A Sharp Separation Bound for Triple-Occurrence Cyclic Tree Words

Augusto Rehfeldt

Text generated by language models in the mathforge pipeline; not reviewed by a person.

Abstract

Let w be a cyclic word in which each of n labeled letters occurs three times. Join two letters when their restriction alternates, and count a pair as separated when its restriction is *aaabbb* up to rotation. If the alternating graph is a tree, the maximum number of separated pairs is $(n-2)(n-3)/2$. An explicit family with path alternating graphs attains this bound; for $n \geq 4$, every maximizing word has a path alternating graph. The supplied exhaustive computations support the result for $3 \leq n \leq 5$. The Lean 4 formalization compiles with `sorry`, so it is not a completed machine-checked proof.

1. Introduction

Word-representable graphs encode adjacency through alternation in words. Here each letter occurs exactly three times, words are considered cyclically, and attention is restricted to representations whose alternating graph is a tree.

The additional statistic studied here counts letter pairs whose occurrences form two separate cyclic blocks. The result gives its sharp maximum and shows that maximizing alternating graphs must be paths. Its key observation is local: two letters sharing an alternating neighbor cannot be separated.

The supplied novelty assessment identifies this extremal statistic and path rigidity as apparently new, rather than establishing priority. Although the seed topic also asks for leaf-count refinements and admissible insertion-gap identities, the verified result does not provide those identities.

2. Notation

A **cyclic word** is a word considered modulo rotation, but not modulo reversal or relabeling. Its alphabet is $\{1, \dots, n\}$, and each letter occurs exactly three times.

The **restriction** to distinct letters a, b deletes every other letter. Define:

- $A(w)$: the simple graph on $\{1, \dots, n\}$ whose edges are the pairs with restriction *ababab* up to rotation;
- $S(w)$: the number of unordered pairs with restriction *aaabbb* up to rotation;
- $\deg_T(v)$: the degree of v in a graph T.

A tree is connected and has no cycles. The labeled path has edges $\{i, i+1\}$ for $1 \leq i < n$.

3. Results

Theorem 1 (Sharp separation bound and path rigidity)**.** *For every integer $n \geq 3$, the maximum of $S(w)$, over all cyclic words w on the labeled alphabet $\{1, \dots, n\}$ in which each letter occurs exactly three times and $A(w)$ is a tree, equals $(n-2)(n-3)/2$. Moreover, for every $n \geq 4$, every word attaining this maximum has $A(w)$ isomorphic to the path on n vertices.*

The following proof is reproduced verbatim from the supplied verified proof.

Proof.

Call a pair of distinct letters **alternating** if its restriction is *ababab* up to rotation, and **separated** if its restriction is *aaabbb* up to rotation. These two possibilities are mutually exclusive: in the alternating restriction every cyclic run has length one, whereas in the separated restriction there are two cyclic runs, each of length three.

Lemma 1 (A common neighbor prevents separation)**.** *If distinct letters a, b, c satisfy*

$$\{a, c\}, \{b, c\} \in E(A(w)),$$

then $\{a, b\}$ is not separated.

Proof. The three occurrences of c divide the cyclic word into three open arcs. Since a alternates with c, each arc contains exactly one occurrence of a. Likewise, each arc contains exactly one occurrence of b.

Consequently, after deleting c and every other letter, the restriction to a, b is a cyclic concatenation of three blocks, each equal to ab or ba. No such cyclic concatenation has three consecutive occurrences of the same letter. Indeed, any three consecutive positions contain an entire block: if the first position begins a block, the first two positions form that block; otherwise, the last two positions form the next block. An entire block contains both letters.

A separated restriction has three consecutive occurrences of a, so this restriction cannot be separated.

□

Lemma 2 (Counting pairs at distance two)**.** *For a tree T, the number of unordered vertex pairs at distance two is*

$$\sum_{v \in V(T)} \binom{\deg_T(v)}{2}.$$

Proof. Every choice of two distinct neighbors of a vertex v gives a path of length two with middle vertex v. Its endpoints have distance two: an edge between them would create a cycle.

Conversely, each pair at distance two has a unique middle vertex, because two distinct paths between the same vertices would create a cycle. Thus the choices counted in the displayed sum correspond bijectively to the pairs at distance two.

