Caterpillars and alternating paths

Rain Jiang ☞  Kai Jiang ☞  Minghui Jiang ☞
Home School, USA

Abstract

Let \( p(m) \) (respectively, \( q(m) \)) be the maximum number \( k \) such that any tree with \( m \) edges can be transformed by contracting edges (respectively, by removing vertices) into a caterpillar with \( k \) edges. We derive closed-form expressions for \( p(m) \) and \( q(m) \) for all \( m \geq 1 \). The two functions \( p(n) \) and \( q(n) \) can also be interpreted in terms of alternating paths among \( n \) disjoint line segments in the plane, whose \( 2n \) endpoints are in convex position.

1 Introduction

For any family \( S \) of disjoint line segments in the plane, an alternating path among \( S \) is a simple polygonal chain in which every other segment is in \( S \). An alternating path among \( S \) is compatible with \( S \) if the chain does not cross any segment in \( S \) that is not in the chain.

For \( n \geq 1 \), let \( \hat{p}(n) \) (respectively, \( \hat{q}(n) \)) be the maximum number \( k \) such that any family \( S \) of \( n \) disjoint segments in the plane admits an alternating path among \( S \) (respectively, compatible with \( S \)) going through \( k \) segments in \( S \).

Urrutia [4] showed that \( \hat{p}(n) = O(n^{1/2}) \) and \( \hat{q}(n) = O(\log n) \), and conjectured that \( \hat{p}(n) = \Theta(n^{1/2}) \) and \( \hat{q}(n) = \Theta(\log n) \). Subsequently, Hoffmann and Tóth [2] proved that, indeed, \( \hat{q}(n) = \Omega(\log n) \), and Pach and Pinchasi [3] proved that \( \hat{p}(n) = \Omega(n^{1/3}) \). The follow-up work [2, 3] also refined Urrutia’s initial upper bounds.

Refer to Figure 1. Hoffmann and Tóth [2] refined the upper bound of \( \hat{q}(n) = O(\log n) \) to \( \hat{q}(n) \leq 6 \log_3 n + O(1) \) for all \( n \), by constructing for every \( k \geq 1 \) a family of \( 3(3^k - 1)/2 \) segments such that any alternating path compatible with these segments can go through at most \( 6k - 3 \) segments. Pach and Pinchasi [3] refined the upper bound of \( \hat{p}(n) = O(n^{1/2}) \) to \( \hat{p}(n) \leq \sqrt{8n} - 2 \) for \( n = 2k^2 \), by constructing for every \( k \geq 1 \) a family of \( 2k^2 \) segments such that any alternating path among these segments can go through at most \( 4k - 2 \) segments; they remarked that “it seems likely that the order of magnitude of this bound is not far from optimal.”

In both constructions [2, 3], as illustrated in Figure 1, the endpoints of the segments are in convex position. It is natural to ask whether these upper bounds [2, 3] are close to optimal at least in this restricted setting.

For \( n \geq 1 \), let \( p(n) \) (respectively, \( q(n) \)) be the maximum number \( k \) such that any family \( S \) of \( n \) disjoint segments in the plane, whose \( 2n \) endpoints are in convex position, admits an alternating path among \( S \) (respectively, compatible with \( S \)) going through \( k \) segments in \( S \).

We clearly have \( \hat{p}(n) \leq p(n) \) and \( \hat{q}(n) \leq q(n) \) for all \( n \geq 1 \). Effectively, Pach and Pinchasi [3] proved that \( p(n) \leq \sqrt{8n} - 2 \) for \( n = 2k^2 \), and Hoffmann and Tóth [2] proved that \( q(n) \leq 6 \log_3 n + O(1) \) for all \( n \geq 1 \).
Our two theorems in the following imply that the upper bound on $p(n)$ by Pach and Pinchasi [3] is indeed optimal for $n = 2k^2$, and that the upper bound on $q(n)$ by Hoffmann and Tóth [2] is best possible apart from an additive constant for all $n \geq 1$.

**Theorem 1.** For $n \geq 1$, $p(n) = \lceil \sqrt{8n} - 2 \rceil$.

**Theorem 2.** For $1 \leq n \leq 170$,

| $n$  | $q_0(n)$ | $q_1(n)$ | $q_2(n)$ | $q_3(n)$ | $q_4(n)$ | $q_5(n)$ |
|------|----------|----------|----------|----------|----------|----------|
| $[1, 4]$ | $[5, 6]$ | $[7, 8]$ | $[9, 10]$ | $[11, 12]$ | $[13, 15]$ | $[16, 20]$ | $[21, 25]$ | $[26, 30]$ |
| $n$  | $13$ | $14$ | $15$ | $16$ | $17$ | $18$ | $19$ | $20$ | $21$ |
| $q(n)$ | $[31, 35]$ | $[36, 44]$ | $[45, 55]$ | $[56, 66]$ | $[67, 80]$ | $[81, 96]$ | $[97, 115]$ | $[116, 138]$ | $[139, 170]$ |

For $n \geq 171$, $q(n) = \max \{ q_r(n) \mid 0 \leq r \leq 5 \}$, where

- $q_0(n) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{2n}{55} + \frac{1}{11} \right) + 11 \right] \right\rfloor$
- $q_1(n) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{n}{33} + \frac{1}{11} \right) + 11 \right] \right\rfloor + 1$
- $q_2(n) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{2n}{235} + \frac{1}{47} \right) + 17 \right] \right\rfloor + 2$
- $q_3(n) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{n}{141} + \frac{1}{47} \right) + 17 \right] \right\rfloor + 3$
- $q_4(n) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{2n}{115} + \frac{1}{23} \right) + 11 \right] \right\rfloor + 4$
- $q_5(n) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{n}{69} + \frac{1}{23} \right) + 11 \right] \right\rfloor + 5$.

Recall that “a caterpillar is a tree which metamorphoses into a path when its cocoon of endpoints is removed” [1]. In other words, a caterpillar is a tree whose non-leaf vertices induce a path. Our proofs of Theorem 1 and Theorem 2 use the following equivalent characterizations of $p(n)$ and $q(n)$ in terms of...
caterpillars and trees. From this perspective, the closed-form expressions for \( p(n) \) and \( q(n) \) that we obtained may be of interest beyond the topic of alternating paths through disjoint segments.

**Proposition 1.** For \( m \geq 1 \), \( p(m) \) (respectively, \( q(m) \)) is the maximum number \( k \) such that any tree with \( m \) edges can be transformed into a caterpillar with \( k \) edges by contracting edges (respectively, by removing vertices).

Throughout the paper, the *size* of a graph refers to the number of edges in it.

## 2 Connection between alternating paths and caterpillars in trees

In this section we prove Proposition 1 by showing that the two functions \( p(n) \) and \( q(n) \) for \( n \geq 1 \) can be interpreted in terms of caterpillars and trees.

Refer to Figure 2. Let \( S \) be a family of \( n \) disjoint segments in the plane, whose \( 2n \) endpoints are in convex position. Then there exists a convex body \( C(S) \) with the \( 2n \) endpoints on its boundary, and with the
segments as chords, which partition \( C(S) \) into \( n+1 \) cells. The intersection graph with \( n+1 \) vertices for
the \( n+1 \) cells, and with two vertices connected by an edge if and only if the corresponding cells have one
of the \( n \) segments as their shared boundary, is a tree, which we denote by \( T(S) \).

Conversely, for any tree \( T \) with \( n \) edges, it is easy to construct a family \( S \) of \( n \) disjoint segments as
chords of a convex body \( C \) such that the intersection graph of the resulting \( n+1 \) cells of \( C \) is \( T \).

The following lemma shows that alternating paths compatible with \( S \) are equivalent to caterpillars obtained
from \( T(S) \) by removing vertices:

**Lemma 1.** \( S \) admits an alternating path compatible with \( S \) going through \( k \) segments if and only if \( T(S) \)
contains a caterpillar with \( k \) edges as a subgraph.

**Proof.** The lemma is trivially true when \( k = 1 \). Now assume that \( k \geq 2 \). We first prove the direct implication. Suppose that there is an alternating path compatible with \( S \) that goes through a sequence \( S' \) of \( k \) segments in \( S \). Since every segment bordering two adjacent cells of \( C(S) \) corresponds to an edge between two vertices in \( T(S) \), and every two consecutive segments in \( S' \) are on the boundary of some common cell, it follows that every maximal subsequence of consecutive segments in \( S' \) around a common cell in \( C(S) \) corresponds to a star of edges incident to a common vertex in \( T(S) \). Moreover, every three consecutive segments in \( S' \) not around a common cell must span two adjacent cells with the middle segment as their shared boundary, which corresponds to an edge connecting two stars in \( T(S) \). Thus the \( k \) segments in \( S' \) correspond to a path of stars in \( T(S) \), which is a caterpillar with \( k \) edges.

