Automaton Minimization and Canonical Reduction for PLRS Numeration

Abstract

Canonical positive linear recurrence numeration systems admit multiple coefficient presentations generating the same legality language. Paper III showed that every such presentation belongs to a unique equivalence class determined by an intrinsic periodic boundary sequence and that each class possesses a distinguished canonical reduction of minimal order.

In this paper we give an automata-theoretic interpretation of canonical reduction. To every coefficient presentation we associate a deterministic threshold-reset automaton accepting exactly its legality language. We prove that the natural automaton of every presentation minimizes to the natural automaton of the canonical reduction. Consequently the minimal deterministic automaton of the legality language is uniquely determined by the canonical reduction, and the canonical order is exactly the number of live Myhill–Nerode classes of the legality language.

Thus canonical reduction is identified with automaton minimization, providing an intrinsic automata-theoretic characterization independent of recurrence order.

---

1. Introduction

Canonical positive linear recurrence numeration systems admit many distinct coefficient presentations generating the same place-value sequence. Paper I established that every equivalence class possesses a unique canonical reduction of minimal order. Paper II identified the maximal boundary word governing greedy representations and carry propagation. Paper III showed that both the canonical reduction and every admissible presentation are recovered from a single intrinsic periodic sequence determined solely by the place-value sequence.

The present paper gives a different interpretation of the same phenomenon.

Instead of studying coefficient vectors or intrinsic sequences directly, we study the recursive legality grammar as a deterministic finite automaton. The recursive legality definition naturally produces a threshold-reset automaton whose states record the current matched prefix of the coefficient vector. Every presentation therefore determines a canonical deterministic automaton accepting its legality language.

The principal result of this paper is that these automata contain exactly the same redundancy as the coefficient presentations themselves.

If two presentations generate the same place-value sequence, then their legality languages coincide by the Language Invariance Theorem of Paper III. We prove that the natural automaton of every presentation minimizes to the natural automaton of the canonical reduction. Consequently canonical reduction admits a purely automata-theoretic description:

$$\boxed{\text{canonical reduction} = \text{minimal threshold-reset presentation of the legality language.}}$$

The automaton construction and the minimality argument use only Papers I and II: the threshold-reset automaton is built from the recursive legality grammar of Paper I, and state distinguishability

(Theorem 5.1) rests on the periodic maximal boundary words of Paper II. The reduction of an arbitrary presentation to the iterated-lift form $\mathbf{c}^{(k)} = P^{k-1}\mathbf{c}_{\min}$ — needed in Section 4 to state the Canonical Quotient Theorem for *every* presentation, not merely the canonical one — uses the presentation classification of Paper III (its Theorem 4.5 and Corollary 4.7). That is still a substantially trimmed dependency compared with treating Paper III's intrinsic sequence as load-bearing throughout; only this one input is imported, and the intrinsic sequence e itself appears nowhere in the proofs, only in the closing interpretive remark identifying $\sigma^r(P^\infty)$ with its shifts.

A natural objection deserves to be addressed before it is raised. Myhill–Nerode minimization is standard, and one might ask why identifying canonical reduction with a minimal DFA is more than an application of a textbook theorem to a particular language. The answer is not that we minimize an automaton — minimization itself is routine once the right automaton and the right quotient are in hand. The contribution is the prior identification: that canonical reduction, previously characterized only by recurrence-order minimality (Paper I) and by the primitive period of an intrinsic algebraic sequence (Paper III), coincides exactly with a standard, general-purpose automata invariant — the number of live Myhill–Nerode classes, and the specific coefficient thresholds recovered from the minimal automaton's transition structure. That two a priori unrelated invariants (recurrence order; DFA state count) turn out to be the same number, for a reason traceable to the periodicity of maximal legal words, is the substance of the paper. The bulk of the actual new mathematics accordingly lies in Section 4, where the periodic quotient structure of a lifted presentation is constructed and identified with the canonical automaton; Section 5's minimality argument, once that quotient is available, follows from Myhill–Nerode by a comparatively short distinguishability argument.

A second objection is more specific: automata for linear numeration systems are themselves classical, and a referee working in this area may ask whether the quotient identified here is already implicit in that literature. It is not, though the distinction is worth stating precisely rather than asserted. Frougny's work on linear and Pisot numeration systems addresses whether *normalization* — converting an arbitrary digit string to greedy form, or computing addition — is realizable by a finite automaton for a fixed presentation; her 1988 paper on order-two systems, the closest in spirit to the present one, fixes a single recurrence $u_{n+2} = au_{n+1} + bu_n$ and asks for which digit-alphabet size the resulting system is "perfect," constructing subsequential transducers for normalization and addition once that alphabet is fixed — but the recurrence order itself is never varied there, so the specific redundancy addressed below (Section 4) does not arise in that setting. Bruyere–Hansel's theory of Bertrand numeration systems characterizes which *sets* of integers are recognizable within a fixed system; Charlier, Rampersad, Rigo, and Waxweiler study the minimal automaton for testing divisibility within one fixed presentation, and separately exhibit a morphism from a general such automaton onto the canonical automaton of the Bertrand system sharing its dominant root. Each of these results moves either within a single fixed presentation or across presentations of *different* numeration systems (distinct dominant roots β). The present paper addresses a different redundancy: distinct coefficient presentations of the *same* place-value sequence (H_n) — the non-uniqueness already exhibited by Paper I's example $\mathbf{c} = (2)$ versus $\mathbf{d} = (1, 2)$, both generating $H_n = 2^{n-1}$ — and shows that their legality automata minimize to one another. This identifies canonical reduction with DFA minimization across the equivalence class of a single sequence, which is not the question the morphism, recognizability, and order-two normalization results above are answering.

