Monoid Pictures and Finite Derivation Type

by

David Gains, B. Math.

A thesis submitted to

the faculty of Graduate Studies and Research

in partial fulfilment of

the requirements for the degree of

Master of Science.

School of Mathematics and Statistics

Ottawa-Carleton Institute for Mathematics and Statistics

Carleton University

Ottawa, Ontario, Canada

September 2005

©copyright

2005, David Gains
Abstract

A property of monoids called finite derivation type (FDT) is defined. The major results are that FDT is an invariant, and that every monoid with a finite complete presentation has FDT. Monoid pictures are developed and are used to define and prove the relevant terms and theorems. The thesis is a synthesis of work done by Stephen Pride, Victor Guba, Mark Sapir, Craig Squier, Friedrich Otto and Yuji Kobayashi, who are all responsible for original definitions of FDT and monoid pictures, and proofs of the subsequent theorems.
Acknowledgements

To my supervisors Dr. John Poland and Dr. Benjamin Steinberg. To my wife for pushing me forward and for endless patience.
Table of Contents

Abstract ii
Acknowledgements iii
Table of Contents iv
List of Figures v
1 Introduction 1
2 Preliminaries 4
  2.1 Directed graphs and 2-complexes 4
  2.2 String rewriting systems 7
  2.3 Monoid invariants 12
3 Monoid pictures 14
  3.1 Definition 16
  3.2 Monoid pictures and monoid presentations 20
  3.3 Homotopy relation on the Squier complex 33
  3.4 Monoids of finite derivation type 38
  3.5 A result from homology 56
4 References 61
**List of Figures**

3.1.1: an atomic monoid prepicture. 17
3.1.2: equivalent atomic monoid prepictures. 19
3.2.1: an atomic monoid picture. 21
3.2.2: a positive atomic picture. 22
3.2.3: an atomic picture. 23
3.2.4: inverse of the atomic picture displayed in figure 3.2.3. 23
3.2.5: an atomic picture featuring the rewriting rule $a^3 = 1$. 24
3.2.6: a monoid picture. 24
3.2.7: inverse of the monoid picture displayed in figure 3.2.6. 26
3.2.8: a spherical picture. 26
3.2.9: an empty path. 27
3.2.10: concatenation with the empty picture. 28
3.2.11: distinct pictures with identical initial and terminal vertices. 29
3.2.12: $w.P$. 30
3.2.13: the sum of two atomic pictures. 31
3.2.14: the 2-cell $[P, Q]$. 32
3.3.1: pulling and pushing rectangles in disjoint pictures. 36
3.3.2: $P$ and $Q$. 37
3.3.3: interchanging disjoint rectangles. 37
3.4.1: a critical pair. 47
3.4.2: not a critical pair. 47
3.4.3: a resolved critical pair. 48

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
3.4.4: pictures in the lemma.

3.4.5: the 2-cell described by the second case.

3.4.6: picture for the new lemma.

3.4.7: pictures for the proof of lemma 3.4.8.

3.4.8: splitting $P$.

3.4.9: applying lemma 3.4.8.
Chapter 1

Introduction

In 1987 Professor Craig C. Squier wrote about an invariant property of monoids. Tragically he passed away in 1992 and his paper lay unpublished until 1994 when it was revised by Professors Friedrich Otto and Yuji Kobayashi for the journal *Theoretical Computer Science* [15]. According to these authors (whom we shall refer to collectively as SOK), monoids exhibiting the invariant property are said to be of finite derivation type (FDT). Discovery of FDT was an early step in the (ongoing) attempt of SOK to characterise monoids with finite complete presentations in terms of a set of invariant properties. SOK study the string rewriting systems associated with these presentations, and their definitions and theorems are stated and proved in a combinatorial language.

Professor Stephen J. Pride established monoid theoretic results analogous to some known in combinatorial group theory, and a few of these found their way into his 1995 publication [10]. In that paper Pride introduced innovative diagrammatic devices called monoid pictures, and with them he proved his theorems and, independent of SOK, discovered the FDT invariant. Because of his
monoid pictures, Pride’s formulation of FDT is different than that of SOK, and his exposition is unique and elegant.

After an introduction to Squier’s work, Pride and his student X. Wang published his account of FDT and a new monoid invariant called finite homological type (FHT) [18]. The paper included additional original contributions to monoid theory all resulting from experiments involving monoid pictures.

Pride’s pictures constitute the 1-skeleton of a certain 2-complex that has come to be known as the Squier complex. 2-complexes associated with semigroup presentations are a specialty of the team comprised of Professors Victor Guba and Mark Sapir, and some of these are exposed in their landmark studies of the so called diagram groups [4] [5] (but especially [4]). In particular, Guba and Sapir give their own definitions of monoid picture and Squier complex in these papers.

This thesis is a synthesis of the work of Pride, SOK, and Guba and Sapir, but the focus is on the ideas of Pride. In chapter 2 we state common definitions and facts concerning graphs, string rewriting systems and 2-complexes relevant to a study of monoid pictures and monoids of FDT. Chapter 3 gives a new definition of monoid picture, defines FDT as Pride does, and supplies new picture proofs of the FDT theorems.

My motivation for this project does not coincide with that of any of the researchers referenced above. My interests are in logic and constructivist mathematics, and Pride’s work seems to provide
1. an excellent example of a diagrammatic system of reasoning, and
2. an arena in which to build constructivist proofs.

With respect to the second point I have tried to

1. give the proofs of chapter 3 an algorithmic or procedural feel,
2. avoid the use of a controversial set theory (in particular I have tried not to invoke anything like a ZF version of the axiom of choice),
3. avoid reasoning involving a law of excluded middle.

With respect to the first point, diagrammatic logic is currently a bit of a hot (if not niche) topic [6] [12] [19]. I would have enjoyed shaping Pride’s diagrammatic techniques into a type of formal logic, but time and space did not permit.
Chapter 2

Preliminaries

Here we state definitions and theorems about graphs, string rewriting systems, monoid presentations and 2-complexes. The account is by no means complete - it is sufficient only as background to chapter 3, and theorems are often given without proof. The results are common however and sources are referenced.

Credit must be given to Serre as we adopt his definition of directed graph [16]. Similarly we acknowledge Guba and Sapir for their definition of 2-complex which differs slightly from the usual one [3] and significantly from Pride’s [18]. We have chosen Guba and Sapir’s version because they have conducted a thorough analysis of its consequences in particular its bearing upon their formulation of monoid picture [4] [5].

2.1 Directed graphs and 2-complexes

A directed graph $\Gamma = (v, e, \iota, \tau, -1)$ consists of a set $v$ of vertices, a set $e$ of edges, functions $\iota$ and $\tau$ mapping edges to vertices, and an involution $-1$ on edges called inverse. For any edge $e$, $\iota e$ is called its initial vertex and $e\tau$ is called
its terminal vertex. Inverse must satisfy \( e \neq e^{-1}, e = (e^{-1})^{-1} \), and \( e^{-1} \epsilon = e \tau \) for every edge \( e \).

A path in a graph is a finite sequence of edges \( p = e_1 e_2 \ldots e_n \) with \( e_{i+1} \epsilon = e_i \tau \) for each \( 1 \leq i < n \). We can talk about initial and terminal vertices of paths by defining \( p \epsilon \) and \( p \tau \) to be \( e_1 \epsilon \) and \( e_n \tau \) respectively. In particular, a path with identical initial and terminal vertices is identified as closed. The notion of inverse can also be extended to paths by setting \( p^{-1} = e_n^{-1} \ldots e_2^{-1} e_1^{-1} \). At every vertex \( v \) there is the empty path \( 1_v \) with no edges. The initial and terminal vertices of \( 1_v \) are both \( v \), and \( 1_v^{-1} = 1_v \).

Paths \( p \) and \( q \) of a graph with \( p \tau = q \epsilon \) can be concatenated to form a new path \( pq \): if \( e_1 \ldots e_n \) and \( f_1 \ldots f_m \) are the edge sequences of \( p \) and \( q \) respectively, then \( e_1 \ldots e_n f_1 \ldots f_m \) is the edge sequence of \( pq \). Note that the concatenation operation is associative.

Let \( \Gamma \) be a graph and consider a subset \( e^+ \) of the edge set consisting of exactly one of \( e \) or \( e^{-1} \) for each edge \( e \). Such a set is called an orientation of \( \Gamma \). \( \Gamma \) is said to be oriented when a specific orientation \( e^+ \) of \( \Gamma \) is chosen, and the edges of the orientation are called positive edges. A positive path of \( \Gamma \) is a path with positive edges only.

Suppose that \( \Gamma \) and \( \Gamma_1 \) are graphs. A mapping of graphs \( \Gamma \rightarrow \Gamma_1 \) is a pair of functions \( f = (f_v, f_e) \) where \( f_v \) maps each vertex of \( \Gamma \) to a vertex of \( \Gamma_1 \), and \( f_e \) takes each edge of \( \Gamma \) to a path of \( \Gamma_1 \) in such a way that

\[
e_{\epsilon} f_v = e_{\epsilon} \epsilon \tau
\]

\[
e_{\tau} f_v = e_{\tau} \tau
\]
The edge function $f_e$ can be extended to a function $f_p$ on paths of $\Gamma$ in the following way: if $p = e_1 e_2 \ldots e_n$ is a path of $\Gamma$, then take its image under the extension to be $p f_p = e_1 f_e e_2 f_e \ldots e_n f_e$. It is typical to blur the distinction between $f_v$, $f_e$ and $f_p$ by writing $f$ for any of them, and for economy of notation we observe this practice.

A directed 2-complex $\mathcal{K} = (\Gamma, s, \lceil, \rceil, ^{-1})$ consists of a directed graph $\Gamma$, a set $s$ of elements called 2-cells, functions $\lceil$ and $\rceil$ mapping 2-cells to non-empty paths of $\Gamma$, and an involution on 2-cells $^{-1}$ called inverse. If $s$ is a 2-cell of $\mathcal{K}$, then $[s]$ is called its top and $\_s$ is called its bottom. The top and bottom of a 2-cell must have common initial and terminal vertices, and inverse must satisfy the properties $s^{-1} \neq s$, $(s^{-1})^{-1} = s$, and $[s^{-1}] = [s]$ for every 2-cell $s$. If $s$ is a 2-cell, then define its initial vertex $s_\ell$ to be $s_\ell = [s] \ell$ and its terminal vertex $s_\tau$ to be $s_\tau = [s] \tau$.

A 1-path in $\mathcal{K}$ is a path of the graph $\Gamma$. Suppose that $p$ and $q$ are 1-paths of $\mathcal{K}$, and that $s$ is a 2-cell of $\mathcal{K}$ such that $p \tau = s_\ell$ and $s_\tau = q \ell$. The ordered triple $(p, s, q)$ is then called an atomic 2-path of $\mathcal{K}$. For an atomic 2-path $\pi = (p, s, q)$ define its top and bottom 1-paths to be $[\pi^\ell] = (p, [s], q)$ and $[\pi^\tau] = (p, [s], q)$ respectively. A 2-path in $\mathcal{K}$ is thus a sequence $\pi_1 \pi_2 \ldots \pi_n$ of atomic 2-paths with $[\pi_{i+1}] = [\pi_i]$ for each $i < n$. A 1-path $p$ is also called a trivial 2-path, and we set its top and bottom both to $p$.

