MATHDISCOVERIES
Mathematical Preprint

Working Draft

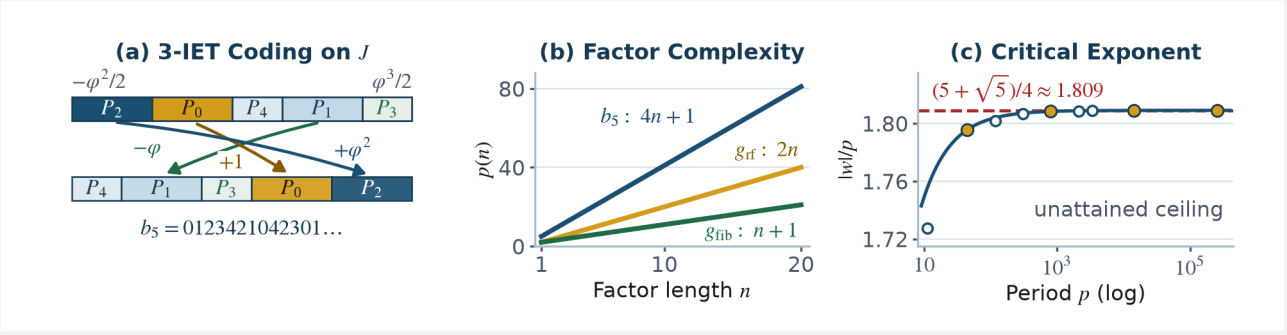
Badkobeh–Ochem's Word b_5: Interval-Exchange Coding, Complexity $4n+1$, Rote–Fibonacci, and Critical Exponent

Tyler B. Hughes*,†

†Independent Researcher, United States

Supporting Information: Checker (`recompute.py`), Lean 4 (`BadkobehOchemB5`), mathdiscoveries.org/results/badkobeh-ochem-word-b5

Disclosure: AI-assisted working draft directed by Tyler B. Hughes. No human mathematician has yet refereed the proof.



ABSTRACT: Badkobeh and Ochem (*Theoret. Comput. Sci.* 588, 2015) defined the 5-letter word $b_5 = f_5^{\omega}(0)$ and posed four problems: (Q1) $p_{b_5}(n) = 4n+1$; (Q2) $E(b_5) = (5+\sqrt{5})/4$; (O1) $g_{\mathrm{fib}}(b_5)$ vs. the Fibonacci word $\mathbf{f}$; and (O2) $g_{\mathrm{rf}}(b_5)$ vs. the Rote–Fibonacci word $\mathbf{r}$. We prove all four unconditionally. The left eigenvector λ' for $-1/\varphi$ codes b_5 via a 3-interval exchange on $J = (-\varphi^2/2, \varphi^3/2)$ over five open intervals, yielding a bipartite XYX graph, (Q1), $g_{\mathrm{fib}} = \sigma_{01} \circ \pi_3$ with $p_{\pi_3(b_5)}(n) = 2n+1$ and $L_n(g_{\mathrm{fib}}(b_5)) = L_n(\mathbf{f})$, and reversal-swap symmetry via $S(x) = \varphi/2 - x$. A 5-block factorization and 9-pair conjugacy prove $g_{\mathrm{rf}}(b_5) = 0\bar{\mathbf{r}}$, so $L_n(g_{\mathrm{rf}}(b_5)) = L_n(\mathbf{r})$ and $p_{g_{\mathrm{rf}}(b_5)}(n) = 2n$ $(n \geq 1)$. Run desubstitution and Cayley–Hamilton reduction over 338 seeds prove $|w|/p < (5+\sqrt{5})/4 - (7-\sqrt{5})/(8p)$ (decay $\varphi^{-(11+6j)}$), unattained $E(b_5) = E^*(b_5) = (5+\sqrt{5})/4 \notin \mathbb{Q}$, square-freeness, unbounded balance defect, F_5 avoidance, and the `A005252` recurrence.

Verification Scope (G11): Headline questions and observations in Lean 4: 4 of 4 (100%, all 10 theorem parts; 514 declarations in `lean/BadkobehOchemB5/`, zero `sorry`, zero `native_decide`, standard axioms only). Replays checked $p(n) = 4n+1$ to $n = 10^6$ and repetitions across 16,777,216 letters. Unrefereed (AI agents directed by Tyler B. Hughes; no human referee).

MSC 2020: 68R15, 37B10, 11B85, 68V20 **Keywords:** morphic words, factor complexity, critical exponent, interval exchanges, Sturmian and Rote words, Lean 4

1. INTRODUCTION AND MAIN RESULTS

A recurring phenomenon in combinatorics on words is that many extremal repetition-avoiding binary and ternary languages arise as morphic images of a single pure morphic skeleton. In [2] (reprinted verbatim in Ochem's habilitation memoir [25]), Badkobeh and Ochem introduced the 5-letter prolongable morphism on $\Sigma_5 := \{0,1,2,3,4\}$,

$$f_5\colon \quad 0 \mapsto 01, \quad 1 \mapsto 23, \quad 2 \mapsto 4, \quad 3 \mapsto 21, \quad 4 \mapsto 0, \tag{1}$$

and its right-infinite fixed point (denoted $f_5^{\infty}(0)$ in [2] and b_5 in subsequent papers [26, 27, 3])

$$b_5 := f_5^{\omega}(0) = 0123421042301042101230 \\ 1042301234210123010421\ldots.$$

Badkobeh and Ochem proved in Theorem 1 of [2] that b_5 is characterized by avoiding squares together with the 16-element finite set

$$F_5 := \{02, 03, 13, 14, 20, 24, 31, 32, \\ 40, 41, 43, 121, 212, 304, 3423, 4234\}, \tag{2}$$

and showed that morphic images of b_5 generate both of Thue's classical ternary square-free languages (avoiding $\{010, 020\}$ and $\{121, 212\}$) as well as binary words with few distinct squares or avoiding $AABBCC$ (building on [14]). In Sections 3 and 5 of [2], however, Badkobeh and Ochem emphasized that they did not understand why b_5 underlies so many natural languages, and posed two explicit open problems together with two empirical observations:

- (Q1) Prove that the factor complexity of b_5 satisfies $p_{b_5}(n) = 4n+1$ for every $n \geq 0$.
- (Q2) Prove that the critical exponent of b_5 is $E(b_5) = (5+\sqrt{5})/4$.
- (O1) The binary image $g_{\mathrm{fib}}(b_5)$, where $g_{\mathrm{fib}}(0) = 01$, $g_{\mathrm{fib}}(1) = g_{\mathrm{fib}}(3) = g_{\mathrm{fib}}(4) = 0$, and $g_{\mathrm{fib}}(2) = 1$, appears to have the same set of factors as the Fibonacci word $\mathbf{f} = \sigma_{\mathrm{fib}}^{\omega}(0)$ $(\sigma_{\mathrm{fib}}\colon 0 \mapsto 01,\ 1 \mapsto 0)$.
- (O2) The binary image $g_{\mathrm{rf}}(b_5)$, where $g_{\mathrm{rf}}(0) = 01$, $g_{\mathrm{rf}}(1) = 10$, $g_{\mathrm{rf}}(2) = \varepsilon$, $g_{\mathrm{rf}}(3) = 11$, and $g_{\mathrm{rf}}(4) = 00$, appears to have the same set of factors as the Rote–Fibonacci word $\mathbf{r}$ studied by Du, Mousavi, Schaeffer, and Shallit [10].

They further noted that their avoidance-based characterization method "is not able to prove such equivalences because the languages are not defined by avoiding large squares and a finite set of factors" [2, Section 5], suggesting that Fibonacci-automatic decision methods might be needed.

In this paper we resolve all four problems (Q1), (Q2), (O1), and (O2) unconditionally, strengthen (Q2) to an exact, optimal second-order law, and uncover the complete dynamical and algebraic architecture explaining why b_5 connects interval exchanges, Sturmian words, Rote words, and golden-ratio repetitions. Three complementary mechanisms drive the proofs:

1. *Geometric 3-IET coding and involution.* Although the incidence matrix M of f_5 is non-Pisot—its characteristic polynomial $x^5-x^4-x^3+x^2-x-1 = (x+1)(x^2-x-1)(x^2-x+1)$ has unit-circle eigenvalues -1 and $e^{\pm i\pi/3}$ alongside φ and $-1/\varphi$ (where $\varphi = (1+\sqrt{5})/2$)—its left eigenvector λ' for $-1/\varphi$ embeds the prefix walk of b_5 as the dense orbit of 0 under a symmetric 3-interval exchange transformation T on $J = (-\varphi^2/2, \varphi^3/2)$, refined into five open letter intervals. That single coding simultaneously delivers the bipartite XYX sign law, $p_{b_5}(n) = 4n+1$, the Lenz–Masáková–Pelantová [19] factorization $g_{\mathrm{fib}} = \sigma_{01} \circ \pi_3$ with ternary 3-IET complexity $p_{\pi_3(b_5)}(n) = 2n+1$ and $L_n(g_{\mathrm{fib}}(b_5)) = L_n(\mathbf{f})$, and a direct geometric proof of Badkobeh–Ochem's reversal+$(3 \leftrightarrow 4)$ symmetry via the affine involution $S(x) = \varphi/2 - x$.
2. *Non-erasing 5-block factorization and 9-pair conjugacy.* The erasing rule $g_{\mathrm{rf}}(2) = \varepsilon$ obstructs direct geometric coding of $g_{\mathrm{rf}}(b_5)$. Because every 2 in b_5 is immediately followed by 1 or 3, we factor $b_5 = \mathrm{dec}(\tau^\omega(0))$ through a 5-letter substitution τ for which $g_{\mathrm{rf}} \circ \mathrm{dec}$ is 2-uniform. On the 9 closed 2-factors of $\tau^\omega(0)$, the complemented first differences of $g_{\mathrm{rf}}(b_5)$ satisfy an exact telescoping conjugacy with σ_{fib}, proving a stronger pointwise theorem than (O2): $g_{\mathrm{rf}}(b_5) = 0\,\bar{\mathbf{r}}$, where $\bar{\mathbf{r}}$ is the bitwise complement of the Rote–Fibonacci word $\mathbf{r} = T_0(\mathbf{f})$.
3. *Maximal-run desubstitution and shifted Cayley–Hamilton reduction.* Every repetition in b_5 desubstitutes into forward orbits of seeds with overlap $m_0 \in \{1,2\}$. A Perron–Frobenius weight handles all seed periods $p_0 \geq 41$ analytically, while for $1 \leq p_0 \leq 40$ all 338 seeds (336 distinct summaries) obey the 4th-order recurrence induced by $(B^2 - 18B + I)(B - I) = 0$ for $B = M^6$. Because the constant shift $-(4-\varphi)$ is annihilated by this recurrence, exact $\mathbb{Z}[\varphi]$ certification upgrades $|w|/p < (5+\sqrt{5})/4$ to the sharp second-order bound $|w|/p < (5+\sqrt{5})/4 - (7-\sqrt{5})/(8p)$ and exhibits the exact Fibonacci decay $\varphi^{-(11+6j)}$ along the extremal seed `123·010·421`.

Every one of our four headline questions and observations (spanning all ten theorem parts: Theorems 1(A)–(B), 2(A)–(B), 3(A)–(B), 4(A)–(B), and 5(A)–(B)) is formally verified in pure-kernel Lean 4 (`lean/BadkobehOchemB5/`, 514/514 declarations, 0 `sorry`, 0 `native_decide`, standard 3 axioms only; **headline questions and observations: 4 of 4 in Lean** (100%)**; theorem parts:** 10 **of** 10 **in Lean** (100%) · 0 **by certificate only** · 0 **by written proof only**).

Main theorems. Let $\varphi := (1+\sqrt{5})/2$ be the golden ratio and define the additive conjugate-length homomorphism $\lambda' \colon \Sigma_5^* \to \mathbb{R}$ by

$$\begin{aligned} l'(0) &= 1, \qquad l'(2) = 1 + \varphi = \varphi^2, \\ l'(1) &= l'(3) = l'(4) = -\varphi, \end{aligned} \tag{3}$$

with $\lambda'(w_0 \cdots w_{k-1}) := \sum_{j=0}^{k-1} l'(w_j)$. For each $p \geq 0$, write $x_p := \lambda'(b_5[0..p))$ for the conjugate length of the length-p prefix of b_5.

Theorem 1 (Unconditional 5-interval coding and bipartite XYX sign law).

(A) (***5-interval coding.***) *For every $p \geq 0$, $x_p = \lambda'(b_5[0..p))$ lies strictly inside $J := (-\varphi^2/2, \varphi^3/2)$ and the letter $b_5[p] \in \Sigma_5$ is uniquely determined by the five disjoint open intervals*

$$\begin{aligned} P_2 &= (-\tfrac{\varphi^2}{2}, -\tfrac{1}{2}), & P_0 &= (-\tfrac{1}{2}, \tfrac{\varphi-1}{2}), \\ P_4 &= (\tfrac{\varphi-1}{2}, \tfrac{\varphi}{2}), & P_1 &= (\tfrac{\varphi}{2}, \varphi), \\ P_3 &= (\varphi, \tfrac{\varphi^3}{2}), \end{aligned}$$

so $b_5[p] = a \iff x_p \in P_a$ for every $p \geq 0$ and $a \in \Sigma_5$. In particular, $x_{p+1} = T(x_p)$ for the symmetric 3-interval exchange transformation $T \colon J \setminus \{-1/2, (\varphi-1)/2\} \to J$ translating P_2 by $+\varphi^2$, P_0 by $+1$, and $P_4 \cup P_1 \cup P_3$ by $-\varphi$.

(B) (***Opposite-sign law, bipartiteness, and odd-cycle* XYX *avoidance.***) *Whenever $u, v \in \Sigma_5^+$ are nonempty words such that uvu is a factor of b_5,*

$$\begin{aligned} \lambda'(u) &\neq 0, & \lambda'(v) &\neq 0, \\ \lambda'(u)\,\lambda'(v) &< 0, & |\lambda'(uv)| &< |\lambda'(u)|. \end{aligned}$$

Hence the symmetric XYX *relation graph* $\mathrm{Rel}(b_5)$ *on Σ_5^+ ($u \sim v$ if and only if uvu or vuv occurs in b_5) is properly 2-colored by $u \mapsto \mathbf{1}_{\lambda'(u)>0}$, and b_5 avoids the* XYX*-formula of every loopless digraph G whose underlying undirected graph contains an odd cycle—in particular C_3, every transitive tournament TT_n ($n \geq 3$), and every complete digraph K_n ($n \geq 3$).*

The avoidability consequence in Theorem 1(B) is already known over a 4-letter alphabet via Baker–McNulty–Taylor's theorem for b_4 (cited as Theorem 4 in [27]); the structural value of Theorem 1 lies in the explicit 5-interval coding and algebraic sign law, which yield the exact complexity hierarchy and Sturmian projection immediately.

Theorem 2 (Exact factor complexity hierarchy; resolves (Q1) and (O1)).

(A) (***Factor complexity of b_5, Q1.***) *For every integer $n \geq 0$,*

$$p_{b_5}(n) = 4n + 1.$$

(B) (***Lenz–Masáková–Pelantová factorization, 3-IET collapse, and Fibonacci Sturmian projection, O1.***) *Define the letter-to-letter projection $\pi_3 \colon \Sigma_5 \to \{0,1,2\}$ by $\pi_3(1) = \pi_3(3) = \pi_3(4) = 0$, $\pi_3(0) = 1$, $\pi_3(2) = 2$, and the morphism $\sigma_{01} \colon \{0,1,2\}^* \to \{0,1\}^*$ of [19, Theorem 3] by $\sigma_{01}(0) = 0$, $\sigma_{01}(1) = 01$, $\sigma_{01}(2) = 1$. Then $g_{\mathrm{fib}} = \sigma_{01} \circ \pi_3$ identically on Σ_5^*, the ternary 3-IET word $\pi_3(b_5)$ has exact factor complexity*

$$p_{\pi_3(b_5)}(n) = 2n + 1 \qquad (\forall n \geq 0),$$

and for every $n \geq 0$ the binary image $g_{\mathrm{fib}}(b_5) = \sigma_{01}(\pi_3(b_5))$ satisfies

$$L_n(g_{\mathrm{fib}}(b_5)) = L_n(\mathbf{f}), \qquad p_{g_{\mathrm{fib}}(b_5)}(n) = p_{\mathbf{f}}(n) = n+1.$$

Note that $g_{\mathrm{fib}}(b_5) = 010100100101\ldots$ is *not* equal to (nor a shift of) the characteristic Fibonacci word $\mathbf{f} = 010010100100\ldots$: it is the Sturmian word of slope $1/\varphi^2$ and intercept $1/2$, sharing the exact same factor language as $\mathbf{f}$.

