Contents

Canonical Reduction and Inverse Reconstruction in PLRS Numeration

---

Abstract

We study the inverse problem for carry propagation in canonical positive linear recurrence sequence (PLRS) numeration systems: given only the histogram of carry depths arising from incrementation, can the underlying recurrence be recovered? We develop a purely combinatorial description of carry propagation, derived entirely from the recursive legality definition and independent of any numerical valuation. The legal words of each width form a rooted ordered tree with canonical minimal and maximal extensions at every node; these identify the immediate lexicographic successor of any legal word, which we then show coincides with arithmetic incrementation once the valuation map is introduced. Carry depth is shown to correspond not to a disjoint partition of legal words, as a naive count might suggest, but to a family of *nested tail classes* that biject with legal prefixes via a maximum-cylinder construction; exact-depth counts are recovered as first differences of these tails. We then address reconstruction directly, and show by explicit counterexample — the coefficient vectors (2) and $(1, 2)$, which generate an identical sequence of place values — that the histogram does *not* determine the defining coefficient vector uniquely. We identify the correct invariant, the **canonical reduction** (the presentation of least admissible order generating the observed sequence), and prove that the carry histogram determines this reduction exactly, recovering the original coefficient vector whenever it is already reduced. The four principal contributions are: (i) a fully combinatorial account of carry propagation via the tree structure of legal words; (ii) the prefix-cylinder tail identity governing carry-depth statistics; (iii) the identification and proof of the canonical-reduction phenomenon, including the non-uniqueness example; and (iv) a deterministic, finite reconstruction procedure recovering the canonical reduction from carry data.

---

1. Introduction

Positive linear recurrence sequences (PLRS) provide a broad class of numeration systems generalising the classical Zeckendorf representation. Every non-negative integer possesses a unique legal expansion

with respect to a canonical PLRS, and arithmetic operations in these systems induce non-trivial digit transformations governed by the recursive legality constraints.

One of the simplest arithmetic operations is incrementation. Given the legal representation of an integer N, incrementing to $N+1$ produces a new legal representation after a finite carry propagation. The number of digits changed during this process — the carry depth — determines the limiting carry distribution studied in prior work on Fibonacci, Tribonacci, and more general PLRS numeration systems.

The present paper studies the inverse problem: given only the histogram of carry depths, can the underlying numeration system be reconstructed?

Our approach is entirely combinatorial. We begin by analysing the recursive language of legal representations independently of their numerical interpretation. The legal words form a rooted ordered tree whose structure is determined solely by the recursive legality definition (Section 3). This tree carries canonical minimal and maximal extensions at every node (Section 4), which together identify the immediate lexicographic successor of any non-maximal legal word.

The arithmetic structure enters only afterwards (Section 5). We establish the Maximum Suffix Theorem and the Counting Theorem directly from recursive legality, and use these — together with a state-dependent monotonicity argument — to prove that lexicographic and numerical order coincide on legal words of fixed width. This identifies the combinatorial successor with arithmetic incrementation.

These results allow carry propagation to be described entirely through the geometry of maximal legal suffixes (Section 6). Rather than partitioning legal words by *exact* carry depth — which is not, as it turns out, a disjoint partition — we show that the natural objects are the *nested tail classes* $\{T_m \geq t\}$, which biject cleanly with legal prefixes via a maximum-cylinder construction. Exact-depth counts are then recovered as first differences of the tail counts.

Section 7 addresses reconstruction. We show, by explicit counterexample, that the carry-cardinality data do **not** determine the original coefficient vector uniquely: distinct canonical presentations of different orders can generate identical legal-word counts at every width. We identify the correct invariant — the **canonical reduction**, the unique presentation of least admissible order generating the observed sequence — and prove that the histogram determines it exactly, recovering the original vector whenever that vector is already reduced.

Section 8 works through the Fibonacci case and the (2)-vs-$(1,2)$ reduction example in full, states the reconstruction algorithm explicitly, and discusses complexity and possible extensions. Section 9 collects concluding remarks.

---

2. Canonical Positive Linear Recurrence Numeration

Let

$$\mathbf{c} = (c_1, \ldots, c_L), \qquad c_i \in \mathbb{Z}_{\geq 0}, \quad c_1 > 0, \quad c_L > 0.$$

(Coefficients are required to be non-negative integers. This is needed by Lemma 7.2, whose reconstruction formulas presuppose an integer-valued candidate vector.)

2.1. The canonical PLRS

The associated positive linear recurrence sequence (PLRS) $(H_n)_{n\geq 1}$ is defined by

$$H_1 = 1,$$

$$H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_n H_1 + 1, \qquad 1 \leq n < L,$$
$$H_{n+1} = c_1 H_n + \cdots + c_L H_{n+1-L}, \qquad n \geq L.$$

2.2. Representations

A finite digit string $a_1 a_2 \cdots a_m$ represents the value

$$\mathrm{val}(a_1 \cdots a_m) = \sum_{i=1}^{m} a_i H_{m+1-i}.$$

For each width m, $\mathcal{L}_m$ denotes the set of legal words of length exactly m.

2.3. Recursive legality

A finite word is legal if and only if:

1. **Initial case.** The word has length $m < L$ and equals $c_1 c_2 \cdots c_m$.
2. **Recursive case.** There exists $1 \leq s \leq L$ such that $a_i = c_i$ for $1 \leq i < s$, $0 \leq a_s < c_s$, and the remaining suffix consists of zero or more zeros followed by a legal word (or is entirely zero).

This recursive definition is the only legality assumption used throughout the paper.

