THE ELEMENTARY ANDREWS–CURTIS BOTTLENECK OF THE AKBULUT–KIRBY PRESENTATION AK(2) IS EXACTLY 13

KOREA SUPERINTELLIGENCE LABS

ABSTRACT. For a balanced two-generator presentation $R = \langle x, y \mid r_0, r_1 \rangle$ write $\mathcal{E}(R) = |r_0| + |r_1|$ for the total length of its freely reduced relators, and call a sequence of elementary Andrews–Curtis moves *peak-bounded by* B if every presentation it visits, endpoints included, has $\mathcal{E} \le B$. We determine the least B for which the Akbulut–Kirby presentation AK(2), with relators $xyxy^{-1}x^{-1}y^{-1}$ and x^2y^{-3}, can be carried to $\langle x, y \mid x, y \rangle$ by a peak-bounded sequence: it is exactly 13, against $\mathcal{E}(\mathrm{AK}(2)) = 11$. The upper half is a 32-move sequence. The lower half is a certificate of a different kind — a set of 1340 presentations that contains AK(2), contains no trivial presentation, and is closed under every elementary move whose result still has total relator length at most 12. Verifying such a set is a single pass with no queue, no visited-set bookkeeping and no termination argument, so a reachability lower bound of this shape is *checkable*, not merely computable. Read for its floor rather than its ceiling, the same set shows that inside budget 12 the presentation AK(2) cannot be shortened at all, not by a single letter; with the peak-13 sequence this pins the wall of its basin at height exactly 2. The quantity is the elementary analogue of a bottleneck distance defined by Carreras on a restricted substitution graph, whose scope paragraph explicitly excludes elementary-path peak bounds. Every theorem below is machine-checked in Lean 4 by kernel evaluation alone, and we say exactly which computations lie outside that guarantee. Nothing here is about AK(3) or about the Andrews–Curtis conjecture.

1. INTRODUCTION

The objects. Let F_2 be the free group on x, y. A *balanced presentation* of rank 2 is an ordered pair $R = (r_0, r_1)$ of elements of F_2, written $\langle x, y \mid r_0, r_1 \rangle$. The *elementary Andrews–Curtis moves* are

$$(1) \qquad \text{(AC1)}\ \ r_i \mapsto r_i r_j^{\pm 1}\ (i \neq j), \qquad \text{(AC2)}\ \ r_i \mapsto r_i^{-1}, \qquad \text{(AC3)}\ \ r_i \mapsto g\, r_i\, g^{-1}\ (g \in F_2),$$

the other relator being left alone. This is the stabilisation-free variant: the number of relators never changes. Two presentations are *AC-equivalent* when a finite sequence of these moves and their inverses joins them, and R is *AC-trivializable* when it is AC-equivalent to $\langle x, y \mid x, y \rangle$. The Andrews–Curtis conjecture [2] asserts that every balanced presentation of the trivial group is AC-trivializable, and it is open. Its best known potential counterexamples are the Akbulut–Kirby presentations [1]

$$\mathrm{AK}(n) = \langle x, y \mid xyxy^{-1}x^{-1}y^{-1},\ x^n y^{-(n+1)} \rangle$$

and the Miller–Schupp presentations [8]

$$\mathrm{MS}(n, w) = \langle x, y \mid x^{-1}y^n x y^{-(n+1)},\ x^{-1}w \rangle.$$

Here AK(2), of total relator length 11, is AC-trivializable, and this is classical; AK(3), of total relator length 13, is the open frontier and is not touched in this paper.

Everything below is quantitative rather than existential, so we fix one measure of size at the outset. The *energy* of a presentation is the total length of its freely reduced relators, $\mathcal{E}(R) = |r_0| + |r_1|$; along a sequence of moves it fluctuates, and the largest value it takes — the *peak* of the sequence, endpoints included — is the quantity this paper computes.

The source, and the question it declines. Carreras [3] supplies machine-checkable certificates at the length-14 Andrews–Curtis frontier: explicit move sequences establishing four equivalences among the six hard Miller–Schupp presentations left by Shehper et al. [11], two of them recorded

there as, to that paper's knowledge, the first public explicit certificates for any of the four AK(3)-equivalences asserted without proof in [11]. That paper also introduces, for a restricted graph of substitution moves it calls the GS graph, exactly the quantity we use. Its definition reads, verbatim,[1]

> "the *energy* of a vertex is its total relator length. For a path γ, define its *bottleneck* $B(\gamma) = \max_{P\in\gamma} E(P)$; for states S, T define $d_{\mathrm{GS}}(S,T) = \min_{\gamma:S\rightsquigarrow T} B(\gamma)$."