The paper is organized as follows.

Section 2 recalls the necessary notation from Papers I and II and defines the threshold-reset automaton.

Section 3 proves that the automaton accepts exactly the recursive legality language.

Section 4 establishes the quotient structure of lifted presentations.

Section 5 proves that the canonical automaton is minimal and derives the Automaton Minimization Theorem.

Section 6 discusses uniqueness, canonical order, and algorithmic reconstruction.

---

2. Threshold-Reset Automata

Throughout the paper let

$$\mathbf{c} = (c_1, \dots, c_L)$$

be a canonical PLRS coefficient vector satisfying

$$c_1 > 0, \qquad c_L > 0,$$

with all coefficients nonnegative integers.

We assume the recursive legality criterion of Paper I together with the associated place-value sequence

$$(H_n)_{n \geq 1}.$$

Only the recursive legality definition and the Exact Periodicity Theorem of Paper II will be required.

2.1 Recursive legality

Recall that a finite digit string is legal precisely when it admits a recursive decomposition into deficient coefficient blocks separated by arbitrary runs of zeros.

Equivalently, every legal block begins by matching the coefficient vector and terminates at the first position where a strictly smaller digit is chosen.

The Restart Property of Paper I then begins a fresh legal block.

This recursive grammar is the starting point of the present paper.

2.2 The threshold-reset automaton

To every coefficient vector

$$\mathbf{c} = (c_1, \dots, c_L)$$

we associate a deterministic finite automaton

$$\mathcal{A}(\mathbf{c}).$$

The automaton has accepting states

$$q_0, q_1, \dots, q_{L-1}$$

together with a single rejecting sink state

$$\perp.$$

The state q_r records that the current block has matched exactly the first r coefficients of $\mathbf{c}$.

The initial state is q_0.

For $0 \le r < L-1$,

$$\delta(q_r, a) = \begin{cases} q_0, & 0 \le a < c_{r+1}, \\ q_{r+1}, & a = c_{r+1}, \\ \perp, & a > c_{r+1}. \end{cases}$$

At the final matching state,

$$\delta(q_{L-1}, a) = \begin{cases} q_0, & 0 \le a < c_L, \\ \perp, & a \ge c_L. \end{cases}$$

The sink satisfies

$$\delta(\perp, a) = \perp$$

for every digit.

We call $\mathcal{A}(\mathbf{c})$ the natural threshold-reset automaton associated with the presentation $\mathbf{c}$.

Remark (alphabet). The digit alphabet is $Sectionigma = \mathbb{Z}_{\ge 0}$, which is infinite. At each state, however, δ is constant on each of at most three intervals of $Sectionigma$ (a low interval returning to q_0, at most one exact threshold value advancing the chain, and a high interval mapping to $\perp$). All finiteness statements in what follows — regularity, minimization, counting of Myhill–Nerode classes — therefore concern the finite state set, not the alphabet, exactly as in the standard treatment of numeration-system automata.

This completes the introduction and Section 2. The next section proves that $\mathcal{A}(\mathbf{c})$ accepts exactly the recursive legality language, establishing the automaton as a faithful reformulation of Paper I's recursive grammar.

---

3. The Legality Language

In this section we prove that the threshold-reset automaton introduced in Section 2 is an exact automata-theoretic realization of the recursive legality grammar of Paper I.

The result is purely structural. No properties of the place-value sequence or the intrinsic sequence are required.

3.1 Accepted Language

Let

$$\mathcal{L}(\mathbf{c})$$

denote the legality language determined recursively by the coefficient vector $\mathbf{c} = (c_1, \ldots, c_L)$, and let $L(\mathcal{A}(\mathbf{c}))$ denote the language accepted by the threshold-reset automaton.

Theorem 3.1 (Legality Automaton Theorem). For every canonical coefficient vector,

$$L(\mathcal{A}(\mathbf{c})) = \mathcal{L}(\mathbf{c}).$$

Proof. The recursive legality predicate of Section 2.3 is defined relative to a matching chain that starts at position 0; write $w \in \mathcal{L}_0$ for this predicate. Reading that same clause with the chain already advanced to position r gives, for each $0 \le r < L$, the predicate $\mathcal{L}_r$: a word w satisfies $\mathcal{L}_r$ if either $w = \varepsilon$, or $w = a\,w'$ where, writing $d_{\max}(r) = c_{r+1}$ if $r < L-1$ and $d_{\max}(r) = c_L - 1$ if $r = L-1$ (the admissible interval at position r is $\{0, \ldots, d_{\max}(r)\}$, by Consecutive Children, Theorem 3.5 of the companion paper):

- if $r < L-1$ and $a = c_{r+1}$ (the *continuing* digit), then $w' \in \mathcal{L}_{r+1}$;
- otherwise, if $a \le d_{\max}(r)$ (a genuine *deficiency*), then $w' \in \mathcal{L}_0$ (Restart Property, Corollary 3.3);
- if $a > d_{\max}(r)$, $w \notin \mathcal{L}_r$.