2.4. Canonical Expansions

A legal representation of an integer N is a legal word whose value equals N. The classical generalized Zeckendorf theorem states that every non-negative integer possesses a unique legal representation. This fact provides the motivating background for the theory below; however, the principal structural results of this paper (Sections 3–6) are derived directly from the recursive legality definition of §2.3, rather than assumed from uniqueness.

2.5. Maximal Words

For each width m, define M_m to be the lexicographically largest element of $\mathcal{L}_m$. Existence and an explicit structural description of M_m are established in Section 5.

2.6. Successors

Given a non-maximal legal word $w \in \mathcal{L}_m$, its canonical successor is the legal word representing $\mathrm{val}(w) + 1$. At this stage no description of the successor operation is assumed; Sections 3–4 develop the combinatorial structure of the legal language needed to construct it, and Section 5 identifies it with arithmetic incrementation.

---

3. Structural Properties of the Legal Language

Theorem 3.1 (Unique Recursive Decomposition). Every non-zero legal word decomposes uniquely into initial or recursive form, with a unique breakpoint s.

Corollary 3.2 (Prefix Closure). Every prefix of a legal word is legal.

Corollary 3.3 (Restart Property). In the recursive decomposition $w = c_1 \cdots c_{s-1} a_s X$, the suffix X is itself a legal word.

Corollary 3.4 (Trailing-Zero Extension). If u is legal, then $u0^r$ is legal for every $r \geq 0$.

Theorem 3.5 (Consecutive Children). Let p be a legal prefix, and let s denote p's position within its current recursively-restarted matching block — positions are indexed $1, \ldots, L$ within each block, resetting to 1 immediately after every breakpoint (Corollary 3.3). The admissible digits following p are

$$\{0, 1, \ldots, c_s\} \text{ if } s < L, \qquad \{0, 1, \ldots, c_s - 1\} \text{ if } s = L.$$

In particular this set is always a consecutive interval $\{0, \ldots, d_{\max}\}$.

Proof. Every digit $d < c_s$ is admissible independently of the others: the word $c_1 \cdots c_{s-1} d$ is legal by the recursive clause with empty residual suffix, and by Trailing-Zero Extension (Corollary 3.4) extends to a legal word of any required length. This accounts for $\{0, \ldots, c_s - 1\}$ in both cases.

It remains to determine whether $d = c_s$ (continuing the match rather than breaking here) is admissible. If $s < L$, the word $c_1 \cdots c_s$ is itself a legal initial-case word (since $s < L$), and again extends to any required length by Corollary 3.4; so $d = c_s$ is admissible, giving the interval $\{0, \ldots, c_s\}$.

If $s = L$, the word $c_1 \cdots c_L$ (matching every coefficient of the current block exactly) satisfies neither legality clause: it fails the initial case because $m = L \not< L$, and it fails the recursive case because no position exhibits a deficiency $a_i < c_i$. Hence $d = c_s$ is *not* admissible when $s = L$, and the admissible set is exactly $\{0, \ldots, c_s - 1\}$.

No digit exceeds c_s in either case, directly from the legality definition (clause 2 requires $a_s < c_s$; no clause admits a digit exceeding the required coefficient). ■

(Remark: the case $s = L$ genuinely differs from $s < L$ and must be treated separately — matching every coefficient of the current block satisfies neither legality clause, so the "continue matching" digit is never admissible precisely at a block's final position.)

Note on Theorem 3.5's role: this is the theorem that forces a genuine "choice node" at any position where two legal words first diverge — it is the load-bearing fact behind the proof of Theorem 5.3 below.

Corollary 3.6 (Breakpoint Locality). Let p be a legal prefix ending in a genuine deficient digit — i.e. $p = c_1 \cdots c_{s-1} a$ with $a < c_s$ at the current block position s. Then, for every $r \geq 0$,

$$pZ \in \mathcal{L}_{|p|+r} \iff Z \in \mathcal{L}_r.$$

Equivalently, the legal completions of p to residual length r are exactly the members of $\mathcal{L}_r$, with no dependence on p beyond the fact that p ends in a deficiency.

Proof. ($\Leftarrow$) This is a direct instance of the recursive legality clause (§2.3): with breakpoint s, deficient digit $a < c_s$, and zero intervening zeros, "the remaining suffix consists of zero or more zeros

followed by a legal word" is satisfied by taking the suffix to be Z itself, for any $Z \in \mathcal{L}_r$. Hence pZ is legal.

($\Rightarrow$) Suppose pZ is legal. By Theorem 3.1, pZ has a *unique* recursive decomposition with breakpoint at some position s'. Since p already exhibits a deficiency at position s (with $a_i = c_i$ for $i < s$), and legal words match the coefficient vector exactly up to their own breakpoint, uniqueness forces $s' = s$ — the breakpoint of pZ coincides with the deficiency already present in p. The Restart Property (Corollary 3.3) then gives that the suffix following this breakpoint, namely Z, is legal: $Z \in \mathcal{L}_r$. ■

This sharpens Restart Property (Corollary 3.3), which only supplied the ($\Rightarrow$) direction, to a full equivalence. It recurs throughout Section 4 below.

---

4. Maximal and Minimal Extensions

Let p be a legal prefix and fix a target length $m \geq |p|$. Write

$$\mathcal{E}_m(p) = \{w \in \mathcal{L}_m : w \text{ begins with } p\}.$$

Lemma 4.0 (Unconditional Completability). For every legal word u and every target length $m \geq |u|$, $\mathcal{E}_m(u) \neq \varnothing$.