It computes exact values of d_{GS} — notably $d_{\mathrm{GS}}(\mathrm{MS}(3, yx^2y), \mathrm{AK}(3)) = 19$ — and then, in a scope paragraph, refuses the elementary analogue in as many words:

> "**Scope restrictions (essential).** These are statements about the GS-substitution graph: they are not AC-distance lower bounds, *not elementary-path peak bounds*, not move-count bounds, and not evidence of AC-inequivalence."

So the notion is published and the elementary values are not. That is the gap this paper enters, for one presentation.

Question 1.1. For a balanced presentation R, what is the least B such that some sequence of elementary Andrews–Curtis moves carries R to a trivial presentation while every presentation visited has total relator length at most B?

Three different quantities in this subject are called 13, and it is worth separating them before any theorem is stated. The total relator length of AK(3) is 13. The minimal total length of relators among potential counterexamples surviving the length-≤ 12 verification of Miasnikov and Myasnikov [9] is also 13. The number proved here is neither: it is a *maximum of total relator length along a path*, not the length of a presentation and not a count of moves.

Results. Write $\mathcal{T}$ for the set of trivial presentations in the strong sense used throughout — both relators single letters on different generators, so eight presentations in all — and say that R *trivializes within* B if some sequence of elementary moves carries R into $\mathcal{T}$ with peak at most B. Our two results are the following; both are sharp, and both are about the packaged move system of Section 2, in which AC3 conjugates by a single generator.

- **Theorem 4.3.** 13 is the least B for which AK(2) trivializes within B. Equivalently, AK(2) trivializes within B if and only if $B \geq 13$.
- **Theorem 4.4.** Every sequence of elementary moves that starts at AK(2) and has peak at most 12 ends at a presentation of total relator length at least $11 = \mathcal{E}(\mathrm{AK}(2))$. Inside budget 12 the presentation cannot be shortened at all; together with Theorem 4.3 the wall around AK(2) has height exactly 2.

The upper half of Theorem 4.3 is an explicit 32-move sequence (Theorem 4.1). The lower half, and all of Theorem 4.4, rest on a certificate whose shape is the methodological point of the paper: a set S of presentations with $\mathrm{AK}(2) \in S$, $S \cap \mathcal{T} = \emptyset$, and S closed under every elementary move whose result still has energy at most 12 (Section 3). Such a set proves that no peak-bounded-by-12 path leaves it, and checking it is a single pass over S: no queue, no visited-set discipline, no termination argument, no search. A reachability lower bound in this subject is therefore a checking problem, not a search problem, even though finding the set is a search. The set used here has 1340 elements and the pass costs 34,480 move applications.

What is not claimed. The bounds are about the packaged move system: AC3 by a *single generator*. An abstract derivation conjugating by an arbitrary $g \in F_2$ refines into $|g|$ packaged conjugations, so a peak bound here bounds the refined path and must not be read as a bound on "the AC distance" in any other sense. Nothing here is a statement about move counts: the 32-move sequence of Theorem 4.1 is not claimed short, only peak-optimal. Nothing here is about AK(3), and nothing

[1] Both quotations from [3] in this paper were read from the e-print LaTeX source of arXiv:2607.23611v1 and re-verified against it on 2026-08-30, at lines 399–403 and 449–452 of `manuscript.tex`. That v1 is the version of record: arXiv holds no later version of the preprint and no journal version is announced. The source's own abbreviating macros are expanded here; the emphasis in the second quotation is ours.

here is evidence for or against the Andrews–Curtis conjecture. Finally, the value 13 proved here and the value $d_{\mathrm{GS}} = 19$ of [3] live on different graphs with different energy functions and are not comparable.

Prior art. The notion is one month old at the time of writing and belongs to [3]; the value does not appear there — AK(2) occurs nowhere in that paper — and we could not find it or any statement of its kind elsewhere. What was searched, on 2026-08-30, was the following. Fagan et al. [12], a recent large computational campaign on this problem, use "the maximum length attained by any intermediate presentation" along a found trajectory as a reinforcement-learning reward feature, and plot presentation length along trivialization paths for AK(2) and two automorphic images; that is an achieved maximum along found paths, never a minimum over paths, and no number is asserted. The 13 of [9] is the different quantity separated above. Miasnikov [7] prints what appears to be the only explicit elementary trivialization of AK(2) in the literature, in 19 steps; recomputing total relator length at each of its own printed steps gives peak 25, which that paper does not state (Remark 4.5). In Panteleev and Ushakov [10] the bound L on harvested conjugate length is a search parameter; AK(2) appears in their results only among the already-settled presentations, trivialized in two of their supermoves, with no expansion into elementary moves and no peak recorded. Shehper et al. [11] record that AK(2) "is known to admit the shortest AC trivialization path of length 14", which is a move count and is not in conflict with anything here. A local full-text arXiv corpus with coverage through the announcement window of 2026-08-28 was swept: 1075 occurrences of "Andrews–Curtis", and a window filter after every occurrence of "Andrews–Curtis" or "Akbulut–Kirby" for bottleneck-, peak- and maximum-length phrasings returned one hit, an incidental conference title.

