Math. Struct. in Comp. Science (1997), *vol.* 7, *pp.* 207–240. Printed in the United Kingdom

String rewriting and homology of monoids

DANIEL E. COHEN

Queen Mary and Westfield College, Mile End Rd., London E1 4NS
Email: d.e.cohen@qmw.ac.uk

Received 10 March 1995; revised 1 February 1996

Results of Anick (1986), Squier (1987), Kobayashi (1990), Brown (1992b), and others, show that a monoid with a finite convergent rewriting sytem satisfies a homological condition known as FP_∞.

In this paper we give a simplified version of Brown's proof, which is conceptual, in contrast with the other proofs, which are computational.

We also collect together a large number of results and examples of monoids and groups that satisfy FP_∞ and others that do not. These may provide techniques for showing that various monoids do not have finite convergent rewriting systems, as well as explicit examples with which methods can be tested.

1. Introduction

Kapur and Narendran (1985) showed that the rewriting system $\{(aba, bab)\}$ has no equivalent finite convergent rewriting system. However, the monoid it presents can also be presented by the finite convergent rewriting system $\{(ab, c), (ca, bc), (bcb, cc), (ccb, acc)\}$ on three generators a, b, c.

It is therefore natural to ask what conditions a monoid must satisfy if it can be presented on some generating set by a finite convergent rewriting system.

An obvious necessary condition is that the monoid has a solvable word problem. Squier (1987) showed that such a monoid must satisfy the condition FP_3, which is well known in homology theory. It was later realised that the results of Anick (1986) contained, in a different language and with some conditions, which can be weakened, the stronger result that such monoids satisfy the homological condition FP_∞. These conditions, and the related conditions FP_n for all n, will be defined in Section 6.

Anick proved his result by showing that any convergent, but not necessarily finite, rewriting system gives rise to a resolution for the monoid. A proof given in Kobayashi (1990), which is very similar to Anick's, is rather easier to follow. Another proof was given by Groves (1991); see also Farkas (1992). A version of Squier's argument is also given in Lafont and Prouté (1991); their argument is closer to ours than the other cited proofs, and is a useful introduction to our method.

All these proofs are computational, and so provide no insight as to why they work. By contrast, a proof due to Brown (1992b) is conceptual, and helps to explain what is happening. Also, Brown's method enables us to compute certain important homomorphisms, called *boundary operators*. The other proofs in principle allow these boundary operators to be computed, but do so only by a complicated inductive definition, which is likely to be troublesome in practice.

Brown's proof, however, requires some detailed knowledge of topology and homology theory to understand it. It therefore seems worthwhile to give a proof of Anick's theorem using the essential ideas of Brown's method in a form that does not require detailed topological or homological knowledge. I gave a version of the proof in Cohen (1993), but the algebra can be simplified further by slightly weakening Brown's conclusion, and we look at the simpler version here.

Squier also proved (the result was published after his death in Squier *et al.* (1994); see also Lafont (1994)) that a monoid with a finite convergent rewriting system satisfies a stronger condition, which he referred to as having *finite derivation type*. He showed that there was a monoid that did not have finite derivation type but is FP_∞. Recently, Cremans and Otto (1994) showed that a group has finite derivation type iff it is FP_3.

Much recent work in group theory has a very geometric flavour, and Hermiller and Meier (1994b) have shown a connection between groups with finite convergent rewriting systems and certain geometric properties. Precisely, they show that such a group has the geometric property known as a *tame 1-combing*. It is known that any group with a tame 1-combing is FP_∞, so we have another proof of Anick's Theorem, but only for groups. The weaker property known as a *bounded combing* is discussed in Alonso (1992), where groups with this property are shown to be FP_∞. Many groups with tame 1-combings are known, some of which have additional properties. It would be of interest to know whether such groups must have a finite convergent rewriting system.

Hermiller and Meier (1994b) also show that if the finite convergent rewriting system has additional properties (for instance, if it is length-reducing), further geometric properties of the group can be obtained.

The relevant properties of rewriting systems are recalled in Section 2. Sections 4 and 5 contain the homological results needed to discuss the FP_n and FP_∞ properties, with their topological background sketched in Section 3. Section 6 lists all the results I know of about these properties. This is intended as a source of information, which can be used in conjunction with Anick's theorem to obtain interesting monoids or groups that have no finite convergent rewriting system (or as examples in which one may want to look for such a system). Section 7 contains the proof of Brown's theorem, and Section 8 compares the different approaches.

2. String rewriting

A *rewriting system* on a set X is a set $\mathscr{R}$ of pairs of elements of the free monoid X^*. We define a relation $\rightarrow$ on X^* by $ulv \rightarrow urv$ for any pair $(l, r) \in \mathscr{R}$ and any (possibly empty) words u and v of X^*, and we then say that *ulv rewrites to urv*. We denote by $\overset{+}{\rightarrow}$ and $\overset{*}{\rightarrow}$ the transitive closure and the reflexive transitive closure of $\rightarrow$. The monoid *presented by* $\langle X ; \mathscr{R}\rangle$ is defined to be the quotient of X^* by the equivalence relation generated by $\rightarrow$.

We say that a word w is *reducible* if there is some word z with $w \rightarrow z$; if there is no such z we call w *irreducible*.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

An arbitrary relation $\succ$ on a set S is called *noetherian* (or *well-founded*) if any sequence $s_1 \succ s_2 \succ \ldots \succ s_n \succ s_{n+1} \succ \ldots$ is finite. In particular, if $\rightarrow$ is noetherian, we see that any sequence of rewritings must eventually stop with an irreducible word.

We say that $\mathscr{R}$ is *convergent* (also called *canonical* or *complete*) if it is noetherian and there is only one irreducible word in each equivalence class. Results of Newman (1943) and others give conditions equivalent to this. In particular, if a finite system is noetherian (for instance, if r is shorter than l for all $(l,r) \in \mathscr{R}$), one can decide whether or not it is convergent. For general properties of rewriting systems see the books by Jantzen (1988) and Book and Otto (1993).

Two rewriting systems on the same set are called equivalent if the equivalence relations they define are the same. Knuth and Bendix (1970) proved that to every rewriting system there is an equivalent convergent rewriting system (which may, however, be infinite even if the original system is finite).

We say that the convergent rewriting system $\mathscr{R}$ is *minimal* if, for each $(l,r) \in \mathscr{R}$, r is irreducible with respect to $\mathscr{R}$ and l is irreducible with respect to $\mathscr{R} - \{(l,r)\}$. It is well known (see Kapur and Narendran (1985) or Theorem 2.4 of Squier (1987)) that to every convergent rewriting system there is an equivalent minimal convergent rewriting system, which is finite if the original system is finite. Indeed, as shown in the references cited, the most obvious modification of $\mathscr{R}$ is suitable. That is, we begin by replacing each pair (l,r) by $(l,\hat{r})$ where $\hat{r}$ is the irreducible word obtained from r by repeated reduction. We then omit any pair (l_1,r_1) for which there is a pair (l,r) and words u and v with $l_1 = ulv$. Finally, for each l we omit all but one pair with left-hand side l. If $\mathscr{R}$ is not convergent, however, the resulting system need not be equivalent to $\mathscr{R}$.

We say that $\mathscr{R}$ is *strongly minimal* if it is minimal and each element of X is irreducible. Any minimal convergent rewriting system $\mathscr{R}$ on X can be modified to a strongly minimal convergent rewriting system presenting the same monoid. We define X_2 to be $\{x \in X;$ there is a pair $(x,r_x) \in \mathscr{R}\}$, and let $X_1 = X - X_2$. Since $\mathscr{R}$ is minimal, no member of X_2 occurs on the right-hand side of any member of $\mathscr{R}$. Also, the only member of $\mathscr{R}$ that has the element x of X_2 occurring in its left-hand side is (x,r_x). It follows that the monoid presented by $\langle X;\mathscr{R}\rangle$ is also presented by $\langle X_1;\mathscr{R}_1\rangle$, where $\mathscr{R}_1 = \mathscr{R} - \{(x,r_x); x \in X_2\}$. This latter system is easily seen to be convergent and strongly minimal. The two systems will not be equivalent, since they are defined over different generating sets.

We conclude this section by proving some results about noetherian relations, which will be used in Section 6.

Let $\succ$ be a relation on a set S. We define the *height* of an element s as the supremum of all k for which there is a sequence $s \succ s_1 \succ \ldots \succ s_k$. This height is either a non-negative integer or is infinite.

There are two reasons why s may have infinite height. The first is that there is an infinite sequence $s \succ s_1 \succ \ldots \succ s_n \succ \ldots$. The second is that there is no such infinite sequence, but, for all n there is a finite sequence $s \succ s_{n1} \succ \ldots \succ s_{nn}$. The following result is well known, and its proof is easy. We make explicit use of it in the proof of Brown's theorem. It is needed in Anick's proof, and in all the other proofs, but its use there is implicit, and readers have to look closely to see that it plays a crucial role in the performance of an inductive process.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

Lemma 1. (König's Lemma) Suppose that for each s there are only finitely many t with $s \succ t$. Then an element s has infinite height only if there is an infinite sequence $s \succ s_1 \succ \ldots \succ s_n \succ \ldots$. In particular, if, in addition, $\succ$ is noetherian, every element has finite height.

Lemma 2. Let $\gg$ and $\Rightarrow$ be noetherian relations on a set S such that whenever $u \Rightarrow v \gg w$ there is some z with $u \gg z \Rightarrow w$. Define $\succ$ by $u \succ v$ if either $u \gg v$ or $u \Rightarrow v$. Then $\succ$ is noetherian.

Proof. Let $\overset{*}{\Rightarrow}$ be the reflexive transitive closure of $\Rightarrow$. An easy induction shows that if $u \overset{*}{\Rightarrow} v \gg w$, there is some z such that $u \gg z \overset{*}{\Rightarrow} w$.

Suppose we have an infinite sequence $v_1 \succeq v_2 \succeq \ldots$. We want to show that this sequence is ultimately constant. It will be enough to find elements u_i such that, for all i, $u_i \overset{*}{\Rightarrow} v_i$ and either $u_i \gg u_{i+1}$ or $u_i = u_{i+1}$ and one of $v_i \Rightarrow v_{i+1}$ and $v_i = v_{i+1}$ holds. For, as $\gg$ is noetherian, the sequence $\{u_i\}$ will be ultimately constant. Once this constancy is achieved, the sequence $\{v_i\}$ will also be ultimately constant, since $\Rightarrow$ is also noetherian.

We begin by defining $u_1 = v_1$, and suppose that we have defined u_i for all $i \leqslant k$. If $v_k \Rightarrow v_{k+1}$ or $v_k = v_{k+1}$, we let $u_{k+1} = u_k$. If $v_k \gg v_{k+1}$, the first paragraph of the proof shows that there is a suitable u_{k+1}. □

3. The topological background

Our algebraic constructions and results first occurred in a topological setting, and closely mimic the corresponding topological constructions and results. It may therefore be helpful to begin with an account of the topological material. Because this is just for background, I leave many of the terms undefined, with the related algebraic concepts precisely defined later. Also, I omit technical conditions necessary for some of the results to hold (for instance, there is usually a need for the spaces considered to be reasonably well behaved).

Let G be a group. Let X and Y be spaces, both of which have trivial higher homotopy groups and have fundamental group G. Then X is homotopy equivalent to Y. Consequently, these spaces have the same homology groups (and cohomology groups). It is natural to refer to these groups (which, for spaces with a geometric or combinatorial structure, are easily defined in terms of the geometry) as the homology groups of the group G.

This result was proved in 1935, but it was not until about ten years later that an explicit construction for the homology groups was made. This construction, and the earlier result, are the sources from which the wide river of homological algebra sprung.

For a simplicial complex (that is, a space made up of simplexes – triangles in dimension 2, tetrahedra in dimension 3, and so on), the homology groups are fairly easy to describe. The chain group in dimension n is defined to be the free abelian group with basis the n-simplexes. A boundary operator from the n-chains to the $(n-1)$-chains is defined by requiring the boundary of a simplex to be the signed sum of its faces. The homology groups then measure how much the subgroup of all n-chains with zero boundary differs from the subgroup of all boundaries of $(n+1)$-chains. For general spaces, we have a related construction, but the chain groups are much harder to describe.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

How can we find such a space X? We can start with a contractible space C on which G acts nicely, and define X to be the quotient C/G.

We then want to obtain the homology groups of X. Because G acts on C, it will also act on the chain groups of C, and the quotient of these chain groups by the action of G gives us the chain groups of X. Because C is contractible, the chain groups of C form a resolution of G (this algebraic notion is defined in the next section). Then the construction of the homology groups of X from the chain groups of X, and the construction of these chain groups from those of C, is exactly the algebraic construction of the homology groups of G given in the next section.

We still have to construct a suitable space C. We build C out of simplexes (we can regard the constructed object either as something abstract and combinatorial, or as a topological space). We take one n-simplex $[y_1, \ldots, y_{n+1}]$ to each $(n+1)$-tuple $y_1, \ldots, y_{n+1}$ of elements of G. This simplex has $n+1$ $(n-1)$-dimensional faces, each obtained by omitting one of the y_i. Also, G acts on C by the obvious rule $[y_1, \ldots, y_{n+1}]z = [y_1 z, \ldots, y_{n+1} z]$. It is not difficult to show that C is contractible.

This construction defines the simplexes in what may be called a homogeneous fashion. There is an alternative approach, which may be called non-homogeneous. For this we take one n-simplex, denoted by $(x_1, \ldots, x_n)x_{n+1}$ to each $(n+1)$-tuple of elements of G. We have an obvious action of G on these simplexes, and we have a bijection between the simplexes in the two constructions under which $(x_1, \ldots, x_n)x_{n+1}$ corresponds to $[y_1, \ldots, y_{n+1}]$ if $y_i = x_i \ldots x_{n+1}$ for $1 \leqslant i \leqslant n+1$; equivalently, $x_{n+1} = y_{n+1}$ and $x_i = y_i y_{i+1}^{-1}$ for $1 \leqslant i \leqslant n$.

In this construction, the non-homogeneous $(n-1)$-dimensional simplexes corresponding to the faces of $[y_1, \ldots, y_{n+1}]$ are easily seen to be $(x_2, \ldots, x_n)x_{n+1}$ for the first face, $(x_1, \ldots, x_{n-1})x_n x_{n+1}$ for the last face, and $(x_1, \ldots, x_{i-2}, x_{i-1}x_i, x_{i+2}, \ldots, x_n)x_{n+1}$ for the i-th face.

This description of the faces explains the definition of the *bar resolution* in a later section. When we are dealing with a group G, the homogeneous definition is simpler. But for an arbitrary monoid M, rather than a group, we no longer get a bijection between the homogeneous simplexes and the non-homogeneous ones, and the complex constructed using the homogeneous simplexes does not provide a resolution of M; also, the action of the monoid on the homogeneous simplexes is no longer free.

