A CATEGORICAL DESCRIPTION OF BASS-SERRE THEORY

M. V. LAWSON AND A. R. WALLIS

Abstract. Self-similar group actions may be encoded by a class of left cancellative monoids called left Rees monoids, a result obtained by combining pioneering work by Perrot with later work by the first author. Left Rees monoids that are also right cancellative are called Rees monoids. Irreducible Rees monoids have the striking property that they are embedded in their universal groups and these universal groups are HNN extensions by a single stable letter. In this paper, we generalize the above theory from monoids to categories. We study the structure of left Rees categories and prove that Rees categories embed into their universal groups. Furthermore, we show that from each graph of groups, we may construct a Rees category and prove that the fundamental group of the former is the universal group of the latter. In this way, Bass-Serre theory may be viewed as a special case of the general problem of embedding categories into groupoids.

2000 AMS Subject Classification: 20M10, 20M50.

The first author’s research was partially supported by an EPSRC grant (EP/I033203/1) and the second by an EPSRC Doctoral Training Account (EP/P504945/1). Some of the material in this paper appeared in the second author’s PhD thesis [33].

1. Introduction and terminology

In this paper, we generalize the theory developed in [16, 18] from monoids to categories. In particular, we shall show how to construct a cancellative category from each graph of groups in such a way that the universal groupoid of the category is the fundamental groupoid of the graph of groups.

Any undefined terms from category theory we use may be found in [22]. For us, categories are small and objects are replaced by identities; thus we view them as ‘monoids with many identities’. The elements of a category $C$ are called arrows and the set of identities of $C$ is denoted by $C_0$. Each arrow $a$ has a domain, denoted by $d(a)$, and a codomain denoted by $r(a)$, both of which are identities and $a = d(a)a = ar(a)$. The product $ab$ exists if and only if $r(a) = d(b)$. Thus our products should be conceived thus

$e \rightarrow f \rightarrow i$

where $e$, $f$ and $i$ are identities. We shall sometimes write $\exists ab$ to mean that the product $ab$ exists in the category. Given identities $e$ and $f$, the set of arrows $eCf$ is called a hom-set and $eCe$ is a monoid called the local monoid at $e$. The group of units of the local monoid at $e$ is called the local group at $e$. We say that arrows $x$ and $y$ are parallel if they belong to the same hom-set. A category is called left cancellative if whenever $ax = ay$ we have that $x = y$. We define right cancellative categories

\footnote{This is the reverse of the way that the first author usually treats category products but is the most natural one in the light of the applications we have in mind.}
dually. A cancellative category is one which is both left and right cancellative. An arrow \( a \) is invertible or an isomorphism if there is an arrow \( a^{-1} \), called an inverse and perforce unique, such that \( aa^{-1} = d(a) \) and \( a^{-1}a = r(a) \). An element \( a \in C \) of a category is said to be an atom if it is not invertible and if \( a = bc \) then either \( b \) or \( c \) is invertible.

**Lemma 1.1.** If \( a \) is an atom in an arbitrary category, \( g \) is invertible and \( \exists ga \) then \( ga \) is an atom.

**Proof.** Suppose that \( ga = bc \). Then \( a = (g^{-1}b)c \). Thus \( g^{-1}b \) or \( c \) is invertible; that is, \( b \) or \( c \) is invertible. It follows that \( ga \) is also an atom. \( \square \)

A category in which every arrow is invertible is called a groupoid. If a groupoid is just a disjoint union of its local monoids then we say that it is totally disconnected. The set of invertible elements of a category forms a groupoid with the same set of identities. Later, we shall deal only with categories \( C \) having the following property: any isomorphism belongs to a local monoid. We shall say that such categories are skeletal. The following is well-known.

**Lemma 1.2.** Every category is equivalent to a skeletal category.

**Proof.** Let \( C \) be a category. Let \( \{ e_i : i \in I \} \) be a transversal of the isomorphism classes of the identities. Thus if \( i \neq j \) then \( e_i \) is not isomorphic to \( e_j \) and every identity in \( C \) is isomorphic to some \( e_i \). Let \( C' \) be the full subcategory determined by the \( e_i \). Then \( C' \) is a skeletal category equivalent to \( C \). \( \square \)

Given any directed graph \( D \), we may construct the free category on \( D \) denoted by \( D^* \). This has one identity for each vertex and consists of all finite directed paths in the graph. Multiplication is concatenation of paths. The elements of \( D^* \) will be written \( a_1 \cdots a_m \) where the \( a_i \) are edges of the graph such that \( a_i \cdots a_2 a_1 \).

With each category \( C \), we may associate its universal or fundamental groupoid \( U(C) \) [21]. To construct this groupoid, first regard \( C \) as a directed graph. For each edge \( e \), \( a \) to \( f \) attach a new edge \( f \leftarrow e \). Form the free category \( (C \cup C^{-1})^* \) on the set \( C \cup C^{-1} \). We denote elements of this category by \( a_1 \cdots a_n \) where \( a_i \in C \cup C^{-1} \) and \( r(a_i) = d(a_{i+1}) \) for \( i = 1, 2, \ldots, n - 1 \). Now define a congruence \( \equiv \) on this free category generated by

\[
\begin{align*}
    a \cdot a^{-1} & \equiv d(a), \\
    a^{-1} \cdot a & \equiv r(a), \\
    a \cdot b & \equiv ab \text{ if } \exists ab
\end{align*}
\]

where \( a, b \in C \). The first two identifications ensure that the quotient category is a groupoid and the third ensures that there will be a functor from \( C \) to the quotient. This gives us our groupoid \( U(C) \) with an associated functor \( \iota : C \to U(G) \). We prove that this has the correct universal property. Let \( \theta : C \to G \) be a functor to a groupoid. We may extend \( \theta \) to a function \( \tilde{\theta} : C \cup C^{-1} \to G \) where we define \( \tilde{\theta}(a) = \theta(a) \) and \( \tilde{\theta}(a^{-1}) = \theta(a)^{-1} \). We may therefore extend \( \tilde{\theta} \) to a functor \( \theta' \) from the free category \( (C \cup C^{-1})^* \) to \( G \). Observe that if \( x \equiv y \) then \( \tilde{\theta}(x) \equiv \tilde{\theta}(y) \). Thus we may define a functor \( \Theta \) from \( U(C) \) to \( G \). The uniqueness property comes from the fact that \( U(C) \) is generated by \( C \).

2. Levi categories

In this section, we define a class of categories that contains the left Rees categories which we are primarily interested in.
The following definition is generalized from semigroup theory [23]. A category \( C \) is said to be \textit{equidivisible} if for every commutative square

\[
\begin{array}{ccc}
\downarrow & & \downarrow \\
\downarrow & & \downarrow \\
\end{array}
\]

we either have an arrow \( u \) making the following diagram commute

\[
\begin{array}{ccc}
\downarrow & & \downarrow \\
\downarrow & & \downarrow \\
\end{array}
\]

or one, \( v \), making the following diagram commute

\[
\begin{array}{ccc}
\downarrow & & \downarrow \\
\downarrow & & \downarrow \\
\end{array}
\]

A \textit{length functor} is a functor \( \lambda: C \to \mathbb{N} \) from a category \( C \) to the additive monoid of natural numbers satisfying the following conditions:

(LF1): If \( xy \) is defined then \( \lambda(xy) = \lambda(x) + \lambda(y) \).

(LF2): \( \lambda^{-1}(0) \) consists of all and only the invertible elements of \( C \).

(LF3): \( \lambda^{-1}(1) \) consists of all and only the atoms of \( C \).

Sometimes we shall write \( \lambda_C \) if we wish to emphasize the fact that the length functor belongs to \( C \).

A \textit{Levi category} is an equidivisible category equipped with a length functor that contains at least one noninvertible element (to rule out groupoids). These generalize the Levi monoids first introduced in [18].

A \textit{principal right ideal} in a category \( C \) is a subset of the form \( aC \) where \( a \in C \). We may similarly define \textit{principal left ideals} and \textit{principal ideals}. Consequently, Greens relations \( \mathcal{L}, \mathcal{R}, \mathcal{H}, \mathcal{D}, \) and \( \mathcal{J} \) can also be defined in categories. Thus, for example, in the category \( C \) we define \( a \mathcal{L} b \) if and only if \( Ca =Cb \). Let \( aC \subseteq bC \). We use the notation \([aC, bC]\) to mean the set of principal right ideals \( xC \) such that \( aC \subseteq xC \subseteq bC \). The proof of the following is straightforward.

\textbf{Lemma 2.1.} Let \( C \) be a category equipped with a length functor having \( G \) as its groupoid of invertible elements.

1. \( a \mathcal{L} b \iff Ga = Gb \). In particular, \( r(a) = r(b) \).
2. \( a \mathcal{R} b \iff aG = bG \). In particular, \( d(a) = d(b) \).
3. \( a \mathcal{J} b \iff GaG = GbG \).

