Trace Fibres of Marked Interval Heaps

Marc Craig

Independent researcher, United Kingdom

2 September 2026

Version 7

Abstract

Let a marked interval pass append one letter to each site in a nonempty interval of a finite path. Two histories are identified when adjacent passes with disjoint supports are interchanged. A history therefore defines a Cartier–Foata trace, or equivalently a Viennot heap, and determines at each site an ordered word of letters. We call the resulting tuple of words a stratigraphy.

For a fixed stratigraphy Λ, let $Q_\Lambda(t)$ count trace classes inducing Λ, graded by their number of passes. We prove an exact factorisation of Q_Λ into polynomials associated with adjacent pairs of words. Their coefficients count matching pairs of common-subsequence embeddings. If $\Lambda = (w_1, \ldots, w_n)$ has total length M, the minimum number of passes is

$$M - \sum_{i=1}^{n-1} \mathrm{LCS}(w_i, w_{i+1}).$$

The same argument gives a product over the edges of a tree for passes supported on connected subtrees.

For two equal alternating words, the local coefficients coincide with the checkerboard-chain triangle OEIS A079213, created by Dean Hickerson from a question of Cees H. Elzinga. We correct a sign error in its rectangular generating function and prove two identities that the public entry, as it stood on 31 August 2026, recorded as conjectural: the bivariate generating function for the diagonal triangle and the row-sum identity. This connects the total fibre count to A025565 and directed animals. The alternating rows fail log-concavity in every length at least seven. Finally, an exhaustive computation shows that the least mass of a binary counterexample is 14; up to exchanging the two letters, it is uniquely $(\texttt{abababa}, \texttt{abababa})$.

1 Introduction

Cartier–Foata traces and Viennot heaps encode sequences of partially commuting operations [6, 15]. When the pieces are intervals of a path and disjoint intervals commute, the underlying model is the classical heap-of-segments model [5, 11]. We consider an inverse enumeration problem for this model.

Each interval carries a letter. A heap is observed locally by reading, from bottom to top, the letters of the pieces meeting each site. This observation forgets which occurrences at different sites belong to the same interval piece. For prescribed local words, we enumerate the fibre of compatible heaps by number of pieces.

For words u and v, let $\mathrm{CS}_k(u, v)$ be the number of pairs of k-subsets of positions inducing the same subword. These are the matching joint embeddings studied by Elzinga, Rahmann and Wang [8]. Define

$$R(u, v; t) = \sum_{k \geq 0} \mathrm{CS}_k(u, v) t^{|v|-k}.$$

Our main result is the factorisation

$$Q_\Lambda(t) = t^{|w_1|} \prod_{i=1}^{n-1} R(w_i, w_{i+1}; t). \tag{1}$$

It reduces the complete fibre enumeration to independent common-subsequence calculations at adjacent boundaries. The same product holds over the edges of a tree when passes are supported on connected subtrees.

Two related compositional frameworks provide context. Berkemer, Höner zu Siederdissen and Stadler formalise pairwise alignments as order-preserving matching relations, study progressive composition, and describe tree-indexed relational composition [3, Sections 3, 5 and 8]. Abbes studies synchronization of sequences on overlapping alphabets in trace monoids, including path and ring topologies [1, Sections 3.4 and 5.4]. The present model makes a different observation: at each site it retains the ordered word of marks and forgets the interval supports of the pieces carrying them. A prescribed tuple of local words can therefore have a nontrivial fibre; Theorem 4.2 and Proposition 4.3 give its exact graded enumeration on paths and trees.

The contribution here is to use the known numbers $\mathrm{CS}_k(u, v)$ as local factors for a prescribed heap fibre and to prove the product formula above. Chase's theorem [7] counts distinct subsequence words of one word; Vatter has recently supplied a new proof [14]. By contrast, $\mathrm{CS}_k(u, v)$ counts pairs of embeddings in two words.

Section 2 defines the model. Section 3 identifies traces with consistent partitions of a cell diagram. Section 4 proves the factorisation and its tree extension. Section 5 records its principal consequences. Section 6 identifies the alternating specialisation with a previously recorded checkerboard triangle, corrects its rectangular generating function, proves the two identities recorded there as conjectural on 31 August 2026, and determines its first failure of log-concavity. Section 7 gives the computer-assisted binary minimality result.

2 Marked interval heaps and stratigraphies

Fix a finite alphabet A and the path of sites $[n] = \{1, \ldots, n\}$.