We next prove the reverse implication. Suppose that \( T(S) \) contains a caterpillar with \( k \) edges, whose path
of non-leaf vertices corresponds to a sequence \( C \) of cells of \( C(S) \). Then the edges of the caterpillar incident
to each non-leaf vertex correspond to a subfamily of segments in \( S \) on the boundary of some common cell
in \( C \), and each edge in the non-leaf path corresponds to a segment in \( S \) bordering two adjacent cells in \( C \).
Since the cells of \( C(S) \) are all convex, the disjoint segments on the boundary of each cell can be linked into
a non-crossing path within the cell, starting and ending at any two endpoints of any two segments. Then
along the sequence \( C \) of cells, these paths can be concatenated into a single alternating path compatible with
\( S \).

The following lemma shows another equivalence between alternating paths among \( S \) and caterpillars
obtained from \( T \) by contracting edges:

**Lemma 2.** \( S \) admits an alternating path among \( S \) going through \( k \) segments if and only if a caterpillar with
\( k \) edges can be obtained from \( T(S) \) by contracting edges.

**Proof.** The deletion of a segment in \( S \) corresponds to the merging of two adjacent cells of \( C(S) \), and to
the contraction of an edge in \( T(S) \). For any subfamily \( S' \) of \( k \) segments in \( S \), any alternating path among
\( S \) that goes through all \( k \) segments of \( S' \) is an alternating path compatible with \( S' \), and conversely, any
alternating path compatible with \( S' \) is an alternating path among \( S \). By Lemma 1, \( S' \) admits an alternating path compatible with \( S' \) going through all \( k \) segments if and only if \( T(S') \) is a caterpillar with \( k \) edges.

Recall that for \( n \geq 1 \), we defined \( p(n) \) (respectively, \( q(n) \)) as the maximum number \( k \) such that any family \( S \) of \( n \) disjoint segments in the plane whose endpoints are in convex position admits an alternating path among \( S \) (respectively, compatible with \( S \)) going through \( k \) segments in \( S \). By Lemma 1 and Lemma 2, it follows that for \( m \geq 1 \), \( p(m) \) (respectively, \( q(m) \)) is the maximum number \( k \) such that any tree with \( m \) edges can be transformed into a caterpillar with \( k \) edges by contracting edges (respectively, by removing vertices). This completes the proof of Proposition 1.

In the next two sections, we prove Theorem 1 and Theorem 2 in terms of caterpillars and trees. By
convention, we will use the symbol \( m \) to denote the number of edges in a graph, and determine \( p(m) \) and
\( q(m) \) for \( m \geq 1 \).
3 Contracting a tree into a caterpillar

In this section we prove Theorem 1 that for \( m \geq 1 \), \( p(m) = \lceil \sqrt{8m - 2} \rceil \). Recall that \( p(m) \) is the maximum number \( k \) such that any tree with \( m \) edges can be transformed into a caterpillar with \( k \) edges by contracting edges. We first prove three lemmas.

For any tree \( T \), denote by \( \kappa(T) \) the number of leaves in \( T \), plus the diameter of \( T \), minus 2.

**Lemma 3.** For any tree \( T \), the maximum size of a caterpillar that can be obtained from \( T \) by contracting edges is \( \kappa(T) \). Moreover, for any \( 1 \leq k \leq \kappa(T) \), a caterpillar of size \( k \) can be obtained from \( T \) by contracting edges.

**Proof.** When \( \kappa(T) = 1 \), \( T \) must be a tree with a single edge. The lemma obviously holds in this case. Now assume that \( \kappa(T) \geq 2 \). Let \( T' \) be a caterpillar of the maximum size obtained from a tree \( T \) by contracting edges. By the maximality of \( T' \), each leaf of \( T' \) must be a leaf of \( T \). Moreover, the edges of the non-leaf path of \( T' \), if any, must come from some non-leaf path in \( T \), whose length is at most the diameter of \( T \) minus 2. Thus the number of edges in \( T' \) is at most the number of leaves in \( T \) plus the diameter of \( T \) minus 2, which is \( \kappa(T) \). This number is actually attainable: simply keep all leaf edges and all edges of a diameter path, and contract the rest. From such a caterpillar with \( \kappa(T) \) edges, a caterpillar with \( k < \kappa(T) \) edges can then be obtained by contracting any \( \kappa(T) - k \) edges.

For \( k \geq 1 \), denote by \( e(k) \) the maximum size of a tree \( T \) with \( \kappa(T) = k \). For \( d \geq 2 \) and \( l \geq 2 \), denote by \( e(d, l) \) the maximum size of a tree with diameter \( d \) and with \( l \) leaves.

A spider is a tree with at most one vertex of degree greater than two. In the next two lemmas, we determine \( e(d, l) \) and \( e(k) \) by showing that the extremal cases are realized by spiders.

**Lemma 4.** For \( d \geq 2 \) and \( l \geq 2 \), \( e(d, l) = \frac{1}{2} \cdot d \cdot l \) when \( d \) is even, \( e(d, l) = \frac{1}{2} \cdot (d - 1) \cdot l + 1 \) when \( d \) is odd, and there exists a spider \( R_{d,l} \) with diameter \( d \), \( l \) leaves, and \( e(d, l) \) edges.

**Proof.** Consider two cases for any tree with diameter \( d \), \( l \) leaves, and \( m \) edges:

1. \( d \) is even. Then there is an internal node \( v \) in the tree whose distance to any other node is at most \( d/2 \).
   
   The number \( m \) of edges in the tree is maximized when the paths from \( v \) to the \( l \) leaves all have the same length \( d/2 \) and are edge-disjoint, that is, when the tree is a spider \( R_{d,l} \) centered at \( v \) with \( l \) legs of equal length \( d/2 \). Thus \( m \leq \frac{1}{2} \cdot d \cdot l \).

2. \( d \) is odd. Then there is an edge \( \{u, v\} \) between two internal nodes \( u \) and \( v \) in the tree such that the distance from this edge to any other node is at most \( (d - 1)/2 \).
   
   The number \( m \) of edges in the tree is maximized when the shortest paths from \( \{u, v\} \) to the \( l \) leaves all have the same length \( (d - 1)/2 \) and are edge-disjoint. Thus \( m \leq \frac{1}{2} \cdot (d - 1) \cdot l + 1 \). There are multiple extremal cases when \( l > 2 \). In particular, if one of the \( l \) leaves is on one side of the edge \( \{u, v\} \), and the other \( l - 1 \) leaves are all on the other side, then the tree is a spider \( R_{d,l} \) with one leg of length \( (d + 1)/2 \) and \( l - 1 \) legs of length \( (d - 1)/2 \).

**Lemma 5.** For \( k \geq 1 \), \( e(k) = \frac{1}{4} k(k + 4) \) when \( k \) mod 4 = 0, \( e(k) = \frac{1}{8} (k + 2)^2 \) when \( k \) mod 4 = 2, \( e(k) = \frac{1}{8} (k + 1)(k + 3) \) when \( k \) mod 2 = 1, and there exists a spider \( R_k \) with \( \kappa(R_k) = k \) and with \( e(k) \) edges. In particular, \( \frac{1}{8} (k + 2)^2 - \frac{1}{7} \leq e(k) \leq \frac{1}{8} (k + 2)^2 \), and \( e(k) \) is strictly increasing for \( k \geq 1 \).

