

Sharp Graham-Sequence Reconstruction for Spiders

Alper Ferudun*

17 October 2026

Abstract

For a graph G , write $a_j(G) = |V(L^j(G))|$, where L is the line-graph operator. Let S be an n -vertex spider with $r \geq 3$ legs. We prove that its isomorphism class within the class of spiders is determined by the prefix ending at index

$$K = \left\lfloor \frac{n - r - 1}{2} \right\rfloor + 2.$$

The bound is sharp whenever $n - r - 1 \geq 2$, even when n and r are known. The proof recursively recovers the number of legs exceeding each length from the orders of a marked core; symmetry removes the unknown assignment of surviving legs. Trees with exactly three leaves are recognizable from a_0, a_2 , so they are reconstructed among all trees by the prefix through $\lfloor n/2 \rfloor$. Separately, two exact implementations verify Graham-sequence separation for all 1,346,024 unlabelled trees with at most twenty vertices. Neither the spider theorem nor the finite computation resolves Graham's conjecture for arbitrary trees.

1 The reconstruction question

All graphs in this paper are finite, simple and undirected. The line graph $L(G)$ has one vertex for each edge of G , with adjacency when the two edges have a common endpoint. Put $L^0(G) = G$ and

$$a_j(G) = |V(L^j(G))| \quad (j \geq 0).$$

The empty graph has all orders zero. Graham's tree reconstruction question asks whether the infinite sequence $(a_j(T))_{j \geq 0}$ determines a tree T up to isomorphism; see [1, 2]. The index-zero term, and hence the order of the original tree, is part of the input.

A *spider* $S(\ell_1, \dots, \ell_r)$, with $r \geq 3$ and positive integer leg lengths, consists of r internally disjoint paths meeting at one common endpoint. Its order is

$$n = 1 + \sum_{i=1}^r \ell_i.$$

The common endpoint is its unique vertex of degree at least three. Two spiders are isomorphic exactly when their multisets of leg lengths agree.

*Mercury Software GmbH. alper@mercurycodelab.com; github.com/AlperTheKing.

Theorem 1. *Let S be a spider of order n with $r \geq 3$ legs, and put*

$$s = n - r - 1, \quad q = \lfloor s/2 \rfloor, \quad K = q + 2.$$

The prefix $a_0(S), \dots, a_K(S)$ determines S up to isomorphism among all spiders. The number r is itself determined by $a_0(S)$ and $a_2(S)$.

For every fixed $r \geq 3$ and $s \geq 2$, two nonisomorphic spiders of order $n = r + s + 1$ agree through index $K - 1$ and first differ at index K .

A prefix through index K contains $K + 1$ terms; these two quantities should not be interchanged. The small cases $s = 0, 1$ are unique when n and r are specified, and are excluded from the sharpness assertion.

Corollary 2. *Every tree with exactly three leaves and n vertices is determined among all trees by its prefix through index $\lfloor n/2 \rfloor$. For every $n \geq 6$, this uniform bound is sharp.*

The comparison class in Theorem 1 is the class of spiders, not all trees. The stronger comparison class in Corollary 2 comes from a separate recognition argument in Section 5.

Cooper, Kay and Swifton [2] obtain a superpolynomial lower bound on the number of distinguishable Graham classes using caterpillars and partition methods. Weatherspoon and Zeilberger [3] give computations and Maple code for small trees. Balch [4] studies regular and biregular graphs and a star with one subdivision vertex. The abstract of King, Lane-Harvard and Milligan [5] describes further counting methods applied to a class of trees.¹ The argument below concerns inversion of the order sequence and a sharp prefix cutoff; the elementary line-graph counting identities it uses are classical.

2 Marked cores and tail lengths

For the fixed spider $S = S(\ell_1, \dots, \ell_r)$, define nested sets and their cardinalities by

$$A_j = \{i : \ell_i > j\}, \quad s_j = |A_j|, \quad A_0 = \{1, \dots, r\}.$$

We construct a core graph H_j with designated, labelled vertices called *ports*. The port set of H_j has labels A_{j-1} .

Start with $H_1 = K_r$, whose vertices are the ports with labels $1, \dots, r$. To pass from H_j to H_{j+1} , attach one new leaf to each port whose label belongs to A_j , and denote the resulting graph by H_j^+ . Set

$$H_{j+1} = L(H_j^+).$$

The vertex corresponding to the new leaf edge at port i is the port of H_{j+1} labelled i . Other port labels are retired, but the vertices of the underlying graph are not deleted.