Proof. $u0^{m-|u|} \in \mathcal{E}_m(u)$, legal by Trailing-Zero Extension (Corollary 3.4), which places no restriction on u or on the target length. ■

This settles, once and for all, the distinction between two notions of "admissible" that might otherwise be conflated in the greedy constructions below: Theorem 3.5 characterizes admissibility of a digit d following a legal prefix p purely as *local legality* — the finite word pd is itself legal, with no reference to any target length. Lemma 4.0 shows that in this numeration system, local legality of a finite word is automatically sufficient for completability to *any* greater target length, because Trailing-Zero Extension is unconditional. Consequently "d is admissible after p" and "pd admits a legal completion to length m" coincide here for every $m \geq |p| + 1$ — but this coincidence is a genuine fact about this system (traceable to Corollary 3.4), not a tautology, and the greedy proofs below invoke it explicitly at each step rather than assuming it.

Theorem 4.1 (Minimal Extension). For every legal prefix p and every admissible length m, the set $\mathcal{E}_m(p)$ contains a unique lexicographically least element, denoted $E_{\min}(p)$ (extended to length m).

Proof. Construct a word recursively, starting from p. At each step, given a current legal partial word q (initially p) with $|q| < m$: since q is legal, Lemma 4.0 gives $\mathcal{E}_m(q) \neq \varnothing$, so some legal completion of q to length m exists, and in particular q has at least one admissible next digit (namely the first digit of any such completion, which is admissible by Theorem 3.5 since q followed by it is a prefix of a legal word, hence legal by Prefix Closure). Append the least admissible digit, 0 (or the least element of the admissible interval, which contains 0 by Theorem 3.5's construction). This produces a new legal partial word q' with $|q'| = |q| + 1$, again satisfying $\mathcal{E}_m(q') \neq \varnothing$ by Lemma 4.0, so the construction may continue. Repeating until length m is reached produces a legal word $w_0 \in \mathcal{E}_m(p)$.

For uniqueness, suppose $z \in \mathcal{E}_m(p)$ with $z \neq w_0$, and let i be the first position after $|p|$ at which z and w_0 differ. Both extend the same legal prefix up to position $i - 1$. At position i, w_0 was constructed by choosing the least admissible digit; since z's digit at i is admissible (as z is legal) and

different, it is strictly greater. Hence $z >_{\text{lex}} w_0$. As z was arbitrary, w_0 is the unique least element. ■

Theorem 4.2 (Maximal Extension). For every legal prefix p and every admissible length m, the set $\mathcal{E}_m(p)$ contains a unique lexicographically greatest element, denoted $E_{\max}(p)$.

Proof. Construct a word recursively as above, but at each step, given a current legal partial word q with $|q| < m$: by Lemma 4.0, $\mathcal{E}_m(q) \neq \varnothing$, so q has at least one admissible next digit (by the same argument as in Theorem 4.1's proof), and by Theorem 3.5 the admissible digits form a nonempty finite interval $\{0, \ldots, d_{\max}\}$; append its greatest element, $d_{\max}$. The result $q' = qd_{\max}$ is legal (admissibility, Theorem 3.5, means precisely that $qd_{\max}$ is legal), and by Lemma 4.0 again, $\mathcal{E}_m(q') \neq \varnothing$, so the construction may continue to the next step regardless of how much length remains. Repeating until length m is reached produces a legal word $w_1 \in \mathcal{E}_m(p)$.

Suppose $z \in \mathcal{E}_m(p)$ with $z >_{\text{lex}} w_1$, and let i be the first position after $|p|$ at which they differ. Both extend the same legal prefix through position $i-1$, so the set of admissible digits at position i is the same interval for both. At position i, w_1's digit is the greatest admissible one, $d_{\max}$; since z's digit is admissible and z's word is lexicographically larger, z's digit at i must exceed $d_{\max}$ — contradicting that $d_{\max}$ is the maximum of the admissible interval. Hence no such z exists, and $w_1 = E_{\max}(p)$ is the unique greatest element. ■

Lemma 4.3 (Active suffix is maximal). Let $w \in \mathcal{L}_m$ be non-maximal, let a occupy the rightmost position of w that can be increased by 1 while yielding a legal prefix (this position exists because $w \neq M_m$), and write $w = UaX$. Then $X = E_{\max}(Ua)$.

Proof. Since $X \in \mathcal{E}_{m-|Ua|}(Ua)$ by construction, and $E_{\max}(Ua)$ is by Theorem 4.2 the unique lexicographically greatest element of that set, suppose toward contradiction $X \neq E_{\max}(Ua)$; then $X <_{\text{lex}} E_{\max}(Ua)$. Let j be the first position within X at which X and $E_{\max}(Ua)$ differ; at that position $E_{\max}(Ua)$'s digit exceeds X's digit, and both extend the same legal prefix $Ua(X_1 \cdots X_{j-1})$ through position $j-1$ (since $E_{\max}(Ua)$ agrees with X up to there by minimality of j). By Theorem 3.5, the admissible digits at position j form a consecutive interval containing both X_j and $E_{\max}(Ua)_j > X_j$; in particular $X_j + 1$ is admissible there. But then the position within w corresponding to index j of X is itself incrementable while preserving legality — and this position lies strictly to the right of the pivot a, contradicting that a was chosen as the *rightmost* incrementable position of w.

Hence no such j exists, so $X = E_{\max}(Ua)$. ■

(Remark: "incrementable" requires not merely that $X_j + 1$ is locally admissible in the sense of Theorem 3.5, but that the resulting prefix $Ua(X_1 \cdots X_{j-1})(X_j + 1)$ admits a legal completion to the full target length m. Since this prefix is legal (Theorem 3.5), Lemma 4.0 guarantees exactly such a completion exists — so the position is genuinely incrementable, not merely admissible as an isolated digit.)