Orientations are defined for 2-complexes: let $s$ be the set of 2-cells of a 2-complex $\mathcal{K}$ and distinguish in a set $s^+$ exactly one of $s$ or $s^{-1}$ for each 2-cell $s$. 

\[ e^{-1} f_e = (e f_e)^{-1} \]
Thus $s^+$ is called an orientation of $K$, and when we fix such a set we say that $K$ is oriented. The elements of an orientation are called positive 2-cells.

We delay examples of directed 2-complexes until we encounter the central objects of study in chapter 3.

### 2.2 String rewriting systems

Let $x$ be a set and let $r$ be a binary relation on the free monoid $xF$ containing no reflexive and no symmetric pair. In the current context, the ordered pairs of $r$ are called rewriting rules, $x$ is called an alphabet, its members are called letters or generators, and elements of the free monoid $xF$ are known as words or strings. We often write $x^+ = x^{-1}$ for the rewriting rules $(x^+, x^{-1})$ of $r$.

From $r$ we obtain a new relation $\rightarrow$ on $xF$ called the single step derivation by saying that $u \rightarrow v$ if and only if there are words $y$ and $z$, and some rewriting rule $x^+ = x^{-1}$ such that $u$ and $v$ can be written $yx^+z$ and $yx^{-1}z$ respectively. This relation gives rise to a graph called a string rewriting system with vertices the elements of $xF$, edges pairs from $\rightarrow$ or $\rightarrow^{-1}$, and for each edge $e = (u, v)$, $eu = u$, $ev = v$, and $e^{-1} = (v, u)$. For an immediate orientation of the graph, let positive edges be members of the set $\rightarrow$. A word is called irreducible when none of its substrings are the initial vertex of any positive edge.

The reflexive, transitive closure of $\rightarrow$ is written $\Rightarrow$, and the equivalence relation on $xF$ obtained by taking the symmetric, reflexive, transitive closure is denoted $\leftrightarrow$. We say that $v$ is a derivation of $u$, or that $v$ is derived from $u$ if the pair $(u, v)$ belongs to $\leftrightarrow$. Derived pairs are characterised by a theorem.

**Theorem 2.2.1.** $u \leftrightarrow v$ if and only if there is a path with initial vertex $u$ and
CHAPTER 2. PRELIMINARIES

terminal vertex \( v \) in the string rewriting system.

Proof. Define a relation \( \equiv \) on the free monoid by \( u \equiv v \) if and only if there is a path from \( u \) to \( v \) in the string rewriting system. This new relation contains \( \rightarrow \), and it is an equivalence: if \( w \) is a word, then its equivalence class is given by all of the vertices in the connected component containing \( w \) in the string rewriting system. Clearly \( \leftrightarrow \) contains \( \equiv \), so as \( \leftrightarrow \) is the smallest equivalence relation on \( xF^* \) containing \( \rightarrow \), we have that \( \leftrightarrow \) coincides with \( \equiv \). The result follows.

\( \Box \)

An equivalence relation \( \equiv \) on a monoid \( M \) is called a congruence if its equivalence classes are preserved under the multiplication of \( M \). That is, if \([mn] = [m'\cdot n]\) and \([nm] = [nm']\) whenever \( m \equiv m' \) for elements \( m, n \) and \( m' \) of \( M \). This sort of relation gives us a new monoid called the factor monoid denoted \( M/\equiv \) as demonstrated by proof of the following.

**Theorem 2.2.2.** If \( \equiv \) is a congruence relation on a monoid \( M \), then its equivalence classes are the underlying set of some other monoid.

Proof. For equivalence classes \([m]\) and \([n]\) we take their product \([m][n]\) to be the equivalence class \([mn]\). This is a well defined operation since \( \equiv \) is a congruence. Associativity follows from associativity of the multiplication in \( M \), and the monoid unit is \([1]\).

\( \Box \)

By definition the relation \( \leftrightarrow \) is an equivalence on \( xF^* \), but it is also a congruence. To see this, suppose that \( m \leftrightarrow m' \) whence there is a path in the string rewriting system from \( m \) to \( m' \) with edge sequence \( e_1e_2\ldots e_n \). If \( e_i = (u_i, v_i) \), then for
any given free monoid element \( n \), put \( f_i = (u_i, v_i, n, n) \) and \( g_i = (u_i, v_i, n, n) \) to get paths with edge sequences \( f_1 f_2 \ldots f_n \) and \( g_1 g_2 \ldots g_n \) from \( m n \) to \( m'n \) and \( n m \) to \( n m' \) respectively. From theorem 2.2.1 we conclude that \( m n \leftrightarrow m'n \) and \( n m \leftrightarrow n m' \) as required.

Hence given any set \( x \), a large number of monoids are easily obtained from \( x F \) simply by specifying sets of rewriting rules. For a set \( r \) of rewriting rules and the associated relation \( \leftrightarrow \) we write \( P M \) for any monoid isomorphic to \( x F / \leftrightarrow \), and we call the ordered pair \( P = (x, r) \) a (monoid) presentation for \( P M \). By convention \( P \) is written \( \langle x : r \rangle \). A presentation \( P \) is finite if both \( x \) and \( r \) are finite.

Of particular interest are factor monoids arising from complete presentations. A presentation is complete when the underlying string rewriting system contains no infinite path of positive edges, and when for each situation

\[
\begin{array}{c}
\text{\( \ast \)} \\
\text{\( u \)} \\
\text{\( \ast \)} \\
\text{\( w \)} \\
\text{\( \ast \)} \\
\text{\( v \)}
\end{array}
\]

there is a word \( w' \) such that \( u \ast w' \) and \( v \ast w' \). Presentations exhibiting the former property are called Noetherian, and ones respecting the latter are called confluent. A standard result concerning complete monoid presentations is that each congruence class of the factor monoid has a unique, irreducible representative. The more impressive corollary is that the word problem is always solvable in such a monoid.

We say that a presentation is locally confluent when there is a word \( w' \) such
that $u \xrightarrow{*} w'$ and $v \xrightarrow{*} w'$ in each situation

\[
\begin{array}{c}
\text{u} \\
\rightarrow \\
\text{w} \\
\downarrow \\
\text{v}
\end{array}
\]

Occasionally local confluence is no weaker than confluence, and this is supported by the following standard result [1].

**Theorem 2.2.3.** Every Noetherian, locally confluent presentation is confluent.

This is useful in proving theorems because local confluence of a presentation is often more easily demonstrated than confluence.

**Example 2.2.4.** The presentation $\mathcal{P} = \langle a : a^2 = a \rangle$ is both Noetherian and confluent.

**Example 2.2.5.** The presentation $\mathcal{P} = \langle a : a = a^2 \rangle$ is neither Noetherian nor confluent.

This example is due to Pride [11].

**Example 2.2.6.** The presentation $\mathcal{P} = \langle a, b : ab = b, ba^2 = a^3 \rangle$ is Noetherian since all words are related by a finite, positive path to one of the irreducibles $b^n a$, $b^n$, or $a^n$, but fails to be confluent since two distinct irreducibles may be derived from the word $ba^3 b$. 
A complete monoid presentation \( \mathcal{P} = \langle x : r \rangle \) may be further refined: call it canonical if \( x^{-1} \) is irreducible in \( \mathcal{P} \) and \( x^{+1} \) is irreducible in \( \langle x : r - \{x^{+1} = x^{-1}\} \rangle \) for each rewriting rule \( x^{+1} = x^{-1} \) of \( r \).

**Example 2.2.7.** The monoid presentation \( \langle a, b : a^2 = 1, b^2 = b \rangle \) is canonical.

Suppose that \( \mathcal{P} = \langle x : r \rangle \) and \( \mathcal{P}_1 = \langle x_1 : r_1 \rangle \) are monoid presentations. A mapping of presentations \( \mathcal{P} \rightarrow \mathcal{P}_1 \) is a homomorphism \( f : xF \rightarrow x_1F \) such that for each rewriting rule \( x^{+1} = x^{-1} \) of \( \mathcal{P} \) we get that \( x^{+1}f = x^{-1}f \) in the monoid \( \mathcal{P}_1M \), i.e. \( x^{+1}f \leftrightarrow x^{-1}f \) in the string rewriting system for \( \mathcal{P}_1 \). In particular, \( f \) is a retraction if \( x_1 \) is a subset of \( x \), \( r_1 \) is a subset of \( r \), and if \( xf = x \) for each generator \( x \) in \( x_1 \). \( \mathcal{P}_1 \) is called a retract of \( \mathcal{P} \) in this special case.
In general a relation $\rightarrow$ on a set $X$ is said to be Noetherian when there is no infinite chain of relations $x_1 \rightarrow x_2 \rightarrow \cdots$ with $x_i$ in $X$ ($i = 1, 2, \ldots$). Relevant to Noetherian relations and often employed in the study of string rewriting systems is the principal of Noetherian induction. The following statement of this principal is lifted directly from [14].

**Theorem 2.2.8.** Let $X$ be a set, let $\rightarrow$ be a Noetherian relation on $X$, and let $P$ be a predicate on $X$. Suppose that whenever $x \in X$ has the property that every $y \in X$ with $x \rightarrow y$ satisfies $P$, it follows that $x$ satisfies $P$. Then every $x \in X$ satisfies $P$.

### 2.3 Monoid invariants

There are four operations that produce new monoid presentations from old ones called Tietze transformations. For any presentation $\mathcal{P} = \langle x : r \rangle$ these are defined as follows.

- **(T$_1$)** If $u$ and $v$ are words of $xF$ such that $u \leftrightarrow v$, then obtain a new presentation from $\mathcal{P}$ by adding a rewriting rule $u = v$ to $r$.

- **(T$_2$)** If $u = v$ is a rewriting rule of $r$, and $u \leftrightarrow v$ in the string rewriting system for $\mathcal{P}_1 = \langle x : r - \{u = v\} \rangle$, then obtain $\mathcal{P}_1$ from $\mathcal{P}$ by deleting a rewriting rule from $r$.

- **(T$_3$)** If $a$ is a letter not in $x$, but $w$ is a word of $xF$, then obtain a new presentation from $\mathcal{P}$ by adding a generator $a$ to $x$ and a new rewriting rule $a = w$ or $w = a$ to $r$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
• \( (T_4) \) If \( a \) is a letter of \( x \) and \( w \) is a word of the free monoid on \( x \), with \( a = w \) or \( w = a \) a rewriting rule of \( r \), then obtain a new presentation \( \langle x_1 : r_1 \rangle \) from \( P \) by deleting a generator from \( x \). The rewriting rules of \( r_1 \) are obtained from \( r \) by deleting the rule \( a = w \) or \( w = a \), and by replacing each occurrence of the letter \( a \) in the remaining rules with the word \( w \).