4. Complexes and resolutions

Let M be a monoid. Let $\mathbf{Z}M$ be the *monoid ring* of M. This is defined to be the set of all finite formal sums $\sum n_u u$, where u runs over all elements of M, $n_u \in \mathbf{Z}$ and $n_u = 0$ for all but finitely many u. Addition and multiplication are defined by $(\sum n_u u) + (\sum p_u u) = \sum (n_u + p_u)u$ and $(\sum n_u u)(\sum p_u u) = \sum q_u u$ where $q_u = \sum_{vw=u} n_v p_w$. It is easy to check that this makes $\mathbf{Z}M$ a ring. We do not use any special properties of $\mathbf{Z}$, and the whole theory would be identical if we chose an arbitrary commutative ring for coefficients. However, $\mathbf{Z}$ is the most important case, so we do not use the more general notation.

A *free chain complex over* M consists of a sequence of free right $\mathbf{Z}M$-modules P_n for all $n \geqslant 0$ and homomorphisms $\partial_n : P_n \to P_{n-1}$ for all $n > 0$ such that $\partial_n \partial_{n+1} = 0$ for

$n > 0$. We shall usually just refer to a *complex*, since we do not consider other kinds of complex. We can just as well work with left modules, but choose right modules so as to be consistent with Brown's notation (and Anick's, but not Squier's).

We say that a complex is *augmented* if there is also a homomorphism $\epsilon : P_0 \to \mathbf{Z}$ such that $\epsilon\partial_1 = 0$. Here, $\mathbf{Z}$ is regarded as a module on which M acts trivially.

An augmented complex is called a *resolution* of M if we have $\operatorname{im}\partial_{n+1} = \ker\partial_n$ for $n > 0$, and $\operatorname{im}\partial_1 = \ker\epsilon$ and $\epsilon P_0 = \mathbf{Z}$. Note that for any complex the left-hand side is contained in the right-hand side. Strictly speaking, we should refer to a *free resolution of* $\mathbf{Z}$, but this is the only kind we shall consider.

Every monoid has a resolution. For, suppose we have a partial resolution; that is, the modules and corresponding homomorphisms are defined only for $n \leqslant k$. Every module is the quotient of a free module. So we can find a free module P_{k+1} that has a homomorphism onto $\ker\partial_k$, and define ∂_{k+1} to be this homomorphism. Inductively, we can extend any partial resolution to a resolution. To start the induction, we may choose P_0 to be $\mathbf{Z}M$ and $\epsilon(\sum n_u u) = \sum n_u$. Resolutions constructed like this are too big to be useful, but a smaller resolution will be constructed in the next section, and an even smaller resolution is obtained by Anick's and Brown's theorems.

Theorem 1. (Anick's Theorem) Any strongly minimal convergent rewriting system on a monoid M gives rise to a resolution of M. This resolution will be finitely generated in each dimension if the rewriting system is finite.

This will be proved using Brown's method in Section 6.

Given a sequence of modules and homomorphisms, it is usually straightforward to see that it forms an (augmented) complex. A direct proof that it is a resolution may be quite hard. Fortunately, to show that a complex is a resolution it is enough to prove the existence of other homomorphisms satisfying certain conditions.

Suppose we have an augmented complex with modules P_n and homomorphisms ∂_n. A *contracting homotopy* consists of a sequence of homomorphisms of abelian groups (not of modules) $\sigma_n : P_n \to P_{n+1}$ and $\eta : \mathbf{Z} \to P_0$ such that, for $n > 0$ and $x \in P_n$ we have $\partial_{n+1}\sigma_n x + \sigma_{n-1}\partial_n x = x$, and for $x \in P_0$ and $r \in \mathbf{Z}$ we also have $\partial_1\sigma_0 x + \eta\epsilon x = x$ and $\epsilon\eta r = r$. In examples we usually have $P_0 = \mathbf{Z}M$ with $\epsilon(\sum n_u u) = \sum n_u$, and we then take η to be given by $\eta r = r1$, where 1 is the identity element of M.

Lemma 3. An augmented complex with a contracting homotopy is a resolution.

Proof. Let x be in P_n, where $n > 0$ (the other cases are similar but simpler). Then (omitting subscripts for ease of notation) $x = \partial\sigma x + \sigma\partial x$, so if $\partial x = 0$, we have, as required, $x = \partial(\sigma x)$. □

It is not hard to see that every resolution has a contracting homotopy, which is defined inductively using the fact that the modules are free abelian groups. Since we do not need this result, the proof is left to the reader.

Let P be a free $\mathbf{Z}M$-module with basis S. We denote by $\tilde{P}$ the free abelian group with basis S. Let Q be another free module with basis T, and let $\phi : P \to Q$ be a homomorphism. Then, for all $s \in S$, we can write $\phi s = \sum tc_{st}$, where $c_{st} \in \mathbf{Z}M$ and, for each s, $c_{st} = 0$ for all but finitely many t. We define $\tilde{\phi} : \tilde{P} \to \tilde{Q}$ by $\tilde{\phi}s = \sum t(\epsilon c_{st})$, where $\epsilon(\sum n_u u) = \sum n_u$. The modules and homomorphism can be defined without mention of a

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

basis. Specifically, we can define $\tilde{P}$ to be $P \otimes_{\mathbf{Z}M} \mathbf{Z}$, where $\otimes$ denotes the tensor product. This is in principle a better approach, but it requires knowledge of tensor products, which readers may not have and do not need for the properties that concern us.

It is easy to see that $\widetilde{\operatorname{im}\phi} = \operatorname{im}\tilde{\phi}$, and that $\widetilde{\ker\phi} \subseteq \ker\tilde{\phi}$. In general this inclusion is not an equality.

Suppose we have a resolution with modules P_n and homomorphisms ∂_n. Then we have the corresponding abelian groups $\widetilde{P_n}$ and homomorphisms of abelian groups $\widetilde{\partial_n}$. Since $\partial_n\partial_{n+1} = 0$, the previous paragraph tells us that $\operatorname{im}\widetilde{\partial_{n+1}} \subseteq \ker\widetilde{\partial_n}$. The quotient group $\ker\widetilde{\partial_n}/\operatorname{im}\widetilde{\partial_{n+1}}$ is called the nth *homology group* of M.

These homology groups appear to depend on the choice of resolution. But it is well-known (see Hilton and Stammbach (1971), Mac Lane (1963), or Lafont and Prouté (1991)) that in fact they only depend on M. This is, of course, a key result, without which homology theory would be impossible. Also, these groups are the homology groups with integer coefficients. It is possible to use other coefficients, and also to define cohomology groups. Any of these groups can be used to check if Anick's condition holds.

5. The bar resolution

Let M be a monoid. We will construct two resolutions of M, the *unnormalised bar resolution* and the *normalised bar resolution*.

An n-tuple of elements of M will be called an *n-cell* or a *cell of dimension n*. Let B_n be the free right $\mathbf{Z}M$-module with basis the n-cells. In particular, B_0 has a single generator (), and we usually identify B_0 with $\mathbf{Z}M$ and () with 1_M, where 1_M is the identity of M (which we usually denote just by 1). We then define $\epsilon : B_0 \to \mathbf{Z}$ by $\epsilon \sum n_u u = \sum n_u$. We also define $\eta : \mathbf{Z} \to \mathbf{Z}M$ by $\eta n = n1_M$,

We define homomorphisms $\partial_{ni} : B_n \to B_{n-1}$ for $n > 0$ and $0 \leqslant i \leqslant n$ by

$$\partial_{nn}(u_1, \ldots, u_n) = (u_1, \ldots, u_{n-1})u_n,$$

$$\partial_{n0}(u_1, \ldots, u_n) = (u_2, \ldots, u_n),$$

and, for $0 < i < n$,

$$\partial_{ni}(u_1, \ldots, u_n) = (u_1, \ldots, u_{i-1}, u_iu_{i+1}, u_{i+2}, \ldots, u_n).$$

We then define ∂_n to be $\sum_0^n(-1)^{n-i}\partial_{ni}$. In particular, $\partial_1(u) = u-1$ (identifying B_0 with $\mathbf{Z}M$ and () with 1), $\partial_2(u, v) = (v)-(uv)+(u)v$, and $\partial_3(u, v, w) = -(v, w)+(uv, w)-(u, vw)+(u, v)w$. Plainly, $\epsilon\partial_1 = 0$.

It is easy to check that $\partial_{ni}\partial_{n+1,j} = \partial_{n,j-1}\partial_{n+1,i}$ for $0 \leqslant i < j \leqslant n+1$, from which we see that $\partial_n\partial_{n+1} = 0$ for all n. Thus we have a complex.

We may define a homomorphism σ_n of abelian groups from B_n to B_{n+1} that sends $(u_1, \ldots, u_n)u_{n+1}$ to $(u_1, \ldots, u_n, u_{n+1})$. It is easy to check that this is a contracting homotopy. Thus our complex is a resolution, which we call the *unnormalised bar resolution*.

We call the cell $(u_1, \ldots, u_n)$ degenerate if $u_i = 1$ for some i, and we let D_n be the submodule of B_n with basis the set of degenerate n-cells. Plainly, ∂_n maps D_n into D_{n-1} and σ_n maps D_n into D_{n+1}. It follows that ∂_n induces a homomorphism of modules from

B_n/D_n to B_{n-1}/D_{n-1}, and σ_n induces a homomorphism of abelian groups from B_n/D_n to B_{n+1}/D_{n+1}. Thus the modules B_n/D_n, which are plainly free with basis the non-degenerate n-cells, together with the homomorphisms induced by ∂_n also form a resolution.

We call this the *normalised bar resolution.* For notational reasons, we usually think of it as generated by all n-cells $(u_1, \ldots, u_n)$ with the relations $(u_1, \ldots, u_n) = 0$ if $u_i = 1$ for some i. We shall usually call this resolution the bar resolution, since we shall not need the unnormalised resolution.

6. The FP_∞ property

We say that a monoid is FP_∞ if it has a free resolution that is finitely generated in all dimensions. More generally, we say that it is FP_n if it has a free resolution that is finitely generated in all dimensions $\leqslant n$.

We have seen that any partial resolution can be extended to a resolution. Thus M will be FP_n if it has a partial resolution that is defined in dimensions $\leqslant n$ and finitely generated in each dimension $\leqslant n$.

The theorem of Anick, Brown and Kobayashi tells us that a monoid with a finite convergent rewriting system will be FP_∞. Thus it is of interest to give examples of monoids and groups that satisfy this condition and other examples that do not satisfy it. I will give here most of the results I am aware of concerning this property, so as to provide a resource for anyone interested in examples. As one of the major references (Bieri 1976) may not be widely available (having been published in a departmental lecture notes series), it seems worthwhile to record many of the results given there. Readers wanting to look at particular examples that cannot be treated by the methods discussed here will probably need to ask assistance from an expert in homology theory. The explicit presentations given (or obtainable) for some of these groups provide a source of examples to test.

As remarked in the introduction, in the rapidly developing field of geometric group theory, important classes of groups are FP_∞ and satisfy the stronger necessary condition given in Hermiller and Meier (1994b). I do not give examples of such groups; the interested reader should check the book by Epstein *et al.* (1992).

Let M be a monoid that is FP_∞ with a resolution by free modules P_n that are finitely generated for all n. Let $\tilde{P}_n$ be the corresponding free abelian group discussed in Section 3; this is finitely generated. Since every subgroup of a finitely generated free abelian group is finitely generated, $\ker \widetilde{\partial}_n$ will be finitely generated. Thus the homology group $H_n(M)$, being a quotient of this, will also be finitely generated. So we have seen that a monoid that is FP_∞ will have finitely generated homology groups in all dimensions. We can use any resolution to evaluate the homology groups. The bar resolution is too large for efficient calculation of these groups. However, Brown's theorem gives us a smaller resolution from any convergent rewriting system, even an infinite one. So we may be able to show, using this smaller resolution, that some homology group is not finitely generated; it will then follow that the monoid is not FP_∞, so cannot be given by any finite convergent rewriting system. (More generally, we could use coefficients other than $\mathbf{Z}$, and we could use the cohomology groups rather than the homology groups.) Here, Brown's theorem is of more

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

use than Anick's and Kobayashi's theorems, as it gives the boundary operators by a simple formula (of geometric origin), whereas they only give the boundary operators by a complicated inductive formula.

Squier (1987) cites several examples of finitely generated groups with a solvable word problem that are not FP_3, and hence cannot be presented by any finite convergent rewriting system. These examples belong to families of finitely presented groups with a solvable word problems, such that each family contains, for every $n \geqslant 2$, a group that is FP_n but not FP_{n+1}. These examples, which will be given later, satisfy Squier's condition if $n > 2$, but do not satisfy Anick's condition.

Squier (1987) also explicitly gives a family of monoids S_k for all $k \geqslant 0$, and proves the following facts. Each S_k is given by an infinite convergent rewriting system in a form from which one can easily see that the monoid has a solvable word problem. S_0 is not FP_2, and so (as remarked below) cannot be presented by any finite rewriting system, even one that is not convergent. For $k \geqslant 1$, S_k can be presented by a finite rewriting system (but this system is not convergent). For $k \geqslant 2$, S_k is not FP_3 (this is shown by using the infinite convergent rewriting system to calculate the homology groups), and so S_k cannot be presented by any finite convergent rewriting system. The monoid S_1 is FP_∞, but Squier showed in a later paper (published after his death as Squier *et al.* (1994); see also Lafont (1994)) that S_1 cannot be presented by any finite convergent rewriting system.

It is shown in Squier (1987) (and in Lafont and Prouté (1991); the result was known much earlier than Squier's paper) that any finitely generated monoid is FP_1 and any finitely presented monoid is FP_2. It is shown in Bieri (1976) (and elsewhere) that a group that is FP_1 is finitely generated. This does not hold for monoids, since a monoid that is right FP_∞ but not left FP_1 cannot be finitely generated. I do not know if there is a monoid that is both left and right FP_1 but is not finitely generated. For groups, the property FP_2 is equivalent, as shown in Bieri (1976), to a property known as *almost finitely presented*.

6.1. *Monoids and groups; differences and similarities*

Our definitions and proofs have been given in terms of right modules, so, strictly speaking, we should refer to the property of being right FP_∞. There will be a similar concept of left FP_∞. A monoid with a finite convergent rewriting system will be both right FP_∞ and left FP_∞.

For a group G, any left G-module can be regarded as a right G-module, and conversely, by defining $ug = g^{-1}u$ for any $g \in G$ and any u in the module. Thus a group is right FP_∞ if and only if it is left FP_∞, and the same holds for FP_n. For monoids, the two concepts can differ. It is shown in Cohen (1992) that there is a monoid that is right FP_∞ but not even left FP_1.

Most of the results on the FP_∞ and FP_n properties hold only for groups. We begin with some results relating the monoid M with presentation $\langle X; \mathscr{R} \rangle$ to the group G with the same presentation. There is a homomorphism i from M to G induced by the inclusion of X^* into F, the free group on X. More precisely, there are, by definition of monoid and group presentations, homomorphisms $\pi : X^* \to M$ and $\rho : F \to G$, and there is a natural inclusion j from X^* to F; we shall sometimes regard j as inclusion, so that it need not be

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

explicitly mentioned. Because the relations of M hold in G, we can define i by $i\pi = \rho j$. However, i need not be an inclusion, as is seen by taking $X = \{a\}$ and $\mathscr{R} = \{(a^2, a)\}$.

