hexagon:2611.00655v1

Flip-connectivity of octonion bracketing classes

A characterization for all lengths, its Bell-number count,
and a Lean 4 formalization

Jesús Vilela Jato

Independent researcher, Madrid

github.com/Dojo-1/Lemma-W-count-determinant-witness-

November 2026

Abstract

Fix a word $w = w_1 \cdots w_n$ in letters and substitute octonions for the letters. Each of the C_{n-1} bracketings of w is a polynomial map, and bracketings inducing the same map form a *kernel class*. The bracketings are the vertices of the associahedron K_n, whose edges are the flips $(AB)C \leftrightarrow A(BC)$. We prove that every kernel class of w induces a connected subgraph of K_n if and only if w has at most two distinct letters or every letter of w occupies a single contiguous run. Consequently, among the Bell(n) words of length n up to renaming of letters, exactly Bell$(n) - 2^n + n$ have a kernel class that is not flip-connected.

The proof has a local and a global part. Locally, a flip that preserves the map must have a letter-count matrix of rank at most two (a count-determinant witness), and for interval words the converse holds by Artin's theorem. Globally, for interval words the components of the graph of rank-≤ 2 flips are separated by an explicit invariant that octonion moments detect; for non-interval words a family of *caterpillar* certificates exhibits two Moufang-equivalent bracketings in different such components. The main theorem, the count and every lemma they use are theorems of a Lean 4 formalization (no `sorry`, standard axioms only); the one informal step is the remark that polynomial identities over $\mathbb{Z}^8$ and over $\mathbb{R}^8$ coincide.

MSC2020: Primary 17D05; Secondary 17A35, 05A18, 05A15, 20N02, 05C40, 52B05, 68V20, 68V15.

Keywords: octonions, alternative algebras, Moufang identities, bracketings, associahedron, flip graph, associative spectrum, Bell numbers, Lean 4, formalized mathematics.

Contents

1 Introduction **2**

2 Preliminaries **3**

2.1 Octonions 3

2.2 Words, bracketings, flips 4

2.3 Run sequences 4

3 Local lemmas **4**

4 Interval words: Theorem G2 **6**
4.1 Components and triple shapes 6
4.2 Three-block words 6
4.3 Moments 7
4.4 Theorem G2 8

5 Non-interval words: Theorem G3 **8**
5.1 Stable nodes 8
5.2 Caterpillars 8
5.3 Existence of caterpillars 9

6 The main theorem and the count **10**

7 Formal verification **11**
7.1 Design 11
7.2 Further checks 12

8 Remarks and questions **12**

1 Introduction

In an associative algebra all bracketings of a product agree. In a nonassociative algebra they need not, and the natural question is *which* bracketings agree. For a groupoid $(G, *)$, Csákány and Waldhauser [2] introduced the *associative spectrum*: the number $s_n(G)$ of distinct n-ary term operations induced by the bracketings of $x_1 \cdots x_n$ with pairwise distinct variables. Associative spectra have since been determined for many operations, including graph algebras [4] and linear quasigroups [5], and extended to p-ary operations [6] and to the associative–commutative setting [3]. The spectrum of an *antiassociative* operation is the Catalan sequence.

This paper studies a refinement in which the variables may repeat. Fix a word w; the bracketings of w are the vertices of the associahedron K_n (Tamari [11], Stasheff [10]), and two adjacent vertices differ by one *flip* $(AB)C \leftrightarrow A(BC)$ applied inside a context. Over the octonions $\mathbb{O}$, bracketings inducing the same polynomial map partition the vertex set. Repeated letters create identities such as the Moufang laws [8], so the classes are no longer singletons, and it becomes meaningful to ask about their *shape* inside K_n.

Definition 1.1. A kernel class of w is *flip-connected* if any two of its members are joined by a path of flips all of whose vertices lie in the class. A word is *interval* if every letter occupies one contiguous run, e.g. $x^a y^b z^c$.

Theorem A (Main theorem)**.** *Every kernel class of bracketings of w over $\mathbb{O}$ is flip-connected if and only if w has at most two distinct letters or w is an interval word.*

Theorem B (Count)**.** *For $n \geq 1$, among the* Bell(n) *words of length n up to renaming of letters, exactly* Bell$(n) - 2^n + n$ *have a kernel class that is not flip-connected. The sequence begins* $0, 0, 0, 3, 25, 145, 756, 3892, 20644, 114961$ *for $n = 1, \ldots, 10$.*

Both are proved below, and both are theorems of the Lean 4 formalization described in Section 7: `MainThm.main` and `Count.count_disconnected`. The count runs over restricted growth strings, and `Canon.canonical` shows that these contain exactly one word per renaming class.

The local criterion. Let a flip $(AB)C \to A(BC)$ act on consecutive segments a, b, c of w, and let N be the $3 \times |\Lambda|$ matrix of letter counts of a, b, c (Λ the alphabet). By context cancellation (Theorem 3.1) the flip stays inside a kernel class exactly when the local associator $[A, B, C]$ vanishes identically. The count-determinant witness (Theorem 3.2) shows that rank $N = 3$ forces $[A, B, C] \not\equiv 0$. We write D_N for the set of flips whose count matrix is degenerate (rank $N \le 2$) and call them D_N*-flips*. So in-class flips are D_N-flips, and two equal maps in different components of the D_N-flip graph certify a disconnected class. The converse fails in general (in $xyzyx$ the flip on $xy \,|\, z \,|\, yx$ has rank $N = 2$ and a nonzero associator), but it holds for interval words (Theorems 3.4 and 3.5).