One item in that sweep is a gap and we state it as such. The breadth-first search of Havas and Ramsay [4] is the standard technique for length-bounded searches in this subject. Their paper is behind a paywall, no open-access copy of it exists that we could find, and *we have not read it*; we therefore assert nothing about its contents, we do not cite it for a peak bound, and a reader with access should treat it as the one channel of our prior-art search that was not closed.

2. Preliminaries

Words, states, packaged moves. A *letter* is a generator with a sign; a *word* is a list of letters, and $|w|$ is its length. Free reduction is written ρ. A *state* is an ordered pair $R = (r_0, r_1)$ of freely reduced words; it presents $\langle x, y \mid r_0, r_1 \rangle$, and we use the two words interchangeably with the group elements they represent when no confusion arises.

The move system we compute with is the one implemented by the verifier of [3], and we transcribe it exactly. A *packaged move* is one of

$$\mathsf{inv}_i : r_i \mapsto \rho(r_i^{-1}), \qquad \mathsf{conj}_{i,g} : r_i \mapsto \rho(g\, r_i\, g^{-1}) \quad (g \text{ a single signed generator}),$$
$$\mathsf{mul}_{i,j,\varepsilon} : r_i \mapsto \rho(r_i r_j^{\varepsilon}) \quad (i \neq j,\ \varepsilon = \pm 1), \qquad \mathsf{cyc}_{i,k} : r_i \mapsto \rho(r_i \text{ rotated left by } k),$$

each undefined (*illegal*) outside its stated side conditions and outside the index range. Two fidelity points matter, because an over-permissive move system is the unsoundness trap here. The condition $i \neq j$ in mul is enforced: without it $r_i \mapsto r_i^2$ would be admitted, which is not an Andrews–Curtis move. Conjugation is by a single generator, which is a restriction of AC3 and hence the safe direction. Rotation is not a fourth kind of move: writing $r = uv$ with $|u| = k$, rotating left by k gives $vu = u^{-1}ru$, an instance of AC3. A *ledger* is a finite list μ of packaged moves; its *replay* from R is defined when every move in turn is legal, and is then written $R \cdot \mu$.

Proposition 2.1 (transcription and soundness)**.** *Every legal packaged move carries a state to an AC-equivalent one, and hence so does every legal replay: if $R \cdot \mu$ is defined then $\langle x, y \mid R \rangle$ and $\langle x, y \mid R \cdot \mu \rangle$ are AC-equivalent. Moreover the relator swap $(a, b) \mapsto (b, a)$, which is not a packaged move, is AC-derivable in five elementary moves, through (ab, b), (ab, a^{-1}), (aba^{-1}, a^{-1}), (aba^{-1}, a).*

The swap identity is what lets a certificate land on a rearranged copy of its target and still prove a statement about the original. Combined with relator inversion and rotation — both themselves elementary — it generates the orbit group Σ under which endpoints are compared in [3].

Proposition 2.2 (the determinant obstruction)**.** *Abelianising (r_0, r_1) gives a 2×2 integer matrix of exponent sums. AC2 negates a row, AC3 fixes the matrix and AC1 adds ± 1 times one row to another, so the absolute determinant is an invariant of AC-equivalence. Consequently $\langle x, y \mid x^2,\, y\rangle$, of absolute determinant 2, is not AC-trivializable.*

Proposition 2.2 is the control that constrains the *definition* of the move system rather than any particular ledger: had mul been allowed to take $i = j$, the induced row operation would be "double a row", the invariant would fail, and every presentation would become trivializable while every certificate in this paper still replayed correctly.

Energy, peak, and the bottleneck.

Definition 2.3. The *energy* of a state is $\mathcal{E}(r_0, r_1) = |r_0| + |r_1|$. For a ledger $\mu = (m_1, \ldots, m_N)$ whose replay from R is defined, with intermediate states $R = R_0, R_1, \ldots, R_N = R \cdot \mu$, the *peak* is

$$\mathrm{pk}(R, \mu) = \max_{0 \le t \le N} \mathcal{E}(R_t),$$

the endpoints included; it is undefined exactly when some move of μ is illegal. Let $\mathcal{T}$ be the set of states $([p], [q])$ whose two relators are single letters on different generators. We say R *trivializes within B* if there is a ledger μ with $R \cdot \mu \in \mathcal{T}$ and $\mathrm{pk}(R, \mu) \le B$.