Lemma 4. Suppose that for any u_1 and u_2 in M there are v_1 and v_2 in M such that $u_1v_1 = u_2v_2$. Then $G = (iM)(iM)^{-1}$.

Proof. Since G is generated by iX, it is enough to show that $(iM)(iM)^{-1}$ is a subgroup. It is plainly closed under taking inverses. So look at a product $(iw_1)(iu_1)^{-1}(iw_2)(iu_2)^{-1}$. By hypothesis, we can find p and q in M such that $u_1p = w_2q$. Let $u = u_1p$, $v = w_1p$, and $z = u_2q$. Then the product we are looking at equals $(iv)(iu)^{-1}(iu)(iz)^{-1} = (iv)(iz)^{-1}$, as required. □

Lemma 5. To any homomorphism α from M into a group H there is a homomorphism β from G into H such that $\alpha = \beta i$. If there is a one–one homomorphism from M into some group, then i is one–one. If α is one–one and M satisfies the conditions of the previous lemma, then β is one–one.

Proof. Define a homomorphism $\hat{\beta} : F \to H$ by $\hat{\beta}x = \alpha\pi x$. For any $(l, r) \in \mathscr{R}$, we have $\pi l = \pi r$, so $\hat{\beta}l = \hat{\beta}r$. It follows that there is a homomorphism $\beta : G \to H$ such that $\hat{\beta} = \beta\rho$. Then $\beta i\pi = \beta\rho j = \hat{\beta}j = \alpha\pi$, and the first part of the lemma follows, as π maps onto M.

When α is one–one, we see at once from $\alpha = \beta i$ that i is one–one.

Suppose the condition of Lemma 4 holds. Then to any $g \in G$ there are u and v in G such that $g = (iu)(iv)^{-1}$. Then $\beta g = (\alpha u)(\alpha v)^{-1}$. If $\beta g = 1$, we get $\alpha u = \alpha v$. Thus, if α is one–one, $\ker \beta = \{1\}$, and thus β is one–one. □

Lemma 6. Suppose the condition of Lemma 4 holds. Then i is one–one if and only if, for all $u, v, w \in M$, $uw = vw$ implies $u = v$.

Proof. Let i be one–one, and suppose that $uw = vw$. Then $(iu)(iw) = (iv)(iw)$. Since G is a group, this implies that $iu = iv$, and so $u = v$.

Conversely, let M be a monoid satisfying the condition of Lemma 4. Define a relation $\sim$ on $M \times M$ by $(u, v) \sim (u_1, v_1)$ if there are w and w_1 such that $uw = u_1w_1$ and $vw = v_1w_1$. Then $\sim$ is obviously reflexive and symmetric, and it is easy to check, using the condition, that it is transitive. Furthermore, to any two equivalence classes A and B there are u, v, w such that A and B are the equivalence classes of (u, v) and (v, w), respectively. Also, the class of (u, w) depends only on A and B, not on the choices of u, v, w. Thus we have a multiplication on the set of equivalence classes that (it is easy to see) makes it into a group. This group is isomorphic to G, but we do not need this fact.

There is a homomorphism from M into this group that sends u into the class of $(u, 1)$. If u and v have the same image, there is some w such that $uw = vw$. Thus our second hypothesis tells us that this homomorphism is one–one, and so, by the previous lemma, i is one–one. □

Lemma 7. Let i be one–one. Then M has a solvable word problem if G has a solvable word problem. Also, if, in addition, M satisfies the condition of Lemma 4, then G has a solvable word problem if M has a solvable word problem.

Proof. Let u and v be in X^*. Then $\pi u = \pi v$ iff $i\pi u = i\pi v$ iff $\rho u = \rho v$. If G has a solvable word problem, we can decide whether or not this holds, and so M has a solvable word problem.

Conversely, suppose that the condition of Lemma 4 holds, and that M has a solvable word problem. By the condition, to any a and b in X^*, there are c and d in X^* such that $\pi(ac) = \pi(bd)$. Since M has a solvable word problem, such c and d can be computed from a and b (by testing all pairs until a suitable pair is found). Evidently, we then have $\rho(a^{-1}b) = \rho(cd^{-1})$. Iterating this procedure, for any w in the free group F we can compute y and z in X^* such that $\rho w = \rho(yz^{-1})$. Thus $\rho w = 1$ iff $\rho y = \rho z$, that is, iff $i\pi y = i\pi z$. Since i is one–one, this holds iff $\pi y = \pi z$, and we can decide whether or not this holds, since M has a solvable word problem. □

The proof of the next lemma is rather complicated, because it is a special case of a general homological result. Readers may prefer to note the statement, but omit its proof.

Lemma 8. Let i be one–one, and let the condition of Lemma 4 hold. If M is FP_m for some m with $1 \leqslant m \leqslant \infty$, then G is FP_m.

Proof. For convenience of notation, we shall regard i as inclusion, so that we can omit mention of it.

An easy induction shows that, for any k and any $u_1, \ldots, u_k \in M$ there are $v_1, \ldots, v_k \in M$ such that $u_1 v_1 = \ldots = u_k v_k$. It then follows that, for any k and any $g_1, \ldots, g_k \in G$, there are $w_1, \ldots, w_k \in M$ and $t \in M$ such that $g_i = w_i t^{-1}$ for all i. It is then easy to see that any element of $\mathbf{Z}G$ can be written as cu^{-1} for some $c \in \mathbf{Z}M$ and $u \in M$, and, finally, that, for any k, any k elements of $\mathbf{Z}G$ can be written as $c_i u^{-1}$ for some $u \in M$ and $c_1, \ldots, c_k \in \mathbf{Z}M$.

Let P be a free $\mathbf{Z}M$-module with basis S. The previous paragraph shows that any element of the free $\mathbf{Z}G$-module with basis S can be written as pu^{-1} for some $p \in P$ and $u \in M$. We therefore denote this free $\mathbf{Z}G$-module by PM^{-1}.

Let Q be another free $\mathbf{Z}M$-module with basis T, and let $\phi : P \to Q$ be a homomorphism. Then there are elements c_{st} of $\mathbf{Z}M$ such that, for all s, $\phi s = \sum tc_{st}$, and for each s there are only finitely many t with $c_{st} \neq 0$. We may define a $\mathbf{Z}G$-homomorphism $\phi M^{-1} : PM^{-1} \to QM^{-1}$ by $(\phi M^{-1})s = \sum tc_{st}$. Also, if ϵ is a $\mathbf{Z}M$-homomorphism from P to the trivial module $\mathbf{Z}$, we have integers k_s such that $\epsilon s = k_s$, and again we define ϵM^{-1} to be the $\mathbf{Z}G$-homomorphism from PM^{-1} to the trivial module $\mathbf{Z}$ that sends s to k_s. It is possible to define PM^{-1} and ϕM^{-1} without reference to a basis, but for our purposes it seems easier to use this definition.

It is easy to check that $\mathrm{im}(\phi M^{-1}) = (\mathrm{im}\,\phi)M^{-1}$, and that $\ker(\phi M^{-1}) = (\ker\phi)M^{-1}$, with similar results for ϵ.

It follows that any resolution of M by modules P_n and homomorphisms ∂_n gives rise to a resolution of G by modules $P_n M^{-1}$ and homomorphisms $\partial_n M^{-1}$. The lemma is now immediate. □

6.2. *Examples of monoids and groups that are FP_∞*

6.2.1. *Some easy examples* The simplest case is that of finite monoids. The bar resolution shows at once that any finite monoid is FP_∞.

Let M be the free monoid on X. Let $P_0 = \mathbf{Z}M$ and let P_1 be free on a basis $\{t_x\}$

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

bijective with X. Define $\partial_1 : P_1 \to P_0$ by $\partial_1 t_x = x - 1$, and define ϵ as usual. Plainly, $\epsilon\partial_1 = 0$. Let $0 \neq p \in P_1$. We can write $p = \sum t_x u_x$, where $u_x \in \mathbf{Z}M$ and $u_x = 0$ for all but finitely many x. Each u_x is a finite sum $\sum n_{xi} m_{xi}$ with $n_{xi} \in \mathbf{Z}$ and $m_{xi} \in M$, where i runs over some index set depending on x. Let m_{yj} have the maximum length among all the elements m_{xi} for all x and all associated i. Then $\partial_1 p \neq 0$, since it contains the term $n_{yj} y m_{yj}$, which cannot cancel against any other term.

Thus M has a resolution with P_0 and P_1 as above, with $P_n = 0$ for $n > 1$. Plainly, then, M is FP_∞ if X is finite.

A similar, but more complicated, analysis of elements of maximum length can be applied to the free group G on X. It gives a resolution with P_0 being $\mathbf{Z}G$, P_1 the free $\mathbf{Z}G$-module with basis $\{t_x\}$, and $P_n = 0$ for $n > 1$. Thus the free group on a finite set X is FP_∞.

Alternatively, we can regard the free group on X as the monoid with presentation $\langle X \cup \bar{X}; (x\bar{x}, 1), (\bar{x}x, 1) \text{ for all } x\rangle$. This rewriting system is convergent, so Brown's theorem tells us that the group is FP_∞ when X is finite. However, this approach does not show that we can take $P_n = 0$ for $n > 1$.

The free abelian monoid on $\{x_1, \ldots, x_n\}$ is presented by the convergent rewriting system $\{(x_j x_i, x_i x_j)\}$ for all $j > i$. Brown's theorem tells us that this monoid has a resolution for which P_k has basis the set of all k-tuples $(x_{j_1}, \ldots, x_{j_k})$ with $j_1 > \ldots > j_k$. In particular, this monoid is FP_∞ (and even has $P_k = 0$ for $k > n$).

By Lemma 8, the free abelian group on a finite set is also FP_∞.

Lyndon (1950) shows that a finitely generated group with a single defining relation is FP_∞. (Note that here we are looking at group presentations, not monoid presentations, when we say that there is a single defining relation.)

6.2.2. *Some simple constructions* We now look at how various group-theoretic constructions behave with regard to the FP_n properties. We will find that various results about FP_∞ have parallel results about convergent rewriting systems.

The next two results are shown in Bieri (1976).

Let H be a subgroup of the group G, and let H have finite index in G. Then, for $1 \leqslant n \leqslant \infty$, G is FP_n iff H is FP_n.

Let K be a normal subgroup of the group G, and let K be FP_∞. Then, for $1 \leqslant n \leqslant \infty$, G is FP_n iff G/K is FP_n. In particular, the direct product of FP_∞ groups is FP_∞.

The following analogues of these for rewriting systems are found in the unpublished paper Groves and Smith (1989); some of these results are published in Groves and Smith (1993).

Let H be a subgroup of the group G, and let H have finite index in G. If H has a finite convergent rewriting system, so does G. However, it is not known if H must have a finite convergent rewriting system when G has a finite convergent rewriting system.

Let K be a normal subgroup of the group G. If both K and G/K have finite convergent rewriting systems, so does G. In particular, the direct product of groups with finite convergent rewriting systems has a finite convergent rewriting system.

The following results are immediate from those given above. *Let $G_0 = \{1\}$, and, for*
$1 \leqslant i \leqslant k$, let G_i be a group containing G_{i-1}. Suppose that, for $1 \leqslant i \leqslant k$, either G_{i-1} has

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

finite index in G_i or G_{i-1} is a normal subgroup of G_i with G_i/G_{i-1} finitely generated and either free or free abelian or with one defining relation. Then G_k is FP_∞. If the last case is omitted, G_k has a finite convergent rewriting system.

The following criterion can be found in Bieri (1980).

Let G be a finitely presented group that is isomorphic to a subgroup of $GL_n(\mathbf{Q})$, the group of invertible $n \times n$ matrices with rational coefficients. If G is FP_∞, the centre of G, which is defined to be $\{g \in G; gh = hg$ for all $h \in G\}$, is finitely generated.

6.2.3. *Further constructions* Two constructions of importance in group theory are the amalgamated free product and the *HNN* extension. The former constructs a group $A *_C B$ from two groups A and B and a group C, which is a subgroup of both. In particular, if C is trivial, we get the free product $A * B$; this has presentation $\langle X \cup Y ; \mathscr{R} \cup \mathscr{S} \rangle$ when A and B have presentations $\langle X ; \mathscr{R} \rangle$ and $\langle Y ; \mathscr{S} \rangle$ with X and Y disjoint. The latter constructs a group from a group A, a subgroup C and a one–one homomorphism $\phi : C \to A$.

The following results are found in Bieri (1976). Here we have $1 \leqslant n \leqslant \infty$, with $n-1$ denoting ∞ if $n = \infty$.

Let G be the amalgamated free product $A *_C B$. If A and B are FP_n and C is FP_{n-1}, then G is FP_n. If G and C are FP_n, so are A and B. If G is FP_n and A and B are FP_{n-1}, then C is FP_{n-1}.

Let G be the HNN extension $\langle A, t; t^{-1}Ct = \phi C \rangle$. If A is FP_n and C is FP_{n-1}, then G is FP_n. If G and C are FP_n, so is A. If G is FP_n and A is FP_{n-1}, then C is FP_{n-1}.

Baumslag and Bieri (1976) define the class of *constructible* groups to be the smallest class $\mathscr{C}$ closed under isomorphisms and such that:

— $\{1\} \in \mathscr{C}$,
— $G \in \mathscr{C}$ if G has a subgroup H of finite index with $H \in \mathscr{C}$,
— an amalgamated free product $A *_C B$ is in $\mathscr{C}$ if A, B, and C are in $\mathscr{C}$,
— an *HNN* extension $\langle A, t; t^{-1}Ct = \phi C \rangle$ is in $\mathscr{C}$ if A and C are in $\mathscr{C}$.

From the previous discussions, it is immediate that all constructible groups are FP_∞. Those constructible groups that are solvable (also called soluble; this means that some term of the derived series is trivial, and it has no connection with having a solvable word problem) are characterised in Baumslag and Bieri (1976). It is shown in Groves and Smith (1989) that all constructible solvable groups have finite convergent rewriting systems. This paper also gives sufficient conditions for an amalgamated free product or an *HNN* extension to have a finite convergent rewriting system.

Let Γ be a finite graph, and suppose that, for each vertex $i \in \Gamma$, we have a group G_i with presentation $\langle X_i; R_i \rangle$. The *graph product* of the groups is defined to be the group with presentation $\langle \bigcup X_i; \bigcup R_i \cup S \rangle$, where $S = \{xyx^{-1}y^{-1}$ for all $x \in X_i$, $y \in X_j$ and i adjacent to $j\}$. This can be shown to depend only on the graph and the groups, not on the presentations chosen. The similar construction for monoids is called a partially commutative monoid.

The graph product can be constructed from direct products and amalgamated free products in a simple way, giving the next result (Cohen 1995).

The graph product of FP_∞ groups is FP_∞. It is shown in Hermiller and Meier (1994a) that convergent rewriting systems on the vertex groups (or monoids) give rise, in a natural way, to a convergent rewriting system on the graph product, and this is finite if the original rewriting systems are finite.