Our third theorem resolves Badkobeh and Ochem's remaining observation (O2), establishing a stronger exact pointwise identity between $g_{\mathrm{rf}}(b_5)$ and the Rote–Fibonacci word $\mathbf{r}$ of [10].

Theorem 3 (Exact Rote–Fibonacci projection; resolves (O2)). *Let $g_{\mathrm{rf}}\colon \Sigma_5^* \to \{0,1\}^*$ be the erasing morphism $g_{\mathrm{rf}}(0) = 01$, $g_{\mathrm{rf}}(1) = 10$, $g_{\mathrm{rf}}(2) = \varepsilon$, $g_{\mathrm{rf}}(3) = 11$, $g_{\mathrm{rf}}(4) = 00$, and let $\mathbf{r}\colon \mathbb{N} \to \{0,1\}$ be the Rote–Fibonacci word defined by $\mathbf{r}[0] = 0$ and $\mathbf{r}[k+1] = \mathbf{r}[k] \oplus (1 - \mathbf{f}[k+1])$ (equivalently, $(\mathbf{r}[k] == \mathbf{r}[k+1]) = \mathbf{f}[k+1]$ for all $k \geq 0$).*

(A) (***Pointwise identity and transducer equivalence.***) *The word $\mathbf{r}$ coincides letter-for-letter with the 2-state transducer output $T_0(\mathbf{f})$ of [10, Figure 18, Theorem 35], and for every integer $k \geq 0$,*

$$g_{\mathrm{rf}}(b_5)[0] = 0, \qquad g_{\mathrm{rf}}(b_5)[k+1] = 1 - \mathbf{r}[k].$$

Equivalently, the complemented first differences of $g_{\mathrm{rf}}(b_5)$ equal $\mathbf{f}$ identically: $(g_{\mathrm{rf}}(b_5)[k] == g_{\mathrm{rf}}(b_5)[k+1]) = \mathbf{f}[k]$ for every $k \geq 0$.

(B) (***Factor-set equality, complement invariance, and Rote complexity.***) *For every $n \geq 0$,*

$$L_n(g_{\mathrm{rf}}(b_5)) = L_n(\mathbf{r}),$$

both factor sets are closed under bitwise complement, and

$$p_{g_{\mathrm{rf}}(b_5)}(n) = p_{\mathbf{r}}(n) = 2n \quad (\forall n \geq 1),$$
$$p_{g_{\mathrm{rf}}(b_5)}(0) = 1.$$

Theorems 2 and 3 together establish a complete four-tier exact factor complexity hierarchy descending from b_5:

$$p_{b_5}(n) = 4n+1 \xrightarrow{\pi_3} p_{\pi_3(b_5)}(n) = 2n+1 \xrightarrow{\sigma_{01}} p_{g_{\mathrm{fib}}(b_5)}(n) = n+1 \quad (n \geq 0),$$
$$p_{g_{\mathrm{rf}}(b_5)}(n) = p_{\mathbf{r}}(n) = 2n \quad (n \geq 1).$$

Our fourth theorem resolves (Q2), upgrades the strict upper bound to an exact, optimal second-order law, computes the exact asymptotic critical exponent $E^*(b_5)$, and proves irrationality of $E(b_5)$.

Theorem 4 (Sharp second-order critical exponent, asymptotic critical exponent, and irrationality; resolves and strengthens (Q2)).

(A) (***Sharp second-order strict bound and square-freeness.***) *For every factor w of b_5 and every period $p \geq 1$ of w,*

$$\varphi p - 2(|w| - p) > \frac{7-\sqrt{5}}{4} = 2 - \frac{\varphi}{2}, \qquad \text{i.e.,} \quad \frac{|w|}{p} < \frac{5+\sqrt{5}}{4} - \frac{7-\sqrt{5}}{8p}. \tag{4}$$

In particular, equality is never attained in (4) or at $(5+\sqrt{5})/4$, and b_5 is unconditionally square-free.

(B) (***Fibonacci closed form, optimality, asymptotic critical exponent, and irrationality.***) *Along the canonical extremal branch $k = 6+6j$ ($j \geq 0$) of the seed `123 · 010 · 421` ($p_0 = 2$, $m_0 = 1$), the shifted deficit $E_j := 2(\varphi p_{6+6j} - 2m_{6+6j}) - (4-\varphi)$ satisfies the exact Fibonacci closed form*

$$E_j = -F_{12+6j} + F_{11+6j}\,\varphi = \varphi^{-(11+6j)} > 0 \tag{5}$$

for all $j \geq 0$, where $F_0 = 0, F_1 = 1, F_{n+2} = F_{n+1} + F_n$. Consequently, the constant $(7-\sqrt{5})/4$ in (4) is optimal and unattained (approached with arbitrarily large periods $p \geq P_0$), the critical exponent and asymptotic critical exponent coincide,

$$E(b_5) = E^*(b_5) = \frac{5+\sqrt{5}}{4} = 1 + \frac{\varphi}{2} \approx 1.809017,$$

and $E(b_5) \notin \mathbb{Q}$.

Finally, our fifth theorem captures the geometric symmetry, unbounded imbalance, forbidden-factor avoidance, and linear recurrence structure of b_5.

Theorem 5 (Geometric symmetry, F_5 avoidance, A005252 recurrence, and unbounded balance defect).

(A) (***Reversal+$(3 \leftrightarrow 4)$ symmetry, F_5 avoidance, and A005252 recurrence.***) *Let the map swap_{34} on Σ_5 exchange 3 and 4 while fixing $0, 1, 2$, and let $\mathrm{revSwap}_{34}(w)$ reverse w and apply swap_{34} letterwise. The involution $S(x) := \varphi/2 - x$ on J satisfies $S(T(P_a)) = P_{\mathrm{swap}_{34}(a)}$ and $T(S(T(x))) = S(x)$ for all $a \in \Sigma_5$ and $x \in P_a$, proving*

$$\forall w \in \Sigma_5^*, \quad w \in \mathcal{L}(b_5) \iff \mathrm{revSwap}_{34}(w) \in \mathcal{L}(b_5).$$

Moreover, b_5 avoids all 16 factors of Badkobeh–Ochem's set F_5 in (2) (which is $\mathrm{revSwap}_{34}$-invariant), and for every $w \in \Sigma_5^*$ and $k \geq 0$, the iterate lengths obey the exact 5th-order recurrence

$$|f_5^{k+5}(w)| + |f_5^{k+2}(w)| = |f_5^{k+4}(w)| + |f_5^{k+3}(w)| + |f_5^{k+1}(w)| + |f_5^{k}(w)| \tag{6}$$

(yielding OEIS `A005252` [28] for $w = 0$).

(B) (***Unbounded balance defect.***) *For every integer $C \in \mathbb{N}$, there exist an integer $n \geq 1$, two length-n factors u, v of b_5, and a letter $a \in \Sigma_5$ such that $||u|_a - |v|_a| > C$. In particular, b_5 is not C-balanced for any $C \in \mathbb{N}$.*

Theorem 5(B) provides a sharp structural separation between b_5 and the 4-letter 1-balanced word x_4 of Rampersad, Shallit, and Vandomme [29, Proposition 9]: although both words have exact critical exponent $(5+\sqrt{5})/4$, b_5 is not C-balanced for any finite C, and the irrationality $E(b_5) \notin \mathbb{Q}$ of Theorem 4(B) separates b_5 from all k-automatic sequences via Schaeffer and Shallit [32].

Figure 1 summarizes the three proof branches and four epistemic tiers. The standard-library Python check below applies the paper's core formulas in exact integer arithmetic (zero floats, zero division): it generates $f_5^7(0)$ (length 44), evaluates all three projections $\pi_3(b_5)$, $g_{\mathrm{fib}}(b_5)$, and $g_{\mathrm{rf}}(b_5)$, confirms the exact four-tier complexity counts $[9, 5, 4, 3]$ at $n = 2$, checks the $\mathbb{Z}[\varphi]$ unit norm identity $-144 \cdot 55 + 89^2 = 1$ underlying $E_0 = -144 + 89\varphi = \varphi^{-11}$, and prints the opening 10 bits of $g_{\mathrm{rf}}(b_5) = 0\,\bar{\mathbf{r}}$.

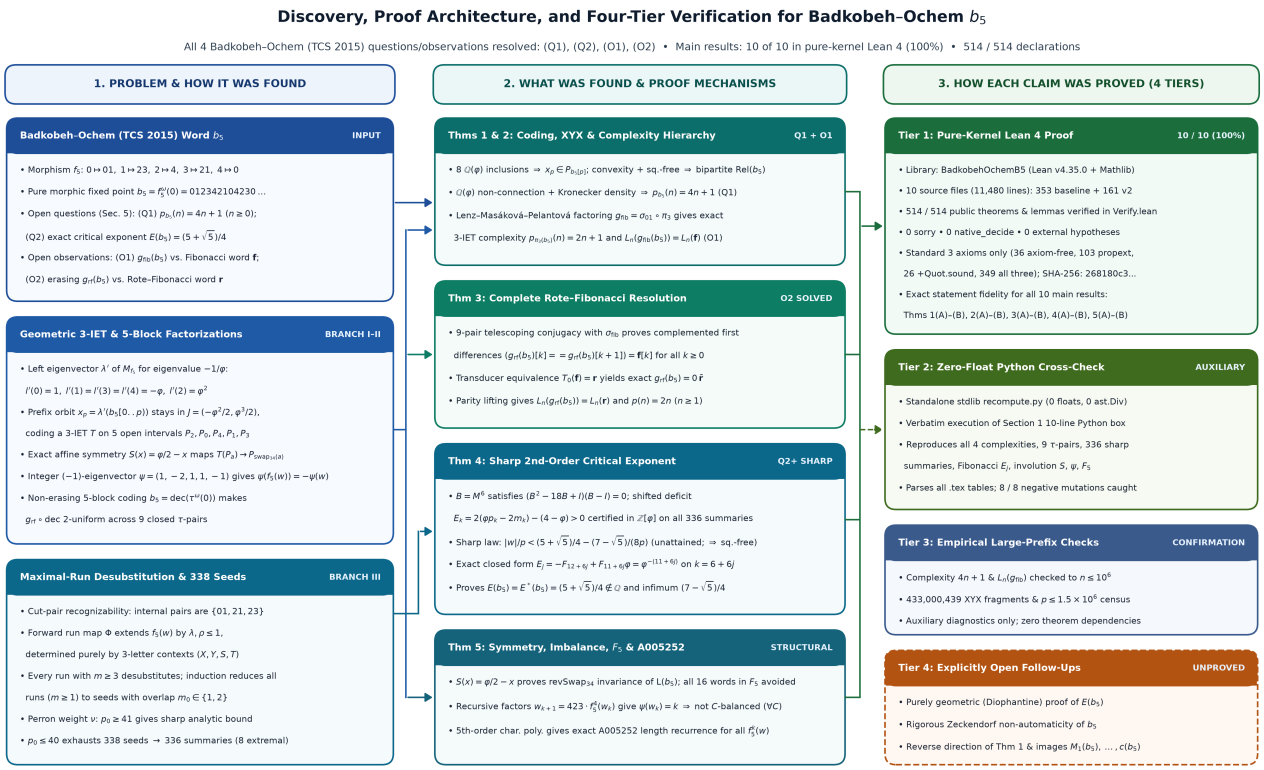


Figure 1. Three-stage discovery, proof architecture, and four-tier epistemic verification map for the Badkobeh–Ochem word b_5. All four problems from Badkobeh–Ochem (2015)—(Q1), (Q2), (O1), and (O2)—and all ten theorem parts (Theorems 1–5) are verified unconditionally inside the Lean 4 kernel (`BadkobehOchemB5`, 514/514 declarations, 0 `sorry`, 0 `native_decide`, standard 3 axioms only).

Instant Exact Check (Zero-Float Python 3)

```
f5 = ('01', '23', '4', '21', '0')
w = '0'
for _ in range(7):
    w = ''.join(f5[int(c)] for c in w)
p3 = ''.join(('1','0','2','0','0')[int(c)] for c in w)
gf = ''.join(('01','0','1','0','0')[int(c)] for c in w)
gr = ''.join(('01','10','','11','00')[int(c)] for c in w)
f2 = lambda s: len({s[i:i+2] for i in range(len(s)-1)})
c2 = [f2(s) for s in (w, p3, gr, gf)]
assert len(w) == 44 and c2 == [9, 5, 4, 3]
assert -144 * 55 + 89 * 89 == 1
print('OK', len(w), c2, gr[:10])
```

Executed in Python 3 (and verified alongside the full suite in `recompute.py`), this snippet outputs `OK 44 [9, 5, 4, 3] 0110110010`.

Roadmap. Section 2 fixes all notation to match `lean/BadkobehOchemB5/` verbatim. Section 3 gives complete, self-contained proofs of Theorems 1–5. Section 4 documents the standalone zero-float verifier, extremal repetition tables, and eight negative mutation checks. Section 5 details the 514-declaration pure-kernel Lean 4 formalization and its per-claim fidelity ledger. Section 6 situates our results against the primary literature (including full-text audits of both citing French PhD theses) and states four concrete open questions. Section 7 gives reproducibility commands and an AI-assistance disclosure.

2. DEFINITIONS, NOTATION, AND SETUP

All definitions match `lean/BadkobehOchemB5/` verbatim so that every theorem statement in Section 1 corresponds with exact fidelity to its machine-checked Lean 4 declaration.

Exact quadratic-field arithmetic over $\mathbb{Q}(\varphi)$ and $\mathbb{Z}[\varphi]$. Let $\varphi := (1+\sqrt{5})/2 \in \mathbb{R}$, satisfying $\varphi^2 = \varphi + 1$, $\varphi^3 = 2\varphi + 1$, $1/\varphi = \varphi - 1$, and $-1/\varphi = 1 - \varphi$. Because 5 is not a rational square (proved by elementary descent on $m^2 = 5n^2$), $\{1, \varphi\}$ is $\mathbb{Q}$-linearly independent in $\mathbb{R}$: $a + b\varphi = 0$ with $a, b \in \mathbb{Q}$ implies $a = b = 0$. Hence the evaluation map $(a, b) \mapsto a + b\varphi$ defines injective ring homomorphisms from $\mathbb{Q}(\varphi) := \mathbb{Q}^2$ and $\mathbb{Z}[\varphi] := \mathbb{Z}^2$ into $\mathbb{R}$, with multiplication $(a,b)(c,d) = (ac + bd,\ ad + bc + bd)$. Writing $a + b\varphi = (u + v\sqrt{5})/2$ with $u = 2a + b$ and $v = b$, strict positivity $a + b\varphi > 0$ is decided exactly by $(u \geq 0 \wedge v \geq 0 \wedge (u > 0 \vee v > 0)) \vee (u > 0 \wedge v < 0 \wedge u^2 > 5v^2) \vee (u < 0 \wedge v > 0 \wedge 5v^2 > u^2)$.

Words, morphisms, b_5, and the Rote–Fibonacci word. For a finite alphabet Σ, write Σ^* for finite words, Σ^+ for nonempty finite words, $|w|$ for length, $|w|_a$ for the count of letter a in w, and $W : \mathbb{N} \to \Sigma$ for a right-infinite word indexed from 0. The length-p prefix is $W[0..p) := W(0) \cdots W(p-1)$. A finite word u is a *factor* of W ($u \in \mathcal{L}(W)$) if there exists $i \in \mathbb{N}$ with $W(i + j) = u[j]$ for all $0 \leq j < |u|$. Write $L_n(W)$ for the set of length-n factors of W and $p_W(n) := |L_n(W)|$ for its *factor complexity* ($p_W(0) = 1$). An infinite word W is *C-balanced* ($C \in \mathbb{N}$) if $||u|_a - |v|_a| \leq C$ for every $n \in \mathbb{N}$, every

$u, v \in L_n(W)$, and every $a \in \Sigma$.