Principal right ideals of the form \( eC \) where \( e \) is an identity are maximal such ideals because if \( eC \subseteq aC \) then \( e = d(a) \) and \( aC = eaC \subseteq eC \) giving \( eC = aC \). A
principal right ideal \( aC \) is said to be \textit{submaximal} if \( aC \neq d(a)C \) and there are no proper principal right ideals between \( aC \) and \( d(a)C \).

**Lemma 2.2.** Let \( C \) be a Levi category.

1. The element \( a \) is an atom if and only if \( aC \) is submaximal if and only if \( Ca \) is submaximal.
2. Each non-invertible element \( a \) of \( C \) can be written \( a = a'b \) where \( a' \) is an atom.
3. Every non-invertible element of \( C \) can be written as a product of atoms.

**Proof.** (1). Suppose that \( a \) is an atom and that \( aC \subseteq bC \). Then \( a = bc \). But \( a \) is an atom and so either \( b \) is invertible or \( c \) is invertible. Suppose that \( c \) is invertible then \( aC = bC \). Suppose that \( b \) is invertible then \( bC = d(b)C \). We have proved that \( aC \) is submaximal.

Conversely, suppose that \( aC \) is submaximal. Let \( a = bc \). Then \( aC \subseteq bC \). It follows that either \( aC = bC \) or \( bC = d(b)C \). If the latter occurs then \( b \) is invertible. If the former occurs then \( \lambda(a) = \lambda(b) \). It follows that \( \lambda(c) = 0 \) and so \( c \) is invertible.

(2). Let \( a \) be a non-invertible element. If \( a \) is an atom then we are done. If not, then \( a = a_1b_1 \) for some \( b_1 \) where neither \( a_1 \) nor \( b_1 \) are invertible. Observe that \( \lambda(a_1) < \lambda(a) \). If \( a_1 \) is an atom then we are done, otherwise we may write \( a_1 = a_2b_2 \) again where neither \( a_2 \) nor \( b_2 \) are atoms. Observe that \( \lambda(a_2) < \lambda(a_1) \). This process can only continue in a finite number of steps and will end with \( a = a_nb' \) for some \( b' \) where \( a_n \) is an atom.

(3). If \( a \) is an atom then there is nothing to prove. Otherwise by (2), we may write \( a = a_1b_1 \) where \( a_1 \) is an atom and \( \lambda(b_1) = \lambda(a) - 1 \). If \( b_1 \) were invertible then \( a \) would have been an atom. If \( b_1 \) is an atom then we are done. Otherwise, we may repeat the above procedure with \( b_1 \).

\[ \square \]

In categories equipped with length functors, we can write every element in terms of atoms and invertible elements. The obvious next question is what kind of uniqueness we can expect. Under the additional assumption of equidivisibility, we can retrieve a kind of uniqueness.

**Lemma 2.3.** Let \( C \) be a Levi category and suppose that

\[ x = a_1 \ldots a_m = b_1 \ldots b_n \]

where the \( a_i \) and \( b_j \) are atoms.

1. \( m = n \).
2. There are invertible elements \( g_1, \ldots, g_{n-1} \) such that

\[ a_1 = b_1g_1, \quad a_2 = g_1^{-1}b_2g_2, \quad \ldots \quad a_n = g_{n-1}^{-1}b_n. \]

This data is best presented by means of the following interleaving diagram.
(3) If $C$ is in addition skeletal then $a_i$ is parallel to $b_i$ for $i = 1, \ldots, m$.

Proof. (1). This is immediate from the properties of length functors.

(2). We bracket as follows

$$a_1(a_2 \ldots a_m) = b_1(b_1 \ldots b_m).$$

By equidivisibility, $a_1 = b_1u$ and $b_2 \ldots b_m = ua_2 \ldots a_m$ for some $u$ or $b_1 = a_1v$ and $a_2 \ldots a_m = vb_2 \ldots b_m$ for some $v$. In either case, $u$ and $v$ are invertible since both $a_1$ and $b_1$ are atoms using the length function. Thus $a_1 = b_1g_1$, where $v = g_1$ is an isomorphism, and $b_2 \ldots b_m = g_1a_2 \ldots a_m$.

We now repeat this procedure bracketing thus

$$b_2(b_3 \ldots b_m) = g_1a_2(a_3 \ldots a_m).$$

By the same argument as above, we get that $g_1a_2 = b_2g_2$ for some isomorphism $g_2$ and $b_3 \ldots b_m = g_2a_3 \ldots a_m$.

The process continues and we obtain the result.

(3). The result is immediate from the assumption that the category is also skeletal.

We shall be interested in left cancellative Levi categories and ultimately those that are also skeletal. But we shall start with a slightly different definition and show that it is equivalent to this one. A category $C$ is said to be right rigid if $aC \cap bC \neq \emptyset$ implies that $aC \subseteq bC$ or $bC \subseteq aC$. A left Rees category is a left cancellative, right rigid category in which each principal right ideal is properly contained in only finitely many distinct principal right ideals.

Example 2.4. Free categories are left Rees categories: the atoms are the edges; the length functor simply counts the number of edges in a path; the groupoid of invertible elements is trivial.

Lemma 2.5. Let $C$ be a left cancellative category with groupoid of invertible elements $G$.

(1) $a \mathcal{R} b$ if and only if $aG = bG$.

(2) $C$ is a right rigid category if and only if it is equidivisible.

(3) If $e = xy$ is an identity then $x$ is invertible with inverse $y$.

Proof. (1). Only one direction needs proving. Suppose that $a \mathcal{R} b$. Then $a = bx$ and $b = ay$ for some $x, y \in C$. Thus $a = ayx$ and $b = bxy$. By left cancellation both $xy$ and $yx$ are identities and so $x$ is invertible with inverse $y$. 


(2). Only one direction needs proving. Suppose that \( ab = cd \). Then \( aC \cap cC \neq \emptyset \). Without loss of generality, we may suppose by right rigidity that \( aC \subseteq cC \). Thus \( a = cu \). But then \( cuv = cd \) and so \( uv = d \), as required.

(3). We have that \( xyx = x \) and by left cancellation this shows that \( yx \) is an identity and so \( x \) is invertible with inverse \( y \). \( \square \)

In the light of the lemma above, the following says that left cancellative equidivisible categories equipped with length functors are left Rees categories.

**Lemma 2.6.** Let \( C \) be a right rigid, left cancellative category equipped with a length functor \( \lambda \). Then \( C \) is a left Rees category.

**Proof.** Let \( a \in C \) be any element with \( e = d(a) \). We need to prove that the set \([aC, eC]\) is finite. Let \( bC \in [aC, eC] \). Then \( a = bx \) for some \( x \) and so \( \lambda(a) \geq \lambda(b) \). There is therefore an upper bound on the lengths of those elements \( b \) such that \( aC \subseteq bC \). Let \( b_1C, b_2C \in [aC, eC] \) and suppose that \( \lambda(b_1) = \lambda(b_2) \). By right rigidity, we may assume, without loss of generality that \( b_1C \subseteq b_2C \). Thus \( b_1 = b_2 x \) for some \( x \). But \( b_1 \) and \( b_2 \) have the same length and so \( x \) must have length zero. It follows that \( x \) is invertible. Hence \( b_1C = b_2C \). It follows that the set \([aC, eC]\) is finite, as claimed. \( \square \)

We shall now prove the converse to the above result. This involves proving that every left Rees category is equipped with a length functor. Let \( a \in C \) be an element of a left Rees category. We shall define the length, \( \lambda(a) \), of \( a \). Put \( d(a) = e \). By assumption, the set \([aC, eC]\) is finite and linearly ordered. The proofs of the following are straightforward.

**Lemma 2.7.** Let \( C \) be a left Rees category.

1. The set \([aC, eC]\) contains one element if and only if \( a \) is an atom.
2. The set \([aC, eC]\) contains two elements if and only if \( a \) is invertible.

Let \( a \in C \) where \( aC \subseteq eC \) where \( e \) is an identity. Define

\[ \lambda(a) = |[aC, eC]| - 1. \]

**Lemma 2.8.** Let \( C \) be a left Rees category and let \( a, b \in C \) such that \( ab \) is defined.

1. \( [abC, d(b)C] = [abC, aC] \).
2. \( [abC, d(a)C] = [abC, d(b)C] \cup [aC, d(a)C] \).
3. If \( bC \subseteq cC, dC \subseteq C \) and \( acC = adC \) then \( cC = dC \).

**Proof.** (1). Let \( xC \in [bC, d(b)C] \). Then \( bC \subseteq xC \subseteq d(b)C \). It is immediate that \( abC \subseteq axC \subseteq aC \). Thus the lefthand side is contained in the righthand side. Now let \( abC \subseteq yC \subseteq aC \). Then \( y = au \) for some \( u \in C \). Thus \( yC = auC = a(uC) \). Now \( ab = yv \) for some \( v \). Hence \( ab = auv \) and so by left cancellation, we have that \( b = uv \). Thus \( bC \subseteq uC \subseteq d(b)C \).

(2). Let \( c = ab \). Then \( cC \subseteq aC \subseteq d(a)C \). Let \( xC \in [cC, d(c)C] \). Since this set is linearly ordered either \( xC \subseteq aC \) or \( aC \subseteq xC \). It follows that the lefthand side is contained in the righthand side. The proof of the reverse containment is immediate.