Interval words (Theorem G2, Theorem 4.6). We classify the D_N-components of an interval word by the shapes that a bracketing induces on leaf triples with three distinct letters (Theorem 4.1), reduce separation to three-block words $x^a y^b z^c$, and describe their components by an explicit invariant (Theorem 4.2). Substituting $x = 1 + se_1$, $y = 1 + qe_2$, $z = 1 + te_4$ turns every bracketing into a polynomial whose $s^p q t^r$ coefficients are *moments* of the invariant (Theorem 4.4); these moments determine the invariant (Theorem 4.5). Hence distinct components are distinct maps.

Non-interval words (Theorem G3, Theorem 5.4). For a three-letter non-interval word we build a bracketing T_0 along a *caterpillar*, a sequence of peels of end runs, and a Moufang move $T_0 \to T_1$ deleting one of its nodes (Fig. 1). The caterpillar's nodes form a *self-certifying* set: every rotation that would delete one of them has a rank-3 count matrix (Theorem 5.2), so no D_N-path from T_0 can lose them (Theorem 5.1). An induction on run sequences shows that every non-interval three-letter word has such a caterpillar (Theorem 5.3). Specialising extra letters to 1 lifts the result to words with more letters (Theorem 3.6).

Formalization. The development is a Lean 4 [7] project of 9,252 lines in 30 files using only the core library. Octonion identities are proved by a verified polynomial normaliser, the finite base of the run-sequence induction by kernel evaluation, and everything else by ordinary proofs. The theorems of Table 2 depend only on `propext`, `Classical.choice` and `Quot.sound` (log in the repository). A second implementation in Prolog runs the constructions and checks their outputs on small words (Section 7.2).

Scope of valuations. "Equal maps" means equal as polynomial maps, equivalently equal on all real octonions. The formal development quantifies over *integer* octonions $\mathbb{Z}^8$. Every bracketing is a polynomial with integer coefficients in the $8|\Lambda|$ coordinates, and a polynomial vanishing on $\mathbb{Z}^{8|\Lambda|}$ is zero, so the notions coincide; this is the only step outside the formal proof.

2 Preliminaries

2.1 Octonions

Let $\mathbb{O}$ be the real algebra with basis $e_0 = 1, e_1, \ldots, e_7$ and product $e_i e_j = \varepsilon(i, j)\, e_{i \oplus j}$, where $\oplus$ is bitwise exclusive or and $\varepsilon(i, j) \in \{\pm 1\}$ is a Cayley–Dickson sign table [1] with $e_i^2 = -1$ for $i \ge 1$. $\mathbb{O}$ is an alternative division algebra; its norm form $\nu(x) = x\bar{x}$ is multiplicative. We write $[a, b, c] = (ab)c - a(bc)$; the associator is trilinear and alternating, and vanishes when one argument is real. We use:

- the Moufang identities [8] $a(b(ac)) = ((ab)a)c$, $((ca)b)a = c(a(ba))$, $(ab)(ca) = a((bc)a)$;
- Artin's theorem: the subalgebra generated by any two elements is associative [9, 12];
- for $X, Y, Z \in \operatorname{span}(e_1, e_2, e_4)$ with coordinate vectors $x, y, z \in \mathbb{R}^3$, $[X, Y, Z] = 2\det(x, y, z)\, e_7$; in particular $[e_1, e_2, e_4] = 2e_7$.

All three are polynomial identities with integer coefficients; in the formalization they are proved for $\mathbb{Z}^8$ by a normaliser.

2.2 Words, bracketings, flips

A word $w = w_1 \cdots w_n$ is a finite sequence over an alphabet Λ. A *bracketing* is a planar binary tree T with leaves $1, \ldots, n$ in order. For a valuation $\rho\colon \Lambda \to \mathbb{O}$, $\mathrm{ev}_T(\rho)$ is the product of $\rho(w_1), \ldots, \rho(w_n)$ bracketed by T, and $T \sim T'$ if $\mathrm{ev}_T = \mathrm{ev}_{T'}$ for all ρ. A subtree spans a *segment*, a set of consecutive positions. A *context* $C[\,\cdot\,]$ is a tree with one hole. A *flip* replaces $C[(AB)C]$ by $C[A(BC)]$ or back; bracketings and flips form the associahedron K_n.

For a flip on segments a, b, c let $N \in \mathbb{Z}^{3\times|\Lambda|}$ be the matrix of letter counts. The flip is a D_N*-flip* if $\operatorname{rank} N \le 2$. Since N depends only on the segments, this does not depend on A, B, C or on the context.

2.3 Run sequences

Collapsing maximal runs of equal letters gives the *run sequence* of w; w is interval iff its run sequence has no repeated letter. A word with at least three letters that is not interval contains a pattern $p \cdots q \cdots p$ with $q \neq p$; with any third letter r, its restriction to $\{p, q, r\}$ is again non-interval.

3 Local lemmas

Lemma 3.1 (context cancellation, L1)**.** *For every context C and bracketings S_1, S_2 of the same segment, $C[S_1] \sim C[S_2]$ iff $S_1 \sim S_2$. Consequently a flip $C[(AB)C] \to C[A(BC)]$ stays inside a kernel class iff $[A, B, C] \equiv 0$.*

Proof. $C[S]$ is S composed with left and right multiplications by values of the other subtrees of C. If all letter values are nonzero these values are nonzero, and multiplication by a nonzero octonion is injective since ν is multiplicative. So $C[S_1] = C[S_2]$ at a valuation with nonzero letters forces $S_1 = S_2$ there; such valuations are Zariski dense. The converse is clear. □

Lemma 3.2 (count-determinant witness, W)**.** *If $\operatorname{rank} N = 3$, then $[A, B, C] \not\equiv 0$ for all bracketings A, B, C of a, b, c. Precisely, if the minor of N on letters x, y, z is $d \neq 0$ and $M = 2|d| + 1$, the integer valuation $x \mapsto 1 + Me_1$, $y \mapsto 1 + Me_2$, $z \mapsto 1 + Me_4$, other letters $\mapsto 1$, gives*

$$[A, B, C] \equiv 2dM^3 e_7 \pmod{M^4\mathbb{Z}^8},$$

which is nonzero. All letter values have real part 1.