Two conventions in Definition 2.3 are chosen so that the lower bounds proved below are the stronger of the available readings. First, the peak includes both endpoints, so no budget below $\mathcal{E}(R)$ can ever be met. Second, the target is the whole eight-element set $\mathcal{T}$ — the full orbit of $\langle x, y \mid x, y\rangle$ under swapping and inverting relators — and not the single presentation $\langle x, y \mid x, y\rangle$, so "no trivialization within B" rules out more.

Since "trivializes within B" is monotone in B, the set of admissible budgets is an up-set of $\mathbb{N}$ and Question 1.1 asks for its least element, when there is one.

3. Closed-set certificates for reachability lower bounds

A ledger is a certificate that two presentations *are* connected. Nothing of that shape can witness that they are *not* connected inside an energy budget, and it is tempting to price such a statement as needing a verified bounded enumerator — a search, reproduced with its queue and its termination argument. It does not. The following definition and the two-line induction after it are the whole of the lower-bound theory used in this paper.

Definition 3.1. A set P of states is *closed inside the sublevel $\{\mathcal{E} \le B\}$* if for every $R \in P$, every packaged move m, and every state R' with $R \cdot m = R'$ and $\mathcal{E}(R') \le B$, we have $R' \in P$. Moves whose result leaves the sublevel are unconstrained, and that is what makes the condition a one-pass check rather than a search.

Theorem 3.2 (soundness of a closed set)**.** *Let P be closed inside $\{\mathcal{E} \le B\}$ and let $R \in P$. If μ is a ledger whose replay from R is defined with $\mathrm{pk}(R, \mu) \le B$, then $R \cdot \mu \in P$. Consequently, if moreover $P \cap \mathcal{T} = \emptyset$, then R does not trivialize within B.*

Proof sketch. Induction on μ. The peak dominates the energy of every state the replay visits, in particular of the state reached after the first move, which is therefore admissible input to the closure hypothesis; the remaining ledger has peak at most $\mathrm{pk}(R, \mu) \le B$, so the induction hypothesis applies. The second statement is the contrapositive against the target set. No finiteness of P is used. □

Theorem 3.2 shifts the burden entirely onto exhibiting P, and two obstacles stand between a set of states and a mechanically checkable certificate.

The first is that Definition 3.1 quantifies over the infinite supply of moves: $\mathsf{conj}_{i,g}$ and $\mathsf{cyc}_{i,k}$ range over unbounded indices and rotation amounts. For a two-relator state R the following finite list suffices: the fourteen index- and generator-bounded moves inv_i (2), $\mathsf{conj}_{i,g}$ (8) and $\mathsf{mul}_{i,j,\varepsilon}$ (4), together with $\mathsf{cyc}_{i,k}$ for $0 \le k < |r_i|$, so $14 + \mathcal{E}(R)$ moves in all.

Lemma 3.3. *Every legal packaged move out of a two-relator state has the same effect as one of the* $14 + \mathcal{E}(R)$ *listed moves.*

Proof sketch. The index bounds are the legality conditions themselves. The rotation case is the one with content, and it is settled by an identity rather than by an elimination: rotation already reduces its argument modulo the relator length, so $\mathsf{cyc}_{i,k}$ and $\mathsf{cyc}_{i,k \bmod |r_i|}$ are literally the same move. The rotation opcode is therefore *kept* in the closure check, and the certificates below are honest for the full four-opcode system that the replay checker uses — there is no side lemma claiming that rotations are redundant. □

The second obstacle is cost: deciding membership of a state in a set of 1340 states by structural comparison, once per move per state, is the dominant expense. We encode a word as a numeral, little-endian base 5 with digits $1, \dots, 4$, and a state of two relators of length at most 13 as $\mathrm{code}(r_0) + 5^{13}\,\mathrm{code}(r_1)$.

Lemma 3.4. *Decoding inverts encoding: for words of length at most* 13*, the decoder applied to the numeral of a state returns that state.*

Lemma 3.4 is what keeps the *statements* about presentations while the *checks* are about integers: a numeral in the shipped table can be turned back into the state it stands for, so no separate injectivity argument is needed anywhere. With it, a closed set ships as a list of numerals bucketed by residue, and membership is a short scan.

4. The bottleneck of AK(2)

Throughout this section $\mathrm{AK}(2) = (xyxy^{-1}x^{-1}y^{-1},\ x^2y^{-3})$, of relator lengths 6 and 5 and energy 11.

The upper half.

