *Every element of  $\mathbf{U}(P)$  can be represented by a  $P$ -reduced sequence;*
- 2) *any two  $P$ -reduced sequences representing the same element are  $\approx$  equivalent, in particular they have the same length;*
- 3)  *$P$  embeds into  $\mathbf{U}(P)$ .*

If  $\mathbf{a} = (a_1, \dots, a_i, a_{i+1}, \dots, a_n)$  is a word in  $P^*$  and  $a_i a_{i+1} = b \in P$  then we say that the word  $(a_1, \dots, b, \dots, a_n)$  is obtained from  $\mathbf{a}$  by *elementary  $P$ -reduction*. We also say that the empty word 1 is obtained from the word  $(\varepsilon)$ , of length one, by elementary  $P$ -reduction. If  $\mathbf{b}$  is obtained from a word  $\mathbf{a}$  by a finite sequence of elementary  $P$ -reductions then we say that  $\mathbf{b}$  is obtained from  $\mathbf{a}$  by  *$P$ -reduction* and that  $\mathbf{b}$  is a  *$P$ -reduction* of  $\mathbf{a}$ . (Note that, unless  $\mathbf{b} = 1$ , such a reduction can always be carried out without ever replacing an occurrence of  $\varepsilon$  with 1.) In particular the following lemma follows from statement 2 of Theorem 7.1.

**Lemma 7.2.** Let  $\mathbf{c} = (c_1, \dots, c_m)$  and  $\mathbf{d} = (d_1, \dots, d_n)$  be  $P$ -reduced words. Then  $\mathbf{c} = \mathbf{d}$  in  $\mathbf{U}(P)$  if and only if  $n = m$  and 1 is a  $P$ -reduction of the word  $(d_m^{-1}, \dots, d_1^{-1}, c_1, \dots, c_m)$ .

If  $\mathbf{a} \in P^*$  and both  $\mathbf{b}$  and  $\mathbf{c}$  are  $P$ -reduced,  $P$ -reductions of  $\mathbf{a}$  then it follows from Theorem 7.1 that  $\mathbf{b} \approx \mathbf{c}$ .

Rimlinger [Rim87a] (see also Rimlinger [Rim87b] and Hoare [Hoa88]) has shown that,

- given any graph of groups with fundamental group  $G$ , a pregroup  $P$  may be constructed such that  $\mathbf{U}(P) = G$ ; and conversely that
- given a pregroup  $P$  which satisfies a condition called “finite height” then a graph of groups with fundamental group  $\mathbf{U}(P)$  may be constructed.

Not all pregroups have finite height, so there exist universal groups of pregroups which are not the fundamental groups of any graph of groups.

## 7.2 Pregroups and the Grzegorzcyk hierarchy

Let  $P = (P, D, m, \varepsilon, {}^{-1})$  be a pregroup and let  $i$  be an injective function  $i : P \longrightarrow \mathbb{Z}_{\geq 0}$ . Define

- $D' = \{(x, y) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} : x = i(p), y = i(q), (p, q) \in D\}$ ,
- the function  $m' : D' \longrightarrow \mathbb{Z}_{\geq 0}$ , such that  $m'(i(p), i(q)) = i(m(p, q))$ , for all  $(p, q) \in D$ , and
- the function  $j$  given by  $j(i(p)) = i(p^{-1})$ , for all  $p \in P$ .

Then  $P$  is said to have *index*  $(i, m', j)$ .

**Definition 7.3.** The pregroup  $P$  is  $\mathcal{E}^n(A)$ -computable, with respect to index  $(i, m', j)$ , if

1.  $i(P)$  is  $\mathcal{E}^n(A)$ -decidable;
2.  $D'$  is  $\mathcal{E}^n(A)$ -decidable;
3.  $m'$  is  $\mathcal{E}^n(A)$ -computable and
4.  $j$  is  $\mathcal{E}^n(A)$ -computable.

*AJD Should results be stated for  $\mathcal{E}^n(A)$  or just  $\mathcal{E}^n$ ? From now on I'll drop  $A$ .*

**Theorem 7.4.** Let  $P$  be a  $\mathcal{E}^n$ -computable pregroup. Then the universal group  $\mathbf{U}(P)$  of  $P$  is  $\mathcal{E}^{n+1}$ -computable.

In order to prove this theorem we first introduce some functions and then use these to construct an index for  $\mathbf{U}(P)$ .