This brings us to the statement of a well known result analogous to one from group theory. A proof is not given, but one can refer to [15].

**Theorem 2.3.1.** Let \( P \) and \( P_1 \) be finite monoid presentations. The monoids \( P_M \) and \( P_1M \) are isomorphic if and only if there is a sequence

\[
P = Q_1, Q_2, \ldots, Q_n = P_1
\]

of presentations such that \( Q_{i+1} \) is obtained from \( Q_i \) by a \( T_1 \)-, \( T_2 \)-, \( T_3 \)- or \( T_4 \)-Tietze transformation for each \( i < n \).

The monoids involved are said to be Tietze equivalent, and the content of the theorem justifies the following definition: a property of finitely presented monoids is invariant when it is preserved under Tietze equivalence.

Some popular examples of Tietze equivalent monoids are identified by the following theorem. For proof consult [14].

**Theorem 2.3.2.** There is an effective procedure that determines for each finite complete monoid presentation \( P = \langle x : r \rangle \) a finite canonical monoid presentation \( P_1 = \langle x : r_1 \rangle \) such that \( P_M \) is Tietze equivalent (isomorphic) to \( P_1M \).
Chapter 3

Monoid pictures

This section begins with a new definition of monoid picture. Existing ones are due to Pride [10] [11] [18] and Guba and Sapir [4]. The former is not used here because it is informal in the sense that it is to be inferred from examples, but in Pride’s defence this suits his needs as he develops monoid theory by combining techniques of combinatorial group theory with the illustrative power of his pictures. However, we require a formal definition of monoid picture here as we aim to reason (as much as possible) exclusively with and about pictures (note that we fail to do so in, for example, lemmas 3.4.3 and 3.4.5).

Guba and Sapir’s definition is formal but it is difficult for us to use as it relies on the additional notions of semigroup diagram and plane isotopy. Our version favours more elementary (read: linear algebraic) concepts.

Our model of monoid picture i.e. the way in which we draw the pictures, is a hybrid of Pride’s and Guba and Sapir’s. Like the definition we have tried to make the realisation as simple as possible. To do so the features of Pride’s pictures enabling a Pride-style discussion of the Squier complex and its homotopy are
retained while the rough form of our pictures is reminiscent of Guba and Sapir's minimalis wire/transistor diagrams.

After the definition we talk about operators on monoid pictures. The sum and product operators and the actions of free monoids on pictures are all as conceived by Pride.

Next the Squier complex (a particular 2-complex with 1-skeleton the monoid pictures over a monoid presentation) is constructed and its homotopy is defined. We prefer Pride's notion of the complex to Guba and Sapir's as the homotopy in his version arises purely from picture operations. Recall however that we use Guba and Sapir's definition of 2-complex. This means that Pride's Squier complex has been gently massaged to make itself amenable to that definition.

We conclude the section with a definition of finite derivation type (FDT) and proofs of the major theorems: FDT is an invariant property of monoids; every monoid with a finite complete presentation has FDT. FDT was discovered by both Pride and SOK, but we choose Pride's account because it is simple and very neat, and because it is stated in terms of monoid pictures (SOK do not use diagrammatic aids. They work directly with the combinatoric properties of string rewriting systems). While our proofs are essentially new, many of them are structured in the same way as the SOK proofs of [15]. In particular, lemma 3.4.4 is nothing but a translation of a SOK combinatorial proof into a picture proof. We consider the SOK proofs before Pride's because the latter rely on facts about monoid pictures derived from combinatorial group theory. This approach
makes for elegant proof (far more elegant than ours), but it is not conducive to a self contained treatment of monoid pictures and FDT. The work of SOK by contrast is directly applicable according to theorem 3.2.3.

In general we have supplied to the theory some missing pieces, and we have tried to consolidate the results of the individuals cited above in a way that (hopefully) respects all of their excellent work.

### 3.1 Definition

An *atomic (monoid) prepicture* (or *AMPP*) in the plane consists of a dash box and a solid rectangle both with sides parallel to the coordinate axes, and labeled lines parallel to the vertical axis. Sides of the box and the rectangle are defined to be *top*, *bottom*, *left* or *right* sides in the obvious way. For example, if an AMPP resides in the right upper half plane, then call the side of its box parallel to and furthest from the horizontal axis the top of the box, and call the opposite side the bottom of the box. Meanwhile the side of its box parallel to and furthest from the vertical axis is the right side of the box while the opposite side is the left side of the box. Similarly for the rectangle.

The box, rectangle and lines of any AMPP must satisfy the following conditions.

1. The rectangle is contained completely within the interior of the box.
2. The lines are contained completely within the box.
3. Lines connect either a single point at the top of the box to a single point at the bottom of the box, or a single point at the top of the box to a single
point at the top of the rectangle, or a single point at the bottom of the box to a single point at the bottom of the rectangle.

4. Lines do not intersect one another.

5. Any line meets the rectangle in at most one point.

Suppose that \( p \) is an AMPP, and that we read from left to right the labels of the lines connected to the top of the box of \( p \). This sequence is called the top or initial label of \( p \) and it is denoted \( p_t \). Similarly we obtain the bottom or terminal label \( p_r \) of \( p \) by reading from left to right the labels of the lines connected to the bottom of the box of \( p \).

Example 3.1.1. Figure 3.1.1 shows an atomic monoid prepicture \( p \). Its top and bottom labels are \( p_t = abbbbaa \) and \( p_r = abcba \) respectively.

Let us place the line labels of an AMPP \( p \) in four sequences. The first, \( p_s_1 \), contains the labels of all lines that lie to the left of the rectangle of \( p \). The second, \( p_s_2 \), contains the labels of all lines connected to the top of the rectangle
of $p$ while the third, $p s_3$, contains the labels of all lines connected to the bottom of the rectangle of $p$. The final sequence $p s_4$ contains the labels of all lines that lie to the right of the rectangle of $p$. Labels in each of the four sequences are indexed according to the order in which they appear in $p$ when read from left to right. Establish a convention by assuming that the indices are zero-based and that each nonzero index is the natural successor of the previous index.

We can use these sequences to partition the class $T$ of all AMPPs. For any $p$ and $q$ in $T$, identify $p$ and $q$ (write $p \simeq q$) if and only if

1. $p u = q t$ and $p r = q r$, and

2. $p s_i = q s_i$ for each $i = 1, 2, 3, 4$.

It is easy to check that $\simeq$ is indeed an equivalence relation on $T$, and we call each equivalence class of $\simeq$ an atomic (monoid) picture. If $P$ is an atomic monoid picture, then define its initial and terminal labels $P_i$ and $P_r$ respectively by $P_i = p i$ and $P_r = p r$ with $p$ any representative of $P$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The given partition of $T$ indicates that dilations and linear translations play no part in the definition of atomic monoid picture. This corresponds to our intuition that two atomic prepictures such as those featured in figure 3.1.2 are essentially the same.

A (monoid) picture is a finite sequence $P = P_1P_2\ldots P_n$ of atomic monoid pictures with $P_{i+1} = P_iP_i$ for $1 \leq i < n$. To draw such a picture $P$ in the plane we choose representatives for $P_1, P_2, \ldots, P_n$ such that

1. the horizontal lengths of their boxes are equal,
2. the top of the box of $P_{i+1}$ coincides with the bottom of the box of $P_i$,
3. the lines connected to the top of the box of $P_{i+1}$ meet the like labeled lines connected to the bottom of the box of $P_i$. 

Figure 3.1.2: equivalent atomic monoid prepictures.
CHAPTER 3. MONOID PICTURES

An example is given in figure 3.2.6.

From now on, unless required for the resolution of a specific problem, we make no further explicit reference to AMPPs, and we do not distinguish particular representatives of monoid pictures (atomic or otherwise). It is assumed that suitable representatives are at hand for given applications.

3.2 Monoid pictures and monoid presentations

In what follows let \( P = \langle x : r \rangle \) be a finite monoid presentation. Let \( A \) be the set of all 4-tuples \( (y, x^+ = x^-, \varepsilon, z) \) with \( y \) and \( z \) elements of the free monoid \( xF \), \( x^+ = x^- \) a rewriting rule of \( P \), and \( \varepsilon = \pm 1 \). A quadruple like this shows how the word \( yx^-z \) may be derived from \( yx^z \) in the string rewriting system associated with \( P \), and it can be related to an atomic monoid picture \( P \) with the following characteristics: for each letter \( y \) of \( y \), there is a line to the left of the rectangle of \( P \) with label \( y \); similarly, for each letter \( z \) of \( z \), there is a line to the right of the rectangle of \( P \) with label \( z \); every letter \( x^\varepsilon \) of \( x^\varepsilon \) is represented by a line with label \( x^\varepsilon \) connecting the top of the box of \( P \) to the top of its rectangle, and every letter \( x^-\varepsilon \) of \( x^-\varepsilon \) is represented by a line with label \( x^-\varepsilon \) connecting the bottom of the box of \( P \) to the bottom of its rectangle; all lines of \( P \) are ordered from left to right so that the top and bottom labels of the atomic picture are \( P_t = yx^\varepsilon z \) and \( P_r = yx^-\varepsilon z \) respectively. See figure 3.2.1.
It is easy to show that the association between 4-tuples of $A$ and atomic monoid pictures like those described above is in fact a one-to-one map. Hence pictures in the codomain of the map are called \textit{atomic (monoid) pictures (over $P$)}.

This gives rise to a graph $P\Gamma$ for $P$ called the \textit{graph of derivations} [9] with vertices the words of $xF$ and edges atomic pictures over $P$. An atomic picture $P = (y, x^{+1} = x^{-1}, \varepsilon, z)$ has initial and terminal vertices $P_l = yx^\varepsilon z$ and $P_T = yx^{-\varepsilon} z$ respectively, and $P^{-1}$ is its mirror image $(y, x^{+1} = x^{-1}, -\varepsilon, z)$. An orientation on $\Gamma$ can be obtained by taking as positive edges all atomic pictures of the form $(y, x^{+1} = x^{-1}, +1, z)$. Some examples further describe the situation.

\textbf{Example 3.2.1.} Consider the monoid presentation $\langle a, b : b^2a = a, a^3 = 1 \rangle$. Figure 3.2.2 is the positive atomic picture $(1, b^2a = a, +1, 1)$ corresponding to the rewriting of $b^2a$ to $a$. Figure 3.2.3 is the atomic picture...
\[ P = (ab, b^2a = a, -1, bab), \]

i.e. a display of \( ababab = abbabab \). Figure 3.2.4 is its inverse

\[ P^{-1} = (ab, b^2a = a, +1, bab), \]

and figure 3.2.5 is a positive atomic picture \((a, a^3 = 1, +1, 1)\) corresponding to the rewriting of \(aaaa\) to \(a\).