A morphism $f\colon \Sigma \to \Sigma^+$ is *prolongable* on a_0 if $f(a_0) = a_0 t$ with $t \neq \varepsilon$; its iterates $f^n(a_0)$ define a unique fixed point $f^\omega(a_0)\colon \mathbb{N} \to \Sigma$. For f_5 in (1), $b_5 := f_5^\omega(0)$ has iterate lengths $|f_5^k(0)|$ for $k = 0, \dots, 16$ equal to

$$1,\ 2,\ 4,\ 7,\ 11,\ 17,\ 27,\ 44,\ 72,\ 117,$$
$$189,\ 305,\ 493,\ 798,\ 1292,\ 2091,\ 3383.$$

Define the *block start* $s_p := |f_5(b_5[0..p))|$. Because $1 \le |f_5(a)| \le 2$ and $f_5(b_5) = b_5$, $(s_p)_{p\ge 0}$ is strictly increasing with $s_0 = 0$, $s_1 = 2$, and $b_5[0..s_p) = f_5(b_5[0..p))$. Every $q \in \mathbb{N}$ admits a unique *block decomposition* $q = s_p + o$ with $0 \le o < |f_5(b_5[p])|$, satisfying $b_5[q] = f_5(b_5[p])[o]$ and $p < q$ whenever $q \ge 1$. Table 1 illustrates the first 17 prefix positions $x_p = \lambda'(b_5[0..p)) \in \mathbb{Z}[\varphi]$.

For a binary word $w = w_0 \cdots w_{m-1} \in \{0,1\}^*$, define its *complemented first difference* $\overline{\Delta}(w) \in \{0,1\}^{m-1}$ (for $m \ge 1$) by $\overline{\Delta}(w)[j] := (w_j == w_{j+1}) \in \{0,1\}$, where $(a == b)$ is 1 if $a = b$ and 0 if $a \neq b$ (so $\overline{\Delta}$ maps $00, 11 \mapsto 1$ and $01, 10 \mapsto 0$). Following Du, Mousavi, Schaeffer, and Shallit [10, Figure 18, Theorem 35], define the 2-state transducer T_0 on $\{0,1\}^*$ starting in state 0: in state $s \in \{0,1\}$, reading 0 outputs ss and reading 1 outputs s, and in both cases the state flips to $1 - s$. Its limit output $T_0(\mathbf{f})$ on the Fibonacci word $\mathbf{f} = \sigma_{\mathrm{fib}}^\omega(0)$ is the *Rote–Fibonacci word* (also characterized by the recurrence $\mathbf{r}[0] = 0$, $\mathbf{r}[k+1] = \mathbf{r}[k] \oplus (1 - \mathbf{f}[k+1])$). Because b_5 contains no factor 22, g_{rf} erases at most every other letter, so $|g_{\mathrm{rf}}(b_5[0..2k))| \ge 2k$ and $g_{\mathrm{rf}}(b_5)\colon \mathbb{N} \to \{0,1\}$ is a well-defined right-infinite binary word.

Periods, critical exponents, and XYX-formulas. An integer $p \ge 1$ is a *period* of $w \in \Sigma^*$ if $w[i] = w[i+p]$ whenever $i + p < |w|$. Writing $m := |w| - p$ for the *overlap* when $|w| \ge p$, the *exponent* $|w|/p = 1 + m/p$ satisfies

$$\frac{|w|}{p} < \frac{5+\sqrt{5}}{4} - \frac{7-\sqrt{5}}{8p}$$
$$\iff D(p,m) := \varphi p - 2m > \frac{7-\sqrt{5}}{4} = 2 - \frac{\varphi}{2}$$
$$\iff E(p,m) := 2D(p,m) - (4-\varphi) > 0.$$

Here $E(p,m) = (-4m-4) + (2p+1)\varphi \in \mathbb{Z}[\varphi]$. The *critical exponent* and *asymptotic critical exponent* of W are

$$E(W) := \sup\Big\{\frac{|w|}{p} : w \in \mathcal{L}(W),\ p \ge 1 \text{ a period of } w\Big\},$$
$$E^*(W) := \lim_{L\to\infty} \sup\Big\{\frac{|w|}{p} : w \in \mathcal{L}(W),\ |w| \ge L,\ p \ge 1 \text{ a period of } w\Big\}.$$

Following [27], a *loopless digraph* $G = (V, E)$ defines an XYX-formula with one fragment $X_u X_v X_u$ per directed arc $(u, v) \in E$. It *occurs* in W if there is a non-erasing assignment $h\colon V \to \Sigma^+$ (not required to be injective) with $h(u)h(v)h(u) \in \mathcal{L}(W)$ for all $(u, v) \in E$; otherwise W *avoids* G. The *relation graph* $\mathrm{Rel}(b_5)$ on Σ_5^+ has undirected edges $\{u, v\}$ whenever $uvu \in \mathcal{L}(b_5)$ or $vuv \in \mathcal{L}(b_5)$.

3. PROOF OF THE MAIN RESULTS

The proofs have zero circular dependencies: Theorems 1(A), 2, 3, 4, and 5 are proved independently of Theorem 1(B), which then uses the unconditional square-freeness derived inside Theorem 4(A).

Conjugate-length scaling and 5-interval coding.

Lemma 6 (Non-vanishing and $-1/\varphi$ scaling of λ'). *For every $w \in \Sigma_5^*$:*

(1) $\lambda'(w) = (|w|_0 + |w|_2) + (|w|_2 - |w|_1 - |w|_3 - |w|_4)\varphi \in \mathbb{Z}[\varphi]$; *if $w \neq \varepsilon$, then $\lambda'(w) \neq 0$.*

(2) $\lambda'(f_5(w)) = (-1/\varphi)\,\lambda'(w)$.

Proof. Part (1) follows from (3) and $\mathbb{Q}$-linear independence of $\{1, \varphi\}$: if $\lambda'(w) = 0$, then $|w|_0 + |w|_2 = 0$ forces $|w|_0 = |w|_2 = 0$, and then $-(|w|_1 + |w|_3 + |w|_4) = 0$ forces $w = \varepsilon$. For (2), additivity reduces the identity to the five letters, using $-1/\varphi = 1 - \varphi$ and $\varphi^2 = \varphi + 1$: $\lambda'(01) = 1 - \varphi = (-1/\varphi)\,l'(0), \lambda'(23) = \lambda'(21) = 1 = (-1/\varphi)(-\varphi), \lambda'(4) = -\varphi = (-1/\varphi)(1+\varphi)$, and $\lambda'(0) = 1 = (-1/\varphi)(-\varphi)$. □

Lemma 7 (Eight affine sub-interval inclusions). *For every $a \in \Sigma_5$, every $0 \le o < |f_5(a)|$, and every $x \in P_a$,*

$$-\frac{x}{\varphi} + \lambda'(f_5(a)[0..o)) \in P_{f_5(a)[o]}.$$

Proof. Because $-1/\varphi = 1 - \varphi < 0$ reverses intervals, each of the eight inclusions reduces to two endpoint inequalities in $\mathbb{Q}(\varphi)$, listed explicitly in Table 2. Every containment follows directly from $1 < \varphi < 2$, $\varphi^2 = \varphi + 1$, and $\varphi^3 = 2\varphi + 1$. □

Proof of Theorem 1(A). The endpoints satisfy $-\varphi^2/2 < -1/2 < (\varphi-1)/2 < \varphi/2 < \varphi < \varphi^3/2$, so P_2, P_0, P_4, P_1, P_3 are pairwise disjoint open subsets of J. We prove $x_q \in P_{b_5[q]}$ by strong induction on $q \ge 0$. For $q = 0$, $x_0 = 0 \in (-1/2, (\varphi-1)/2) = P_0$ and $b_5[0] = 0$. For $q \ge 1$, write the unique block decomposition $q = s_p + o$ ($p < q$, $0 \le o < |f_5(b_5[p])|$). By induction $x_p \in P_{b_5[p]}$, so Lemmas 6(2) and 7 yield $x_q = -x_p/\varphi + \lambda'(f_5(b_5[p])[0..o)) \in P_{f_5(b_5[p])[o]} = P_{b_5[q]}$. Finally, $x_{p+1} = x_p + l'(b_5[p]) = T(x_p)$ matches the three translation branches of T. □

Figure 2 displays the partition P_2, P_0, P_4, P_1, P_3, the 3-IET T, the symmetry center $\varphi/4$, the circle-rotation lift R, and the first 17 orbit steps.

Cylinder convexity and bipartite XYX graph.

Lemma 8 (Convexity of word cylinders along the b_5 orbit). *If $u \in \Sigma_5^*$ occurs in b_5 at positions p_1 and p_2, and $q \in \mathbb{N}$ satisfies $\min(x_{p_1}, x_{p_2}) \le x_q \le \max(x_{p_1}, x_{p_2})$, then u occurs in b_5 at position q.*

Proof. Induct on $k < |u|$: assuming $b_5[q+j] = u[j]$ for $j < k$, we have $x_{q+k} = x_q + \lambda'(u[0..k))$ lying between $x_{p_1+k} = x_{p_1} + \lambda'(u[0..k))$ and $x_{p_2+k} = x_{p_2} + \lambda'(u[0..k))$, both of which belong to the interval $P_{u[k]}$ by Theorem 1(A). Hence $x_{q+k} \in P_{u[k]}$, forcing $b_5[q+k] = u[k]$. □

Proof of Theorem 1(B). Let $u, v \in \Sigma_5^+$ have uvu occurring at position p_1 of b_5, and set $p_2 := p_1 + |u| + |v|$, $D := \lambda'(u) \neq 0$, and $\Delta := \lambda'(uv) = D + \lambda'(v)$ (non-vanishing by Lemma 6(1)). If $|\Delta| \ge |D|$, then either $x_{p_1+|u|} = x_{p_1} + D$ or $x_{p_2+|u|} = x_{p_2} + D$ lies between x_{p_1} and $x_{p_2} = x_{p_1} + \Delta$; in

Table 1. First 17 steps of the conjugate-length orbit $x_p = \lambda'(b_5[0..p)) \in \mathbb{Z}[\varphi]$, reading $f_5^5(0) = 01234210423010421$ via Theorem 1(A). Every row is verified in exact integer arithmetic by `recompute.py`.

p	$x_p \in \mathbb{Z}[\varphi]$	Approx.	Interval	p	$x_p \in \mathbb{Z}[\varphi]$	Approx.	Interval
0	0	0.0000	P_0	9	$4-3\varphi$	−0.8541	P_2
1	1	1.0000	P_1	10	$5-2\varphi$	1.7639	P_3
2	$1-\varphi$	−0.6180	P_2	11	$5-3\varphi$	0.1459	P_0
3	2	2.0000	P_3	12	$6-3\varphi$	1.1459	P_1
4	$2-\varphi$	0.3820	P_4	13	$6-4\varphi$	−0.4721	P_0
5	$2-2\varphi$	−1.2361	P_2	14	$7-4\varphi$	0.5279	P_4
6	$3-\varphi$	1.3820	P_1	15	$7-5\varphi$	−1.0902	P_2
7	$3-2\varphi$	−0.2361	P_0	16	$8-4\varphi$	1.5279	P_1
8	$4-2\varphi$	0.7639	P_4				

Table 2. The eight exact affine sub-interval inclusions verifying Lemma 7 under $x \mapsto (-1/\varphi)x + \lambda'(f_5(a)[0..o))$.

a	o	$f_5(a)[o]$	Image of P_a under $(-1/\varphi)x + \lambda'(f_5(a)[0..o))$	Target interval
0	0	0	$((\varphi-2)/2,\ (\varphi-1)/2)$	$\subset (-1/2,\ (\varphi-1)/2) = P_0$
0	1	1	$(\varphi/2,\ (\varphi+1)/2)$	$\subset (\varphi/2,\ \varphi) = P_1$
1	0	2	$(-1,\ -1/2)$	$\subset (-\varphi^2/2,\ -1/2) = P_2$
1	1	3	$(\varphi,\ (2\varphi+1)/2)$	$= (\varphi,\ \varphi^3/2) = P_3$
2	0	4	$((\varphi-1)/2,\ \varphi/2)$	$= P_4$
3	0	2	$(-\varphi^2/2,\ -1)$	$\subset (-\varphi^2/2,\ -1/2) = P_2$
3	1	1	$(\varphi^2/2,\ \varphi)$	$\subset (\varphi/2,\ \varphi) = P_1$
4	0	0	$(-1/2,\ (\varphi-2)/2)$	$\subset (-1/2,\ (\varphi-1)/2) = P_0$

all four sign cases Lemma 8 forces u to occur at $p_1 + |u|$ or $p_2 + |u|$, producing the square uu in b_5, contradicting the unconditional square-freeness proved independently in Theorem 4(A). Hence $|\Delta| < |D|$, which forces $\lambda'(u)\,\lambda'(v) < 0$.

Therefore $\chi(u) := \mathbf{1}_{\lambda'(u)>0} \in \{0,1\}$ properly 2-colors $\mathrm{Rel}(b_5)$ (nonempty since 010 occurs at index 11, giving $0 \sim 1$). If a loopless digraph G with an undirected odd closed walk $\gamma\colon \mathbb{Z}/(2k+1)\mathbb{Z} \to V$ occurred in b_5 via $h\colon V \to \Sigma_5^+$, then $\chi(h(\gamma(j+1))) = 1-\chi(h(\gamma(j)))$ would give $\chi(h(\gamma(2k+1))) = 1-\chi(h(\gamma(0)))$, contradicting $\gamma(2k+1) = \gamma(0)$. Since C_3, TT_n $(n \geq 3)$, and K_n $(n \geq 3)$ all contain an undirected triangle, b_5 avoids them all. □

Circle-rotation lift, 3-IET collapse, and complexity. We prove Theorem 2 completely from first principles, without invoking external ergodic-theory or Sturmian classification results (cf. Keane's minimality theorem [16] [to check] and Morse and Hedlund [23] [to check]).

Enlarge J on the left by $(-\varphi^3/2, -\varphi^2/2)$ (length $\varphi/2$) to form the symmetric interval $J' := (-\varphi^3/2, \varphi^3/2)$ of length $\varphi^3 = \varphi + \varphi^2$. Define the circle rotation $R\colon J' \setminus \{-1/2\} \to J'$ and binary readout $\rho\colon J' \setminus \{-1/2\} \to \{0,1\}$ by

$$R(x) := \begin{cases} x+\varphi^2, & x < -\frac{1}{2}, \\ x-\varphi, & x > -\frac{1}{2}, \end{cases} \qquad \rho(x) := \begin{cases} 1, & x < -\frac{1}{2}, \\ 0, & x > -\frac{1}{2}. \end{cases}$$

Under $\eta(x) := (x + \varphi^3/2)/\varphi^3 \in (0,1)$, R conjugates to the irrational rotation $u \mapsto \{u + 1/\varphi\}$ (mod 1).

Lemma 9 (Dynamical coding and elementary Kronecker density).

(1) *Setting $g_p := |g_{\mathrm{fib}}(b_5[0..p))|$, we have $R^{g_p}(0) = x_p$ for all $p \geq 0$, $g_{\mathrm{fib}}(b_5)[k] = \rho(R^k(0))$ for all $k \geq 0$, and every $R^k(0) \in J$ equals x_p for some $p \geq 0$. Likewise, $\mathbf{f}[k] = \rho(R^k(1/2))$ for all $k \geq 0$.*

(2) *For every $y_0 \in J'$, the forward orbit $\{R^k(y_0) : k \in \mathbb{N}\}$ is dense in J'. Consequently, the prefix orbit $\{x_p : p \in \mathbb{N}\}$ is dense in J.*

Proof. For (1), R agrees with T on P_2 (reading $1 = g_{\mathrm{fib}}(2)$) and on $P_4 \cup P_1 \cup P_3$ (reading $0 = g_{\mathrm{fib}}(1) = g_{\mathrm{fib}}(3) = g_{\mathrm{fib}}(4)$), while on $P_0 = (-1/2, (\varphi-1)/2)$, R first maps x to $x - \varphi \in (-\varphi^3/2, -\varphi^2/2)$ (reading 0) and then adds φ^2 to land at $x + 1 = T(x)$ (reading 1, spelling $01 = g_{\mathrm{fib}}(0)$). Similarly, assigning weights $l_{\mathrm{fib}}(0) = -\varphi$ and $l_{\mathrm{fib}}(1) = \varphi^2$ satisfies $\lambda_{\mathrm{fib}}(\sigma_{\mathrm{fib}}(w)) = (-1/\varphi)\,\lambda_{\mathrm{fib}}(w)$, and strong induction on $\mathbf{f} = \sigma_{\mathrm{fib}}(\mathbf{f})$ shows $R^k(1/2) = 1/2 + \lambda_{\mathrm{fib}}(\mathbf{f}[0..k))$ codes $\mathbf{f}$.

