hexagon:2611.00554v1

Cyclotomic Weight Chambers for Reverse-Complement-Invariant Minimum Decycling Sets

Oleksiy Babanskyy

Abstract

For DNA order eight, the fixed labelled weights $(s(A), s(C), s(G), s(T)) = (-\rho, -1, 1, \rho)$ in $\mathcal{W}$, $\rho > 1$, with fixed $A < C < G < T$ ties meet 21 ordinary complete-move components. Their entire union has optimum residual path length 141, including non-RC endpoints. A certified reverse-complement-invariant minimum decycling set of 8230 words instead attains 107 and is optimal in a different component. The atlas classifies $\mathcal{W}$, not all labelled antisymmetric weight maps. Certified 8232-word RC feedback sets attain length 105; the global cardinality problem remains $8230 \leq \tau_{\mathrm{RC}}(106) \leq 8232$.

A two-family selector determines equality strata, and an exact atlas has 86 walls, 173 strata and 114 terminals. Boolean classical retiming optimizes entire components; prescribed RC mate sums give two paired-move sectors, each costing at most one additional path unit. Beyond a component, conditional stable completion is exactly 2-SAT; bounded coordinates and antisymmetric heights describe endpoint and cardinality problems. Balanced closed-walk loads have an exact signed-real finite-pool closure. A certified exterior passes one such relaxation but requires at least 8235 selections in its stated real-gradient class at bound 106, without a global lower-bound claim.

Classical cyclotomic tensor relations and bounded signed-one-hot realization also give an all-order Farey spectrum of rational DNA embedding resonances. A real algebraic ratio is positionally realizable at some even order exactly when it belongs to a cyclotomic field. Complete arithmetic proofs and finite local exclusions are retained in the technical supplement.

1 Problem and main results

Minimum decycling sets (MDS) remove all directed cycles from a de Bruijn graph at minimum cardinality. They support sequence sketching; the ordinary minimum follows from Mykkeltveit's construction and the unavoidable-set formulation of Champarnaud, Hansel and Perrin [20, 5]. Recent work studies residual paths, move components and reverse-complement-aware sketches [24, 21, 17, 18].

The imported parity construction [2] admits arbitrary injective, nonzero, antisymmetric real weights. Our complete DNA atlas and its 141 bound concern precisely the fixed labelled family

$$\mathcal{W} = \{(s(A), s(C), s(G), s(T)) = (-\rho, -1, 1, \rho) : \rho > 1\},$$

with the fixed lexicographic tie order $A < C < G < T$. The general labelled weight domain is larger and is not classified here. Length L counts residual arcs in the lower graph $B(q, k-1)$; Section 2.1 specifies the equivalent upper-graph and avoiding-word conventions.

Theorem 1.1 (Fixed-family separation at DNA order eight)**.** *With these labels and ties, the terminals of $\mathcal{W}$ meet exactly 21 ordinary complete-move components. The minimum*

of L over the union of these entire components is 141, including non-RC endpoints. A reverse-complement-invariant MDS with 8230 words instead has $L = 107$, lies in a different component, and attains that component's optimum. There are also RC feedback sets with 8232 words and $L = 105$. If $\tau_{\mathrm{RC}}(106)$ is the minimum cardinality of an RC feedback set at residual path bound 106, then

$$8230 \le \tau_{\mathrm{RC}}(106) \le 8232.$$

Neither the global minimum length of an RC MDS nor the exact value of $\tau_{\mathrm{RC}}(106)$ is determined.

Proof. The atlas and component coverage are Proposition 3.9; the complete-component optima are Proposition 4.2. The 107 witness and its component certificate are Theorem 4.4. The cardinality statements and their evidence are developed in Proposition 5.11. Since 107 is below 141, its component is outside the union. □

Domain	Cardinality	Certified residual length
Union of 21 components from $\mathcal{W}$	8230	optimum 141
Separate RC MDS component	8230	optimum 107
RC feedback sets	8232	attaining 105
Global RC problem at bound 106	8230–8232	exact cost open

Thus varying ρ in $\mathcal{W}$ and then applying complete moves and inverses cannot reach the 107 witness. This does not exclude all other antisymmetric weight assignments.