**Proof.** The only tree \( T \) with \( \kappa(T) = 1 \) is \( P_2 \), a path with a single edge. Thus \( e(1) = 1 \), which equals \( \frac{1}{8} (1 + 1)(1 + 3) \), and we have \( R_1 = P_2 \). Also, the only tree \( T \) with \( \kappa(T) = 2 \) is \( P_3 \), a path with two edges. Thus \( e(2) = 2 \), which equals \( \frac{1}{8} (2 + 2)^2 \), and we have \( R_2 = P_3 \). For \( k \geq 3 \), we have

\[
e(k) = \max\{ e(d, l) \mid d \geq 2, l \geq 2, l + d - 2 = k \}.
\]
To determine \( e(k) \), we evaluate the maxima of \( e(d, l) \) for even \( d \) and for odd \( d \) separately, and apply Lemma 4. Consider four cases:

- \( k \mod 4 = 2 \). Then
  
  \[
  \max_{d \text{ even}} e(d, l) = e\left( \frac{k+2}{2}, \frac{k+2}{2} \right) = \frac{(k+2)^2}{8},
  \]
  
  \[
  \max_{d \text{ odd}} e(d, l) = e\left( \frac{k+4}{2}, \frac{k}{2} \right) = \frac{k(k+2)}{8} + 1.
  \]

  Since \( \frac{1}{8}(k+2)^2 - \frac{1}{8}k(k+2) - 1 = \frac{1}{4}(k+2) - 1 \geq 1 \) for \( k \geq 6 \), we have \( e(k) = \frac{1}{8}(k+2)^2 \).

- \( k \mod 4 = 0 \). Then
  
  \[
  \max_{d \text{ even}} e(d, l) = e\left( \frac{k}{2}, \frac{k+4}{2} \right) = e\left( \frac{k+4}{2}, \frac{k}{2} \right) = \frac{k(k+4)}{8},
  \]
  
  \[
  \max_{d \text{ odd}} e(d, l) = e\left( \frac{k+2}{2}, \frac{k+2}{2} \right) = \frac{k(k+2)}{8} + 1.
  \]

  Since \( \frac{1}{8}k(k+4) - \frac{1}{8}k(k+2) - 1 = \frac{1}{8}k - 1 \geq 0 \) for \( k \geq 4 \), we have \( e(k) = \frac{1}{8}k(k+4) = \frac{1}{8}(k+2)^2 - \frac{1}{2} \).

- \( k \mod 4 = 3 \). Then
  
  \[
  \max_{d \text{ even}} e(d, l) = e\left( \frac{k+1}{2}, \frac{k+3}{2} \right) = \frac{(k+1)(k+3)}{8},
  \]
  
  \[
  \max_{d \text{ odd}} e(d, l) = e\left( \frac{k+3}{2}, \frac{k+1}{2} \right) = \frac{(k+1)^2}{8} + 1.
  \]

  Since \( \frac{1}{8}(k+1)(k+3) - \frac{1}{8}(k+1)^2 - 1 = \frac{4}{8}(k+1) - 1 \geq 0 \) for \( k \geq 3 \), we have \( e(k) = \frac{1}{8}(k+1)(k+3) = \frac{1}{8}(k+2)^2 - \frac{1}{8} \).

- \( k \mod 4 = 1 \). Then
  
  \[
  \max_{d \text{ even}} e(d, l) = e\left( \frac{k+3}{2}, \frac{k+1}{2} \right) = \frac{(k+1)(k+3)}{8},
  \]
  
  \[
  \max_{d \text{ odd}} e(d, l) = e\left( \frac{k+1}{2}, \frac{k+3}{2} \right) = e\left( \frac{k+5}{2}, \frac{k-1}{2} \right) = \frac{(k-1)(k+3)}{8} + 1.
  \]

  Since \( \frac{1}{8}(k+1)(k+3) - \frac{1}{8}(k-1)(k+3) - 1 = \frac{1}{8}(k+3) - 1 \geq 1 \) for \( k \geq 5 \), we again have \( e(k) = \frac{1}{8}(k+1)(k+3) = \frac{1}{8}(k+2)^2 - \frac{1}{8} \).

The last two cases \( k \mod 4 = 3 \) and \( k \mod 4 = 1 \) can be combined into a single case \( k \mod 2 = 1 \). In any case, there exists by Lemma 4 a spider \( R_k = R_{d,1} \) with \( \kappa(R_k) = \kappa(R_{d,1}) = l + d - 2 = k \) and with \( e(k) = e(d, l) \) edges, where \( d \) is even. In particular, we can let \( R_3 = R_{2,3}, R_4 = R_{2,4}, R_5 = R_{4,3}, R_6 = R_{4,4}, R_7 = R_{4,5}, R_8 = R_{4,6}, R_9 = R_{6,5}, R_{10} = R_{6,6}, R_{11} = R_{6,7}, R_{12} = R_{6,8} \). Refer to Figure 3 for illustrations of \( R_k \) for \( 1 \leq k \leq 12 \).

Note that \( \frac{1}{8}(k+2)^2 - \frac{1}{2} \leq e(k) \leq \frac{1}{8}(k+2)^2 \) for all \( k \geq 1 \). It follows that

\[
\frac{1}{8}(k+3)^2 - \frac{1}{2} - \frac{1}{8}(k+2)^2 = \frac{1}{8}(2k+5) - \frac{1}{2} > 0.
\]

Thus \( e(k) \) is strictly increasing for \( k \geq 1 \). \( \square \)
We are now ready to prove that \( p(m) = \lceil \sqrt{8m - 2} \rceil \) for \( m \geq 1 \). For \( m = 1 \), we clearly have \( p(1) = \lceil \sqrt{8 - 2} \rceil = 1 \). Now fix \( m \geq 2 \), and let \( k = \lceil \sqrt{8m - 2} \rceil \). Then \( k \geq 2 \), and \( \sqrt{8m - 2} \leq k < \sqrt{8m - 1} \).

From \( k \geq \sqrt{8m - 2} \), it follows that \( e(k) \geq \frac{1}{8}(k + 2)^2 - \frac{1}{2} \geq \frac{1}{8} \cdot (\sqrt{8m - 2} + 2)^2 - \frac{1}{2} = m - \frac{1}{2} \), and hence \( e(k) \geq m \) since \( e(k) \) is an integer. From \( k < \sqrt{8m - 1} \), it follows that \( e(k - 1) \leq \frac{1}{8}(k + 1)^2 < \frac{1}{8} \cdot (\sqrt{8m - 1} + 1)^2 = m \). In summary, we have \( e(k - 1) < m \leq e(k) \).

By Lemma 5, \( e(k) \) is strictly increasing for \( k \geq 1 \). Thus the inequality \( e(k - 1) < m \) implies that any tree with \( \kappa(T) \leq k - 1 \) must have fewer than \( m \) edges. In other words, any tree \( T \) with \( m \) edges must have \( \kappa(T) \geq k \), and hence can be transformed by contracting edges into a caterpillar with \( k \) edges, by Lemma 3.

On the other hand, corresponding to the extremal cases of Lemma 4 and Lemma 5, there exists a spider \( R_k \) with \( \kappa(R_k) = k \) and with \( e(k) \) edges. Since \( \kappa(R_k) = k \), it follows by Lemma 3 that \( R_k \) cannot be transformed by contracting edges into a caterpillar with more than \( k \) edges. By the inequality \( m \leq e(k) \), there exists a subgraph of \( R_k \) with \( m \) edges which cannot be thus transformed either. Thus \( p(m) = k = \lceil \sqrt{8m - 2} \rceil \). This completes the proof of Theorem 1.

### 4 Pruning a tree into a caterpillar

In this section we prove Theorem 1. Recall that \( q(m) \) is the maximum number \( k \) such that any tree with \( m \) edges can be transformed into a caterpillar with \( k \) edges by removing vertices. In other words, \( q(m) \) is the maximum number \( k \) such that any tree with \( m \) edges contains a caterpillar with \( k \) edges as a vertex-induced subgraph.

For \( k \geq 1 \), let \( e(k) \) be the maximum size \( m \) of a tree in which the maximum size of a caterpillar is exactly \( k \). To derive \( q(m) \) for \( m \geq 1 \), we first determine the exact values of \( e(k) \) for \( k \geq 1 \).
4.1 Single branch: $f(k)$

In a rooted tree, we call any caterpillar consisting of all edges incident to a path from the root to a leaf a very hungry caterpillar. For $k \geq 1$, let $f(k)$ be the maximum size of a rooted tree in which the root is incident to a single edge, and the maximum size of a very hungry caterpillar is $k$.

To determine $f(k)$, we only need to consider symmetric trees in which all nodes of the same depth have the same number of children. Let $T$ be a rooted tree in which the root is incident to a single edge, and the maximum size of a very hungry caterpillar is $k$. Suppose that $T$ is not symmetric. Let $d \geq 0$ be the minimum depth at which there are nodes of different degrees in $T$. Let $v$ be a node at depth $d$ such that the subtree of $T$ rooted at $v$ has the maximum size. By replacing every subtree rooted at a node at depth $d$ by a copy of the subtree rooted at $v$, we either make $T$ symmetric, or increase the minimum depth at which there are nodes of different degrees in $T$. Moreover, the size of $T$ does not decrease, and the maximum size of a very hungry caterpillar in $T$ does not increase. By repeating such replacements, we can make $T$ symmetric. Then, by splitting the single edge incident to the root of $T$ into a path if necessary, we can ensure that the size of $T$ does not decrease, and the maximum size of a very hungry caterpillar in $T$ is exactly $k$.