Definition 2.1. A *pass* is a pair (I, a), where $I = [\ell, r] \subseteq [n]$ is a nonempty interval and $a \in A$. A *history* is a finite sequence of passes.

Two passes are independent when their supports are disjoint. Histories are trace-equivalent when one can be obtained from the other by interchanging adjacent independent passes. A *trace* is an equivalence class of histories.

A history H determines a word w_x at each site x: read chronologically the letters of all passes whose supports contain x. Its *stratigraphy* is

$$\Lambda(H) = (w_1, \ldots, w_n).$$

We assume throughout that every w_x is nonempty. Put

$$d_x = |w_x|, \qquad M = |\Lambda| = \sum_{x=1}^{n} d_x.$$

The map $H \mapsto \Lambda(H)$ is constant on trace classes.

Definition 2.2. Let $\mathcal{T}(\Lambda)$ be the set of trace classes inducing Λ. The *trace-fibre polynomial* is

$$Q_\Lambda(t) = \sum_{T \in \mathcal{T}(\Lambda)} t^{|T|},$$

where $|T|$ is the number of passes in any representative of T.

The observation map is not injective. The stratigraphy (a, a) is induced either by one pass on $[1, 2]$ or by two singleton passes, so $Q_{(a,a)}(t) = t + t^2$.

Proposition 2.3. *For every stratigraphy of mass M,*

$$\deg Q_\Lambda = M, \qquad [t^M]Q_\Lambda(t) = 1.$$

Proof. Every pass contributes at least one local letter, so no realisation has more than M passes. Equality forces every pass to have singleton support. Their order at each site is fixed by w_x, while passes at distinct sites commute. Hence there is exactly one trace of size M. □

3 A partition model

For $\Lambda = (w_1, \ldots, w_n)$ define

$$\mathrm{Cells}(\Lambda) = \{(x, j) : 1 \le x \le n,\ 1 \le j \le d_x\}.$$

The cell (x, j) has letter $w_x[j]$.

Definition 3.1. A *piece* contains exactly one cell from each site of a nonempty interval $I \subseteq [n]$, all carrying the same letter. Its support is I. A partition of $\mathrm{Cells}(\Lambda)$ into pieces is *consistent* if, whenever two pieces meet at more than one site, their vertical order is the same at every common site.

Theorem 3.2. *Trace classes inducing Λ are in bijection with consistent partitions of* $\mathrm{Cells}(\Lambda)$ *into pieces. The number of passes equals the number of parts.*

Proof. A history partitions the cells according to the pass that created them. Every block is a piece. Two passes meeting at several sites have the same chronological order at all those sites, so the partition is consistent. Interchanging disjoint passes does not alter it.

Conversely, let $\mathcal{P}$ be a consistent piece partition. On its pieces form a directed graph: if distinct pieces P and Q meet, orient $P \to Q$ when the cell of P lies below the cell of Q at a common site. Consistency makes this orientation well defined. Exactly one orientation is assigned to each intersecting pair, so a directed 2-cycle is impossible.

The graph is acyclic. Otherwise choose a shortest directed cycle, of length at least three. The intersection graph of intervals is chordal. A cycle of length at least four has a chord; whichever way that chord is oriented, it closes one of the two directed paths between its endpoints into a shorter directed cycle. A shortest directed cycle must therefore be a triangle. Its three interval supports pairwise intersect and hence, by the Helly property for intervals, share a site. At that site their three cells are totally ordered. Consistency identifies these local orientations with the orientations defining the directed graph, so a directed triangle is impossible.

Choose a linear extension of the resulting partial order and perform the pieces in that order. At each site this gives exactly the order prescribed by Λ. Any two linear extensions differ by interchanges of adjacent incomparable elements. Overlapping pieces are comparable, so these interchanges involve disjoint supports and preserve the trace. Thus the partition determines one trace class, and the two constructions are inverse. □

Remark 3.3 (Why the path hypothesis matters). The direct cyclic analogue of Theorem 3.2 is false if pieces are allowed on proper connected arcs of a cycle while the same local consistency condition is retained. On C_3, give every site the word aa. Let

$$P_1 = \{(1,2),(2,1)\}, \quad P_2 = \{(2,2),(3,1)\}, \quad P_3 = \{(3,2),(1,1)\}.$$

These pieces form a consistent partition because each pair meets at only one site. However, P_1 lies below P_2 at site 2, P_2 lies below P_3 at site 3, and P_3 lies below P_1 at site 1. Their forced local orders therefore form a directed 3-cycle, so no history realises the partition. A cyclic theory requires an additional global acyclicity condition. A cycle-sensitive synchronization obstruction also appears in Abbes's probabilistic trace-monoid setting [1, Proposition 5.5].