Figure 3.2.2: a positive atomic picture.
CHAPTER 3. MONOID PICTURES

Figure 3.2.3: an atomic picture.

Figure 3.2.4: inverse of the atomic picture displayed in figure 3.2.3.
Note that the initial vertex of an edge (= atomic picture) $P$ of $\mathcal{P} \Gamma$ coincides with the top label of $P$. Hence the notation $P_t$ for both is unambiguous (and intentional). The same can be said for the terminal vertex and bottom label $P_T$ of $P$. It follows immediately that the paths of $\mathcal{P} \Gamma$ are monoid pictures, and when the distinction is required these are called (monoid) pictures over $\mathcal{P}$. As with any graph, functions $t$, $r$ and inverse are extended to paths. In particular, a path in the graph of derivations with identical initial and terminal vertices is called a spherical (monoid) picture. If the rectangles of two adjacent atomic pictures in a path are not connected by any lines, then the pictures are said to be disjoint. Otherwise they overlap. The number of rectangles, hence the number of atomic pictures in a picture $P$ over $\mathcal{P}$ is defined to be the area of $P$.

**Example 3.2.2.** Let $\mathcal{P}$ be the monoid presentation of example 3.2.1. Figure
3.2.6 is a monoid picture $P$ with initial vertex $a b^2 a^3 = P_i$ and terminal vertex $b^4 a = P_T$ while figure 3.2.7 is its inverse $P^{-1}$. Figure 3.2.8 displays a spherical picture in which the last two atomic pictures are disjoint while the first two overlap.

Figure 3.2.6: a monoid picture.
Figure 3.2.7: inverse of the monoid picture displayed in figure 3.2.6.

Figure 3.2.8: a spherical picture.
Suppose that $w$ is a word of the monoid with presentation $\mathcal{P}$. The empty path $1_w$ is given by an atomic monoid picture containing no rectangle with lines labeled with the letters of $w$. For example, continuing with the presentation of 3.2.1, the empty path at $a^2ba$ is shown below.

![Figure 3.2.9: an empty path.](image)

By convention, when $1_w$ is concatenated with a picture we may erase the dashed line separating their boxes. This is demonstrated in figure 3.2.10. We may also draw the same dashed line into the picture to make explicit the presence of the empty picture.
The following may be clear from the definition of the graph $\mathcal{P}$, but we make
the statement for future reference.

**Theorem 3.2.3.** There is a mapping of graphs $f$ from the graph of derivations
$\mathcal{P}$ onto the string rewriting system $S$ for $\mathcal{P}$ that is the identity on vertices.

**Proof.** Let $f : \mathcal{P} \rightarrow S$ be the identity on vertices of $S$. For each atomic monoid
picture $P = (y, x+1 = x^{-1}, \varepsilon, z)$ over $\mathcal{P}$, i.e. for each edge $P$ of $\mathcal{P}$, there is an
edge $Pf = (yx+1z \rightarrow yx^{-1}z)^f$ of $S$. The map $f$ defined in this way is onto.

In general the function $f$ described above is not an isomorphism. For example,
there are two monoid pictures over $\mathcal{P} = \langle a : a^2 = 1 \rangle$ with initial vertex $a^4$ and
terminal vertex $a^2$. See figure 3.2.11. What we do see however is that $f$ identifies

![Figure 3.2.10: concatenation with the empty picture.](image)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
all pictures \( P \) and \( Q \) with \( P_i = Q_i \) and \( P_T = Q_T \), and that the string rewriting system \( S \) is the result of this identification. This implies that the connected components of the graph of derivations and the string rewriting system for a presentation have the same vertices.

There is a left action of the free monoid \( x F \) on atomic pictures of \( P \Gamma \): if \( w \) is a word of \( x F \) and \( P = (y, x^+ = x^{-1}, \varepsilon = \pm 1, z) \) is an atomic picture from \( P \Gamma \), then \( w.P \) is obtained from \( P \) by drawing within the dashbox a sequence of lines labeled with the letters of \( w \) to the left of all lines of \( P \). That is, \( w.P = (wy, x^+ = x^{-1}, \varepsilon, z) \). A right action of \( x F \) on edges of \( P \Gamma \) is similarly defined so that \( P.w = (y, x^+ = x^{-1}, \varepsilon, zw) \). Drawing the relevant pictures it is clear that if \( u \) is another word of \( x F \), then \((w.P).u = w.(P.u)\). Hence the expression \( w.P.u \) is unambiguous.
These left and right actions on atomic pictures extend themselves to actions on pictures. If \( w \) and \( u \) are elements of the free monoid, and if \( \mathcal{P} = P_1 P_2 \ldots P_n \) is a picture with \( P_1, P_2, \ldots, P_n \) atomic, then \( w.\mathcal{P}.u \) is the picture

\[
    w.\mathcal{P}.u = (w.P_1.u)(w.P_2.u) \cdots (w.P_n.u).
\]

The summary in an algebraic language is that there is a compatible biaction of \( xF \) on the paths of \( \mathcal{P} \Gamma \).

If \( \mathcal{P} \) and \( \mathcal{Q} \) are pictures, then define their sum \( \mathcal{P} + \mathcal{Q} \) to be the picture \( (\mathcal{P} . (Q\ell))((\mathcal{P} \ell).\mathcal{Q}) \). For \( n \) pictures \( \mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_n \), define inductively their sum \( \mathcal{P}_1 + \mathcal{P}_2 + \cdots + \mathcal{P}_n \) to be \( (\mathcal{P}_1 + \cdots + \mathcal{P}_{n-1}) + \mathcal{P}_n \).
For each pair of atomic pictures $P$ and $Q$ of $\mathcal{P}\Gamma$, let $[P, Q]$ be the spherical picture $(P + Q)(P^{-1} + Q^{-1})$. A diagram is shown in figure 3.2.14. If we let $s$ be the set of all such pictures $[P, Q]$ and their inverses, then define maps $\llbracket \rrbracket$ and $\llbracket \rrbracket$ from $s$ into nonempty paths of $\mathcal{P}\Gamma$ by putting

$$\llbracket [P, Q] \rrbracket = P + Q$$

$$\llbracket [P, Q] \rrbracket = (P^{-1} + Q^{-1})^{-1}$$

$$\llbracket [P, Q]^{-1} \rrbracket = (P^{-1} + Q^{-1})^{-1}$$

$$\llbracket [P, Q]^{-1} \rrbracket = (P + Q)^{-1}$$

Restricting the path version of $^{-1}$ to pictures of $s$, we claim that $\mathcal{PD} = (\mathcal{P}\Gamma, s, \llbracket, \rrbracket, \llbracket, \rrbracket, ^{-1})$ is a directed 2-complex. To see this, note first that for each
s = [P, Q] in s, its top and bottom paths ([s] and [s] respectively) both have initial vertex (P\textsuperscript{t})(Q\textsuperscript{t}) and terminal vertex (P\textsuperscript{r})(Q\textsuperscript{r}). Next one can verify that the maps |, ] and \(^{-1}\) are as they should be for a 2-complex. This exercise is mostly pedestrian: the requirement that \(s \neq s^{-1}\) is the only one that is not a direct consequence of the definitions, but it is certainly the case since the atomic pictures \((P\textsuperscript{r}).Q\) and \(P^{-1}.(Q\textsuperscript{r})\) are disjoint.

A similar proof holds for each \(s = [P, Q]^{-1}\), so the claim is established. The directed 2-complex \(PD\) is called the Squier complex of \(\mathcal{P}\) and \(s\) is its set of 2-cells. Impose an orientation on the complex by taking 2-cells of the form \([P, Q]\) to be positive.

![2-cell diagram](image)

**Figure 3.2.14:** the 2-cell \([P, Q]\).

There are well defined left and right actions of \(xF\) on 2-cells of the Squier complex \(PD\): if \(w\) is a word of the free monoid and if \([P, Q]\) is a 2-cell, then put
CHAPTER 3. MONOID PICTURES

\[ w.[P, Q] = [w.P, Q] \text{ and } [P, Q].w = [P, Q.w]. \]

3.3 Homotopy relation on the Squier complex

Suppose that \( \mathcal{P} = \langle \mathcal{X} \rangle \) is a monoid presentation and that \( \mathcal{S} \) is a set of spherical pictures over \( \mathcal{P} \). We define two types of operations on pictures over \( \mathcal{P} \):

- **Type I.** Insert or delete an inverse pair \( PP^{-1} \) from a picture (where \( P \) is atomic).

- **Type II.** Insert or delete paths \( u.P.v \) from a picture (where \( P \) is a picture from \( \mathcal{S} \), and \( u \) and \( v \) are from \( \mathcal{X.F} \)).

It is clear that the insertion operations can only be performed at appropriate vertices. That is, if \( Q \) is an inverse pair or a picture \( u.P.v \) with \( P \) in \( \mathcal{S} \) and \( u, v \) in \( \mathcal{X.F} \), and if \( Q \) has initial (hence terminal) vertex \( w \), then it can be inserted in a picture \( R \) only if \( R \) contains the vertex \( w \) and only at an occurrence of that vertex.

If it turns out that \( \mathcal{P} \Gamma \) and \( \mathcal{S} \) are part of some directed 2-complex \( \mathcal{K} = (\mathcal{P} \Gamma, \mathcal{S} \}, \{ \}, ^{-1}) \), then two pictures are said to be homotopic (in \( \mathcal{K} \)) if one can be obtained from the other by a finite sequence of operations of type I or II. The relation \( \simeq_{\mathcal{K}} \) (or just \( \simeq \) when it is unambiguous) consisting of homotopic pairs of pictures is called the homotopy relation (on \( \mathcal{K} \)). Use of this terminology is justified by the following two results.

**Proposition 3.3.1.** The homotopy relation is an equivalence relation.

**Proof.** Let \( P, Q, \) and \( R \) be pictures over \( \mathcal{P} \). \( P \) is homotopic to \( P \) by an empty (hence finite) sequence of operations, so the relation is reflexive. If \( P \simeq Q \) by
CHAPTER 3. MONOID PICTURES

a sequence \((o_1, o_2, \ldots, o_n)\) of picture operations, then \(Q \simeq P\) by the sequence
\((o_n^{-1}, \ldots, o_2^{-1}, o_1^{-1})\) of picture operations (here \(o_i^{-1}\) is an insertion operation if
\(o_i\) is a deletion operation and vice versa for each \(i \leq n\)). This shows that \(\simeq\) is
symmetric. Suppose that \(P \simeq Q\) and \(Q \simeq R\) by sequences \((o_1, o_2, \ldots, o_n)\) and
\((p_1, p_2, \ldots, p_m)\) of picture operations respectively. We get that \(P\) is homotopic
to \(R\) by the composite sequence \((o_1, o_2, \ldots, o_n, p_1, p_2, \ldots, p_m)\). Hence \(\simeq\) is
transitive and the relation is an equivalence.

\[\Box\]

An equivalence class of the homotopy relation is called a homotopy class.

