Abstract

Given a finite presentation of a group $G$, proving properties of $G$ can be difficult. Indeed, many questions about finitely presented groups are unsolvable in general. Algorithms exist for answering some questions while for other questions algorithms exist for verifying the truth of positive answers. An important tool in this regard is the Todd-Coxeter coset enumeration procedure. It is possible to extract formal proofs from the internal working of coset enumerations. We give examples of how this works, and show how the proofs produced can be mechanically verified and how they can be converted to alternative forms. We discuss these automatically produced proofs in terms of their size and the insights they offer. We compare them to hand proofs and to the simplest possible proofs. We point out that this technique has been used to help solve a longstanding conjecture about an infinite class of finitely presented groups.

Keywords: finitely presented group, proof, Todd-Coxeter coset enumeration, Hilbert’s 24th problem, Fibonacci group, van Kampen diagram, trivial group, Andrews-Curtis conjecture.

1 Introduction

Many theorems in group theory are based on the results of computations. Indeed, often, the computations are now done on machines. Comprehensive details about computing with finitely presented groups appear in [31, 16]. Of particular relevance to our considerations are the chapters on coset enumeration and also a brief review of computability issues.

Proofs of theorems which include machine computations are opaque. This is especially true for proofs which rely on the collapse of a coset enumeration to one coset. Indeed, one of the most likely consequences of an error in a hand computation or of a bug in a computer application of coset enumeration is an incorrect total collapse. Leech [21] discusses this difficulty.

We address this lack of transparency of proofs based on direct coset enumeration by studying them carefully and deriving more usual proofs. We compare and contrast various machine and human proofs. As regards coset enumeration, we focus on the least transparent situation, where total collapse occurs.

---

1 Both authors were partially supported by the Australian Research Council
Questions regarding the length and elegance of proofs are difficult in general. Thiele [32] tells us that Hilbert considered including another problem in his famous list.

The twenty-fourth problem belongs to the realm of foundations of mathematics. In a nutshell, it asks for the simplest proof of any theorem.

Grattan-Guinness [9] suspects it was ultimately omitted since “simplicity is an extremely complicated notion to capture in a general way”.

In the context of a limited range of proofs in finitely presented groups we can address this problem with reasonable success. We introduce the notion of a proof certificate for some results based on coset enumeration. This enables us to classify proofs with respect to sensible criteria.

Finally, the Andrews-Curtis [1, 2] conjecture addresses the nature of proofs of triviality for balanced presentations and has been studied computationally [12, 13]. Work on this conjecture provides alternative proof methods. As a result we are able to guarantee some proofs are simplest in specific terms.

We give examples of how our ideas work by investigating various proofs of results about some interesting finitely presented groups. Throughout this paper, we use the convention that lower case letters denote group generators and upper case letters their inverses; thus, $A = a^{-1}$, etc. Coset number 1 will always represent the subgroup and all cosets are right cosets.

2 Coset enumeration

Coset enumeration is the basis of important techniques for investigating finitely presented groups with manifold applications. Its use dates back at least to Moore [25]. Given defining relations for a group $G$ and words generating a subgroup $H$ of finite index, coset enumeration programs implement systematic procedures for enumerating the cosets of $H$ in $G$. Computer implementations are based on methods initially described by Todd and Coxeter [34]. Many details are given in [31, 16]. Implementations are available in the computer algebra systems GAP [8] and Magma [3] and as a stand-alone program, the ACE coset enumerator, [12]. A particularly useful tool for experimenting with coset enumerations is the Interactive Todd Coxeter package, ITC [7].

The end result of a successful coset enumeration is a coset table which gives a permutation representation for $G$, corresponding to the action by multiplication of $G$ on the cosets of $H$. When the index is not one we can readily check that the nontrivial representation obtained is indeed a representation for the group, which gives significant comfort that the process worked correctly. However, when the index is one, nothing reassuring is left except for our confidence in the process.

Implicit in the underlying working of an enumeration are formal proofs that particular words in the generators are in the subgroup, as shown by Leech [22]. The utility PEACE [28] (proof extraction after coset enumeration) has been developed to automate the production of such proofs. PEACE produces proof-words which can be regarded as certificates attesting to subgroup membership.
The natural measure of simplicity for a coset enumeration based proof is the total number of cosets defined in the process. Once we have a proof-word, then natural measures are based on its length.

3 Proof certificates

A successful coset enumeration, in its workings, embodies proofs of many embedded results. However its very mechanical nature makes it both unpalatable to read and also error-prone. In spite of this, very many computer implementations are regularly used to correctly build coset tables for all sorts of purposes.

Our focus here is on proving theorems based upon coset enumeration. Given a successful coset enumeration of the cosets of the subgroup $H$ in the group $G$, we may claim that this proves that some word $\omega$ is in $H$. (The fact that $\omega \in H$ is checked by applying coset 1 to $\omega$ and tracing it through the coset table back to coset 1.) This claim rests on the validity of the coset enumeration strategy employed and the correctness of its implementation.

Instead of making such a claim on this basis we can provide a stand-alone proof, as follows. We extract from the workings of a coset enumeration a proof-word which can be verified mechanically and which explicitely gives $\omega$ in terms of the generators of $H$. This proof-word, together with the presentation for $G$ and the generators of $H$, forms a (proof) certificate that $\omega \in H$. The validity of this certificate is independent of how it was generated, and does not depend on anything other than the group axioms and the definitions of $G$ and $H$. Such proof-words are implicit in the workings of a coset enumeration, and their generation by considering circuits in the Schreier diagram is explained by Leech [22]. Many further details are given in [28], where the PEACE package is documented.