(3). This is immediate by left cancellation. \( \square \)

The proof of the following is now immediate.

**Lemma 2.9.** Let \( C \) be a left Rees category.
(1) $\lambda(ab) = \lambda(a) + \lambda(b)$.
(2) $\lambda(a) = 0$ if and only if $a$ is invertible.
(3) $\lambda(a) = 1$ if and only if $a$ is an atom.

We have therefore proved the following.

**Proposition 2.10.** A left cancellative, right rigid category is a left Rees category if and only if it is equipped with a length functor. In other words, the left Rees categories are precisely the left cancellative Levi categories.

The following result is worth noting here.

**Lemma 2.11.** Let $C$ be a left Rees category. Then each local monoid of $C$ is a left Rees monoid.

**Proof.** Let $S = eCe$ be a local monoid. It is immediate that it is a left cancellative monoid. Let $aS \cap bS \neq \emptyset$ where $a, b \in S$. Then clearly $aC \cap bC \neq \emptyset$. Without loss of generality, it follows that $aC \subseteq bC$. Thus $a = bc$ for some $c \in C$. Now $eae = a$ and $b = ebe$ so that $a = b(ece)$ and $c = ece$. We have therefore proved that $aS \subseteq bS$. Thus $S$ is right rigid. We denote by $[aS, bS]$ the obvious set of principal right ideals in $S$. The set $[aS, eS]$ is linearly ordered. We shall prove that it is finite. Let $aS \subseteq b_1S \subseteq b_2S \subseteq eS$ where $b_1, b_2 \in S$. We shall prove that $b_1C \neq b_2C$. Suppose on the contrary that $b_1C = b_2C$. Then $b_1 = b_2g$ for some invertible element $g$. Then $b_1 = b_2(ege)$ and $ege = g$. Thus $g$ is an invertible element in $S$ and so $b_1S = b_2S$, a contradiction. Since, by assumption, the set $[aC, bC]$ is finite it follows that the set $[aS, bS]$ is finite. □

**Remark 2.12.** Let $C$ be a left Rees category. It is important to observe that although $S = eCe$ is a left Rees monoid, its length function need not be the restriction of the one in $C$. The length of the element $a \in S$, viewed as an element of the monoid $S$, is defined to be one less than the number of elements of $[aS, eS]$. The length of the element $a \in S$ viewed as an element of $C$ is defined to be one less than the number of elements in $[aC, eC]$. But the latter set may contain more elements than the former. Thus $\lambda_S(a) \leq \lambda_C(a)$. Consider the following example. Let $C$ be the free category defined by the following directed graph

![Directed Graph]

The local monoid at $e$ is just $S = (ab)^*$, the free monoid on one generator $ab$. Here $\lambda_S(ab) = 1$ but $\lambda_C(ab) = 2$ since both $a$ and $b$ are atoms in $C$.

**Remark 2.13.** It can be shown that condition (LF3) is not needed in the definition of a length functor. If Levi categories are defined with respect to this apparently weaker definition, a length functor may be easily constructed that satisfies this condition using the analogue of Lemma 2.8. This was carried out explicitly in the monoid case described in [18], and the proof there may be easily extended to the category case.

### 3. Constructing Levi categories

The goal of this section is to show that each Levi category is isomorphic to a tensor category over what we call a bimodule: that is, sets on which a given groupoid acts on the left and the right in such a way that the two actions associate.
Let $G$ be a groupoid and let $X$ be a set equipped with two functions

$$G_0 \overset{s}{\leftarrow} X \overset{t}{\rightarrow} G_0.$$ 

We suppose that there is a left groupoid action $G \times X \to X$ and a right groupoid action $X \times G \to X$ such that the two actions associate meaning $(gx)h = g(wh)$ when defined. We write $\exists gx$ and $\exists xg$ if the actions are defined. Observe that $\exists gx$ iff $r(g) = s(x)$ and $\exists xg$ iff $t(x) = d(g)$. We call the structure $(G, X, G)$ a bimodule or a $(G, G)$-bimodule. If whenever $\exists xg$ and $xy = x$ we have that $g$ is an identity, then we say the action is right free. A bimodule which is right free is called a covering bimodule. We define left free dually. A bimodule which is both left and right free is said to be bifree. We define homomorphisms and isomorphisms between $(G, G)$-bimodules in the usual way.

Our first result shows that bimodules arise naturally from Levi categories. The proof is straightforward; in particular, the fact that the actions are well-defined follows from Lemma 1.1

Lemma 3.1. Let $C$ be a Levi category. Denote by $X$ the set of all atoms of $C$ equipped with the maps $d, r: X \to G_0$. Denote by $G$ the groupoid of invertible elements of $C$. Define a bimodule $(G, X, G)$ where the left and right actions are defined via multiplication in $C$ when defined. We obtain a covering bimodule if $C$ is left cancellative and a bifree bimodule if $C$ is cancellative.

We call $(G, X, G)$ constructed as in the above lemma, the bimodule associated with $C$ or the bimodule of atoms of $C$.

Our goal now is to show that from each bimodule we may construct a Levi category. Our tool for this will be tensor products and the construction of a suitable tensor algebra analogous to the ones defined in module theory; see Chapter 6 of [20], for example. We recall the key definitions and results we need first.

Let $G$ be a groupoid that acts on the set $X$ on the right and the set $Y$ on the left. We consider the set $X \ast Y$ consisting of those pairs $(x, y)$ where $t(x) = s(y)$. A function $\alpha: X \ast Y \to Z$ to a set $Z$ is called a bi-map or a 2-map if $\alpha(xy, y) = (x, yy)$ for all $(xy, y) \in X \ast Y$ where $y \in G$. We may construct a universal such bimap $\lambda: X \ast Y \to X \otimes Y$ in the usual way [5]. However, there is a simplification in the theory due to the fact that we are acting by means of a groupoid. The element $x \otimes y$ in $X \otimes Y$ is the equivalence class of $(x, y) \in X \ast Y$ under the relation $\sim$ where $(x, y) \sim (x', y')$ if and only if $(x', y') = (xy^{-1}, yy)$ for some $y \in G$. Observe that we may define $s(x \otimes y) = s(x)$ and $t(x \otimes y) = t(y)$ unambiguously.

Suppose now that $X$ is a $(G, G)$-bimodule. We may therefore define the tensor product $X \otimes X$ as a set. This set is equipped with maps $s, t: X \otimes X \to G_0$. We define $g(x \otimes y) = gx \otimes y$ and $(x \otimes y)g = x \otimes yg$ when this makes sense. Observe that $x \otimes y = x' \otimes y'$ implies that $gx \otimes y = gx' \otimes y'$, and dually. It follows that $X \otimes X$ is also a bimodule. Put $X^\otimes 2 = X \otimes X$. More generally, we may define $X^\otimes n$ for all $n \geq 1$ using $n$-maps, and we define $X^\otimes 0 = G$ where $G$ acts on itself by multiplication on the left and right. The proof of the following lemma is almost immediate from the definition and the fact that we are acting by a groupoid.

Lemma 3.2. Let $n \geq 2$. Then

$$x_1 \otimes \ldots \otimes x_n = y_1 \otimes \ldots \otimes y_n$$

if and only if there are elements $g_1, \ldots, g_{n-1} \in G$ such that $y_1 = x_1g_1$, $y_2 = g_1^{-1}x_2g_2$, $y_3 = g_2^{-1}x_3g_3$, \ldots, $y_n = g_{n-1}^{-1}x_n$. 


Define
\[ T(X) = \bigcup_{n=0}^{\infty} X^\otimes n. \]
We shall call this the tensor category associated with the bimodule \((G, X, G)\). Observe that we may regard \(X\) as a subset of \(T(X)\). The justification for this terminology will follow from (1) below.

**Theorem 3.3.**

1. \(T(X)\) is a Levi category whose associated bimodule is \((G, X, G)\).
2. The category \(T(X)\) is left cancellative if and only if \((G, X, G)\) is right free, and dually.
3. Let \(C\) be a category whose groupoid of invertible elements is \(G\). Regard \(C\) as a \((G, G)\)-bimodule under left and right multiplication. Let \(\theta: X \rightarrow C\) be any bimodule morphism to \(C\). Then there is a unique functor \(\Theta: T(X) \rightarrow C\) extending \(\theta\).
4. Every Levi category is isomorphic to the tensor category of its associated bimodule.

**Proof.**

(1). The identities of the category are the same as the identities of \(G\). The element \(x_1 \otimes \cdots \otimes x_n\) has domain \(d(x_1)\) and codomain \(r(x_n)\). Multiplication is tensoring of sequences that begin and end in the right places and left and right actions by elements of \(G\). We define \(\lambda(g) = 0\) where \(g \in G\) and \(\lambda(x_1 \otimes \cdots \otimes x_n) = n\). Formally, we are using the fact that there is a canonical isomorphism
\[ X^\otimes p \otimes X^\otimes q \cong X^\otimes (p+q). \]

The proof of equidivisibility is essentially the same as that of Proposition 5.6 of [16]. The elements of length 0 are precisely the elements of \(G\) and so the invertible elements; the elements of length 1 are precisely the elements of \(X\).