6.3. *Examples of groups that are not FP_∞*

6.3.1. *Bieri's groups A_n and B_n* We need to begin with the notion of the *split extension* of one group by another. We define an *action* of the group G on the group H as a function from $G \times H$ to H, denoted by $\cdot$, such that

— $g \cdot (hh_1) = (g \cdot h)(g \cdot h_1)$ for all $g \in G$ and $h, h_1 \in H$,
— $(gg_1) \cdot h = g \cdot (g_1 \cdot h)$ for all $g, g_1 \in G$ and $h \in H$,
— $g \cdot 1_H = 1_H$ for all $g \in G$, where 1_H is the identity element of H,
— $1_G \cdot h = h$ for all $h \in H$, where 1_G is the identity element of G.

Equivalently, an action of G on H is a homomorphism from G to the group of automorphisms of H.

Given an action of G on H, the *split extension* of H by G is the group whose underlying set is $H \times G$ with the multiplication defined by $(h, g)(h_1, g_1) = (h(g \cdot h_1), gg_1)$. It is easy to see that this multiplication makes $H \times G$ a group, and that this group contains a subgroup isomorphic to G and a normal subgroup isomorphic to H. The intersection of these subgroups is trivial, and their product is the whole group.

We now apply this construction. Let $D_n = \langle x_1, y_1 \rangle \times \ldots \times \langle x_n, y_n \rangle$, the direct product of n free groups of rank 2. Let F be the free group of countable rank with basis $\{a_k; k \in \mathbf{Z}\}$, and let $\mathbf{Z}[1/d]$, where d is an integer greater than 1, be the set of rational numbers of the form r/d^s for some $r \in \mathbf{Z}$ and $s \in \mathbf{N}$. We may specify an action of D_n on either of these groups by specifying how the generators act, and any action of these generators extends to an action of D_n if it respects the relations. In particular, we have an action of D_n on F given by $x_i \cdot a_k = a_{k+1} = y_i \cdot a_k$ for all $k \in \mathbf{Z}$ and $1 \leqslant i \leqslant n$, and an action of D_n on $\mathbf{Z}[1/d]$ by $x_i \cdot q = dq = y_i \cdot q$ for all $q \in \mathbf{Z}[1/d]$ and $1 \leqslant i \leqslant n$. Let A_n and B_n be the corresponding split extensions.

Then A_n and B_n have normal subgroups isomorphic to F and $\mathbf{Z}[1/d]$, respectively, the quotient groups being D_n in each case. It is then easy to see that both A_n and B_n have a solvable word problem. It is also easy to check that (for instance), for any j, the relation $x_2 \cdot a_j = a_{j+1}$ can be derived from the relations $x_2 \cdot a_0 = a_1$, $x_2x_1 = x_1x_2$, and, for all k, $x_1 \cdot a_k = a_{k+1}$. As a result, we find that A_n has a finite presentation with generators $x_1, y_1, \ldots, x_n, y_n$, and a with the relations $x_ix_j = x_jx_i$, $y_iy_j = y_jy_i$, $x_iy_j = y_jx_i$ for $1 \leqslant i < j \leqslant n$, and $x_iax_i^{-1} = x_1ax_1^{-1}$, $y_iay_i^{-1} = x_1ax_1^{-1}$ for $1 \leqslant i \leqslant n$. Similarly B_n has a finite presentation with the same generators as A_n, and with the relations of A_n and one additional relation $z_1ax_1^{-1} = a^d$. It is shown in Bieri (1976) that both A_n and B_n are FP_n but not FP_{n+1}. It is these examples that Squier uses.

6.3.2. *Abels' group G_n* Let p be a prime, and let $\mathbf{Z}[1/p]$ be the ring consisting of all rational numbers of form r/p^n.

The group $GL_n(\mathbf{Z}[1/p])$ consists of those $n \times n$ matrices with entries in $\mathbf{Z}[1/p]$ that have

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

an inverse of the same form. Let G_n be the subgroup of $GL_{n+1}(\mathbf{Z}[1/p])$ that consists of those matrices $A = (a_{ij})$ with $a_{11} = 1 = a_{n+1,n+1}$ and $a_{ij} = 0$ for $i > j$. As a matrix group, G_n has a solvable word problem. Brown shows that, for $n > 2$, G_n is finitely presented and FP_{n-1}, but not FP_n.

We can find an explicit finite presentation of G_n. It is a split extension of the group T of unitriangular matrices (that is, matrices with $a_{ii} = 1$ for all i and $a_{ij} = 0$ for $i > j$) by the group of those diagonal matrices that are in G_n (that is, $a_{ij} = 0$ for $i \neq j$ and $a_{11} = 1 = a_{n+1,n+1}$). For $1 \leqslant k \leqslant n$, let T_k be the subgroup of T consisting of those matrices in T such that $a_{ij} = 0$ for $1 \leqslant j - i \leqslant k$. Then T_{k+1} is a normal subgroup of T_k such that the quotient T_k/T_{k+1} is the direct sum of $n+1-k$ copies of $\mathbf{Z}[1/p]$, regarded as a group under addition.

Now, if P is a group with a normal subgroup Q, it is easy to find a presentation of P from presentations of Q and P/Q. Thus we can find inductively an (infinite) presentation of each T_k. This leads to an infinite presentation of T, and then of G_n. A finite presentation of G_n can be found from its infinite presentation by routine but tedious manipulations. Details are left to the reader. Abels (1979), which first discussed the group G_3, gives a group presentation (not a monoid presentation) of G_3 with five generators and thirteen relations.

6.3.3. *Brown's group H_n* Brown (1987) gives conditions under which a group acting on a space is FP_∞, and under which it is FP_n but not FP_{n+1}. In Brown (1984) he discusses circumstances under which such a group is finitely presented. When the action is given in sufficient detail, a finite presentation can be explicitly determined, but it is often quite complicated to do so. The proofs of his results require a considerable knowledge of homological methods. The statements, however, are not difficult to understand. In particular, the statements of Corollary 3.3 in Brown (1987) and Theorem 3 in Brown (1984) require very little technical knowledge. His examples include the following ones.

Let H_n be the group of those permutations α of $\mathbf{N} \times \{0, \ldots, n-1\}$ for which there are integers $m_0, \ldots, m_{n-1}$ such that, for $0 \leqslant i \leqslant n-1$ and all sufficiently large $x \in \mathbf{N}$, we have $\alpha(x, i) = (x + m_i, i)$. It is easy to see that the function sending α to $(m_0, \ldots, m_{n-1})$ is a homomorphism. The kernel of this homomorphism is isomorphic (by means of the bijection from $\mathbf{N} \times \{0, \ldots, n-1\}$ to $\mathbf{N}$, which sends (k, i) to $kn + i$) to the group of those permutations of $\mathbf{N}$ that fix all but finitely many elements.

Given α, choose k so that $\alpha(x.i) = (x + m_i, i)$ for $x > k$ and $0 \leqslant i \leqslant n-1$. Since α is a permutation, it maps $\bigcup_i \{(x, i); x \leqslant k\}$ bijectively onto $\bigcup_i \{(x, i); x \leqslant k + m_i\}$. Hence the image of the homomorphism is contained in $\{(m_0, \ldots, m_{n-1}; \sum m_i = 0\}$. That the image is precisely this set is easily seen, using the permutations α_i for which

$$\alpha_i(x, 0) = (x + 1, 0), \alpha_i(0, i) = (0, 0), \alpha_i(x + 1, i) = (x, i), \alpha_i(x, j) = (x, j) \text{ for all } i \neq 0, i.$$

From this we can easily see that H_n has a solvable word problem. It is shown in Brown (1987) that, for $n > 2$, H_n is finitely presented and is FP_{n-1} but not FP_n. In principle, Brown's results could be used to find an explicit finite presentation of H_n, but it may be easier to use the technique sketched below.

We begin with presentations $\langle X_k; R_k \rangle$ of the finite symmetric groups S_k (for instance, that in 6.28(1) of Coxeter and Moser (1958), or one of the other presentations given by them). If these are chosen so that $X_k \subseteq X_{k+1}$ and $R_k \subseteq R_{k+1}$, then $\langle \bigcup X_k ; \bigcup R_k \rangle$ is an infinite presentation of the group of all permutations of $\mathbf{N}$ that move only finitely many elements. This is a normal subgroup of H_n, and we have described the quotient group, which has an obvious presentation. As in the discussion of Abels' group, this leads easily to an infinite presentation of H_n, and routine but complicated calculations enable us (with the aid of the permutations α_i defined above) to replace this by a finite presentation. The details are left to the reader.

6.4. *Further monoids and groups that are FP_∞*

The first of these is the monoid M presented on $X = \{x_i, \text{all } i \in \mathbf{N}\}$ by the rewriting system $\{(x_j x_i, x_i x_{j+1})$ for all i, j with $i < j\}$, and the group F with the same presentation. The history of this group is discussed in Brown (1987; 1992b). The rewriting system is evidently convergent, and so M has a solvable word problem. It is easy to check that the irreducible words of X^* are exactly the words $x_{i_1} \dots x_{i_k}$ for all k and all $i_1, \dots, i_k$ with $i_1 \leqslant \dots \leqslant i_k$. We can then see easily that M satisfies both hypotheses of Lemma 6. It follows, by Lemma 7, that F also has a solvable word problem. Brown shows by topological methods in Brown (1987) that F is FP_∞, and he indicates how the topological proof can be turned into an algebraic one. A slightly different proof is given in Brown and Geoghegan (1984). Further discussion in Brown (1992b) makes it clear (though this is not stated explicitly) that M is FP_∞; the proof in Brown (1992b) that F is FP_∞ is that of Lemma 8, stated in a more abstract homological version.

We will show that F can be finitely presented. This result is stated in the references cited, but the proof is left to the reader. As it is slightly messy, but still managable by hand, I give the details here. By contrast, M is right FP_∞ but is shown in Cohen (1993) not to be left FP_1, so it is not even finitely generated.

For $k > 0$ let r_k be the relation $x_0^{-1} x_k x_0 = x_{k+1}$, for $k > 1$ let s_k be the relation $x_1^{-1} x_k x_1 = x_{k+1}$, and for $k \geqslant 0$ let t_k be the relations $x_k^{-1} x_{k+1} x_k = x_{k+2}$; in particular, r_1 is the same as t_0 and s_2 is the same as t_1.

It is easy to see that, for $0 < i < j$, the relation $x_{i+1}^{-1} x_{j+1} x_{i+1} = x_{j+2}$ can be derived from r_i, r_j, r_{j+1} and the relation $x_i^{-1} x_j x_i = x_{j+1}$. It follows that all the relations can be derived from $\{r_k ; k > 0\} \cup \{s_k ; k > 1\}$. It is also clear that, for $k > 0$, r_{k+2} can be derived from r_k, r_{k+1}, t_k and t_{k+1}, and that, for $k > 1$, s_{k+2} can be derived from s_k, s_{k+1}, t_k and t_{k+1}. Thus we may take for the defining relations of F the relations t_k for all $k \geqslant 0$, together with the relations r_2 and s_3. The relations t_k for $k > 2$ can be used to eliminate the generators x_k for $k > 4$. We get a finite presentation of F with generators x_i for $0 \leqslant i \leqslant 4$ and relations r_2, s_3, t_0, t_1 and t_2. If we prefer, the relation s_3 can be replaced by the relation $x_1^{-1} x_3 x_1 = x_0^{-1} x_3 x_0$. Each of these finite presentations can easily be replaced by presentations on the two generators x_0, x_1.

Guba and Sapir (1995) show that this group, regarded as a monoid on the generators $x_0, x_1, \bar{x}_0, \bar{x}_1$ can be given by a convergent rewriting system that is regular (it is still unknown whether it has a finite convergent rewriting system).

There is a family of groups, which arise as groups of homeomorphisms or as automorphism groups, all of which are finitely presented infinite simple groups that are FP_∞. Some of the properties of these groups are obtained in Brown (1992b). In Brown (1992a) one of these groups is exhibited explicitly in such a way that a finite presentation can easily be derived from it.

Brown (1992a) defines a *triangle of groups* to consist of three groups G_i for $i = 1, 2, 3$, together with a subgroup G_{ij} of G_i for $i \neq j$ and isomorphism from G_{ij} to G_{ji} for $1 \leqslant i < j \leqslant 3$. Let G_i have presentation $\langle X_i; R_i \rangle$, where $X_i \cap X_j = \varnothing$ for $i \neq j$. Let Y_{ij} be a subset of the free group on X_i whose image in G_i generates G_{ij}, and let these sets be chosen so that the isomorphism from G_{ij} to G_{ji} is induced by a bijection from Y_{ij} to Y_{ji}. Brown defines the *triangle product* in a way that depends on the groups but not on their presentations. However, it is easy to see from his definition that this triangle product has presentation $\langle X_1 \cup X_2 \cup X_3; R_1 \cup R_2 \cup R_3 \cup S \rangle$, where S consists of all $y_{ij} y_{ji}^{-1}$ where $i < j$, $y_{ij} \in Y_{ij}$, and y_{ji} corresponds to y_{ij} in the bijection between Y_{ij} and Y_{ji}.

The triangle product is defined in a similar way to the amalgamated free product (which uses two groups rather than three, with the obvious modifications). But they behave very differently. The amalgamated free product always contains copies of the original groups, whereas the triangle product of non-trivial groups can be trivial.

Brown shows that one of the groups he considered can be described as a triangle product of symmetric groups of degrees 5, 6, and 7 in the following way. Let G_1 be the group of all permutations of the set $\{a, b, c, d, e\}$, let G_2 be the group of all permutations of $\{a, b, c, d, e_0, e_1\}$, and let G_3 be the group of all permutations of $\{a, b, c, d_0, d_1, e_0, e_1\}$. Let G_{12} be the subgroup of G_1 fixing e, G_{21} the subgroup of G_2 fixing both e_0 and e_1, G_{23} the subgroup of G_2 fixing d, and G_{32} the subgroup of G_3 fixing both d_0 and d_1, the isomorphisms being the obvious ones. Let G_{13} be the subgroup of G_1 that preserves the sets $\{a, b, c\}$ and $\{d, e\}$. Thus G_{13} is the direct product of the group of permutations of $\{a, b, c\}$ by a cyclic group of order 2 whose non-trivial element is the permutation $(d\ e)$. Finally, G_{31} is to be the direct product of the group of permutations of $\{a, b, c\}$ by the cyclic group of order 2 whose non-trivial element is $(d_0\ e_0)(d_1\ e_1)$.

From this description, it is routine to obtain a presentation of the triangle product. We could use various presentations of the symmetric group to get a number of presentations of the triangle product.

7. Brown's Theorem

Brown showed that, in circumstances to be given shortly, a complex (either in the algebraic sense of this paper or in a topological sense) can be replaced by a smaller complex that is equivalent, in a suitable sense, to the original complex. In particular, in the algebraic setting, if the original complex is a resolution, so is the smaller complex. Brown proved his result in a topological setting. A translation of his proof to the algebraic setting was given in Cohen (1993). If we are content to prove the result for resolutions, which will be enough to obtain Anick's result, the algebraic proof can be further simplified, and this simplified proof will be given in this section.

7.1. *The topological background*

Let T be a triangle (edges and interior), let T_1 be the union of two of its edges, and let T_2 be the third edge. It is easy to see that T can be deformed (pushed down, collapsed) into T_1, by moving each point in a direction perpendicular to T_2.