4 Adjacent-boundary factorisation

If $u = u_1 \cdots u_p$ and $B = \{b_1 < \cdots < b_k\} \subseteq [p]$, write $u|_B = u_{b_1} \cdots u_{b_k}$.

Definition 4.1. For words u and v, put

$$\mathrm{CS}_k(u,v) = \#\{(B,C) : |B| = |C| = k,\ u|_B = v|_C\}$$

and

$$R(u,v;t) = \sum_{k=0}^{\min(|u|,|v|)} \mathrm{CS}_k(u,v) t^{|v|-k}. \tag{2}$$

Equivalently,

$$\mathrm{CS}_k(u,v) = \sum_{z \in A^k} \mathrm{emb}(z,u)\,\mathrm{emb}(z,v),$$

where $\mathrm{emb}(z,u)$ is the number of embeddings of z as a subsequence of u.

Theorem 4.2 (Factorisation theorem). *For every stratigraphy $\Lambda = (w_1, \ldots, w_n)$,*

$$Q_\Lambda(t) = t^{d_1} \prod_{x=1}^{n-1} R(w_x, w_{x+1}; t). \tag{3}$$

Proof. Let $\mathcal{P}$ be a consistent piece partition. At the boundary between sites x and $x+1$, record the cells belonging to pieces crossing that boundary. If there are k_x such pieces, their cells determine subsets

$$B_x \subseteq [d_x], \qquad C_{x+1} \subseteq [d_{x+1}], \qquad |B_x| = |C_{x+1}| = k_x.$$

Consistency forces their matching to preserve order, while constancy of the letter on a piece gives $w_x|_{B_x} = w_{x+1}|_{C_{x+1}}$. Thus there are $\mathrm{CS}_{k_x}(w_x, w_{x+1})$ choices at boundary x.

Conversely, choose such a pair of subsets independently at every boundary and match them in increasing order. Join the corresponding cells across adjacent sites. A cell has at most one matching edge across either adjacent boundary. A cycle in the cell graph would project to a closed walk in the underlying path; every closed walk in a tree contains an immediate backtrack, but such a backtrack would traverse the same matching edge twice and cannot occur in a simple cycle. Hence every component is a path. The same argument shows that a component meets any site at most once. Its support is therefore a contiguous interval, and its letters agree. It is a piece. Order preservation at adjacent boundaries prevents the vertical order of two overlapping components from changing, so the partition is consistent.

The constructions are inverse. The cell graph has M vertices and $\sum_x k_x$ edges and is a forest, so its number of components is

$$M - \sum_{x=1}^{n-1} k_x = d_1 + \sum_{x=1}^{n-1} (d_{x+1} - k_x).$$

Summing t to this power over the independent boundary choices gives (3). $\square$

Proposition 4.3 (Tree extension). *Let G be a finite tree rooted at r. Replace interval passes by marked passes supported on nonempty connected subtrees of G, and define local words and trace equivalence as above. Then*

$$Q_\Lambda^G(t) = t^{|w_r|} \prod_{(p(v),v)\in E(G)} R(w_{p(v)}, w_v; t), \tag{4}$$

where each edge is oriented from a parent $p(v)$ to its child v.

Proof. Families of subtrees of a tree have the Helly property, and their intersection graphs are chordal [9]; hence the proof of Theorem 3.2 applies to the static piece partitions.

Choose an order-preserving common-subsequence matching independently across every oriented edge of G. In the resulting cell graph, a cell has at most one incident matching edge over any fixed edge of G. A simple path in the cell graph therefore cannot project to a walk that immediately backtracks along a tree edge. If a component contained a cycle, or met one tree vertex twice, a simple path within that component would project to a nontrivial closed walk in G without an immediate backtrack, which is impossible. Every component consequently meets each tree vertex at most once and has connected support, so it is a valid subtree piece. Order preservation gives consistency. The inverse construction is the same boundary restriction as before.