(2). Suppose that the category is left cancellative and that \(xg = x = xd(g)\) in the bimodule. But this can also be interpreted as a product in the category and so \(g = d(g)\), as required. Conversely, suppose that the bimodule is right free. Let \(x \otimes y = x \otimes z\). From Lemma [3.2] and the fact that lengths match, we have that \((x, y) = (xg, g^{-1}z)\) for some \(g \in G\). But using the fact that the action is right free, we get that \(g\) is an identity and so \(y = z\), as required.

(3). Define \(\Theta(g) = g\) when \(g \in G\) and \(\Theta(x_1 \otimes x_2 \otimes \cdots \otimes x_n) = \theta(x_1) \otimes \theta(x_2) \otimes \cdots \otimes \theta(x_n)\). This is well-defined by Lemma [3.2] It is routine to check that this defines a functor.

(4). This now follows from (3) above, part (3) of Lemma [2.2] and part (2) of Lemma [2.3].

**Remark 3.4.** It follows by the above theorem that left Rees categories are described by covering bimodules.

### 4. Left Rees Categories

Left cancellative Levi categories are what we call left Rees categories. Such Levi categories admit an alternative description from the tensor category description of the previous section. Here, we shall describe the structure of arbitrary left Rees categories in terms of free categories using Zappa-Szép products generalized to categories; see [3]. This will show that they can be regarded as the categories...
associated with self-similar groupoid actions. The material in this section can be regarded as a special case of \cite{17}. However, we have included it for the sake of completeness.

Let $G$ be a groupoid with set of identities $G_0$ and let $C$ be a category with set of identities $C_0$. We shall suppose that there is a bijection between $G_0$ and $C_0$ and, to simplify notation, we shall identify these two sets. Denote by $G * C$ the set of pairs $(g, x)$ such that $r(g) = d(x)$. We shall picture such pairs as follows:

\[
\begin{array}{c}
g \\
\downarrow \\
x
\end{array}
\]

We suppose that there is a function $G * C \to C$ denoted by $(g, x) \mapsto g \cdot x$, which gives a left action of $G$ on $C$ and a function $G * C \to G$ denoted by $(g, x) \mapsto g|_x$, which gives a right action of $C$ on $G$ such that these two functions satisfy the following conditions:

(C1): $d(g \cdot x) = d(g)$.
(C2): $r(g \cdot x) = d(g|_x)$.
(C3): $r(x) = r(g|_x)$.

This information is summarized by the following diagram

\[
\begin{array}{c}
g \cdot x \\
\downarrow \\
g|_x \\
\downarrow \\
x
\end{array}
\]

We also require that the following axioms be satisfied:

(SS1): $d(x) \cdot x = x$. Observe that this is the action, not the category product.
(SS2): If $gh$ is defined then $(gh) \cdot x = g \cdot (h \cdot x)$.
(SS3): $d(g) = g \cdot r(g)$.
(SS4): $d(x)|_x = r(x)$.
(SS5): $g|_{r(g)} = g$.
(SS6): If $xy$ is defined and $r(g) = d(x)$ then $g|_{xy} = (g|_x)|_y$.
(SS7): If $gh$ is defined and $r(h) = d(x)$ then $(gh)|_x = g|_{h \cdot x}|_x$.
(SS8): If $xy$ is defined and $r(g) = d(x)$ then $g \cdot (xy) = (g \cdot x)(g|_x \cdot y)$.

If there are maps $(g, x) \mapsto g \cdot x$ and $(g, x) \mapsto g|_x$ satisfying (C1)–(C3) and (SS1)–(SS8) then we say that there is a self-similar action of $G$ on $C$. Put

\[ C \bowtie G = \{(x, g) \in C \times G : r(x) = d(g)\}. \]
We represent \((x, g)\) by the diagram
\[
\begin{array}{ccc}
  & x \downarrow & \\
  & g \downarrow & \\
  & y \downarrow & \\
  & h &
\end{array}
\]

Given elements \((x, g)\) and \((y, h)\) satisfying \(r(g) = d(y)\) we then have the following diagram
\[
\begin{array}{ccc}
  & x \downarrow & \\
  & g \downarrow & \\
  & g \cdot y \downarrow & \\
  & g|v \downarrow & \\
  & y \downarrow & \\
  & h &
\end{array}
\]

Completing the square, as shown, enables us to define a partial binary operation on \(C \rtimes G\) by
\[(x, g)(y, h) = (x(g \cdot y), g|y h).
\]

**Lemma 4.1.** Let \(G\) be a groupoid having a self-similar action on the category \(C\).

1. If \(y\) is an invertible element of \(C\) then so too is \(g \cdot y\).
2. If \(g \in G\) and \(x \in X\) is an atom then \(g \cdot x\) is an atom.
3. If \(C\) is a left Rees category with length function \(\lambda\) then \(\lambda(g \cdot x) = \lambda(x)\).

**Proof.** (1). Suppose that \(e \xrightarrow{g} f\). We have that \(g \cdot e = d(g)\), by (SS3), and \(e = yy^{-1}\). Thus \(d(g) = (g \cdot y)(g|v \cdot y^{-1})\). It follows that \(g \cdot y\) is invertible with inverse \(g|v \cdot y^{-1}\).

(2). Let \(g \cdot x = uv\), where \(u, v \in C\). By axioms (SS1), (SS2) and (SS8), we have that
\[x = (g^{-1} \cdot u)(g^{-1}|u \cdot v)\]
By assumption, \(x\) is an atom and so at least one of the elements in the product is invertible. Suppose that \(g^{-1} \cdot u\) is invertible. Then by our result above \(g \cdot (g^{-1} \cdot u) = u\) is invertible. Suppose now that \(g^{-1}|u \cdot v\) is invertible then we may deduce that \(v\) is invertible. We have therefore proved that \(g \cdot x\) is an atom.

(3). Write \(x = x_1 \ldots x_n\) a product of atoms where \(n = \lambda(x)\). Now use (SS8). \(\Box\)

**Proposition 4.2.** Let \(G\) be a groupoid having a self-similar action on the category \(C\).

1. \(C \rtimes G\) is a category.
2. \(C \rtimes G\) contains copies \(C'\) and \(G'\) of \(C\) and \(G\) respectively such that each element of \(C \rtimes G\) can be written as a product of a unique element from \(C'\) followed by a unique element from \(G'\).
3. If \(C\) is left cancellative then so too is \(C \rtimes G\).
4. If \(C\) is left cancellative then the set of invertible elements of \(C \rtimes G\) consists of all those elements \((x, g)\) where \(x\) is invertible in \(C\).
(5) If $C$ is left cancellative then the set of atoms in $C \bowtie G$ consists of all those elements $(x, g)$ where $x$ is an atom in $C$.

(6) If $C$ is left cancellative and right rigid then so too is $C \bowtie G$.

(7) If $C$ is a left Rees category then so too is $C \bowtie G$.

Proof. (1) Define $d(x, g) = (d(x), d(x))$ and $r(x, g) = (r(g), r(g))$. The condition for the existence of $(x, g)(y, h)$ is that $r(x, g) = d(y, h)$. Axioms (C1), (C2) and (C3) then guarantee the existence of $(x(g \cdot y), g \cdot h)$. We therefore have a partially defined multiplication. We next locate the identities. Suppose that $(u, a)$ is an element such that if $(u, a)(x, g)$ is defined then $(u, a)(x, g) = (x, g)$. Now $(u, a)(r(a), r(a))$ is defined. We deduce that $r(a) = u(a \cdot r(a))$ and $r(a) = a|_{r(a)} r(a)$. By (SS5), we have that $a = r(a)$ and by (SS3) that $a \cdot r(a) = d(a)$. Thus $(u, a) = (r(a), r(a))$. By (SS1) and (SS4), we deduce that the identities are the elements of the form $(e, e)$ where $e \in C_o = G_o$. Observe that $d[(x, g)(y, h)] = d(x, g)$ and $r[(x, g)(y, h)] = r(y, h)$. It remains only to prove associativity.