## 7.3 Functions for pregroups

In this section we shall use the notation and constructions of Section 2.6. As we shall mostly be working with encoded lists we use  $\mathbf{a}$  to represent both an encoded list  $\mathbf{a}$  or an arbitrary non-negative integer  $\mathbf{a}$ , as appropriate to the context. Let  $i : P \longrightarrow \mathbb{Z}_{\geq 0}$  be an injective function and assume that  $i(P)$  is  $\mathcal{E}^n$ -decidable. We shall assume that  $i(\varepsilon) = 1$ . Elements of  $P^*$  are regarded as finite sequences of elements of  $P$ , and usually referred to as *words*. Define an indexing function  $i_1 : P^* \longrightarrow \mathbb{Z}_{\geq 0}$  by

$$i_1(a_0, \dots, a_m) := [i(a_0), \dots, i(a_m)]. \quad (70)$$

Then  $i_1$  is injective; in particular, by definition of the Gödel number of the empty sequence,  $i_1(1) = 1$ , while, on the other hand,  $i_1(\varepsilon) = 4$ .

The function  $i_1$  is  $\mathcal{E}^n$ -decidable: the characteristic function of  $i_1(P^*)$  is the function  $\chi$  such that

$$\chi(\mathbf{a}) := \bigwedge_{s=0}^{|\mathbf{a}|-1} (\mathbf{a})_s \in i(P).$$

We also define functions corresponding to multiplication in  $P^*$  and to the extension of the inversion operation on  $P$  to  $P^*$ . That is we define  $m_1 : \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  and  $j_1 : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  by

$$m_1(i_1(\mathbf{u}), i_1(\mathbf{v})) := i_1(\mathbf{u}) \dot{+} i_1(\mathbf{v}) \quad (71)$$

and

$$j_1(i_1(u_0, \dots, u_m)) := i_1(j(u_m), \dots, j(u_0)). \quad (72)$$

Then  $m_1$  and  $j_1$  are  $\mathcal{E}^n$ -computable functions,  $m_1((i_1(\mathbf{u}), i_1(\mathbf{v}))) = i_1(\mathbf{uv})$  and  $j_1(i_1(u_0, \dots, u_m)) = i_1(u_m^{-1}, \dots, u_0^{-1})$ , as required.

Next we define functions to work with  $P$ -reduction of encoded words. Define  $\text{Dfind} : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  by

$$\text{Dfind}(\mathbf{a}) := \min_y (y < |\mathbf{a}| \wedge ((\mathbf{a})_y, (\mathbf{a})_{y+1}) \in D'). \quad (73)$$

As  $D'$  is  $\mathcal{E}^n$ -decidable and the minimisation is bounded this is an  $\mathcal{E}^n$ -computable function.

Next define the predicate  $\text{Reduced}$  on  $\mathbb{Z}_{\geq 0}$  by

$$\text{Reduced}(\mathbf{a}) := (\text{find}(\mathbf{a}, i(\varepsilon)) = |\mathbf{a}|) \wedge (\text{Dfind}(\mathbf{a}) = |\mathbf{a}|). \quad (74)$$

Again this is  $\mathcal{E}^n$ -computable and we have  $\text{Reduced}(i_1(\mathbf{u})) = 1$  if and only if  $\mathbf{u}$  is a  $P$ -reduced word in  $P^*$ .

Now we construct a function to perform  $P$ -reduction on an encoded word. This function will perform elementary  $P$ -reduction at the leftmost position where it is possible in a given word. As pointed out in Section 7.1 the order in which  $P$ -reductions are performed does not affect the equivalence class of the  $P$ -reduced word that is produced, so there will be no need to consider any other possible sequences of  $P$ -reduction.

Define a function  $\text{leftm} : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  by

$$\text{leftm}(\mathbf{a}) := \begin{cases} 1, & \text{if } \mathbf{a} = [1], \text{ or } [\varepsilon] \\ \mathbf{a}[0 \cdots \text{Dfind}(\mathbf{a}) - 1] \dot{+} \\ m'((\mathbf{a})_{\text{Dfind}(\mathbf{a})}, (\mathbf{a})_{\text{Dfind}(\mathbf{a})+1}) \dot{+} \\ \mathbf{a}[\text{Dfind}(\mathbf{a}) + 2 \cdots |\mathbf{a}| - 1], & \text{otherwise.} \end{cases} \quad (75)$$

This is an  $\mathcal{E}^n$ -computable function.