This is not a new object: $\mathcal{L}_r$ is simply the recursive legality predicate of Section 2.3 relative to q_r, and $\mathcal{L}_0$ recovers the ordinary predicate exactly, since an initial-case word is precisely a chain of continuing digits with no restart triggered before length $L-1$, and every other legal word contains at least one restart. So it suffices to prove, for every r:

$$w \text{ is accepted starting from } q_r \quad \Longleftrightarrow \quad w \in \mathcal{L}_r,$$

by strong induction on $n = |w|$.

Base case $n = 0$. $w = \varepsilon$ is accepted from every q_r, since every q_r is accepting; and $\varepsilon \in \mathcal{L}_r$ by definition. Equivalence holds.

Inductive step. Let $w = a\,w'$, $|w'| = n-1$, and suppose the equivalence holds for all words of length $n-1$ and every position.

If $r < L-1$: by definition of δ, $\delta(q_r, a) = q_0$ for $a < c_{r+1}$, $= q_{r+1}$ for $a = c_{r+1}$, $= \perp$ for $a > c_{r+1}$. Compare digit by digit with $\mathcal{L}_r$: for $a > c_{r+1} = d_{\max}(r)$, $w \notin \mathcal{L}_r$, and $\delta(q_r, a) = \perp$, from which no word is ever accepted (the sink is absorbing and non-accepting) — both sides fail. For $a = c_{r+1}$,

$w \in \mathcal{L}_r$ iff $w' \in \mathcal{L}_{r+1}$, which by the induction hypothesis holds iff w' is accepted from $q_{r+1} = \delta(q_r, a)$ — both sides agree. For $a < c_{r+1}$, $w \in \mathcal{L}_r$ iff $w' \in \mathcal{L}_0$, which by the induction hypothesis holds iff w' is accepted from $q_0 = \delta(q_r, a)$ — both sides agree.

If $r = L-1$: by definition, $\delta(q_{L-1}, a) = q_0$ for $a < c_L$, $= \perp$ for $a \geq c_L$, with no continuing transition — matching exactly the block-final case of Consecutive Children ($d_{\max}(L-1) = c_L - 1$, no continuing digit available). The same digit-by-digit comparison as above, with only the reset and sink clauses present, gives the equivalence.

This closes the induction. Specializing at $r = 0$, and noting $\mathcal{L}_0$ coincides with the recursive legality definition of Section 2.3, gives $L(\mathcal{A}(\mathbf{c})) = \mathcal{L}(\mathbf{c})$. ■

Corollary 3.2. The legality language is regular.

Proof. The threshold-reset automaton is finite. Therefore its accepted language is regular. By Theorem 3.1 this language is precisely $\mathcal{L}(\mathbf{c})$. ■

3.2 Residual Languages

The automaton viewpoint allows legality to be studied through residual languages.

For each live state q_r, define

$$R_r = q_r^{-1}\mathcal{L} = \{\, x : \delta(q_r, x) \text{ is accepting} \,\}.$$

Thus R_r consists of all continuations that remain legal after the first r coefficients of the current block have already been matched.

The languages $R_0, \ldots, R_{L-1}$ are the residual languages of the live states of $\mathcal{A}(\mathbf{c})$. Whether distinct states can share a residual language — and hence whether these residual languages represent distinct Myhill–Nerode classes — is addressed in Sections 4 and 5; for a lifted presentation ($L > \ell$) some of them coincide, and identifying exactly which is the content of the quotient construction below.

3.3 Maximal Continuations

Paper II proves that greedy maximal words are eventually periodic. The automaton provides a local version of that result.

For each live state q_r, let $M_n(q_r)$ denote the lexicographically maximal accepted continuation of length n beginning from q_r, constructed by the greedy rule: at each step, append the largest digit admissible at the current state.

That this greedy construction genuinely produces the lexicographic maximum, and that the resulting family is compatible across lengths — so that an infinite limit $M_\infty(q_r)$ is actually well-defined, rather than merely a family of separately-maximal finite words — requires an explicit argument. Finite-length maximality of the greedy construction is already established: it is precisely the Maximal Extension Theorem of Paper I, applied at the state q_r rather than at an explicit prefix, which is licensed by Consecutive Children (Theorem 3.5 of Paper I), since the admissible digit interval there is shown to depend only on the chain-position r and not on which legal prefix reached it — exactly the determinism the automaton is built to express. What Paper I and Paper II do not state for a general starting state q_r is the prefix-compatibility across lengths needed to pass to an infinite limit. We supply this directly.

Lemma 3.3 (Greedy Well-Definedness). For every live state q_r of $\mathcal{A}(\mathbf{c})$:

(i) *(Unconditional completability.)* Every live state admits an accepted continuation of every length $n \geq 0$.

(ii) *(Prefix compatibility.)* $M_n(q_r)$ is the length-n prefix of $M_{n+1}(q_r)$, for every $n \geq 0$. Consequently the family $(M_n(q_r))_{n\geq 0}$ determines a well-defined infinite word $M_\infty(q_r)$.