Suppose first that $[(x, g)(y, h)](z, k)$ exists. The product $(x, g)(y, h)$ exists and so we have the following diagram

```
  x
 / \
|   |
|   |
g
|   |
|   |
y
 / \
|   |
|   |
h
```

similarly $[(x, g)(y, h)](z, k)$ exists and so we have the following diagram

```
  x(g \cdot y)
 / \
|   |
|   |
g\cdot h
 / \
|   |
|   |
y
 / \
|   |
|   |
h
```

resulting in the product

$$(x(g \cdot y)((g|_y h) \cdot z], (g|_y h) |_z k).$$

By assumption, $x(g \cdot y)((g|_y h) \cdot z]$ exists and so $(g \cdot y)((g|_y h) \cdot z]$ exists. We now use (SS8) and (SS4) and (SS7), to get that $y(h \cdot z)$ exists and we use (SS7) and (SS6) to show that

$$(g|_y h) |_z k = g|_{y(h \cdot z)} |_z k.$$

By (SS2),

$$x(g \cdot y)((g|_y h) \cdot z] = (x(g \cdot y))((g|_y h) \cdot (h \cdot z)).$$
It now follows that
\[(y, h)(z, k) = (g(h \cdot z), h \cdot z, k)\]
eexists. It also follows that \((x, g)[(y, h)(z, k)]\) exists and is equal to
\[[(x, g)(y, h)](z, k).\]

Next suppose that
\[(x, g)[(y, h)(z, k)]\]
eexists. This multiplies out to give \((x[g \cdot (y(h \cdot z))], g[y(h \cdot z)](z, k))\). By (SS6) and (SS7) we get that
\[g[y(h \cdot z)](z, k) = (g[yh]z, k),\]
and by (SS8) and (SS2) we get that \(x(g \cdot (y(h \cdot z))) = x(g \cdot y)(y(h \cdot z)).\) This completes the proof that \(C \bowtie G \) is a category.

(2) Define \(\iota_C : C \rightarrow C \bowtie G \) by \(\iota_C(x) = (x, r(x)).\) This is well-defined. Suppose that \(xy\) exists. Then in particular \(r(x) = d(y).\) It is easy to check using (SS4) and (SS1) that \(\iota_C(x)\iota_C(y) = \iota_C(xy).\) In fact, \(\iota_C(x)\iota_C(y) \) exists iff \(xy\) exists. Thus the categories \(C \) and \(C'\) are isomorphic.

Now define \(\iota_G : G \rightarrow C \bowtie G \) by \(\iota_G(g) = (d(g), g).\) Then once again the categories \(G \) and \(G'\) are isomorphic.

Finally, pick an arbitrary non-zero element \((x, g)\). Then \(r(x) = d(g).\) We may write \((x, g) = (x, r(x))(d(g), g)\) using the fact that \(r(x) \cdot d(g) = d(g)\) by (SS1) and \(r(x)d(g) = d(g)\) by (SS4).

(3) Suppose that \(C\) is left cancellative. We prove that \(C \bowtie G \) is left cancellative. Suppose that \((x, g)(y, h) = (x, g)(z, k)\). Then \(x(g \cdot y) = x(g \cdot z)\) and \(g[yh]g[k].\)

By left cancellation in \(C\) it follows that \(g \cdot y = g \cdot z\) and by (SS1) we deduce that \(y = z\). Hence \(h = k\). We have therefore proved that \((y, h) = (z, k),\) as required.

(4) We know by (3), that the resulting category is left cancellative. Suppose that \((x, g)\) is invertible. Then there is an element \((y, h)\) such that \((d(x), d(x)) = (x, g)(y, h)\). In particular, \(d(x) = x(g \cdot y)\) and so \(x\) is invertible. Conversely, if \(x\) is invertible, it can be verified that
\[(x, g)^{-1} = (g^{-1}, x^{-1}, (g[yh^{-1}, x^{-1}]^{-1}).\]

(5) Suppose that \(x\) is an atom. Let \((x, g) = (u, h)(v, k)\). Then \(x = u(h \cdot v)\) and \(g = h \cdot k\). If \(u\) is invertible then \((u, h)\) is invertible, whereas if \(h \cdot v\) is invertible then \(v\) is invertible and so \((v, k)\) is invertible. It follows that \((x, g)\) is an atom. To prove the converse, suppose that \((x, g)\) is an atom. It is immediate that \(x\) is not invertible. Suppose that \(x\) is not an atom. Then we may write \(x = uv\) where neither \(u\) nor \(v\) is invertible. But then \((x, r(x)) = (u, r(u))(v, r(x))\) and this leads to a non-trivial factorization of \((x, g),\) which is a contradiction.

(6) Suppose now that \(C\) is left cancellative and right rigid. By (3), we know that \(C \bowtie G \) is left cancellative so it only remains to be proved that \(C \bowtie G \) is right rigid. Suppose that
\[(x, g)(y, h) = (u, k)(v, l)\]

From the definition of the product it follows that \(x(g \cdot y) = u(k \cdot v)\) and \(g[yh]h[l]l.\) From the first equation, we see that \(xC \cap uC \neq \emptyset.\) Without loss of generality, suppose that \(x = uw.\) Then by left cancellation \(w(g \cdot y) = k \cdot v.\) Observe that \(k^{-1} \cdot (k \cdot v)\) is defined and so \(k^{-1} \cdot (w(g \cdot y))\) is defined by (SS2). Thus by (SS8), \(k^{-1} \cdot w\) is defined. It is now easy to check that
\[(x, g) = (u, k)(k^{-1} \cdot w, (k[k^{-1} \cdot w]^{-1} g)).\]
(7) Let $C$ be a left Rees category. It is enough to prove that $C \triangleright\triangleleft G$ is equipped with a length function. Let the length function on $C$ be denoted by $\lambda$. Let $(x, g) \in C \triangleright\triangleleft G$. Define $\mu(x, g) = \lambda(x)$. By the above, $\mu(x, g) = 0$ if and only if $(x, g)$ is invertible, and $\mu(x, g) = 1$ if and only if $x$ is an atom. The fact that $\mu$ is a functor follows from the fact that $\lambda(g \cdot y) = \lambda(y)$. □

We call $C \triangleright\triangleleft G$ the Zappa-Szép product of the category $C$ by the groupoid $G$.

Let $C$ be a left Rees category. A transversal of the generators of the submaximal principal right ideals is called a basis for the category. From Section 3, a basis is therefore a subset of the set of atoms of $C$.

**Theorem 4.3.** A category is a left Rees category if and only if it is isomorphic to the Zappa-Szép product of a free category by a groupoid.

**Proof.** We shall sketch the proof. Let $C$ be the left Rees category. Choose a basis $X$ for $C$. Every element of $C$ can be written uniquely as a product of elements of $X$ followed by an invertible element. The subcategory $X^\ast$ generated by $X$ is free. Thus $C = X^\ast G$. The Zappa-Szép product representation then readily follows. □

The following theorem describes the precise circumstances under which Zappa-Szép products arise.

**Theorem 4.4.** Let $A$ be a category with subcategories $B$ and $C$. We suppose that $A_0 = B_0 = C_0$ and that $C = AB$, uniquely. Then $C$ is isomorphic to the Zappa-Szép product $A \triangleleft\triangleleft B$.

**Proof.** We sketch out the proof. Suppose that $ba$ is defined where $b \in B$ and $a \in A$. Then $ba = a'b'$ for uniquely determined elements $a' \in A$ and $b' \in B$. We define $ba = (b \cdot a)b|_a$. It is immediate that axioms (C1), (C2) and (C3) hold. The equality $a = d(a)a$ yields the axioms (SS1) and (SS4); the equality $b = br(b)$ yields axioms (SS3) and (SS5); the equality $b_1(b_2a) = (b_1b_2)a$ yields axioms (SS2) and (SS7); the equality $b(a_1a_2) = (ba_1)a_2$ yields axioms (SS6) and (SS8). We may therefore construct the category $A \triangleleft\triangleleft B$. We define a map $C \rightarrow A \triangleleft\triangleleft B$ by $c \mapsto (a, b)$ if $c = ab$. It is now straightforward to check that this determines an isomorphism of categories. □

The following result shows that the theory in this paper can be regarded as a generalization of the theory of free categories.

**Theorem 4.5.** The free categories are precisely the left Rees categories with a trivial groupoid of invertible elements and precisely the equidivisible categories with length functors having trivial groupoids of invertible elements.

**Proof.** We need only prove that an equidivisible category with length functor having a trivial groupoid of invertible elements is left cancellative. But this essentially follows from Lemma 2.3. □

**Remark 4.6.** Cancellative equidivisible monoids with trivial groups of units do not have to be free; see, for example, Example 1.8 of Chapter 5 of [14]. This example also suggests that studying equidivisible categories with more general kinds of length functors may be interesting.
5. Diagrams of partial homomorphisms

From now on, the groupoids involved in bimodules will always be totally disconnected.

We showed in the previous section, that covering bimodules are the building blocks on left Rees monoids. In this section, we want to describe how covering bimodules may be constructed. This generalizes a construction due to Nekeshevych in the case where the acting groupoid is actually a group [24].

Let \( D \) be a directed graph. An edge \( x \) from the vertex \( e \) to the vertex \( f \) will be written \( e \xrightarrow{x} f \). With each vertex \( e \) of \( D \) we associate a group \( G_e \), called the vertex group, and with each edge \( e \xrightarrow{x} f \), we associate a surjective homomorphism \( \phi_e : (G_e)_x^+ \to (G_f)_x^- \) where \((G_e)_x^+ \leq G_e \) and \((G_f)_x^- \leq G_f \). In other words, with each edge \( e \xrightarrow{x} f \), we associate a partial homomorphism \( \phi_e \) from \( G_e \) to \( G_f \). We call this structure a diagram of partial homomorphisms. If all the \( \phi_e \) are isomorphisms then we shall speak of a diagram of partial isomorphisms. For brevity, we shall say that \( D \) is the diagram of partial homomorphisms though of course it is defined only by the totality of data.