Theorem 4.1. *There is a ledger of* 32 *packaged Andrews–Curtis moves carrying* $\mathrm{AK}(2)$ *to* $\langle x, y \mid x, y\rangle$ *with peak exactly* 13*. In particular* $\mathrm{AK}(2)$ *trivializes within* 13*, and* $\mathrm{AK}(2)$ *is AC-trivializable.*

Proof sketch. The ledger is exhibited and replayed. Its energy sequence is

$$11, 11, 11, 11, 11, 11, 11, 12, 12, 12, 12, 12, 12, 12, 13, 13, 13, 13, 13, 13, 13, 11, 11, 8, 8, 8, 8, 5, 5, 3, 3, 3, 2,$$

so the peak is 13 and the largest single relator along the way has length 7. The replay lands on (x, y) exactly, not merely somewhere in $\mathcal{T}$, so Proposition 2.1 upgrades it to AC-trivializability of $\mathrm{AK}(2)$ — the classical fact, here refined rather than restated. No search occurs inside the proof: the ledger is data, and the verification is one replay of 32 moves. □

The ledger was found by a bounded-peak search written for this purpose, whose correctness is not part of any proof below; how it was found and how it was cross-checked is described in Section 8.

The lower half.

Theorem 4.2. *No ledger of packaged Andrews–Curtis moves carries* $\mathrm{AK}(2)$ *into* $\mathcal{T}$ *with peak at most* 12*.*

Proof sketch. Let S be the explicit set of 1340 states listed as numerals in the formal development — the connected component of $\mathrm{AK}(2)$ in the graph on $\{\mathcal{E} \le 12\}$. Three facts are checked about it, and nothing else about it is used:

(1) $\mathrm{AK}(2) \in S$;

(2) no state of the shape $([p],[q])$ lies in S — checked for all sixteen such states, not merely for the eight that constitute $\mathcal{T}$, so the bound survives any widening of the target convention;
(3) S is closed inside $\{\mathcal{E} \le 12\}$.

Item (3) is the exhaustive part, and it is a single pass: for each of the 1340 states, Lemma 3.3 reduces the infinitely many moves to at most 26 (namely $14 + \mathcal{E}(R)$, with $\mathcal{E}(R) \le 12$), each is applied, and the result is either of energy greater than 12 or a member of S. That is 34,480 move applications, of which 23,840 stay inside the sublevel. There is no queue, no frontier, no visited set and no termination argument, because nothing is being searched: the set is given, and closure is a property of it. Theorem 3.2 then applies. □

Theorem 4.3. 13 *is the least* B *for which* AK(2) *trivializes within* B*; equivalently,* AK(2) *trivializes within* B *if and only if* $B \ge 13$.

Proof. Theorem 4.1 gives membership, and monotonicity in B propagates it upwards. If some $B \le 12$ admitted a trivialization then so would 12, contradicting Theorem 4.2. □

Since $\mathcal{E}(\mathrm{AK}(2)) = 11$, the sharp statement is that trivializing AK(2) requires going *two* units above its own length, and that two units are enough.

The floor of the same basin. The set S was built to miss the trivial presentations. It happens to be squeezed from below as well, and that is a second theorem rather than a remark, because it is the direction in which the literature has only qualitative statements.

Theorem 4.4. *Every state of* S *has energy at least* 11*. Hence every ledger of packaged Andrews–Curtis moves that starts at* AK(2) *and has peak at most* 12 *ends at a presentation of total relator length at least* 11*: inside budget* 12*, the presentation* AK(2) *cannot be shortened at all, not by a single letter, let alone trivialized. Together with Theorem 4.1, the wall around* AK(2) *has height exactly* 2*.*

Proof sketch. One pass over the same 1340 states verifies $\mathcal{E} \ge 11$ throughout; combined with Theorem 3.2 and $\mathrm{AK}(2) \in S$, every peak-$\le$ 12 replay from AK(2) stays in S and so ends at energy at least 11. The last sentence is the conjunction of this with the peak-13 ledger, which does reach energy 2. □

An inspection of S outside the formal development — not needed for the theorem, and recorded for shape only — finds that the only relator-length shapes occurring in it are $(6,5)$, on 360 states, and $(7,5)$, on 980; so the energies occurring in S are exactly 11 and 12, and the floor of Theorem 4.4 is attained. The reason to state the floor separately is that it speaks to the one qualitative observation we could find on this point: [12] reports that a greedy heuristic "struggles on automorphisms of AK(2) whose short trivialization paths … involve length-increasing steps", with no number attached. For AK(2) itself the necessary increase is now pinned at exactly 2.

Remark 4.5. Outside the formal development, and recorded for calibration only. The results section of [7] prints an explicit 19-step elementary trivialization of $\langle a, b \mid aba = bab,\ a^2 = b^3\rangle$ with $r_0 = a^2b^{-3}$ and $r_1 = abab^{-1}a^{-1}b^{-1}$ — that is, of AK(2), with the two relators in the other order. It states no peak. Reading off the total relator length of each of its own printed intermediate relators gives

$$11, 11, 14, 12, 12, 12, 15, 17, 25, 19, 19, 11, 8, 5, 4, 4, 3, 2, 2, 2,$$

so its peak is 25. Its step 9 conjugates by the length-3 word $ba^{-1}b$; refining that into three single-generator conjugations, innermost first, passes through total lengths $23, 21, 19$ and so adds no new maximum. The refinement is what fixes the reading of that step: it returns, letter for letter, the relator the paper itself prints for step 9. This arithmetic is ours, was re-derived here directly from the relators printed in the e-print of arXiv:math/0304306, and is not machine-checked.