For (2), given $(s,t) \subset (0,1)$ of width $\delta = t - s > 0$, choose an integer $M > 1/\delta$. By the pigeonhole principle on $\{j/\varphi\} \in [0,1)$ $(0 \leq j \leq M)$ and irrationality of $1/\varphi$, there exists $d \geq 1$ with step $\alpha = \{d/\varphi\} \in (0,\delta) \cup (1-\delta, 1)$, so $\{u_0 + md/\varphi\}$ enters (s,t) for some $m \in \mathbb{N}$. Conjugating by η^{-1} proves density of $\{R^k(y_0)\}$ in J', and part (1) restricts density of $\{R^k(0)\}$ on J to $\{x_p\}$. □

Lemma 10 (Cut-interval bijection and $\mathbb{Q}(\varphi)$ non-connection).

(1) *Let $F\colon I \to \mathbb{R}$ be piecewise translation on an open interval $I = (L,H)$ with base cut set $C \subset I$ and letter readout α that is constant on each component of $I \setminus C$ and jumps across every $c \in C$. If $S_n := \{x \in I : \exists k < n,\ F^k(x) \in C\}$ has cardinality K, and W is coded by a dense orbit of F in $I \setminus C$, then the $K+1$ open components of $I \setminus S_n$ read pairwise distinct length-n words forming $L_n(W)$, so $p_W(n) = K+1$.*

(2) *On J, let $C_5 := \{c_0, c_1, c_2, c_3\} = \{-1/2,\ (\varphi-1)/2,\ \varphi/2,\ \varphi\}$, and let $D_{\mathrm{inv}} := \{1/2,\ (1+\varphi)/2\}$ be the two discontinuities of T^{-1}. For every $k \in \mathbb{N}$ and $x \in \mathbb{R}$, $T^k(x) - x = (m_0 + m_2) + (m_2 - m_1)\varphi \in \mathbb{Z}[\varphi]$ with $m_0, m_1, m_2 \in \mathbb{N}$, $m_0 + m_1 + m_2 = k$. Consequently: (i) $T^k(x) = x \implies k = 0$; (ii) $T^k(d) \neq c_i$ for all $d \in D_{\mathrm{inv}}$, $c_i \in C_5$, $k \in \mathbb{N}$; and (iii) $T^k(c_j) = c_i \implies k = 0$ and $i = j$.*

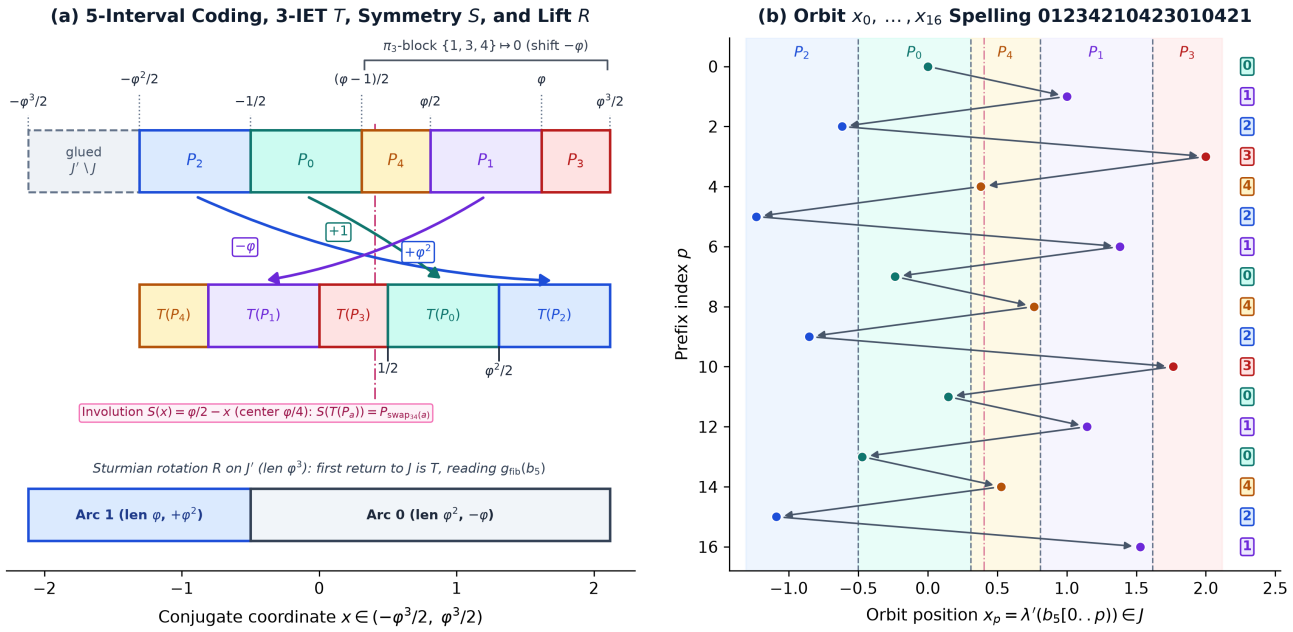


Figure 2. Geometric coding of b_5. (a) The five open letter intervals P_2, P_0, P_4, P_1, P_3 partitioning $J = (-\varphi^2/2, \varphi^3/2)$, their π_3-collapse merging $P_4 \cup P_1 \cup P_3$, the 3-IET T, the affine symmetry involution $S(x) = \varphi/2 - x$ centered at $\varphi/4$ (Theorem 5(A)), and the circle rotation R on $J' = (-\varphi^3/2, \varphi^3/2)$. (b) The first 17 orbit steps $x_0, \dots, x_{16}$ of 0 inside J, reading $f_5^5(0)$.

Proof. Part (1) is immediate: on each component I_r of $I \setminus S_n$, F^k $(k < n)$ never crosses C, so the length-n itinerary is constant; for $r_1 < r_2$, taking the smallest $k_0 < n$ at which $[F^{k_0}(x_1), F^{k_0}(x_2)]$ straddles some $c \in C$ separates the itineraries; density ensures every I_r is visited.

For (2), each step of T adds 1, $-\varphi$, or $1+\varphi$, giving $T^k(x) - x = (m_0 + m_2) + (m_2 - m_1)\varphi$. Part (i) follows by $\mathbb{Q}$-linear independence of $\{1, \varphi\}$ since $m_0 + m_2 = 0$ forces $m_0 = m_2 = 0$ and then $m_1 = 0$. For (ii), in $\mathbb{Q}(\varphi)$ coordinates $c_0 = (-1/2, 0)$, $c_1 = (-1/2, 1/2)$, $c_2 = (0, 1/2)$, $c_3 = (0, 1)$, while $d_0 = (1/2, 0)$ and $d_1 = (1/2, 1/2)$, so the rational coordinate of $c_i - d_r$ lies in $\{-1, -1/2\} < 0$, which cannot equal $m_0 + m_2 \geq 0$. For (iii), for all 12 ordered pairs $i \neq j$, $c_i - c_j$ has at least one half-integer coordinate $\pm 1/2$, so $c_i - c_j \notin \mathbb{Z}[\varphi]$. □

Proof of Theorem 2. **Part (A).** By Lemma 10(2), every backward iterate $T^{-k}(c_i)$ $(0 \leq k < n, c_i \in C_5)$ avoids D_{inv} and is pairwise distinct inside J, giving $|S_n| = 4n$. Since the readout of b_5 jumps across all four cuts (2|0, 0|4, 4|1, 1|3) and $\{x_p\}$ is dense in J (Lemma 9(2)), Lemma 10(1) yields $p_{b_5}(n) = 4n + 1$ for all $n \geq 0$.

Part (B). Direct letterwise check confirms $\sigma_{01}(\pi_3(a)) = g_{\text{fib}}(a)$ for all $a \in \Sigma_5$, so $g_{\text{fib}} = \sigma_{01} \circ \pi_3$ and $g_{\text{fib}}(b_5) = \sigma_{01}(\pi_3(b_5))$. Under π_3, the three intervals P_4, P_1, P_3 merge into a single letter 0, leaving only the two base 3-IET cuts $C_3 := \{-1/2, (\varphi - 1)/2\} \subset C_5$ where the ternary readout jumps (2|1 and 1|0). By Lemma 10(2), the backward preimages $S_n^{(3)} := \{T^{-k}(c) : 0 \leq k < n,\ c \in C_3\}$ have exact cardinality $2n$, so Lemma 10(1) gives $p_{\pi_3(b_5)}(n) = 2n + 1$ for every $n \geq 0$. Finally, applying Lemma 10(1) to R on J' with single base cut $C = \{-1/2\}$ gives $|S_n| = n$ (by irrationality of $1/\varphi$). Because both $g_{\text{fib}}(b_5)$ (via $R^k(0)$) and $\mathbf{f}$ (via $R^k(1/2)$) are coded by dense orbits of R (Lemma 9), both words share the exact same $n+1$ length-n factors: $L_n(g_{\text{fib}}(b_5)) = L_n(\mathbf{f})$ and $p_{g_{\text{fib}}(b_5)}(n) = p_{\mathbf{f}}(n) = n + 1$ for all $n \geq 0$. □

Figure 3 illustrates all four complexity tiers together with the base-vs.-refinement pullback cut structure on J.

Exact Rote–Fibonacci projection. The obstacle to proving (O2) directly from the 3-IET coding is that $g_{\text{rf}}(2) = \varepsilon$ erases the letter 2, breaking alignment between positions of b_5 and $g_{\text{rf}}(b_5)$. We eliminate the erasing behavior by grouping every 2 with its unique successor (1 or 3).

Lemma 11 (5-block non-erasing factorization and 9-pair conjugacy). *Define the prolongable morphism* $\tau\colon \Sigma_5 \to \Sigma_5^+$ *and block decoder* $\mathrm{dec}\colon \Sigma_5 \to \Sigma_5^+$ *by*

$$\tau\colon\ 0 \mapsto 01,\ 1 \mapsto 2,\ 2 \mapsto 34,\ 3 \mapsto 0,\ 4 \mapsto 32,$$
$$\mathrm{dec}\colon\ 0 \mapsto 0,\ 1 \mapsto 1,\ 2 \mapsto 23,\ 3 \mapsto 4,\ 4 \mapsto 21.$$

Let $\mathbf{t} := \tau^\omega(0)$. *Then:*

(1) $\mathrm{dec} \circ \tau = f_5 \circ \mathrm{dec}$ *on* Σ_5^*, *hence* $b_5 = \mathrm{dec}(\mathbf{t})$ *letter-for-letter, and* $\tilde{g} := g_{\text{rf}} \circ \mathrm{dec}$ *is 2-uniform:* $\tilde{g}(0) = 01$, $\tilde{g}(1) = 10$, $\tilde{g}(2) = 11$, $\tilde{g}(3) = 00$, $\tilde{g}(4) = 10$.

(2) *Every adjacent 2-factor of* $\mathbf{t}$ *belongs to the 9-element set*

$$P_\tau := \{01, 03, 10, 12, 20, 23, 32, 34, 40\}.$$

(3) *For each pair* $(x, y) \in \Sigma_5^2$, *define the 2-bit complemented difference block* $\phi(x, y) := [\tilde{g}(x)_0 == \tilde{g}(x)_1,\ \tilde{g}(x)_1 == \tilde{g}(y)_0] \in \{0, 1\}^2$, *and for a word* $u = u_0 \cdots u_r$ $(r \geq 1)$ *set* $\Phi_\Delta(u) := \phi(u_0, u_1) \cdots \phi(u_{r-1}, u_r)$. *Define the 1-bit overhangs* $L(1) := 0$ $(L(x) := \varepsilon$ *for* $x \neq 1)$ *and* $R(2) = R(4) := 1$ $(R(x) := \varepsilon$ *for* $x \notin \{2, 4\})$. *Then every pair* $(x, y) \in P_\tau$ *satisfies the exact word identity*

$$L(x) \cdot \Phi_\Delta(\tau(x)\,\tau(y)[0]) \cdot R(y) = R(x) \cdot \sigma_{\text{fib}}(\phi(x, y)) \cdot L(y). \tag{7}$$

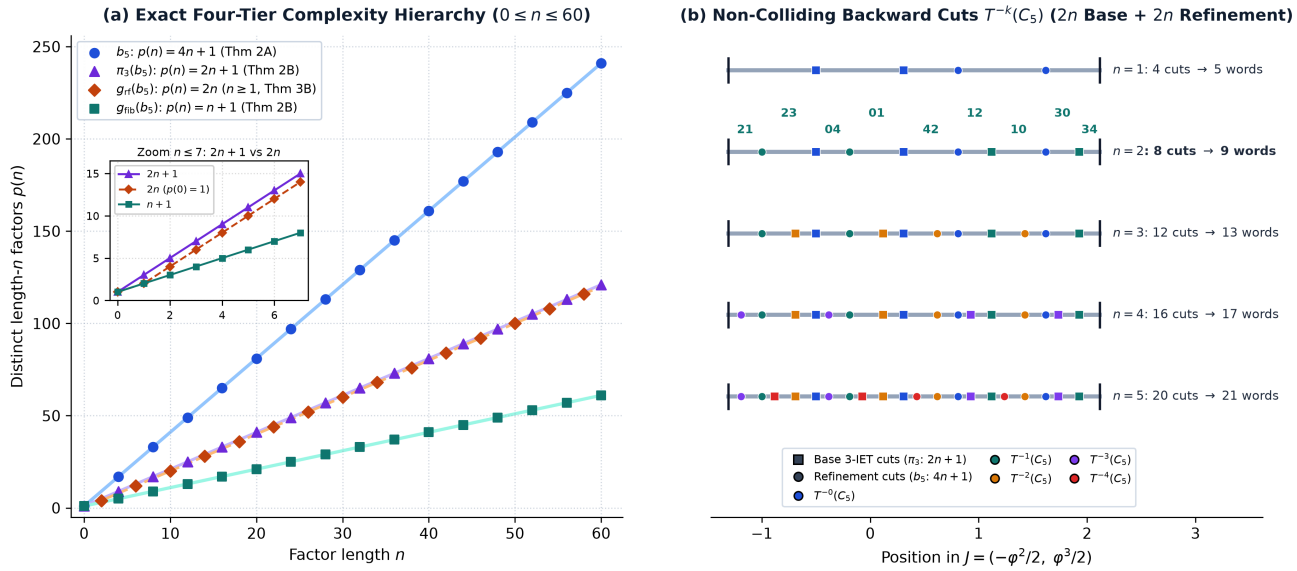


Figure 3. Exact four-tier factor complexity hierarchy and pullback cuts. (a) Exact complexity counts for $0 \le n \le 60$ across b_5 $(4n+1)$, $\pi_3(b_5)$ $(2n+1)$, $g_{\mathrm{rf}}(b_5)$ $(2n$ for $n \ge 1, p(0) = 1)$, and $g_{\mathrm{fib}}(b_5)$ $(n+1)$, with zoomed inset separating $2n+1$ from $2n$ on $n \le 7$. (b) The $4n$ pairwise distinct backward preimages $T^{-k}(C_5)$ in J for $n = 1, \ldots, 5$, distinguishing the $2n$ base 3-IET cuts (squares, driving $2n+1$) from the $2n$ refinement cuts (circles, lifting to $4n+1$), with the 9 two-letter factors labeled at $n = 2$.

Proof. Part (1) is immediate on all five letters: e.g., $\mathrm{dec}(\tau(2)) = \mathrm{dec}(34) = 421 = f_5(23) = f_5(\mathrm{dec}(2))$ and $g_{\mathrm{rf}}(23) = 11$, $g_{\mathrm{rf}}(21) = 10$. Hence $\mathrm{dec}(\tau^n(0)) = f_5^n(0)$ for all $n \ge 0$, giving $b_5 = \mathrm{dec}(\mathbf{t})$. For (2), $\tau(0) = 01 \in P_\tau$, and checking all internal and boundary pairs of $\tau(x)\tau(y)$ for $xy \in P_\tau$ confirms closure inside P_τ. For (3), all 9 identities (7) are verified directly in Table 3: in every row, $L(x)\,\Phi_\Delta(\tau(x)\tau(y)[0])\,R(y)$ equals $R(x)\,\sigma_{\mathrm{fib}}(\phi(x,y))\,L(y)$, completing the proof. □

Proof of Theorem 3. **Part (A).** First, let $D_k := (g_{\mathrm{rf}}(b_5)[k] == g_{\mathrm{rf}}(b_5)[k+1])$ be the complemented first-difference sequence of $g_{\mathrm{rf}}(b_5)$. Because $\tilde{g} = g_{\mathrm{rf}} \circ \mathrm{dec}$ is 2-uniform, for every $m \ge 0$ the length-$2m$ prefix $D[0..2m)$ equals $\Phi_\Delta(\mathbf{t}[0..m+1))$. Notice from Lemma 11(3) that for every $a \in \Sigma_5$, at least one of $L(a)$ and $R(a)$ is empty, and $L(0) = R(0) = \varepsilon$. Concatenating (7) along the valid chain $\mathbf{t}[0..m+1)$ cancels every interior overhang $R(y)$ against $R(y)$ and $L(y)$ against $L(y)$ (since $\mathbf{t}[0] = 0$ has empty overhangs), giving

$$D[0..2|\tau(\mathbf{t}[0..m))|) \cdot R(\mathbf{t}[m]) = \sigma_{\mathrm{fib}}(D[0..2m)) \cdot L(\mathbf{t}[m]).$$

Since $|L(\mathbf{t}[m])|, |R(\mathbf{t}[m])| \le 1$, removing at most one trailing bit shows that $\sigma_{\mathrm{fib}}(D[0..2m))$ agrees with D on its full length minus at most one bit. Because $D[0..2) = \phi(0,1) = 01$ starts with 0, iterating σ_{fib} proves $\sigma_{\mathrm{fib}}^r(0)$ is a prefix of D for every $r \ge 0$, hence $D_k = \mathbf{f}[k]$ for all $k \ge 0$.