Let \( D_1 \) and \( D_2 \) be two diagrams of partial homomorphisms having the same vertex sets, and edge sets that differ only in labelling and the same vertex groups. We say these two diagrams of partial homomorphisms are conjugate, if for each edge \( e \xrightarrow{\alpha} f \) in \( D_1 \) and corresponding edge \( e \xrightarrow{\beta} f \) in \( D_2 \) there are inner automorphisms \( \alpha_{x,y} : G_e \to G_e \) and \( \beta_{x,y} : G_f \to G_f \) such that \( \alpha_{x,y}((G_e)_x^+) = (G_e)_y^+ \) and \( \beta_{x,y}((G_f)_x^-) = (G_f)_y^- \) and \( \beta_{x,y}\phi_x = \phi_y\alpha_{x,y} \).

**Remark 5.1.** The above definition could be generalized a little by assuming only that the underlying graphs were isomorphic.

The goal of this section is to prove that covering bimodules and diagrams of partial homomorphisms are different ways of describing the same object.

Let \( (G, X, G) \) be a covering bimodule where \( G = \bigcup_{e \in V} G_e \). Define a relation \( \mathcal{C} \) on \( X \) by \( x \mathcal{C} y \) if and only if \( y = gxh \) for some \( g, h \in G_e \). Observe that because \( G \) is a disjoint union of groups if \( x \mathcal{C} y \) then \( x \) and \( y \) are parallel. Clearly, \( \mathcal{C} \) is an equivalence relation.

**Lemma 5.2.** Let \( (G, X, G) \) be a covering bimodule where \( G = \bigcup_{e \in V} G_e \).

1. With each transversal of the \( \mathcal{C} \)-classes, we may associate a diagram of partial homomorphisms.
2. Different transversals yield conjugate diagrams of partial homomorphisms.
3. The bimodule is bifree if and only if the associated diagram is a diagram of partial isomorphisms.

**Proof.**

(1). Choose a transversal \( E \) of the \( \mathcal{C} \)-classes. Define the directed graph \( D \) to have as vertices the set of identities of \( G \) and edges the elements of \( E \). The group associated with the vertex \( e \) is the group \( G_e \). Consider now the edge \( e \xrightarrow{x} f \). We define

\[
(G_e)_x^+ = \{ g \in G_e : gx = xh \text{ for some } h \in G_f \}
\]

and

\[
(G_f)_x^- = \{ h \in G_f : gx = xh \text{ for some } g \in G_e \}
\]

This relation was used by Paul Cohn whence the choice of notation.
and $\phi_x(g) = h$ if $g \in (G_e)^+_x$ and $gx = xh$ where $h \in G_f$. In a covering bimodule the right action is free and so $\phi_x$ is a function and, in fact, a homomorphism. We have therefore constructed a diagram of partial homomorphisms.

(2). Let $E'$ be another transversal. We denote the elements of $E$ by $x_i$ where $i \in I$ and the elements of $E'$ by $y_i$ where $i \in I$ and assume that $x_i \in C y_i$. Choose $g_i, h_i \in G$ such that $x_i = g_i y_i h_i$. Let $e \overset{x_i}{\rightarrow} f$ and $g \in (G_e)^+_x$. Then

$$(g_i^{-1} g g_i) y_i = y_i (h_i \phi_x(g) h_i^{-1}).$$

Thus

$$h_i \phi_x(g) h_i^{-1} = \phi_y(g_i^{-1} g g_i).$$

Define inner automorphisms $\beta_i(-) = h_i - h_i^{-1}$ and $\alpha_i(-) = g_i^{-1} - g_i$ where $\alpha_i: G_e \to G_e$ and $\beta_i: G_f \to G_f$. We therefore have $\beta_i \phi_x = \phi_y \alpha_i$ and $\alpha_i ((G_e)^+_x) = (G_e)^+_y$ and $\beta_i((G_f)^+_x) = (G_f)^+_y$. Thus the two diagrams are conjugate

(3). Suppose that the bimodule is bifree. By definition $gx = x\phi_x(g)$ when $g \in G^+_x$. Suppose that $g_1 x = g_2 x$. Then $x = (g_1^{-1} g_2) x$ and so by left freeness we have that $g_1^{-1} g_2$ is an identity and so $g_1 = g_2$. Thus $\phi_x$ is injective and so an isomorphism. The proof of the converse is straightforward.

We now show how to go in the opposite direction. Let $D$ be a diagram of partial homomorphisms. Let $G = \bigcup_{e \in V} G_e$ be the disjoint union of the vertex groups regarded as a groupoid. Denote the set of edges by $E$. Let $G * E * G$ be the set of triples $(g, x, h)$ where $e \overset{x_i}{\rightarrow} f$ and $g \in G_e$ and $h \in G_f$. We define a relation $\equiv$ on the set $G * E * G$ as follows: $(g_1, x_1, h_1) \equiv (g_2, x_2, h_2)$ if and only if $x_1 = x_2 = x$, say, $g_2^{-1} g_1 \in (G_e)^+_x$ and $\phi_x(g_2^{-1} g_1) = h_2 h_1^{-1}$.

**Lemma 5.3.** With the above definition, $\equiv$ is an equivalence relation. Denote the $\equiv$-class containing $(g, x, h)$ by $[g, x, h]$. Then we get a covering bimodule $B(D)$ when we define $g [g_1, x, h_1] = [g g_1, x, h_1]$ when $\exists g_1$ and $[g_1, x, h_1] h = [g_1, x, h_1] h$ when $\exists h_1 h$. The diagram of partial homomorphisms associated with this covering module is conjugate to $D$.

**Proof.** We shall just prove the last part since the other proofs are routine. We choose the transversal $[e, x, f]$ where $e \overset{x_i}{\rightarrow} f$ is an edge of the diagram $D$. It is now easy to check that we get back exactly the diagram of partial homomorphisms we started with.

**Lemma 5.4.** Let $(G, X, G)$ be a covering bimodule, let $E$ be a transversal of the $C$-classes and let $D$ be the associated diagram of partial homomorphisms. Then the bimodule $X$ is isomorphic to the bimodule $B(D)$.

**Proof.** Let $x \in X$. Then $x \in C y$ for a unique $y \in E$. By definition, $x = g_1 y h_1$ for some $g_1, h_1 \in G$. Suppose also that $x = g_2 y h_2$. Then $g_2^{-1} g_1 y = y h_2 h_1^{-1}$. It follows that $(g_1, y, h_1) \equiv (g_2, y, h_2)$. We may therefore define a function $\theta: X \to B(D)$ by $\theta(x) = [g_1, x, h_1]$. It remains to show that this is a bijection and an isomorphism of bimodules both of which are now straightforward.

We summarize the results of this section in the following theorem.

**Theorem 5.5.** With each diagram of partial homomorphisms $D$ we may associate a covering bimodule $B(D)$ and every covering bimodule is isomorphic to one of this form. Diagrams of partial isomorphisms correspond to bifree covering bimodules.
6. Presentations of skeletal left Rees categories

In this section, we shall prove that every skeletal left Rees category has a presentation of a particular form. This presentation is then the final link in showing that the theory of graphs of groups is a special case of the theory of skeletal left Rees categories.

Let $C$ be a skeletal left Rees category. Then for each atom $x \in \epsilon CF$, we have as before the following definitions:

$$(G_c)_x^+ = \{ g \in G_c : gx = xh \text{ for some } h \in G_f \},$$

and

$$(G_f)_x^- = \{ h \in G_f : gx = xh \text{ for some } g \in G_c \},$$

and

$$\phi_x(g) = h \text{ if } g \in (G_c)_x^+ \text{ and } h \in (G_f)_x^- \text{ and } gx = xh.$$ 

We have that $(G_c)_x^+ \leq G_c$ and $(G_f)_x^- \leq G_f$ and $\phi_x$ is a surjective homomorphism.

Lemma 6.1. Let $C$ be a skeletal left Rees category. Then $C$ is right cancellative if and only if all the homomorphisms $\phi_x$ defined above are also injective and so in fact isomorphisms.

Proof. If $C$ is right cancellative, it is immediate that all the homomorphisms $\phi_x$ are injective. We prove the converse using Section 5. Choose a basis for $C$ so that $C = X^*G$ where $X^*$ is a free category and $G$ the groupoid of isomorphisms. Let $a, b, c \in C$ such that $ab = cb$. We shall prove that $a = c$. We have that $a = xg$, $b = yh$ and $c = zk$ where $x, y, z \in X^*$ and $g, h, k \in G$. Thus

$$x(g \cdot y)g|_y h = z(k \cdot y)k|_y h.$$ 

From length considerations, $x = z$. We also have that $g|h = k|y$. It remains only to prove that $g = k$ and we are done. We know that $gy = ky$. Thus $(g^{-1}k)y = y$. Therefore, to prove our result it is enough to prove that $gy = y$ implies that $g$ is an identity. If $y$ is an atom then the result is immediate. We assume the result holds for all elements $y$ whose length is at most $n$ and prove it for those elements of length $n + 1$. Let $y$ have length $n + 1$. Let $y = uv$ where $v$ has length 1. Then $gy = guv = (g \cdot u)(g|_u \cdot v)g|_w = y = uv$. Thus $g|_y = g|_w = r(g)$. Also $u = g \cdot u$ and $v = g|_u \cdot v$. We have that $g|_u v = vg|_w$. It follows that $g|_u$ is an identity. We then have $gu = u$ and so $g$ is an identity, as required. \qed

The proof of the following is immediate from the definitions and Lemma 2.1

Lemma 6.2. Let $C$ be a skeletal left Rees category. If $x$ and $y$ are atoms then $x \not\in y$ if and only if $x \not\in y$.