Let X be a space and let Y be the space obtained by attaching T to X along T_1. More precisely, let $f : T_1 \to X$ be a map, and let Y be obtained from the disjoint union of X and T by identifying t with ft for all $t \in T_1$. Then the collapsing of T into T_1 can be extended to a collapsing of Y into X. Thus Y and X are homotopy equivalent, and so have the same homology. We call T_2 a *redundant* edge with corresponding *collapsible* triangle T.

This can obviously be extended to higher dimensions; T could be a tetrahedron, T_2 one of its faces, and T_1 the union of its other faces, or, more generally, T could be a cell (simplex) of arbitrary dimension and T_1 the union of all but one of its faces. Furthermore, there is no need to add just one cell (simplex); we can add an arbitrary number of cells (possibly infinite) at once, and Y will still collapse into X.

We can proceed further, attaching new collapsible and redundant cells to Y (by means of maps into Y that are not maps into X), and so on. We get a sequence of spaces,

$$X = Y_0 \subseteq Y_1 \subseteq Y_2 \subseteq \ldots,$$

where each Y_n is obtained by adding collapsible and redundant cells to Y_{n-1}.

Then, as before, each Y_n collapses into Y_{n-1}, and, inductively, Y_n collapses into X. Let $Y_\infty = \bigcup Y_n$. A subtler argument shows that Y_∞ is homotopy equivalent to X.

A further generalisation is now needed. We are often given a space Y that we want to regard as coming from a simpler space X by adding cells in various stages. That is, we have a sequence

$$X = Y_0 \subseteq Y_1 \subseteq Y_2 \subseteq \ldots,$$

where each Y_n is obtained by adding cells to Y_{n-1}, and $Y = \bigcup Y_n$. In this situation, we will not expect that all the cells added to Y_{n-1} to form Y_n will be collapsible or redundant; we call the remaining cells *essential*. Brown's topological theorem states that Y is homotopy equivalent to a space Z obtained from X by adding, for each k, one k-cell for each essential k-cell occurring somewhere in the sequence (and, furthermore, the attaching map can be constructed).

This topological idea translates to the algebraic notion of a *collapsing scheme*, which will be defined shortly. All our proofs are given in the algebraic setting.

7.2. *Collapsing schemes and resolutions*

Let $\mathbf{P}$ be an augmented free chain complex. We specify a basis for each P_n, and we call the elements of this basis *n-cells*. A *collapsing scheme* for $\mathbf{P}$ consists of the following:

(1) a division of the cells into three disjoint classes, called the *essential*, *redundant*, and *collapsible* cells, with all 0-cells being essential, and all 1-cells being either essential or redundant,

(2) a function, called *height*, from the set of all redundant cells into $\mathbf{N}$,

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

(3) a bijection between the set of redundant n-cells and the set of collapsible $(n+1)$-cells such that, when c is the collapsible cell corresponding to the redundant cell r, for one of the two choices of sign any redundant cell occurring in $r \pm \partial c$ has smaller height than r.

A collapsing scheme on a resolution (not on an arbitrary complex) is called *strong* if it satisfies the following extra condition:

(4) there is a contracting homotopy σ such that, for any collapsible chain x and any $m \in M$ the chain $\sigma(xm)$ is collapsible.

We shall see later how this notion arises in the bar resolution.

Usually the height function is given implicitly and not explicitly. That is, we replace conditions (2) and (3) by

(2′) for each n, a noetherian relation $>$ on the set of redundant n-cells,
(3′) a bijection between the set of redundant n-cells and the set of collapsible $(n+1)$-cells such that, when c is the collapsible cell corresponding to the redundant cell r, for one of the two choices of sign any redundant cell r' occurring in $r \pm \partial c$ has $r > r'$.

For the rest of the theory we will assume that the sign in $\pm$ is always $-$. This can be achieved simply by replacing c by $-c$ for some collapsible cells to get a new basis. In the explicit construction given later, however, it is preferable to use $\pm$ because there is a natural choice of the cells.

König's Lemma and its preceding remarks, applied to the relation $\succ$ given by $r \succ r'$ if r' occurs in $r - \partial c$, show that, when (1), (2′) and (3′) hold, there will be a height function satisfying (2) and (3). Note that we only need to compare redundant cells of the same dimension.

Theorem 2. (Brown's Theorem) If a free resolution has a (strong) collapsing scheme, there is a free resolution that in dimension n has as basis the essential n-cells. Also, the boundary operator in dimension n is determined by the original boundary operator in dimension n and the collapsing scheme in dimensions $n-1$ and n. Every strongly minimal convergent rewriting system $\langle X ; \mathscr{R} \rangle$ gives rise to a strong collapsing scheme for the normalised bar resolution of the monoid presented by the rewriting system in such a way that the essential 1-cells and 2-cells are bijective with X and $\mathscr{R}$, respectively, and there are only finitely many essential cells in each dimension if the rewriting system is finite.

The theorem is true for any collapsing scheme, but the proof for a strong collapsing scheme is slightly easier, and is all that we need, so we will only prove this case. The proof of the theorem will take the rest of this section.

Let $\mathbf{P}$ be a resolution with a collapsing scheme. We call a chain essential, redundant or collapsible if all its cells are essential, redundant or collapsible, respectively. We denote by E_n the set of all essential n-chains. The resolution we construct will have E_n as its set of n-chains, but it will take some work to define the boundary operators.

We begin by defining a homomorphism $\theta_n : P_n \to P_n$. For a redundant cell r with corresponding collapsible cell c, we let $\theta r = r - \partial c$ (as usual, we omit subscripts wherever this does not cause confusion), we let $\theta e = e$ for any essential cell e, and we let $\theta c' = 0$ for any collapsible cell c'.

We define the height of a chain to be -1 if the chain has no redundant cells, and to be the maximum height of its redundant cells if it has any such cells. It is easy to see that if x is a chain of height $k \geqslant 0$, then θx has height less than k. It follows that (even if $k = -1$) $\theta^{k+1}x$ has no redundant cells, $\theta^{k+2}x$ is essential, and $\theta^m x = \theta^{k+2}x$ for $m \geqslant k+2$.

We define ϕx to be $\theta^m x$ for any $m \geqslant k+2$. Then ϕ_n can be regarded either as a homomorphism from P_n to itself or as a homomorphism from P_n to E_n; we rely on the context to determine which is meant.

We define $\delta_n : E_n \to E_{n-1}$ by $\delta_n = \phi_{n-1}\partial_n$.

Proposition 1. The modules E_n and homomorphisms δ_n form an augmented free chain complex.

Proof. Observe first that to any n-chain x there is a collapsible n-chain y and a collapsible $(n+1)$-chain z such that $x - \theta x = \partial y + z$. For this holds for all chains if it holds for all cells. However, if x is essential, we may take $y = 0$ and $z = 0$, if x is redundant, we may take $z = 0$ and y the collapsible cell corresponding to x, while if x is collapsible, we take $y = 0$ and $z = x$.

Since $\phi x = \theta^m x$ for some m, it is immediate that for each x

$$x - \phi x = \partial y + z \text{ for some collapsible } y, z \tag{†}$$

Since z is collapsible, there will be a redundant u such that $\partial z = u - \theta u$. Hence $\phi\partial(\partial y + z) = \phi\partial z = \phi(u - \theta u) = 0$, since $\partial\partial = 0$, and, by definition of ϕ, $\phi\theta = \phi$.

It follows that

$$\phi\partial x = \phi\partial\phi x. \tag{‡}$$

Since all 0-cells are essential, ϕ_0 is the identity, and $\epsilon\delta_0 = \epsilon\partial_0 = 0$.

Also, by the above, $\delta\delta = \phi\partial\phi\partial = \phi\partial\partial = 0$, as required. □

The first part of Brown's theorem is completed by the following proposition.

Proposition 2. Suppose that the collapsing scheme is strong. Define $\tau_n : E_n \to E_{n+1}$ by $\tau = \phi\sigma$. Then τ is a contracting homotopy for the complex of essential chains, so this complex is a resolution.

Proof. We begin by showing that $\phi\sigma\phi = \phi\sigma$. Take any chain u. By † there are collapsible chains y and z such that $u - \phi u = \partial y + z$. Thus $\phi\sigma u - \phi\sigma\phi u = \phi\sigma(\partial y + z)$. By hypothesis, σz is a collapsible chain, so $\phi\sigma z = 0$. Because σ is a contracting homotopy, we have $\phi\sigma\partial y = \phi y - \phi\partial\sigma y$. Now $\phi y = 0$, because y is collapsible. By hypothesis, σy is a collapsible chain, so there is a redundant chain w with $\partial\sigma y = w - \theta w$. As $\phi\theta = \phi$, we see that $\phi\sigma u - \phi\sigma\phi u = 0$, as required.

Now, for any essential chain x, $(\tau\delta + \delta\tau)x = (\phi\sigma\phi\partial + \phi\partial\phi\sigma)x$, by definition, and this, by the result just shown and ‡, equals $(\phi\sigma\partial + \phi\partial\sigma)x = \phi x$, because σ is a contracting homotopy. Since x is essential, we have $\phi x = x$, which shows that τ is a contracting homotopy. □

It is not difficult to prove the result for all collapsing schemes. Since we do not need it, the details will be left to the reader. One needs to use (†), and to begin by proving that for a non-zero collapsible chain x the chain ∂x is non-zero.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

7.3. *Rewriting systems and collapsing schemes*

Let $\mathscr{R}$ be a strongly minimal convergent rewriting system on the set X. Let M be the monoid with presentation $\langle X\,;\mathscr{R}\rangle$. Brown shows how to obtain from $\mathscr{R}$ a collapsing scheme on the normalised bar resolution of M. We shall follow the account in Brown (1992b). The details are given here only to make the current paper self-contained.

Even when $\mathscr{R}$ is infinite, the collapsing scheme may be simple enough to enable calculations to be made in the resolution by essential cells. We may then be able to prove that M has an infinitely generated homology group in some dimension, and so cannot be FP_∞, and therefore cannot be given by any finite convergent rewriting system.

We will regard the elements of M as the irreducible elements of X^*. If u and v are in M the product uv may be reducible. We will denote by $u \times v$ the irreducible element obtained from uv; thus $u \times v$ is the product in M of u and v.

The construction is somewhat complicated, so we will look explicitly at the low-dimensional cases first.

Let $v \in M - X$. Then in X^* we have $v = xu$ for some $x \in X$ and $u \in M$. As indicated in the topological background material, we are thinking of the cells as being added in stages. Since u is simpler than v, it is natural to expect that u will already be present when we want to add v. The 2-cell (x, u) has boundary $(x)u - (x \times u) + (u)$. Since xu is irreducible, $v = x \times u$, and the second term in the boundary is the 1-cell (v), while the third term is (u), which is already present. We therefore say that (v) is a redundant cell whose corresponding collapsible cell is (x, u). The essential 1-cells will be the cells (x) for $x \in X$.

By analogy with the 1-dimensional case, with v as above and any w, the 2-cell (v, w) is to be redundant, with corresponding collapsible 3-cell (x, u, w). The boundary of this cell is $(x, u)w - (x, u \times w) + (x \times u, w) - (u, w)$. Here the third term is our 2-cell (v, w), while (x, u) is collapsible. The other two terms may be collapsible, essential, or redundant (or zero, since we are working in the normalised bar resolution, if $u \times w = 1$). We will have to define the partial ordering on redundant cells so that these two precede (u, v) if they are redundant.

We still have to consider 2-cells of the form (x, w). Such a cell will, by our earlier discussion, be collapsible if xw is irreducible.

Suppose that xw is reducible. If no proper subword is reducible (equivalently, if xw is the left-hand side of an element of $\mathscr{R}$), there is no natural way of associating a 3-cell with (x, w), and we therefore call this cell essential.

Now suppose that some proper subword of xw is reducible. Since w is irreducible, we can write $w = pq$, where xp is the left-hand side of an element of $\mathscr{R}$. We now define (x, w) to be redundant, with corresponding collapsible 3-cell (x, p, q). Its boundary is $(x, p)q - (x, p \times q) + (x \times p, q) - (p, q)$. Here the second term is our redundant cell (x, w), while (x, p) is essential. The other two terms can be of any type. We must define our partial ordering so that they precede (x, w) if they are redundant.

We will now look at arbitrary dimensions. Recall that the n-cells are the n-tuples $(u_1, \ldots, u_n)$ with $u_i \in M - \{1\}$ for all i, and that, if $u_i = 1$ for some i, then $(u_1, \ldots, u_n)$ denotes 0.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

The cell $(u_1, \dots, u_n)$ is defined to be essential if

(i) $u_1 \in X$,
(ii) for all $i < n$, $u_i u_{i+1}$ is reducible,
(iii) for all $i < n$, no proper prefix of $u_i u_{i+1}$ is reducible.

It is obvious that there are only finitely many essential cells in each dimension if $\mathcal{R}$ is finite, so Brown's analysis provides a proof that a monoid with a finite convergent rewriting system is FP_∞.

Let $u_1 \notin X$, and write $u_1 = xv$ for some $x \in X$. Then the cell $(u_1, \dots, u_n)$ is to be redundant, with corresponding collapsible cell $(x, v, u_2, \dots, u_n)$. We say that the *level* of this redundant cell is 1.

Now let $u_1 \in X$. We say that $(u_1, \dots, u_n)$ has level i if i is as large as possible subject to $(u_1, \dots, u_{i-1})$ being essential; note that $i > 1$. Thus an essential cell has level $n+1$, and a non-essential cell has level at most n. If the cell is not essential, $(u_1, \dots, u_{i-1})$ is essential but $(u_1, \dots, u_i)$ is not, so either $u_{i-1}u_i$ is irreducible or some prefix of it is reducible. In the first case we call the cell collapsible, in the second case we call it redundant.

Suppose that $(u_1, \dots, u_n)$ is redundant of level i. Write u_i as vw, where v is as short as possible subject to $u_{i-1}v$ being reducible. Then the cell $(u_1, \dots, u_{i-1}, v)$ is essential, so $(u_1, \dots, u_{i-1}, v, w, u_{i+1}, \dots, u_n)$ is easily seen to be collapsible. We define a function from redundant cells to collapsible cells by letting this collapsible cell correspond to the redundant cell $(u_1, \dots, u_n)$.

It is easy to see that this is a bijection between the set of redundant n-cells and the set of collapsible $(n+1)$-cells, and that every collapsible cell comes from a redundant cell of one level lower.

We still need to define a relation $>_n$ on the n-cells (it is convenient to define $>_n$ on all n-cells, though we only need it on the redundant ones), and show that (3′) holds.

In X^*, we say that v is a subword of w, and write $w \Rightarrow v$, if $w = pvq$ for some words p and q that are not both empty (though one may be). We define $w > v$ to mean that either $w \overset{+}{\to} v$ (in the sense of the rewriting system) or $w \Rightarrow v$. Since $\mathcal{R}$ is noetherian, and $\Rightarrow$ is obviously noetherian, Lemma 2 tells us that $>$ is noetherian.

When b denotes the cell $(u_1, \dots, u_n)$, we let $W(b)$ be the element $u_1 \dots u_n$ of X^*, and we let $i(b)$ be the level of b.