Proof. (i) The digit 0 always lies in the admissible interval at q_r (the interval is $\{0, \ldots, d_{\max}(r)\}$ with $d_{\max}(r) \geq 0$, by Consecutive Children), and $\delta(q_r, 0) \in \{q_0, q_{r+1}\}$ in every case (the "advance" clause fires exactly when the threshold itself is 0, in which case 0 is the continuing digit; otherwise 0 is a genuine deficiency and resets to q_0). Either target is again a live state. Hence reading an arbitrary run of zeros from q_r remains live at every step, giving an accepted word of any prescribed length n.

(ii) Let $w = M_{n+1}(q_r)$ and let p be its length-n prefix. Since the residual language of every reachable state is prefix-closed (inherited directly from prefix-closedness of $\mathcal{L}(\mathbf{c})$ itself, Corollary 3.2 of the companion paper, via the determinism of δ), p is accepted from q_r. Suppose $p \neq M_n(q_r)$; then some z of length n, accepted from q_r, satisfies $z >_{\text{lex}} p$. By part (i), z extends to an accepted word zb of length $n+1$ for some digit b. Since z and p already differ, and $z >_{\text{lex}} p$, we get $zb >_{\text{lex}} pa = w$ for every digit a — in particular for the digit a with $w = pa\,(\cdots)$, contradicting the maximality of $w = M_{n+1}(q_r)$ among accepted words of length $n+1$. Hence $p = M_n(q_r)$. ∎

With Lemma 3.3 in hand, $M_\infty(q_r)$ is genuinely well-defined, and by construction $M_n(q_r)$ is its length-n prefix for every n.

Proposition 3.4 (Shifted Maximal Continuations). Let $P_{\mathbf{c}} = c_1 c_2 \cdots c_{L-1}(c_L - 1)$ denote the boundary block of the given presentation $\mathbf{c}$, of length L, and let $P = p_1 \cdots p_\ell$ denote the primitive maximal boundary block (of length $\ell \leq L$), so that $P_{\mathbf{c}}^\infty = P^\infty$ by the Exact Periodicity Theorem of Paper II. Then, for every live state q_r of $\mathcal{A}(\mathbf{c})$,

$$M_\infty(q_r) = \sigma^r(P^\infty)\,, \qquad 0 \leq r < L,$$

where σ denotes cyclic shift. (Since P^∞ has period ℓ, the right-hand side depends on r only through $r \bmod \ell$.)

Proof. Beginning from q_r, the greedy rule always chooses the largest admissible digit. At every interior matching position ($r < L-1$) the largest admissible digit is the threshold digit c_{r+1} itself (Consecutive Children, Paper I), advancing the automaton one step along the matching chain. At the final position of the block ($r = L-1$) the largest admissible digit is the deficient digit $c_L - 1$, which returns the automaton to q_0. Thus, starting from q_0, the greedy computation traverses the block $P_{\mathbf{c}} = c_1 \cdots c_{L-1}(c_L - 1)$ once and then repeats it indefinitely, so $M_\infty(q_0) = P_{\mathbf{c}}^\infty$. Starting from q_r instead simply omits the first r symbols of this same repeating computation, giving $M_\infty(q_r) = \sigma^r(P_{\mathbf{c}}^\infty)$. Since $P_{\mathbf{c}}^\infty = P^\infty$, the claim follows. ∎

Section 3 establishes the automaton as an exact reformulation of the recursive legality grammar, secures the well-definedness of the greedy infinite continuation from every live state (Lemma 3.3), and identifies that continuation explicitly (Proposition 3.4), for states of an arbitrary (possibly lifted) presentation. These continuations provide the key invariant used in the minimization argument of the next section.

4. The Canonical Quotient

The preceding section showed that the threshold-reset automaton is an exact automata-theoretic realization of the recursive legality grammar. We now compare the automata associated with different presentations of the same place-value sequence.

Throughout this section let

$$P = p_1 p_2 \cdots p_\ell$$

denote the primitive maximal boundary block of the canonical reduction, so that every presentation of the same place-value sequence has the form

$$\mathbf{c}^{(k)} = P^{k-1}\mathbf{c}_{\min},$$

where $\mathbf{c}_{\min}$ is the canonical reduction and $k \geq 1$. This is the one point in the paper at which a result from Paper III is used rather than merely cited for interpretation: the classification of presentations as periods of the intrinsic sequence (Paper III, Theorem 4.5) and the resulting iterated-lift form of every non-minimal presentation (Paper III, Corollary 4.7) together justify this equation. Everything from here through Section 5 uses only this classification fact, together with Papers I and II.

The natural automaton of $\mathbf{c}^{(k)}$ therefore contains $k\ell$ live states.

Our goal is to show that all copies of the primitive period collapse to a single copy under automaton minimization. This is the substantive new construction of the paper: the periodic congruence identified below (Theorem 4.2) and its identification with the canonical automaton (Theorem 4.4) are what give the abstract minimization machinery of Section 5 something specific — and previously unknown — to act on. Once the quotient is in hand, minimality is comparatively routine.

4.1 The Periodic Transition Structure

Proposition 4.1. Let $q_0, q_1, \ldots, q_{k\ell-1}$ be the live states of the natural automaton of $\mathbf{c}^{(k)}$. For every $0 \leq r < \ell - 1$ and every $i \equiv r \pmod{\ell}$ with $q_i \neq q_{k\ell-1}$ (i.e. q_i is not the last live state of the automaton), the threshold at q_i is p_{r+1}. The last live state, $q_{k\ell-1}$, has threshold $p_\ell + 1$ regardless of its residue.