The following two results will enable us to refine the way that we choose a basis for a left Rees category. We refer the reader to Section 5 for the structure theory of left Rees categories which we use here.

Lemma 6.3. Let $C$ be a left Rees category and let $X$ be a transversal of the generators of the submaximal principal right ideals so that $C = X^*G$. Let $x, y \in X$. Then $x \not\in y$ if and only if $y = g \cdot x$.

Proof. Suppose that $y = g \cdot x$. Then $CyC = Cg \cdot xC = C(g \cdot x)g|_x C = CgxC = CxC$. Thus $x \not\in y$. Conversely, suppose that $x \not\in y$. Then $y = gxh$ for some $g, h \in G$. Thus $y = (g \cdot x)g|_x h$. But by the uniqueness of the factorization, we must have that $y = g \cdot x$. \qed
Lemma 6.4. Let $C$ be a skeletal left Rees category and let $x$ be an atom where $e \xrightarrow{f} e$. Choose a coset decomposition $G_e = \bigcup_{i \in I} g_i(G_e)^+$. The set $\{g_i x : i \in I\}$ consists of pairwise $\mathcal{R}$-inequivalent elements and every atom $f$-related to $x$ is $\mathcal{R}$-related to one of these elements.

Proof. Suppose that $g_i x \not\sim g_j x$. Then $g_i x = g_j x y$ for some $y \in G_f$. Then $x = (g_i^{-1} g_j) x y$. But $g_i^{-1} g_j \in (G_e)^+$. By the definition of coset representatives, we have that $g_i = g_j$, as required. Let $y \sim x$ where $y$ is any atom. Then $y = g x h$ for some $g, h \in G$. Write $g = g_i a$ where $a \in (G_e)^+_x$. Then $y = g_i a x h = g_i x \phi_x(a) h$. Hence $y \not\sim g_i x$.

We now define a special type of basis for a left Rees category $C$. For each hom-set $eCf$, choose a transversal $Y$ of the $f$-classes of the atoms in $eCf$. For each $x \in Y$, choose a coset decomposition $G_x = \bigcup_{i \in I} g_i(G_x)^+$. Denote by $T_x^+$ the transversal $\{g_i : i \in I\}$. We shall assume that the appropriate identities are always elements of these transversals. For each such $x$, we have a set of atoms $\{kx : k \in T_x^+\}$ and the totality of those atoms as $x$ varies over $Y$ then provides a basis for $C$. We call a basis constructed in this way a co-ordinatization. Observe that it contains two components: a transversal of the $f$-classes of the set of atoms of $C$, we call this an atomic transversal, and for each atom $x$ in that transversal a set of coset representatives $T_x^+$.

Theorem 6.5 (Skeletal left Rees categories and their diagrams).

(1) Let $D$ be a diagram of partial homomorphisms. Then we may construct a skeletal left Rees category $C$ equipped with an atomic transversal such that the diagram of partial homomorphisms constructed from the bimodule of atoms of $C$ relative to that transversal is equal to $D$. If $D$ is a diagram of partial isomorphisms then $C$ is a Rees category.

(2) Let $C$ be a skeletal left Rees category equipped with an atomic transversal. Then we may construct a diagram of partial homomorphisms $D$ using that transversal such that the tensor category of the covering bimodule associated with $D$ is isomorphic to $C$. If $C$ is a Rees category then $D$ is a diagram of partial isomorphisms.

Proof. (1). Let $D$ be a diagram of partial homomorphisms. Then by Lemma 5.3 we may construct a covering bimodule $B(D)$. This bimodule has a transversal with respect to which the associated diagram of partial homomorphisms is $D$. By Theorem 3.3 we may construct a left cancellative, skeletal equidivisible category equipped with a length functor $T(B(D))$. By Proposition 2.10 this is a skeletal left Rees category whose associated bimodule is $B(D)$. It follows by our results proved earlier that if we start with a diagram of partial isomorphisms we obtain a Rees category.

(2). Let $C$ be a skeletal left Rees category equipped with an atomic transversal. Then by Proposition 2.10 and Lemma 3.1 we may construct a covering bimodule $B$ from the set of atoms acted on by the groupoid of invertible elements. From such a bimodule and an atomic transversal, we may construct a diagram $D$ of homomorphisms by Lemma 5.2. But then by part (3) of Theorem 3.3 the category $C$ is isomorphic to the tensor category $T(B)$ where $B$ is essentially $B(D)$. It follows by our results proved earlier that if we start with a Rees category we obtain a diagram of partial isomorphisms. □
Remark 6.6. We have now seen that diagrams of partial homomorphisms and skeletal left Rees categories are two ways of viewing the same structure, the path from diagram to category taking in both the construction of bimodules from diagrams and then tensor categories from bimodules. We shall now describe a direct construction of left Rees categories from diagrams of partial homomorphisms.

Let $D$ be a diagram of partial homomorphisms. We shall define a category $\langle D \rangle$ by means of a presentation constructed from $D$. We first construct a new directed graph $D'$ from $D$. This contains the directed graph $D$ but at each vertex we adjoin additional loops labelled by the elements of $G_e$. We construct the free category $(D')^*$. We denote elements of this category thus $x_1 \cdot \ldots \cdot x_m$ where the $x_i$ are edges that match. We now factor out by two kinds of relations: those of the form $g \cdot h = gh$ where $g, h \in G_e$; and those of the form $g \cdot x = x \cdot \phi_x(g)$ where $g \in (G_e)^+_x$. The resulting category is denoted by $\langle D \rangle$.

Theorem 6.7 (Presentation theorem). Let $D$ be a diagram of partial homomorphisms. Then the category $\langle D \rangle$ is a skeletal left Rees category isomorphic to the category obtained from $D$ by constructing the tensor category of the covering bimodule associated with $D$. The category $\langle D \rangle$ is a skeletal Rees category if and only if $D$ is a diagram of partial isomorphisms.

Proof. Denote by $C$ the skeletal left Rees category constructed from the diagram of partial homomorphisms $D$ according to part (1) of Theorem 6.5. This category satisfies all the defining relations of the category $\langle D \rangle$ and so $C$ is a functorial image of $\langle D \rangle$.

It remains therefore only to show that this functor is injective. We work first in the category $C$. In the category $C$ we choose an atomic transversal whose elements can be identified with the edges of the diagram $D$. For each such atom $x$ choose a coset decomposition $G_e = \bigcup_{i \in I} g_i(G_e)_x^+$. This leads to a co-ordinatization for $C$ as described earlier in this section. We now appeal to the structure theory of Section 5, and deduce that every element of $x$ can be written as a unique product of atoms in this co-ordinatization followed by an invertible element. Call this a normal form.

We may paraphrase the above theorem as saying that every skeletal left Rees category may be presented by a diagram of partial homomorphisms.

The following theorem was originally suggested by [31, 32] and [4] but the proof is a straightforward generalization of Higgins’s main result [7] and at the same time this shows how our approach is related to his.

Theorem 6.8. Every skeletal Rees category may be embedded in its universal groupoid.

Proof. Let $C$ be a skeletal Rees category. By Theorem 6.7 we may assume that $C = \langle D \rangle$ where $D$ is a diagram of partial isomorphisms. Denote by $\mathcal{G}$ the universal groupoid of $C$. Our goal is to obtain a normal form for the elements of $\mathcal{G}$. To that end, let $e \xrightarrow{\gamma} f$. Choose a coset decomposition $G_e = \bigcup_{i \in I} g_i(G_e)_x^+$ to obtain the transversal $T^+_x$ as before. However, now that we want to work in a groupoid, we
shall also need a coset decomposition $G_f = \bigcup_{j \in J} h_j(G_f)$ to obtain the transversal $T_\sigma^-$. In both transversals, we assume that the identity elements of their respective groups have been chosen. We may now follow the proof of the theorem in Section 3 of [7] by defining suitable normal forms since finiteness plays no role in Higgins’s proof. This shows that $C$ is in fact embedded in $G$. □

Example 6.9. Free monoids on $n$ generators are Rees monoids. The group constructed according to the above theorem is the free group on $n$ generators.

Remark 6.10. Alternative ways of proving the above theorem are suggested by the results of Cohn [4] and von Karger [31, 32] though we do not pursue these here.

Remark 6.11. Higgins uses the relations $g \cdot x = x \cdot \phi_x(g)$ in two directions to construct the fundamental groupoid of a diagram of partial isomorphisms. We, on the other hand, use these relations in one direction only to construct a Rees category. It is then evident that the Rees category sits inside the fundamental groupoid. What is perhaps surprising is that these cancellative categories can be abstractly characterized.