Lemma 3. *For each $j \geq 1$, the graph $L^j(S)$ is obtained from H_j by attaching, at its port labelled i , a pendant path of length $(\ell_i - j)_+$. A path of length zero adds no vertices or edges. All designated ports of H_j are distinct and have degree $r - 1$ inside H_j .*

Proof. For $j = 1$, the first edge of each leg becomes a vertex of the clique K_r , and the rest of that leg becomes a pendant path of length $\ell_i - 1$. The assertions about distinctness and degree follow.

¹The full text of [5] was not available for the present comparison. Possible overlap with its detailed methods or results has not been excluded. This paper does not certify historical priority.

Assume the description holds at step j . For each nonempty pendant path, put its first edge into H_j^+ . In the line graph, that first edge is the new port. The remaining edges of its pendant path give a pendant path one edge shorter, attached at that port. Edges of the old core and the first edges just added supply exactly $L(H_j^+)$; there are no other adjacencies between different tails.

The new ports correspond to distinct edges. A newly attached leaf edge shares an endpoint with exactly the $r - 1$ old core edges at its old port. Its other endpoint is a leaf in H_j^+ . Thus its vertex in $L(H_j^+)$ has degree $r - 1$. This remains true if the tail had length one and is now exhausted. $\square$

Write

$$R_j = \sum_i (\ell_i - j)_+, \quad B(H) = \sum_{v \in V(H)} \binom{\deg_H(v)}{2} - |E(H)|.$$

Each attached tail has equally many new vertices and edges. Consequently,

$$a_j(S) = |V(H_j)| + R_j, \quad a_{j+1}(S) = |E(H_j)| + R_j. \quad (1)$$

Lemma 4 (Inversion identity). *For every $j \geq 1$,*

$$s_j = \frac{a_{j+2}(S) - a_{j+1}(S) - B(H_j)}{r - 2}. \quad (2)$$

Proof. For every simple graph F ,

$$|V(L(F))| = |E(F)|, \quad |E(L(F))| = \sum_{v \in V(F)} \binom{\deg_F(v)}{2}.$$

Adding a leaf at a port of core degree $r - 1$ increases the latter sum by $\binom{r}{2} - \binom{r-1}{2} = r - 1$. The new leaf contributes zero. Hence

$$|V(H_{j+1})| = |E(H_j)| + s_j, \quad (3)$$

$$|E(H_{j+1})| = \sum_{v \in V(H_j)} \binom{\deg_{H_j}(v)}{2} + (r - 1)s_j. \quad (4)$$

Apply (1) at index $j + 1$ and subtract:

$$a_{j+2}(S) - a_{j+1}(S) = |E(H_{j+1})| - |V(H_{j+1})| = B(H_j) + (r - 2)s_j.$$

Since $r \geq 3$, division by $r - 2$ is valid. $\square$

3 Why the inverse is determined by counts

Formula (2) would not by itself be a reconstruction algorithm if constructing H_j required the unknown assignment of leg lengths. Symmetry is the essential extra ingredient.

Lemma 5. *Every permutation of the initial labels preserving $A_1, \dots, A_{j-1}$ is induced by an automorphism of H_j on its ports. In particular, every permutation of its current port set A_{j-1} is induced by such an automorphism.*

Proof. All permutations act on $H_1 = K_r$. A permutation preserving the set of ports where a new leaf is attached extends to the graph with those leaves, by mapping the corresponding new leaves in the same way. It then induces an automorphism of the line graph and maps the new labelled ports accordingly. Induction proves the first assertion.

For the second assertion, extend a permutation of A_{j-1} by the identity on all other initial labels. As the active sets are nested, this extension preserves each earlier active set. $\square$

Thus, if $s_1, \dots, s_{j-1}$ have already been recovered, an isomorphic marked core H_j can be constructed without the individual leg lengths. After computing s_j from (2), choose *any* s_j of its current ports for the next leaf attachments. Any two such choices are related by Lemma 5, and yield isomorphic marked cores. It follows that the prefix through index $m+2$ determines $s_1, \dots, s_m$.

For completeness, the initial number of legs is also recoverable. The spider has r leaves, one vertex of degree r , and $n - r - 1$ vertices of degree two. Therefore

$$a_2(S) = n - r - 1 + \binom{r}{2}, \quad a_2(S) - n + 2 = \binom{r-1}{2}. \quad (5)$$

The last expression is strictly increasing for integers $r \geq 3$.

The inversion gives the following explicit procedure. It is a finite reconstruction algorithm; no polynomial-time bound is asserted.

1. Recover $n = a_0$ and r from (5), and form $H_1 = K_r$.
2. For $j = 1, \dots, q$, compute s_j by (2). If $j < q$, attach leaves at any s_j current ports and take the next line graph.
3. The number of legs of length exactly $j \leq q$ is $s_{j-1} - s_j$, where $s_0 = r$.
4. Determine the remaining lengths from their count and total sum, as proved in the next section.