Proof. The lifted presentation consists of $k - 1$ repeated copies of the primitive boundary block $P = p_1 \cdots p_\ell$, followed by the canonical reduction $\mathbf{c}_{\min}$ itself. Every state belonging to one of the first $k - 1$ copies, at position r within its copy, therefore has threshold p_{r+1} directly. Within the final copy — the states of $\mathbf{c}_{\min}$ — every threshold except the last agrees with p_{r+1} as well, since $\mathbf{c}_{\min} = (p_1, \ldots, p_{\ell-1}, p_\ell + 1)$ by construction (Paper III, Corollary 4.6). Only the truly last state, $q_{k\ell-1}$, departs from this pattern, carrying threshold $p_\ell + 1$ instead of p_ℓ. ■

Note that Proposition 4.1 does *not* assert that the transition function of $\mathcal{A}(\mathbf{c}^{(k)})$ is itself periodic with period ℓ: the literal transition rule at $q_{k\ell-1}$ differs from the transition rule at every other state of the same residue class $\ell - 1$. What is true, and what is needed for the minimization argument, is the weaker fact that these differing rules still send every digit into the *same* equivalence class — this is established next.

4.2 The Periodic Congruence

Motivated by Proposition 4.1, define an equivalence relation on the live states by

$$q_i \equiv q_j \quad \Longleftrightarrow \quad i \equiv j \pmod{\ell}.$$

The rejecting sink forms a singleton equivalence class.

Theorem 4.2 (Periodic Quotient). The congruence modulo ℓ is a right congruence of the threshold-reset automaton.

Proof. Suppose $q_i \equiv q_j$, say $i = j'\ell + r$ and $j = j''\ell + r$ for the common residue r, $0 \le r < \ell$. We must show that for every digit a, $\delta(q_i, a) \equiv \delta(q_j, a)$.

Case $r < \ell - 1$, neither q_i nor q_j is the final live state. By Proposition 4.1 both states share threshold p_{r+1}. For $a < p_{r+1}$, both transitions reset to q_0, and $q_0 \equiv q_0$. For $a = p_{r+1}$, both transitions advance: $q_i \mapsto q_{j'\ell+r+1}$ and $q_j \mapsto q_{j''\ell+r+1}$, and since $(j'\ell + r + 1) \equiv (j''\ell + r + 1) \equiv r + 1 \pmod{\ell}$, these targets are congruent. For $a > p_{r+1}$, both transitions enter $\perp \equiv \perp$.

Case $r = \ell - 1$, neither state is the final live state. By Proposition 4.1 both thresholds equal p_ℓ. For $a < p_\ell$, both reset to $q_0 \equiv q_0$. For $a = p_\ell$, both advance to the start of the next copy of the block: $q_i \mapsto q_{(j'+1)\ell}$ and $q_j \mapsto q_{(j''+1)\ell}$; since $(j' + 1)\ell \equiv (j'' + 1)\ell \equiv 0 \pmod{\ell}$, these targets are both congruent to q_0, matching the reset target of the interior case exactly. For $a > p_\ell$, both enter $\perp$.

Case one of q_i, q_j is the final live state $q_{k\ell-1}$. Then necessarily $r = \ell - 1$ and, since there is only one state at index $k\ell - 1$, we must have (say) $j = k\ell - 1$ and $i < j$ an interior state of residue $\ell - 1$. At q_i, threshold p_ℓ: digits $a < p_\ell$ reset to q_0, digit $a = p_\ell$ advances to $q_{(j'+1)\ell} \equiv q_0$, digits $a > p_\ell$ enter $\perp$. At $q_j = q_{k\ell-1}$, threshold $p_\ell + 1$: digits $a < p_\ell + 1$ — that is, $a \le p_\ell$ — reset to q_0, and digits $a \ge p_\ell + 1$ enter $\perp$. Comparing digit by digit: for $a < p_\ell$ both go to q_0; for $a = p_\ell$ — the single digit at which the two states invoke different clauses of δ, advance for q_i versus reset for q_j — both nonetheless land in the class of q_0; for $a > p_\ell$ both enter $\perp$. The three ranges $\{a < p_\ell\}$, $\{a = p_\ell\}$, $\{a > p_\ell\}$ exhaust $\mathbb{Z}_{\ge 0}$ with no overlap, so this is a complete case analysis, and the transitions agree modulo $\equiv$ at every digit.

In every case, corresponding transitions enter equivalent states. Hence congruence modulo ℓ is preserved by every transition, i.e. it is a right congruence. ■

Corollary 4.3. The quotient automaton has precisely ℓ live states together with one rejecting sink.

4.3 Identification of the Quotient

The quotient automaton is not merely smaller; it is exactly the natural automaton associated with the canonical reduction.

Theorem 4.4. The quotient of the natural automaton of $\mathbf{c}^{(k)}$ under congruence modulo ℓ is isomorphic to the natural threshold-reset automaton of the canonical reduction.