A fully expanded proof-word consists of a product of subgroup generators and of conjugates of group relators (by group generators). The subgroup generators appear as given in the presentation for $H$, or as the formal inverses thereof. There is no such requirement on relators for $G$, and they or their formal inverses may be cycled in proofs. By construction, $\omega$ and the proof-word are equivalent and, since conjugates of relators are trivial in $G$, the proof-word is also equivalent to a product of subgroup generators. Thus, free reduction of the proof word produces $\omega$, while reduction after cancelling the conjugates of relators produces a product of subgroup generators. Such reductions are easy to verify by hand for short proof-words and by computer for longer words.

Although the content of a proof-word depends on the details of how the coset enumeration is performed, it stands alone as a proof. Its validity depends only on the presentation used for the group, the generators of the subgroup, the word to be proved, and the general axioms of group theory. It is easy to understand and, in this sense, provides a simple proof that a word is in a subgroup (and gives an explicit construction). The PEACE package includes a stand-alone verification utility, which both checks that a proof-word is properly formed and performs the two required reductions.
4 Pruned enumeration

Starting with a successful coset enumeration where the total number of cosets used, \( T \), exceeds the subgroup index, \( I \), it is often possible to prune the sequence of \( T - 1 \) coset definitions by eliminating cosets which do not contribute to the final table. (It would be nice to reduce \( T \) to \( I \) but, in general, this is not possible.) Early workers pruned enumerations by hand, but ITC automates the process by repeatedly reordering the sequence of definitions and rerunning the enumeration, while omitting definitions which become redundant. Such pruning techniques have been incorporated into PEACE and, combined with the ability to consider equivalent presentations (that is, presentations where the relators have been cycled, inverted and/or reordered), have proved very effective at generating short enumerations.

The pruned enumerations that ITC and PEACE yield are frequently the shortest known. To prove that no shorter enumeration exists an exhaustive search technique can be used, although this is only feasible for small values of \( T \). PEACE has the ability to generate and test all possible valid sequences of definitions of a given length, in a standardised form; e.g., ordered by coset number and coset-table column number. See §6.2 for more details, and an example.

We are now ready to consider some proofs. We focus on four examples which use coset enumerations that collapse to one coset. This is the case where the proof based on straight coset enumeration is least perspicuous.

5 Some Fibonacci groups

The Fibonacci group (see [19]) \( F(r, n) \) is generated by \( n \) generators \( x_1, \ldots, x_n \), which satisfy \( n \) relations \( x_j x_{j+1} \cdots x_{j+r-1} = x_{j+r}, 1 \leq j \leq n \), where the subscripts are reduced modulo \( n \) to lie in the range \( 1, \ldots, n \). Not all of the Fibonacci groups have been identified, see [33]. They have been a fertile area for investigation. Some are known to be cyclic, and it is from these that we draw our examples since many proofs by coset enumeration for these groups are based on total collapses.

5.1 \( F(2, 5) \)

Conway [4] raised interest in the Fibonacci groups by asking for a proof that the Fibonacci group \( F(2, 5) = \langle a, b, c, d, e | abC, bcD, cdE, deA, eaB \rangle \) is cyclic of order 11. Various solutions are mentioned in [5], each utilising a few lines of algebraic manipulation and some group theory.

A coset enumeration in \( F(2, 5) \) over a cyclic subgroup yields a single coset, and forms the basis of a short proof of the result. Such an enumeration is easily carried out by hand; for example, using the subgroup \( \langle a \rangle \), with cosets 2 and 3 defined as \( 1b \) and \( 1B \), the required collapse to a single coset occurs after processing some nine deductions. It is straightforward to confirm that at least two definitions (not counting coset 1) are required for a successful enumeration. Thus \( F(2, 5) \) is cyclic, and hence abelian, and its order is easily found. For example, the presentation’s exponent sum matrix, along with its Hermite normal form, is shown in the first
Figure 1. Exponent sums and Hermite normal forms

<table>
<thead>
<tr>
<th>Group</th>
<th>Exponent sums</th>
<th>Hermite normal form</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 1 -1 0 0</td>
<td>1 0 0 0 7</td>
</tr>
<tr>
<td>F(2,5)</td>
<td>0 1 1 -1 0</td>
<td>0 1 0 0 6</td>
</tr>
<tr>
<td></td>
<td>0 0 1 1 -1</td>
<td>0 0 1 0 2</td>
</tr>
<tr>
<td></td>
<td>0 0 0 0 1  1</td>
<td>0 0 0 1 8</td>
</tr>
<tr>
<td></td>
<td>1 -1 0 0 1</td>
<td>0 0 0 0 11</td>
</tr>
<tr>
<td>F(3,5)</td>
<td>1 1 1 -1 0</td>
<td>1 0 0 0 13</td>
</tr>
<tr>
<td></td>
<td>0 1 1 1 -1</td>
<td>0 1 0 0 7</td>
</tr>
<tr>
<td></td>
<td>0 0 1 0 1  1</td>
<td>0 0 1 0 19</td>
</tr>
<tr>
<td></td>
<td>0 0 0 1 1</td>
<td>0 0 0 1 17</td>
</tr>
<tr>
<td></td>
<td>0 0 0 0 1</td>
<td>0 0 0 0 22</td>
</tr>
<tr>
<td>F(2,7)</td>
<td>1 1 -1 0 0 0</td>
<td>1 0 0 0 0 0 5</td>
</tr>
<tr>
<td></td>
<td>0 1 1 -1 0 0</td>
<td>0 1 0 0 0 0 4</td>
</tr>
<tr>
<td></td>
<td>0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 9</td>
</tr>
<tr>
<td></td>
<td>0 0 1 1 -1 0</td>
<td>0 0 1 0 0 0 13</td>
</tr>
<tr>
<td></td>
<td>0 0 0 0 0 1</td>
<td>0 0 0 0 0 0 22</td>
</tr>
<tr>
<td></td>
<td>0 0 0 0 0 1</td>
<td>0 0 0 0 0 0 29</td>
</tr>
</tbody>
</table>

This not only tells us that $F(2, 5) \cong C_{11}$, but also that $a = E^7$, $b = E^6$, etc.