4 The sharp cutoff

Proof of Theorem 1. The excess lengths $e_i = \ell_i - 1$ are nonnegative integers of total $s = n - r - 1$. If $q = \lfloor s/2 \rfloor = 0$, their multiset is unique: all are zero when $s = 0$, and precisely one is one when $s = 1$.

Suppose $q \geq 1$. The preceding inversion, through s_q , determines every leg of length at most q . Equivalently, it determines every excess $e_i < q$. The remaining excesses are all at least q , and both their number and their sum are known. If $q \geq 2$, there can be at most two: three would have sum at least $3q > 2q + 1 \geq s$. If two remain, their sum is either $2q$ or $2q + 1$, so their multiset is respectively $\{q, q\}$ or $\{q, q + 1\}$. If only one remains, its value is its known sum. If $q = 1$, the remaining positive integers have sum at most three; for a specified number of them their multiset is again unique. This reconstructs all leg lengths using terms through index $q + 2$.

To prove sharpness, let $s = 2q + \varepsilon \geq 2$ with $\varepsilon \in \{0, 1\}$. Consider the leg multisets

$$X = (\underbrace{1, \dots, 1}_{r-2}, q, q + 2 + \varepsilon), \quad (6)$$

$$Y = (\underbrace{1, \dots, 1}_{r-2}, q + 1, q + 1 + \varepsilon). \quad (7)$$

Their common total is $r + s$, and their multisets differ, including when $q = 1$. The counts s_j agree for $0 \leq j < q$, so their marked cores can be identified through H_q . The terms a_0, a_1, a_2 agree. Using (2) in its forward form proves agreement through index $q + 1$. At step $j = q$, the spider with lengths Y has exactly one more surviving leg. Therefore

$$a_{q+2}(S(Y)) - a_{q+2}(S(X)) = r - 2.$$

The first separating index is precisely $K = q + 2$. $\square$

These sharp pairs demonstrate the need for the last term. They do not give equal *infinite* sequences and hence are not counterexamples to Graham's conjecture.

5 Recognition of three-leaf trees

For a tree T with $n \geq 2$ and vertex degrees d_v , the handshaking identity gives

$$a_2(T) - n + 2 = \sum_v \left(\binom{d_v}{2} - d_v + 1 \right) = \sum_v \binom{d_v - 1}{2}. \quad (8)$$

Every summand is nonnegative. Degree three contributes one; degree at least four contributes at least three.

Proof of Corollary 2. Equation (8) equals one exactly when T has one degree-three vertex and all remaining degrees are one or two. This is equivalent to having exactly three leaves: the usual tree identity is

$$\#\{\text{leaves}\} = 2 + \sum_{d_v \geq 3} (d_v - 2).$$

Such a tree is a three-legged spider. Thus a tree sharing a_0, a_2 with a three-leaf tree must itself be in this class. Theorem 1 applies with $r = 3$, giving

$$K = \left\lfloor \frac{n - 4}{2} \right\rfloor + 2 = \lfloor n/2 \rfloor.$$

For $n \geq 6$, the excess $s = n - 4$ is at least two, and the sharp pairs (6)–(7) prove optimality even within this class. $\square$

For general trees, the ports of a core need not be interchangeable. Neither the symmetry lemma nor the class-recognition argument above extends merely by knowing vertex degrees. This is why the theorem is not stated for arbitrary trees.

6 An exact finite computation

Separately from the all-length proof, we enumerated all unlabelled trees of order at most twenty. The source supplement contains their canonical tree encodings and distinguishing prefixes, together with two implementations. Table 1 records k_n , the smallest index such that prefixes through that index distinguish all trees of order n in this finite computation. For $n = 1, 2, 3$ there is one tree, so $k_n = 0$; for $n = 4, 5$ it is two.

The total is 1,346,024 trees. Computed first-four-term class counts agree with the sixteen published entries of OEIS A383998 [6], and with the overlapping tables in [3].

n	Trees	k_n	n	Trees	k_n
6	6	3	14	3,159	7
7	11	3	15	7,741	7
8	23	4	16	19,320	8
9	47	4	17	48,629	8
10	106	5	18	123,867	9
11	235	5	19	317,955	9
12	551	6	20	823,065	10
13	1,301	7			

Table 1: Exact finite separation; k_n is an index, not a number of terms.