□

Lemma 3 (The degree bound and its equality case)**.** *If T is a tree on $n \geq 3$ vertices, then*

$$\sum_{v \in V(T)} \binom{\deg_T(v)}{2} \geq n - 2.$$

Equality holds if and only if T is a path.

Proof. Every vertex has positive degree, since T is connected and has more than one vertex. For every integer $d \geq 1$,

$$\binom{d}{2} - (d - 1) = \frac{(d-1)(d-2)}{2} \geq 0,$$

with equality precisely when $d \in \{1, 2\}$.

A tree on n vertices has $n - 1$ edges, and summing degrees counts each edge twice. Therefore

$$\sum_{v} \binom{\deg_T(v)}{2} \geq \sum_{v} (\deg_T(v) - 1) = 2(n - 1) - n = n - 2.$$

Equality holds precisely when every degree is at most two.

To verify the resulting characterization, choose a longest simple path in T. If a vertex lay outside this path, connectedness would supply a path from it to the chosen path, meeting the chosen path for the first time at some vertex x. If x were an endpoint, this would extend the chosen path, contradicting maximality. If x were an internal vertex, it would have at least three neighbors, contradicting the degree bound. Thus the chosen path contains every vertex. There can be no additional edge between its vertices, since such an edge would create a cycle. Hence T is a path.

Conversely, every vertex of a path has degree one or two, so equality holds.

□

Lemma 4 (An attaining construction)**.** *For $n \geq 3$, let w_n be the concatenation, for $k = 1, \dots, n+2$, of the letters*

$$k,\ k-1,\ k-2$$

that belong to $\{1, \dots, n\}$, in that order, regarded as a cyclic word. Then

$$A(w_n) = (\{1, \dots, n\}, \{\{i, i+1\} : 1 \leq i < n\}),$$

and a pair $\{i, j\}$, with $i < j$, is separated exactly when $j - i \geq 3$.

Proof. Call the contribution indexed by k block k. A letter i occurs exactly in blocks $i, i+1, i+2$, so each letter occurs three times.

Fix $i < j$.

If $j = i + 1$, the two-letter restriction, in the displayed linear order of the construction, is

$$i\, j\, i\, j\, i\, j.$$

Indeed, block i contributes i, blocks $i+1$ and $i+2$ each contribute ji, and block $i+3$ contributes j. Thus the pair alternates.

If $j = i + 2$, the restriction is

$$i\, i\, j\, i\, j\, j.$$

Blocks i and $i+1$ each contribute i, block $i+2$ contributes ji, and blocks $i+3$ and $i+4$ each contribute j. This cyclic restriction has four runs, so it is neither alternating nor separated.

If $j - i \geq 3$, all occurrences of i precede all occurrences of j: the last block containing i is $i+2$, while the first containing j is j. The restriction is therefore

$$i\, i\, i\, j\, j\, j,$$

which is separated and not alternating.

These cases exhaust all pairs and prove both assertions.

□

Proof of the theorem. Let w satisfy the hypotheses, and put $T = A(w)$. A separated pair cannot be an edge of T, because alternating and separated restrictions are mutually exclusive. By Lemma 1, it also cannot have distance two in T.

There are $n - 1$ pairs at distance one, and Lemma 2 counts the pairs at distance two. Hence

$$S(w) \leq \binom{n}{2} - (n-1) - \sum_{v} \binom{\deg_T(v)}{2}.$$

Applying Lemma 3 gives

$$S(w) \leq \binom{n}{2} - (n-1) - (n-2) = \frac{(n-2)(n-3)}{2}.$$

By Lemma 4, w_n has alternating graph equal to the path, and its separated pairs are precisely those with $j - i \geq 3$. Their number is

$$\sum_{i=1}^{n-3}(n-i-2) = \frac{(n-2)(n-3)}{2},$$

where the sum is empty when $n = 3$. Thus the bound is attained for every $n \geq 3$.

Finally, suppose $n \geq 4$ and w attains this bound. The two inequalities above force

$$\sum_{v} \binom{\deg_T(v)}{2} = n - 2;$$

otherwise the upper bound for $S(w)$ would be strictly smaller than the attained value. Lemma 3 therefore implies that $A(w)$ is isomorphic to the path on n vertices.

□

4. Computational verification

The following reports summarize the supplied verification records; no new computations are claimed here.