Why the labels matter. At $k = 4$, keep the same ties and compare $s_0 = (-2, -1, 1, 2)$ with $s_1 = (-2, 1, -1, 2)$. On the simple directed cycle

$$\texttt{AAAC} \to \texttt{AACA} \to \texttt{ACAG} \to \texttt{CAGA} \to \texttt{AGAA} \to \texttt{GAAA} \to \texttt{AAAC},$$

the first terminal selects only `AGAA`; the second selects `ACAG` and `AGAA`. Cycle sums of a gradient telescope, so these terminals belong to different ordinary complete-move components. A weight relabelling cannot silently be absorbed into $\mathcal{W}$ while the ties stay fixed.

Contribution and mathematical dependencies. The chain is weights, selector, chambers, components, escape from the fixed-family union, and limits of exterior relaxations. The two-family selector includes equality strata. Cycle-signature component invariance is due to Marçais, DeBlasio and Kingsford [17]; our component optimization specializes classical retiming [15]. The two RC sectors use the tight-edge SCC criterion for UTVPI systems [13], with Boolean endpoint reconstruction and legal paired moves. Stable completion specializes classical 2-SAT [1]. Finite-pool closure specializes classical LP duality [4] and yields signed real lifts, rather than Boolean endpoints. Polynomial complexity concerns the explicitly represented graph, whose size is exponential in word order.

The arithmetic results use classical cyclotomic tensor relations [19, 8] and bounded signed-one-hot realization to obtain an all-order rational Farey spectrum and eventual algebraic realization. These are positional embedding statements: raw equalities need not be active selector walls. We use exact algebraic sign evaluation [9] and realizable sign arrangements [3]. General root-of-unity relations provide context [16, 7, 14, 22, 12, 6, 11]; these targeted comparisons do not certify exhaustive novelty.

Reading path. Section 2 fixes the construction and chambers; Section 3 gives retiming and RC sectors; Section 4 proves the finite separation. Sections 5 and 6 develop completions, heights and balanced loads. Section 7 states the arithmetic interfaces, with full proofs and ancillary counts in the technical supplement. Section 8 records limitations and evidence access. Earlier escapes and finite local exclusions remain in the same complete PDF.

2 Construction, the two-family selector, and finite chambers

2.1 Imported construction and normalized coordinates

Let $\mathcal{A}$ be an alphabet of cardinality $q = 2m$, with a fixed-point-free involution $a \mapsto \bar{a}$. For a word $w = w_0 \cdots w_{k-1}$, define

$$\mathrm{RC}(w)_j = \overline{w_{k-1-j}}.$$

The directed vertex de Bruijn graph $B(q,k)$ has vertex set $\mathcal{A}^k$ and an arc $u \to v$ when $u_1 \cdots u_{k-1} = v_0 \cdots v_{k-2}$. A decycling set is a set of vertices whose deletion leaves a directed acyclic graph. An ordinary minimum decycling set is abbreviated to MDS.

Rotation $R(w_0 \cdots w_{k-1}) = w_1 \cdots w_{k-1} w_0$ partitions $\mathcal{A}^k$ into pure cycling registers (PCRs). Their number is the necklace count

$$N_q(k) = \frac{1}{k} \sum_{d \mid k} \varphi(d) q^{k/d}. \tag{1}$$

Every MDS has size $N_q(k)$ and contains exactly one distinct representative from each PCR [20, 5]. The converse is false: a one-per-PCR set can leave a cycle that traverses multiple PCRs.

We use the line-graph representation. A length-k word is the labelled arc

$$\mathrm{tail}(w) = w_0 \cdots w_{k-2} \longrightarrow \mathrm{head}(w) = w_1 \cdots w_{k-1}$$

of $B(q, k-1)$. For an integer function f on lower states, set

$$(\nabla f)(w) = f(\mathrm{tail}(w)) - f(\mathrm{head}(w)). \tag{2}$$

Let $k \geq 4$ be even. Choose distinct positive magnitudes $\beta_1, \ldots, \beta_m$ and label each complement pair so that

$$s(a_i) = \beta_i, \qquad s(\bar{a}_i) = -\beta_i.$$

Thus s is real, injective, nonzero, and antisymmetric. Put $\zeta_k = e^{2\pi i/k}$ and

$$P_s(w) = \sum_{j=0}^{k-1} s(w_j) \zeta_k^{j+1}, \tag{3}$$

$$W_s(w) = \sum_{j=0}^{k-1} s(w_j), \tag{4}$$

$$\lambda_s(u) = \sum_{j=0}^{k-2} s(u_j) \sin\left(\frac{2\pi(j+1)}{k}\right). \tag{5}$$

The weighted seed M_0 selects, on every nonzero embedded PCR, the negative real vertex if present and otherwise the unique strict lower-to-upper crossing. On a zero-embedding PCR it uses the fixed deterministic reverse-complement-equivariant tie rule of [2].