**Proposition 3.3.2.** Fix a word \(w\) of \(xF\) and consider the spherical pictures with
initial and terminal vertex \(w\). The homotopy classes of all such pictures form the
underlying set of a group.

**Proof.** If \([P]\) and \([Q]\) are homotopy classes, then take their group product \([P] \cdot [Q]\)
to be \([PQ]\). This multiplication is everywhere defined since every picture is
spherical at the same vertex \(w\), and well defined since if \(P'\) is homotopic to \(P\)
by a sequence of picture operations, then \(P'Q\) is homotopic to \(PQ\) by the same
sequence. If \([R]\) is a third homotopy class, then because picture concatenation
is associative we get that

\[
([P] \cdot [Q]) \cdot [R] = [PQ] \cdot [R] = ([PQ]R) = [P(QR)] = [P] \cdot [QR] = [P] \cdot ([Q] \cdot [R]).
\]

It follows that \(\cdot\) is associative. Take \([1_w]\) to be the group unit since \(1_wP = P = P1_w\).
Finally we see that \(PP^{-1}\) and \(P^{-1}P\) both collapse to \(1_w\) under \(\simeq\), so take
\([P]^{-1}\) to be \([P^{-1}]\).
The group is called the *fundamental group with base point* \( w \) and is denoted \((\mathcal{K}, w)\pi_1\).

Tying some of the preceding ideas together, the next theorem shows us how some spherical pictures can be broken into two pieces that might informally be described as top and bottom paths.

**Proposition 3.3.3.** Let \( w \) be a word of \( xF \), let \( P \) and \( Q \) be pictures with \( P_t = Q_t = w \), and suppose that \( PQ \) is defined. Then \( PQ \simeq 1_w \) if and only if \( P \simeq Q^{-1} \).

**Proof.** The picture \( PQQ^{-1} \) is certainly homotopic to \( P \), but given \( PQ \simeq 1_w \) we see that it is also homotopic to \( Q^{-1} \). Transitivity of the homotopy relation then yields \( P \simeq Q^{-1} \). For the converse, if \( P \simeq Q^{-1} \), then \( PQ \simeq Q^{-1}Q \simeq 1_w \), and the result follows again from the transitivity of \( \simeq \).

There is a nice consequence to certain applications of type I and II operations to pictures of the Squier complex. Suppose that \( PQ \) is part of a picture with \( P \) and \( Q \) atomic, and suppose that these two pictures are disjoint. We can define an operation where we push the rectangle of \( P \) down while pulling the rectangle of \( Q \) up (figure 3.3.1).
In one case there are atomic pictures $A$ and $B$ such that $P = A \cdot (B$ and $Q = (A \tau) \cdot B$. In other words, the rectangle of $P$ is to the left of the rectangle of $Q$. Here we perform a type II operation and insert the 2-cell $[A, B]^{-1}$ below $Q$. The picture is now

$$PQ[A, B]^{-1} = (A \cdot (B \iota))((A \tau) \cdot B)((A \tau) \cdot B)^{-1}(A \cdot (B \iota))^{-1}((A \iota) \cdot B)\cdot (A \cdot (B \tau)).$$

To obtain the desired result, follow this with two type I operations: deletion of $((A \tau) \cdot B)((A \tau) \cdot B)^{-1}$ and deletion of $(A \cdot (B \iota))(A \cdot (B \iota))^{-1}$. $P$ and $Q$ are shown in figure 3.3.2 and the sequence of operations just described appears in figure 3.3.3.
CHAPTER 3. MONOID PICTURES

Figure 3.3.2: P and Q.

\[
\begin{align*}
A &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} & B &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\
\mathcal{P} &= A \cdot (B_i) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} & \mathcal{Q} &= (A \cdot c) \cdot B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}
\end{align*}
\]

Figure 3.3.3: interchanging disjoint rectangles. Step 1: insert the 2-cell \([A, B]^{-1}\). Step 2: delete an inverse pair. Step 3: delete another inverse pair.

In the other case the rectangle of \(\mathcal{P}\) is to the right of the rectangle of \(\mathcal{Q}\), so there are atomic pictures \(A\) and \(B\) such that \(\mathcal{P} = (A_i) \cdot B\) and \(\mathcal{Q} = A \cdot (B \tau)\). By turning
Thus following from previously defined picture operations, we identify this new operation on pictures of the Squier complex:

- **Type III.** Replace a picture of the form \((A.i.B)(A.r.B)\) with the picture \(((A.t).B)(A.(B.t))\) or vice versa (where \(A\) and \(B\) are atomic).

Conversely, in the Squier complex the type II operation can be derived from type I and type III operations. If \(P\) and \(Q\) are pictures of the complex, we can delete the 2-cell \([P, Q]\) from a picture by first pulling the rectangle due to \(P^{-1}\) in \((P^{-1} + Q^{-1})\) up through \((P + Q)\) while pushing the rectangle due to \(Q\) in \((P + Q)\) down through \((P^{-1} + Q^{-1})\) so that \([P, Q]\) is transformed into \((P.(Q.t))(P.(Q.i))^{-1}(P.r.Q)((P.r).Q)^{-1}\). We then complete the manoeuvre with two type I operations. Reversing these steps inserts the 2-cell \([P, Q]\). Deleting or inserting the 2-cell \([P, Q]^{-1}\) is similar.

### 3.4 Monoids of finite derivation type

Here we introduce our monoid invariant, and supply proof of its invariance. The property is a finiteness condition on the homotopy of the Squier complex.

Let \(P = \langle x : r \rangle\) be a monoid presentation and let \(PD = (P\Gamma, s, [], [], ^{-1})\) be its Squier complex. We describe a final operation on pictures over \(P\) given any set \(t\) of spherical pictures over \(P\):

- **Type IV.** Insert or delete paths \(w.T.u\) from a picture (where \(T\) or \(T^{-1}\) is a picture from \(t\), and \(w\) and \(u\) are from \(xF\)).
Hence define a new homotopy relation $\simeq_t$ by saying that two pictures are homotopic in $\mathcal{P}D$ if one can be obtained from the other by a finite sequence of operations of type I, II, III, or IV. This relation is indeed a homotopy in the sense of proposition 3.3.1 and proposition 3.3.2 - to show it one can even supply the same proofs. We also retain the result of proposition 3.3.3. When $t$ is empty or $t$ is a subset of $s$ it is clear that $\simeq_t$ coincides with $\simeq$, so in the presence of the Squier complex and a given $t$ we often write $\simeq$ for both.

This $t$ is of particular interest when it is a trivialiser of the Squier complex $\mathcal{P}D$: say that $t$ is such a set when every spherical picture over $\mathcal{P}$ is homotopic to the empty picture, i.e. when the Squier complex has trivial fundamental groups. We say that a finite monoid presentation $\mathcal{P}$ is of finite derivation type (FDT) when the Squier complex $\mathcal{P}D$ has a finite trivialiser.

**Lemma 3.4.1.** If $f : \mathcal{P}_1 \to \mathcal{P}$ is a retraction, and if $\mathcal{P}_1$ has FDT, then so does $\mathcal{P}$.

**Proof.** Let $\mathcal{P}_1 = \langle x_1 : r_1 \rangle$ and let $t_1$ be a finite trivialiser of $\mathcal{P}_1 D$. The map $f$ is a retraction, so for each rewriting rule $x^{+1} = x^{-1}$ of $\mathcal{P}_1$ we have a path $p$ with initial and terminal vertices $pv = x^{+1}f$ and $pt = x^{-1}f$ respectively in the string rewriting system for $\mathcal{P}$. Hence by theorem 3.2.3 there is a monoid picture over $\mathcal{P}$ with the same intial and terminal vertices. Choose such a picture for each rewriting rule $x^{+1} = x^{-1}$ of $\mathcal{P}_1$ and denote it $(x^{+1} = x^{-1})f$ so that $f$ is extended to a mapping of graphs $\mathcal{P}_1 \Gamma \to \mathcal{P} \Gamma$. Explicitly: on vertices (words of $x_1 \Gamma$) $f$ is already defined; on edges (atomic pictures) $\mathcal{P} = \langle y, x^{+1} = x^{-1}, \varepsilon = \pm 1, z \rangle$ of $\mathcal{P}_1 \Gamma$ we set $\mathcal{P}f = yf.(x^{+1} = x^{-1})f^\varepsilon zf$. Thus assert that the image $t$ of $t_1$ under $f$ is a trivialiser for $\mathcal{P}D$. To see this, consider a spherical picture $\mathcal{P}$ over $\mathcal{P}$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
This can be reduced to the empty picture by adding and removing inverse pairs, 2-cells and pictures of $t_1$ over $P_1$. As $f$ fixes monoid pictures over $P$ it fixes $P$, and verify that the picture can be reduced to the empty picture by adding and removing the images of the same inverse pairs, 2-cells and elements of $t_1$ under $f$.

Lemma 3.4.2. Finite derivation type is preserved by $T_1$-Tietze transformations.

Proof. Suppose that $P = \langle x : r \rangle$ is a monoid presentation, $P$ has FDT, and that $t$ is a finite trivialiser of the Squier complex $PD$. Let $P_1 = \langle x : r, u = v \rangle$ be obtained from $P$ by a $T_1$-Tietze transformation. Identify $B$ as a picture of $P\Gamma$ with $B_t = u$ and $B_r = v$, and $E$ as the atomic picture $(1, u = v, +1, 1)$ in $P_1\Gamma$. We can then define a mapping of graphs $f : P_1\Gamma \rightarrow P\Gamma$ that is the identity on $P\Gamma$, and that maps atomic pictures $x.E.y$ to pictures $x.B.y$. Assert that $t_1 = t \cup \{EB^{-1}\}$ is a (finite) trivialiser of $P_1D$, and verify by checking the following claims.

Claim. If $P$ is a picture of $P_1\Gamma$, then $P \simeq Pf$ in $P_1D$. This is obvious if $P$ is also a picture of $P\Gamma$, so suppose that it is not. Then $P$ contains edges of the form $x.E^\varepsilon.y$ with $\varepsilon = \pm 1$, and we replace each with a picture $(x.E^\varepsilon.y)(x.B^{-\varepsilon}.y)(x.B^\varepsilon.y)$ via a sequence of type I picture operations. Deletion of $(x.E^\varepsilon.y)(x.B^{-\varepsilon}.y)$ by type IV picture operations gives $Pf$, so the claim holds.

Claim. If $P$ is a spherical picture over $P_1$ with $P_t = P_r = w$, then $P$ is homotopic to the empty picture $1_w$ in $P_1D$. Indeed, the first claim helps us see that $P \simeq t_1 Pf \simeq t_1 t_1 1_w$. 

□
Lemma 3.4.3. Finite derivation type is preserved by $T_2$-Tietze transformations.