Now define  $\text{Predn} : \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \longrightarrow \mathbb{Z}_{\geq 0}$  by

$$\begin{aligned} \text{Predn}(\mathbf{a}, 0) &:= \mathbf{a}, \\ \text{Predn}(\mathbf{a}, n+1) &:= \text{leftm}(\text{Predn}(\mathbf{a}, n)). \end{aligned} \quad (76)$$

As  $\text{Predn}$  is defined by primitive recursion using  $\mathcal{E}^n$ -computable functions it is  $\mathcal{E}^{n+1}$ -computable. The recursion requires up to  $n$  applications of  $m'(x, y)$ , and we do not have a  $\mathcal{E}^n$ -computable bounding function for the recursion; so cannot say whether or not  $\text{Predn}(\mathbf{a}, n)$  is in  $\mathcal{E}^n$ .

Finally define  $\text{Preduction} : \mathbb{Z}_{\geq 0} \longrightarrow \mathbb{Z}_{\geq 0}$  by

$$\text{Preduction}(\mathbf{a}) := \text{Predn}(\mathbf{a}, |\mathbf{a}|). \quad (77)$$

Now if  $\mathbf{u} \in P^*$  and  $\mathbf{v}$  is a  $P$ -reduced word obtained from  $\mathbf{u}$  by  $P$ -reduction, where reduction is always made at the leftmost possible position, then  $\text{Preduction}(i_1(\mathbf{u})) = i_1(\mathbf{v})$ .

In order to recognise when two encoded words represent elements of the same  $\approx$  equivalence class we define the relation  $\text{Interleaven} \subseteq \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$  defined as

$$\begin{aligned} \text{Interleaven}(\mathbf{a}, \mathbf{b}) &:= \text{Preduced}(\mathbf{a}) \wedge \text{Preduced}(\mathbf{b}) \\ &\wedge (|\mathbf{a}| = |\mathbf{b}|) \wedge (\text{Preduction}(m_1(j_1(\mathbf{b}), \mathbf{a})) = 1). \end{aligned} \quad (78)$$

This is an  $\mathcal{E}^{n+1}$ -decidable relation and from Theorem 7.1 and Lemma 7.2 we have, for  $P$ -reduced words  $\mathbf{u}$  and  $\mathbf{v}$ ,  $\mathbf{u} = \mathbf{v} \in \mathbf{U}(P)$  if and only if  $|\mathbf{u}| = |\mathbf{v}|$  and  $\mathbf{v}^{-1}\mathbf{u}$   $P$ -reduces to the empty word, if and only if  $\text{Interleaven}(i_1(\mathbf{u}), i_1(\mathbf{v}))$ . We shall write  $m \approx n$ , for integers  $m, n$  such that  $\text{Interleaven}(m, n)$ .

## 7.4 Proof of Theorem 7.4

We define a function  $i_2 : \mathbb{Z}_{\geq 0} \longrightarrow \mathbb{Z}_{\geq 0}$  which will be used to construct an index for  $\mathbf{U}(P)$ .

$$i_2(\mathbf{a}) := \min_y (y < \mathbf{a} \wedge (y \in i_1(P)) \wedge \text{Preduced}(y) \wedge (y \approx \mathbf{a})). \quad (79)$$

Now define functions  $i_G : P^* \longrightarrow \mathbb{Z}_{\geq 0}$ ,  $m_G : i_G(P^*) \times i_G(P^*) \longrightarrow \mathbb{Z}_{\geq 0}$  and  $j_G : i_G(P^*) \longrightarrow \mathbb{Z}_{\geq 0}$  by

$$i_G(\mathbf{u}) := i_2 i_1(\mathbf{u}), \quad (80)$$

$$m_G(\mathbf{a}, \mathbf{b}) := i_2(\text{Preduction}(m_1(\mathbf{a}, \mathbf{b}))) \text{ and} \quad (81)$$

$$j_G(\mathbf{a}) := i_2 j_1(\mathbf{a}), \quad (82)$$

for all for all  $P$ -reduced words  $\mathbf{u} \in P^*$  and  $\mathbf{a}, \mathbf{b} \in i_G(P^*)$ . Theorem 7.4 then follows from the next Proposition.