Theorem 4.4 (Canonical Lexicographic Successor). With $w = UaX$ as in Lemma 4.3, the immediate lexicographic successor of w within $\mathcal{L}_m$ is

$$w^+ = U(a+1)\, E_{\min}(U(a+1)).$$

Proof. First, Ua is a legal prefix by construction (it is a prefix of the legal word w, hence legal by Prefix Closure), and $a+1$ is admissible after U by hypothesis, so $U(a+1)$ is itself a legal prefix

and $E_{\min}(U(a+1))$ is well-defined by Theorem 4.1. Since $U(a+1)E_{\min}(U(a+1))$ and $w = UaX$ first differ at the position of a, where $a < a+1$, we have $w^+ >_{\text{lex}} w$ by construction.

It remains to show no legal word lies strictly between them. Let $z \in \mathcal{L}_m$ with $z >_{\text{lex}} w$. If z and w first differ at a position before the pivot, then since $z > w$ there, z's digit there is at least one more than w's, so $z \geq_{\text{lex}} U(a+1)0^{\cdots} \geq_{\text{lex}} U(a+1)E_{\min}(U(a+1)) = w^+$ (the last inequality because $E_{\min}$ is by definition the smallest legal completion). If instead z and w agree through the pivot position (both have digit a there) and first differ afterward, then since $X = E_{\max}(Ua)$ (Lemma 4.3), z sharing the prefix Ua cannot exceed X there — so this case is impossible for $z > w$. Finally, if z agrees with w through the pivot and is *smaller* afterward, then $z <_{\text{lex}} w$, contradicting $z > w$.

Hence every $z >_{\text{lex}} w$ satisfies $z \geq_{\text{lex}} w^+$, so w^+ is the immediate successor. ■

Corollary 4.5 (Global Maximal-Suffix Normal Form). Let $w \in \mathcal{L}_m$ be non-maximal, and let $r = T_m(w)$ (carry depth, formally defined in §6.2, but usable here as $|X|$ from Theorem 4.4's decomposition $w = UaX$). Then there exist a legal prefix U and digit a such that

$$w = UaM_r, \qquad w^+ = U(a+1)0^r.$$

In particular, the suffix erased during any successor transition is the global lexicographically maximal legal word of its length — not merely the maximal completion of the specific prefix Ua.

Proof. Since a is (by construction, Theorem 4.4) the rightmost incrementable digit, $a+1$ is admissible at the pivot, so $a < d_{\max}$ at that position. By Theorem 3.5, $d_{\max} = c_s$ (if $s < L$) or $c_s - 1$ (if $s = L$); in either case $a < c_s$, so a is a genuine deficient digit at its block position s. By Breakpoint Locality (Corollary 3.6), the legal completions of Ua to residual length r are exactly $\mathcal{L}_r$ — independent of U and a beyond this deficiency. Lexicographic comparison among words sharing the fixed prefix Ua reduces to lexicographic comparison of their length-r tails, so the maximum of this cylinder is Ua followed by the lexicographic maximum of $\mathcal{L}_r$, namely M_r (§2.5). Combined with Lemma 4.3 ($X = E_{\max}(Ua)$), this gives $X = M_r$.

For the successor suffix: $U(a+1)$ is legal (Theorem 4.4's proof), so by Trailing-Zero Extension (Corollary 3.4), $U(a+1)0^r$ is legal. By Theorem 3.5, the admissible interval at every position always contains 0 as its minimum (in the interior case trivially; in the block-final case because $c_L > 0$ is guaranteed by the standing hypothesis of §2, keeping $\{0, \ldots, c_L - 1\}$ non-empty), so the greedy least-completion construction of Theorem 4.1 selects 0 at every step, giving $E_{\min}(U(a+1)) = 0^r$. Theorem 4.4 then gives $w^+ = U(a+1)0^r$. ■

(In the binary case $a = 0$, this reduces to the familiar form $U0M_r \mapsto U10^r$.)

---

5. Arithmetic Structure of the Legal Language

5.1. The Valuation Map — unchanged.

5.2. Maximum Suffix Theorem

Theorem 5.1. For every $m \geq 0$, $S_m := \max_{w \in \mathcal{L}_m} \text{val}(w) = H_{m+1} - 1$.

Proof. By strong induction on m. For $m = 0$, $S_0 = 0 = H_1 - 1$ trivially.

Case $m < L$ (upper bound and achievability together). The unique length-m initial word $w_0 = c_1 \cdots c_m$ has value $\sum_{i=1}^{m} c_i H_{m+1-i} = H_{m+1} - 1$ exactly, by the canonical initial condition

(§2.1). For any other legal word w of length m, Theorem 3.1 gives a recursive decomposition $w = c_1 \cdots c_{s-1} a_s X$ with $s \le m$, $a_s < c_s$, $X \in \mathcal{L}_{m-s}$. By the induction hypothesis, $\mathrm{val}(X) \le H_{m-s+1} - 1$, so

$$\mathrm{val}(w) \le \sum_{i=1}^{s-1} c_i H_{m+1-i} + (c_s - 1) H_{m+1-s} + (H_{m-s+1} - 1) = \sum_{i=1}^{s} c_i H_{m+1-i} - 1,$$

using $H_{m+1-s} = H_{m-s+1}$ to cancel the two middle terms. Since $s \le m < L$, this is a partial sum of the same identity defining $H_{m+1} - 1 = \sum_{i=1}^{m} c_i H_{m+1-i}$ (§2.1), and every term $c_i H_{m+1-i}$ is non-negative, so $\sum_{i=1}^{s} c_i H_{m+1-i} \le \sum_{i=1}^{m} c_i H_{m+1-i} = H_{m+1} - 1$. Hence $\mathrm{val}(w) \le H_{m+1} - 2 < \mathrm{val}(w_0)$. So w_0 achieves the maximum and $S_m = H_{m+1} - 1$.