Let $T$ be a symmetric tree of height $h \geq 1$ in which the root is incident to a single edge. For $0 \leq d \leq h$, let $c_d$ be the number of children of each node with depth $d$ in $T$, with $c_0 = 1$ for the root, and $c_h = 0$ for the leaves. Then all very hungry caterpillars in $T$ have the same size $\sum_{d=0}^{h} c_d$.

We can assume that $c_1 \geq \ldots \geq c_h$ because, if $c_d < c_{d+1}$ for some $1 \leq d < h$, then we can swap the two elements $c_d$ and $c_{d+1}$ in the sequence $\langle c_0, c_1, \ldots, c_h \rangle$, such that the resulting sequence corresponds to a symmetric tree with more edges between nodes of depths $d$ and $d + 1$, and with the same number of edges in other layers, while the sum $\sum_{d=0}^{h} c_d$ remains the same.

Moreover, we can assume that $c_1 \leq 3$ because, if $c_1 \geq 4$, then we can replace the element $c_1$ in the sequence $\langle c_0, c_1, \ldots, c_h \rangle$ by two elements $c_1 - 2$ and 2, such that the resulting sequence of $h + 2$ elements corresponds to a symmetric tree with height $h + 1$ and with more edges, but still has the same sum.

In summary, we can assume that $T$ is a symmetric tree of height $h \geq 1$, with $c_0 = 1$ and $3 \geq c_1 \geq \ldots \geq c_h = 0$, and we call such trees beautiful trees. Thus for $k \geq 1$, $f(k)$ is the maximum size of a beautiful tree $T$ with $\sum_{d=0}^{h} c_d = k$.

Lemma 6. $f(1) = 1$, $f(2) = 2$, $f(3) = 3$, $f(4) = 5$, $f(5) = 7$.

- For $k = 3i$ with $i \geq 2$, $f(k) = \frac{1}{2}(23 \cdot 3^{(k-6)/3} - 1)$.
- For $k = 3i + 1$ with $i \geq 2$, $f(k) = \frac{1}{2}(33 \cdot 3^{(k-7)/3} - 1)$.
- For $k = 3i + 2$ with $i \geq 2$, $f(k) = \frac{1}{2}(47 \cdot 3^{(k-8)/3} - 1)$.

In particular, (a) for $k \geq 2$, $\frac{f(k)}{f(k-1)} \geq \frac{7}{5}$, (b) for $k \geq 7$, $\frac{f(k)}{f(k-1)} < \frac{3}{2}$. Moreover, for $k \geq 1$, there exists a beautiful tree $B_k$ with exactly $f(k)$ edges such that the maximum size of a very hungry caterpillar in $B_k$ is $k$, and the maximum size of any caterpillar in $B_k$ is less than $2k$.

Proof. Clearly, $f(1) = 1$, and $B_1$ is the tree consisting of a single edge.

For $k \geq 2$, let $T$ be a beautiful tree and height $h \geq 2$, with $\sum_{d=0}^{h} c_d = k$, where $c_0 = 1$ and $1 \leq c_1 \leq 3$. Then the sequence $\langle c_0, c_1, \ldots, c_{h-1} \rangle = \langle 1, c_2, \ldots, c_h \rangle$ corresponds to a beautiful tree $T'$ of height $h - 1$, with $\sum_{d=0}^{h-1} c_d = 1 + \sum_{d=2}^{h} c_d = k - c_1$, which is a subtree of $T$. Indeed $T$ contains $c_1$ edge-disjoint copies of $T'$ with a shared root at the lower end of the single edge from the root of $T$. If $T$ has the maximum size $f(k)$, then $T'$ must have the maximum size $f(k - c_1)$. By enumerating $c_1$, we have the recurrence

$$f(k) = \max \{ c_1 \cdot f(k - c_1) + 1 \mid 1 \leq c_1 < k \}.$$  \hspace{1cm} (1)
Moreover, since we can assume that \( c_1 \leq 3 \),
\[
f(k) = \max\{ c_1 \cdot f(k - c_1) + 1 \mid 1 \leq c_1 \leq \min\{k - 1, 3\} \}.
\]

By the recurrence, we can derive \( f(k) \) for \( k = 2, \ldots, 11 \) sequentially:
\[
\begin{align*}
f(1) &= 1, \\
f(2) &= 1 \cdot f(1) + 1 = 2, \\
f(3) &= 1 \cdot f(2) + 1 = 2 \cdot f(1) + 1 = 3, \\
f(4) &= 2 \cdot f(2) + 1 = 5, \\
f(5) &= 2 \cdot f(3) + 1 = 3 \cdot f(2) + 1 = 7, \\
f(6) &= 2 \cdot f(4) + 1 = 11, \\
f(7) &= 3 \cdot f(4) + 1 = 16, \\
f(8) &= 2 \cdot f(6) + 1 = 23, \\
f(9) &= 3 \cdot f(6) + 1 = 34, \\
f(10) &= 3 \cdot f(7) + 1 = 49, \\
f(11) &= 3 \cdot f(8) + 1 = 70.
\end{align*}
\]

Let \( \hat{c}(k) \) be the minimum \( c_1, 1 \leq c_1 \leq \min\{k - 1, 3\} \), such that \( f(k) = c_1 \cdot f(k - c_1) + 1 \). Then \( \hat{c}(k) = 1, 1, 2, 2, 2, 3, 2, 3, 3, 3 \) for \( k = 2, \ldots, 11 \). In particular, \( \hat{c}(k) = 3 \) for \( 9 \leq k \leq 11 \). Recall that a beautiful tree \( T \) with sequence \( \langle 1, c_1, c_2, \ldots, c_h \rangle \) contains \( c_1 \) copies of a beautiful tree \( T' \) with sequence \( \langle 1, c_2, \ldots, c_h \rangle \). Since \( 3 \geq c_1 \geq c_2 \), we must have \( \hat{c}(k) = 3 \), and hence \( f(k) = 3 \cdot f(k - 3) + 1 \), for \( k \geq 9 \). Thus we can determine \( f(k) \) for \( k \geq 6 \) as follows.

- For \( k = 3i \) with \( i \geq 2 \), it follows from \( f(6) = 11 \) that \( f(k) = \frac{23}{2} \cdot 3^{(k-6)/3} - \frac{1}{2} \).
• For $k = 3i + 1$ with $i \geq 2$, it follows from $f(7) = 16$ that $f(k) = \frac{33}{2} \cdot 3^{(k-7)/3} - \frac{1}{2}$.

• For $k = 3i + 2$ with $i \geq 2$, it follows from $f(8) = 23$ that $f(k) = \frac{47}{2} \cdot 3^{(k-8)/3} - \frac{1}{2}$.

Refer to Figure 4. For $k \geq 2$, let $B_k$ be the beautiful tree with a single edge incident to the root, whose lower end is the shared root of $c(k)$ copies of $B_{k'}$, where $k' = k - \hat{c}(k)$. Then the maximum size of a very hungry caterpillar in $B_k$ is exactly $k$. Moreover, since any caterpillar in $B_k$ can be covered by two very hungry caterpillars, which always contain the single edge incident to the root, the maximum size of any caterpillar in $B_k$ is at most $2k - 1$.

It remains to prove the two inequalities (a) and (b).