5. Controls

A sharp bound proved by a certificate deserves controls on the certificate, and four ways this could have been wrong are excluded mechanically.

Proposition 5.1.

(1) The budget is load-bearing. *The same set S is* not *closed inside $\{\mathcal{E} \le 13\}$. The state*
$$\left(y^{-2}x^{-1}yxy^{-1}x,\ x^{-1}y^3x^{-1}\right) \in S$$
admits the move $r_0 \mapsto r_0 r_1$, whose result $\left(y^{-2}x^{-1}yxy^2x^{-1},\ x^{-1}y^3x^{-1}\right)$ has energy 13 and does not lie in S. A set closed at every budget would prove nothing about 12 in particular.

(2) The set is neither everything nor nothing. *It has exactly 1340 elements, and it contains neither* AK(3) *nor $\langle x, y \mid x, y\rangle$.*

(3) The budget notion is not degenerate. *The trivial presentation trivializes within 2, its own energy, and does not trivialize within 1; more generally no presentation trivializes within a budget below its own energy, since the peak includes the endpoints.*

(4) Not every presentation trivializes at some budget. *For every B, the presentation $\langle x, y \mid x^2, y\rangle$ does not trivialize within B. This routes through Proposition 2.2, which is also the check that the peak-bounded layer is correctly wired to the abstract move relation of* (1).

Two further checks guard the apparatus: the finite move list of Lemma 3.3 has exactly $25 = 14+6+5$ entries at AK(2)*, against a mis-sized list; and the previously known ledger for* AK(2) *used in this development has peak 18, so the ledger of Theorem 4.1 is a genuine improvement rather than a restatement.*

Item (4) is worth a sentence of emphasis. No amount of tampering with ledgers can detect an over-permissive move system, because a broken move system accepts honest ledgers for false statements; only an invariant can, and Proposition 2.2 is that invariant.

6. The published material replayed here

The peak-bounded layer above sits on a replay checker for elementary-move certificates, and that checker was built to reproduce the results of [3]. We record what it reproduces, since those statements are the source's and not ours, and since they fix the meaning of the numbers quoted in Section 7. The six hard length-14 presentations left by [11] are, in the labelling of [3],

$$\begin{aligned} &P_1 = \mathrm{MS}(2, x^{-2}y^{-1}x^2y), \quad P_2 = \mathrm{MS}(2, x^{-2}y^{-1}x^2y^{-1}), \quad P_3 = \mathrm{MS}(3, yx^2y), \\ &P_4 = \mathrm{MS}(3, y^{-1}x^2y^{-1}), \quad P_5 = \mathrm{MS}(2, yx^2yx^{-2}), \qquad P_6 = \mathrm{MS}(2, yx^2y^{-1}x^{-2}). \end{aligned} \tag{2}$$

Proposition 6.1 ([3], Theorems 1–4)**.** *Write $C = \langle x, y \mid x^{-1}y^2xy^{-3},\ y^{-1}x^2yx^{-3}\rangle$ for the classical candidate. Each of the pairs*
$$(C, P_6), \qquad (P_5, P_2), \qquad (P_3, \mathrm{AK}(3)), \qquad (P_4, \mathrm{AK}(3))$$
is AC-equivalent, by explicit ledgers taken verbatim from the ancillary files of [3]*, the third by two independent ledgers. Moreover C equals P_1 up to a rotation of its second relator, itself a one-move certificate.*

Proposition 6.2. *Consequently P_3 is AC-trivializable if and only if* AK(3) *is, and likewise for P_4: those two length-14 questions are not new open problems but the length-13 one in disguise. The remaining four presentations of* (2) *fall into the two blocks $\{P_1, P_6\}$ and $\{P_5, P_2\}$, whose members stand or fall together and which remain open.*

Proposition 6.3. *Both accountings of* [3] *are reproduced for all five of its ledgers: packaged counts 36, 85, 66, 71, 13 and classical counts 58, 155, 101, 106, 22, where a rotation by k of a relator of length L is charged $\min(k \bmod L,\ L - (k \bmod L))$ single-generator conjugations, evaluated at the moment of the move. In particular the difference between 66 and 71 is two different certificates for the same statement, not two accountings of one.*

Proposition 6.4. *Explicit ledgers, found here rather than taken from* [3] *and not claimed short, trivialize* AK(2) *and the five Miller–Schupp presentations*

$$\mathrm{MS}(1, y), \quad \mathrm{MS}(1, x^2y), \quad \mathrm{MS}(1, yx^2y), \quad \mathrm{MS}(1, y^{-1}x^2y^{-1}), \quad \mathrm{MS}(2, y).$$

These facts are classical; the ledgers exist to exercise the checker on data the source did not supply.