In the context of the above 5-generator, 5-relator presentation, the 'best' proof-word we have been able to extract from any coset enumeration is

$$\Psi = [E](eaB)[E](eDC)c(deA)(abC)(Db)c(cBA)$$

$$[E](eaB)[E](eDC)c(deA)(abC)(cdE)(eAd)(Db)c.$$  

($\Psi$ is a single string, which has been written as shown to illustrate its structure; see below). Items in square brackets, [ ], are subgroup generators (in this case inverted); items in round brackets, ( ), are relators (perhaps inverted or cycled); and the remaining items are conjugating group generators. The word $\Psi$ was extracted from an enumeration over $\langle e \rangle$, and proves that $b = E^6$ (i.e., $b \in \langle e \rangle$). The entire word, freely reduced, equals $b$, while if the relators are first deleted (being trivial in $F(2, 5)$), reduction yields $E^6$.

Although our coset enumeration proves that $F(2, 5)$ is cyclic, our proof-word by itself does not. However, substituting $b = E^6$ into the other relations rapidly leads to that conclusion and also enables us to find the group order.

The overall length of the proof-word compares favorably with that of the other published proofs (see §7). It is not a conventional proof, but we can convert it in
various ways. For this example, first note that $\Psi$ can be written as $\alpha \beta \alpha \gamma$, where

$$
\alpha = [E](eaB)[E](eDC)c(deA)(abC),
\beta = (cDb)C(cBA),
\gamma = (cdE)(eAd)(Dbc)C.
$$

Now $\alpha$, $\beta$ and $\gamma$ prove, respectively, that $aC = E^3c$, $cDA = C$ and $cdAb = C$. So $\alpha \beta$ proves $aDA = E^3$, while $\alpha \gamma$ proves $adAb = E^3$. Thus $\Psi$ can be split into a series of lemmas, which can be combined to yield the result. To rewrite, say, $\beta$ as a more usual type of proof, start with $cDA$ and insert the trivial relator $bB$ to give $cDbBA$. Now $cDb$ is trivial in $F(2, 5)$, being a cyclic conjugate of the relator $bcD$, and so we can cancel to $BA$. Now insert the trivial relator $Cc$ to give $CcBA$, and cancel the relator $cBA$ (inverse of the relator $abC$) to yield $C$.

Note that $\alpha$, $\beta$ and $\gamma$ are not proof-words (since the conjugating symbols are not properly paired within them). However, $\alpha C$, $c\beta$ and $c\gamma$ are; they prove, respectively, $aC^2 = E^3$, $c^2DA = 1$ and $c^2dAb = 1$. This is equivalent to rewriting $\Psi$ as $\Psi_1 = \alpha C c^2 \alpha c^2 c\gamma$. Now $c\beta$ and $c\gamma$ make no use of $[e]$ or $[E]$, and so can be regarded as proof-words in $F(2, 5)$ over the trivial subgroup, while $\alpha C$ makes use of the subgroup $\langle e \rangle$. However, if the square brackets are deleted, the word can be rewritten as the proof-word $e^3\alpha C$, which proves that $e^3aC^2$ is trivial in $F(2, 5)$. We can now rewrite $\Psi$ as $\Psi_2 = e^3\alpha C c^2 \beta E^3 e^3\alpha C c^2 e^3$, proving that $e^3be^3$ is trivial in $F(2, 5)$.

Various graphical representations of a group have been described, and these have numerous applications [30, 24]. One useful pictorial representation of proofs in a group $G = \langle X \mid R \rangle$ is provided by van Kampen diagrams. These can be thought of as portions of the Cayley diagram of $G$ drawn in the plane, where the words corresponding to the boundaries of internal faces are members of $R$. Such a diagram demonstrates that the boundary of the graph is trivial in $G$ (see [17, 18]).

Each relator or conjugated relator yields a component of such a graph, as shown in the left part of Figure 2 for $c\beta$. Then the (conjugated) relator can be recovered by tracing the boundary of the component and recording the arc labels, where a traversal against the arrow yields the inverse. These components are combined into a single reduced diagram (the right part of the figure), where there is at most one in-arc and one out-arc at each node for each generator. The proof-word $c\beta$ corresponds to the traversal $123421241$ of the diagram; note that internal edges are
traversed an equal number of times in opposite directions, while boundary arcs are traversed once more in the direction of the external traversal than against it.

Van Kampen diagrams for $e^3\alpha C$ and $c\gamma$ can be constructed similarly. Copies of all these diagrams can be combined to yield a composite diagram for $\Psi_2$, which provides an alternative proof that $be^6$ is trivial in $F(2, 5)$. Figure 3 is an ‘exploded’ version of such a diagram, illustrating how the various sub-diagrams (i.e., lemmas) combine to yield the result.