If k_e cells are matched across edge e, the cell graph is a forest with $M - \sum_e k_e$ components. The identity

$$M - \sum_e k_e = |w_r| + \sum_{(p(v),v)\in E(G)} (|w_v| - k_{p(v),v})$$

now gives (4). $\square$

Remark 4.4 (Root independence). The right side of (4) does not depend on the root. Moving the root across one edge reverses only that edge, and

$$t^{|u|} R(u, v; t) = t^{|v|} R(v, u; t)$$

because $\mathrm{CS}_k(u, v) = \mathrm{CS}_k(v, u)$.

Corollary 4.5 (Constant words on trees). *If every vertex of a tree G carries the same word w, then*

$$Q_\Lambda^G(t) = t^{|w|} R(w, w; t)^{|V(G)|-1}.$$

Thus, in this specialisation, the polynomial depends only on the number of vertices and not on the shape of the tree.

4.1 Computing a local factor

Let

$$F_{i,j}(z) = \sum_{k\geq 0} \mathrm{CS}_k(u_1\cdots u_i, v_1\cdots v_j)z^k.$$

Then $F_{i,0} = F_{0,j} = 1$ and

$$F_{i,j} = F_{i-1,j} + F_{i,j-1} - F_{i-1,j-1} + \mathbf{1}_{u_i=v_j} z F_{i-1,j-1}. \tag{5}$$

Thus every local factor is computable by a standard length-graded common-subsequence dynamic programme [8].

5 Consequences

Corollary 5.1 (Minimum number of passes)**.** *For* $M = \sum_x |w_x|$,

$$\mu_\Lambda := \min\{|T| : T \in \mathcal{T}(\Lambda)\} = M - \sum_{x=1}^{n-1} \mathrm{LCS}(w_x, w_{x+1}). \tag{6}$$

It is computable in $O(\sum_x d_x d_{x+1})$ *time.*

Proof. In (3), minimising the exponent is equivalent to maximising the common-subsequence length independently at every boundary. The largest k for which $\mathrm{CS}_k(u,v) > 0$ is $\mathrm{LCS}(u,v)$. □

For two sites, (6) is the shortest-common-supersequence identity $|u| + |v| - \mathrm{LCS}(u,v)$. In contrast, shortest common supersequence for an unbounded number of input words is NP-complete even over a binary alphabet [12].

Corollary 5.2 (Support and endpoint coefficients)**.** *Put* $L_x = \mathrm{LCS}(w_x, w_{x+1})$*. Then*

$$\operatorname{supp} Q_\Lambda = \{\mu_\Lambda, \mu_\Lambda + 1, \ldots, M\},$$

and

$$[t^{\mu_\Lambda}]Q_\Lambda(t) = \prod_{x=1}^{n-1} \mathrm{CS}_{L_x}(w_x, w_{x+1}), \quad [t^M]Q_\Lambda(t) = 1,$$

$$Q_\Lambda(1) = \prod_{x=1}^{n-1} \sum_{k\geq 0} \mathrm{CS}_k(w_x, w_{x+1}).$$

Proof. Deleting letters from a common subsequence of length L gives common subsequences of every length from 0 to L. Each factor therefore has contiguous positive support. The formulas follow from (3). □

Corollary 5.3 (Symmetries)**.** Q_Λ *is invariant under a common relabelling of the alphabet, simultaneous reversal of all words, and reflection of the path.*

Proof. The first two operations preserve every CS_k. Reflection follows from symmetry of CS_k and the identity in Remark 4.4. □

Proposition 5.4 (Visible complexity does not bound stratigraphic complexity)**.** *Let* V_Λ *be the word formed by the last letter of each stack, and let* $\pi(V)$ *be the minimum number of ordinary interval overwrites needed to produce* V *from a blank path. For every* $m \geq 1$ *there is a stratigraphy* Λ_m *on* $2m$ *sites such that*

$$\pi(V_{\Lambda_m}) = 1, \qquad \mu_{\Lambda_m} = 2m + 1.$$

Proof. Take $\Lambda_m = (ba, aa, ba, aa, \ldots, ba, aa)$. Its mass is $4m$, and every adjacent pair has longest common subsequence of length one. Corollary 5.1 gives $\mu_{\Lambda_m} = 4m-(2m-1) = 2m+1$. Every final letter is a, so one full-width overwrite produces the visible word. $\square$

Remark 5.5 (Chain interpretation). Let $P(u,v) = \{(i,j) : u_i = v_j\}$, ordered by $(i,j) < (i',j')$ when $i < i'$ and $j < j'$. Then $\mathrm{CS}_k(u,v)$ is the number of k-element chains in $P(u,v)$. If both words are permutations of the same alphabet, these are increasing-subsequence counts for the associated permutation. This per-permutation statistic differs from the distribution, over all permutations, of the longest increasing-subsequence length studied in [4].