We first prove (a) that for $k \geq 2$, $\frac{f(k)}{f(k-1)} \geq \frac{7}{5}$. It is easy to verify this for $2 \leq k \leq 6$. For $k \geq 7$, consider three cases:

• For $k = 3i + 1$ with $i \geq 2$,

\[
\frac{f(k)}{f(k-1)} = \frac{33 \cdot 3^{(k-7)/3} - \frac{1}{2}}{23 \cdot 3^{(k-7)/3} - \frac{1}{2}} = \frac{33 \cdot 3^{(k-7)/3} - 1}{23 \cdot 3^{(k-7)/3} - 1} \geq \frac{33 \cdot 3^{(k-7)/3}}{23 \cdot 3^{(k-7)/3}} = \frac{33}{23} > \frac{7}{5}.
\]

• For $k = 3i + 2$ with $i \geq 2$,

\[
\frac{f(k)}{f(k-1)} = \frac{47 \cdot 3^{(k-8)/3} - \frac{1}{2}}{33 \cdot 3^{(k-8)/3} - \frac{1}{2}} = \frac{47 \cdot 3^{(k-8)/3} - 1}{33 \cdot 3^{(k-8)/3} - 1} \geq \frac{47 \cdot 3^{(k-8)/3}}{33 \cdot 3^{(k-8)/3}} = \frac{47}{33} > \frac{7}{5}.
\]

• For $k = 3i$ with $i \geq 3$,

\[
\frac{f(k)}{f(k-1)} = \frac{23 \cdot 3^{(k-6)/3} - \frac{1}{2}}{17 \cdot 3^{(k-6)/3} - \frac{1}{2}} = \frac{69 \cdot 3^{(k-9)/3} - 1}{47 \cdot 3^{(k-9)/3} - 1} \geq \frac{69 \cdot 3^{(k-9)/3}}{47 \cdot 3^{(k-9)/3}} = \frac{69}{47} > \frac{7}{5}.
\]

We next prove (b) that for $k \geq 7$, $\frac{f(k)}{f(k-1)} < \frac{3}{2}$. Again consider three cases:

• For $k = 3i + 1$ with $i \geq 2$,

\[
\frac{f(k)}{f(k-1)} = \frac{33 \cdot 3^{(k-7)/3} - 1}{23 \cdot 3^{(k-7)/3} - 1} \leq \frac{33 \cdot 3^{(k-7)/3} - 3^{(k-7)/3}}{23 \cdot 3^{(k-7)/3} - 3^{(k-7)/3}} = \frac{33 - 1}{23 - 1} < \frac{3}{2}.
\]

• For $k = 3i + 2$ with $i \geq 2$,

\[
\frac{f(k)}{f(k-1)} = \frac{47 \cdot 3^{(k-8)/3} - 1}{33 \cdot 3^{(k-8)/3} - 1} \leq \frac{47 \cdot 3^{(k-8)/3} - 3^{(k-8)/3}}{33 \cdot 3^{(k-8)/3} - 3^{(k-8)/3}} = \frac{47 - 1}{33 - 1} < \frac{3}{2}.
\]

• For $k = 3i$ with $i \geq 3$,

\[
\frac{f(k)}{f(k-1)} = \frac{69 \cdot 3^{(k-9)/3} - 1}{47 \cdot 3^{(k-9)/3} - 1} \leq \frac{69 \cdot 3^{(k-9)/3} - 3^{(k-9)/3}}{47 \cdot 3^{(k-9)/3} - 3^{(k-9)/3}} = \frac{69 - 1}{47 - 1} < \frac{3}{2}.
\]
4.2 Multiple branches: \( g(k) \)

Let \( T \) be a tree with at least two edges, and let \( C \) be a caterpillar with the maximum number \( k \geq 2 \) edges in \( T \). Then there exists a non-leaf vertex \( v \) in \( C \), of the same degree \( r \geq 2 \) in both \( C \) and \( T \), that connects \( r \) edge-disjoint parts of \( C \), including \( r - 2 \) single edges, and two smaller caterpillars of \( x \geq 1 \) and \( y \geq 1 \) edges, where \( x + y + r - 2 = k \), and \( x \leq y \leq \lceil k/2 \rceil \). These \( r \) parts of \( C \) are in \( r \) subtrees of \( T \), respectively; in particular, the two smaller caterpillars are very hungry caterpillars in two subtrees of \( T \) rooted at \( v \). The size of the subtree containing the very hungry caterpillar of size \( x \) is at most \( f(x) \). The size of the subtree containing the very hungry caterpillar of size \( y \) is at most \( f(y) \). Moreover, each of the other \( r - 2 \) subtrees can have size at most \( f(x) \), because otherwise we can replace the caterpillar of size \( x \) by a larger caterpillar. Thus the size of \( T \) is at most \( f(y) + (r - 1) \cdot f(x) \).

For \( k \geq r \geq 2 \), define

\[
g(r, k) = \max \{ f(y) + (r - 1) \cdot f(x) \mid x + y + r - 2 = k, \ 1 \leq x \leq y \leq \lceil k/2 \rceil \}.
\]

For \( k \geq 2 \), define

\[
g(k) = \max \{ g(r, k) \mid 2 \leq r \leq k \}.
\]

Then \( g(k) \) is an upper bound on \( e(k) \). The following lemma shows that \( e(k) = g(k) \) for \( k \geq 2 \).

**Lemma 7.** \( g(2) = 2, g(3) = 3, g(4) = 4, g(5) = 6, g(6) = 8, g(7) = 10, g(8) = 12, g(9) = 15, g(10) = 20, g(11) = 25, g(12) = 30, g(13) = 35, g(14) = 44. \)

- For \( k = 2i \) with \( i \geq 8 \), \( g(k) = 6f\left(\frac{k-4}{2}\right) \).
- For \( k = 2i + 1 \) with \( i \geq 7 \), \( g(k) = 5f\left(\frac{k-3}{2}\right) \).

Moreover, for \( k \geq 2 \), there exists a tree \( T_k \) with exactly \( g(k) \) edges such that the maximum size of a caterpillar in \( T_k \) is \( k \).

**Proof.** Write \( h(r, x, y) = f(y) + (r - 1) \cdot f(x) \). Then

\[
g(k) = \max \{ h(r, x, y) \mid 2 \leq r \leq k, \ x + y + r - 2 = k, \ 1 \leq x \leq y \leq \lceil k/2 \rceil \}.
\]

With this formula, we can determine \( g(k) \) for small values of \( k \) without much difficulty. Incidentally, for each \( k = 2, \ldots, 14 \), \( g(k) \) can be realized by \( h(r, x, y) \) for some tuple \((r, x, y)\) with \( x = y \):

- \( g(2) = h(2, 1, 1) = 2 \cdot f(1) = 2, \)
- \( g(3) = h(3, 1, 1) = 3 \cdot f(1) = 3, \)
- \( g(4) = h(4, 1, 1) = 4 \cdot f(1) = 4, \)
- \( g(5) = h(3, 2, 2) = 3 \cdot f(2) = 6, \)
- \( g(6) = h(4, 2, 2) = 4 \cdot f(2) = 8, \)
- \( g(7) = h(5, 2, 2) = 5 \cdot f(2) = 10, \)
- \( g(8) = h(4, 3, 3) = 4 \cdot f(3) = 12, \)
- \( g(9) = h(3, 4, 4) = 3 \cdot f(4) = 15, \)
- \( g(10) = h(4, 4, 4) = 4 \cdot f(4) = 20, \)
- \( g(11) = h(5, 4, 4) = 5 \cdot f(4) = 25, \)
- \( g(12) = h(6, 4, 4) = 6 \cdot f(4) = 30, \)
- \( g(13) = h(5, 5, 5) = 5 \cdot f(5) = 35, \)
- \( g(14) = h(4, 6, 6) = 4 \cdot f(6) = 44. \)
In the following, we assume that $k \geq 15$. Note that $5f(\frac{k-3}{2}) = h(5, \frac{k-3}{2}, \frac{k-3}{2})$ and $6f(\frac{k-4}{2}) = h(6, \frac{k-4}{2}, \frac{k-4}{2})$. To find the maximum value of $h(r, x, y)$ for $2 \leq r \leq k$, $x + y + r - 2 = k$, and $1 \leq x \leq y \leq \lfloor k/2 \rfloor$, we can assume that either $x = y$, or $x < y$ and $2y > k$, because if $x < y$ and $2y \leq k$, then at least one of the following two inequalities must hold:

$$h(r, x, y) \leq h(r + (y - x), x, x), \quad (2)$$  
$$h(r, x, y) \leq h(r - (y - x), y, y). \quad (3)$$

Inequality (2) is equivalent to

$$(r - 1)f(x) + f(y) \leq (r + (y - x))f(x)$$

$$f(y) \leq (y - x + 1)f(x)$$

$$f(x) \geq \frac{1}{y - x + 1}. \quad (4)$$

Inequality (3) is equivalent to

$$(r - 1)f(x) + f(y) \leq (r - (y - x))f(y)$$

$$(r - 1)f(x) \leq (r - 1 - (y - x))f(y)$$

$$f(x) \leq 1 - \frac{y - x}{r - 1}. \quad (5)$$

Since $x + y + r - 2 = k$, we have $r - 1 = k - x - y + 1 = k - 2y + y - x + 1$. If $2y \leq k$, then $r - 1 \geq y - x + 1$, and hence

$$1 - \frac{y - x}{r - 1} \geq 1 - \frac{y - x}{y - x + 1} = \frac{1}{y - x + 1}.$$

Thus at least one of (4) and (5) must hold. It follows that at least one of (2) and (3) must hold.