Let b and b' be n-cells. We say that $b >_n b'$ if either $W(b) > W(b')$ or $W(b) = W(b')$ and $i(b) < i(b')$. Since $>$ is noetherian, and $i(b)$ and $i(b')$ are at most $n+1$, we see easily that $>_n$ is also noetherian.

Let $c = (u_1, \dots, u_n)$ be a collapsible n-cell of level $i+1$ with corresponding redundant $(n-1)$-cell r of level i. We have $\partial_{n0}c = (u_2, \dots, u_n)$ and $\partial_{nn}c = bu_n$, where $b = (u_1, \dots, u_{n-1})$. Both $W(\partial_{n0}c)$ and $W(b)$ are subwords of $W(c)$.

For $0 < j < n$, $\partial_{nj}c = (u_1, \dots, u_{j-1}, u_j \times u_{j+1}, \dots, u_n)$. By the definition of level and of essential, $u_i u_{i+1}$ is irreducible, while $u_j u_{j+1}$ is reducible for $0 < j < i$. In particular, we have $W(r) = W(c)$. Since $u_j u_{j+1}$ reduces to $u_j \times u_{j+1}$, we see that if $u_j u_{j+1}$ is reducible, then $W(c) \overset{+}{\to} W(\partial_{nj}c)$; in particular, this holds for $0 < j < i$, and may hold for some $j > i$. When $j > i$ and $u_j u_{j+1}$ is irreducible, we have $W(c) = W(\partial_{nj}c)$. Also, for $j > i$, $\partial_{nj}c$ begins with $u_1, \dots, u_i$, as does c itself, so its level is at least $i+1$.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

Thus ∂c has one entry $\pm r$, while all its other cells (whether redundant or not) are $<_n r$, and we have the final condition (3′) for a collapsing scheme.

Brown goes on to discuss methods of calculating the boundary maps, giving several examples. This material is easy to follow, now that we have made the general algebraic construction; readers are strongly recommended to look at these examples in Brown (1992b).

Brown and Geoghegan (1984) discuss the monoid M presented on generators $\{x_n; n \in \mathbf{N}\}$ by the convergent rewriting system $\{(x_j x_i, x_i x_{j+1});$ all i, j with $i < j\}$. The general analysis of the collapsing scheme associated with a convergent rewriting system shows that M has a resolution whose basis in dimension n is the set of all cells $(x_{j_1}, \ldots, x_{j_n})$. They show that there is a further collapsing scheme on this resolution for which there are exactly two essential cells in each dimension. Thus M is FP_∞, as previously remarked.

7.4. *An alternative ordering*

There is a variation of the definitions and proofs, which will be useful in the next section.

When b is an n-cell, and $u \in M$, we call bu an *n-term*. We call a term redundant, essential, or collapsible if the corresponding cell is redundant, essential, or collapsible. We modify the definition of a collapsing scheme to allow for a height function, or a partial ordering, on terms rather than cells. Precisely, we replace (2′) and (3′) by

(3″) for each n, a noetherian relation $>$ on the set of redundant n-terms,

(4″) a bijection between the set of redundant n-cells and the set of collapsible $(n+1)$-cells such that, when c is the collapsible cell corresponding to the redundant cell r, for one of the two choices of sign any redundant term $r'u'$ occurring in $ru \pm \partial cu$ has $ru > r'u'$.

It is easy to check that, with this variant definition, all the previous results apply, and once again a collapsing scheme gives rise to a resolution whose chains have as basis the essential cells.

We now associate with a convergent rewriting system a collapsing scheme in this new sense.

We change the definition of $\Rightarrow$, and now define $w \Rightarrow v$ if $w = vq$ for some non-empty q. As before, we define $w > v$ to mean that either $w \xrightarrow{+} v$ (in the sense of the rewriting system) or $w \Rightarrow v$, using this new sense of $\Rightarrow$. Then $>$ will be noetherian. Let $\gg$ be the transitive closure of $>$.

Fo a term bu, we define $W(bu)$ and $i(bu)$ to be $W(b)u$ and $i(b)$. We get a noetherian relation $>_n$ on the n-terms by $bu >_n b'u'$ if either $W(bu) > W(b'u')$ or $W(bu) = W(b'u')$ and $i(bu) < i(b'u')$.

As before, let $c = (u_1, \ldots, u_n)$ be a collapsible n-cell of level $i+1$ with corresponding redundant $(n-1)$-cell r of level i; let $u \in M$. With our new definition, we still have $\partial_{ni}(cu) = ru$ and $W(ru) = W(cu)$, and we have $ru >_{n-1} \partial_{nj}(cu)$ for $0 < j < n$ and $j \neq i$. If $u_n u$ is reducible, we have $W(cu) > W(\partial_{nn}(cu))$, while if $u_n u$ is irreducible, we have $W(cu) = W(\partial_{nn}(cu))$ and $i(\partial_{nn}(cu)) \geqslant i+1 > i(ru)$, so we also have $ru >_{n-1} \partial_{nn}(cu)$. Thus we have a collapsing scheme in the new sense.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

We also note that any term t, of whatever kind, occurring in $\theta(ru)$ satisfies $W(ru) \gg W(t)$ or $W(ru) = W(t)$.

It is then immediate that, for any term t and any term t' occurring in ϕt we have either $W(t) \gg W(t')$ or $W(t) = W(t')$.

We now consider an essential n-term eu, where $e = (u_1, \ldots, u_n)$. Suppose that $v = u_n u$ is irreducible, and let $e' = (u_1, \ldots, u_{n-1})$, which is an essential cell. Then $e'v = \partial_{nn}(cu)$, and $W(e'v) = W(eu)$. Also, because e is essential, $u_j u_{j+1}$ is reducible for $0 < j < n$, and we see that $W(eu) > W(\partial_{nj}(eu))$. Also, $W(eu) \Rightarrow W(\partial_{n0}(eu))$. Hence $W(e'v) > W(\partial_{nj}(eu))$ for $0 \leqslant j < n$.

Since $\delta = \phi\partial$, we see, from the remarks previously made about ϕ, that $\delta(eu) = e'v + q$, where any term t in q satisfies $W(e'v) \gg W(t)$, and $W(e'v) = W(eu)$.

Now suppose that $u_n u$ is reducible. Then $\partial_{nn}(eu) = (u_1, \ldots, u_{n-1})w$, where $w = u_n \times u$. Thus $W(eu) > W(\partial_{nn}(eu))$. So any term t in $\partial(eu)$ has $W(eu) > W(t)$. As before, it follows that any term t' in $\delta(ey)$ has $W(eu) \gg W(t')$.

We shall need these results in the next section.

8. Comparison with other approaches

In this section, we shall show that our approach leads to the same resolutions as the ones found by Kobayashi, Anick, and (up to dimension 3) Squier. Note that we will use brackets [and] as a means of grouping expressions together, because we are using parentheses (and) in the notation for essential cells; this will be done except for the few occasions when we need several levels of grouping. We need a notation because some of our functions are homomorphisms of abelian groups, not of modules; if f is such a function and e is a cell, we need to distinguish between $f(eu)$ and $(fe)u$.

8.1. *Kobayashi's method*

Kobayashi's definition uses as basis for the chains what we have called the essential cells. However, he uses an inductive definition for the boundary operator d_n and the contracting homotopy t_n. (Our notation is slightly different from his, in order to be consistent with that of Section 7.) We shall use the definition of $>$ given in Subsection 7.3. He defines

$$d_1(x) = x - 1$$

and, when $u = x_1 \ldots x_n$,

$$t_0 u = \sum (x_i) x_{i+1} \ldots x_n.$$

Suppose that d_i has been defined for $i \leqslant n$ and t_i has been defined for $i < n$. He then defines

$$d_{n+1}(u_1, \ldots, u_{n+1}) = (u_1, \ldots, u_n)u_{n+1} - t_{n-1} d_n[(u_1, \ldots, u_n)u_{n+1}].$$

Let $(u_1, \ldots, u_n)$ be an essential n-cell, and let $u \in M$. When $u_n u$ is irreducible (in particular, when $u = 1$), he defines $t_n[(u_1, \ldots, u_n)u] = 0$.

When $u_n u$ is reducible, we can write $u = vw$ where v is as short as possible with $u_n v$

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

reducible. Then $(u_1, \ldots, u_n, v)$ is an essential cell. He then defines

$$t_n[(u_1, \ldots, u_n)u] = (u_1, \ldots, u_n, v)w + t_n[(t_{n-1}d_n[(u_1, \ldots, u_n)v])w].$$

This last is shown to be a valid definition by noetherian induction, using other properties of the d_i and t_i that are themselves proved inductively. We want to identify his boundary operators and contracting homotopy with ours. For this we need to know that t_n satisfies the stated formula, but we do not have to show that this formula provides a definition of t_n.

Note that this approach (and Anick's) requires us to know both d_n and t_{n-1} in order to compute d_{n+1}. In turn, d_n requires knowledge of d_{n-1} and t_{n-2}. Thus we cannot use either Kobayashi's formula or Anick's to compute d_{n+1} until we have computed the boundary operators and contracting homotopy in all lower dimensions.

By contrast, our method starts with the bar resolution, in which the boundary operators and contracting homotopies are given by simple formulas. The division of the cells into essential, collapsible and redundant cells is also fairly simple, and the boundary operator for our resolution in dimension $n+1$ is determined using only these data in dimensions $n+1$ and n; there is no need to look at lower dimensions.

Let $u = x_1 \ldots x_n$, and, if $n > 1$, let $v = x_2 \ldots x_n$. We have $\sigma_0 1 = 0$, since we are regarding degenerate cells as zero. We have $\sigma_0 u = (u)$. If $n > 1$, the cell (u) is redundant, with corresponding collapsible 2-cell (x, v). We have $\partial(x, v) = (x)v - (xv) + (v)$, so $\theta(u) = (x)v + v$; similarly if $n > 2$, we obtain θv. Hence

$$\tau_0 u = \phi\sigma_0 u = \sum (x_i)x_{i+1} \ldots x_n = t_0 u.$$

Plainly, $d_1 = \delta_1$.

Now suppose that $d_i = \delta_i$ for $i \leqslant n$, and $t_i = \tau_i$ for $i < n$. To show that $d_{n+1} = \delta_{n+1}$, it is evidently enough to show that

$$\delta_{n+1}(u_1, \ldots, u_{n+1}) = (u_1, \ldots, u_n)u_{n+1} - \tau_{n-1}\delta_n[(u_1, \ldots, u_n)u_{n+1}].$$

However,

$$\tau_{n-1}\delta_n[(u_1, \ldots, u_n)u_{n+1}] = (u_1, \ldots, u_n)u_{n+1} - \delta_{n+1}\tau_n[(u_1, \ldots, u_n)u_{n+1}],$$

and $\sigma_n[(u_1, \ldots, u_n)u_{n+1}] = (u_1, \ldots, u_n, u_{n+1})$. As this is essential, we have

$$\tau_n[(u_1, \ldots, u_n)u_{n+1}] = (u_1, \ldots, u_n, u_{n+1}),$$

and thus we have the required formula.

To show that $t_n = \tau_n$, we use noetherian induction on $W[(u_1, \ldots, u_n)u]$. If $u = 1$, then $\sigma_n[(u_1, \ldots, u_n)u] = 0$, by definition, so $\tau_n[(u_1, \ldots, u_n)u] = 0$. If $u_n u$ is irreducible, then $\sigma_n[(u_1, \ldots, u_n)u]$ is collapsible, so $\tau_n[(u_1, \ldots, u_n)u] = 0$. Thus $t_n = \tau_n$ on terms of these kinds.

Now suppose that $u_n u$ is reducible, and write $u = vw$ as before. We have

$$\tau_{n-1}\delta_n[(u_1, \ldots, u_n)v] = (u_1, \ldots, u_n)v - \delta_{n+1}\tau_n[(u_1, \ldots, u_n)v].$$

Also, $\sigma_n[(u_1, \ldots, u_n)v] = (u_1, \ldots, u_n, v)$, which is essential, so

$$\tau_n[(u_1, \ldots, u_n)v] = (u_1, \ldots, u_n, v).$$

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

It then follows that

$$\tau_n[(\tau_{n-1}[\delta_n(u_1,\ldots,u_n)v])w] = \tau_n[(u_1,\ldots,u_n)u] - \tau_n\delta_{n+1}[(u_1,\ldots,u_n,v)w].$$

However,

$$\tau_n\delta_{n+1}[(u_1,\ldots,u_n,v)w] = (u_1,\ldots,u_n,v)w - \delta_{n+2}\tau_{n+1}[(u_1,\ldots,u_n,v)w].$$

As already remarked, because vw is irreducible, $\tau_{n+1}[(u_1,\ldots,u_n,v)w] = 0$. That is, we have

$$\tau_n[(\tau_{n-1}[\delta_n(u_1,\ldots,u_n)v])w] = \tau_n[(u_1,\ldots,u_n)u] - (u_1,\ldots,u_n,v)w.$$

This formula is obtained from that for t_n, by replacing t by τ and d by δ.

Inductively, we assume that $d_i = \delta_i$ for $i \leqslant n$, and that $t_i = \tau_i$ for $i < n$. As a subsidiary noetherian induction, we assume that $t_n[e'u'] = \tau_n[e'u']$ for any term $e'u'$ with $W[(u_1,\ldots,u_n)u] \gg W[e'u']$, and we want to prove that $t_n[(u_1,\ldots,u_n)u] = \tau_n[(u_1,\ldots,u_n)u]$. By the discussion of the previous paragraphs, we may assume that u_nu is reducible, write $u = vw$, and we need only show that in $[\tau_{n-1}[\delta_n(u_1,\ldots,u_n)v]]w$ any term $e'u'$ has $W[(u_1,\ldots,u_n)u] \gg W[e'u']$.

Now, for any essential cell e'' and any $m \in M$, we plainly have $W[e''m] = W[\sigma[e''m]]$. By the discussion in Subsection 6.3, it follows that for any term $e'u'$ in $\tau[e''m]$ we have $W[e''m] \gg W[e'u')]$ or $W[e''m] = W[e'u']$. Because u_nv is reducible, we know that $W[(u_1,\ldots,u_n)v] \gg W[e''m]$ for any term $e''m$ in $\delta_n[(u_1,\ldots,u_n)v]$. Hence, by the previous paragraph, in $\tau_{n-1}[\delta_n(u_1,\ldots,u_n)v]$ any term $e'u'$ has $W[(u_1,\ldots,u_n)v] \gg W[e'u']$. Since the relation $\gg$ is preserved by multiplication on the right, we have the required property.

Thus we have identified Kobayashi's boundary operators and contracting homotopy with ours.

Kobayashi states his result in a more general from, involving a second rewriting system $\mathscr{S}$, for which we can rewrite lv to rv, but cannot rewrite ulv to urv if $u \neq 1$. Brown's method generalises to this situation. We begin by changing the definition of ∂_{n0} in the unnormalised bar resolution to $\partial_{n0}(u_1,\ldots,u_n) = (\hat{u}_1u_2,\ldots,u_n)$, where $\hat{u}_1$ is the normal form of u_1, leaving ∂_{ni} unchanged for $n > 0$. We then define a cell to be degenerate if either u_1 is irreducible or $u_i = 1$ for some $i > 1$, and get a normalised resolution (in a slightly more general sense) by factoring out the degenerate cells. The details of the collapsing scheme are left to the reader.