**Proposition 7.5.** *The group  $G = \mathbf{U}(P)$  has  $\mathcal{E}^{n+1}$  index  $(i_G, m_G, j_G)$ .*

*Proof.* The group  $G$  consists of equivalence classes of  $P$ -reduced words under the



relation  $\approx$ . If  $\mathbf{u}$  and  $\mathbf{v}$  are  $P$ -reduced words with  $\mathbf{u} \approx \mathbf{v}$  then  $i_G(\mathbf{u}) = i_2 i_1(\mathbf{u}) = i_2 i_1(\mathbf{v}) = i_G(\mathbf{v})$ , by construction. Therefore  $i_G$  is a well defined function on  $G$ . Moreover if  $\mathbf{u}$  and  $\mathbf{v}$  are  $P$ -reduced words such that  $i_G(\mathbf{u}) = i_G(\mathbf{v})$  then both  $\mathbf{u}$  and  $\mathbf{v}$  are equivalent to  $i_2 i_1(\mathbf{u})$ , so  $i_G$  is injective. Let  $R$  be the set of  $P$ -reduced words and let  $S$  be the image  $i_G(R)$  of  $R$  under  $i_G$ . Then, for  $\mathbf{x} \in \mathbb{Z}_{\geq 0}$ ,

$$\chi_S(\mathbf{x}) = (\mathbf{x} \in i_1(P_1)) \wedge \text{Preduced}(\mathbf{x}) \wedge (i_2(\mathbf{x}) = \mathbf{x}),$$

so  $\chi_S$  is  $\mathcal{E}^{n+1}$ -computable. Hence  $i_G(R) = i_G(G)$  is  $\mathcal{E}^{n+1}$ -decidable.

As  $i_2$  is  $\mathcal{E}^{n+1}$ -computable, so are  $m_G$  and  $j_G$ . It therefore remains to show that, for all  $\mathbf{u}, \mathbf{v} \in R$ ,

$$m_G(i_G(\mathbf{u}), i_G(\mathbf{v})) = i_G(\mathbf{u} \cdot \mathbf{v}) \text{ and } j_G(i_G(\mathbf{u})) = i_G(\mathbf{u}^{-1}),$$

where  $\mathbf{u} \cdot \mathbf{v}$  denotes the product of  $\mathbf{u}$  and  $\mathbf{v}$  in  $G$ .

Let  $\mathbf{u}_0$  and  $\mathbf{v}_0$  be  $P$ -reduced words such that  $i_G(\mathbf{u}_0) = i_2(i_1(\mathbf{u}_0)) = i_1(\mathbf{u}_0) = i_G(\mathbf{u})$  and similarly  $i_2(i_1(\mathbf{v}_0)) = i_1(\mathbf{v}_0) = i_G(\mathbf{v})$ . Let  $\mathbf{a} = i_1(\mathbf{u}_0)$  and  $\mathbf{b} = i_1(\mathbf{v}_0)$ ; so  $i_2(\mathbf{a}) = \mathbf{a}$  and  $i_2(\mathbf{b}) = \mathbf{b}$ . The product  $\mathbf{u} \cdot \mathbf{v}$  is the  $\approx$  equivalence class of the  $P$ -reduction of the concatenation of  $\mathbf{u}$  and  $\mathbf{v}$ . Therefore  $i_G(\mathbf{u} \cdot \mathbf{v}) = i_2(\text{Preduction}(m_1(i_1(\mathbf{u}), i_1(\mathbf{v}))))$ . Since  $\mathbf{u} =_G \mathbf{u}_0$  and  $\mathbf{v} =_G \mathbf{v}_0$ , we have  $\mathbf{u} \cdot \mathbf{v} =_G \mathbf{u}_0 \cdot \mathbf{v}_0$ , so  $\text{Preduction}(m_1(i_1(\mathbf{u}), i_1(\mathbf{v}))) = \text{Preduction}(m_1(\mathbf{a}, \mathbf{b}))$ . Hence  $i_G(\mathbf{u} \cdot \mathbf{v}) = i_2(\text{Preduction}(m_1(\mathbf{a}, \mathbf{b}))) = m_G(\mathbf{a}, \mathbf{b}) = m_G(i_G(\mathbf{u}), i_G(\mathbf{v}))$ , as required.

A similar argument shows that  $j_G(i_G(\mathbf{u})) = i_G(\mathbf{u}^{-1})$ .  $\square$

## 7.5 Pregroups and Bass-Serre groups

We shall use the results of this section to give an alternative proof that under certain conditions on the vertex and edge groups, a graph of  $\mathcal{E}^n$ -computable groups is  $\mathcal{E}^{n+1}$ -computable.