Since $F(2, 5) \cong C_{11}$, an enumeration over the trivial subgroup must make at least ten definitions (not counting coset 1). We do not know how few suffice, and our best result is fifteen definitions. One example is the sequence: $1c, 2c, 1A, 4B, 1C, 6e, 7b, 7C, 9e, 6d, 11A, 2B, 4E, 2c, 2b$. That is: coset 2 is defined as coset 1 times $c$, coset 3 is defined . . . , coset 16 is defined as coset 2 times $b$. Note that there is a coset table involving eleven cosets implicit in the diagram of Figure 3. However, this table is incomplete, and only contains the table entries corresponding to the diagram’s edges (after making all possible deductions). Furthermore, two of the nodes represent the same coset. (Since the group is abelian, the relator $bcD$ implies that $Dcb$ – the path 1234 – should trace out a cycle, which it does not.)

We were unable to find any proof-word showing that $F(2, 5)$ is cyclic which used less than thirteen relators of our initial presentation. We found many proof-words with thirteen relators, and $\Psi$ has the smallest number of subgroup generators plus group generators (six and four respectively) and the ‘nicest’ structure. There are many possible ways to extract a proof which readily implies the group is cyclic. For example, we can enumerate over the trivial subgroup and prove that, say, $ae^7$ or $be^6$ is trivial, or that the commutator $[a, b]$ or $[a, c]$ is trivial. Or we could enumerate over a subgroup $\langle e^m \rangle$, for some $m$, and extract a proof that $a$ or $b$ is in the subgroup. We tried a variety of these, and none yielded a better proof-word than $\Psi$. We did find various other proofs using only thirteen relators, but these
were all the same as $\Psi$, in the sense that they proved that $b = E^6$, or an equivalent of this under the action of the cyclic automorphism of the generators.

5.2 $F(3, 5)$

The Fibonacci group $F(3, 5) = \langle a, b, c, d, e \mid abcD, bcdE, cdeA, deaB, eabC \rangle$ is cyclic of order 22. The proof of this relies on coset enumeration, and Johnson [17] noted that “some 200 cosets are required and no short manual proof has yet been found”.

A coset enumeration in $F(3, 5)$ over one of the generators gives index one, and establishes that $F(3, 5)$ is cyclic. The exponent sum and normal form matrices can now be used to complete the proof (see Figure 1). PEACE can complete such an enumeration using 30 definitions. One such sequence of definitions, over $\langle a \rangle$, is: 1d, 1c, 3d, 4A, 3D, 6A, 6c, 1b, 9d, 6b, 11C, 9b, 13b, 1C, 15d, 6D, 17c, 13c, 19d, 20a, 19A, 9A, 23b, 24A, 11c, 26d, 27b, 3E, 29E, 30a. This is a substantial improvement on the number of cosets used by Johnson.

The shortest proof-words which we have been able to extract via PEACE all contained 103 occurrences of the group’s relators. They all have a similar structure, and the example chosen came from a (pruned) enumeration in 46 total cosets; it proves that $a \in \langle e \rangle$. Consider the words $\delta$ and $\varepsilon$ shown in Figure 4. These prove, respectively, that $CB = AE^6C$ and $dce = cEa$. Now consider the proof-word $a\delta \varepsilon(bCea)(AEbD)b_c$. This proves $a = E^{13}$, as required, and the proof can now be completed by substitution in the presentation. Note that there is considerable further structure in the proof-word which we have not exploited; e.g., the first two-thirds of the first two lines of $\delta$ are the same.

We also attempted to extract proofs that some commutators are trivial from enumerations over the trivial subgroup. The best we found had 126 group relators, and did not have any obvious substructure. However we did note that the coset enumeration over the trivial subgroup could be done using a total of only 98 cosets.
5.3  \( F(2, 7) \)

The Fibonacci group

\[
F(2, 7) = \langle a, b, c, d, e, f, g \mid abC, bcD, cdE, deF, efG, fgA, gaB \rangle
\]

is cyclic of order 29. The first hand proof of this was given by Havas [10]. His proof that \( F(2, 7) \) is cyclic was extracted from the workings of a series of index 1 coset enumerations over a one-generator subgroup. The first of these took 327 total cosets, and subsequent enumerations incorporated the results of previous ones as additional (redundant) subgroup generators. Taking into account the multiplicity with which the various results are used, Havas’s proof (implicitly) uses the group’s relators a total of some 24742 times. Subsequently, using a single enumeration in 55 cosets, Edeson [6] was able to extract a similar proof which used the relators only 2278 times.

At the time of writing, the best proof-word we have found using PEACE and a cyclic subgroup uses the group’s relators a total of 595 times, and was derived from an enumeration using a total of 74 cosets. The proof-word produced by PEACE contains many repeated substrings and could be rewritten as a series of lemmas, in a similar fashion to the proofs given by Havas and Edeson. Although much shorter than the previous proofs, our PEACE proof is no less opaque, so we refrain from reproducing it here.

Edeson’s proof is some four times longer than ours, despite the fact that the enumeration she used is approximately three-quarters the length of ours. Edeson had available other, shorter enumerations, but she rejected these since “the path to the primary coincidence was longer and more convoluted in the smaller sets”. It is interesting to note that Leech [23] also extracted, but did not publish, a proof from a 55 coset enumeration, but its length only bettered that of Havas’s proof by a factor of about two.