Case $m \ge L$. Every legal word $w \in \mathcal{L}_m$ decomposes (Theorem 3.1) as $w = c_1 \cdots c_{s-1} a_s X$ for some $1 \le s \le L$, $a_s < c_s$, $X \in \mathcal{L}_{m-s}$. As above, using the induction hypothesis on X (valid since $m - s < m$) and cancelling $H_{m+1-s} = H_{m-s+1}$,

$$\mathrm{val}(w) \le \sum_{i=1}^{s} c_i H_{m+1-i} - 1.$$

Now s ranges up to L (not merely m, since $m \ge L$ here), so we bound against the homogeneous recurrence identity valid for $m \ge L$: $H_{m+1} = \sum_{i=1}^{L} c_i H_{m+1-i}$ (§2.1). Since $s \le L$ and every term is non-negative, $\sum_{i=1}^{s} c_i H_{m+1-i} \le \sum_{i=1}^{L} c_i H_{m+1-i} = H_{m+1}$, giving $\mathrm{val}(w) \le H_{m+1} - 1$ for every legal word w of length m. This proves $S_m \le H_{m+1} - 1$.

For achievability, take $s = L$ (so the bound above is tight with no dropped terms) and $a_L = c_L - 1$. By the induction hypothesis applied to width $m - L < m$, there exists some $W^* \in \mathcal{L}_{m-L}$ with $\mathrm{val}(W^*) = H_{m-L+1} - 1$ (the induction hypothesis asserts only the *existence* of such a word achieving the maximum value at that width — no claim about which specific word it is, or about lexicographic order, is needed here). Define

$$w^* = c_1 \cdots c_{L-1} (c_L - 1)\, W^*.$$

This is legal: it matches the recursive legality clause exactly, with breakpoint $s = L$, deficient digit $a_L = c_L - 1 < c_L$, and residual suffix W^* itself legal. Computing its value,

$$\mathrm{val}(w^*) = \sum_{i=1}^{L-1} c_i H_{m+1-i} + (c_L - 1) H_{m+1-L} + \mathrm{val}(W^*) = \sum_{i=1}^{L} c_i H_{m+1-i} - H_{m+1-L} + (H_{m-L+1} - 1).$$

Since $H_{m+1-L} = H_{m-L+1}$, these two terms cancel, leaving $\mathrm{val}(w^*) = H_{m+1} - 1$, matching the upper bound exactly. Hence $S_m = H_{m+1} - 1$. ■

(This proof deliberately avoids naming M_{m-L} — the lexicographically maximal word of §2.5 — as the witness at width $m - L$, since at this point in the paper lexicographic and numerical order have not yet been shown to coincide; that identification is Theorem 5.3 onward. The induction hypothesis of Theorem 5.1 itself is enough to supply some* value-maximizing word W^*, without needing to know anything about its lexicographic status. Corollary 5.8 below closes this loop once Monotonicity is available, confirming that M_m and the value-maximizer coincide after all.)*

5.3. Counting Theorem

Theorem 5.2. For every $m \ge 0$, $|\mathcal{L}_m| = H_{m+1}$.

5.4. Order Theory

The order theory is developed as four logically separate results, each citing only earlier ones, so that no step assumes what a later step is meant to prove.

Theorem 5.3 (Lexicographic Monotonicity). Let $u, v \in \mathcal{L}_m$. If $u <_{\text{lex}} v$, then $\text{val}(u) < \text{val}(v)$.

Proof. Let i be the first position at which u, v differ, so $u_j = v_j$ for $j < i$ and $u_i < v_i$. Let p be their common prefix of length $i-1$; by Prefix Closure (3.2), p is legal.

By the Consecutive-Children description (Theorem 3.5), if the recursive block underlying p has not yet reached a deficient-choice position, only one digit is admissible at position i — so u and v could not differ there, contrary to hypothesis. Hence position i **is** a deficient-choice position, where by Theorem 3.5 the admissible digits form a consecutive interval $\{0, \ldots, d_{\max}\}$ (with $d_{\max} = c_s$ or $c_s - 1$, depending on whether the position is interior or block-final — the argument below does not need to know which). In either case, every digit strictly below $d_{\max}$ is a breakpoint digit: at an interior position the sole non-breakpoint digit is $d_{\max} = c_s$ itself (Theorem 3.5's proof), so digits below it are exactly the breakpoint digits $\{0, \ldots, c_s - 1\}$; at a block-final position *every* admissible digit, including $d_{\max}$ itself, is a breakpoint digit, so in particular all digits below $d_{\max}$ are too.

Since v_i is admissible, $v_i \le d_{\max}$; combined with $u_i < v_i$, this gives $u_i < d_{\max}$. Hence u_i is a breakpoint digit **for** u. (Nothing is claimed about v at this position — v may either break here as well or continue matching; the argument below does not need to know which.) By the Restart Property (3.3), the suffix of u strictly after position i is a genuine, freestanding legal word $X \in \mathcal{L}_{m-i}$. By the Maximum Suffix Theorem, $\text{val}(X) \le S_{m-i} = H_{m-i+1} - 1$.