6 Alternating words and failure of log-concavity

Let $a_n = abab\cdots$ be the alternating word of length n, beginning with a, and put

$$T(n,k) = \mathrm{CS}_k(a_n, a_n).$$

Proposition 6.1 (Checkerboard identification)**.** *$T(n,k)$ is the number of k-element chains of black squares in an $n \times n$ checkerboard whose upper-left square is black, ordered strictly downwards and to the right. Thus $T(n,k)$ is the triangle recorded as OEIS A079213 [16], using its indexing in which row n is the $n \times n$ board, $0 \le k \le n$, with offset $(0,5)$. The entry was created by Dean Hickerson from a question posed by Cees H. Elzinga in December 2002.*

Proof. The square (i,j) is black exactly when $i \equiv j \pmod 2$, equivalently when the ith and jth letters of a_n agree. By Remark 5.5, a pair of matching k-embeddings is exactly a k-chain of such squares. $\square$

For unequal lengths define $f(m,n,k) = \mathrm{CS}_k(a_m, a_n)$ and

$$\mathcal{F}(x,y,z) = \sum_{m,n,k\ge 0} f(m,n,k) x^m y^n z^k.$$

Proposition 6.2 (Generating functions)**.** *The rectangular and diagonal generating functions are*

$$\mathcal{F}(x,y,z) = \frac{(1+x)(1+y)}{(1-x^2)(1-y^2) - xyz(1+xy)} \tag{7}$$

and

$$\sum_{n,k\ge 0} T(n,k) s^n z^k = \sqrt{\frac{1+s}{(1+s-sz)((1-s)^2 - sz(1+s))}}. \tag{8}$$

Consequently,

$$\sum_{n\ge 0} \left(\sum_{k=0}^{n} T(n,k) \right) s^n = \sqrt{\frac{1+s}{1-3s}}. \tag{9}$$

Equation (8) *proves the bivariate generating function for the diagonal triangle $T(n,k) = f(n,n,k)$ that the public A079213 entry, as it stood on 31 August 2026, recorded as conjectural, while* (9) *gives the row-sum generating function. The same series occurs in a different combinatorial model. In the notation of Baccherini, Merlini and Sprugnoli, the diagonal coefficient $F^{[101]}_{n,n}$ counts binary words with n zeroes and n ones that avoid* 101 *as a contiguous factor. Their Example 6.1, equation (6.6), gives its generating function [2]:*

$$d(s) = \frac{\sqrt{1-2s-3s^2}}{1-3s} = \sqrt{\frac{1+s}{1-3s}}.$$

Thus their balanced-word sequence and our row sums agree through their generating functions; no direct bijection is asserted here. Separately, the A025565 entry [17] has offset 1 *and records*

$$\sum_{m\geq 1} \mathrm{A025565}(m)s^m = s\sqrt{\frac{1+s}{1-3s}}.$$

Taking that entry as the definition of the OEIS label, comparison with (9) *gives*

$$\sum_{k=0}^{n} T(n,k) = \mathrm{A025565}(n+1).$$

Thus (9) *proves the row-sum generating-function identity, while the label and shift are taken explicitly from A025565. Together these establish the row-sum identification that the public A079213 entry recorded as conjectural on 31 August 2026. The OEIS entries are used for historical attribution and sequence nomenclature; the generating-function derivations are given here.*

Proof. A chain of black squares is specified by successive positive horizontal and vertical increments of equal parity. The generating function for one increment is

$$S(x,y) = \sum_{\substack{p,q\geq 1\\ p\equiv q\ (2)}} x^p y^q = \frac{xy(1+xy)}{(1-x^2)(1-y^2)}.$$

After its final selected square, the board may extend arbitrarily to the right and down. Hence

$$\mathcal{F}(x,y,z) = \frac{1}{(1-x)(1-y)} \sum_{k\geq 0} (zS(x,y))^k,$$

which simplifies to (7). This derivation also determines the required sign: a plus sign in the second denominator factor is incompatible with the positive coefficients. At the time of this version, the public A079213 entry still displays a plus sign at this point. The correction to the minus sign in (7) was submitted to the OEIS by the present author and remains under editorial review.