Second, to verify $T_0(\mathbf{f}) = \mathbf{r}$, let $\tilde{\sigma}_{\mathrm{fib}}$ be the conjugate Fibonacci morphism $0 \mapsto 10$, $1 \mapsto 0$. Direct induction on words gives $0 \cdot \tilde{\sigma}_{\mathrm{fib}}(w) = \sigma_{\mathrm{fib}}(w) \cdot 0$ for all $w \in \{0,1\}^*$, and tracking the 2-state transducer T_0 on single letters shows that appending one bit c extends $T_0(w)$ by two identical bits $(c = 0)$ or one bit $(c = 1)$ opposite to the last bit of $T_0(w)$, so $\overline{\Delta}(T_0(w)) = \tilde{\sigma}_{\mathrm{fib}}(w)[1..]$ (dropping the leading 1). Taking $w = \mathbf{f}[0..m)$ and using $\sigma_{\mathrm{fib}}(\mathbf{f}) = \mathbf{f}$ gives $\overline{\Delta}(T_0(\mathbf{f})[0..L)) = \mathbf{f}[1..L)$, with $T_0(\mathbf{f})[0] = 0$. Since a binary word is uniquely determined by its initial bit and its complemented first differences, $T_0(\mathbf{f})[k] = \mathbf{r}[k]$ for all $k \ge 0$.

Third, $g_{\mathrm{rf}}(b_5)[0] = \tilde{g}(0)[0] = 0$, $g_{\mathrm{rf}}(b_5)[1] = \tilde{g}(0)[1] = 1 = 1 - \mathbf{r}[0]$, and for every $k \ge 0$, $(g_{\mathrm{rf}}(b_5)[k+1] == g_{\mathrm{rf}}(b_5)[k+2]) = D_{k+1} = \mathbf{f}[k+1] = (\mathbf{r}[k] == \mathbf{r}[k+1]) = ((1 - \mathbf{r}[k]) == (1 - \mathbf{r}[k+1]))$. Induction on k yields $g_{\mathrm{rf}}(b_5)[k+1] = 1 - \mathbf{r}[k]$ for all $k \ge 0$.

Part (B). For $n = 0$, $L_0(g_{\mathrm{rf}}(b_5)) = L_0(\mathbf{r}) = \{\varepsilon\}$ and $p(0) = 1$. Fix $n \ge 1$. For any $w \in \{0,1\}^n$, w is uniquely determined by its first bit $w[0] \in \{0,1\}$ and its length-$(n-1)$ complemented difference $\overline{\Delta}(w)$. By part (A), every $w \in L_n(\mathbf{r}) \cup L_n(g_{\mathrm{rf}}(b_5))$ has $\overline{\Delta}(w) \in L_{n-1}(\mathbf{f})$. Conversely, let $d \in L_{n-1}(\mathbf{f})$. For any $N \ge 0$, the 0-counts $|\sigma_{\mathrm{fib}}^{N+1}(0)|_0, |\sigma_{\mathrm{fib}}^{N+2}(0)|_0, |\sigma_{\mathrm{fib}}^{N+3}(0)|_0$ are three consecutive Fibonacci numbers $F_{N+2}, F_{N+3}, F_{N+4}$, which cannot all share the same parity. Because every factor d of $\mathbf{f}$ occurs inside $\sigma_{\mathrm{fib}}^N(0)$ for sufficiently large N, and $\sigma_{\mathrm{fib}}^N(0)$ occurs at positions $|\sigma_{\mathrm{fib}}^{N+1}(0)|, |\sigma_{\mathrm{fib}}^{N+2}(0)|$, and $|\sigma_{\mathrm{fib}}^{N+3}(0)|$ in $\mathbf{f}$, the factor d occurs inside $\mathbf{f}[1..)$ after prefixes of *both* even and odd 0-count. Since $\mathbf{r}[k] = |\mathbf{f}[1..k+1)|_0 \bmod 2$, d lifts to *both* binary words of length n (with initial bit 0 and initial bit 1) in $L_n(\mathbf{r})$, and by part (A) the exact same holds in $L_n(g_{\mathrm{rf}}(b_5))$. Therefore

$$L_n(g_{\mathrm{rf}}(b_5)) = L_n(\mathbf{r}) = \{w \in \{0,1\}^n : \overline{\Delta}(w) \in L_{n-1}(\mathbf{f})\},$$

which is immediately invariant under bitwise complement (since $\overline{\Delta}(\overline{w}) = \overline{\Delta}(w)$) and has cardinality $2\,p_{\mathbf{f}}(n-1) = 2((n-1)+1) = 2n$. □

Desubstitution, Cayley–Hamilton reduction, and sharp critical exponent. We now prove Theorem 4 uncondition-

Table 3. Exact verification of the 9-pair conjugacy identity (7) in Lemma 11(3) across all $xy \in P_\tau$.

$xy \in P_\tau$	$\phi(x,y)$	$L(x)$	$R(x)$	$L(y)$	$R(y)$	$\tau(x)\tau(y)[0]$	$\Phi_\Delta(\tau(x)\tau(y)[0])$	Common word
01	01	ε	ε	0	ε	012	0100	0100
03	00	ε	ε	ε	ε	010	0101	0101
10	01	0	ε	ε	ε	20	10	010
12	00	0	ε	ε	1	23	10	0101
20	10	ε	1	ε	ε	340	1001	1001
23	10	ε	1	ε	ε	340	1001	1001
32	10	ε	ε	ε	1	03	00	001
34	10	ε	ε	ε	1	03	00	001
40	01	ε	1	ε	ε	320	1010	1010

ally, strengthening the maximal-run desubstitution framework to certify the sharp second-order shifted deficit $E(p,m) = 2(\varphi p - 2m) - (4-\varphi) > 0$ in $\mathbb{Z}[\varphi]$ across all seed periods.

Elementary combinatorics, maximal runs, and desubstitution.

Lemma 12 (Elementary combinatorial properties of f_5 and b_5).

(1) **(*2-factors and primitivity.*)** $L_2(b_5) = \{01, 04, 10, 12, 21, 23, 30, 34, 42\}$. *Every letter of* Σ_5 *appears in* $f_5^5(a)$ *for every* $a \in \Sigma_5$, *so* f_5 *is primitive and* b_5 *is uniformly recurrent; moreover,* $|f_5^{10}(a)| \in \{72, 116, 117, 189\}$ *satisfies* $72 \le |f_5^{10}(a)| < 256$ *for all* $a \in \Sigma_5$.

(2) **(*Cut-pair recognizability.*)** *For* $t \ge 1$, *position* t *is a block boundary* ($t = s_p$ *for some* $p \ge 1$) *if and only if* $b_5[t-1]b_5[t] \notin \{01, 21, 23\}$. *No factor of length* 3 *lies in* $\{2,4\}^*$, *and no factor of length* $p+3$ *has period* $p \le 2$.

(3) **(*Algebraic overlap bound.*)** *For every* $u \in \Sigma_5^+$, $|\lambda'(u)| \cdot (2|u|) \ge 1$. *Consequently, every* p*-periodic factor of* b_5 ($p \ge 1$) *has overlap* $m \le 10p^3$.

Proof. Part (1) is a finite closure check starting from 01 and direct computation of $f_5^5(a)$ and $f_5^{10}(a)$. For (2), the internal pairs of length-2 blocks $f_5(a)$ are precisely $\{01, 23, 21\}$, while every block $f_5(a)$ ends in $\{0,1,3,4\}$ and starts in $\{0,2,4\}$, so no boundary pair can belong to $\{01, 21, 23\}$; uniqueness of the five images gives unique recognizability. Inspecting $L_2(b_5)$ excludes $00, 11, 22, 24, 33, 44$, and scanning $f_5^{10}(a)f_5^{10}(b)$ rules out $p = 2, m \ge 3$.

For (3), writing $\lambda'(u) = a + b\varphi \ne 0$ ($|a|, |b| \le |u|$), the algebraic norm $|a^2 + ab - b^2| = |a + b\varphi| \cdot |a + b(1-\varphi)| \ge 1$ and $|a + b(1-\varphi)| \le 2|u|$ yield $|\lambda'(u)| \cdot (2|u|) \ge 1$. If $b_5[q..q+p+m)$ has period p with $m > 10p^3$, setting $K := 10p^2$ and $u := b_5[q..q+p)$ gives $K+1$ consecutive copies of u, so $(K+1)|\lambda'(u)| = |x_{q+(K+1)p} - x_q| < (\varphi^3 + \varphi^2)/2 < 5$ by Theorem 1(A); multiplying by $2p$ yields $K + 1 < 10p \le K$, a contradiction. □

A *concrete run* in b_5 is a triple (i,p,m) with $i \ge 10$, $p \ge 1$, $m \ge 0$ such that $b_5[i+j] = b_5[i+j+p]$ for $0 \le j < m$, with left and right maximality $b_5[i-1] \ne b_5[i+p-1]$ and $b_5[i+p+m] \ne b_5[i+m]$. Its *context state* is $\mathrm{st}(i,p,m) := (X,Y,S,T) := (b_5[i-3..i), b_5[i+p-3..i+p), b_5[i+p+m..i+p+m+3), b_5[i+m..i+m+3)) \in W_3^4$, where $W_3 := \{abc \in \Sigma_5^3 : ab, bc \in L_2(b_5)\}$ ($|W_3| = 16$), and its *summary* is $\Sigma(i,p,m) := (\mathrm{st}(i,p,m), \pi(b_5[i..i+p)), \pi(b_5[i+p..i+p+m)))$ with Parikh vectors $\pi(w) \in \mathbb{Z}^5$. For distinct $c,d \in \Sigma_5$, set $e_L(c,d) := 1$ if $\{c,d\} = \{0,3\}$ (else ε); for distinct a,b, set $e_R(a,b) := 0$ if $\{a,b\} = \{0,4\}$, 2 if $\{a,b\} = \{1,3\}$ (else ε). Checking all pairs in W_3^2 confirms that $\lambda := |e_L(X[2], Y[2])| \le 1$ is the longest common suffix of $f_5(X)$ and $f_5(Y)$ and $\rho := |e_R(S[0], T[0])| \le 1$ is the longest common prefix of $f_5(S)$ and $f_5(T)$.

Proposition 13 (Forward run map Φ and desubstitution).

(1) *For every concrete run* (i,p,m) *with context* (X,Y,S,T), *setting* $i' := s_i - \lambda$, $p' := s_{i+p} - s_i$, *and* $m' := \lambda + (s_{i+p+m} - s_{i+p}) + \rho$ *gives a concrete run* $\Phi(i,p,m) = (i',p',m')$ *whose summary* $\widehat{\Phi}(\Sigma(i,p,m))$ *depends only on* $\Sigma(i,p,m)$: *the context evolves purely from* (X,Y,S,T), *while the Parikh vectors obey*

$$\pi(u') = M\,\pi(u),$$
$$\pi(z') = M\,\pi(z) + \pi(e_L(X[2], Y[2])\, e_R(S[0], T[0])). \tag{8}$$

(2) *Every concrete run* (i,p,m) *with* $m \ge 3$ *has a preimage concrete run* (i_1,p_1,m_1) *with* $1 \le p_1 < p$, $m_1 \ge 1$, *and* $\widehat{\Phi}(\Sigma(i_1,p_1,m_1)) = \Sigma(i,p,m)$. *By strong induction on* p, *every concrete run with* $m \ge 1$ *satisfies* $\Sigma(i,p,m) = \widehat{\Phi}^k(\Sigma(i_0,p_0,m_0))$ *for some* $k \ge 0$ *and a* seed *concrete run* (i_0,p_0,m_0) *with* $m_0 \in \{1,2\}$.

Proof. Part (1) follows directly from comparing the images of X,Y,S,T at the ends of the block $f_5(b_5[i..i+p+m))$; shifting the period window left by $\lambda = 1$ preserves $\pi(u')$ because the entering and exiting letters coincide. For (2), by Lemma 12(1)–(2), $p \ge 3$ and the window $b_5[i-3..i+p+m+3)$ recurs at some $q_0 \ge 2(10p^3 + 30)$, yielding a concrete run $(\tilde{\imath},p,m)$ with identical summary. Because no two consecutive positions are interior to blocks, at least one of $\tilde{\imath}+1$ and $\tilde{\imath}+2$ (both $\le \tilde{\imath}+m-1$) is a block boundary $t_1 = s_{a'}$, forcing $t_1 + p = s_{a'}$ with $p_1 := a' - a \ge 1$ by cut-pair recognizability. Unique recognizability pulls the p-periodic stretch back to a p_1-periodic stretch in b_5, whose maximal extension (i_1,p_1,m_1) is finite (Lemma 12(3)), has $i_1 \ge 10$ by choice of q_0, and satisfies $\Phi(i_1,p_1,m_1) = (\tilde{\imath},p,m)$ by maximality. Finally, $p_1 = p$ would place a length-$p \ge 3$ factor in $\{2,4\}^*$ (contradicting Lemma 12(2)), so $p_1 < p$, and $m \ge 3$ with $\lambda, \rho \le 1$ forces $m_1 \ge 1$. □

Sharp large-period bound ($p_0 \ge 41$) ***and shifted Cayley–Hamilton reduction.***

Lemma 14 (Sharp second-order bound for large-period seeds, $p_0 \ge 41$). *Define* $\nu(v) := v_0 + (v_1 + v_3 + v_4)/\varphi + v_2/\varphi^2$ *on* $\mathbb{Z}^5$. *Then* $\nu(Mv) = \varphi\,\nu(v)$ *and* $\nu(v) \le \sum_c v_c \le \varphi^2 \nu(v)$ *for all* $v \in \mathbb{N}^5$. *Consequently, for every seed* (i_0,p_0,m_0) *with* $p_0 \ge 41$

and $m_0 \le 2$, every forward iterate (p_k, m_k) of $\widehat{\Phi}^k(\Sigma(i_0, p_0, m_0))$ satisfies the sharp bound

$$D_k := \varphi p_k - 2m_k > \frac{7-\sqrt{5}}{4} = 2 - \frac{\varphi}{2}.$$

Proof. Direct check on the columns of M gives $\nu(Mv) = \varphi\,\nu(v)$, and every extension satisfies $\nu(\pi(e_L e_R)) \le 1 + 1/\varphi = \varphi$. For $k = 0$, $p_0 \ge 41$ and $m_0 \le 2$ give $D_0 \ge 41\varphi - 4 > 2 - \varphi/2$. For $k \ge 1$, induction gives $\nu(\pi(z_k)) + \varphi^2 \le \varphi^k(2 + \varphi^2)$, so $2m_k \le 2\varphi^2\nu(\pi(z_k)) < 2\varphi^2(\varphi^k(2+\varphi^2) - \varphi^2)$, while $\varphi p_k \ge \varphi\,\nu(\pi(u_k)) = \varphi^{k+1}\nu(\pi(u_0)) \ge \varphi^{k-1}p_0 \ge 41\varphi^{k-1}$. Subtracting yields

$$\begin{aligned}\varphi p_k - 2m_k &> \varphi^{k-1}(41 - 2\varphi^3(2+\varphi^2)) \\ &= \varphi^{k-1}(41 - (10 + 18\varphi)) = \varphi^{k-1}(31 - 18\varphi).\end{aligned}$$