Therefore, to find the maximum value of $h(r, x, y)$ for $2 \leq r \leq k$, $x + y + r - 2 = k$, and $1 \leq x \leq y \leq \lfloor k/2 \rfloor$, we can assume that either $x = y$, or $x < y$ and $2y > k$. Consider three cases:

Case 1: $2 \leq r \leq k$, $x + y + r - 2 = k$, $1 \leq x = y \leq \lfloor k/2 \rfloor$, $k \geq 16$ is even. Then $r$ is also even, and $x = y = \frac{k-r+2}{2}$. We next show that $rf(\frac{k-r+2}{2}) \leq 6f(\frac{k-4}{2})$ for all even $r$, $2 \leq r \leq k$:

- $r = 2$. Since $k \geq 16$, we have $\frac{k}{2} \geq 8$, and it follows by Lemma 6(b) that $\frac{f(\frac{k}{2})}{f(\frac{k-4}{2})} < (\frac{3}{2})^{2} < \frac{6}{2}$. Thus $2f(\frac{k}{2}) < 6f(\frac{k-4}{2})$.

- $r = 4$. Since $k \geq 16$, we have $\frac{k-2}{2} \geq 7$, and it follows by Lemma 6(b) that $\frac{f(\frac{k-2}{2})}{f(\frac{k-4}{2})} < \frac{3}{2} = 4$. Thus $4f(\frac{k-2}{2}) < 6f(\frac{k-4}{2})$.

- $6 \leq r \leq k$. We prove by induction that $rf(\frac{k-r+2}{2}) \leq 6f(\frac{k-4}{2})$. For the base case when $r = 6$, we have $rf(\frac{k-r+2}{2}) = 6f(\frac{k-4}{2})$. For the inductive step, fix $8 \leq r \leq k$. By Lemma 6(a), we have $\frac{f(\frac{k-r+4}{2})}{f(\frac{k-r+2}{2})} \geq \frac{7}{8} > \frac{r}{r-2}$, and hence $rf(\frac{k-r+2}{2}) < (r-2)f(\frac{k-r+4}{2})$. By induction, $(r-2)f(\frac{k-r+4}{2}) \leq 6f(\frac{k-4}{2})$. Thus $rf(\frac{k-r+2}{2}) < 6f(\frac{k-4}{2})$.

Case 2: $2 \leq r \leq k$, $x + y + r - 2 = k$, $1 \leq x = y \leq \lfloor k/2 \rfloor$, $k \geq 15$ is odd. Then $r$ is also odd, and $x = y = \frac{k-r+2}{2}$. We next show that $rf(\frac{k-r+2}{2}) \leq 5f(\frac{k-3}{2})$ for all odd $r$, $3 \leq r \leq k$:
• $r = 3$. Since $k \geq 15$, we have $\frac{k+1}{2} \geq 7$, and it follows by Lemma 6(b) that $\frac{f(k+1)}{f(k/2)} < \frac{3}{2} < \frac{5}{3}$. Thus $3f(\frac{k+1}{2}) < 5f(\frac{k-3}{2})$.

• $5 \leq r \leq k$. We prove by induction that $rf(\frac{k-r+2}{2}) \leq 5f(\frac{k-3}{2})$. For the base case when $r = 5$, we have $rf(\frac{k-3}{2}) = 5f(\frac{k-3}{2})$. For the inductive step, fix $7 \leq r \leq k$. By Lemma 6(a), we have $\frac{f(k-r+2)}{f(\frac{k-r+2}{2})} \geq \frac{7}{5} \geq \frac{r-2}{2}$, and hence $rf(\frac{k-r+2}{2}) \leq (r-2)f(\frac{k-r+4}{2})$. By induction, $(r-2)f(\frac{k-r+4}{2}) \leq 5f(\frac{k-3}{2})$. Thus $rf(\frac{k-r+2}{2}) \leq 5f(\frac{k-3}{2})$.

Case 3: $2 \leq r \leq k$, $x + y + r - 2 = k$, $1 \leq x < y \leq \lfloor k/2 \rfloor$, $2y > k \geq 15$. Since $y \leq \lfloor k/2 \rfloor$ and $2y > k$, we must have $y = \frac{k+1}{2}$ for an odd $k$. Then $x = k + 1 - y - r + 1 = y - r + 1$. With fixed $y = \frac{k+1}{2}$, we have

$$\max \{ f(y) + (r-1) \cdot f(x) \mid 2 \leq r \leq k, \ x + y + r - 2 = k, \ 1 \leq x \leq y \leq \lfloor k/2 \rfloor \}$$

$$= f(y) + \max \{ (r-1) \cdot f(y-r+1) \mid 2 \leq r \leq k, \ 1 \leq y-r+1 < y \}$$

$$= f(y) + \max \{ (r-1) \cdot f(y-(r-1)) \mid 2 \leq r \leq y \}$$

$$= f(y) + \max \{ c \cdot f(y-c) + 1 \mid 1 \leq c < y \} - 1$$

$$= f(y) + f(y) - 1,$$

where the last step follows from the recurrence (1) in the proof of Lemma 6. Since $k \geq 15$, we have $\frac{k+1}{2} \geq 8$, and it follows by Lemma 6(b) that $\frac{f(k+1)}{f(k/2)} < \left(\frac{3}{2}\right)^2 < \frac{9}{4}$. Thus, $2f(\frac{k+1}{2}) - 1 < 2f(\frac{k+1}{2}) < 5f(\frac{k-3}{2})$.

We have proved that $g(k) = 5f(\frac{k-3}{2})$ for $k = 2i + 1$ with $i \geq 7$, and $g(k) = 6f(\frac{k-1}{2})$ for $k = 2i$ with $i \geq 8$. In particular, for $15 \leq k \leq 26$, we have

$$g(15) = 5 \cdot f(6) = 55,$$

$$g(16) = 6 \cdot f(6) = 66,$$

$$g(17) = 5 \cdot f(7) = 80,$$

$$g(18) = 6 \cdot f(7) = 90,$$

$$g(19) = 5 \cdot f(8) = 115,$$

$$g(20) = 6 \cdot f(8) = 138,$$

$$g(21) = 5 \cdot f(9) = 170,$$

$$g(22) = 6 \cdot f(9) = 204,$$

$$g(23) = 5 \cdot f(10) = 245,$$

$$g(24) = 6 \cdot f(10) = 294,$$

$$g(25) = 5 \cdot f(11) = 350,$$

$$g(26) = 6 \cdot f(11) = 420.$$

Let $T_1$ be the tree consisting of a single edge. For $k \geq 2$, we have $g(k) = rf(x)$ for some $r$ and $x$ where $2 \leq r \leq k$ and $x = \frac{k-r+2}{2}$. Let $T_k$ be the tree consisting of $r$ copies of $B_x$ sharing the root. Then the size of $T_k$ is exactly $g(k)$. There exists a caterpillar in $T_k$ with exactly $k$ edges, including $2x$ edges of two very hungry caterpillars in two copies of $B_x$, and $r - 2$ other edges incident to the root that are not in these two very hungry caterpillars. Any caterpillar in $T_k$ must include, either at most one, or at least two, of the $r$ edges incident to the root. In the first case, the caterpillar is contained completely in one copy of $B_x$, and hence can have at most $2r - 1 = k - r + 1 < k$ edges by Lemma 6. In the second case, the caterpillar can extend into at most two copies of $B_x$ composing $T_k$, and hence can have at most $2x + r - 2 = k$ edges.

Refer to Figures 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 for illustrations of $T_k$ for $2 \leq k \leq 26$. \qed
Figure 5: $T_2$ with 2 edges, $T_3$ with 3 edges, $T_4$ with 4 edges, $T_5$ with 6 edges, $T_6$ with 8 edges, and $T_7$ with 10 edges.

Figure 6: $T_8$ with 12 edges, $T_9$ with 15 edges, and $T_{10}$ with 20 edges.

Figure 7: $T_{11}$ with 25 edges and $T_{12}$ with 30 edges.
Figure 8: $T_{13}$ with 35 edges and $T_{14}$ with 44 edges.