The diagonal is the constant term in y of $\mathcal{F}(s/y, y, z)$. Put

$$C = 1 + s^2 - sz(1+s).$$

The denominator after substitution is $C - y^2 - s^2y^{-2}$. Its reciprocal is even in y, so the terms $y + s/y$ in the numerator contribute no constant term. Therefore

$$\sum_{n,k\geq 0} T(n,k)s^n z^k = (1+s)\,\mathrm{CT}_y\,\frac{1}{C - y^2 - s^2y^{-2}} = \frac{1+s}{\sqrt{C^2 - 4s^2}}.$$

Here the final identity follows by expanding geometrically and collecting the terms with equal powers of y^2 and s^2y^{-2}. Finally,

$$C + 2s = (1+s)(1+s-sz), \qquad C - 2s = (1-s)^2 - sz(1+s),$$

which gives (8); setting $z = 1$ gives (9). $\square$

Corollary 6.3 (Enumeration and growth)**.** *Let* $S_n = Q_{(a_n,a_n)}(1)$. *Then* $S_0 = 1$, $S_1 = 2$, *and for* $n \geq 2$,

$$nS_n = 2nS_{n-1} + 3(n-2)S_{n-2}.$$

Moreover,

$$S_n \sim \frac{2}{\sqrt{3\pi n}}\, 3^n.$$

For $n \geq 1$, $S_n = 2\,\mathrm{A005773}(n)$, connecting these totals to the directed-animal sequence A005773 [18].

Proof. The recurrence follows by differentiating (9) and comparing coefficients. The asymptotic follows from the dominant square-root singularity at $s = 1/3$. The final identity is the relation between A025565 and A005773 recorded in the cited entries. □

The formulas nearest the top of the checkerboard triangle were recorded in A079213. We include their short derivation because it makes the log-concavity consequence transparent.

Lemma 6.4. *For every $n \geq 2$,*

$$T(n,n) = 1, \qquad T(n,n-1) = n, \qquad T(n,n-2) = \frac{(n-1)(3n-4)}{2}.$$

Proof. The first identity is immediate. Deleting one position from an alternating word produces a word that identifies the deletion: an internal deletion creates a doubled adjacent letter, while an endpoint deletion changes the relevant endpoint. Thus the n resulting words are distinct, giving $T(n,n-1) = n$.

For length $n-2$, each of the $n-1$ adjacent deletion pairs produces the same alternating word and therefore contributes $(n-1)^2$ ordered pairs. Every nonadjacent deletion pair produces a distinct word and contributes only its diagonal pair. Hence

$$T(n,n-2) = (n-1)^2 + \binom{n}{2} - (n-1) = \frac{(n-1)(3n-4)}{2}.$$

□

Theorem 6.5. *For every $n \geq 7$, the coefficients of $Q_{(a_n,a_n)}(t)$ are not log-concave.*

Proof. By Theorem 4.2,

$$Q_{(a_n,a_n)}(t) = t^n \sum_{k=0}^{n} T(n,k) t^{n-k}.$$

Its first three nonzero coefficients are therefore

$$1, \qquad n, \qquad \frac{(n-1)(3n-4)}{2}.$$

The first log-concavity margin is

$$n^2 - \frac{(n-1)(3n-4)}{2} = \frac{-n^2 + 7n - 4}{2},$$

which is negative for every integer $n \geq 7$. □

Thus the natural two-word joint-embedding analogue of the Chase–Vatter log-concavity theorem fails, even for two identical words [7, 14].

At $n = 7$ this gives

$$Q_{(abababa,abababa)}(t) = t^7 + 7t^8 + 51t^9 + 135t^{10} + 197t^{11} + 117t^{12} + 25t^{13} + t^{14}, \tag{10}$$

and $7^2 < 51$. At $n = 6$ the first margin is $6^2 - 35 = 1$, the sharp preceding near-miss.

For comparison, if both words are constant then

$$\mathrm{CS}_k(a^p, a^q) = \binom{p}{k}\binom{q}{k},$$

which is log-concave in k. Thus the failure is not an artefact of length alone; it requires a nontrivial equality pattern.

7 The least binary counterexample

The following result is computer-assisted.