The developers of ITC used \( F(2, 7) \) as one of their test cases, and they were able to complete a coset enumeration in \( F(2, 7) \) over one of the generators with 50 cosets (i.e., 49 definitions). We were able to duplicate this value using PEACE, and one sequence of 49 definitions, for an enumeration over \( \langle b \rangle \), is: \( 1c, 2a, 3F, 3B, 5E, 6g, 7B, 8F, 1a, 10B, 11C, 12d, 1A, 14F, 15C, 16g, 17E, 11D, 19a, 20B, 12C, 22E, 1F, 14g, 10c, 14E, 1D, 28F, 1E, 30C, 30B, 11B, 11f, 11a, 28E, 36A, 36d, 10g, 39F, 40a, 41F, 39D, 10C, 44A, 45E, 40C, 36C, 45D, 15c. \)

It is straightforward to verify this definition sequence works. For example, this can be done independently using ITC. The result now follows from the exponent sum and normal form matrices (see Figure 1). Given the relative sizes of the coset table involved in this enumeration and of the proof-word, it is not clear which of these is the better proof. The sequence of definitions needed to complete the coset enumeration is certainly much shorter than the proof-word. However, verifying that the enumeration completes correctly is more involved than checking the reductions of the proof-word. Neither proof seems to offer any insight, so their relative merit in that sense is a moot point.
6 The trivial group

The presentation of the group $E_1 = \langle a, b, c \mid CacAA, AbaBB, BcbCC \rangle$ is a well-known example for the trivial group. That $E_1$ is trivial is not obvious; see Higman [15] and Neumann [26]. Neumann [27] notes that this presentation was very difficult for early computer implementations of the Todd-Coxeter process, and that it can be used to build an infinite series of presentations for the trivial group of ever-increasing difficulty (see [11] for more details and solutions for the next presentation in the series, $E_2$).

6.1 $E_1$ and 2-generator subgroups

Higman [15] proved that $E_1$ is trivial by expressing one of the generators in terms of the other two, and then using an argument based on derived groups. Replacing his $a_1, a_2, a_3$ by our $a, b, c$, the first part of his proof runs as follows. $AbaBB = 1$ implies that $baB = ab$, and conjugating this by $c$ gives $CbcCacCBc = CacCbc$. Now $CacAA = 1$ and $BcbCC = 1$ imply that $Cac = aa$ and $Cbc = bC$, and substituting these yields $bCaacB = aabC$. Finally, $CacAA = 1$ also implies that $Ca = aaC$ and $ac = caa$, and substitution now gives $baaCcB = baaaaB = aabC$. Thus, $c = bAAAAABaab$. (Note that Higman’s paper says $c = ba^4b^{-1}a^2b$, but this is presumably just a typographical error.)

Coset enumerations in $E_1$ over two-generator subgroups yield a single coset and can be completed with two definitions. (For the subgroup $\langle a, b \rangle$, there are precisely two such definitions sequences: $1C, 2b$ and $1C, 2C$.) The best proof-words we have been able to extract from such enumerations have eight group relators, and a typical example is the word

$$
\Upsilon = (CbccB)[b] [A][A](aaCAc) [A][A](aaCAc) [B](bCCBc) 
Cb(caaCA)(aBBAb)Bc (CacAA)[a][a] (CbccB)[b].
$$

This word freely reduces to $c$ and, after deleting the relators, reduces to the product $bAAAAABaab$ of subgroup generators $a$ and $b$. Higman’s proof also contains eight uses of the relators and, in fact, each of the relators presenting $E_1$ is used the same number of times in Higman’s proof as in $\Upsilon$. We obtained various other proof-words of a similar length, but these were all analogues of $\Upsilon$, over different subgroups and with different arrangements of the relators and varying amounts of conjugation.

Given a balanced presentation of the trivial group (i.e., with the same number of relators as generators), one way of proving it to be trivial is to reduce it to the standard presentation $\langle X \mid X \rangle$ by a sequence of relator inversions, conjugations and multiplications. Andrews and Curtis [1, 2] have conjectured that this is always possible. This conjecture is still open, and $E_1$ is a long-standing potential counter-example for the 3-generator case; in fact, the shortest 3-generator possibility.

In some circumstances we can extract from a proof-word a sequence of Andrews-Curtis moves on the relators which prove the same result (see [13] for more details). Since $\Upsilon$ contains the relator $AbaBB$ once only, the simple technique described in [13] can be used. We first rewrite $\Upsilon$ as $\Upsilon BAAbaaaB$ so that it is a proof,
over the trivial subgroup, that $cBABaaaBBBB$ is trivial (cf. Section 5.1). It is now straightforward to build up the proof-word, one relator at a time, by first using $AbaBB$ and thereafter multiplying by appropriately conjugated versions of the other two relators. Counting conjugation by a single group generator as a single move, one sequence of moves produced was 34 moves long, and transformed $(CacAA, AbaBB, BcbCC)$ into $(CacAA, bAAAABaabC, bCCBc)$.

A much shorter sequence can be found by searching directly for a sequence of Andrews-Curtis moves which produce a relator of the desired form. As part of the work described in [12] we produced a utility ACME (Andrews-Curtis move enumerator) which performs a breadth-first search through a tree of Andrews-Curtis equivalent presentations. It is a simple matter to modify this to check each new relator as it is generated to see whether it contains some group generator once only. If the intermediate presentations generated are allowed to grow to a total length of 21, then ACME can produce a 20-move sequence. For a length bound of 22, ACME can produce the 19-move sequence illustrated in Figure 5. (These searches took about 92 and 517 minutes, respectively, of CPU time on a 733 MHz Itanium 1 machine, and used about 3 and 12 gigabytes of memory.)