8.2. *Anick's method*

Anick's account is both more and less general than ours. It is more general because he works with an associative algebra rather than a monoid; we shall not make this generalisation. It is less general, because on the free monoid he takes a well-ordering compatible with multiplication and assumes that any $(l,r) \in \mathscr{R}$ has r less than l. It is possible to modify his approach, using a noetherian partial ordering as before, and using noetherian induction rather than transfinite induction. Since we already have a better proof, we shall not do this. However, we will show that his resolution is also the same as ours.

Anick's definition of the elements that form a basis for the chains is very different from

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

ours, and the main work is in showing that there is a natural bijection between his basis and the essential cells. We will leave this till later, and for the moment assume that they are the same.

We denote his boundary operator in dimension n (but with the essential cells as basis using the bijection) by β_n. He does not define a contracting homotopy, but only defines a homomorphism of abelian groups α_n from $\ker \beta_n$ to the group of $(n+1)$-chains such that $\beta_{n+1}\alpha_n c = c$ when $\beta_n c = 0$.

With the notation of the previous subsection, he defines $\beta_1 = d_1$ and $\alpha_0(u-1) = t_0 u$. Thus we have $\beta_1 = \delta_1$ and $\alpha_0 = \tau_0$ on $\ker \epsilon$.

Suppose we have defined β_i for $i \leqslant n$ and α_i for $i < n$, and that $\beta_i = \delta_i$ for $i \leqslant n$ and $\alpha_i = \tau_i$ on $\ker \beta_i$ for $i < n$.

He defines

$$\beta_{n+1}(u_1, \ldots, u_{n+1}) = (u_1, \ldots, u_n)u_{n+1} - \alpha_{n-1}[\beta_n(u_1, \ldots, u_n)u_{n+1}].$$

As in Kobayshi's approach, we find that $\beta_{n+1} = \delta_{n+1}$.

Before defining α_n we need a lemma.

Lemma 9. Let $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ be essential cells, and let u and v be in M. If $W[(u_1, \ldots, u_n)u] = W[(v_1, \ldots, v_n)v]$, then $u_i = v_i$ for all i and $u = v$.

Proof. Since $u_1 \ldots u_n u = v_1 \ldots v_n v$ and both u_1 and v_1 are in the generating set X, we have $u_1 = v_1$. Suppose we have $u_i = v_i$ for $i \leqslant r$ and let $r+1 \leqslant n$. Since $u_r u_{r+1}$ and $v_r v_{r+1}$ are prefixes of the same word, either one is a prefix of the other or they are equal. By the definition of an essential cell, both $u_r u_{r+1}$ and $v_r v_{r+1}$ are reducible, but neither has a reducible prefix. Thus they must be equal. As $u_r = v_r$, we see that $u_{r+1} = v_{r+1}$. So, inductively, $u_i = v_i$ for all i, and then $u = v$. □

We define a quasi-ordering of essential terms that may have different dimensions by saying that the essential term eu is higher than the essential term $e'u'$ if $W[eu]$ is higher than $W[e'u']$. From the lemma, this is a well-ordering on essential terms of a fixed dimension. From the way the well-ordering is defined, eu is higher in this well-ordering than $e'u'$ if $eu \gg e'u'$.

Now consider an essential n-chain c with $\delta_n c = 0$. Let $(u_1, \ldots, u_n)u$ be its highest term, so that $c = k(u_1, \ldots, u_n)u + c'$, where $k \in \mathbf{Z}$ and the terms in c' are lower than $(u_1, \ldots, u_n)u$. Then, by our analysis in Subsection 6.3, all the terms in $\delta_n c'$ will be lower than $(u_1, \ldots, u_n)u$, as will all but one of the terms in $\delta_n(u_1, \ldots, u_n)u$. If $u_n u$ is irreducible, the remaining term is $(u_1, \ldots, u_{n-1})U$ with $U = u_n u$, and $W[(u_1, \ldots, u_{n-1})U] = W[(u_1, \ldots, u_n)u]$, so this term is higher than all the other terms of $\delta_n c$. This is impossible, because $\delta_n c = 0$.

Thus $u_n u$ is reducible, and, as usual, we write $u = vw$. We have $\beta_{n+1}(u_1, \ldots, u_n, v)w = \delta_{n+1}(u_1, \ldots, u_n, v)w = (u_1, \ldots, u_n)u + c'$, where, as before, all terms in c' are lower than $(u_1, \ldots, u_n, v)w$ and so lower than $(u_1, \ldots, u_n)u$. It follows that $c - k\beta_{n+1}(u_1, \ldots, u_n, v)w$ is a chain in $\ker \beta_n$, all of whose terms are lower than $(u_1, \ldots, u_n)u$; we denote it by c''.

Anick now defines

$$\alpha_n c = k(u_1, \ldots, u_n, v)w + \alpha_n c'',$$

a definition by transfinite induction.

Since $\beta_{n+1} = \delta_{n+1}$ and, as shown in the previous subsection, $\tau_n \delta_{n+1}[(u_1, \ldots, u_n, v)w] =$

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

$(u_1, \ldots, u_n, v)w$, we also have

$$\tau_n c = k(u_1, \ldots, u_n, v)w + \tau_n c''.$$

Thus we have proved by transfinite induction that $\alpha_n = \tau_n$ on $ker\beta_n$. We now proceed to give Anick's account of the basis elements. We shall define *n-overlaps* and *leftmost n-overlaps*. (These correspond to Anick's notion of *n-prechains* and *n-chains*. His names do not indicate what the objects are; also, we need the word *chain* to refer to the elements of our complex.) His resolution has as basis in dimension n the leftmost $(n-1)$-overlaps.

Let L be the set of left-hand sides of elements of $\mathcal{R}$. When $l_i \in L$, we denote by r_i the word such that $(l_i, r_i) \in \mathcal{R}$. We call the elements of X *0-overlaps* and the elements of L *1-overlaps*. Every member of X or L is to be a leftmost overlap.

An *n-overlap* is defined to be a word $x_1 \ldots x_k$ for some k and some $x_i \in X$, together with integers i_r and j_r for $1 \leqslant r \leqslant n$ such that

$$1 = i_1 < i_2 \leqslant j_1 < i_3 \leqslant j_2 < \ldots < i_n \leqslant j_{n-1} < j_n = k,$$

and $x_{i_r} \ldots x_{j_r} \in L$ for $1 \leqslant r \leqslant n$. Less formally, an n-overlap consists of a word w and words $l_1, \ldots, l_n \in L$ such that

— l_1 is a prefix of w, l_n is a suffix of w, and l_r is a subword of w for $1 < r < n$,

— for $i < n$ the subword l_{i+1} of w overlaps the subword l_i and starts later than the start of l_i (it will also end later than the end of l_i, because $\mathcal{R}$ is reduced; conversely, any subword that is in L and ends later than the end of l_i must start later than the start of l_i),

— the subword l_{i+2} does not overlap the subword l_i.

A 2-overlap is often referred to as an *overlap ambiguity*, which gives rise to a critical pair (since $\mathcal{R}$ is reduced, there are no inclusion ambiguities).

We sometimes refer to the word w as an n-overlap, without explicit mention of the corresponding integers. Strictly speaking, this is incorrect, as there may be more than one way of assigning the integers so as to get an n-overlap. Because $\mathcal{R}$ is reduced, this cannot happen if $n = 2$. But suppose that $\{abc, bcd, cde, def\} \subseteq L$. Then the word *abcdef* is a 3-overlap in two ways. one using the subwords *abc*, *cde*, and *def*, and the other using *abc*, *bcd*, and *def*. Note that the word *abcde*, in which the subwords *abc* and *bcd* overlap and the subwords *bcd* and *cde* also overlap, is not a 3-overlap, as the subwords *abc* and *cde* overlap, which is forbidden by our definition. In this example, the second expression as an overlap is further to the left than the first, in an obvious sense. More precisely, the first expression gives rise to a 2-overlap of *abc* and *cde* in the word *abcde*, and the second expression gives rise to a 2-overlap of *abc* and *bcd* in the word *abcd*; the second word is a prefix of the first. This explains the notion of a leftmost overlap, given in the next paragraph.

To each n-overlap and $m < n$ there is a corresponding m-overlap given by the word $x_1 \ldots x_{j_m}$ and the integers i_r and j_r for $r \leqslant m$. We say that an n-overlap is *leftmost* if, for all $m \leqslant n$ and all $r < j_m$ the word $x_1 \ldots x_r$ is not an m-overlap.

It is easy to see that, although some words may be n-overlaps in more than one way, a word can be a leftmost n-overlap in at most one way. For suppose it had two such

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

expressions, corresponding to integers i_r and j_r for one and i'_r and j'_r for the other. Because each corresponding m-overlap is leftmost for each expression, we must have $j'_m = j_m$ for all m. Because $\mathscr{R}$ is reduced, this requires that we also have $i'_m = i_m$.

Note that a word that can be expressed as an n-overlap may not have an expression as a leftmost n-overlap. For instance, let $L = \{abc, bcd, cde, efg\}$. Then the word $abcdefg$ is expressed as a 3-overlap using the subwords abc, cde and efg. The corresponding 2-overlap is $abcde$, which has the 2-overlap $abcd$ as prefix. So this expression is not leftmost. But if we try to find an expression that is further to the left, we must use the subwords abc, bcd and efg, which express our word as the product of a 2-overlap and a 1-overlap, not as a 3-overlap.

Anick constructs his resolution using the leftmost $(n-1)$-overlaps as the basis for the chains in dimension n. We shall show how to obtain a bijection between essential n-cells and leftmost $(n-1)$-overlaps.

The leftmost 0-overlaps are the elements of X, and the essential cell corresponding to x is (x). The leftmost 1-overlaps are the elements of L. Since $\mathscr{R}$ is strongly reduced, any $l \in L$ can be written $l = xu$ for some non-empty word u, and the corresponding essential 2-cell is (x, u).

Now let $x_1 \ldots x_k$ be a word that is a leftmost $(n-1)$-overlap, for $n > 1$. We define words $u_1, \ldots, u_n$ inductively, by setting $u_1 = x_1$ and $u_1 \ldots u_{r+1} = x_1 \ldots x_{j_r}$ for $0 < r < n$. By definition of an n-overlap, $u_m u_{m+1}$ is reducible for $0 < m < n$. Because our overlap is leftmost, no prefix of $u_m u_{m+1}$ is reducible. Hence the n-cell $(u_1, \ldots, u_n)$ is essential.

Conversely, take an essential n-cell $(u_1, \ldots, u_n)$, and write $u_1 \ldots u_n = x_1 \ldots x_k$ with $x_i \in X$ for all i. Define j_r for $0 < r < n$ by $u_1 \ldots u_{r+1} = x_1 \ldots x_{j_r}$. Because $u_r u_{r+1}$ is reducible but no prefix of it is reducible, there will be a suffix of $u_r u_{r+1}$ in L; this suffix will be unique, since $\mathscr{R}$ is reduced. We define i_r by requiring $x_{i_r} \ldots x_{j_r} \in L$. This expresses $x_1 \ldots x_k$ as an $(n-1)$-overlap. If it is not leftmost, take the smallest value of m such that the corresponding m-overlap has a prefix that is also an m-overlap. We would then find that $u_m u_{m+1}$ had a reducible prefix, contrary to hypothesis.

So we have the required bijection, and we have identified Anick's resolution with ours. As already remarked, the boundary operator is much easier to calculate using our approach. However, in some cases the structure of the overlaps is more transparent than that of the essential cells, so Anick's method may be better if we do not need to compute the boundary operators.

8.3. *Squier's method*

We conclude the detailed comparison by considering Squier's method. Squier works with left modules, while we use right ones. As already remarked, we can use either, and we will look at the right module version of Squier's result. We begin by defining a function $D_x : X^* \to \mathbf{Z}X^*$ for each $x \in X$. Let $u = x_1 \ldots x_n$. Then we define $D_x u = \sum x_{r+1} \ldots x_n$, the sum being taken over those r for which $x_r = x$. For example, if $u = xyxyxz$, then $D_x u = yxyxz + yxz + z$, $D_y u = xyxz + xz$, $D_z u = 1$. We extend D_x to $\mathbf{Z}X^*$ linearly. It is easy to see that, for any $u, v \in X^*$, we have $D_x[uv] = [D_x u]v + D_x v$ and $\sum [x-1] D_x u = u - 1$.

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

From our previous description, we see that $\tau_0 u = \sum(x)D_x u$. We already know that $\phi_1(u) = \tau_0 u$.

For any $w \in X^*$, let πw be the irreducible word equivalent to w; we extend π linearly to $\mathbf{Z}X^*$.

Let $l \in L$. For any u and v, we have $D_x[ulv - urv] = [D_x u][lv - rv] + [D_x l - D_x r]v$, so $\pi D_x[ulv - urv] = \pi[[D_x l - D_x r]v]$.

Write l as xu. Then there is a corresponding essential 2-cell (x, u). To focus attention on l, we will also denote this cell by $\langle l \rangle$.

Now $\partial_2(x, u) = (x)u + (u) - (\pi[xu])$. From our discussion of ϕ_1 and τ_0, we have $\delta_2\langle l \rangle = \sum(x)D_x[l - r]$. This identifies our δ_2 with Squier's. Note that we also have $\delta_2[\langle l \rangle w] = \sum(x)\pi D_x[[l - r]w]$; the formula of the previous sentence does not need π because the elements involved are irreducible. Furthermore, we have already seen that $\pi D_x[[l - r]w] = \pi[[D_x l - D_x r]w]$. For each $w \in X^*$, choose (arbitrarily) a sequence $w = w_1 \to w_2 \ldots \to w_n = \pi w$. For $i < n$, let $w_i = u_i l_i v_i$ and $w_{i+1} = u_i r_i v_i$, where $l_i \in L$ with corresponding right-hand side r_i. Denote this sequence by F, and define $\langle F \rangle = \sum\langle l_i \rangle v_i$. From our previous remarks, we see that $\delta_2\langle F \rangle = \sum(x)\pi D_x[w - \pi w]$.

In particular, if we always make the rightmost reduction possible, the resulting function will be denoted by ρ. Note that $\rho w = \langle l_1 \rangle v_1 + \rho w_2$. Also, $\rho[ww'] = \rho w' + \rho[wz]$, where $z = \pi w'$.

We now show that $\phi_2(u, v) = \rho[uv]$ for any 2-cell (u, v). To do this, we use noetherian induction on $>_2$, defined as in Subsection 6.3. First suppose that $u = xt$ for some $x \in X$ and non-empty t. Then (u, v) is redundant, and, looking at the boundary of the corresponding collapsible cell, we have $\phi_2(u, v) = \phi_2(t, v) + \phi_2(x, \pi[tv]) = \rho[tv] + \rho[x[\pi[tv]]$, inductively, and this equals $\rho[xtv]$, as required.