Proof. Suppose that a monoid presentation $\mathcal{P} = \langle x : r \rangle$ is obtained from another presentation $\mathcal{P}_1 = \langle x : r_1 \rangle$ by a $T_2$-Tietze transformation so that $r = r_1 - \{u = v\}$ for some rewriting rule $u = v$ of $r_1$, and $u \mapsto v$ in the string rewriting system for $\mathcal{P}$. We establish that $f : \mathcal{P}_1 \to \mathcal{P}$ is a retraction when $f$ is the identity on $xF$: first it is obvious that $f$ is a homomorphism, $x$ is a subset of $x$, and $r$ is a subset of $r_1$; second we have $x^+f = x^{-1}f$ for each rewriting rule $x^+ = x^{-1}$ of $\mathcal{P}_1$ - in particular $uf = vf$ in $\mathcal{P}M$ since $uf = u$, $vf = v$, and $u \mapsto v$ in the string rewriting system for $\mathcal{P}$. Hence the result follows from lemma 3.4.1 when $\mathcal{P}_1$ has FDT.

\[ \square \]

Lemma 3.4.4. Finite derivation type is preserved by $T_3$-Tietze transformations.

Proof. Suppose that $\mathcal{P} = \langle x : r \rangle$ is a monoid presentation, $\mathcal{P}$ has FDT, and that $t$ is a finite trivialiser of the Squier complex $\mathcal{P}D$. Let $\mathcal{P}_1 = \langle x_1 : r_1 \rangle$ where $x_1 = x \cup \{a\}$ and $r_1 = r \cup \{a = w\}$ be obtained from $\mathcal{P}$ by a $T_3$-Tietze transformation. In what follows we show that $t$ is also a trivialiser of $\mathcal{P}_1D$.

There is a mapping of presentations $f : \mathcal{P}_1 \to \mathcal{P}$ that is the identity on $x$, and that maps $a$ to $w$. This is extended to a mapping of graphs $f : \mathcal{P}_1\Gamma \to \mathcal{P}\Gamma$ that is again the identity on pictures over $\mathcal{P}$, and that maps atomic pictures of the form $(u, a = w, \pm 1, v)$ to the empty picture $1_{uwv}$. We now establish some intermediate results. Throughout let $\Delta\Gamma$ be the subgraph of $\mathcal{P}_1\Gamma$ containing only edges (atomic pictures) of the form $(u, a = w, \pm 1, v)$ with $u$ and $v$ words of $x_1F$, and let $\simeq_\epsilon$ be the homotopy relation $\simeq_\epsilon$.

First claim. If $u$ is a word of $x_1F$, then there is a positive picture of $\Delta\Gamma$
with initial vertex \( u \) and terminal vertex \( uf \). Moreover, any two monoid pictures like this are homotopic.

**Proof.** If \( u = \left( \prod_{i=1}^{n} u_i a \right) v \) with \( u_i \) and \( v \) words of \( xF \) for \( i = 1, \ldots, n \), and \( A \) is the atomic picture \( A = (1, a = w, +1, 1) \), then the picture

\[
P = \left( u_1 A \left( \prod_{i=2}^{n} u_i a \right) v \right) \left( u_1 w u_2 A \left( \prod_{i=3}^{n} u_i a \right) v \right) \cdots \left( \prod_{i=1}^{n-1} u_i a \right) u_n A_v
\]

is appropriate. A positive picture of \( \Delta \) has initial vertex \( u \) and terminal vertex \( uf \) if and only if it can be obtained from \( P \) via a finite sequence of type III or type I picture operations.

**Second claim.** Let \( P \) be a picture of \( \Delta \). There are positive pictures \( P_1 \) and \( P_2 \) of \( \Delta \) with \( P_1 = P_1 \cdot P_2 \), \( P_2 = P_2 \cdot \tau \), and \( P_1 \cdot \tau = P_2 \cdot \tau \) such that \( P \cong P_1 P_2 \).

**Proof.** Suppose that \( P = A_1 A_2 \cdots A_n \) with each \( A_i \) atomic for \( i = 1, \ldots, n \). If \( P \) is not already as described, then there is a least \( i \leq n \) such that \( A_i \) is not positive while \( A_{i+1} \) is. It must be that either \( A_{i+1} = A_i^{-1} \) or that \( A_i \) and \( A_{i+1} \) are disjoint. In the case the former we remove the inverse pair with a type I picture operation to obtain \( Q = A_1 \cdots A_{i-1} A_i A_{i+1} A_{i+2} \cdots A_n \). Otherwise a type III picture operation exchanges the rectangles of \( A_i \) and \( A_{i+1} \) to produce a positive atomic picture \( B \) and a non positive atomic picture \( C \), and set \( Q = A_1 \cdots A_{i-1} B C A_{i+2} \cdots A_n \). Repeat the procedure above on \( Q \) until the resulting picture is of the required form.

**Third claim.** Let \( A = (1, x^+ = x^{-1}, \varepsilon = \pm 1, 1) \) be an atomic picture over \( \mathcal{F} \), and take words \( u \) and \( v \) of \( x_1 F \). There are positive pictures \( P_1 \) and \( P_2 \) in \( \Delta \) with \( P_1 = (u A_v) \), \( P_2 = (u A_v) \), \( P_1 \tau = (u A_v) \), \( P_2 \tau = (u A_v) \), \( P_1 uf = A_v A_v P_1^{-1} \).

**Proof.** If \( u \) and \( v \) are from \( xF \), then the claim is trivially true. Otherwise
the first claim tells us that there is a positive picture \( \mathbb{P} \) in \( \Delta \Gamma \) with \( \mathbb{P}_t = (u.A.v)_t \) and \( \mathbb{P}_r = (u.A.v)_r f \). Drawing \( \mathbb{P}^{-1}(u.A.v) \) we notice that no line connects a rectangle of \( \mathbb{P}^{-1} \) to the rectangle of \( A \). Hence with a sequence of type III picture operations we pull the rectangle of \( A \) up past all rectangles due to \( \mathbb{P}^{-1} \) to obtain a picture \( Q = Pu.f.A.vfP_2^{-1} \) with \( P_2 \) positive. Set \( P_1 = \mathbb{P} \). Combined with the observation that \( \mathbb{P}P^{-1} \simeq 1_{(u.A.v)} \), whence \( Q \simeq \mathbb{P}P^{-1}(u.A.v) \simeq u.A.v \) we acknowledge the claim.

**Fourth claim.** Let \( \mathbb{P} \) be a picture over \( \mathcal{P} \). There are positive pictures \( P_1 \) and \( P_2 \) of \( \Delta \Gamma \) and a picture \( Q \) over \( \mathcal{P} \) with \( P_1t = P_t, P_1r = Pu.f, Q_t = Pu.f, Q_r = Pt.f, P_2t = P_t, \) and \( P_2r = Pt.f \) such that \( \mathbb{P} \simeq P_1QP_2^{-1} \).

**Proof.** From the third claim we know that \( \mathbb{P} \simeq \mathbb{P}' = R_0Q_1R_1 \ldots Q_nR_n \) where \( R_i \) is a picture of \( \Delta \Gamma \) for each \( i = 0, \ldots, n \), and \( Q_i \) is a picture over \( \mathcal{P} \) for each \( i = 1, \ldots, n \). If \( n = 0 \), then the second claim gives \( \mathbb{P}' \simeq K_1K_2^{-1} \) for some positive pictures \( K_1 \) and \( K_2 \) of \( \Delta \Gamma \). If \( K_1r \) contains the letter \( a \), then by first claim there is a positive picture \( K_3 \) of \( \Delta \Gamma \) with \( K_3t = K_1t \) and \( K_3r = K_1r f \). Otherwise take \( K_3 \) to be empty. Thus \( \mathbb{P} \simeq \mathbb{P}' \simeq K_1K_31_{K_1rf}K_3^{-1}K_2^{-1} \) as required (we are setting \( P_1 = K_1K_3, P_2 = K_2K_3, \) and \( Q = 1_{K_1rf} \)). When \( n > 0 \), \( Q_0 = R_0 f \) is a word of \( aF \). From this and the fact that \( R_0t = P_t \) it follows that \( R_0r = P_t f \). Hence by second claim there is a positive picture \( P_1 \) of \( \Delta \Gamma \) with \( P_1t = P_t, P_1r = Pu.f, \) and \( P_1 \simeq R_0 \). Similarly we find a positive picture \( P_2 \) of \( \Delta \Gamma \) with \( P_2t = Pt, P_2r = Pt.f, \) and \( P_2^{-1} \simeq R_n \). Hence \( P \simeq P_1QP_1 \ldots Q_nP_2^{-1} \).

For each \( R_i \) there are positive pictures \( K_{i_1} \) and \( K_{i_2} \) of \( \Delta \Gamma \) with \( R_i \simeq K_{i_1}K_{i_2}^{-1} \) by second claim. Note however that \( K_{i_1}t = Q_i t \) and \( K_{i_2}t = Q_{i+1}t \) are words of \( aF \), thus do not contain the letter \( a \), so we are forced to conclude that \( K_{i_1} \) and \( K_{i_2} \) are both empty. Therefore, \( P \simeq P_1Q_1 \ldots Q_nP_2^{-1} \). The claim is then
completely verified by setting $Q = Q_1 \ldots Q_n$.

Continuing with the notation used in the statement of the fourth claim, if $P$ is spherical, then $Q_i = P_i f = P \tau f = Q \tau$, so $Q$ is homotopic to the empty picture since it is a spherical picture over $P$. Also, $P_1 i = P_2 i = P i$ and $P_1 \tau = P_2 \tau = P f$ so that $P_1 \sim P_2$ by first claim which amounts to $P_1 P_2^{-1} \sim 1_P$, by proposition 3.3.3.

\[ \square \]

**Lemma 3.4.5.** Finite derivation type is preserved by $T_4$-Tietze transformations.

**Proof.** Let $P_1 = \langle x : r \rangle$ be a monoid presentation such that $P_1$ is of finite derivation type, and let $P = \langle x : r \rangle$ be a presentation obtained from $P_1$ by deleting a generator $a$ from $P_1$. Suppose that $a = w$ is the rewriting rule deleted from $r$ in the process, so the relations in $r$ are those in $r_1 - \{a = w\}$ with all occurrences of $a$ replaced with the word $w$. Now, the monoid presentation $Q = \langle x : r, a = w \rangle$ can be obtained from $P_1$ by a finite sequence of $T_1$- and $T_2$-Tietze transformations in two steps.

**First step.** All rewriting rules of $r - r_1$ are of the form

\[ u \prod_{i=1}^{n} wx_i = v \prod_{j=1}^{m} wy_j \]  
(3.4.1)

where $u$, $v$, $x_i$, and $y_j$ are words of $xF$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. However, in $r_1$ we do have the rules $a = w$ and

\[ u \prod_{i=1}^{n} ax_i = v \prod_{j=1}^{m} ay_j \]

which yields
in the string rewriting system for \( \mathcal{P}_1 \). It follows that
\[
u \prod_{i=1}^{n} wx_i \rightarrow^* v \prod_{j=1}^{m} wy_j
\]
in the same string rewriting system. We can then add the rule (3.4.1) to \( r_1 \) via a \( T_1 \)-Tietze transformation. Doing this for each such rewriting rule we obtain a set \( \overline{r_1} = r \cup r_1 \).