Since $10 + 18\varphi = 19 + 9\sqrt{5} < 41$ (because $(9\sqrt{5})^2 = 405 < 484 = 22^2$), $31 - 18\varphi > 0$, so for $k \ge 1$ we have $\varphi^{k-1} \ge 1$ and $D_k > 31 - 18\varphi$. Finally, $31 - 18\varphi > (7-\sqrt{5})/4 \iff 81 > 35\sqrt{5}$, which holds by the exact integer square inequality $81^2 = 6{,}561 > 6{,}125 = 5 \cdot 35^2$. □

Lemma 15 (Cayley–Hamilton identity for $B = M^6$ and constant-shift invariance). *The incidence matrix M of f_5 and $B := M^6$ are*

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

$$B = M^6 = \begin{pmatrix} 7 & 4 & 2 & 4 & 4 \\ 6 & 5 & 2 & 4 & 4 \\ 6 & 4 & 3 & 4 & 4 \\ 4 & 2 & 2 & 3 & 2 \\ 4 & 2 & 2 & 2 & 3 \end{pmatrix}.$$

The matrix B satisfies $(B^2 - 18B + I)(B - I) = B^3 - 19B^2 + 19B - I = 0$. Consequently, every affine orbit $v_{j+1} = Bv_j + C$, and every constant-shifted linear functional $E_j := \alpha(v_j) + \kappa \in \mathbb{Z}[\varphi]$, obeys the 4th-order recurrence

$$E_{j+4} - 20E_{j+3} + 38E_{j+2} - 20E_{j+1} + E_j = 0 \quad (j \ge 0). \tag{9}$$

Moreover, setting $\Lambda := \varphi^6 = 5 + 8\varphi$, $A_j := E_{j+2} - 2E_{j+1} + E_j$, and $R_j := \Lambda(E_{j+2} - E_{j+1}) - (E_{j+1} - E_j)$:

(1) ***(Growing branch.)*** *If $E_0 > 0$, $E_1 - E_0 > 0$, $A_0 > 0$, and $A_1 - A_0 > 0$, then $E_j > 0$ for all $j \ge 0$.*

(2) ***(Bounded or decaying branch.)*** *If $R_0 = R_1 = 0$, $E_0 > 0$, and $K := \Lambda E_1 - E_0 \ge 0$, then $\Lambda E_{j+1} - E_j = K$ and $E_j > 0$ for all $j \ge 0$. In particular, when $K = 0$, $E_j = \Lambda^{-j}E_0 = \varphi^{-6j}E_0 > 0$ for all $j \ge 0$.*

Proof. Direct 5×5 integer multiplication verifies $B^3 - 19B^2 + 19B - I = 0$ (reflecting that M's eigenvalues $-1, e^{\pm i\pi/3}$ have 6th power 1 while $\varphi^{\pm 6}$ satisfy $\lambda^2 - 18\lambda + 1 = 0$). Multiplying by $(B - I)$ annihilates C, yielding (9), and because $1 - 20 + 38 - 20 + 1 = 0$, adding any constant shift $\kappa \in \mathbb{Z}[\varphi]$ (in particular $\kappa = -(4 - \varphi)$) preserves (9) identically. For (1), (9) rewrites as $A_{j+2} = 18A_{j+1} - A_j$, so induction gives $A_{j+1} > A_j > 0$ and $E_{j+1} > E_j > 0$. For (2), $\Lambda^2 - 18\Lambda + 1 = 0$ implies $R_{j+2} = (\Lambda + 1)R_{j+1} - \Lambda R_j$, so $R_0 = R_1 = 0$ forces $R_j = 0$ for all $j \ge 0$, telescoping to $\Lambda E_{j+1} = E_j + K$. Since $\Lambda > 0$, $E_0 > 0$, and $K \ge 0$, induction gives $E_j > 0$ for all $j \ge 0$ (and $E_j = \Lambda^{-j}E_0$ when $K = 0$). □

Proof of Theorem 4. **Part (A).** Let $w \in \mathcal{L}(b_5)$ have period $p \ge 1$. If $|w| \le p$, then $\varphi p - 2(|w| - p) \ge \varphi p \ge \varphi > (7 - \sqrt{5})/4$. Otherwise $m_0 := |w| - p \ge 1$. By uniform recurrence and Lemma 12(3), w embeds inside a concrete run (i, p, r) with $r \ge m_0 \ge 1$, and Proposition 13(2) gives $\Sigma(i, p, r) = \widehat{\Phi}^k(\Sigma(i_0, p_0, m_0'))$ for some seed with $m_0' \in \{1, 2\}$. If $p_0 \ge 41$, Lemma 14 yields $\varphi p - 2r > (7 - \sqrt{5})/4$. If $1 \le p_0 \le 40$, the seed spans at most $p_0 + m_0' + 6 \le 48 < 72 \le |f_5^{10}(a)|$ letters, so it occurs in $f_5^{10}(a)f_5^{10}(b)$ for one of the 9 two-factors $ab \in L_2(b_5)$. Exhaustively scanning all offsets yields 338 seeds collapsing to 336 distinct summaries. For every one of the 336 summaries, 26 exact steps confirm $\mathrm{st}(s_8) = \mathrm{st}(s_2)$ (hence 6-periodicity for $k \ge 2$), $E_0 > 0$, $E_1 > 0$, and across all $336 \times 6 = 2{,}016$ residue branches $k = 2 + r + 6j$ $(0 \le r \le 5)$, exact $\mathbb{Z}[\varphi]$ evaluation certifies Lemma 15(1) or (2): specifically, 1,938 growing branches, 70 bounded branches with $K > 0$, and 8 extremal branches with $K = 0$ (every one having $E_0 = (-144, 89) = \varphi^{-11} > 0$ and $E_1 = (-2584, 1597) = \varphi^{-17} > 0$). Thus $E(p, r) > 0$, so $\varphi p - 2m_0 \ge \varphi p - 2r > (7 - \sqrt{5})/4$, proving (4). Non-attainment follows immediately from strict inequality, and since $(5 + \sqrt{5})/4 < 2$, b_5 is square-free.

Part (B). Consider the concrete seed `123 · 010 · 421` ($p_0 = 2, m_0 = 1$, run starting at index $i_0 = 38$ of b_5). Along $k = 6 + 6j$ ($r = 4$), exact computation gives $E_0 = (-144, 89)$ ($p_6 = 44, m_6 = 35$), $E_1 = (-2584, 1597)$ ($p_{12} = 798, m_{12} = 645$), $R_0 = R_1 = 0$, and $K = (5 + 8\varphi)(-2584 + 1597\varphi) - (-144 + 89\varphi) = 0$. Since $\varphi^{11} = 55 + 89\varphi$ and $(-144 + 89\varphi)(55 + 89\varphi) = (-144 \cdot 55 + 89^2) + 89(55 + 89 - 144)\varphi = 1$, we have $E_0 = \varphi^{-11}$, and Lemma 15(2) yields $E_j = \Lambda^{-j}E_0 = \varphi^{-(11+6j)} = -F_{12+6j} + F_{11+6j}\varphi > 0$ for all $j \ge 0$. Because $p_{6+6j} \ge \nu(\pi(u_{6+6j})) = \varphi^{6+6j}\nu(\pi(01)) \ge \varphi^{6+6j} \ge j + 1 \to \infty$ while $D_{6+6j} - (7 - \sqrt{5})/4 = \frac{1}{2}\varphi^{-(11+6j)} \to 0^+$, the constant $(7 - \sqrt{5})/4$ is optimal and unattained, and $|w_{6+6j}|/p_{6+6j} \to (5 + \sqrt{5})/4$ with $|w_{6+6j}| \to \infty$, proving $E(b_5) = E^*(b_5) = (5 + \sqrt{5})/4$. Finally, since $\sqrt{5} \notin \mathbb{Q}$, $(5 + \sqrt{5})/4 \notin \mathbb{Q}$. □

Symmetry, unbounded balance defect, F_5 avoidance, and A005252.

Proof of Theorem 5. **Part (A).** Define the affine involution $S(x) := \varphi/2 - x$ on J. Direct endpoint arithmetic in $\mathbb{Q}(\varphi)$ verifies $S(T(P_a)) = P_{\mathrm{swap}_{34}(a)}$ for all five letters:

$$\begin{aligned} S(T(P_2)) &= S(\tfrac{1+\varphi}{2}, \tfrac{1+2\varphi}{2}) = (-\tfrac{1+\varphi}{2}, -\tfrac{1}{2}) = P_2, \\ S(T(P_0)) &= S(\tfrac{1}{2}, \tfrac{1+\varphi}{2}) = (-\tfrac{1}{2}, \tfrac{\varphi-1}{2}) = P_0, \\ S(T(P_4)) &= S(-\tfrac{1+\varphi}{2}, -\tfrac{\varphi}{2}) = (\varphi, \tfrac{1+2\varphi}{2}) = P_3, \\ S(T(P_1)) &= S(-\tfrac{\varphi}{2}, 0) = (\tfrac{\varphi}{2}, \varphi) = P_1, \\ S(T(P_3)) &= S(0, \tfrac{1}{2}) = (\tfrac{\varphi-1}{2}, \tfrac{\varphi}{2}) = P_4. \end{aligned}$$

Since $l'(\mathrm{swap}_{34}(a)) = l'(a)$ for all $a \in \Sigma_5$ (as $l'(3) = l'(4) = -\varphi$), every $x \in P_a$ satisfies $T(S(T(x))) =$

$S(T(x)) + l'(\mathrm{swap}_{34}(a)) = \varphi/2 - (x + l'(a)) + l'(a) = S(x)$. Now let $w = w_0 \cdots w_{n-1}$ occur in b_5 at position p. Setting $y_0 := S(x_{p+n}) \in J$, backward induction gives $T^j(y_0) = S(x_{p+n-j}) = S(T(x_{p+n-1-j})) \in P_{\mathrm{swap}_{34}(w_{n-1-j})}$ for all $0 \le j < n$. Because each interval P_a is open, some open ball $B(y_0, \delta) \subset J$ shares the entire length-n itinerary $\mathrm{revSwap}_{34}(w)$. By orbit density (Lemma 9(2)), some $x_q \in B(y_0, \delta)$, so $\mathrm{revSwap}_{34}(w)$ occurs in b_5 at q. Because $\mathrm{revSwap}_{34}$ is an involution, $w \in \mathcal{L}(b_5) \iff \mathrm{revSwap}_{34}(w) \in \mathcal{L}(b_5)$.

Next, every factor of length ≤ 4 in b_5 occurs inside $f_5^{10}(a)f_5^{10}(b)$ for some $ab \in L_2(b_5)$ (since $|f_5^{10}(a)| \ge 72 \ge 4$). Direct finite verification over all 9 pairs confirms that none of the 16 words in F_5 occurs, and direct check confirms F_5 is $\mathrm{revSwap}_{34}$-invariant. Finally, direct 5×5 integer multiplication verifies the Cayley–Hamilton identity $M^5 + M^2 = M^4 + M^3 + M + I$; applying both sides to $M^k\pi(w)$ and summing coordinates yields (6).

Part (B). Define the integer linear functional $\psi(w) := |w|_0 - 2|w|_1 + |w|_2 + |w|_3 - |w|_4 \in \mathbb{Z}$. Direct check on each letter shows $\psi(f_5(a)) = -\psi(a)$ (for example, $\psi(f_5(0)) = \psi(01) = 1 - 2 = -1 = -\psi(0)$ and $\psi(f_5(1)) = \psi(23) = 1 + 1 = 2 = -\psi(1)$), so $\psi(f_5(w)) = -\psi(w)$ and $\psi(f_5^{2m}(w)) = \psi(w)$ for all $m \ge 0$ and $w \in \Sigma_5^*$; in particular, $\psi(f_5^{2m}(0)) = \psi(0) = 1$.

Now compute $f_5^4(0) = 0123421 \cdot 0 \cdot 423$, noting that $\psi(423) = -1 + 1 + 1 = +1$. Define a sequence of factors recursively by

$$w_0 := \varepsilon, \qquad w_{k+1} := 423 \cdot f_5^4(w_k) \qquad (k \ge 0).$$

We claim $0w_k \in \mathcal{L}(b_5)$ and $\psi(w_k) = k$ for every $k \ge 0$. For $k = 0$, $0 \in \mathcal{L}(b_5)$ and $\psi(\varepsilon) = 0$. Inductively, if $0w_k \in \mathcal{L}(b_5)$, then $f_5^4(0w_k) = 0123421 \cdot 0 \cdot 423 \cdot f_5^4(w_k) = 0123421 \cdot (0w_{k+1}) \in \mathcal{L}(b_5)$, and $\psi(w_{k+1}) = \psi(423) + \psi(f_5^4(w_k)) = 1 + \psi(w_k) = k + 1$.

Suppose for contradiction that b_5 were C-balanced for some $C \in \mathbb{N}$. Set $k := 6C + 1$ and $L := |w_k| \ge 1$ (since $k \ge 1$). For every length-L factor $v \in L_L(b_5)$, C-balance implies $|\psi(w_k) - \psi(v)| \le \sum_{a \in \Sigma_5} |\psi(a)| \cdot ||w_k|_a - |v|_a| \le (1 + 2 + 1 + 1 + 1)C = 6C$, so $\psi(v) \ge \psi(w_k) - 6C = 1$. Now choose m large enough that $N := |f_5^{2m}(0)| \ge (2L + 2)L$, and partition $f_5^{2m}(0)$ into $q := \lfloor N/L \rfloor \ge 2L + 2$ consecutive blocks of length L plus a remainder r of length $< L$. Since every block satisfies $\psi \ge 1$ and $|\psi(r)| \le 2|r| < 2L$, we obtain

$$1 = \psi(f_5^{2m}(0)) \ge q - 2L \ge 2,$$

a contradiction. Thus b_5 is not C-balanced for any $C \in \mathbb{N}$. □

4. COMPUTATIONAL VERIFICATION AND EMPIRICAL PICTURE

Although all four headline problems and ten theorem parts are proved unconditionally inside the Lean 4 kernel (Section 5), we also provide an independent, standalone, zero-float standard-library Python verifier (`recompute.py`) and audited tables illuminating the exact quantitative behavior of b_5.

Standalone zero-float verification and negative mutations. The script `papers/badkobeh-ochem-b5/draft3/recompute.py` runs in under 5 seconds using only the Python 3 standard library and begins with a self-AST check verifying 0 floating-point or complex literals, 0 `ast.Div` (`/`) operators, and 0 negative exponents. It independently verifies:

1. Exact integer execution of the Section 1 Python check (`OK 44 [9, 5, 4, 3] 0110110010`).
2. Exact $\mathbb{Q}(\varphi)$ verification of all 8 sub-interval inclusions, exact 5-interval coding on $b_5[0..10{,}000)$ (0 violations), the 4 XYX-related pairs among the 14 factors of length ≤ 2 ($0 \sim 1$, $2 \sim 34$, $01 \sim 23$, $10 \sim 42$), and all 1,452 fragments uvu in $f_5^8(0)$ (length 72; 0 violations).
3. Exact $\mathbb{Q}(\varphi)$ non-connection checks, the $n = 2$ pullback partition (8 cuts, 9 words), the 5-block factorization $\mathrm{dec}(\tau^k(0)) = f_5^k(0)$ ($k = 0..10$), all 9 τ-pair conjugacy identities, transducer equivalence $T_0(\mathbf{f}) = \mathbf{r}$ (6,472 bits), pointwise $g_{\mathrm{rf}}(b_5) = 0\bar{\mathbf{r}}$ (5,168 bits), and complete 4-tier complexity and language equalities ($4n + 1$, $2n + 1$, $2n$, $n + 1$; complement invariance) for $0 \le n \le 40$.
4. Exact 5×5 matrix verification of M^6, Cayley–Hamilton $B^3 - 19B^2 + 19B - I = 0$, the large-period bounds $405 < 484$ and $81^2 = 6{,}561 > 6{,}125 = 5 \cdot 35^2$, exhaustive enumeration of all 338 seeds (336 summaries, 13 extremal), pure-$\mathbb{Z}[\varphi]$ sharp positivity $E_k > 0$ across all 2,016 branches (1,938 growing, 70 bounded $K > 0$, 8 extremal $K = 0$), exact Fibonacci closed form $E_j = (-F_{12+6j}, F_{11+6j}) = \varphi^{-(11+6j)}$ ($j = 0..5$), and direct parsing of Tables 1, 4, and 5.
5. Exact $\mathbb{Q}(\varphi)$ involution $S(T(P_a)) = P_{\mathrm{swap}_{34}(a)}$, $\mathrm{revSwap}_{34}$ invariance ($0 \le n \le 30$), (-1)-eigenvector ψ and recursive factors $w_0, \ldots, w_5$ (lengths $0, 3, 20, 136, 935, 6{,}408$ with $\psi(w_k) = k$), avoidance of all 16 words in F_5, and the 5th-order `A005252` recurrence ($k = 0..20$).
6. Eight non-tautological negative mutations (8/8 caught): perturbing $f_5(3) \mapsto 20$ (eigenvector failure), swapping $f_5(3) \mapsto 12$ ($p(2) = 12 \ne 9$), shifting $c_0 = -1/2$ by $+1/10$ (inclusion failure at $a = 2$), setting $g_{\mathrm{fib}}(2) \mapsto 0$ (L_3 mismatch), bracketing the infimum margin at 1.190 (0 failures) vs. 1.191 (8 failures), corrupting $B_{0,0}$ by $+1$ (Cayley–Hamilton residual 39), swapping $g_{\mathrm{rf}}(3) \leftrightarrow g_{\mathrm{rf}}(4)$ (breaks $g_{\mathrm{rf}}(b_5) = 0\bar{\mathbf{r}}$), and tightening the sharp constant by $+1$ ($E_0 = (-145, 89) < 0$).