Writing $u = p\,u_i\,X$ and $v = p\,v_i\,Y$ with $\text{val}(Y) \ge 0$:

$$\text{val}(v) - \text{val}(u) = (v_i - u_i)H_{m+1-i} + \text{val}(Y) - \text{val}(X) \ge H_{m+1-i} - S_{m-i} = H_{m+1-i} - (H_{m-i+1} - 1) = 1.$$

Hence $\text{val}(u) < \text{val}(v)$. ■

Corollary 5.4 (Injectivity). The map $\text{val} : \mathcal{L}_m \to \mathbb{N}$ is injective. *(Immediate from Theorem 5.3 applied to either ordering of $u \neq v$.)*

Theorem 5.5 (Interval Theorem). val is a bijection $\mathcal{L}_m \to \{0, \ldots, H_{m+1} - 1\}$. *(Injectivity from 5.4; both sets have cardinality H_{m+1} by the Counting Theorem 5.2; an injection between finite sets of equal cardinality is a bijection.)*

Corollary 5.6 (Order Isomorphism). For legal words of fixed width, $u <_{\text{lex}} v \iff \text{val}(u) < \text{val}(v)$. *(Forward direction is Theorem 5.3; the converse follows because injectivity and trichotomy rule out both equality and the reversed inequality.)*

Theorem 5.7 (Canonical Successor Theorem). Let $w \in \mathcal{L}_m$ be non-maximal. The immediate lexicographic successor of w (Theorem 4.4) is precisely the legal representation of $\text{val}(w) + 1$. *(Immediate from Corollary 5.6.)*

Corollary 5.8 (Maximal word realizes the extremal value). $\text{val}(M_m) = S_m = H_{m+1} - 1$, where M_m is the lexicographically maximal word of §2.5.

Proof. For any $w \in \mathcal{L}_m$ with $w \neq M_m$, maximality of M_m in the (finite, total) lexicographic order on $\mathcal{L}_m$ gives $M_m >_{\text{lex}} w$, so $\text{val}(M_m) > \text{val}(w)$ by Theorem 5.3. Hence M_m is also the value-maximizing word, and $\text{val}(M_m) = S_m = H_{m+1} - 1$ by Theorem 5.1. ■

(This is the corollary promised in §2.5's remark that "existence and an explicit structural description of M_m are established in Section 5"; it also confirms that the witness word w^* constructed inside Theorem 5.1's proof — never named as M_m there, deliberately — is consistent with M_m after the fact, though the two are not asserted identical beyond both achieving value $H_{m+1} - 1$; Corollary 5.4 (Injectivity) shows no other word can share that value, so in fact $w^* = M_m$ necessarily.)

Dependency chain for this section: Prefix Closure → Consecutive Children → Restart Property → Maximum Suffix Theorem → Monotonicity → Injectivity → Interval Theorem → Order Isomorphism → Canonical Successor Theorem. No step depends on any result from Section 6.

---

6. Carry Depth and the Prefix-Cylinder Identity

6.1. Prefix Cylinders

Let $p \in \mathcal{L}_{m-t}$ be a legal prefix of length $m - t$. Define its depth-t cylinder by

$$\mathcal{C}_m(p) = \{w \in \mathcal{L}_m : w \text{ begins with } p\}.$$

By the Maximal Extension Theorem (4.2), $\mathcal{C}_m(p)$ has a unique lexicographically greatest element, denoted $\max \mathcal{C}_m(p)$. Distinct prefixes determine disjoint cylinders.

6.2. Carry Depth

Let $w \in \mathcal{L}_m$ be non-maximal, and let w^+ denote its successor (Theorem 4.4). Write $w = UaX$, $w^+ = U(a+1)Y$, where U is the longest common prefix of w and w^+. The **carry depth** of w is $T_m(w) = |X|$: the number of positions strictly to the right of the incremented pivot. Equivalently, $T_m(w) \geq t$ if and only if the final t positions lie entirely inside the suffix replaced during incrementation.

Convention. $T_m(M_m) := m$. For every t from 0 to m, M_m's own length-$(m-t)$ prefix p_t has M_m as the maximum of its cylinder $\mathcal{C}_m(p_t)$, since M_m is the global maximum of all of $\mathcal{L}_m$. So M_m naturally belongs to every tail class $\{T_m \geq t\}$, $0 \leq t \leq m$ — exactly what this convention encodes, removing any need for a separate overflow exception below.

6.3. Cylinder Maxima and Carry Tails

Theorem 6.1 (Prefix-Cylinder Bijection). For each $0 \leq t \leq m$, the map $\Phi_{m,t} : \mathcal{L}_{m-t} \to \{w \in \mathcal{L}_m : T_m(w) \geq t\}$, $\Phi_{m,t}(p) = \max \mathcal{C}_m(p)$, is a bijection.

Proof. Let $p \in \mathcal{L}_{m-t}$ and $w = \max \mathcal{C}_m(p)$. Since w is the greatest legal word having prefix p, no legal word larger than w can retain the complete prefix p. When w is not the global maximum, its successor w^+ must therefore differ from w at or before position $m - t$, so all of the final t positions belong to the replaced suffix and $T_m(w) \geq t$; when $w = M_m$, $T_m(w) \geq t$ holds directly by the convention above. This shows $\Phi_{m,t}$ maps into the stated tail class. The map is injective because distinct prefixes determine disjoint cylinders (§6.1) and therefore have distinct maxima.

Surjectivity. Let $w \in \mathcal{L}_m$ satisfy $T_m(w) \geq t$; if $w = M_m$ the claim is immediate from the convention above, so assume w non-maximal. Write $w = UaX$, $w^+ = U(a+1)Y$ as in §6.2; since $T_m(w) \geq t$, the pivot lies within the first $m - t$ positions, so U is (a prefix of) p, w's own length-$(m-t)$ prefix.