ACME was able to generate all cyclic permutations and inversions of the word $BaabCbAAAA$, as well as the equivalent proofs for $a$ and $b$ (60 proofs in all), but was unable to produce any shorter word or any presentation of shorter total length. (In [12] we were able to show that some move sequences had minimal length. Here, the searches were non-exhaustive, so we cannot say whether or not any shorter

---

<table>
<thead>
<tr>
<th>move</th>
<th>relator $r$</th>
<th>relator $s$</th>
<th>relator $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r \rightarrow r^C$</td>
<td>acAAC</td>
<td>AbaBB</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^A$</td>
<td>acAAC</td>
<td>baBBA</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow sr$</td>
<td>acAAC</td>
<td>baBBcAAC</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^b$</td>
<td>acAAC</td>
<td>aBBcAACb</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow st$</td>
<td>acAAC</td>
<td>aBBcAbCC</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^{-1}$</td>
<td>acAAC</td>
<td>ccBaaCbbaA</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^a$</td>
<td>acAAC</td>
<td>AccBaaCbb</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^B$</td>
<td>acAAC</td>
<td>bAccBaaCb</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow st$</td>
<td>acAAC</td>
<td>bAccBaabCC</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^b$</td>
<td>acAAC</td>
<td>AcBaaabCbb</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow st$</td>
<td>acAAC</td>
<td>AccBaabCcCC</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^A$</td>
<td>acAAC</td>
<td>ccBaabCbCCA</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow sr$</td>
<td>acAAC</td>
<td>ccBaabCbCAAC</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^c$</td>
<td>acAAC</td>
<td>cBaabCbCAA</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^a$</td>
<td>acAAC</td>
<td>AcBaabCbCA</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow sr$</td>
<td>acAAC</td>
<td>AcBaabCbAAC</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^A$</td>
<td>acAAC</td>
<td>cBaabCbAAACA</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow sr$</td>
<td>acAAC</td>
<td>cBaabCbAAAAAC</td>
<td>BcbCC</td>
</tr>
<tr>
<td>$s \rightarrow s^c$</td>
<td>acAAC</td>
<td>BaabCbAAAAA</td>
<td>BcbCC</td>
</tr>
</tbody>
</table>
word can be generated, or whether there is a shorter sequence of moves.) The sequence of moves in Figure 5 uses each of the relators the same number of times as Higman’s proof and as Υ, and is easily converted to a proof-word similar to ΥBAAAbaaaaB (in fact, slightly shorter).

Note that Higman’s proof that $c = bAAAABaab$ can be extended by using the relator $AbaBB$ twice to yield $c = bAABB$. (However, Andrews-Curtis and ACME proofs cannot be simply extended to reproduce this, since the relator $AbaBB$ has been ‘forgotten’ by the time it is needed.) Continuing on, if we conjugate $c = bAABB$ by $b$ and then use the relator $BcbCC$, we obtain $c^2 = A^2B^2$ (cf. Neumann’s proof in Section 6.2).

We note that Rapaport [29] also reduced the problem from a presentation on three generators to one on two generators; however her proof used automorphisms as well as Andrews-Curtis moves. Automorphisms can always be eliminated from a successful reduction to the standard presentation, but not necessarily from a partial reduction. As far as we are aware, our reduction is the first direct proof that the Andrews-Curtis conjecture for $E_1$ can be reduced to a two generator problem. (Of course, this result is already implicit in Higman’s work, which predated the conjecture.)

The various methods we have illustrated all yield essentially the same result; i.e., a particular relation which holds in $E_1$ and which proves that one of the generators can be written in terms of the other two. This raises the question as to whether or not this relation is the shortest of its kind in $E_1$. Of course, since $E_1$ is trivial any relation is true in it, so this not a well-formed question. However, it is meaningful to ask whether Υ is a shortest proof-word for this relation, or for any relation of this kind. Υ was found as part of a large, non-deterministic search and, in principle, we can establish whether or not it is a minimal proof-word via an exhaustive search of all proof-words containing fewer than 53 group generator symbols. However, the search-space is very large, and we have not investigated this further.

6.2 $E_1$ and cyclic subgroups

Neumann [26] proved $E_1$ trivial by first showing that $c^2 = b^3$, and so $b$ and $c^2$ commute. With this, $BcbCC = 1$ implies that $c$ is trivial, and the result follows. We elaborate the details here to enable easy counting of relator use. To establish $c^2 = b^3$, Neumann started by noting that $B'ab^i = aB^i$ (1). The relator $AbaBB$ yields that $aBB = Ba$ (2) and $aB = Bab$ (3). Conjugating (3) by $b$ yields $Ba = BBabb$, and substitution using (2) gives $aBB = Bbabb$. Repeated conjugation now gives (1) for general $i$. Similarly, $C^ibc^i = bC^i$ (4).

Now consider $CacAA = 1$, in the form $Cac = aa$, and conjugate by $b$ to give $Bcacab = Baab$ (5). The left-hand side $BCa.cb = CCabcc$ using the relator $BcbCC$, and (1) now gives $CCaBcc$. Now $CacAA = 1$ implies $Ca = aaC$, and substituting repeatedly for $Ca$ implies $CCaBcc = aaaaCCBcc$. Now use (4), in the form $C^2Bc^2 = c^2B$, to give $aaaaaccB$. For the right-hand side of (5), $Baab = BabBab = aBaB = aaBBB$, using (1) and (2). Thus $aaaaaccB = aaBBB$, and so $cc = AABB$ (6).
Now take (4), in the form $bc^2 = c^2bC^2$, and substitute using (6). This gives $bcc = AABBBbaa = A.Aba.a = Abba$, using the relator $AbaBB$. But $Abba = Aba.Aba = bbbb$, again using the relator $AbaBB$. Thus, $bcc = bbbb$ and so $cc = bbb$.