AK(3) itself is recorded in the development as a statement, deliberately not as a theorem in either direction, and no search for it was run.

7. A measurement log for four further pairs

This section states no result of this paper. It is a log of measurements made with the same apparatus while the work was being prepared, and it is set apart for one reason: half of each measurement is not a certificate. Nothing in Sections 2–6 depends on anything here, none of it is machine-checked, and the formal ledger to which the theorems above are pinned contains no entry for any of it. We report it because it locates the one proved value in a landscape; a reader who wants only proved statements should stop at Section 6.

Write $d_E(P, Q)$ for the least B such that some ledger of packaged moves joins P to a relator-inverted, relator-rotated copy of Q with peak at most B. For AK(2) against the trivial presentation, $d_E = 13$ is Theorem 4.3, and its lower half is the closed set S — 1340 states, 335 classes under relator inversion. For the four pairs below the two halves have very different standing.

The upper halves are checkable. Each is an explicit ledger, found here, and each was accepted by the verifier distributed with [3] before being recorded; the four ledgers are held in the repository referenced in Section 8, and re-running one against that verifier takes seconds. Column three is therefore a genuine upper bound on d_E.

The lower halves are not. Each is a sublevel component that a breadth-first search written for this purpose reported as exhausted at the budget in column four without meeting the target orbit. No closed-set certificate was extracted from that output, and the search program's correctness is not established anywhere. Column four is therefore a measurement, and we do not assert the lower bound it would give.

pair	ledger	its peak, so $d_E \leq$	budget reported exhausted	source's ledger peak
$\mathrm{AK}(3) \to P_3$	91 moves	19	18 (503,748 classes)	22
$\mathrm{AK}(3) \to P_4$	27 moves	19	18 (503,748 classes)	23
$P_1 \to P_6$	63 moves	21	20 (304,766 classes)	27
$P_5 \to P_2$	122 moves	23	22 (2,737,868 classes)	33

The lower halves were computed against the larger orbit that also allows swapping the two relators and the upper halves land in the smaller one, so both directions err safely; were the four searches confirmed, the four values of d_E would be 19, 19, 21 and 23. We state that as a conditional and not as a result.

One comparison in the table does *not* depend on the searches, because both of its sides are peaks of exhibited ledgers. The last column is the elementary peak of the source's own certificate for the same statement, recomputed here with the source's own move implementation: 22, 23, 27, 33, against its published move counts. So each of the four ledgers of column two lowers the peak of the published certificate for the same equivalence, by 3, 4, 6 and 10 respectively — for P_3 against the better of the source's two certificates, whose peaks are 24 and 22. That much is checkable by replay alone. Two of the four pairs — the last two lines — concern blocks that are still open at the frontier.

Finally, and claiming nothing from it: the first line's ledger has peak 19, and 19 is also the value $d_{\mathrm{GS}}(\mathrm{MS}(3, yx^2y), \mathrm{AK}(3))$ computed in [3]. The two quantities live on different graphs with different energy functions — substitution moves against elementary moves, cyclically reduced length against freely reduced length — so even if the searches above were confirmed the agreement would be a coincidence of two computations, not a restatement of one.

8. Formal verification

Every theorem, proposition and lemma stated above with the exception of Remark 4.5 and Section 7 is formalised and machine-checked in Lean 4 against Mathlib, and the check is by kernel evaluation only: the development uses no compiled-evaluation escape hatch, and the axiom set of every headline statement is the standard one (propositional extensionality, quotient soundness, and choice, the last entering only through library lemmas about orders and free groups). Three things lie outside that guarantee and are worth naming precisely. First, the *discovery* of both certificates: the 32-move ledger of Theorem 4.1 came from a bounded-peak search, and the set S of Theorem 4.2 from a breadth-first search of the sublevel component, both written independently of the formal development; neither program's correctness is used, since the ledger is replayed and the closure of S is re-derived from scratch inside the kernel. Second, the cross-checks that preceded transcription, which are evidence rather than proof: the ledger was accepted by the verifier distributed with [3] in both its equivalence mode and its trivialization mode; the closure of S was re-verified using that same verifier's own move implementation, a different code path from the search that produced S (34,480 move applications, 23,840 inside the sublevel, no violations); the component sizes 30 and 335 at budgets 11 and 12 were reproduced by two independent implementations in different languages, and the ratio 4 between the raw and inversion-quotiented counts held exactly at both, which is the regression gate that caught an earlier miscount. Third, the engineering choices that make the kernel check affordable: the closure sweep is split into eight kernel evaluations glued by one lemma, which cuts peak resident memory from 18.7 GB to 8.8 GB — a sixfold cut in the cost above the 6.8 GB library baseline, from 11.9 GB to 2.0 GB — while the processor time went down rather than up, from 255 to 212 seconds; and the membership table is bucketed by residue modulo 37, chosen by measuring expected lookup cost, which brings a membership test to about 101 reduction steps against about 2000 for a flat scan. The whole of the new material costs under ten minutes of processor time. These figures should be read as engineering estimates and not as reproducible constants: they come from one timed run on a machine shared with other jobs, and a second run of the same chunked check was recorded at 236 seconds and about 2.3 GB above baseline. The two records agree on everything that the argument uses — the sign and rough size of both effects — and we have not tried to reconcile the last ten percent. The repository is private; repository reference to be added at submission.