Suppose toward contradiction some $z \in \mathcal{C}_m(p)$ has $z > w$. Since w^+ is the least legal word numerically exceeding w (Theorem 5.7), $\mathrm{val}(w^+) \le \mathrm{val}(z)$, so by the Interval Theorem / Order Isomorphism (5.5–5.6), $w^+ \le_{\mathrm{lex}} z$. But z shares the full prefix p with w, which includes the pivot position, so z's digit there is a — while w^+'s digit there is $a+1$. Hence $z <_{\mathrm{lex}} w^+$, a contradiction. So no such z exists, and $w = \max \mathcal{C}_m(p)$. ■

6.4. Exact Tail Counts

$$A_{m,t} := \#\{w \in \mathcal{L}_m : T_m(w) \ge t\} = |\mathcal{L}_{m-t}| = H_{m-t+1}.$$

6.5. Exact-Depth Counts

$$B_{m,t} := \#\{w : T_m(w) = t\} = A_{m,t} - A_{m,t+1} = H_{m-t+1} - H_{m-t}, \qquad 0 \le t < m,$$

with $B_{m,m} = H_1 = 1$ (just M_m itself, under the adopted convention — no separate terminal case needed).

6.6. Probability Distribution

Choose w uniformly from $\mathcal{L}_m$. By §6.4, the carry-depth tail satisfies

$$\mathbb{P}(T_m \ge t) = \frac{H_{m-t+1}}{H_{m+1}}.$$

If $H_n \sim \kappa\rho^n$ for the dominant characteristic root $\rho > 1$, then for every fixed t, $\mathbb{P}(T_m \ge t) \to \rho^{-t}$ as $m \to \infty$. Writing $q = \rho^{-1}$, the limiting carry depth has geometric tail $\mathbb{P}(T \ge t) = q^t$, hence $\mathbb{P}(T = t) = (1-q)q^t$.

6.7. Information Recoverable from the Histogram

The tail counts determine the legal-word cardinalities directly: $|\mathcal{L}_r| = A_{m,m-r}$ for $0 \le r \le m$. Equivalently, the exact-depth counts determine the tails by cumulative summation, $A_{m,t} = \sum_{j \ge t} B_{m,j}$. Thus either the tail histogram or the complete exact-depth histogram recovers the finite cardinality sequence $|\mathcal{L}_0|, |\mathcal{L}_1|, \ldots, |\mathcal{L}_m|$, with no overflow exception needed under the convention of §6.2. This is the precise input required by the reconstruction theory of Section 7.

---

7. Canonical Reduction and Inverse Reconstruction

(This section develops the reconstruction theory in full. An unconditional uniqueness claim — that the histogram determines the coefficient vector outright — is false, as Example 7.1 below shows by explicit counterexample; the correct invariant, the canonical reduction, is developed in its place.)

7.1. Equivalent Canonical Presentations

Two canonical coefficient vectors $\mathbf{c}, \mathbf{d}$ are **equivalent** if $H_n^{(\mathbf{c})} = H_n^{(\mathbf{d})}$ for all n.

Example 7.1. $\mathbf{c} = (2)$ and $\mathbf{d} = (1,2)$ both generate $H_n = 2^{n-1}$, since $2^n = 2^{n-1} + 2 \cdot 2^{n-2}$. Their characteristic polynomials are $x - 2$ and $x^2 - x - 2 = (x-2)(x+1)$ respectively — the second properly contains the first as a factor.

7.2. The Candidate of a Fixed Order

Lemma 7.2. For a fixed sequence (H_n) with $H_1 = 1$, there is at most one coefficient vector of each prescribed order L generating it, obtained by solving the canonical initial-condition equations sequentially for $c_1, \ldots, c_L$.

7.3–7.4. Admissible Orders and Canonical Reduction

Definition 7.3. The canonical reduction of a presentation is the coefficient vector of least admissible order generating its sequence.

Theorem 7.4. Every PLRS sequence admitting a canonical presentation has a unique canonical reduction. *(Well-ordering of $\mathbb{Z}_{>0}$ plus Lemma 7.2.)*

7.5. Corrected Reconstruction Theorem

Theorem 7.5. Given the complete carry histogram (hence the cardinality sequence $|\mathcal{L}_0|, |\mathcal{L}_1|, \ldots$ via Section 6), the carry data determine the canonical reduction of the underlying presentation uniquely, and recover the original coefficient vector exactly when that presentation is reduced.

7.6. Reconstruction Procedure

Given the cardinality sequence $|\mathcal{L}_0|, |\mathcal{L}_1|, \ldots, |\mathcal{L}_m|$ (equivalently $H_1, \ldots, H_{m+1}$), test candidate orders $L = 1, 2, 3, \ldots$ in increasing order. For each L, Lemma 7.2 determines a unique candidate coefficient vector from $H_1, \ldots, H_{L+1}$; reject it if any coefficient is negative or if $c_1 \leq 0$ or $c_L \leq 0$; otherwise verify the homogeneous recurrence $H_{n+1} = \sum_{i=1}^{L} c_i H_{n+1-i}$ against every remaining observed term. The first order to pass both tests is the least admissible order, hence — by Theorem 7.4 — the canonical reduction. With only a finite histogram (width m), this procedure identifies every candidate order consistent with the observed range; a genuinely larger true order that happens to agree with a smaller candidate over the observed range cannot be ruled out without further terms. A full worked instance of this procedure is given in §8.3.

7.7. Scope of the Result

Carry-cardinality data distinguish equivalence classes of presentations, not individual presentations; within each class there is a unique reduced representative, which is the natural output of inverse reconstruction.

---

8. Examples, Algorithmic Aspects, and Extensions

8.1. Worked example: the Fibonacci case

Take $\mathbf{c} = (1, 1)$, $L = 2$. Then $H_1 = 1$, $H_2 = c_1 H_1 + 1 = 2$, and $H_{n+1} = H_n + H_{n-1}$ for $n \geq 2$, giving H_n the classical Fibonacci sequence $1, 2, 3, 5, 8, 13, \ldots$ — legal words are exactly Zeckendorf strings avoiding two consecutive 1s.