Second step. All rewriting rules of \( \overline{r_1} - r \cup \{a = w\} \) are of the form
\[
u \prod_{i=1}^{n} ax_i = v \prod_{j=1}^{m} ay_j
\]
where \( u, v, x_i \) and \( y_j \) are words of \( xF \) for \( i = 1, \ldots, n \) and \( j = 1, \ldots, m \). To remove such a rule from \( \overline{r_1} \), note that \( a = w \) and
\[
u \prod_{i=1}^{n} wx_i = v \prod_{j=1}^{m} wy_j
\]
are both in \( \overline{r_1} \). Thus
\[
u \prod_{i=1}^{n} ax_i \rightarrow^* u \prod_{i=1}^{n} wx_i \rightarrow v \prod_{j=1}^{m} wy_j \leftarrow^* v \prod_{j=1}^{m} ay_j
\]
in the string rewriting system for \( \langle x_1 : \overline{r_1} \rangle \) whence
\[
u \prod_{i=1}^{n} ax_i \leftarrow v \prod_{j=1}^{m} ay_j
\]
in the same string rewriting system. The removal is then accomplished via a \( T_2 \)-Tietze transformation. Doing this for each such rewriting rule leaves the set \( r \cup \{a = w\} \).
By lemma 3.4.2 and lemma 3.4.3 we get that $Q$ is of finite derivation type. The presentation $\mathcal{P}$ is a retract of $Q$ under the map that sends $a$ to $w$ and is the identity on other letters of $x_1$, so we draw the desired conclusion from lemma 3.4.1.

Let $\mathcal{P}$ be a finite monoid presentation. Say that the monoid $\mathcal{P}M$ is of finite derivation type if $\mathcal{P}$ is of finite derivation type. At last we are in a position to give one of the main results.

**Theorem 3.4.6.** Finite derivation type is an invariant property of monoids.

**Proof.** Immediate by combination of lemma 3.4.2 through lemma 3.4.5 with theorem 2.3.1

Examples of FDT monoids are abundant after proof of theorem 3.4.9, but we must first know a little bit more about string rewriting systems. Let $\mathcal{P}$ be a monoid presentation, and suppose that $E_1$ and $E_2$ are positive atomic pictures of $\mathcal{P}I$ with $E_1^t = E_2^t$. The ordered pair $(E_1, E_2)$ is called a critical pair when the rectangles of $E_1^{-1}E_2$ overlap, and when every line in $E_1^{-1}E_2$ is connected to a rectangle.
Figure 3.4.1: a critical pair.

Figure 3.4.2: not a critical pair.
A critical pair \((E_1, E_2)\) is said to be resolved if there are positive paths \(R_1\) and \(R_2\) with \(R_1 \tau = E_1 \tau\), \(R_2 \tau = E_2 \tau\), and \(R_1 \tau = R_2 \tau\). The pair \((R_1, R_2)\) is then called a resolution of the critical pair, and the situation is depicted below.

![Figure 3.4.3: a resolved critical pair.](image)

As mentioned the crucial result is theorem 3.4.9, but it relies upon lemma 3.4.7 and lemma 3.4.8. For these statements, we let \(\mathcal{P} = \langle x : r \rangle\) be a finite complete presentation. Some more preparation is required: for each critical pair \((E_1, E_2)\) of \(\mathcal{P} \Gamma\) fix a resolution \((R_1, R_2)\), and let \(t\) be the set of all spherical pictures \(R_1^{-1} E_1^{-1} E_2 R_2\). Note guaranteed existence of at least one resolution for each critical pair because \(\mathcal{P}\) is confluent. We then consider the Squier complex \(\mathcal{P} \Delta\) with homotopy relation \(\simeq = \simeq_t\).

**Lemma 3.4.7.** Suppose that \(w\) is a word of \(xF\), \(z\) is irreducible in \(\mathcal{P}\), and that \(P_1\) and \(P_2\) are positive pictures of \(\mathcal{P} \Gamma\) with \(P_1 \tau = P_2 \tau = w\) and \(P_1 \tau = P_2 \tau = z\).
CHAPTER 3. MONOID PICTURES

Then $\mathbb{P}_1 \simeq \mathbb{P}_2$ (see figure 3.4.4).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure3_4_4}
\caption{pictures in the lemma.}
\end{figure}

Proof. (Please note that the structure and organisation of this proof owes much to the advice of Dr. B. Steinberg - one of the thesis supervisors). Fix an irreducible word $z$ of $xF$, and let $W$ be the set of all words $w$ in $xF$ such that there is some positive picture over $\mathcal{P}$ with initial vertex $w$ and terminal vertex $z$, i.e. the set of all words $w$ such that $w \leftrightarrow z$ (by completeness of $\mathcal{P}$). We show by Noetherian induction that for any $w$ in $W$, whenever there are positive pictures $\mathbb{P}_1$ and $\mathbb{P}_2$ over $\mathcal{P}$ with initial and terminal vertices $w$ and $z$ respectively, it must be that $\mathbb{P}_1 \simeq \mathbb{P}_2$. The relation to consider throughout the induction is the single step derivation on $xF$ restricted to $W$. That it is Noetherian follows from the fact that $\mathcal{P}$ is complete.

Hence suppose that $w$ in $W$ is such that for every $y$ in $W$ with $w \rightarrow y$, we
have the property that whenever $Q_1$ and $Q_2$ are positive pictures over $\mathcal{P}$ with initial and terminal vertices $y$ and $z$ respectively it follows that $Q_1 \simeq Q_2$. We must now show that the property also holds for $w$: whenever $P_1$ and $P_2$ are pictures with $P_1\ell = P_2\ell = w$ and $P_1\tau = P_2\tau = z$ it follows that $P_1 \simeq P_2$.

Indeed, for $i = 1, 2$ take $P_i = A_iQ_i$ in such a way that $A_i$ is positive atomic and $Q_i$ is positive. Suppose that $A_i\tau = Q_i\ell = w_i$, and claim that there are positive pictures $K_1, K_2$, with $K_i\ell = w_i$ and $K_i\tau = v$ for some word $v$ such that $A_1K_1 \simeq A_2K_2$. Three cases are examined to verify this.

The first case is described by $A_1 = A_2$. This means that $w_1 = w_2$, so we take $K_1$ and $K_2$ to be empty.

In the second case the rectangles of $A_1$ and $A_2$ in $A_1^{-1}A_2$ are disjoint. There is then a 2-cell $P = K_1^{-1}A_1^{-1}A_2K_2$ in $\mathcal{PD}$ with $K_1$, $K_2$ atomic, and $K_1\tau = K_2\tau = v$, say. To see this, refer to the picture in figure 3.4.5. In symbols, suppose without loss of generality that the rectangle of $A_2$ is to the left of the rectangle of $A_1$ in $A_1^{-1}A_2$. We can then write

$$A_1 = (px_2^1q, x_1^1 = x_1^{-1}, 1, r)$$

and

$$A_2 = (p, x_2^1 = x_2^{-1}, 1, qx_1^1r)$$

for some words $p, q, r$ of $XF$ and some rewriting rules $x_1^1 = x_1^{-1}$ and $x_2^1 = x_2^{-1}$ of $\mathcal{P}$. Hence take

$$K_1 = (p, x_2^1 = x_2^{-1}, 1, qx_1^{-1}r)$$

and
\[ K_2 = (px_2^{-1}q, x_1^{-1} = x_1^{-1}, 1, r) \]

so that \( v = px_2^{-1}qx_1^{-1}r \). The situation is similar when the rectangle of \( A_2 \) lies to the left of the rectangle of \( A_1 \) in \( A_1^{-1}A_2 \). Conclude from proposition 3.3.3 that \( A_1K_1 \simeq A_2K_2 \) since \( P \simeq 1_v \).

The last case has rectangles of \( A_1 \) and \( A_2 \) in \( A_1^{-1}A_2 \) overlapping, so for some words \( x \) and \( y \) there is a critical pair \((E_1, E_2)\) with \( A_i = x_i.E_i.y \) \((i = 1, 2)\). Hence \( K_1^{-1}E_1^{-1}E_2R_2 \) is an element of \( t \) for some fixed resolution \((R_1, R_2)\) of the critical pair. Put \( K_1 = x.R_i.y \). A type IV picture operation then gives \( K_1^{-1}A_1^{-1}A_2K_2 \simeq 1_v \) in \( PD \) for some word \( v \), whence \( A_1K_1 \simeq A_2K_2 \) by proposition 3.3.3.

We now have \( v \mapsto z \) as these words reside in the same connected component of \( P \Gamma \) hence in the same connected component of the relevant string rewriting system according to theorem 3.2.3. Thus by completeness of \( P \) there is a positive picture \( B \) with \( B_t = v \) and \( B_\tau = z \).

For each \( i = 1, 2 \) the summary so far is that (1) the pictures \( Q_i \) and \( K_iB \) are both positive with initial vertex \( w_i \) and terminal vertex \( z \), and (2) there is a single step derivation \( w \rightarrow w_i \). From (1) deduce that the \( w_i \) belong to \( W \), so (2) combined with the induction hypothesis yields \( Q_i \simeq K_iB \). We now follow the chain of homotopy equivalences

\[
P_1 = A_1Q_1 \simeq A_1K_1B \simeq A_2K_2B \simeq A_2Q_2 = P_2
\]

to the desired conclusion by Noetherian induction.

\[ \square \]
Lemma 3.4.8. Consider the picture of figure 3.4.6 with $z_1, z_2$ irreducible and $P_1, P_2$ positive. It is a fact that $z_1 = z_2$ and $P \cong P_1 P_2^{-1}$.
CHAPTER 3. MONOID PICTURES

Figure 3.4.6: picture for the new lemma.

Proof. Suppose that we take \( P \) as a concatenation of pictures \( P = Q_1Q_2 \) with \( Q_1\tau = Q_2\nu = w \) for some word \( w \). Then \( Q_1^{-1}P_1^{-1} \) and \( Q_2P_2 \) are paths in \( P\Gamma \) with initial vertex \( w \) and terminal vertex \( z_i \) \((i = 1, 2)\) respectively. The words \( z_i \) are irreducible, so completeness of \( P \) gives \( z_1 = z_2 \). We shall refer to this word as \( z \) for the rest of the proof.