Neumann’s proof suggests that we should attempt to extract proofs of the general form $W \in \langle x \rangle$, where $x$ is a generator of $E_1$, and $W$ is a word such as $y^3$ or $y^jz^k$. His proofs that $c^2 = A^2B^2$ and $c^2 = b^3$ use the group relators a total of 12 and 29 times, respectively, and provide convenient metrics against which to compare proof-words produced by PEACE. (Given $c^2 = b^3$, direct substitution in $BcbCC$ yields $c = b^3$, and so a single relator use completes the proof that $c$ is trivial.)

The shortest proof-words that PEACE was able to extract from enumerations over cyclic subgroups had 24 or 26 occurrences of the relators and proved that $c = b^3$ (or an equivalent). Although not the shortest proofs, the ones with 26 relators have an interesting structure. A typical example is $\Phi = [B]c(CbCcB)\zeta \eta \zeta^{-1}C$, where

$$
\zeta = bbb(aaCaC)B(bABab)BAb(acAcA)B(bABab)aB(bABab)(BCbCc)
$$

$$
CC(BAbaB)bAB(acAAC)(ebCCb)b(caaCA)aBc(CbCcB),
$$

$$
\eta = [b]A(AbaBB)a[b](BAbB)a[b](BAbB).
$$

The words $\zeta$ and $\eta$ prove, respectively, that $bbabab = C$ and $bAAbbaaB = bbbb$, so $\zeta \eta \zeta^{-1}$ proves $b = Cb^4c$. The eleven relators in $\zeta$ match those in the extended version of Higman’s proof given in Section 6.1, while the conjugation of $\eta$ by $\zeta^{-1}$ mimics a key step in Neumann’s proof.

Our shortest proof-words, with 24 relators, are 20% shorter than Neumann’s 30, and seem to have a variety of different structures. We have not been able to find any ‘nice’ characterisation of these proofs, and we simply report an example with a minimal number of group generator symbols. This has 163 symbols, and proves that $a = c^3$.

$$
aCbcBA(abbAB)B(bABab)(ACacA)ba(ABabb)(CbcC)'bc(CbCcC)'bc
$$

$$
(CbCcC)(acAAC')[c]B(babaBA)ab(babaBA)B(babaBA)Ab[C](cBbCc)[C]
$$

$$
B(caaCA)a(cbcCB)A(abbAB)bac(CbcC)(cbCCb)(cbCcC)(cbCcC)(CbcC)(CbcC)(CcCcC)(bcBcb)
$$

$$
bcAAC(caaCA)AbAb(cBCbC)(caaca)(Aca[a][c][CbcC][c][c].
$$

### 6.3 Proof variability

To illustrate the extent of variation in the proofs produced by PEACE, we collected statistics from one set of runs for $E_1$ over cyclic subgroups. This involved 1000 coset enumerations over each of $\langle a \rangle$, $\langle b \rangle$ and $\langle c \rangle$, where the relators were randomly permuted, cycled and inverted before each enumeration. For each enumeration the number of cosets was pruned, and then proofs that each of the other two group generators are trivial (i.e., in the subgroup) were extracted. In the 3000 enumerations, the initial number of cosets varied from 17 to 112, and the pruned number varied from 10 to 56; the ratio of initial to pruned varied from 1.125 to 7.385. In the 6000 proofs, the number of occurrences of group relators varied from 26 to 39721; the median value was 961, and the lower and upper quartiles were 393 and 1669 respectively. There were two proofs with 26 relators, and these both
had a pruned coset count of 19. Thus enumerations and proofs are very variable, while succinct proofs are uncommon and are not necessarily associated with short enumerations.

The shortest enumerations we were able to find used ten cosets; i.e., made nine definitions. To check that this is the best possible, we need to enumerate and test all different eight-definition sequences. Now each of these definitions must fall within the first eight rows in the coset-table and, since the table has six columns (there are no involutions), must fall in one of 48 table positions. Each definition uses two positions (the definition and its inverse), and processing the subgroup generator fills an initial pair; thus we have 23 pairs to choose from. Since each pair can be filled in two ways, the total number of possible sets of eight definitions is $2^8 \binom{23}{8} = 125520384$. Each such set corresponds to a unique standardised definition sequence, if we order the definitions. (For example, first by coset number and then by column number: $1a, 1A, 1b, 1B, 1c, 1C, 2a, \ldots, 8c, 8C$. We have to take care that each definition uses an already defined coset, but, other than that, reordering the sequence simply relabels the cosets.) A similar calculation yields a total of 1599769600 possible sets for nine-definition sequences.

These counts are overestimates, since many of the sequences are not valid. For example, each definition in a valid sequence must use an already defined coset, so the first definition must be drawn from \{1a, 1A, 1b, 1B, 1c, 1C\}. It is easy to enumerate valid sequences in order using a backtrack search, and then to test whether or not they yield a complete coset table.

There are 16112057 valid standardised definition sequences of eight definitions (for each choice of subgroup generator), and to generate and test them took about 6 minutes of CPU time on a 400 MHz SPARC machine. None of these sequences generated a complete coset-table. For nine definitions, the corresponding figures are 167710664 sequences and about 68 minutes. 564 of these sequences generate a complete coset-table, and the first and last definition sequences in order are: $1B, 1C, 2a, 2C, 3C, 4C, 5C, 6C, 8B$ and $1C, 2C, 3C, 4C, 5B, 6c, 7c, 8B, 9C$ (for enumerations over $\langle a \rangle$ with columns ordered $a, A, b, B, c, C$).