Proof. Write the letter values as $1 + Mq_\ell$ with $q_x = e_1$, $q_y = e_2$, $q_z = e_4$, $q_\ell = 0$ otherwise. Products respect congruences mod M^4, so a bracketing S of a segment s satisfies $\mathrm{ev}_S \equiv 1 + M\Sigma_s + M^2P_2 + M^3P_3$, the product of truncated power series, where $\Sigma_s = \sum_\ell N_{s\ell} q_\ell$ depends only on the counts. In the associator of three such elements the terms of order M^0, M^1, M^2 cancel because the associator vanishes when an argument is real, and the term of order M^3 is

$[\Sigma_a, \Sigma_b, \Sigma_c] = 2de_7$. If $[A, B, C] = 0$ then $M^4 \mid 2dM^3$, so $M \mid 2d$, contradicting $0 < |2d| < M$. (Over $\mathbb{R}$, $x \mapsto \exp(\theta e_1)$ etc. gives $[A, B, C] = 2d\theta^3 e_7 + O(\theta^4)$.) □

Corollary 3.3. *For every word, every flip that stays inside a kernel class is a* D_N*-flip. Hence two equal maps in different components of the* D_N*-flip graph lie in a kernel class that is not flip-connected.*

Proof. Theorems 3.1 and 3.2; the witness valuation has nonzero letter values, as context cancellation requires. □

Lemma 3.4 (vanishing, V)**.** $[A, B, C] \equiv 0$ *if the segments* a, b, c *together use at most two letters, or if two of* a, b, c *are powers of the same letter.*

Proof. In the first case all three values lie in the subalgebra generated by two letter values. In the second, two values lie in $\mathrm{span}\{1, x\}$ ($\mathbb{O}$ is quadratic), so all three lie in the subalgebra generated by x and the third value. Artin's theorem makes either subalgebra associative. □

Lemma 3.5 (interval rank, I)**.** *Let* w *be interval and* a, b, c *consecutive segments using at least three letters, such that no two of* a, b, c *are powers of the same letter. Then* $\operatorname{rank} N = 3$.

Proof. Order the letters by block, so that $\max \mathrm{supp}(a) \le \min \mathrm{supp}(b)$ and $\max \mathrm{supp}(b) \le \min \mathrm{supp}(c)$.

Case $\mathrm{supp}(b) = \{p\}$. Neither a nor c is a power of p, so a has a letter $x < p$ and c a letter $z > p$. On columns (x, p, z),

$$N_{xpz} = \begin{pmatrix} a_x & a_p & 0 \\ 0 & b_p & 0 \\ 0 & c_p & c_z \end{pmatrix}, \qquad \det N_{xpz} = a_x b_p c_z \neq 0.$$

Case $|\mathrm{supp}(b)| \ge 2$, with $p = \min \mathrm{supp}(b) < q = \max \mathrm{supp}(b)$; then $\mathrm{supp}(a) \cap \mathrm{supp}(c) = \emptyset$. Either $\mathrm{supp}(a) = \{p\}$ or a has a letter $x < p$; either $\mathrm{supp}(c) = \{q\}$ or c has a letter $z > q$; let $y \in \mathrm{supp}(b)$ with $p < y < q$ when it exists. Take the first column x (or p), the last column z (or q), and the middle column y, or, if no such y exists, the letter of $\{p, q\}$ not already used. Row a vanishes outside the first column and the column p; row c vanishes outside the last column and the column q. In every combination one of these rows has a single nonzero entry; expanding along it leaves a triangular 2×2 block, so the determinant is a product of three nonzero counts. For instance, if $\mathrm{supp}(a) = \{p\}$ and $\mathrm{supp}(c) = \{q\}$ (y then exists by the three-letter hypothesis), on columns (p, y, q) the matrix is

$$\begin{pmatrix} a_p & 0 & 0 \\ b_p & b_y & b_q \\ 0 & 0 & c_q \end{pmatrix},$$

not triangular, with determinant $a_p b_y c_q \neq 0$. □

Theorems 3.4 and 3.5 together say that in an interval word every D_N-flip has a vanishing associator.

Lemma 3.6 (killing, K)**.** *Setting a set of letters to* 1 *sends bracketings of* w *to bracketings of the subword* w' *(or to the empty product). It maps* D_N*-flips to* D_N*-flips or to identities, and equal maps to equal maps. Moufang moves whose repeated variable is a single letter* lift*: if* $T_0' \to T_1'$ *is such a move of* w' *joining different* D_N*-components of* w'*, there is such a move* $T_0 \to T_1$ *of* w *restricting to it and joining different* D_N*-components of* w.

Proof. Deleting columns of N cannot raise its rank, and specialisation preserves identities. In the patterns $x(Y(xC))$, $((Cx)Y)x$ and $x((BC)x)$ every gap between consecutive surviving leaves borders a non-repeated variable (Y, B or C) or the context, so the killed leaves can be absorbed there, and the lifted move is again such a Moufang move. A D_N-path between the lifts would restrict to a D_N-path, or a constant path, between T_0' and T_1'. □

4 Interval words: Theorem G2

4.1 Components and triple shapes

Proposition 4.1 (Proposition C)**.** *Let w be interval. For a bracketing T let $\tau(T)$ record, for every leaf triple $i_1 < i_2 < i_3$ with pairwise distinct letters, whether T induces $((i_1 i_2) i_3)$ or $(i_1 (i_2 i_3))$. Then T, T' lie in the same D_N-component iff $\tau(T) = \tau(T')$.*

Proof. ($\Rightarrow$) A flip $(AB)C \to A(BC)$ changes exactly the shapes of the triples in $A \times B \times C$. If it is a D_N-flip, by Theorem 3.5 abc has at most two letters or two of a, b, c are powers of the same letter; in an interval word, if a and c are powers of x then so is b. In every case each triple of $A \times B \times C$ repeats a letter, so τ is unchanged.