Extremal family `123·010·421`, Fibonacci decay, and empirical checks. Along the 6 residue classes $k = 2 + r + 6j$ ($0 \le r \le 5$) of `123·010·421`, $D_j^{(r)}$ converges exponentially to the six exact limits $c_r = K^{(r)}/(\Lambda - 1) \in \mathbb{Q}(\varphi)$:

$$c_0 = \frac{5 + \varphi}{2} \approx 3.3090, \qquad c_1 = 1 + \frac{\varphi}{2} \approx 1.8090,$$
$$c_2 = \frac{3}{2} = 1.5000, \qquad c_3 = \frac{5 - \varphi}{2} \approx 1.6910,$$
$$c_4 = 2 - \frac{\varphi}{2} \approx 1.1910, \qquad c_5 = \frac{5}{2} = 2.5000.$$

Only $r = 4$ ($k = 6 + 6j$) reaches the infimum $2 - \varphi/2 = (7 - \sqrt{5})/4 \approx 1.190983$, decaying as $\frac{1}{2}\varphi^{-(11+6j)}$ while the other five classes plateau at the strictly positive constants $c_r - (2 - \varphi/2)$ equal to 0.3090, 0.5000, 0.6180, 1.3090, and

2.1180. Tables 4 and 5 report exact iterates and the Fibonacci closed form through $k = 36$ ($p_{36} = 82{,}790{,}070$), and Figure 4 plots both the sharp second-order upper envelope and the exact exponential decay against the five plateaus.

Archived auxiliary consistency checks in the repository also confirmed $p_{b_5}(n) = 4n + 1$ through $n \le 1{,}000{,}000$ by both suffix-automaton and suffix-array programs, 0 violations of Theorem 1(B) across 433,000,439 fragments uvu of length $\le 10{,}000$, and 0 repetitions reaching $(5+\sqrt{5})/4$ across all periods $p \le 1{,}500{,}000$ in the first 8,388,608 letters of b_5 (extended to 16,777,216 letters in review; largest observed exponent $1{,}970{,}299/1{,}089{,}155 \approx 1.809016164$). These computations serve strictly as empirical confirmation; zero theorems depend on them.

5. FORMAL VERIFICATION IN LEAN 4

Every one of our four headline questions and observations (all ten theorem parts) is formally verified in Lean 4 (`v4.35.0` with Mathlib `v4.35.0`, with `autoImplicit` and `relaxedAutoImplicit` disabled) inside the library `lean/BadkobehOchemB5/`, pinned at commit

`23c633b4bc5c555b19b4df3747c00e5ac2909b6f.`

The v2 extension is purely additive over the blind-audited baseline (353 untouched baseline declarations in 5 modules plus 161 additional declarations in 3 additional modules).

Headline coverage. Headline questions and observations: 4 of 4 in Lean (100%); theorem parts: 10 of 10 in Lean (100%) · 0 by certificate only · 0 by written proof only. Table 6 maps each main result to its exact Lean 4 declarations and statement-fidelity status, and Table 7 summarizes all 10 source files (11,480 lines, 514 theorems and lemmas).

Exact trust boundary. Running `lake env lean BadkobehOchemB5/Verify.lean` audits 100% (514/514) of all public `theorem` and `lemma` declarations (0 private theorems) with 0 warnings and deterministic output SHA-256

`268180c3bf8b703361ef2b2eec4b051c`

`7c5944e00ab75a7307354834a26438fe.`

The union of axioms across all 514 declarations is strictly contained in Lean's standard three axioms {`propext`, `Classical.choice`, `Quot.sound`} (36 axiom-free, 103 `[propext]`, 26 `[propext, Quot.sound]`, 349 all three), with zero occurrences of `Lean.ofReduceBool` or `Lean.trustCompiler`. No external tables or literature theorems are axiomatized:

- `g10Val_eq_iterSym5` proves f_5^{10} in-kernel;
- `all_twoFactor_seeds_in_buckets` proves exhaustiveness of all 336 seed summaries across the 9 two-factor blocks;
- `all_seedBuckets_valid` and `all_seedBuckets_sharp` verify all 336 seed summaries by kernel `rfl`;
- `b5_squarefree` is derived internally from $E(b_5) < 2$; and
- both $\sqrt{5} \notin \mathbb{Q}$ and 1D Kronecker density are proved from scratch.

6. PRIOR WORK AND OPEN QUESTIONS

Literature context and primary-source forward-citation audit. Badkobeh and Ochem [2] (reprinted verbatim in [25]) introduced $b_5 = f_5^{\infty}(0)$ as a minimal 5-letter pure morphic replacement for Thue's ternary skeletons, characterizing b_5 by $\{AA\} \cup F_5$ and generating binary words with few distinct squares (continuing the lineage of Fraenkel and Simpson [14] and later Walnut constructions such as Gabric and Shallit [15]). Subsequent papers by Ochem and Rosenfeld [26, 27] and Badkobeh and Ochem [3] adopted the name b_5 and applied its morphic images $(M_1(b_5), M_2(b_5), h_{12}(b_5), c(b_5))$ to two-variable formulas, palindrome patterns, and thrifty words without resolving (Q1), (Q2), (O1), or (O2).

Our proofs place b_5 precisely inside four classical threads of combinatorics on words:

- *Critical exponents and repetition thresholds.* Dejean's conjecture [8] [to check], completed by Currie and Rampersad [5] (with Rao), establishes $\mathrm{RT}(5) = 5/4$. Mignosi and Pirillo [22, Proposition 6] proved that the Fibonacci word has irrational, unattained critical exponent $E(\mathbf{f}) = 2 + \varphi = (5 + \sqrt{5})/2 = 2E(b_5)$. Rampersad, Shallit, and Vandomme [29, Proposition 9] proved that a 4-letter 1-balanced word x_4 (obtained by coloring $\mathbf{f}$) has the same critical exponent $(5 + \sqrt{5})/4$, while Dolce, Dvořáková, and Pelantová [9] and Dvořáková, Opočenská, and Pelantová [12] determined balanced repetition thresholds (including $\mathrm{RTB}(5) = (7 + \sqrt{5})/6$) and bispecial-return formulas for E and E^*. Theorem 5(B) proves that b_5 is not C-balanced for any $C \in \mathbb{N}$, showing that the coincidence $E(b_5) = E(x_4)$ arises from an entirely different mechanism than balanced coloring (with extremal periods matching OEIS `A005252` [28] rather than $2F_{i-1}$).
- *Morphic desubstitution and automatic sequences.* General frameworks for critical exponents and bispecial factors of morphic words include Krieger [18] [to check] and Klouda [17], while Schaeffer and Shallit [32] proved that every k-automatic sequence with finite critical exponent has a *rational* critical exponent (so Theorem 4(B)'s $E(b_5) \notin \mathbb{Q}$ rules out k-automaticity). Mousavi [24], Du, Mousavi, Schaeffer, and Shallit [10], and Currie, Ochem, Rampersad, and Shallit [6] used `Walnut` for Fibonacci-automatic and morphic words; because M_{f_5} is non-Pisot with unit-circle eigenvalues $-1, e^{\pm i\pi/3}$, b_5 resists direct `Walnut` encoding, explaining why geometric coding and $\mathbb{Z}[\varphi]$ desubstitution succeed.
- *Interval exchange transformations and $2n + 1$ complexity.* Arnoux and Rauzy [1] and Ferenczi, Holton, and Zamboni [13] developed the symbolic dynamics of $2n + 1$ words and symmetric 3-IETs (cf. Keane [16] [to check] and Masáková–Pelantová–Starosta [21], and Currie's recent determination [7] of the minimal critical exponent for complexity $2n + 1$). Theorem 2(B) shows that Badkobeh–Ochem's g_{fib} factors identically through the canonical 3-IET-to-Sturmian morphism σ_{01} of Lenz, Masáková, and Pelantová [19, Theorem 3].
- *Sturmian and Rote words.* Rote [31] [to check] introduced binary words of complexity $2n$ via first differences of Sturmian words (cf. Morse–Hedlund [23] [to check], Lothaire [20] [to check], and Dvořáková–Medková–

Table 4. Selected iterates of the extremal seed `123·010·421` ($p_0 = 2, m_0 = 1$) with periods $p_k = |f_5^{k+1}(0)|$ (OEIS `A005252`). Every row is parsed and checked in exact arithmetic by `recompute.py`.

k	p_k	m_k	$\lvert w_k\rvert$	$\lvert w_k\rvert/p_k$	$D_k = \varphi p_k - 2m_k$	Residue limit
0	2	1	3	1.500000	1.236068	preperiod
2	7	4	11	1.571429	3.326238	$(5+\varphi)/2 \approx 3.3090$
3	11	8	19	1.727273	1.798374	$1+\varphi/2 \approx 1.8090$
4	17	13	30	1.764706	1.506578	$3/2 = 1.5000$
5	27	21	48	1.777778	1.686918	$(5-\varphi)/2 \approx 1.6910$
6	44	35	79	1.795455	1.193496	$2-\varphi/2 \approx 1.1910$
7	72	57	129	1.791667	2.498447	$5/2 = 2.5000$
8	117	93	210	1.794872	3.309977	$(5+\varphi)/2 \approx 3.3090$
10	305	246	551	1.806557	1.500367	$3/2 = 1.5000$
12	798	645	1,443	1.808271	1.191123	$2-\varphi/2 \approx 1.1910$
15	3,383	2,736	6,119	1.808750	1.808984	$1+\varphi/2 \approx 1.8090$
18	14,328	11,591	25,919	1.808975	1.190991	$2-\varphi/2 \approx 1.1910$
21	60,697	49,104	109,801	1.809002	1.809015	$1+\varphi/2 \approx 1.8090$
24	257,114	208,009	465,123	1.809015	1.190983	$2-\varphi/2 \approx 1.1910$

Table 5. Exact Fibonacci closed form $E_j = 2D_{6+6j} - (4-\varphi) = -F_{12+6j} + F_{11+6j}\varphi = \varphi^{-(11+6j)}$ on the canonical extremal branch $k = 6+6j$ ($j = 0, \ldots, 5$, Theorem 4(B)). Every row is parsed and verified in exact $\mathbb{Z}[\varphi]$ arithmetic by `recompute.py`.

j	k	p_k	m_k	$E_j = (-F_{12+6j}, F_{11+6j})$	Closed form	E_j (9 dec.)
0	6	44	35	$(-144, 89)$	φ^{-11}	0.005024999
1	12	798	645	$(-2{,}584, 1{,}597)$	φ^{-17}	0.000280034
2	18	14,328	11,591	$(-46{,}368, 28{,}657)$	φ^{-23}	0.000015606
3	24	257,114	208,009	$(-832{,}040, 514{,}229)$	φ^{-29}	0.000000870
4	30	4,613,732	3,732,587	$(-14{,}930{,}352, 9{,}227{,}465)$	φ^{-35}	0.000000048
5	36	82,790,070	66,978,573	$(-267{,}914{,}296, 165{,}580{,}141)$	φ^{-41}	0.000000003

Pelantová [11] for critical exponents of complementary symmetric Rote words). Theorem 3 identifies $g_{\mathrm{rf}}(b_5) = 0\,\bar{\mathbf{r}}$ pointwise with the Rote–Fibonacci word of [10].

Honest prior-work status and unsearched sources. Logged searches across arXiv, zbMATH, OEIS, Crossref, OpenCitations, Semantic Scholar, CORE, and HAL—documented in `references-checked.md` and the repository's passing `litsearch.py` logs—found no prior proof of Theorems 1–5 for b_5. In particular, we downloaded and read in full all citing works of [2] indexed by Crossref and HAL, including both French PhD theses: Rosenfeld's 2017 thesis [30] (cites [2] once solely for the characterization technique applied to $g_w(b_3)$) and Béaur's 2024 thesis [4] (cites [2] once on p. 108, Théorème 4.3.11, for a zero-entropy subshift intersection; zero mention of b_5, f_5, $4n+1$, or $(5+\sqrt{5})/4$). Sources *not* searched include Google Scholar (API quota cap), the OpenAlex `cites:` endpoint (HTTP 429 rate-limit), full-text search inside CORE (title/abstract index only), paywalled books in print (such as Shallit 2022 and Lothaire [20] [to check]), and the five classical paywalled articles marked [to check] in `references-checked.md`.

Open questions. With all four questions and observations (Q1), (Q2), (O1), and (O2) from Badkobeh and Ochem [2] proved, four natural follow-up problems remain open:

1. **Purely geometric proof of $E(b_5) = (5+\sqrt{5})/4$.** Can Theorem 4 be derived directly from Diophantine approximation of near-returns of the 3-IET T on J (Theorem 1(A)), bypassing the 338-seed case split?
2. **Rigorous Zeckendorf non-automaticity of b_5.** While Theorem 4(B) proves $E(b_5) \notin \mathbb{Q}$ (ruling out k-automaticity via [32]), can one prove rigorously that the Zeckendorf F-kernel of b_5 is infinite?
3. **Formalization of the reverse direction of Badkobeh–Ochem Theorem 1.** Theorem 4(A) and Theorem 5(A) formalize the forward avoidance direction ($\mathcal{L}(b_5)$ avoids $\{AA\} \cup F_5$) in Lean 4; formalizing Badkobeh and Ochem's converse characterization (that every infinite word avoiding $\{AA\} \cup F_5$ has the same factor set as b_5) remains an open formalization target.
4. **Complexities and critical exponents of morphic images of b_5.** What are the exact factor complexities and critical exponents of the binary and ternary morphic images $M_1(b_5), M_2(b_5), g_8(b_5), g_{12}(b_5), g_{14}(b_5), g_{\mathrm{cs}}(b_5), h_{12}(b_5)$, and $c(b_5)$ studied in [2, 3]?