Figure 9: $T_{15}$ with 55 edges and $T_{16}$ with 66 edges.

Figure 10: $T_{17}$ with 80 edges and $T_{18}$ with 96 edges.
Figure 11: $T_{19}$ with 115 edges and $T_{20}$ with 138 edges.

Figure 12: $T_{21}$ with 170 edges and $T_{22}$ with 204 edges.
Figure 13: $T_{23}$ with 245 edges and $T_{24}$ with 294 edges.

Figure 14: $T_{25}$ with 350 edges and $T_{26}$ with 420 edges.
4.3 Deriving \( e(k) \) from \( f(k) \) and \( g(k) \)

We clearly have \( e(1) = 1 \). By Lemma 7, we have \( e(k) = g(k) \) for \( k \geq 2 \). Recall Lemma 7 that

- For \( k = 2i \) for \( i \geq 8 \), \( g(k) = 6 f \left( \frac{k-4}{2} \right) \).
- For \( k = 2i + 1 \) for \( i \geq 7 \), \( g(k) = 5 f \left( \frac{k-3}{2} \right) \).

Also recall Lemma 6 that

- For \( k = 3i \) with \( i \geq 2 \), \( f(k) = \frac{1}{2} (23 \cdot 3^{(k-6)/3} - 1) \).
- For \( k = 3i + 1 \) with \( i \geq 2 \), \( f(k) = \frac{1}{2} (33 \cdot 3^{(k-7)/3} - 1) \).
- For \( k = 3i + 2 \) with \( i \geq 2 \), \( f(k) = \frac{1}{2} (47 \cdot 3^{(k-8)/3} - 1) \).

Altogether, there are six cases for \( k \geq 15 \):

- For \( k = 6i \) with \( i \geq 3 \), \( e(k) = 6 f \left( \frac{k-4}{2} \right) = 3(33 \cdot 3^{\left( \frac{k-4}{2} \right) / 3 - 7} - 1) = 3(33 \cdot 3^{(k-18)/6} - 1) \).
- For \( k = 6i + 2 \) with \( i \geq 3 \), \( e(k) = 6 f \left( \frac{k-4}{2} \right) = 3(47 \cdot 3^{\left( \frac{k-4}{2} \right) / 3 - 8} - 1) = 3(47 \cdot 3^{(k-20)/6} - 1) \).
- For \( k = 6i + 4 \) with \( i \geq 2 \), \( e(k) = 6 f \left( \frac{k-4}{2} \right) = 3(23 \cdot 3^{\left( \frac{k-4}{2} \right) / 3 - 6} - 1) = 3(23 \cdot 3^{(k-16)/6} - 1) \).
- For \( k = 6i + 1 \) with \( i \geq 3 \), \( e(k) = 5 f \left( \frac{k-3}{2} \right) = \frac{5}{2} (47 \cdot 3^{\left( \frac{k-3}{2} \right) / 3 - 8} - 1) = \frac{5}{2} (47 \cdot 3^{(k-19)/6} - 1) \).
- For \( k = 6i + 3 \) with \( i \geq 2 \), \( e(k) = 5 f \left( \frac{k-3}{2} \right) = \frac{5}{2} (23 \cdot 3^{\left( \frac{k-3}{2} \right) / 3 - 6} - 1) = \frac{5}{2} (23 \cdot 3^{(k-15)/6} - 1) \).
- For \( k = 6i + 5 \) with \( i \geq 2 \), \( e(k) = 5 f \left( \frac{k-3}{2} \right) = \frac{5}{2} (33 \cdot 3^{\left( \frac{k-3}{2} \right) / 3 - 7} - 1) = \frac{5}{2} (33 \cdot 3^{(k-17)/6} - 1) \).

In particular, for \( k = 15, \ldots, 21 \), we have

\[
\begin{align*}
e(15) &= \frac{5}{2} (23 \cdot 3^{(15-15)/6} - 1) = 55, \\
e(16) &= 3 (23 \cdot 3^{(16-16)/6} - 1) = 66, \\
e(17) &= \frac{5}{2} (33 \cdot 3^{(17-17)/6} - 1) = 80, \\
e(18) &= 3 (33 \cdot 3^{(18-18)/6} - 1) = 96, \\
e(19) &= \frac{5}{2} (47 \cdot 3^{(19-19)/6} - 1) = 115, \\
e(20) &= 3 (47 \cdot 3^{(20-20)/6} - 1) = 138, \\
e(21) &= \frac{5}{2} (23 \cdot 3^{(21-15)/6} - 1) = 170.
\end{align*}
\]

In summary, we have the following theorem:

**Theorem 3.** \( e(1) = 1, e(2) = 2, e(3) = 3, e(4) = 4, e(5) = 6, e(6) = 8, e(7) = 10, e(8) = 12, e(9) = 15, \)
\( e(10) = 20, e(11) = 25, e(12) = 30, e(13) = 35, e(14) = 44, e(15) = 55, e(16) = 66, e(17) = 80, \)
\( e(18) = 96, e(19) = 115, e(20) = 138, e(21) = 170. \)

- For \( k = 6i \) with \( i \geq 3 \), \( e(k) = 3 (11 \cdot 3^{(k-12)/6} - 1) \).
- For \( k = 6i + 2 \) with \( i \geq 3 \), \( e(k) = 3 (47 \cdot 3^{(k-20)/6} - 1) \).
- For \( k = 6i + 4 \) with \( i \geq 2 \), \( e(k) = 3 (23 \cdot 3^{(k-16)/6} - 1) \).
- For \( k = 6i + 1 \) with \( i \geq 3 \), \( e(k) = \frac{5}{2} (47 \cdot 3^{(k-19)/6} - 1) \).
- For \( k = 6i + 3 \) with \( i \geq 2 \), \( e(k) = \frac{5}{2} (23 \cdot 3^{(k-15)/6} - 1) \).
- For \( k = 6i + 5 \) with \( i \geq 2 \), \( e(k) = \frac{5}{2} (11 \cdot 3^{(k-11)/6} - 1) \).
4.4 Deriving $q(m)$ from $e(k)$

By Theorem 3, the exact values of $e(k)$ for $1 \leq k \leq 21$ imply the following exact values of $q(m)$ for $1 \leq m \leq 170$:

| $m$ | $q(m)$ |
|-----|--------|
|     | $[1, 4]$ | $[5, 6]$ | $[7, 8]$ | $[9, 10]$ | $[11, 12]$ | $[13, 15]$ | $[16, 20]$ | $[21, 25]$ | $[26, 30]$ |
| $m$ | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| $q(m)$ | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |

In the following, we assume that $m \geq 171$. If a tree $T$ has $m > e(k−1)$ edges, then it contains a caterpillar of $k$ edges. Thus $q(m)$ is the largest integer $k$ such that $e(k−1) < m$.

For $k−1 = 6i$ with $i \geq 3$, $e(k−1) = 3(11 \cdot 3^{(k−13)/6} − 1)$. The inequality $e(k−1) < m$ is equivalent to

\[
3(11 \cdot 3^{(k−13)/6} − 1) < m \\
11 \cdot 3^{(k−13)/6} < \frac{m}{3} + 1 \\
3^{(k−13)/6} < \frac{m}{33} + \frac{1}{11} \\
(k−13)/6 < \log_3 \left( \frac{m}{33} + \frac{1}{11} \right) \\
k < 6 \log_3 \left( \frac{m}{33} + \frac{1}{11} \right) + 13 \\
k < \left[ 6 \log_3 \left( \frac{m}{33} + \frac{1}{11} \right) + 13 \right] \\
k \leq \left[ 6 \log_3 \left( \frac{m}{33} + \frac{1}{11} \right) + 12 \right] \\
k−1 \leq \left[ 6 \log_3 \left( \frac{m}{33} + \frac{1}{11} \right) + 11 \right].
\]