($\Leftarrow$) Induction on n; words with at most two letters have a single component. Let $T = (T_L, T_R)$ and $T' = (T_L', T_R')$ have root splits after positions m and m'. If $m = m'$, τ restricts to both factors and D_N-flips inside a subtree are D_N-flips of T; induction applies. Say $m < m'$. Every triple in $[1, m] \times (m, m'] \times (m', n]$ has shape $i_1(i_2 i_3)$ in T and $(i_1 i_2) i_3$ in T', so no such triple has three distinct letters. With letter sets X, Y, Z of the three segments, ordered by block, every $x \in X, y \in Y, z \in Z$ has $x = y$ or $y = z$. If $|Y| \geq 3$ a middle letter gives a distinct triple; if $Y = \{y_1 < y_2\}$ then $X = \{y_1\}$, $Z = \{y_2\}$ and w has two letters. So $Y = \{y\}$ and $X = \{y\}$ or $Z = \{y\}$. If $X = Y = \{y\}$, let $T'' = (P, T_R')$ on $w[m+1..n]$ with any bracketing P of $(m, m']$; its distinct-letter triples lie in $(m', n]$ or have shape $i_1(i_2 i_3)$ with $i_1 \in (m, m']$, where T agrees with T'. So $\tau(T_R) = \tau(T'')$, and by induction T_R and T'' are D_N-connected. The flip $(T_L, (P, T_R')) \to ((T_L, P), T_R')$ is a D_N-flip (a, b are powers of y) producing root split m', which reduces to the case $m = m'$. The case $Y = Z = \{y\}$ is symmetric. □

4.2 Three-block words

Let $w = x^a y^b z^c$ with $a, b, c \geq 1$. For a bracketing T, the *core* $\kappa = (U, V)$ is the lowest node containing the last x and the first z. Then $U = x^{a'} y^j$ and $V = y^{b-j} z^{c'}$ with $a', c' \geq 1$, and every proper ancestor of κ attaches a pure-x block on the left or a pure-z block on the right. Let W be the word in $\{x, z\}$ of attached letters read from κ outward, and $\Phi(T) = (j, a', c', W)$.

Proposition 4.2 (Proposition T)**.** *On the D_N-components of $x^a y^b z^c$: j is invariant; for $0 < j < b$, Φ is a complete invariant; for $j = 0$ (resp. $j = b$) the complete invariant is the join word from the core outward in which the core x's (resp. z's) merge with the adjacent outer run.*

Proof. By Theorem 4.1, components are the fibres of τ. For a three-block word, $\tau(T)$ is determined by, and determines, the function assigning to each pair (y_k, z_l) the number of x's grouped with y_k before z_l; translating this function into Φ (respectively the join word) gives the stated invariants. The bookkeeping is routine but long, and we rely on its formal version (`PropT.propT`: same D_N-component $\iff$ same invariant, via `propC_iff`, `sig_iff_cx` and `key_iff_cx`). For orientation, the rotations that change Φ have determinants $\pm\, i\, j\, c'$ (an x-block of size i crossing the core, including splitting U), $\pm\, a'(b-j)k$ (the mirror, with a z-block of size k), $\pm\, a'k\, c'$ (moving k letters y across κ) and $\pm\, i\, k\, n_y$ (an x-attachment against

a z-attachment through a node containing $n_y \geq 1$ letters y); merging adjacent attachments on one side is a D_N-flip leaving W unchanged. Hence j never changes, and for interior j neither do a', c'; for $j = 0$ only the core x's can merge outward, and symmetrically for $j = b$. □

Remark 4.3. Counting the realisable invariants gives $(b-1)\big(\binom{a+c}{a} - 1\big) + \binom{a+c-1}{a} + \binom{a+c-1}{c} = 1 + b\big(\binom{a+c}{a} - 1\big)$ components, using $\sum_{i<a,\,k<c} \binom{i+k}{i} = \binom{a+c}{a} - 1$. This count is not used in the proof of Theorem A. The binomial identity is formal (`Octonion.propT_count`); the count of components itself is verified exhaustively for small a, b, c (Section 7.2) but not formalized.

4.3 Moments

Substitute $x = 1 + se_1$, $y = 1 + qe_2$, $z = 1 + te_4$ and let $\mu_T(p, r) \in \mathbb{O}$ be the coefficient of $s^p q t^r$ in ev_T. For a word ω in $\{x, z\}$ let P_ω be the ordered product of $1 + sX$ per x and $1 + tZ$ per z in the Clifford algebra with $X^2 = Z^2 = 1$, $XZ = -ZX$, and write $P_\omega = \sum_{p,r} \Gamma_\omega(p, r)\, s^p t^r\, X^{[p]} Z^{[r]}$ with $\Gamma_\omega(p, r) \in \mathbb{Z}$ and $[p] = p \bmod 2$. Put $\omega_+ = x^{a'} z^{c'} W$ and $\omega_- = z^{c'} x^{a'} W$.