References

[1] S. Akbulut and R. Kirby, *A potential smooth counterexample in dimension* 4 *to the Poincaré conjecture, the Schoenflies conjecture, and the Andrews–Curtis conjecture*, Topology **24** (1985), no. 4, 375–390. doi:10.1016/0040-9383(85)90010-2

[2] J. J. Andrews and M. L. Curtis, *Free groups and handlebodies*, Proc. Amer. Math. Soc. **16** (1965), no. 2, 192–195. doi:10.1090/S0002-9939-1965-0173241-8

[3] J. Carreras, *Machine-checkable equivalence certificates at the length-14 Andrews–Curtis frontier*, arXiv: 2607.23611v1 (26 July 2026), together with the ancillary verifier and the five certificate files distributed with it, archived at doi:10.5281/zenodo.21499081. At the time of writing v1 is the only version of the preprint and no journal version is announced. All quotations and numbers taken from it in this paper were read from that e-print source, retrieved on 2026-08-30.

[4] G. Havas and C. Ramsay, *Breadth-first search and the Andrews–Curtis conjecture*, Internat. J. Algebra Comput. **13** (2003), no. 1, 61–68. doi:10.1142/S0218196703001365 Cited only as the standard length-bounded search technique in this subject, and never for a peak bound. The paper is paywalled, we found no open-access copy, and we did not read it; see Section 1.

[5] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, Lecture Notes in Computer Science, vol. 12699, Springer, 2021, pp. 625–635.

[6] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, pp. 367–381.

[7] A. D. Miasnikov, *Genetic algorithms and the Andrews–Curtis conjecture*, Internat. J. Algebra Comput. **9** (1999), no. 6, 671–686; doi:10.1142/S0218196799000370 e-print arXiv:math/0304306. The trivialization chain discussed in Remark 4.5, and its step numbering, are those of the e-print, read on 2026-08-30.

[8] C. F. Miller III and P. E. Schupp, *Some presentations of the trivial group*, in: Groups, Languages and Geometry (R. H. Gilman, ed.), Contemp. Math. **250**, Amer. Math. Soc., Providence, RI, 1999, pp. 113–115. doi:10.1090/conm/250/03848

[9] A. D. Miasnikov and A. G. Myasnikov, *Balanced presentations of the trivial group on two generators and the Andrews–Curtis conjecture*, in: Groups and Computation III (W. M. Kantor and Á. Seress, eds.), Ohio State Univ. Math. Res. Inst. Publ. **8**, de Gruyter, Berlin, 2001, pp. 257–263; e-print arXiv:math/0304305, read on 2026-08-30. Its Theorem 1 is the verification of the Andrews–Curtis conjecture through total relator length 12, and the statement referred to in Section 1 is its consequence that the minimal total length of relators in potential counterexamples that still stand is 13. The two authors are different people with similar names, and this is not a misprint.

[10] D. Panteleev and A. Ushakov, *Conjugacy search problem and the Andrews–Curtis conjecture*, Groups Complex. Cryptol. **11** (2019), no. 1, 43–60; doi:10.1515/gcc-2019-2005 e-print arXiv:1609.00325, read on 2026-08-30.

[11] A. Shehper, A. M. Medina-Mardones, L. Fagan, B. Lewandowski, A. Gruen, Y. Qiu, P. Kucharski, Z. Wang and S. Gukov, *What makes math problems hard for reinforcement learning: a case study*, arXiv:2408.15332v2 (11 February 2025). The sentence quoted in Section 1 stands unchanged in v1 and v2; both were read on 2026-08-30.

[12] L. Fagan, M. Tarquini, A. Shehper, M. Manko, A. Gruen, C. Huang, G. Butbaia, D. Passaro and S. Gukov, *The Two-Hump Problem: Bridging the Difficulty Gap in Mathematical Reinforcement Learning*, arXiv:2606.21611 (19 June 2026); the arXiv comments field announces acceptance at ICML 2026, which we have not independently confirmed. Title, author list and full source read on 2026-08-30.