For $m = 3$: $\mathcal{L}_3 = \{000, 001, 010, 100, 101\}$, with values $0, 1, 2, 3, 4$ respectively, confirming $|\mathcal{L}_3| = 5 = H_4$ (Theorem 5.2) and val bijects onto $\{0, \ldots, 4\}$ (Theorem 5.5). The maximal word is $M_3 = 101$, value $H_4 - 1 = 4$ (Theorem 5.1).

Carry tails: $A_{3,0} = |\mathcal{L}_3| = 5$, $A_{3,1} = |\mathcal{L}_2| = 3$, $A_{3,2} = |\mathcal{L}_1| = 2$, $A_{3,3} = |\mathcal{L}_0| = 1$. Exact-depth counts by first difference: $B_{3,0} = 2$, $B_{3,1} = 1$, $B_{3,2} = 1$, $B_{3,3} = 1$ (the last by the $T_m(M_m) := m$ convention, §6.5), summing correctly to $5 = |\mathcal{L}_3|$. Direct inspection of successors confirms this: $000 \to 001$ ($T = 0$), $001 \to 010$ ($T = 1$), $010 \to 100$ ($T = 2$), $100 \to 101$ ($T = 0$), and $101 = M_3$ ($T := 3$ by convention).

8.2. Worked example: canonical reduction

Recall Example 7.1: $\mathbf{c} = (2)$ and $\mathbf{d} = (1, 2)$ both generate $H_n = 2^{n-1}$. Given only the sequence $1, 2, 4, 8, 16, \ldots$, the reconstruction procedure of §7.6 tests $L = 1$ first: Lemma 7.2 gives the unique order-1 candidate $c_1 = H_2/H_1 = 2$, which satisfies $c_1 > 0$ and reproduces $H_{n+1} = 2H_n$ for all observed n. Since $L = 1$ already succeeds, the procedure halts there — never reaching the order-2 candidate $(1, 2)$ — and correctly outputs the canonical reduction (2). This illustrates Theorem 7.5 directly: the reduction is recovered regardless of which presentation the histogram actually arose from.

8.3. Reconstruction algorithm

Input: the cardinality sequence $H_1, \ldots, H_{m+1}$ recovered from a carry histogram of width m (via §6.7). **Output:** the canonical reduction, or "insufficient data" if no admissible order is confirmed within the observed range.

```
for L = 1, 2, 3, ... up to m:
    solve for (c_1,...,c_L) via Lemma 7.2's recursive formulas
      using H_1,...,H_{L+1}
    if any c_i < 0, or c_1 <= 0, or c_L <= 0:  reject L
    verify H_{n+1} = sum_i c_i H_{n+1-i} for all n with L <= n <= m-1
    if verification fails:  reject L
    else: return (c_1,...,c_L)   # first admissible order found
return "insufficient data"
```

Because admissible orders are tested in increasing order and the first success is returned, the procedure returns the least admissible order consistent with the observed data — the canonical reduction, when the observed range is long enough to rule out all smaller inadmissible orders (a short histogram may leave some larger order not yet excluded, per the finite-data caveat of §7.6).

8.4. Complexity

For a candidate order L, Lemma 7.2's formulas determine $c_1, \ldots, c_L$ in $O(L^2)$ arithmetic operations (each c_n requires summing $n - 1$ prior terms). Verifying the homogeneous recurrence against the remaining $m - L$ observed terms costs $O(L(m - L))$ operations. Summing over all candidate orders $L = 1, \ldots, m$, the total cost is $O(m^3)$ arithmetic operations on integers whose size grows with H_{m+1} (itself exponential in m), so the true bit-complexity is correspondingly larger; we do not pursue a sharper bound here.

8.5. Extensions and open directions

- **Partial histograms.** The reconstruction procedure as stated assumes the cardinality sequence is known exactly up to width m. It would be useful to characterize the minimal m needed to

certify the canonical reduction for a given family, rather than only exhibiting the (worst-case) $O(m^3)$ procedure above.

- **Beyond canonical presentations.** The present theory is developed entirely from the recursive legality definition of §2.3, without further structural assumptions on $\mathbf{c}$. It would be natural to ask whether the canonical-reduction phenomenon (Example 7.1) is itself classifiable — e.g., whether every non-reduced presentation arises from a reducible characteristic polynomial in the manner of that example, or whether other mechanisms exist.
- **Relation to prior carry statistics.** Berthé, Frougny, Rigo and Sakarovitch study a related but distinct carry statistic (terminal region traversed by the successor map, rather than the Hamming-style tail depth used here); it would be of interest to determine whether their statistic admits an analogous inverse reconstruction theorem, and if so, whether it recovers the same canonical-reduction invariant or a different one.
- **Noisy or sampled histograms.** The present reconstruction is exact and deterministic. A natural probabilistic extension would ask how much of the canonical reduction can be recovered from a *sampled* carry distribution (e.g., the limiting geometric-tail law of §6.6) rather than the exact finite counts.

---

9. Concluding Remarks

We have given a purely combinatorial account of carry propagation in canonical PLRS numeration, built without reference to numerical valuation until Section 5, and shown that carry-depth statistics are governed by a clean nested-tail structure (§6) rather than the naive exact-depth partition one might initially expect. The central result of the paper is the identification of the correct invariant for inverse reconstruction: not the coefficient vector itself, but its canonical reduction (§7), a distinction made necessary by the explicit non-uniqueness example of §7.1 (and illustrated further in §8.2). Every result in Sections 3–7 depends only on results established earlier in the paper, with the dependency structure noted briefly at the start of the relevant sections.