Thus the largest integer $k$ such that $e(k−1) < m$ and $k \mod 6 = 1$ is

\[
q_1(m) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{m}{33} + \frac{1}{11} \right) + 11 \right] \right\rfloor + 1. \tag{6}
\]
For $k - 1 = 6i + 2$ with $i \geq 3$, $e(k - 1) = 3(47 \cdot 3^{(k-21)/6} - 1)$. The inequality $e(k - 1) < m$ is equivalent to

$$3(47 \cdot 3^{(k-21)/6} - 1) < m$$
$$47 \cdot 3^{(k-21)/6} < \frac{m}{3} + 1$$
$$3^{(k-21)/6} < \frac{m}{141} + \frac{1}{47}$$
$$(k - 21)/6 < \log_3 \left( \frac{m}{141} + \frac{1}{47} \right)$$
$$k < 6\log_3 \left( \frac{m}{141} + \frac{1}{47} \right) + 21$$
$$k < \left[ 6\log_3 \left( \frac{m}{141} + \frac{1}{47} \right) + 21 \right]$$
$$k \leq \left[ 6\log_3 \left( \frac{m}{141} + \frac{1}{47} \right) + 20 \right]$$
$$k - 3 \leq \left[ 6\log_3 \left( \frac{m}{141} + \frac{1}{47} \right) + 17 \right].$$

Thus the largest integer $k$ such that $e(k - 1) < m$ and $k \mod 6 = 3$ is

$$q_3(m) = 6 \left[ \frac{1}{6} \left[ 6\log_3 \left( \frac{m}{141} + \frac{1}{47} \right) + 17 \right] \right] + 3. \quad (7)$$

For $k - 1 = 6i + 4$ with $i \geq 2$, $e(k - 1) = 3(23 \cdot 3^{(k-17)/6} - 1)$. The inequality $e(k - 1) < m$ is equivalent to

$$3(23 \cdot 3^{(k-17)/6} - 1) < m$$
$$23 \cdot 3^{(k-17)/6} < \frac{m}{3} + 1$$
$$3^{(k-17)/6} < \frac{m}{69} + \frac{1}{23}$$
$$(k - 17)/6 < \log_3 \left( \frac{m}{69} + \frac{1}{23} \right)$$
$$k < 6\log_3 \left( \frac{m}{69} + \frac{1}{23} \right) + 17$$
$$k < \left[ 6\log_3 \left( \frac{m}{69} + \frac{1}{23} \right) + 17 \right]$$
$$k \leq \left[ 6\log_3 \left( \frac{m}{69} + \frac{1}{23} \right) + 16 \right]$$
$$k - 5 \leq \left[ 6\log_3 \left( \frac{m}{69} + \frac{1}{23} \right) + 11 \right].$$

Thus the largest integer $k$ such that $e(k - 1) < m$ and $k \mod 6 = 5$ is

$$q_5(m) = 6 \left[ \frac{1}{6} \left[ 6\log_3 \left( \frac{m}{69} + \frac{1}{23} \right) + 11 \right] \right] + 5. \quad (8)$$
For $k - 1 = 6i + 1$ with $i \geq 3$, $e(k - 1) = \frac{5}{2}(47 \cdot 3^{(k-20)/6} - 1)$. The inequality $e(k - 1) < m$ is equivalent to

\[
\frac{5}{2}(47 \cdot 3^{(k-20)/6} - 1) \leq m \\
47 \cdot 3^{(k-20)/6} \leq \frac{2m}{5} + 1 \\
3^{(k-20)/6} \leq \frac{2m}{235} + \frac{1}{47} \\
(k - 20)/6 \leq \log_3 \left( \frac{2m}{235} + \frac{1}{47} \right) \\
k < 6 \log_3 \left( \frac{2m}{235} + \frac{1}{47} \right) + 20 \\
k < \left[ 6 \log_3 \left( \frac{2m}{235} + \frac{1}{47} \right) + 20 \right] \\
k \leq \left[ 6 \log_3 \left( \frac{2m}{235} + \frac{1}{47} \right) + 19 \right] \\
k - 2 \leq \left[ 6 \log_3 \left( \frac{2m}{235} + \frac{1}{47} \right) + 17 \right].
\]

Thus the largest integer $k$ such that $e(k - 1) < m$ and $k \mod 6 = 2$ is

\[
q_2(m) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{2m}{235} + \frac{1}{47} \right) + 17 \right] + 2. \right\rfloor (9)
\]

For $k - 1 = 6i + 3$ with $i \geq 2$, $e(k - 1) = \frac{5}{2}(23 \cdot 3^{(k-16)/6} - 1)$. The inequality $e(k - 1) < m$ is equivalent to

\[
\frac{5}{2}(23 \cdot 3^{(k-16)/6} - 1) \leq m \\
23 \cdot 3^{(k-16)/6} \leq \frac{2m}{5} + 1 \\
3^{(k-16)/6} \leq \frac{2m}{115} + \frac{1}{23} \\
(k - 16)/6 \leq \log_3 \left( \frac{2m}{115} + \frac{1}{23} \right) \\
k < 6 \log_3 \left( \frac{2m}{115} + \frac{1}{23} \right) + 16 \\
k < \left[ 6 \log_3 \left( \frac{2m}{115} + \frac{1}{23} \right) + 16 \right] \\
k \leq \left[ 6 \log_3 \left( \frac{2m}{115} + \frac{1}{23} \right) + 15 \right] \\
k - 4 \leq \left[ 6 \log_3 \left( \frac{2m}{115} + \frac{1}{23} \right) + 11 \right].
\]

Thus the largest integer $k$ such that $e(k - 1) < m$ and $k \mod 6 = 4$ is

\[
q_4(m) = 6 \left\lfloor \frac{1}{6} \left[ 6 \log_3 \left( \frac{2m}{115} + \frac{1}{23} \right) + 11 \right] + 4. \right\rfloor (10)
\]
For $k - 1 = 6i + 5$ with $i \geq 2$, $e(k - 1) = \frac{5}{2}(11 \cdot 3^{(k-12)/6} - 1)$. The inequality $e(k - 1) < m$ is equivalent to

$$
\frac{5}{2}(11 \cdot 3^{(k-12)/6} - 1) < m \\
11 \cdot 3^{(k-12)/6} < \frac{2m}{5} + 1 \\
3^{(k-12)/6} < \frac{2m}{55} + \frac{1}{11} \\
(k - 12)/6 < \log_3 \left(\frac{2m}{55} + \frac{1}{11}\right) \\
k < 6 \log_3 \left(\frac{2m}{55} + \frac{1}{11}\right) + 12 \\
k < \left[6 \log_3 \left(\frac{2m}{55} + \frac{1}{11}\right) + 12\right] \\
k \leq \left[6 \log_3 \left(\frac{2m}{55} + \frac{1}{11}\right) + 11\right].
$$

Thus the largest integer $k$ such that $e(k - 1) < m$ and $k \mod 6 = 0$ is

$$q_0(m) = 6 \left[\frac{1}{6} \left[6 \log_3 \left(\frac{2m}{55} + \frac{1}{11}\right) + 11\right]\right]. \quad (11)$$

Thus for $m \geq 171$, $q(m) = \max\{q_r(m) \mid 0 \leq r \leq 5\}$. This completes the proof of Theorem 2.

5 An open question

Recall our definitions of $\hat{p}(n), \hat{q}(n), p(n),$ and $q(n)$ for $n \geq 1$:

- $\hat{p}(n)$ (respectively, $\hat{q}(n)$) is the maximum number $k$ such that any family $S$ of $n$ disjoint segments in the plane admits an alternating path among $S$ (respectively, compatible with $S$) going through $k$ segments in $S$.

- $p(n)$ (respectively, $q(n)$) is the maximum number $k$ such that any family $S$ of $n$ disjoint segments in the plane, whose $2n$ endpoints are in convex position, admits an alternating path among $S$ (respectively, compatible with $S$) going through $k$ segments in $S$.

We clearly have $\hat{p}(n) \leq p(n)$ and $\hat{q}(n) \leq q(n)$ for all $n \geq 1$, but can we have $\hat{p}(n) < p(n)$ or $\hat{q}(n) < q(n)$ for some $n \geq 1$? In other words, could it be true that $\hat{p}(n) = p(n)$ and $\hat{q}(n) = q(n)$ for all $n \geq 1$?

References

[1] F. Harary and A. J. Schwenk. The number of caterpillars. *Discrete Mathematics*, 6:359–365, 1973.

[2] M. Hoffmann and C. D. Tóth. Alternating paths through disjoint line segments. *Information Processing Letters*, 87:287–294, 2003.

[3] J. Pach and R. Pinchasi. A long noncrossing path among disjoint segments in the plane. In *Combinatorial and Computational Geometry*, volume 52 of MSRI Publications, pages 495–500, 2005.

[4] J. Urrutia. Open problems in computational geometry. In *Proceedings of the 5th Latin American Symposium on Theoretical Informatics (LATIN’02)*, pages 4–11, 2002.