Lemma 4.4 (moment formula, M)**.** *There are signs $\varepsilon_{p,r}$ and signed units $E_{p,r} = \pm e_1^{[p]} e_4^{[r]} e_2$, independent of T, with*

$$\mu_T(p, r) = \varepsilon_{p,r}\big[\, j\, \Gamma_{\omega_+}(p, r) + (b - j)\, \Gamma_{\omega_-}(p, r)\big] E_{p,r}.$$

Proof. $\mu_T(p, r)$ sums, over choices of p leaves x, one leaf y_k and r leaves z, the value of T with the chosen leaves set to e_1, e_2, e_4 and the others to 1. Let $H = \langle e_1, e_4 \rangle \cong \mathbb{H}$ and $\ell = e_2 \perp H$, so $\mathbb{O} = H \oplus H\ell$ with $h(c\ell) = (ch)\ell$ and $(c\ell)h = (c\bar{h})\ell$. Subtrees without y_k take values in the associative H. Walking from y_k to the root, the value keeps the form $c\ell$: a left sibling h appends h, a right sibling appends $\bar{h}$. After setting unchosen leaves to 1, right siblings contain only z's, and $\overline{e_4^m} = (-1)^m e_4^m$. So the value is $(-1)^r e^{\omega_k} \ell$, where ω_k is the word of absorbed blocks read bottom-up. By the definition of the core, the ancestors of y_k absorb the core and then the attachments in the order W; if y_k lies in the core's left child (j choices) the core is absorbed as $x^{a'} z^{c'}$, otherwise as $z^{c'} x^{a'}$. Summing over subsets and reordering to $e_1^{[p]} e_4^{[r]}$, with signs that depend only on (p, r), gives the formula; the Clifford relations match those of e_1, e_4 up to these signs. □

Proposition 4.5 (separation, S)**.** *Distinct D_N-components of $x^a y^b z^c$ have distinct maps; already the moments $\mu(p, 1)$ and $\mu(1, r)$ separate them.*

Proof. In the t-linear part of P_ω, a z preceded by α letters x contributes $u^\alpha Z u^{a-\alpha} = Z \bar{u}^\alpha u^{a-\alpha}$ with $u = 1 + sX$, $\bar{u} = 1 - sX$; on the eigenline $X = +1$ these are $(1 - s)^\alpha (1 + s)^{a-\alpha}$, $\alpha = 0, \ldots, a$, a basis of polynomials of degree $\leq a$. So the t-linear moments determine the weighted profile $\zeta(\alpha) = j\, \zeta_{\omega_+}(\alpha) + (b - j)\, \zeta_{\omega_-}(\alpha)$, where $\zeta_\omega(\alpha)$ counts the z's of ω preceded by exactly α letters x; symmetrically the s-linear moments determine the x-profile ξ. Explicitly

$$\zeta = b\, \zeta_W(\cdot - a') + jc'\delta_{a'} + (b - j)c'\delta_0, \qquad \xi = b\, \xi_W(\cdot - c') + ja'\delta_0 + (b - j)a'\delta_{c'}.$$

Since $b \geq 1$, $\zeta(0)$ and $\xi(0)$ are not both 0. If both are positive, j is interior, $a' = \min\{\alpha > 0 : \zeta(\alpha) \neq 0\}$, $c' = \min\{\gamma > 0 : \xi(\gamma) \neq 0\}$, $j = b - \zeta(0)/c'$, and ζ_W follows; a binary word is determined by its z-profile and its number of x's, so W is determined. If $\xi(0) = 0$ then $j = 0$ and $\zeta = b\, \zeta_{\omega_-}$; if $\zeta(0) = 0$ then $j = b$ and $\zeta = b\, \zeta_{\omega_+}$; either way the profile determines the join word, the complete invariant for boundary j (Theorem 4.2). □

In the formalization all moments are read off a single integer valuation: the coefficient of $s^p q t^r$ is recovered from one Kronecker point $s = \beta$, $q = \beta^{a+b+c+1}$, $t = \beta^{(a+b+c+1)^2}$ with $\beta = 2^{a+b+c+2} + 1$, using an explicit bound on all coefficients (`Coeff.coef_of_map`).

4.4 Theorem G2

Theorem 4.6 (G2)**.** *In an interval word, two bracketings are equal maps iff they lie in the same* D_N*-component. Every kernel class is a* D_N*-component and is flip-connected.*

Proof. If T, T' lie in different components, by Theorem 4.1 they differ on a triple with letters $p < q < r$. Kill every other letter (Theorem 3.6); both restrictions are bracketings of $p^a q^b r^c$ that still differ on that triple, hence lie in different components (Theorem 4.1), hence are different maps (Theorem 4.5); killing is a specialisation, so $T \nsim T'$. Conversely, by Theorem 3.5 every D_N-flip of an interval word uses at most two letters or has two segments that are powers of the same letter, so by Theorem 3.4 its associator vanishes and it stays in the class. □

Corollary 4.7. *If all letters of w are distinct, the C_{n-1} bracketings are pairwise distinct maps, so the associative spectrum of $\mathbb{O}$ is the Catalan sequence.*

Proof. This follows from Theorem 4.6, but directly: distinct planar binary trees induce different shapes on some leaf triple; kill every other letter, and $[x, y, z] \not\equiv 0$. □

The direct proof uses only a two-sided identity (to kill letters) and one non-associative triple, so it shows $s_n(G) = C_{n-1}$ for every non-associative magma G with a two-sided identity. We have not found this stated explicitly in the associative-spectrum literature (Huang and Lehtonen [3] treat related questions for groupoids with a neutral element), but given its two-line proof we expect it to be known.

5 Non-interval words: Theorem G3

Throughout this section w uses exactly three letters and is not interval. Write $[i, i']$ for intervals of positions and identify a bracketing with its set of internal nodes.