Proof. The quotient possesses one live state for each residue class $[q_0], \ldots, [q_{\ell-1}]$. By Proposition 4.1 and Theorem 4.2, the threshold at residue class r is precisely p_{r+1} for $0 \le r < \ell - 1$ — every representative of this class other than possibly the final state shares this threshold, and where the final state differs (Theorem 4.2's third case), its transitions still agree with the class's transitions after passing to the quotient. The final residue class $[q_{\ell-1}]$ has, via its representative $q_{k\ell-1}$, terminal threshold $p_\ell + 1$, exactly as required for the canonical reduction.

Consequently the quotient transitions satisfy

$$\delta([q_r], a) = \begin{cases} [q_0], & a < p_{r+1}, \\ [q_{r+1}], & a = p_{r+1}, \\ \perp, & a > p_{r+1}, \end{cases}$$

for $r < \ell - 1$, and

$$\delta([q_{\ell-1}], a) = \begin{cases} [q_0], & a < p_\ell + 1, \\ \perp, & a \geq p_\ell + 1. \end{cases}$$

These are precisely the transition rules defining the threshold-reset automaton of the canonical reduction. Hence the quotient automaton is isomorphic to the canonical automaton. ■

Section 4 identifies the natural quotient arising from any lifted presentation. The remaining task is to prove that the quotient is already minimal. This will be accomplished in the next section by showing that the live states of the canonical automaton have pairwise distinct residual languages.

---

5. Automaton Minimization

Section 4 showed that every natural threshold-reset automaton admits a canonical quotient indexed by the primitive period of the maximal boundary word. It remains to prove that this quotient is already minimal. Given the quotient, this is the standard half of the argument: distinguishability plus Myhill–Nerode. What makes the resulting Automaton Minimization Theorem (Theorem 5.3) more than a routine invocation of that theorem is that Section 4 supplied the specific quotient — with no independent motivation for it beyond the PLRS structure itself, it is not something Myhill–Nerode could have handed us.

The proof separates naturally into two parts. First, we show that the live states of the canonical automaton are pairwise distinguishable in the sense of Myhill–Nerode. Second, we combine this with the quotient constructed in Section 4 to obtain the minimal deterministic automaton.

5.1 State Distinguishability

Let $\mathcal{A}_{\min}$ denote the threshold-reset automaton of the canonical reduction, whose primitive boundary block has length ℓ. Its live states are $q_0, q_1, \ldots, q_{\ell-1}$. For each state q_r, let $R_r = q_r^{-1}\mathcal{L}$ denote its residual language.

The key observation is that equality of residual languages immediately forces equality of their greedy maximal continuations.

Theorem 5.1 (State Distinguishability). The live states of the canonical threshold-reset automaton have pairwise distinct residual languages.

Proof. Suppose $R_r = R_s$ for some $0 \le r < s < \ell$. Since the residual languages are identical sets, their lexicographically maximal words of every finite length are identical: $M_n(q_r) = M_n(q_s)$ for every $n \ge 1$.

By Lemma 3.3(ii), $M_n(q_r)$ is the length-n prefix of $M_\infty(q_r)$ for every n, and likewise for q_s. Since $M_n(q_r) = M_n(q_s)$ for every n, the two infinite continuations agree prefix by prefix, so

$$M_\infty(q_r) = M_\infty(q_s).$$

By Proposition 3.4,

$$M_\infty(q_r) = \sigma^r(P^\infty), \qquad M_\infty(q_s) = \sigma^s(P^\infty),$$

where P is the primitive maximal boundary block of length ℓ. Consequently $\sigma^r(P^\infty) = \sigma^s(P^\infty)$. Thus $s - r$ is a period of P^∞. Since $0 < s - r < \ell$, this contradicts the primitivity of P. Therefore $R_r \neq R_s$ whenever $r \neq s$.

The rejecting sink is distinguished immediately, since its residual language is empty, whereas every live state accepts the empty continuation (every q_r is accepting, by construction, and $\mathcal{L}$ is prefix-closed, so this is not merely consistent but forced).

Hence every state of the canonical automaton belongs to a distinct Myhill–Nerode class. ■

Corollary 5.2. The canonical threshold-reset automaton is minimal.

Proof. By Theorem 5.1 every pair of distinct states is distinguishable. Therefore no two states may be identified by DFA minimization. Hence the automaton is minimal. ■

Remark (a pathological case). Theorem 5.1 asserts only that $M_\infty(q_r)$ and $M_\infty(q_s)$ eventually differ, not that they differ quickly, and this is essential rather than a slack corner of the proof: the presentation $\mathbf{c} = (3, 0, 0, 2)$, with two consecutive interior zero-thresholds, has $M_\infty(q_1) = 0, 0, 1, 3, \ldots$ and $M_\infty(q_2) = 0, 1, 3, 0, \ldots$ agreeing at their first digit and diverging only at the second. An argument that tried to distinguish states by a bounded number of leading digits would fail on this example; the full infinite-continuation-plus-primitivity argument used above does not, since it makes no claim about how soon the divergence occurs.

5.2 The Automaton Minimization Theorem

We now combine Sections 4 and 5. Section 4 produced a quotient of every lifted automaton. Section 5 shows that this quotient is already minimal. The principal theorem follows immediately.

Theorem 5.3 (Automaton Minimization Theorem). Let $\mathbf{c}$ be any presentation of a canonical PLRS. Then the minimal deterministic automaton accepting its legality language is isomorphic to the threshold-reset automaton of the canonical reduction.