Leech [20, §12] gave an example of a coset enumeration wherein, to achieve an enumeration with the smallest number of definitions, a particular definition had to be avoided. He also noted that it would be difficult to “devise a computer procedure” to find such a sequence. It is interesting to note that none of the 564 sequences (for each of $\langle a \rangle$, $\langle b \rangle$ and $\langle c \rangle$) our searches generated used the definitions $1a, 1b$ or $1c$, implying that these definitions must be avoided if a shortest definition sequence is required.

6.4 $E_1$ over the trivial subgroup

For comparison with the previous results, we also performed coset enumerations over the trivial subgroup. The best enumerations, after pruning, completed with only 56 definitions; the best previous result seems to be 59 [7, §6.1]. We also built a variety of proof-words. The shortest that we found contained 49 relators, and proved that one of the group’s generator is trivial. Not surprisingly, we obtained no new insights into why the group is trivial.
7 Conclusions

We have described tools which enable us to investigate proofs in finitely presented groups. We have used these to throw light on various kinds of proofs, in particular proofs based upon coset enumeration. Our aim was to clarify the nature of some proofs and to start addressing Hilbert’s 24th problem about simplest proofs.

We introduced the notion of a proof-certificate for proofs which show that a specific word is constructively in a subgroup of a finitely presented group. We used the length of the certificate as a measure of simplicity. This is readily compared to the total number of cosets defined in an enumeration as an alternative measure. We also showed how proof certificates can be converted into other forms of proof.

We studied four specific examples for which coset enumeration based proofs rely on total collapse. We comment on each of these in turn.

For the Fibonacci group $F(2, 5)$ we illustrated a range of proofs that could be derived from coset enumerations using the initial presentation and we gave some attractive proof-words. We also showed how to translate these into alternative proofs. The best coset enumerations use only 2 definitions, so are indeed quite simple. We have not determined the length of the shortest useful proof-word for this case, but give a proof word which is both concise and has some meaningful structure. This proof-word uses the relators 13 times and contains only 49 generator symbols. We conjecture that it is a shortest useful proof-word for this situation.

To compare it with human proofs we see that Conway [5] explicitly gave two solutions to the problem, by J.A. Wenzel and by R.C. Lyndon, and acknowledged 50 more (one of which explicitly used coset enumeration). Presumably he chose the published solutions for their elegance. We note that both published solutions start by reducing the 5-generator, 5-relator presentation to a 2-generator, 2-relator presentation. If we count relator use in the explicit way it is done for proof-words, this first step uses the original relators 10 times and precedes some more manipulations which take a few lines. All up, they (implicitly) use the relators more than our proof-word. By that measure, our proof is simpler.

A 2-generator, 2-relator presentation for $F(2, 5)$ is $\langle a, b \mid abaBabb, abbAbab \rangle$. With this as a basis for coset enumeration, PEACE produces the following proof-word for $a = B^8$ from a coset enumeration over $\langle b \rangle$:

$$\begin{align*}
[B]|aabb & \quad (abaBabb) & \quad BBAA \\
[B]|B|A & \quad (abbabAbab) & \quad a & \quad [B]|abb(ABaBAbA)BBB \\
[B]|B|AAbab & \quad (ababAbab) & \quad BABaa & \quad [B]|abb(ABaBAbA)BBB & \quad [B]
\end{align*}$$

It uses its relators five times and fits easily on two lines, but is shown spaced out over three to highlight its structure.

For the Fibonacci group $F(3, 5)$ we illustrated how a more difficult problem could be addressed in the same kind of way. Indeed we obtain concise coset definition sequences and moderate length proof-words. For the Fibonacci group $F(2, 7)$, a harder problem again, we demonstrated how much better coset enumeration based proofs could now be done than before. However we obtained no new insights into why these groups are cyclic.
On proofs in finitely presented groups

Consideration of the trivial group $E_1$ introduces additional proof techniques related to the Andrews-Curtis conjecture. We observe that our best proof-word based on 2-generator subgroups corresponds closely to a long established proof by Higman. We conjecture that these indeed provide simplest proofs. We showed that the Andrews-Curtis conjecture for $E_1$ is not a totally genuine 3-generator problem in the sense that resolution of the conjecture for a 2-generator example would resolve it for $E_1$. Moving on to cyclic subgroups, we showed that our best proof-word is substantially simpler than Neumann’s corresponding proof, in terms of its use of relators.

We used coset enumerations over cyclic subgroups in $E_1$ to investigate shortest definition sequences. We were able to determine the length of these by a brute force calculation. On the other hand we have not determined the length of the shortest possible proof-word. Rather we showed that the lengths of proof-words produced by PEACE are very variable and that short ones are hard to find. In spite of this we were successful in finding some quite short proof-words.

Our experiments show that shortest proof-words are not necessarily obtained from shortest coset enumerations. The sizes of the enumerations and proof-words produced by PEACE are very variable. To extract short enumerations and proof-words we had to run large numbers of enumerations, while varying the enumeration parameters, the presentation’s form, the enumeration pruning, and the word to be proved. As yet we have no good methods for directly producing short enumerations or for predicting which enumerations will yield good proofs. However we have been successful in producing short proofs which perhaps meet Hilbert’s simplicity criteria, even when they do not provide natural insights as to why they work.

Finally, we point out that proof certificates can do more than just give us access to another method of proof and to some simple proofs. They can sometimes provide important insights. Automatic generation of proof certificates was a key step in resolving a long-standing conjecture [14] about an infinite family of groups. By studying proof certificates for particular cases we were able to develop completely general proofs.

References