5.1 Stable nodes

Lemma 5.1 (self-certifying sets, R′)**.** *Let T be a bracketing and $\mathcal{S}$ a set of nodes of T such that every rotation that would delete some $J \in \mathcal{S}$ has a rank-3 count matrix; here a rotation deleting J has children J_1, J_2 of J and a sibling I, with $J_1, J_2, I, J \cup I$ laminar with $\mathcal{S}$ and the created node laminar with $\mathcal{S} \setminus \{J\}$. Then every bracketing in the* D_N*-component of T contains $\mathcal{S}$.*

Proof. On a D_N-path from T take the first bracketing that loses some $J \in \mathcal{S}$. Its predecessor contains $\mathcal{S}$, so the deleting rotation satisfies the laminarity conditions, has rank 3, and is not a D_N-flip. □

5.2 Caterpillars

Let the run sequence of w be $s_0 \cdots s_{R-1}$. A *caterpillar* is a sequence of steps peeling runs off the ends of the current interval of runs: **L** peels the left run, **R** the right run, and **D** both end runs when they carry the same letter; it stops at a *bottom* of two runs. Its intervals $J_0 = w \supset J_1 \supset \cdots \supset J_h$ give $\mathcal{S} = \{J_1, \ldots, J_h\}$. The bracketing T_0 has every J_k as a node; peeled runs hang off the path letter by letter, except where a move below needs a peeled run as a single subtree (the block Y of M-left, and the pattern $x((BC)x)$ of M-mid).

Lemma 5.2 (Lemma Cat)**.** *$\mathcal{S}$ is self-certifying for T_0 provided*

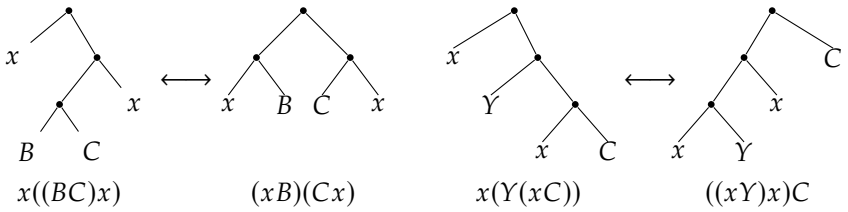


Figure 1: The middle and left Moufang moves used by caterpillar certificates. Each preserves the map and deletes one node (BC, respectively $Y(xC)$) that a self-certifying set protects against D_N-flips.

(C1) every J_k with $1 \le k < h$, entered by peeling letter α and left by peeling letter β, has $\alpha \neq \beta$, and the third letter $\gamma(\alpha,\beta)$ occurs in J_{k+1};

(C2) the bottom J_h, with runs u, v and entry letter α, has α, u, v pairwise distinct.

Proof. Let $J \in \mathcal{S}$ and consider a rotation deleting J. Laminarity with the nearest $\mathcal{S}$-ancestor forces the sibling to be a piece α^i of the run(s) peeled to reach J; laminarity with the next node $J' \in \mathcal{S}$ forces the split of J to cut inside a run peeled next, a piece $\beta^{i'}$, with J' whole on the other side (at the bottom any split is allowed). So two segments are single-letter pieces and $\det N = \pm\, i\, i' \cdot$ (count of $\gamma(\alpha,\beta)$ in the third segment) $\neq 0$, the third segment containing J'. At the bottom a direct 3×3 computation gives $\pm i \cdot (\cdots) \cdot \det(e_\alpha, e_u, e_v) \neq 0$. □

A Moufang move whose repeated variable is a single letter and which deletes a node of $\mathcal{S}$ then joins two different D_N-components (Fig. 1). Three moves arise:

- **M-mid.** A **D** step at $J_k = x^m\, BC\, x^{m'}$, where T_0 contains $x((BC)x)$ with $BC = J_{k+1}$: the move $x((BC)x) \to (xB)(Cx)$ deletes J_{k+1}.
- **M-left.** Two consecutive **L** steps peeling runs u and Y, where the next left run carries the letter of u and the following step is **L** or the bottom: with Y a single subtree, $x(Y(xC)) \to ((xY)x)C$ deletes J_{k+1}.
- **M-right.** The mirror image of M-left.

5.3 Existence of caterpillars

A run sequence is *valid* if it uses three letters and has $R \ge 4$ runs; these are exactly the run sequences of non-interval three-letter words. Let Good(s,η) mean that s has a caterpillar satisfying (C1), (C2) and admitting one of the moves, where the root, *entered* by the letter η, also satisfies (C1) ($\eta = \emptyset$: no entry).

Proposition 5.3 (Proposition E)**.** *For valid s,* Good(s,η) *holds unless (F1) $\eta = s_0 = s_{R-1}$, or (F2) $s_0 = \eta$ and $s_0 \cdots s_{R-2}$ uses only two letters, or (F2′) the mirror image of (F2). In particular* Good$(s,\emptyset)$ *always holds.*

Proof. Induction on R. The base $R \in \{4,5\}$ is a finite check of 240 pairs (s,η). Let $R \ge 6$ and (s,η) not a failure. Once the remainder after a peel is valid it contains all three letters, so the third-letter clause of (C1) holds automatically. For valid s with $R \ge 5$, at least one of $s_1 \cdots s_{R-1}$ and $s_0 \cdots s_{R-2}$ is valid: otherwise both would alternate two letters, and s_0, s_{R-1} would each carry a letter occurring once, which is impossible.

Case $s_0 \neq s_{R-1}$. Peel **L** if $\eta \neq s_0$ and $s_1 \cdots$ is valid: Good$(s_1 \cdots, s_0)$ holds by induction, since (F1), (F2) would need $s_1 = s_0$ and (F2′) would need $s_{R-1} = s_0$. Otherwise peel **R**

symmetrically. Both fail only if, say, $\eta = s_0$ and $\cdots s_{R-2}$ is invalid; then it is two-letter, which is (F2).