For a lower word u, define

$$\psi_s(u) = \sum_{j=0}^{k-2} s(u_j)(\zeta_k^{j+1} - 1), \tag{6}$$

$$r_s(u) = \operatorname{Re} \psi_s(u), \tag{7}$$

$$U_s = \{u : \lambda_s(u) > 0\}, \qquad \vartheta(u) = \mathrm{RC}(u). \tag{8}$$

Fix a lexicographic order on lower words and put

$$\begin{aligned} a_s(u) = 1 \quad \Longleftrightarrow \quad & u \in U_s \text{ and} \\ & \big(r_s(u) < 0 \text{ or } [r_s(u) = 0 \text{ and } u < \vartheta(u)]\big). \end{aligned} \tag{9}$$

We use the following theorem as an imported interface.

Theorem 2.1 (Imported parity-selector interface)**.** *For every fixed-point-free complemented alphabet and every even $k \geq 4$, the seed M_0 is an ordinary MDS. The function*

$$m_s(w) = m_0(w) + a_s(\mathrm{tail}(w)) - a_s(\mathrm{head}(w)) \tag{10}$$

is the indicator of a literal reverse-complement-invariant ordinary MDS. The construction and proof require only real linearity, injectivity, nonvanishing, and antisymmetry of s.

The seed statement is Theorem 4.3, the weighted selector is Theorem 6.3, and its literal reverse-complement invariance and minimum-cardinality consequences are Theorem 7.2 and Corollary 7.3 of [2]. Their legal-descent and acyclicity proofs are not repeated here. We use only the formulas above and the fact that every weight-dependent branch is a sign or zero test of an explicitly named real linear form.

2.1.1 Residual length convention

Throughout, $L(M)$ counts edges in $B(q, k-1) - M$, viewing selected length-k words as arcs. Its line graph is $B(q, k) \setminus M$. When $L(M) \geq 1$, the longest upper residual path has $L(M) - 1$ edges. A longest word avoiding selected k-mers has $L(M) + k - 1$ letters.

2.1.2 Component notation

For each letter write

$$s(x) = \sum_{i=1}^{m} \beta_i \varepsilon_i(x), \qquad \varepsilon_i(x) \in \{-1, 0, 1\},$$

with exactly one nonzero component. Define

$$X_i(w) = \sum_{j=0}^{k-1} \varepsilon_i(w_j) \zeta_k^{j+1}, \tag{11}$$

$$L_i(u) = \sum_{j=0}^{k-2} \varepsilon_i(u_j) \frac{\sin(2\pi(j+1)/k)}{\sin(2\pi/k)}. \tag{12}$$

Then

$$P_s(w) = \sum_i \beta_i X_i(w), \qquad \lambda_s(u) = \sin(2\pi/k) \sum_i \beta_i L_i(u). \tag{13}$$

If $K_k = \mathbb{Q}(\zeta_k)$ and $K_k^+ = \mathbb{Q}(\zeta_k + \zeta_k^{-1})$, then $X_i(w) \in K_k$ and $L_i(u) \in K_k^+$. The normalization by the common positive sine in (12) is essential; the sine itself need not lie in the real cyclotomic subfield used for the coefficient test.

2.2 Only two families are needed for the terminal

Write $V_m = \{\pm e_1, \ldots, \pm e_m\}$ for the signed one-hot alphabet. For DNA take $T = e_1, A = -e_1, G = e_2, C = -e_2$ and $\pi_\rho(h, l) = \rho h + l$. Structural vanishing means componentwise vanishing. All path lengths below count edges of the lower residual $D_s = B(q, k-1) - M_s$. Let $w : u \to v$ be a lower-graph arc, $\alpha = 2\pi/k$, and $b(w)$ the fixed zero-PCR seed membership rule, used only when $P_s(w) = 0$.