For \( i = 1, 2 \) let \( P_i\nu = w_i \). We show that \( P \simeq P_1P_2^{-1} \) by induction on the area \( n \) of \( P \). Indeed, suppose that \( n = 0 \). Then \( w_1 = w_2 \) and lemma 3.4.7 yields \( P_1P_2^{-1} \simeq 1_z = P \) as required. In case \( n > 0 \), there is a picture \( Q \) of area \( n - 1 \) and an atomic picture \( A \) such that \( P = QA \). Set \( w = Q\tau = A\nu \) and note that \( Q_\ell = w_1, A_\tau = w_2 \). The presentation \( P \) is complete so we can produce a positive picture \( K \) with initial vertex \( w \) and terminal vertex \( z \), and all pictures now involved are shown in figure 3.4.7. By induction hypothesis we have \( Q \simeq P_1K^{-1} \). If \( A \) is positive, then \( K \) and \( AP_2 \) are both positive with initial and
terminal vertices $w$ and $z$ respectively. Hence by lemma 3.4.7 $K \simeq \Lambda P_2$ and

$$P_1 P_2^{-1} \simeq P_1 P_2^{-1} A^{-1} A \simeq P_1 K^{-1} A \simeq QA \simeq P$$

If $A$ is not positive, then $A^{-1} K$ and $P_2$ are both positive with initial and terminal vertices $w_2$ and $z$ respectively. Hence by lemma 3.4.7 $A^{-1} K \simeq P_2$, so

$$P_1 P_2^{-1} \simeq P_1 K^{-1} A \simeq QA \simeq P$$

\[\square\]

Figure 3.4.7: pictures for the proof of lemma 3.4.8. We have $Q \simeq P_1 K^{-1}$ by induction hypothesis (compare to figure 3.4.6).

**Theorem 3.4.9.** If a monoid $M$ has a finite complete presentation, then it is of finite derivation type.

*Proof.* We consider a finite complete presentation $\mathcal{P} = \langle x : r \rangle$ of $M$ and a set $t$ of spherical pictures as described in the paragraph preceding the statement of lemma 3.4.7. Claim that $t$ is a finite trivialiser of $\mathcal{P}D$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Assertion: $t$ is finite. The number of critical pairs of $\mathcal{P}T$ is bounded by the number of ways in which pairs of positive atomic pictures corresponding to the rewriting rules of $\mathcal{P}$ overlap, and since there are finitely many of these pictures this number is finite.

To show that $t$ is a trivialiser of $\mathcal{P}D$ we demonstrate that any spherical picture over $\mathcal{P}$ is homotopic to an empty picture using lemma 3.4.8. That result relied only on picture operations of type I through IV, so the approach is valid.

Hence let $\mathcal{P}$ be a spherical picture over $\mathcal{P}$ and let its initial and terminal vertex be $w_1$. Split the picture at some point so that it is $\mathcal{P} = \mathcal{P}_1\mathcal{P}_2^{-1}$ with $\mathcal{P}_1i = \mathcal{P}_2i = w_1$ and $\mathcal{P}_1\tau = \mathcal{P}_2\tau = w_2$ for some word $w_2$ as shown in figure 3.4.8. Completeness of $\mathcal{P}$ gives positive pictures $Q_1$ and $Q_2$ with $Q_1i = w_i$ and $Q_1\tau$ irreducible for $i = 1, 2$. By lemma 3.4.8 we get $Q_1\tau = Q_2\tau$ and $\mathcal{P}_1 \simeq Q_1Q_2^{-1} \simeq \mathcal{P}_2$ whence $\mathcal{P} \simeq \mathcal{P}_1\mathcal{P}_2^{-1} \simeq 1_{w_1}$ as required (see figure 3.4.9).

$\square$

Figure 3.4.8: splitting $\mathcal{P}$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
3.5 A result from homology

Of course one would like to know the status of the converse of theorem 3.4.9: does every monoid of finite derivation type have a finite complete presentation? The answer is negative, but I am unaware of any direct proof in the literature, i.e. exhibition of an appropriate monoid, and the existing indirect one [2] relies on constructions and theorems beyond the scope of this document. We do however present an outline of the reasoning involved.

**Theorem 3.5.1.** A monoid of finite derivation type may not have a finite complete presentation.

**Outline of proof.** Required is a definition of a homological finiteness condition on monoids called FP$_3$. A monoid $M$ is said to be FP$_3$ if there is a partial resolution

$$C_3 \xrightarrow{\delta_3} C_2 \xrightarrow{\delta_2} C_1 \xrightarrow{\delta_1} C_0 \xrightarrow{\varepsilon} \mathbb{Z} \rightarrow 0$$

Figure 3.4.9: applying lemma 3.4.8.
with each $C_i$ ($i = 0, 1, 2, 3$) a finite dimensional free left $\mathbb{Z}M$-module (where $\mathbb{Z}M$ is the integral monoid ring of $M$). Squier gives explicit partial resolutions satisfying the definition for monoids with finite complete presentations in [14]. The paper [2] shows that for groups the conditions FDT and $FP_3$ are equivalent, and confirms the existence of $FP_3$ groups that, considered as monoids, do not have finite complete presentations. The result follows.

Here we almost offer a direct proof of theorem 3.5.1 with leg work done by Lafont and Proute. We say almost because the requirement that a monoid presentation be finite in order to be classified as one of finite derivation type is dropped in the following. This is technically feasible because a version of theorem 2.3.1 holds for infinite presentations [11], but we certainly do not claim that our proof is intuitively satisfying.

Begin by defining four free (left) $\mathbb{Z}$-modules given a canonical monoid presentation $\mathcal{P} = \langle x : r \rangle$:

1. $\mathbb{Z}$ is a $\mathbb{Z}$-module generated by 1;

2. for each letter $a$ of $x$, $C_1$ is the $\mathbb{Z}$-module generated by elements of the form $[a]$;

3. for each rewriting rule $x^+ = x_1$ of $r$, $C_2$ is the $\mathbb{Z}$-module generated by elements of the form $[A]$ where $A$ is the atomic picture $(1, x^+ = x_1, +1, 1)$.

4. for each critical pair $(E_1, E_2)$ of $\mathcal{P}T$, fix a resolution $(R_1, R_2)$, and let $C_3$ be the $\mathbb{Z}$-module generated by elements of the form $[R_1^{-1}E_1^{-1}E_2R_2]$. This use of square brackets can be extended to words, positive atomic pictures, and positive pictures in the following way: if $w = x_1x_2 \ldots x_n$ is a word with $x_i$
CHAPTER 3. MONOID PICTURES

a letter of \( a \) for each \( i = 1, 2, \ldots, n \), then put \([w] = [x_1] + [x_2] + \cdots + [x_n]\); if \( P = u.A.v \) is a positive atomic picture with \( A = (1, x^{+1} = x^{-1}, +1, 1) \) for some rewriting rule \( x^{+1} = x^{-1} \) of \( r \), then put \([P] = [A]\); if \( P = P_1 P_2 \ldots P_n \) is a positive picture with \( P_i \) atomic for each \( i = 1, 2, \ldots, n \), then set \([P] = [P_1] + [P_2] + \cdots + [P_n]\).

Next we define a partial free resolution

\[ C_3 \xrightarrow{\delta_3} C_2 \xrightarrow{\delta_2} C_1 \xrightarrow{\delta_1} \mathbb{Z} \to 0 \]

by describing the behaviour of the boundary maps on the generators of the given \( \mathbb{Z} \)-modules:

\[ [a] \delta_1 = 0 \]

\[ [A] \delta_2 = [A\ell] - [A\tau] \]

\[ [R_1^{-1}E_1^{-1}E_2R_2] \delta_3 = [E_1] + [R_1] - [E_2] - [R_2] \]

It is clear that \( \delta_2 \delta_1 = 0 \). To see that \( \delta_3 \delta_2 = 0 \) it is sufficient to show by induction on the area of \( P \) that \([P] \delta_2 = [P\ell] - [P\tau] \) for any positive picture \( P \). Hence for \( P = R_1^{-1}E_1^{-1}E_2R_2 \) it follows that

\[ [P] \delta_3 \delta_2 = [E_1\ell] - [E_1\tau] + [R_1\ell] - [R_1\tau] - [E_2\ell] + [E_2\tau] - [R_2\ell] + [R_2\tau] \]

\[ = [E_1\ell] + 0 - [R_1\tau] - [E_2\ell] + 0 + [R_2\tau] \]

\[ = 0 \]

Now, as \( \delta_{i+1} \delta_i = 0 \), we have that \( \text{im} \delta_{i+1} \) is a subset of \( \ker \delta_2 \) for each \( i = 1, 2, \ldots, n \).
so the (left) $\mathbb{Z}$-modules $H_i = \ker \delta_i / \text{im} \delta_{i+1}$ are well defined. The module $H_i$ is called the $i$th homology module of the monoid $P_M$.

**Theorem 3.5.2.** The module $H_1$ is finitely generated if $P_M$ has a presentation with a finite set of generators. Similarly, $H_2$ is finitely generated if $P_M$ has a presentation with a finite set of rewriting rules.

The proof of this theorem as well as the definitions of $C_1$, $C_2$, $C_3$ and the associated boundary maps are offered by Lafont and Prouté in [7]. The only original contribution here is to state the definitions of the $\mathbb{Z}$-modules in terms of monoid pictures.

Consider the monoid presentation $P$ consisting of generators $a$, $b$, $c$, and for each natural number $n$ a rewriting rule $ac^n b = 1$. The graph of derivations $P \Gamma$ has no critical pair, and it is easy to see that $P$ is canonical. Hence $P$ gives rise to the partial resolution

$$C_3 = 0 \xrightarrow{\delta_3} C_2 \xrightarrow{\delta_2} C_1 \xrightarrow{\delta_1} \mathbb{Z} \rightarrow 0$$

with $\mathbb{Z}$-modules and boundary maps defined as before. The $\mathbb{Z}$-module $C_3$ has a finite basis while $C_2$ has an infinite basis, so $H_2 = \ker \delta_2$ is not finitely generated. In fact it is generated by elements of the form $[A_n] - n[A_1] + (n-1)[A_0]$ where $n > 1$ and $A_n$ is the atomic picture $(1, ac^n b = 1, +1, 1)$ for each natural number $n$. By theorem 3.5.2 the monoid $P_M$ has no presentation with a finite set of rewriting rules.

Recall lemma 3.4.7. That result holds for pictures over $P$ since, in the absence of critical pairs, we may disregard the third case in its proof, and because its proof

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
does not rely on a finite set of rewriting rules. For this last reason lemma 3.4.8 also holds for pictures over $P$. Hence reasoning like we did for the proof of theorem 3.4.9 we get that $PM$ is a monoid of finite derivation type with an empty trivialiser. Thus we have produced a FDT monoid that has no finite complete presentation and theorem 3.5.1 is proved.

Note that the monoid $PM$ has a solvable word problem. Indeed, Lafont and Prouté used it to answer a long open question (again in [7]): are there monoids with solvable word problem that do not have a finite complete presentation?
Chapter 4

References


[7] Yves Lafont and Alain Prouté, *Church-Rosser property and homology of

61

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