Theorem 7.1. *Over a binary alphabet, every stratigraphy of mass at most* 13 *has a log-concave trace-fibre polynomial. At mass* 14 *there are exactly two labelled failures,*

$$(abababa, abababa), \qquad (bababab, bababab).$$

Consequently (10) *is the unique least binary counterexample up to exchanging the two letters.*

Exhaustive range

For fixed mass M and n nonempty sites, a binary stratigraphy is specified by a composition of M into n positive parts and a binary label on every cell. Its number is $\binom{M-1}{n-1}2^M$. Summing over n gives 2^{2M-1} stratigraphies of mass M. The computation evaluated all

$$\sum_{M=1}^{13} 2^{2M-1} = 44,739,242$$

stratigraphies through mass 13, with no failure, and all $2^{27} = 134,217,728$ stratigraphies of mass 14, with exactly the two failures above. No symmetry reduction or heuristic pruning was used.

Independently, every ordered pair of nonempty binary words of total length at most 14 was evaluated. All $180,228$ pairs of total length at most 13 had log-concave local factors. Among the $212,992$ pairs of total length 14, only the two alternating pairs failed. Products of nonnegative log-concave sequences without internal zeros are log-concave [10, 13]. By Corollary 5.2, every local factor has contiguous positive support. In a mass-14 stratigraphy with at least three sites, every adjacent pair has total length at most 13. The factor scan therefore proves the entire theorem; the full stratigraphy sweep is a logically redundant but independent computational cross-check.

The binary qualification is essential. Over a three-letter alphabet, for example,

$$\Lambda = (abac, cccccaba)$$

has mass 12 and

$$Q_\Lambda(t) = t^9 + 3t^{10} + 10t^{11} + t^{12},$$

so $3^2 < 10$. This example has been reproduced both from (3) and by direct history enumeration. No claim of arbitrary-alphabet minimality is made here.

Cross-checks

Four implementations were compared on overlapping complete ranges:

(M1) direct history enumeration followed by canonical trace reduction;

(M2) dynamic programming over Cartier–Foata layers;

(M3) subset dynamic programming on consistent cell partitions;

(M4) the product formula (3) with recurrence (5).

Methods M2–M4 agreed on $71,690$ stratigraphies with $n \leq 4$ and $M \leq 9$. A separately written checker compared direct history enumeration with the product formula, using brute-force subset pairs rather than (5), on $31,066$ binary and ternary stratigraphies, with no mismatch. The same checker tested the tree formula on $3,344$ stratigraphies on $K_{1,3}$ for all four roots, and compared brute-force CS_k with (5) on $7,684$ ordered word pairs, again with no mismatch. Five deliberately corrupted variants of the recurrences were rejected by the checkpoint suite.

These comparisons are not four independent confirmations of Theorem 4.2. M3 and M4 both use the adjacent-boundary decomposition and are coupled checks of its implementation. The history-first calculation and brute-force common-subsequence calculation are structurally independent of that decomposition. Finite agreement supports the computation but does not replace mathematical or human review.

8 Concluding remarks

The factorisation theorem solves the prescribed-stratigraphy fibre problem for marked interval heaps on a path. It gives the complete distribution by number of pieces and a polynomial-time formula for its minimum. Coefficient questions reduce to the two-word matching-embedding polynomials $R(u, v; t)$.

The tree extension shows how far the argument survives under branching. Remark 3.3 shows that a direct cyclic extension fails already on three sites: local consistency must be supplemented by a global acyclicity condition. Further problems include a structural classification of log-concavity failure, unimodality of the local and global polynomials, and direct bijections linking alternating trace fibres, the 101-avoiding binary-word diagonals of Baccherini–Merlini–Sprugnoli, and the path or directed-animal models counted by A025565 and A005773.

Data and code

The accompanying archive `interval_heap_stratigraphies_verification.zip` contains the C and Python implementations, checkpoint, mutation and ledger checks, sweep and validation logs, and per-mass timing data. It has SHA-256 digest

`0326f08c5a3ab6911bc527663133b101de0497a4d107abb80df0920d10618c02`.