7. A categorical approach to Bass-Serre theory

We shall now explain the connection between the theory we have developed and the theory of graphs of groups. Our references for this theory are [29] and [34]. We start with an observation. A graph of groups equipped with a given orientation is essentially the same thing as a diagram of partial isomorphisms where the directed graph underlying it is finite and weakly connected in the sense that as a graph it is connected. Essentially, graphs of groups represent partial isomorphisms by means of relations and so are unoriented. We shall call the diagrams of partial isomorphisms that arise from graphs of groups equipped with an orientation Serre diagrams of partial isomorphisms.

Let $C$ be a category. A zig-zag joining the identity $e$ to the identity $f$ is determined by a sequence of identities $e = e_1, \ldots, e_n = f$ such that for each consecutive pair of identities $e_i$ and $e_{i+1}$ we have that either $e_i Ce_{i+1}$ or $e_{i+1} Ce_i$ is non-empty. We say that a category $C$ is connected if any two identities $e, f \in C$ are joined by a zig-zag.

A skeletal Rees category $C$ is called a Serre category if it satisfies the following conditions.

(S1): The number of identities in $C$ is finite and nonzero.
(S2): In each hom-set, the number of $\mathcal{J}$-classes of atoms is finite.
(S3): $C$ is connected.

The following theorem is now immediate from the above definitions and what we proved in the previous section.

Theorem 7.1 (Graphs of groups as categories).

1. There is a correspondence between graphs of groups with a given orientation and Serre diagrams of partial isomorphisms.
2. There is a correspondence between Serre diagrams of partial isomorphisms and Serre categories.
3. The fundamental groupoid of a graph of groups with a given orientation is isomorphic to the universal groupoid of the Serre category constructed from
the diagram of partial isomorphisms associated with the oriented graph of groups.

(4) The universal groupoid of a Serre category is connected.

The following two examples are the basic building blocks of Bass-Serre theory.

Examples 7.2.

(1) **HNN extensions.** These are constructed from Rees monoids. If we make the additional assumption that the monoid has the property that any two atoms are \( \mathcal{J} \)-related, then the universal groups are precisely HNN extensions with one stable letter. This case was the subject of our paper [18] and motivated the work of the current paper.

(2) **Amalgamated free products.** The building blocks of these are \((G, H)\)-bimodules \(X\) where \(G\) and \(H\) are both groups. We say that such a bimodule is **irreducible** if there exists \(x \in X\) such that \(GxH = X\). There is a bijective correspondence between conjugacy classes of partial isomorphisms from \(G\) to \(H\) and isomorphism classes of irreducible, bifree \((G, H)\)-bisets. Consider now any irreducible, bifree biset \((G, X, H)\). Choose and fix \(x \in X\). Let \(A = G_x, B = H_x\) and \(\theta = \phi_x\) be the associated partial isomorphism. We may regard \((G, X, H)\) as a cancellative category with two identities in the following way. We let the identity of \(G\), \(1_G\) say, be one of the identities and the identity of \(H\), \(1_H\) say, the other. Thus we take the disjoint union \(G \cup X \cup H\). The products in \(G\) and \(H\) are the group products. The product \(gx\) is the action and the product \(xh\) is the action. We denote by \(C\) the above biset regarded as a category in this way. Then \(C\) is a Rees category and is in fact a Serre category. In this case, the tensor product construction is essentially degenerate. The universal groupoid of \(C\) is connected and any vertex group is isomorphic to \(G *_{\theta} H\) the amalgamated free product of \(G\) and \(H\) via the identifying partial isomorphism \(\theta\).

7.1. **Concluding remarks.** Andrew Duncan (Newcastle, UK) has pointed out that there are some interesting parallels between our theory and Stallings theory of pregroups as described in [8]. We do not know, at this point, whether we can derive that theory from ours.

The Serre tree of a graph of groups may be constructed from the way the associated Rees category is embedded in its universal group [34]. This is strongly reminiscent of the categorical interpretation of the proof of McAlister’s \(P\)-theorem described in [10]. In fact, using [12], every Rees category gives rise to an inverse semigroup which is strongly \(E^*\)-unitary. The maximum enlargement theorem, described in Chapter 8 of [15], may then be used to construct the Serre tree. Nick Gilbert (Heriot-Watt, UK) has a different approach to Bass-Serre theory [6] which uses ordered groupoids. We do not yet know how his approach can be reconciled with the one we have sketched above.

All cancellative equidivisible categories may be embedded in groupoids; see [31] although a direct proof using the ideas of [11] would be useful. Motivated by group theory, one might consider the structure of such categories equipped with other kinds of length functors.
References

[1] J.-C. Birget, Iteration of expansions — unambiguous semigroups, J. Pure Appl. Algebra 34 (1984), 1–55.
[2] M. R. Bridson, A. Haefliger, Metric spaces of non-positive curvature, Springer, 1999.
[3] M. G. Brin, On the Zappa-Szép product, Comm. Algebra 33 (2005), 393–424.
[4] P. M. Cohn, Free rings and their relations, Second Edition, Academic Press, 1985.
[5] B. Elkins, J. A. Zilber, Categories of actions and Morita equivalence, Rocky Mountain Journal of Mathematics 6 (1976), 199–225.
[6] N. D. Gilbert, A P-theorem for ordered groupoids, in Proc. Intl. Conf. Semigroups and Formal Languages (eds. J. M. André et al) (2007), 84–100.
[7] Ph. J. Higgins, The fundamental groupoid of a graph of groups, J. Lond. Math. Soc. (2) 13 (1976), 145–149.
[8] A. H. M. Hoare, Pregroups and length functions, Math. Proc. Camb. Phil. Soc 104 (1988), 21–30.
[9] J. M. Howie, Fundamentals of semigroup theory, Clarendon Press, Oxford, 1995.
[10] H. James, M. V. Lawson, An application of groupoids of fractions to inverse semigroups, Period. Math. Hungar. 38 (1999), 43–54.
[11] P. T. Johnstone, On embedding categories in groupoids, Math. Proc. Camb. Phil. Soc. 145 (2008), 273–294.
[12] D. G. Jones, M. V. Lawson, Graph inverse semigroups: their characterization and completion, arXiv:1106.3644
[13] W. Krieger, On subshifts and semigroups, Bull. Lond. Math. Soc. 38 (2006), 617–624.
[14] G. Lallement, Semigroups and combinatorial applications, John Wiley, 1979.
[15] M. V. Lawson, Inverse semigroups, World Scientific, 1998.
[16] M. V. Lawson, A correspondence between a class of monoids and self-similar group actions I, Semigroup Forum 76 (2008), 489–517.
[17] M. V. Lawson, Semigroups related to subshifts of graphs, in Semigroups, acts and categories with applications to graphs, Math. Stud. (Tartu), 3, Est. Math. Soc., Tartu, (2008), 92–107.
[18] M. V. Lawson, A. Wallis, A correspondence between a class of monoids and self-similar group actions II, preprint.
[19] M. Lothaire, Combinatorics on words, Addison-Wesley Publishing Co., Reading, Mass., 1983.
[20] R. C. Lyndon, P. E. Schupp, Combinatorial group theory, Springer Verlag, Berlin, Heidelberg, 2001.
[21] S. W. Margolis, J. E. Pin, Inverse semigroups and extensions of groups by semilattices, J. Algebra 110 (1987), 277–297.
[22] S. Mac Lane, Categories, second edition, Springer, 1998.
[23] J. D. McKnight, A. J. Storey, Equidivisible semigroups, J. Algebra 12 (1969), 24–48.
[24] V. Nekrashevych, Self-similar groups, AMS, 2005.
[25] M. Nivat, J.-F. Perrot, Une généralisation du monoïde bicyclique, Comptes Rendus de l’Académie des Sciences de Paris 271 (1970), 824-827.
[26] J.-F. Perrot, Contribution a l’étude des monoides syntactiques et de certains groupes associés aux automates finis, Thése Sc. math. Paris, 1972.
[27] J.-F. Perrot, Une famille de monoïdes inversifs 0-bisimples généralisant le monoïde bicyclique, Séminaire Dubreil. Algèbre 25 (1971-1972), 1–15.
[28] D. Rees, On the ideal structure of a semi-group satisfying a cancellation law, Quart. J. Math. Oxford Ser. 19 (1948), 101–108.
[29] J.-P. Serre, Trees, Springer, 1980.
[30] R. Street, Quantum groups, CUP, 2007.
[31] B. von Karger, Temporal algebra, Habilitationsschrift, Christian-Albrechts-Universität, Kiel, 1997.
[32] B. von Karger, Embedding sequential calculus in relational calculus, Preprint dated March 1994.
[33] A. Wallis, Semigroup and category-theoretic approaches to partial symmetry, PhD Thesis, Herriot-Watt University, July 2013.
[34] R. Weidmann, Combinatorial and geometric group theory, lecture notes, Christian-Albrechts-Universität, Kiel, 2013.