Next suppose that $u = x$. If xv is irreducible, then (x, v) is collapsible, so $\phi_2(x, v) = 0$, and $\rho[xv] = 0$. Otherwise we can write $v = pq$ with $xp = l \in L$. Note that xp is the only subword of xpq that is in L. Then $\pi[xp] = r$. We have $\phi_2(x, v) = \langle l \rangle q + \phi_2(r, q) = \langle l \rangle q + \rho[rq]$, inductively, and this equals $\rho[xpq]$, as required.

We see that $\tau_1[(x)u] = \rho[xu]$, since $\sigma_1[(x)u] = (x, u)$.

In dimension 3, Squier's module differs from ours. Before discussing it, we look at some results on 3-cells. Consider a 3-cell (u, v, w) with uv irreducible. If $u \in X$, this cell is collapsible, and $\phi_3(u, v, w) = 0$. Otherwise this cell is redundant, and, looking at the boundary of the corresponding collapsible 4-cell, we easily see that $\phi_3(u, v, w) = 0$, by induction on the length of u. From this we then see that if $uv \in L$ and $u = xt$, we have $\phi_3(u, v, w) = \phi_3(x, tv, w)$.

Squier takes a module S with basis the cells (u, v, w), where $uv = l_1$ and $vw = l_2$ are in L. We consider the boundary $\delta_3 = \phi_2\partial_3$ on S. We know that $\delta_2\delta_3 = 0$, and that $\phi_2\partial_3\phi_3 = \phi_2\partial_3$. So, by the previous paragraph, $\delta_3 S = \delta_3 S'$, where S' has as basis the cells (x, tv, w), with u, v, w as before and $u = xt$. These cells include all essential cells, so we have $\ker \delta_2 \subseteq \delta_3 S'$, as required.

Now $\partial_3(u, v, w) = (u, v)w - (u, r_2) + (r_1, w) - (v, w)$, by definition. Using the connection between ϕ_2 and ρ previously found, and noting that $\rho[uv] = \langle l_1 \rangle$ and $\rho[vw] = \langle l_2 \rangle$, we

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

obtain

$$\delta_3(u,v,w) = \langle l_1\rangle w - \langle l_2\rangle + \rho[r_1 w] - \rho[ur_2].$$

Allowing for the fact that we are using right modules and Squier is using left ones, this is exactly Squier's definition of his boundary operator. Thus Squier's module in dimension 3 is larger than it needs to be, which leads to the difficulties in extending the resolution further. Comparing with Anick's approach, we find that the the bijection between essential 3-cells and leftmost 2-overlaps extends to a bijection between the cells of S' and all 2-overlaps.

Squier extends his theory further by using the function α obtained by using an arbitrary choice of reductions, rather than ρ, which was obtained using the rightmost possible reduction at each step. That is, taking the same set S, he defines

$$\Delta_3(u,v,w) = \langle l_1\rangle w - \langle l_2\rangle + \alpha[r_1 w] - \alpha[ur_2].$$

He then proves that $\ker \delta_2 = \Delta_3 S$.

We have

$$\Delta_3(u,v,w) - \delta_3(u,v,w) = \alpha[r_1 w] - \rho[r_1 w] - \alpha[ur_2] + \rho[ur_2].$$

So $\delta_2\Delta_3 = \delta_2\delta_3 = 0$ if $\delta_2\alpha z = \delta_2\rho z$ for all z. But we have already remarked that both these expressions equal $\sum(x)D_x[z - \pi z]$. If we show that $\alpha z - \rho z \in \Delta_3 S$ for any z, it will be immediate that $\delta_3(u,v,w) \in \Delta_3 S$. The next lemma proves a stronger property. We first need to obtain a simple formula. Let F be a sequence of rewritings from w to πw, and let u and v be any words. Let H be a sequence of rewritings from $u[\pi w]v$ to $\pi[uwv]$. Let uF denote the obvious sequence from uwv to $u[\pi w]v$, and let G consist of uFv followed by H. It is easy to check (noting that the product in M of πv_i and πv is $\pi[v_i v]$) that $\langle G\rangle = \langle F\rangle \pi v + \langle H\rangle$.

Lemma 10. Let F and G be sequences of rewritings from a word z to πz. Then $\langle F\rangle - \langle G\rangle \in \Delta_3 S$.

Proof. We prove the result by noetherian induction.

Let F_1 and G_1 be the sequences obtained from F and G by omitting the first rewriting. If the first rewriting in F is the same as the first in G, then $\langle F\rangle - \langle G\rangle = \langle F_1\rangle - \langle G_1\rangle$, and the result holds by induction.

Otherwise we have the usual two possibilities. The first is that $z = al_1 bl_2 c$, where the first rewriting of F replaces l_1 by r_1, and the first rewriting of G replaces l_2 by r_2. Let H be any sequence of rewritings from $ar_1 br_2 c$ to $\pi[ar_1 br_2 c]$. Then $\langle F\rangle = \langle l_1\rangle \pi[bl_2 c] + \langle F_1\rangle$. Inductively, $\langle l_2\rangle \pi c + \langle H\rangle - \langle F_1\rangle \in \Delta_3 S$. Using the similar formulae concerning G and G_1, the result is immediate.

The second case is that $z = auvwb$ with $uv = l_1$ and $vw = l_2$; we will have u, v, w irreducible and non-empty, and so $w = \pi w$. Then $\langle F\rangle = \langle l_1\rangle \pi[wb] + \langle F_1\rangle$. Let A be that sequence of rewritings from $r_1 w$ to $\pi[r_1 w] = \pi[uvw]$ that defines $\alpha[r_1 w]$, and let H be any sequence of rewritings from $a\pi[uvw]b$ to $\pi[auvwb]$. Then we have the sequence aAb followed by H from $ar_1 wb$ and the sequence F_1 from the same word. Inductively, using the formula before the lemma, $\alpha[r_1 w]\pi b + \langle H\rangle - \langle F_1\rangle \in \Delta_3 S$. Similarly, $\langle G\rangle = \langle l_2\rangle \pi b + \langle G_1\rangle$, and $\alpha[ur_2]\pi b + \langle H\rangle - \langle G_1\rangle \in \Delta_3 S$. So it is enough to show that $\langle l_1\rangle\pi[wb] + \alpha[r_1w]\pi b - \langle l_2\rangle\pi b - \alpha[ur_2]\pi b \in \Delta_3 S$. But this is precisely $\Delta_3(u,v,w)\pi b$. □

A closer look at the proof shows that we have $\langle F\rangle - \langle G\rangle = \Delta_3 q$, where any term t in the chain q has either $z \gg W[t]$ or $z = W[t]$. The same will hold for δ_3, which is just a special case of Δ_3. It then follows that for any chain p there is a chain q such that $\Delta_3 p = \delta_3[p+q]$ and $p+q \neq 0$ if $q \neq 0$. In particular, Δ_3 has trivial kernel if δ_3 has trivial kernel.

Here we are looking at δ_3 as a homomorphism from S. If (in the notation used in the previous section) all overlaps are leftmost, ϕ_3 is plainly a bijection from the basis of S to the set of essential 3-cells. Since $\delta_3\phi_3 = \delta_3$, we see that δ_3 has trivial kernel on S if it has trivial kernel on the set of essential 3-chains. By Brown's Theorem, this will certainly hold if there are no 3-overlaps. We have therefore shown the following result, which is exactly Theorem 3.2 of Squier (1987).

Proposition 3. Suppose that all 2-overlaps are leftmost and that there are no 3-overlaps. Then $\ker \Delta_3 = 0$.

8.4. *Groves' method*

Groves (1991) gives yet another proof of Anick's theorem. This differs in significant details from the approaches we have discussed, because it is based on cubes rather than simplexes (which, for some situations such as free and direct products of groups, is more convenient). Nevertheless, readers who have followed our approach this far will probably feel, on looking at Groves's work, that it can be fitted into our account of collapsing schemes. The notations and concepts he uses seem to me rather complicated, and I have not been able to simplify them enough to give a detailed account. I therefore make just a few comments on his work.

Groves starts with our usual strongly minimal convergent rewriting system $\langle X;\mathscr{R}\rangle$, presenting the monoid M. He then defines a directed graph Γ, with vertex set X^* and with one edge starting at u for each application of a rewrite rule to u. He then defines what he calls *cubes* and *stars* of Γ, and obtains a resolution whose basis consists of certain stars.

We note that Γ is disconnected, with one component for each element of M. Also, each component has a terminal vertex; that is, a vertex that can be reached from any other vertex in the component by a directed path. These vertices are precisely the irreducible words. Because of this, it is easy to construct a resolution

$$\dots C_2 \to C_1 \to C_0 \to \mathbf{Z}M \to 0$$

of the abelian group $\mathbf{Z}M$ by free abelian groups (note that we do not define a module structure yet). The basis of C_n will consist of all n-cubes except that certain cubes are equated to zero.

In the low dimensions it is not too difficult to see what we want the redundant and collapsible cells to be, and we should be able to extend this to a collapsing scheme. It is then necessary to factor out further subgroups, as indicated by Groves, in order to get an action of M. The boundary operators map each such subgroup into the next one down,

https://doi.org/10.1017/S0960129596002149 Published online by Cambridge University Press

and so we get a complex. This will satisfy the condition for a resolution, except in low dimensions, because the contracting homotopies will map each subgroup in the next one up. In low dimensions, we need to make explicit definitions slightly different from the original ones, and check that the required properties hold.

References

Abels, H. (1979) An example of a finitely presented solvable group. In: Wall, C. J. C. (ed.) Homological group theory. *London Mathematical Society Lecture Notes* **36**, Cambridge University Press.

Alonso, J. (1992) Combings of groups. In: Baumslag, G., Cannonito F. and Miller, C. F. (eds.) Proceedings of the workshop on algorithms, word problems, and classification in combinatorial group theory. *MSRI Publications* **23**, Springer-Verlag 165–178.

Anick, D. J. (1986) On the homology of associative algebras. *Transactions of the American Mathematical Society* **296** 641–659.

Baumslag, G. and Bieri, R. (1976) Constructible solvable groups. *Mathematische Zeitschrift* **151** 249–257.

Bieri, R. (1976) *Homological dimension of discrete groups*, Queen Mary College Mathematics Notes, Queen Mary College, London.

Bieri, R. (1980) A connection between the integral homology and the centre of a rational linear group. *Mathematische Zeitschrift* **170** 263–266.

Book, R. V. and Otto, F. (1993) *String-Rewriting Systems*, Springer-Verlag.

Brown, K. S. (1984) Presentations for groups acting on simply-connected complexes. *Journal of Pure and Applied Algebra* **32** 1–10.

Brown, K. S. (1987) Finiteness properties of groups. *Journal of Pure and Applied Algebra* **44** 45–75.

Brown, K. S. (1992a) The geometry of finitely presented infinite simple groups. In: Baumslag, G., Cannonito F. and Miller, C. F. (eds.) Proceedings of the workshop on algorithms, word problems, and classification in combinatorial group theory. *MSRI Publications* **23**, Springer-Verlag 121–136.

Brown, K. S. (1992b) The geometry of rewriting systems; a proof of the Anick–Groves–Squier theorem. In: Baumslag, G., Cannonito F. and Miller, C. F. (eds.) Proceedings of the workshop on algorithms, word problems, and classification in combinatorial group theory. *MSRI Publications* **23**, Springer-Verlag 137–164.

Brown, K. S. and Geoghegan, R. (1984) An infinite-dimensional torsion-free FP_∞ group. *Inventiones Mathematicae* **77** 367–381.

Cohen, D. E. (1992) A monoid which is right FP_∞ but not left FP_1. *Bulletin of the London Mathematical Society* **24** 340–342.

Cohen, D. E. (1993) String rewriting; a survey for group theorists. In: Niblo, G. and Roller, M. (eds.) Proceedings of the conference on Geometric Group Theory, Isle of Thorns 1991. *London Mathematical Society Lecture Notes* **181**, Cambridge University Press.

Cohen, D. E. (1995) Projective resolutions for graph products. *Proceedings of the Edinburgh Mathematical Society* **38** 185–188.

Coxeter, H. S. M. and Moser, W. O. J. (1958) Generators and relations for discrete groups. *Ergebnisse Mathematik* **14**, Springer-Verlag.

Cremans, R. and Otto, F. (1994) For groups the property of having finite derivation type is equivalent to the homological finiteness condition FP_3, Mathematische Schrift Kassel, preprint 8/94.

Epstein, D. B. A., Cannon, J. W., Holt, D. F., Levy, S. V. F., Paterson, M. S. and Thurston, W. P. (1992) *Word Processing in Groups*, Jones and Bartlett, Boston.

Farkas, D. R. (1992) The Anick resolution. *Journal of Pure and Applied Algebra* **79** 159–168.

Groves, J. R. J. (1991) Rewriting systems and homology of groups. In: Kovacs, L. G. (ed.) Groups – Canberra 1989. *Springer-Verlag lecture notes in mathematics* **1456** 114–141.

Groves, J. R. J. and Smith, G. S. (1989) Rewriting systems and soluble groups, Bath computer science technical report 89-19.

Groves, J. R. J. and Smith, G. S. (1993) Soluble groups with a finite rewriting system. *Proceedings of the Edinburgh Mathematical Society* **36** 283–288.

Guba, V. S. and Sapir, M. V. (1995) The Dehn function and a regular set of normal forms for R. Thompson's group F, preprint.

Hermiller, S. and Meier, J. (1994a) Algorithms and geometry for graph products of groups, to appear.

Hermiller, S. and Meier, J. (1994b) Tame combings, almost convexity and rewriting systems for groups, to appear.

Hilton, P. J. and Stammbach, U. (1971) *A course in homological algebra*, Springer-Verlag.

Jantzen, M. (1988) Confluent string rewriting. *European association for theoretical computer science* **14**, Springer-Verlag.

Kapur, D. and Narendran, P. (1985) A finite Thue system with a decidable word problem and without an equivalent finite canonical system. *Theoretical Computer Science* **35** 337–344.

Knuth, D. and Bendix, P. (1970) Simple word problems in universal algebra. In: Leech, J. (ed.) *Computational problems in abstract algebra*, Pergamon, Oxford.

Kobayashi, Y. (1990) Complete rewriting systems and homology of monoid algebras. *Journal of Pure and Applied Algebra* **65** 263–275.

Lafont, Y. (1994) A new finiteness condition for monoids presented by complete rewriting systems (after Craig C. Squier). *Journal of Pure and Applied Algebra* (to appear).

Lafont, Y. and Prouté, A. (1991) Church–Rosser property and homology of monoids. *Mathematical structures in computer science* **1** 297–326.

Lyndon, R. C. (1950) Cohomology theory of groups with a single defining relation. *Annals of Mathematics* **52** 650–665.

Mac Lane, S. (1963) *Homology*, Springer-Verlag.

Newman, M. H. A. (1943) On theories with a combinatorial definition of equivalence. *Annals of Mathematics* **43** 223–243.

Squier, C. C. (1987) Word problems and a homological finiteness condition for monoids. *Journal of Pure and Applied Algebra* **49** 201–217.

Squier, C. C., Otto, F. and Kobayashi, Y. (1994) A finiteness condition for rewriting systems. *Theoretical Computer Science* **131** 271–294.