Proof. By Theorem 4.4, the natural automaton of $\mathbf{c}$ admits a quotient isomorphic to the threshold-reset automaton of the canonical reduction. By Corollary 5.2 that quotient automaton is already minimal. Therefore minimizing the natural automaton of any presentation yields precisely the canonical threshold-reset automaton. ■

5.3 Uniqueness

Corollary 5.4 (Uniqueness). The threshold-reset automaton of the canonical reduction is, up to isomorphism, the unique minimal deterministic automaton accepting the legality language.

Proof. Every state q_r of $\mathcal{A}_{\min}$ is reachable from q_0 (via the prefix $p_1 \cdots p_r$), so $\mathcal{A}_{\min}$ is trim. Every regular language possesses a unique minimal (trim) deterministic automaton up to isomorphism, by the Myhill–Nerode theorem. By Theorem 5.3 the canonical threshold-reset automaton is minimal. Hence every minimal deterministic automaton accepting the legality language is isomorphic to it. ■

Section 5 is the conceptual heart of the paper. The automata associated with different coefficient presentations are no longer merely equivalent recognizers of the same language. They all minimize to one distinguished automaton, namely the threshold-reset automaton of the canonical reduction. Thus canonical reduction acquires an intrinsic automata-theoretic interpretation independent of recurrence order or the intrinsic sequence developed in Paper III.

---

6. Consequences and Reconstruction

The Automaton Minimization Theorem identifies canonical reduction with the unique minimal deterministic realization of the legality language. In this section we record several immediate consequences. None of these requires additional mathematical machinery: they are direct applications of the preceding theorem together with standard automata theory.

6.1 Canonical Order

Recall that a state of a deterministic automaton is live if some continuation from that state is accepted. Since legality languages are prefix-closed, every live state of the threshold-reset automaton is accepting. The unique non-live state is the rejecting sink.

Corollary 6.1. The order of the canonical reduction equals the number of live Myhill–Nerode equivalence classes of the legality language.

Proof. Let $\mathbf{c}_{\min}$ have order ℓ. By construction, its threshold-reset automaton possesses exactly ℓ live states together with one rejecting sink. By Theorem 5.1 these live states are pairwise distinguishable, so each represents a distinct Myhill–Nerode class. By Corollary 5.4 this automaton is the unique minimal DFA. Therefore the minimal DFA possesses exactly ℓ live equivalence classes. Since ℓ is precisely the order of the canonical reduction, the result follows. ■

Paper I defined canonical order by recurrence minimization. The present paper identifies exactly the same quantity as the number of live residual-language classes:

$$\boxed{\text{canonical order} = \text{number of live Myhill–Nerode classes.}}$$

6.2 Canonical Reduction via Automaton Minimization

Corollary 6.2. The canonical reduction is uniquely determined by the minimal deterministic automaton of the legality language.

Proof. Let M be any minimal deterministic automaton accepting $\mathcal{L}(\mathbf{c})$, presented abstractly (i.e. with no a priori labeling of its states as $q_0, \ldots, q_{\ell-1}$). By Corollary 5.4, M is isomorphic to $\mathcal{A}_{\min}$, the canonical threshold-reset automaton. Moreover this isomorphism $\varphi : M \to \mathcal{A}_{\min}$ is *unique*: any language-preserving isomorphism must send each state s of M to a state of $\mathcal{A}_{\min}$ with the same residual language as s, and by Theorem 5.1 the states of $\mathcal{A}_{\min}$ have pairwise distinct residual languages, so at most one target state is possible for each s. This answers directly how the matching chain is to be located inside an arbitrary minimal DFA: it is not chosen, but forced, by matching residual languages against the known canonical automaton.

Fix this unique φ. Under it, the live states of M inherit the labels $q_0, \ldots, q_{\ell-1}$ in the canonical order of Section 2 (i.e. φ sends the initial state of M to q_0, since isomorphisms preserve initial states, and the ordering along the matching chain follows by iterating δ), and the sink inherits $\perp$.

Reconstruction now proceeds on this canonically labeled copy exactly as in the threshold-reset construction itself. For $0 \le r < \ell - 1$, the coefficient c_{r+1} is the unique digit at q_r whose transition is the *advancing* one, leaving the reset/sink pattern and moving to q_{r+1} — unique by construction of the threshold-reset rule (Section 2). At the terminal state $q_{\ell-1}$, no advancing transition exists; c_ℓ is instead recovered as one more than the largest digit still returning to q_0, i.e. $c_\ell = p_\ell + 1$ where p_ℓ is the greatest digit resetting from $q_{\ell-1}$.

Together these recover the full vector $(c_1, \ldots, c_\ell)$, exactly the canonical reduction, and the recovery is canonical in the strict sense that it depended on no choice: φ was unique, and every subsequent step read data directly off it. ■

Thus canonical reduction no longer requires recurrence minimization or intrinsic-sequence reconstruction. It is encoded directly in the unique minimal deterministic grammar of the legality language. This yields the principal conceptual identity of the paper:

$$\boxed{\text{canonical reduction} = \text{minimal threshold-reset presentation of the legality language.}}$$

6.3 Reconstruction Algorithm

Algorithm. Given the natural threshold-reset automaton of any presentation $\mathbf{c}$ of order L:

1. Minimize it (any standard DFA minimization algorithm).
2. Identify the sink, and enumerate the remaining live states along the matching chain from the initial state.
3. At each live state but the last, record the digit advancing to the next live state; at the last, take one more than the largest digit resetting to the initial state.