Case $s_0 = s_{R-1} = p$, so $\eta \neq p$. Let $s' = s_1 \cdots s_{R-2}$. If s' is valid, step **D**: Good(s', p) holds by induction since p is neither end letter of s', and the **D** step supplies an M-mid move. Otherwise s' alternates two letters $a, b \neq p$; peel **L** and use Good$(s_1 \cdots, p)$: (F1), (F2) fail as $s_1 \neq p$, and (F2′) fails as $s_2 \cdots$ contains a, b, p. □

Theorem 5.4 (G3)**.** *Every non-interval word with at least three letters has two bracketings related by a Moufang move (hence equal maps) that lie in different* D_N*-components; their kernel class is not flip-connected.*

Proof. For three letters: Theorem 5.3 gives a caterpillar, Theorem 5.2 a self-certifying set and a Moufang move deleting one of its nodes, and Theorem 5.1 separates the two ends. With more letters, choose p, q from a pattern $p \cdots q \cdots p$ and any third letter, kill the rest and lift (Theorem 3.6). Theorem 3.3 turns separation into disconnection. □

6 The main theorem and the count

Proof of Theorem A. If w has at most two letters, all bracketings are equal maps by Artin's theorem, and K_n is connected. If w is interval, Theorem 4.6 gives flip-connected classes. Otherwise Theorem 5.4 gives a disconnected class. □

Words up to renaming correspond to their canonical forms, the restricted growth strings: $w_1 = 0$ and $w_{i+1} \leq \max(w_1, \ldots, w_i) + 1$. There are Bell$(n)$ of them; refining by the number k of letters gives the Stirling numbers $S(n, k)$, with $S(n+1, k+1) = (k+1)S(n, k+1) + S(n, k)$.

Proof of Theorem B. Among restricted growth strings of length $n \geq 1$, those with at most two letters are the words in $\{0, 1\}^n$ starting with 0: 2^{n-1}. Interval ones are the non-decreasing ones, determined by the set of positions where a new letter starts: 2^{n-1}. Both: 0^n or $0^k 1^{n-k}$, i.e. n words. So $2^{n-1} + 2^{n-1} - n = 2^n - n$ words have only flip-connected classes (Theorem A), and Bell$(n) - 2^n + n$ do not (Table 1). □

n	Bell(n)	≤ 2 letters	interval	both	disconnected
3	5	4	4	3	0
4	15	8	8	4	3
5	52	16	16	5	25
6	203	32	32	6	145
7	877	64	64	7	756
8	4140	128	128	8	3892
9	21147	256	256	9	20644
10	115975	512	512	10	114961

Table 1: Words up to renaming. For $3 \leq n \leq 9$ the last column was also obtained by an exact classifier that decides every word independently of the proof (Section 7.2).

Statement	Lean name
Moufang identities, $[e_1, e_2, e_4] = 2e_7$	`Octonion.moufang_left/right/middle`, `lemmaW_span`
L1, context cancellation	`Octonion.context_cancel`, `flip_iff_assoc`
Lemma W (series form; over $\mathbb{Z}$)	`Octonion.lemmaW_expansion`; `WExist.lemmaW_exists`
Theorem 3.3	`WExist.kerflip_dn`
Artin, Lemma V	`Octonion.artin`, `lemmaV_two_letters`, `lemmaV_powers(_AC,_BC)`
Lemma I	`Octonion.propC_forward`, `Closure.dn_assoc_zero`
Lemma K (restriction, Moufang lift)	`kill_DN`, `kill_reach`, `lift_mouf`, `lift_separates`
Proposition C	`Octonion.propC_iff`
Proposition T (classification)	`PropT.propT`
Lemma M	`LemmaM.lemmaM`
Proposition S	`PropS.propS_interior`, `Profiles.key_of_K`, `Coeff.coef_of_map`
Theorem G2, both directions	`G2Final.g2`, `Closure.reach_ev`, `Closure.g2_iff`
Proposition E	`Caterpillar.propE`
Lemma Cat, Lemma R′	`Caterpillar.chain_selfCert`, `stable_of_reach`
Theorem G3 (three letters)	`Caterpillar.g3_trees`
Main theorem	`MainThm.main`
Count	`Count.count_disconnected`
Canonical forms, Stirling, Bell sum	`Canon.canonical`, `stirling_rec`, `bell_eq_sum`

Table 2: Paper statements and their Lean counterparts. `#print axioms` for each name is committed in `data/lean_axioms.log`: each depends only on (a subset of) `propext`, `Classical.choice` and `Quot.sound`. Not formalized: the count in the remark after Theorem 4.2, and Theorem 4.7 (a direct consequence of `G2Final.g2`).

7 Formal verification

7.1 Design

The formalization is a Lean 4 (v4.15) project using only the core library. Octonions are `Fin 8` → `Int` with a Cayley–Dickson table; bracketings are planar binary trees over positions. Two representations coexist: trees over consecutive positions with letters in $\{0, 1, 2\}$ and interval spans, used for G3, and trees over arbitrary positions with natural-number letters, used on the octonion side; the bridge (`MainThm.dn_bridge`, `reach_bridge`, `relab`, `rn_reach`) transports Moufang moves and D_N-reachability between them.

Polynomial identities. Octonion identities are verified by a normaliser on integer polynomials: both sides are expanded into monomial lists, sorted and collapsed, and the difference is checked to be empty by kernel evaluation (`decide +kernel`); a correctness theorem transports the check to all integer octonions.

Lemma W over $\mathbb{Z}$. The analytic expansion of Lemma W is replaced by congruences modulo M^4 (Theorem 3.2), giving an explicit integer witness.

Separation in G2. The moments are recovered from one Kronecker point (digits in base β), which avoids polynomial rings.