An external verifier should, at minimum:

(i) inspect independently Theorems 3.2 and 4.2, Proposition 4.3, and Remark 4.4;

(ii) rebuild the local factors without reusing recurrence (5);

(iii) verify the checkerboard identification and live A079213 offset $(0, 5)$; distinguish the diagonal generating function (8) from the row-sum generating function (9); check the agreement of (9) with the balanced, contiguous-factor-avoiding model in BMS Example 6.1, equation (6.6), via $1-2s-3s^2 = (1-3s)(1+s)$; and verify from the live A025565 entry its offset 1, generating function, the shift $S_n = \mathrm{A025565}(n+1)$, and the identity $\mathrm{A025565}(n+1) = 2\,\mathrm{A005773}(n)$, valid for $n \geq 1$, used in Corollary 6.3;

(iv) reproduce the two binary failures at mass 14 and the absence of failures below mass 14;

(v) verify the full polynomial (10); and

(vi) confirm that the exhaustive loops cover each positive depth composition and each binary cell labelling exactly once; and

(vii) reproduce the binary and ternary algorithm cross-checks, the $K_{1,3}$ tree tests for all four roots, the brute-force CS_k checks, and the five corrupted-recurrence rejections.

Declaration of generative AI and AI-assisted technologies

During the preparation of this work the author used generative-AI systems to assist with conjecture generation, code development, adversarial checking, literature searching and language revision. All mathematical claims were reconstructed as explicit proofs or identified as computer-assisted statements. AI systems are not authors and are not treated as mathematical authorities. Independent human verification is pending.

References

[1] S. Abbes, Synchronization of Bernoulli sequences on shared letters, *Inform. and Comput.* **255** (2017), 1–26, https://doi.org/10.1016/j.ic.2017.04.002.

[2] D. Baccherini, D. Merlini and R. Sprugnoli, Binary words excluding a pattern and proper Riordan arrays, *Discrete Math.* **307** (2007), 1021–1037.

[3] S. J. Berkemer, C. Höner zu Siederdissen and P. F. Stadler, Compositional properties of alignments, *Math. Comput. Sci.* **15** (2021), 609–630, https://doi.org/10.1007/s11786-020-00496-8.

[4] M. Bóna, M.-L. Lackner and B. E. Sagan, Longest increasing subsequences and log concavity, *Ann. Comb.* **21** (2017), 535–549.

[5] M. Bousquet-Mélou and X. G. Viennot, Empilements de segments et q-énumération de polyominos convexes dirigés, *J. Combin. Theory Ser. A* **60** (1992), 196–224.

[6] P. Cartier and D. Foata, *Problèmes combinatoires de commutation et réarrangements*, Lecture Notes in Mathematics 85, Springer, 1969.

[7] P. J. Chase, Subsequence numbers and logarithmic concavity, *Discrete Math.* **16** (1976), 123–140.

[8] C. Elzinga, S. Rahmann and H. Wang, Algorithms for subsequence combinatorics, *Theoret. Comput. Sci.* **409** (2008), 394–404.

[9] F. Gavril, The intersection graphs of subtrees in trees are exactly the chordal graphs, *J. Combin. Theory Ser. B* **16** (1974), 47–56.

[10] S. G. Hoggar, Chromatic polynomials and logarithmic concavity, *J. Combin. Theory Ser. B* **16** (1974), 248–254.

[11] C. Krattenthaler, The theory of heaps and the Cartier–Foata monoid, appendix to the electronic edition of Cartier and Foata, 2006.

[12] K.-J. Räihä and E. Ukkonen, The shortest common supersequence problem over binary alphabet is NP-complete, *Theoret. Comput. Sci.* **16** (1981), 187–198.

[13] R. P. Stanley, Log-concave and unimodal sequences in algebra, combinatorics, and geometry, *Ann. New York Acad. Sci.* **576** (1989), 500–535.

[14] V. Vatter, Log-concavity of subsequence counts of words, *arXiv:2608.22147 [math.CO]*, 2026.

[15] G. X. Viennot, Heaps of pieces, I: Basic definitions and combinatorial lemmas, in *Combinatoire énumérative*, Lecture Notes in Mathematics 1234, Springer, 1986, pp. 321–350.

[16] D. Hickerson, OEIS A079213, "Triangle read by rows giving $T(n,k)$", The On-Line Encyclopedia of Integer Sequences, entry created from a question posed by Cees H. Elzinga in December 2002, https://oeis.org/A079213 (accessed 2 September 2026).

[17] C. Kimberling, OEIS A025565, The On-Line Encyclopedia of Integer Sequences, https://oeis.org/A025565 (accessed 2 September 2026).

[18] The OEIS Foundation, OEIS A005773, "Number of directed animals of size n", The On-Line Encyclopedia of Integer Sequences, https://oeis.org/A005773 (accessed 2 September 2026).