6.1 Enumeration and exact order formulas

The C++ implementation generates order n by adjoining a leaf to every vertex of every order- $(n-1)$ tree. Every tree is reached because it has a leaf. Duplicates are removed by the canonical parenthesis encoding rooted at a center, taking the lexicographic minimum when there are two centers. Rooted encodings sort the child encodings. This is a complete isomorphism invariant for trees.

For any simple graph G , write $d_v = \deg(v)$ and $q_{uv} = d_u + d_v - 2$ on its edges. Then

$$a_1 = |E(G)|, \quad a_2 = \sum_v \binom{d_v}{2}, \quad a_3 = \sum_{uv \in E(G)} \binom{q_{uv}}{2},$$

and

$$a_4 = \sum_v \sum_{\{u,w\} \subseteq N(v)} \binom{q_{uv} + q_{vw} - 2}{2}. \quad (9)$$

Indeed an edge of $L(G)$ corresponds to a pair of edges incident at v ; its degree as a vertex of $L^2(G)$ is $q_{uv} + q_{vw} - 2$. Applying the line-graph edge-count identity proves (9).

For the C++ calculation, put $S_v = \sum_{u \in N(v)} q_{uv}$ and $Q_v = \sum_{u \in N(v)} q_{uv}^2$. Expanding (9) gives

$$2a_4 = \sum_v \left((d_v - 2)Q_v + S_v^2 - 5(d_v - 1)S_v + 6\binom{d_v}{2} \right).$$

These formulas compute five terms before explicit line-graph construction. Only unresolved prefix classes are extended. At each further step, one line graph is formed and its fourth-order formula supplies the next term. Already distinguished classes need not be extended.

6.2 Replay and limitations

The second implementation enumerates trees using the WROM generator in NetworkX 3.5 [7]. It uses eccentricity-based centers, the unaggregated wedge sum (9), and NetworkX's line-graph implementation. It checks the exact canonical tree sets, recomputes every certified term using Python integers, and verifies that no certificate is a prefix of another. The C++ arithmetic is integral, using signed 128-bit integers; the largest certified term in the retained run is 50,858,095,643. The code is not a proof-assistant formalization.

The spider inversion was also checked on 351 spiders with $3 \leq r \leq 8$ and $0 \leq s \leq 8$: 995 decoding steps and 42 sharp pairs passed. In these tests, the next active ports are selected solely

from their decoded number, without consulting the unknown leg lengths. These bounded checks support the implementation; the proof of Theorem 1 is the induction and symmetry argument, not enumeration.

The value $k_{13} = 7 > \lfloor 13/2 \rfloor$ also shows that the three-leaf cutoff cannot simply be extended to all trees. The certificates contain the two trees with the common prefix

$$(13, 12, 14, 24, 68, 347, 3396)$$

and next terms 65984 and 66018. They separate at that next term; once again this is not an infinite-sequence counterexample.

The computational conclusion is restricted to $n \leq 20$. It does not establish a universal cutoff or a resolution of the conjecture. The source supplement documents dependencies, commands, all certificates and exact replay receipts so that this finite claim can be checked independently.

This is an unrefereed preprint. AI assistance was used in developing the arguments, writing the manuscript and implementing the checks. The author is responsible for the claims. The software checks are distinct from independent peer review.

References

- [1] Open Problem Garden, *Graham’s conjecture on tree reconstruction*, posted 18 March 2007, consulted 17 October 2026. https://www.openproblemgarden.org/op/grahams_conjecture_on_tree_reconstruction.
- [2] J. Cooper, B. Kay and A. Swifton, Graham’s tree reconstruction conjecture and a Waring-type problem on partitions, *J. Combin.* **9** (2018), 469–488. doi:10.4310/JOC.2018.v9.n3.a3.
- [3] K. Weatherspoon and D. Zeilberger, *An experimental note on Graham’s tree reconstruction conjecture*, 2025, author-hosted manuscript and computational supplement. <https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/treeline.html>.
- [4] B. Balch, Iterated line graphs on bi-regular graphs and trees, *Rose-Hulman Undergraduate Mathematics Journal* **21** (2020), no. 1, Article 2. <https://scholar.rose-hulman.edu/rhumj/vol21/iss1/2/>.
- [5] Z. King, L. Lane-Harvard and T. Milligan, Counting vertices in iterated line graphs, in *Combinatorics, Graph Theory and Computing*, Springer Proceedings in Mathematics & Statistics **448** (2024), 305–323. doi:10.1007/978-3-031-52969-6_27.
- [6] K. Weatherspoon, OEIS A383998, *Number of distinct truncated Graham sequences of length 4 on trees of order n*, 2025. <https://oeis.org/A383998>.
- [7] NetworkX developers, *Nonisomorphic trees*, NetworkX 3.5, function implementation and documentation. Official function documentation.