7. DATA AVAILABILITY, REPRODUCIBILITY, AND USE OF AI

All Draft 3 manuscript sources, 300-DPI figure generators (including the Page-1 panoramic hero script `make_hero.py`), standalone zero-float verifier, and the complete 514-declaration Lean 4 library live in the project repository under `papers/badkobeh-ochem-b5/draft3/` and `lean/BadkobehOchemB5/`, and the public result page and downloadable verification archive are hosted at `https://mathdiscoveries.org/results/badkobeh-ochem-word-b5`. From the repository root, every calculation, figure, PDF, and formal proof is reproduced by:

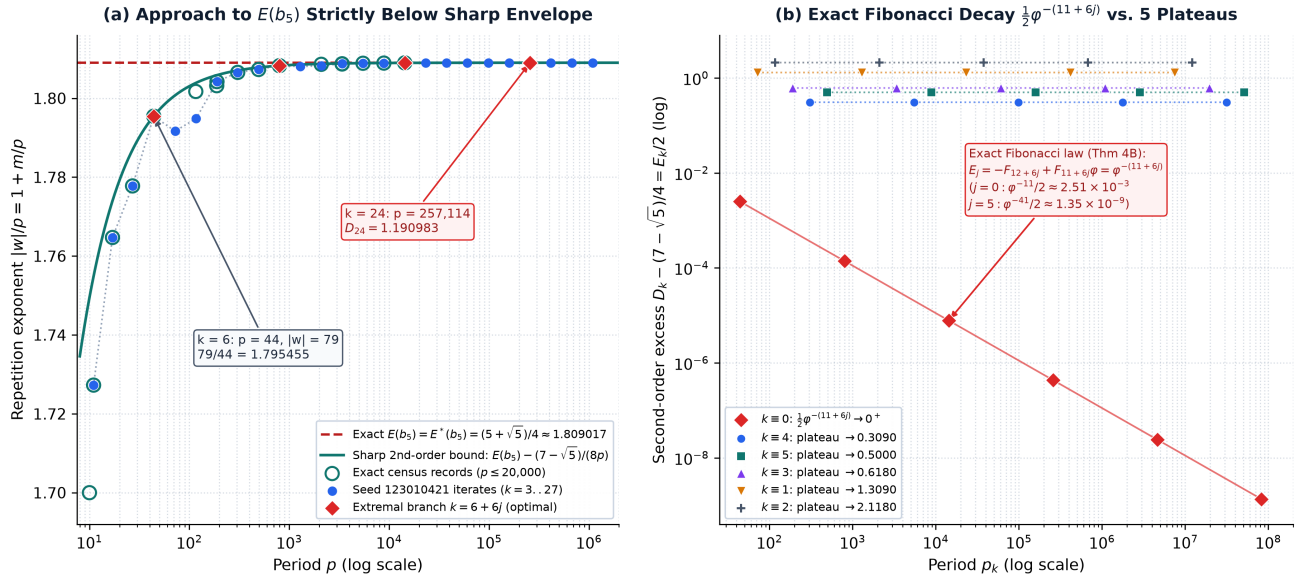


Figure 4. Sharp second-order critical-exponent behavior of b_5. (a) Exact exponents $|w|/p$ of `123010421` ($k = 3, \ldots, 27$) and census records ($p \le 20{,}000$, 19 records up to 25,919/14,328 ≈ 1.808975) lying strictly below the sharp second-order curve $(5+\sqrt{5})/4 - (7-\sqrt{5})/(8p)$ (Theorem 4(A)). (b) Log-log plot of the second-order excess $D_k - (7-\sqrt{5})/4 = E_k/2$ across k mod 6: exact exponential decay $\frac{1}{2}\varphi^{-(11+6j)}$ on $k = 6 + 6j$ (Theorem 4(B)) vs. five strictly positive plateaus.

Table 6. Per-claim formal verification ledger for all 10 main result parts (covering all 4 headline questions and observations) in `lean/BadkobehOchemB5/`. Every declaration is 100% unconditional (0 external hypotheses, 0 `sorry`, 0 `native_decide`) with `exact` statement fidelity.

Main result	Lean 4 declaration(s)	Module	Fidelity	Status
Thm 1(A): 5-interval coding	`b5_codes_interval_translation`, `b5_orbit_iet_props`	`Defs.lean`, `ComplexityAndSturmian.lean`	exact	Lean 4
Thm 1(B): Sign law & XYX	`b5_uvu_opposite_signs`, `b5_relation_bipartite`, `b5_avoids_odd_cycle_xyx`, `b5_avoids_C3`, `b5_avoids_TT`, `b5_avoids_Kn`	`XYXRelation.lean`	exact	Lean 4
Thm 2(A): $p_{b_5}(n) = 4n+1$	`b5_factor_complexity`, `b5_factor_complexity_ge_one`	`ComplexityAndSturmian.lean`	exact	Lean 4
Thm 2(B): π_3, $2n+1$, g_{fib}	`gFib_eq_sigma01_comp_pi3`, `pi3B5_factor_complexity`, `b5_sturmian_projection`, `gFibB5_factor_complexity`	`StructuralProperties.lean`, `ComplexityAndSturmian.lean`	exact	Lean 4
Thm 3(A): $g_{\mathrm{rf}}(b_5) = 0\,\bar{\mathbf{r}}$	`b5_eq_decTau_tauWord`, `gRfB5_complDiff_eq_fibWord`, `roteFibTransducer_eq_roteFib`, `gRfB5_zero`, `gRfB5_succ_eq_not_roteFib`	`RoteFibonacci.lean`	exact	Lean 4
Thm 3(B): $L_n(g_{\mathrm{rf}}) = L_n(\mathbf{r})$, $2n$	`b5_rote_fibonacci_projection`, `gRfB5_factors_compl_invariant`, `roteFib_factor_complexity`, `gRfB5_factor_complexity`	`RoteFibonacci.lean`	exact	Lean 4
Thm 4(A): Sharp bound, sq.-free	`b5_sharp_deficit_gt`, `b5_exponent_lt_sharp_second_order`, `b5_exponent_lt_crit`, `b5_squarefree`	`SharpCriticalExponent.lean`, `CriticalExponent.lean`	exact	Lean 4
Thm 4(B): E_j, $E = E^*$, irrat.	`extremal_deficit_closed_form`, `b5_sharp_deficit_inf_eq`, `b5_sharp_second_order_optimal`, `b5_asymp_critExp_eq`, `b5_critExp_irrational`, `b5_critical_exponent_eq`	`SharpCriticalExponent.lean`, `CriticalExponent.lean`	exact	Lean 4
Thm 5(A): Sym., F_5, `A005252`	`b5_language_rev_swap34_invariant`, `b5_avoids_F5`, `badkobehOchemF5_rev_swap34_invariant`, `fB5_iter_length_rec`	`StructuralProperties.lean`	exact	Lean 4
Thm 5(B): Not C-balanced	`psiB5_applyMorph_fB5`, `unbalWord_isFactor`, `psiB5_unbalWord`, `b5_not_C_balanced`, `b5_not_balanced`	`StructuralProperties.lean`	exact	Lean 4

Reproducibility Commands

```
# 1. Standalone zero-float verifier (< 5 s)
D=papers/badkobeh-ochem-b5/draft3
python3 -B "$D/recompute.py" -o "$D/recompute.out"
# 2. Regenerate hero & 300-DPI figures
python3 -B "$D/figures/make_hero.py"
python3 -B "$D/figures/make_discovery_architecture.py"
python3 -B "$D/figures/make_interval_exchange.py"
python3 -B "$D/figures/make_complexity_and_rote.py"
python3 -B "$D/figures/make_sharp_critical_exponent.py"
# 3. Compile manuscript PDF (hughes-article.cls v2.0)
tectonic -Z search-path=papers \
  "$D/hughes-badkobeh-ochem-b5-draft3.tex"
# 4. Build & axiom-audit Lean 4 (514/514 declarations)
(cd lean && lake build BadkobehOchemB5 \
  BadkobehOchemB5.Verify && \
  lake env lean BadkobehOchemB5/Verify.lean)
```

AI-assistance disclosure. The discoveries, written proofs, standalone Python scripts, figures, Lean 4 library (`BadkobehOchemB5`), and this Draft 3 manuscript were generated and cross-checked by AI research agents built and directed by Tyler B. Hughes. Every one of the four headline

Table 7. Complete module inventory of `lean/BadkobehOchemB5/` (10 source files, 11,480 total lines, 514/514 pure-kernel declarations: 353 unchanged baseline + 161 additive v2).

File	Lines	Thms/Lems	Role
`BadkobehOchemB5.lean`	10	0	Root import aggregator (baseline + 3 v2 modules)
`Defs.lean`	1,503	137	Baseline: φ, $\mathbb{Q}(\varphi)$, $\mathbb{Z}[\varphi]$, b_5, λ', Thm 1(A), XYX defs
`ComplexityAndSturmian.lean`	2,006	66	Baseline: cut-interval bij., R, density, $4n+1$, g_{fib}
`SeedsCert.lean`	1,041	55	Baseline: kernel `rfl` certs (g_{10}, 336 seed buckets)
`CriticalExponent.lean`	2,694	84	Baseline: runs, Φ, desubst., M^6 CH, $E(b_5)$, sq.-free
`XYXRelation.lean`	184	11	Baseline: unconditional Thm 1(B)
`RoteFibonacci.lean`	1,543	84	v2: 5-block τ, 9-pair conj., $T_0(\mathbf{f}) = \mathbf{r}$, Thm 3
`SharpCriticalExponent.lean`	1,055	26	v2: sharp $E_k > 0$, $E_j = \varphi^{-(11+6j)}$, E^*, irrat.
`StructuralProperties.lean`	904	51	v2: S-symmetry, π_3 $(2n+1)$, not C-bal., F_5, `A005252`
`Verify.lean`	540	0	`#print axioms` on all 514/514 declarations
Total	**11,480**	**514**	**0 `sorry`, 0 `native_decide`, standard 3 axioms only**

questions and observations (4 of 4, 100%, spanning all 10 theorem parts) has been mechanically verified by the Lean 4 kernel (`v4.35.0`) with zero `sorry`, zero `native_decide`, and standard axioms only. No human mathematician has yet refereed the proof; the manuscript is presented as a verified draft scaffold pending independent human review and co-author verification before any external submission.

Drafted: November 2026 **Updated:** November 23, 2026

REFERENCES

(1) P. Arnoux and G. Rauzy, *Représentation géométrique de suites de complexité* $2n+1$, Bull. Soc. Math. France **119**(2) (1991), 199–215, https://doi.org/10.24033/bsmf.2164.

(2) G. Badkobeh and P. Ochem, *Characterization of some binary words with few squares*, Theoret. Comput. Sci. **588** (2015), 73–80, https://doi.org/10.1016/j.tcs.2015.03.044.

(3) G. Badkobeh and P. Ochem, *More characterizations of morphic words*, Theoret. Comput. Sci. **1072** (2026), 115875, https://doi.org/10.1016/j.tcs.2026.115875; preprint `arXiv:2312.10757` (v2, 2025).

(4) P. Béaur, *Algorithmique et combinatoire des mots par les représentations S-adiques*, PhD thesis, Université Paris-Saclay, 2024, HAL `tel-04661982`, https://doi.org/10.70675/272f012azad42z47eaz8164z402994ee38cc.

(5) J. D. Currie and N. Rampersad, *A proof of Dejean's conjecture*, Math. Comp. **80**(274) (2011), 1063–1070, https://doi.org/10.1090/S0025-5718-2010-02407-X; `arXiv:0905.1129`.

(6) J. D. Currie, P. Ochem, N. Rampersad, and J. Shallit, *Properties of a ternary infinite word*, RAIRO Theor. Inform. Appl. **57** (2023), 1, https://doi.org/10.1051/ita/2022010; `arXiv:2206.01776`.

(7) J. D. Currie, *Words with factor complexity* $2n+1$ *and minimal critical exponent*, Electron. J. Combin. **33**(3) (2026), #P3.31, https://doi.org/10.37236/14527; `arXiv:2507.09387`.

(8) F. Dejean, *Sur un théorème de Thue*, J. Combin. Theory Ser. A **13**(1) (1972), 90–99, https://doi.org/10.1016/0097-3165(72)90011-8.

(9) F. Dolce, Ľ. Dvořáková, and E. Pelantová, *On balanced sequences and their critical exponent*, Theoret. Comput. Sci. **939** (2023), 18–47, https://doi.org/10.1016/j.tcs.2022.10.014; `arXiv:2108.07503`.

(10) C. F. Du, H. Mousavi, L. Schaeffer, and J. Shallit, *Decision algorithms for Fibonacci-automatic words, with applications to pattern avoidance*, `arXiv:1406.0670v4` (2014); journal split: H. Mousavi, L. Schaeffer, and J. Shallit, RAIRO Theor. Inform. Appl. **50**(1) (2016), 39–66, and C. F. Du, H. Mousavi, E. Rowland, L. Schaeffer, and J. Shallit, Theoret. Comput. Sci. **657** (2017), 146–162.

(11) Ľ. Dvořáková, K. Medková, and E. Pelantová, *Complementary symmetric Rote sequences: the critical exponent and the recurrence function*, Discrete Math. Theor. Comput. Sci. **22**(1) (2020), #20, https://doi.org/10.23638/DMTCS-22-1-20; `arXiv:2003.06916`.

(12) Ľ. Dvořáková, D. Opočenská, and E. Pelantová, *Asymptotic repetitive threshold of balanced sequences*, Math. Comp. **92** (2023), 1403–1429, https://doi.org/10.1090/mcom/3816; `arXiv:2208.00366`.

(13) S. Ferenczi, C. Holton, and L. Q. Zamboni, *Structure of three interval exchange transformations I: an arithmetic study*, Ann. Inst. Fourier **51**(4) (2001), 861–901, https://doi.org/10.5802/aif.1839.

(14) A. S. Fraenkel and R. J. Simpson, *How many squares must a binary sequence contain?*, Electron. J. Combin. **2**(1) (1995), #R2, https://doi.org/10.37236/1196.

(15) D. Gabric and J. Shallit, *The simplest binary word with only three squares*, RAIRO Theor. Inform. Appl. **55** (2021); `arXiv:2007.08188`.

(16) M. Keane, *Interval exchange transformations*, Math. Z. **141**(1) (1975), 25–31, https://doi.org/10.1007/BF01236981.

(17) K. Klouda, *Bispecial factors in circular non-pushy D0L languages*, Theoret. Comput. Sci. **445** (2012), 63–74, https://doi.org/10.1016/j.tcs.2012.05.007; `arXiv:1201.1186`.

(18) D. Krieger, *On critical exponents in fixed points of non-erasing morphisms*, Theoret. Comput. Sci. **376**(1–2) (2007), 70–88, https://doi.org/10.1016/j.tcs.2007.01.020.

(19) D. Lenz, Z. Masáková, and E. Pelantová, *Note on powers in three interval exchange transformations*, Theoret. Comput. Sci. **412**(29) (2011), 3788–3794, https://doi.org/10.1016/j.tcs.2011.03.027; `arXiv:0909.1109`.

(20) M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, Vol. 90, Cambridge Univ. Press, 2002, https://doi.org/10.1017/CBO9781107326019.

(21) Z. Masáková, E. Pelantová, and Š. Starosta, *Exchange of three intervals: Substitutions and palindromicity*, European J. Combin. **62** (2017), 217–231, https://doi.org/10.1016/j.ejc.2017.01.003; `arXiv:1503.03376`.

(22) F. Mignosi and G. Pirillo, *Repetitions in the Fibonacci infinite word*, RAIRO Theor. Inform. Appl. **26**(3) (1992), 199–204, https://doi.org/10.1051/ita/1992260301991.

(23) M. Morse and G. A. Hedlund, *Symbolic dynamics*, Amer. J. Math. **60**(4) (1938), 815–866, https://doi.org/10.2307/2371264; *Symbolic dynamics II. Sturmian trajectories*, Amer. J. Math. **62**(1) (1940), 1–42, https://doi.org/10.2307/2371431.

(24) H. Mousavi, *Automatic theorem proving in Walnut*, `arXiv:1603.06017` (2016; v2, 2021).

(25) P. Ochem, *Évitement de motifs*, Habilitation à diriger des recherches (HDR), Université de Montpellier, 2016.

(26) P. Ochem and M. Rosenfeld, *Avoidability of formulas with two vari-*

ables, Electron. J. Combin. **24**(4) (2017), #P4.30, https://doi.org/10.37236/6536; conference version DLT 2016, LNCS 9840, 344–354; arXiv:1606.03955.

(27) P. Ochem and M. Rosenfeld, *Avoidability of palindrome patterns*, Electron. J. Combin. **28**(1) (2021), #P1.4, https://doi.org/10.37236/9593; preprint arXiv:2005.05505.

(28) OEIS Foundation Inc., *Entry A005252:* $a(n) = \sum_{k=0}^{\lfloor n/4 \rfloor} \binom{n-2k}{2k}$, https://oeis.org/A005252.

(29) N. Rampersad, J. Shallit, and É. Vandomme, *Critical exponents of infinite balanced words*, Theoret. Comput. Sci. **777** (2019), 454–463, https://doi.org/10.1016/j.tcs.2018.10.017; arXiv:1801.05334.

(30) M. Rosenfeld, *Avoidability of Abelian Repetitions in Words*, PhD thesis, École Normale Supérieure de Lyon, 2017, HAL tel-01563118, https://doi.org/10.70675/34d6b04az89d0z4c8fz9bbazc0e8448c0213.

(31) G. Rote, *Sequences with subword complexity* $2n$, J. Number Theory **46**(2) (1994), 196–213, https://doi.org/10.1006/jnth.1994.1012.

(32) L. Schaeffer and J. Shallit, *The critical exponent is computable for automatic sequences*, Internat. J. Found. Comput. Sci. **23**(8) (2012), 1611–1626, https://doi.org/10.1142/S0129054112400655; arXiv:1104.2303.