Exhaustive search and construction tests

The search enumerated cyclic words with three copies of each labeled letter for $3 \leq n \leq 5$, using least rotations as representatives. It retained tree alternating graphs and checked the bound and maximizing-graph condition.

n	Cyclic representatives	Tree cases	Maximum $S(w)$
3	280	9	0
4	15,400	212	1
5	1,401,400	9,470	3

Thus 1,417,080 representatives and 9,691 tree cases were checked, with no counterexample. The explicit construction was tested for every $3 \leq n \leq 100$, giving 98 successful cases. Sanity testing included three known two-letter restrictions and 1,220 direct brute-force cross-checks.

Independent check

The independent checker reported that all checks passed:

Property	Checks
Rotation invariance of definitions	112
Alternating/separated exclusivity	217,698
Tree generator; Lemma 2; Lemma 3	20,547 each
Word multiplicities	42,250
Common-neighbor obstruction	49,372
Theorem upper bound	883
Exhaustive maximum	2
Maximizer characterization	479
Construction multiplicities, graph, separation, and attainment	58 each

The independent output does not specify the underlying ranges for these counts; they should not be inferred from the primary search ranges. The supplied referee assessment also marked the proof **VALID**, with no issues or gaps found.

Lean 4 status

Theorem	Lean 4 status	Back-translation faithfulness
Theorem 1	**compiles with `sorry`**	**FAITHFUL**

The record reports eight `sorry` placeholders and dependence of the main theorem on `sorryAx`. This is not a sorry-free machine-checked proof.

The back-translation expresses the universal bound, existence of an attaining word, and path rigidity using lists over `Fin n`. Its verdict is faithful: zero-based labels merely encode the alphabet, rotation-invariant predicates permit lists to represent cyclic words, and natural-number arithmetic agrees with the formula for $n \geq 3$.

A sorry-free Lean proof certifies the Lean statement, which still has to be read against the informal one. Here that statement-level comparison was favorable, but proof completion remains outstanding.

5. Limitations and open questions

The theorem concerns an extremal statistic, not enumeration refined by leaf count. It does not characterize every maximizing word: having a path alternating graph is necessary, but the theorem does not claim that every such representation maximizes $S(w)$.

The supplied material contains no remaining `## GAP` markers, and the referee reported no mathematical gaps. Nevertheless, the Lean placeholders remain an explicit formal-verification limitation.

Open directions include:

- deriving leaf-deletion and reinsertion identities that track admissible triples of insertion gaps;
- tracking simultaneous changes in leaf count and separated-pair count;
- enumerating maximizing words modulo rotation;
- completing the Lean proof without `sorry`.

Novelty

The supplied assessment is **APPARENTLY_NEW**, with `evidence_base: retrieval`, not `memory-only`. It reports no matching result. Related known results concern tree word-representability and semi-transitive orientations; they do not, by themselves, establish novelty of this extremal theorem.

The exact queries recorded as submitted to arXiv, Crossref, and OpenAlex were:

1. `"3-uniform" "word-representable" trees "separated pairs" extremal path "(n-2)(n-3)/2"`
2. `uniform word representations alternation graphs cyclic words triple-occurrence words noninterleaving letter pairs`
3. `extremal word-representable trees leaf-deletion induction path characterization separated letter pairs`

The retrieved list contains **24 entries: 12 from arXiv and 12 from Crossref**, with **zero identified matching hits**. Per-query response totals are not separately supplied. All three OpenAlex requests failed with HTTP 429 ("Too Many Requests"); no successful OpenAlex hit count is available.

The additional suggested follow-up queries were not recorded as having been run and are therefore not counted as search evidence.

This was a **bounded automated search over those indexes—not a literature review**. It is blind to books, journals outside them, and anything phrased differently. Consequently, apparent novelty is provisional, and neither the search nor model recall establishes priority.

Models

A sharp separation bound from leaf deletion (`verified`)

- Proposed the claim: gpt-6.1-sol
- Counterexample search: gpt-6.1-sol
- Wrote the proof: gpt-6.1-sol
- Referee: gpt-6.1-sol
- Independent check script: gpt-6.1-sol
- Lean formalization: gpt-6.1-sol
- Faithfulness judge: gpt-6.1-sol