**Theorem 7.6.** *Suppose we have  $G = \pi_1(\mathbf{G}, \Gamma, T, v_0)$ , where  $\Gamma$  has  $\nu$  vertices. Suppose all the  $G_v$  are  $\mathcal{E}^n$ -computable groups for  $n \geq 3$ . Assume all the edge groups are  $\mathcal{E}^n$ -decidable, and all the isomorphisms  $\phi_e$  are  $\mathcal{E}^n$ -computable. Then  $G$  is  $\mathcal{E}^{n+1}$ -computable. If the indices for the vertex groups are standard then  $G$  is  $\mathcal{E}^{n+1}$ -standard.*

## 7.6 Questions

1. Under which conditions is  $\mathbf{U}(P)$   $\mathcal{E}^n$ -computable?
2. How does the index found in this section compare with that found for Bass-Serre groups in earlier sections?
3. ....

## References

- [Can66] Frank B. Cannonito. Hierarchies of computable groups and the word problem. *J. Symbolic Logic*, 31:376–392, 1966.
- [CG73] F. B. Cannonito and R. W. Gatterdam. The computability of group constructions. I. In *Word problems: decision problems and the Burnside problem in group theory (Conf., Univ. California, Irvine, Calif. 1969; dedicated to Hanna Neumann)*, pages 365–400. Studies in Logic and the Foundations of Math., Vol. 71. North-Holland, Amsterdam, 1973.
- [DDM10] V. Diekert, A. J. Duncan, and A. G. Miasnikov. Geodesic rewriting systems and pregroups. In O. Bogopolski, I. Bumagin, O. Kharlamovich, and E. Ventura, editors, *Combinatorial and Geometric Group Theory, Dortmund and Carleton Conferences*, Trends in Mathematics, pages 55–91, Basel, 2010. Birkhäuser.
- [Gat73] R. W. Gatterdam. The computability of group constructions. II. *Bull. Austral. Math. Soc.*, 8:27–60, 1973.
- [Grz53] Andrzej Grzegorzcyk. Some classes of recursive functions. *Rozprawy Mat.*, 4:46, 1953.
- [Hoa88] A. H. M. Hoare. Pregroups and length functions. *Math. Proc. Cambridge Philos. Soc.*, 104(1):21–30, 1988.
- [Jan08] Svante Janson. Plane recursive trees, Stirling permutations and an urn model. In *Fifth Colloquium on Mathematics and Computer Science*, Discrete Math. Theor. Comput. Sci. Proc., AI, pages 541–547. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2008.
- [Kle52] S.C. Kleene. *Introduction to metamathematics*. North-Holland Publishing Company, 1952. Co-publisher: Wolters-Noordhoff; 8th revised ed.1980.
- [KWM05] Ilya Kapovich, Richard Weidmann, and Alexei Miasnikov. Foldings, graphs of groups and the membership problem. *Internat. J. Algebra Comput.*, 15(1):95–128, 2005.
- [Rab60] Michael O. Rabin. Computable algebra, general theory and theory of computable fields. *Trans. Amer. Math. Soc.*, 95:341–360, 1960.
- [Rim87a] Frank Rimlinger. Pregroups and Bass-Serre theory. *Mem. Amer. Math. Soc.*, 65(361):viii+73, 1987.
- [Rim87b] Frank Rimlinger. A subgroup theorem for pregroups. In *Combinatorial group theory and topology (Alta, Utah, 1984)*, volume 111 of *Ann. of Math. Stud.*, pages 163–174. Princeton Univ. Press, Princeton, NJ, 1987.
- [Ros84] H. E. Rose. *Subrecursion: functions and hierarchies*, volume 9 of *Oxford Logic Guides*. The Clarendon Press Oxford University Press, New York, 1984.

- [Ser77] Jean-Pierre Serre. *Arbres, amalgames,  $SL_2$* . Société Mathématique de France, Paris, 1977. Avec un sommaire anglais, Rédigé avec la collaboration de Hyman Bass, Astérisque, No. 46.
- [Sta71] John Stallings. *Group theory and three-dimensional manifolds*. Yale University Press, New Haven, Conn., 1971. A James K. Whittemore Lecture in Mathematics given at Yale University, 1969, Yale Mathematical Monographs, 4.
- [Sta87] John R. Stallings. Adian groups and pregroups. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 321–342. Springer, New York, 1987.