Main statements. With `wd wl p` the p-th letter of the list `wl`,

```
theorem main (wl : List Nat) :
   KerConn (wd wl) wl.length ↔ TwoLetters wl ∨ Interval wl
```

where `KerConn w n` says that any two bracketings of $0, \dots, n-1$ with equal values at all integer valuations are joined by flips that keep the map, and

```
theorem count_disconnected (m : Nat) :
   ((rgs (m+1)).filter (fun w => @decide (¬ KerConn (wd w) w.length)
      (Classical.propDecidable _))).length + 2^(m+1)
      = bell (m+1) + (m+1)
```

with `bell n` the number of restricted growth strings of length n. `Canon.canonical` proves that every word is a renaming of exactly one of them, `Canon.kerconn_rename` that `KerConn` is renaming-invariant, and `Canon.bell_eq_sum`, `stirling_rec` identify `bell` with the Bell numbers.

Trusted base. One trusts the Lean kernel, the three standard axioms, and the definitions entering the main statements: the multiplication table and product (`Octonion.lean`), trees (`RPrime.lean`), evaluation, contexts and filling (`Division.lean`), leaf lists, D_N-steps and reachability (`Lift.lean`, `Kill.lean`), flips and in-class flips (`Closure.lean`), and `Main.lean`, `Count.lean`, `Canon.lean`. The equivalence of equality on $\mathbb{Z}^8$ and on $\mathbb{R}^8$ (Section 1) is the only informal step.

7.2 Further checks

A second implementation, in Prolog and developed in the same AI-assisted process, runs the constructions of the proofs as programs and checks their outputs: the 44,670 caterpillar certificates produced by the induction of Theorem 5.3 up to 12 runs; G3 end to end for all 4,330 non-interval three-letter words with $n \le 9$, including exact D_N-components; the integer witness of Theorem 3.2 on all 20,984 rank-3 flips of words with $n \le 6$; Propositions T, S, C and Lemma M exhaustively on small words; and the counts of Table 1 up to $n = 10$. Independently of the proof, an exact classifier decides every word with $3 \le n \le 9$ (26,439 words): for non-interval words it finds a move valid in every alternative algebra between two D_N-components, and for interval words it separates components by invariants or exact integer evaluation. Its output agrees with Theorem A.

8 Remarks and questions

Locality. A stronger, local form of G3 asks for a Moufang move toggling an *absolutely rigid* interval, one whose every split has rank 3 against every adjacent nonempty interval. It holds for $n \le 9$ but fails in general: $xyxyzzyxyx$ and $xyyxzzxyyx$ ($n = 10$) and four words with $n = 12$ have no rigid separator, although G3 holds for them. Self-certifying sets (Theorem 5.1) repair this: their nodes certify each other rather than individually.

Other algebras. Kernel classes are defined by polynomial identities with integer coefficients, which hold over $\mathbb{O}$ iff they hold over $\mathbb{O} \otimes \mathbb{C}$. The split octonions are another real form of $\mathbb{O} \otimes \mathbb{C}$, so they have the same kernel classes, and Theorems A and B hold for them verbatim. The proof uses that $\mathbb{O}$ is alternative (Moufang, Artin), that generic elements are invertible (Theorem 3.1), and that three units have a nonzero associator (Theorem 3.2). The same question for the free alternative algebra [12], where the identities could differ, is open.

Quantitative refinements. A closed form for the number of kernel classes of an arbitrary interval word, i.e. the number of realisable triple-shape functions τ in Theorem 4.1, would refine Theorem 4.7 and the remark after Theorem 4.2.

Acknowledgements and disclosure

The formal proofs, the Prolog checks and much of the exposition were developed with the assistance of Claude (Anthropic), an AI system, under the author's direction. The statements listed in Table 2 are checked by the Lean kernel; the remaining claims are marked as such in the text. The repository contains the full history, the proofs and the logs of every computation. An AI referee report was obtained and addressed before release; it is not a substitute for human review, which this preprint invites.

References

[1] J. C. Baez, The octonions, *Bull. Amer. Math. Soc.* 39 (2002), 145–205.

[2] B. Csákány and T. Waldhauser, Associative spectra of binary operations, *Multiple-Valued Logic* 5 (2000), 175–200; arXiv:1102.2124.

[3] J. Huang and E. Lehtonen, The associative-commutative spectrum of a binary operation, arXiv:2202.11826.

[4] E. Lehtonen and T. Waldhauser, Associative spectra of graph algebras II. Satisfaction of bracketing identities, spectrum dichotomy, *J. Algebraic Combin.*, doi:10.1007/s10801-021-01061-7.

[5] E. Lehtonen and T. Waldhauser, Associativity conditions for linear quasigroups and equivalence relations on binary trees, arXiv:2310.09184.

[6] S. Liebscher and T. Waldhauser, On associative spectra of operations, *Acta Sci. Math. (Szeged)* 75 (2009); arXiv:1102.2337.

[7] L. de Moura and S. Ullrich, The Lean 4 theorem prover and programming language, in *CADE-28*, LNCS 12699, Springer, 2021, 625–635.

[8] R. Moufang, Zur Struktur von Alternativkörpern, *Math. Ann.* 110 (1935), 416–430.

[9] R. D. Schafer, *An Introduction to Nonassociative Algebras*, Academic Press, 1966.

[10] J. D. Stasheff, Homotopy associativity of H-spaces I, II, *Trans. Amer. Math. Soc.* 108 (1963), 275–312.

[11] D. Tamari, The algebra of bracketings and their enumeration, *Nieuw Arch. Wisk.* (3) 10 (1962), 131–146.

[12] K. A. Zhevlakov, A. M. Slin'ko, I. P. Shestakov and A. I. Shirshov, *Rings that are nearly associative*, Academic Press, 1982.