The resulting vector is the canonical reduction, independently of L: correctness is immediate from Theorem 5.3 and Corollary 6.2, since every presentation minimizes to the same canonical automaton regardless of how many elementary lifts of Paper III it carries. Each state's transition function is constant on at most three intervals of the digit alphabet (the alphabet remark, Section 2.2), which is what makes minimization well-defined and finite in the first place — but this per-state property does not by itself bound the *global* number of distinct digit classes across all L states, which may be as large as $O(L)$, so quoting a specific runtime such as $O(L \log L)$ would require a dedicated symbolic-automaton minimization argument not given here. What can be said unconditionally is: standard DFA minimization, applied to the finite threshold partition induced by the coefficients, recovers the canonical reduction effectively.

6.4 Relationship with Previous Papers

The three preceding papers identify canonical reduction from three distinct viewpoints.

Paper I characterizes it as the unique presentation of minimal recurrence order.

Paper III characterizes it as the coefficient vector determined by the primitive period of the intrinsic boundary sequence.

The present paper identifies exactly the same object as the unique minimal deterministic grammar of the legality language.

These three characterizations are equivalent:

$$\boxed{\begin{aligned} \text{minimal recurrence} &= \text{primitive boundary period} \\ &= \text{minimal threshold-reset automaton.} \end{aligned}}$$

Each arises naturally from a different mathematical structure: recurrence theory; combinatorial boundary dynamics; deterministic automata.

This equivalence completes the automata-theoretic interpretation of canonical reduction.

---

7. Discussion

The Automaton Minimization Theorem provides an interpretation of canonical reduction that does not require the reader to already understand the intrinsic sequence of Paper III — the automaton, the quotient, and the minimality argument can all be followed without it. As the corrected dependency account below makes precise, this is not the same as full independence: extending the theorem to cover every presentation, not merely the canonical one, imports one classification fact (Paper III's Theorem 4.5 and Corollary 4.7) whose own proof does rely on that algebraic construction.

The automaton and the minimality argument depend only upon the recursive legality grammar of Paper I and the periodic maximal boundary words of Paper II — state distinguishability in particular (Theorem 5.1) needs nothing beyond these two. The Canonical Quotient Theorem, in the form covering every presentation rather than only the canonical one, additionally uses Paper III's presentation classification to justify the iterated-lift form $\mathbf{c}^{(k)} = P^{k-1}\mathbf{c}_{\min}$ from which the quotient is built. Beyond that one input, the intrinsic sequence introduced in Paper III is not required: it appears only in the closing interpretation, where Paper III identifies the periodic maximal continuations appearing here with the shifts of the intrinsic sequence, providing an independent reading of the same automaton — by Paper III, the periodic maximal continuation $\sigma^r(P^\infty)$ identified in Proposition 3.4 is precisely the r-shift of the intrinsic sequence e.

The threshold-reset automaton therefore forms a bridge between recurrence theory and automata theory. The recursive legality definition, originally introduced to characterize legal digit strings, naturally generates a deterministic automaton whose minimal form is precisely the canonical reduction.

The paper also clarifies the role of non-canonical presentations. Elementary lifts increase the size of the natural automaton by introducing additional copies of the primitive boundary block. These extra states carry no new residual languages. Automaton minimization removes exactly this redundancy, collapsing every lifted presentation to the canonical automaton.

The broader problem of characterizing all regular languages that arise as legality languages remains open. The threshold-reset construction gives a necessary automaton model for canonical PLRS presentations, but a complete intrinsic characterization of the image of this construction among arbitrary regular languages appears to require additional arithmetic conditions beyond automata theory alone.

Thus the present work should be viewed not as a complete classification of legality languages, but as an automata-theoretic characterization of canonical reduction.

It complements the previous three papers by showing that recurrence minimization, intrinsic boundary periodicity, and deterministic automaton minimization are three manifestations of the same underlying structure. This identification provides a natural endpoint to the development of canonical PLRS presentations while opening a direct connection with classical automata theory.

References

[Paper I] *Canonical Reduction and Inverse Reconstruction in PLRS Numeration.*

[Paper II] *Boundary Maps and Carry Costs in Canonical Positive Linear Recurrence Numeration Systems.*

[Paper III] *Intrinsic Boundary Sequences, Canonical Presentations, and Language Invariance in Canonical Positive Linear Recurrence Numeration.*

C. Frougny, *Linear numeration systems of order two*, Information and Computation 77 (1988), 233–259.

C. Frougny, *Linear numeration systems, θ-developments and finite automata*, in STACS 89, Lecture Notes in Computer Science 349, Springer, 1989, 144–155.

C. Frougny, *Representations of numbers and finite automata*, Mathematical Systems Theory 25 (1992), 37–60.

V. Bruyere and G. Hansel, *Bertrand numeration systems and recognizability*, Theoretical Computer Science 181 (1997), 17–43.

E'. Charlier, N. Rampersad, M. Rigo, and L. Waxweiler, *The minimal automaton recognizing* $m\mathbb{N}$ *in a linear numeration system*, INTEGERS 11B (2011), Paper A4.