Theorem 2.2 (Two-family selector)**.** *For every admissible weight map and even $k \geq 4$,*

$$m_0(w) = \begin{cases} b(w), & \lambda_s(u) = \lambda_s(v) = 0, \\ 1, & \lambda_s(u) \leq 0 < \lambda_s(v), \\ 0, & \textit{otherwise.} \end{cases}$$

Consequently the signs of the two lower-word families λ_s, r_s, together with the frozen combinatorial tie rules, determine the seed, midpoint ideal, terminal, and labelled residual graph.

Proof. The suffix letter has exponent k and contributes no imaginary part, so

$$\lambda_s(u) = \Im P_s(w), \qquad \lambda_s(v) = \Im(e^{-i\alpha} P_s(w)).$$

Writing $P_s(w) = x + iy$, the second coordinate equals $y\cos\alpha - x\sin\alpha$. Since $\sin\alpha > 0$, both lower coordinates vanish exactly when $x = y = 0$. If $y = 0$ and the embedding is nonzero, the next imaginary coordinate is positive exactly when $x < 0$. This is the negative-real seed branch. Otherwise the seed is the strict lower-to-upper crossing, giving the displayed rule. The midpoint indicator remains

$$a_s(u) = 1 \iff \lambda_s(u) > 0 \text{ and } [r_s(u) < 0 \text{ or } (r_s(u) = 0 \text{ and } u < \mathrm{RC}(u))].$$

Substitute into $m_s(w) = m_0(w) + a_s(u) - a_s(v)$. □

This is a sufficient decision interface, not a claim that arbitrary raw seed-real signs can be recovered everywhere. The five-role certificate catalogue remains useful for classification and audit. The five-role trace is not an irredundant decision interface: the identities above supply the explicit aliases.

Corollary 2.3 (Symmetry compression)**.** *Let Z_λ^{str} count globally structural lower-lambda words. The terminal is determined by at most*

$$\frac{q^{k-1} - Z_\lambda^{\mathrm{str}}}{2}$$

labelled real-form occurrences, in addition to the structural registry and fixed ties.

Proof. Let C complement letters and J reverse positions. The pairs (λ, r) transform under $1, C, J, CJ$ with signs $(+,+), (-,-), (-,+), (+,-)$ respectively. A J-fixed lower word is palindromic and has componentwise zero lambda. Neither C nor CJ fixes an odd-length word over a fixed-point-free complemented alphabet. Thus the nonstructural-lambda lower words form free four-element orbits. Retain λ and r at one representative per orbit and restore all signs. A globally structural-lambda word always has $a_s = 0$, so its midpoint sign is unnecessary. □

For DNA at $k = 4, 6, 8$, the bound is respectively $24, 440, 7904$. These are occurrence counts before proportional-form deduplication, not minimum hyperplane counts.

2.3 Complete decision covectors and terminal chambers

At fixed $\mathcal{A}$ and k, the selector has finitely many labelled branch occurrences. We retain the occurrence tags rather than deduplicating coincident linear forms. This makes branch completeness auditable.

Definition 2.4 (Complete decision catalog)**.** *The catalog $\mathcal{D}_{\mathcal{A},k}$ contains every weight-linear real form tested by the deterministic construction, with its labelled word, rotation, and branch role. The roles are:*

$$\text{seed-real}, \quad \text{seed-imag}, \quad \text{embedding-zero}, \quad \text{lower-lambda}, \quad \text{lower-midpoint}.$$

The embedding-zero role records both real coordinates needed to decide $P_s(w) = 0$. A form that is identically zero in weight space is retained as a structural tie and is not treated as a hyperplane. The complete covector $\sigma(s)$ *is the sign in $\{-, 0, +\}$ of every nonidentically-zero catalog occurrence, together with the structural-tie registry.*

The catalog can be exponentially large in k; this section claims finiteness and completeness, not a small irredundant arrangement.

Theorem 2.5 (Terminal sign-covector factorization)**.** *Fix a complemented alphabet, even $k \geq 4$, the deterministic zero-PCR rule, and the lexicographic midpoint tie rule. If two admissible weight maps s, t have the same complete decision covector, then they have identical:*

1. *seed indicators and zero-PCR representatives;*
2. *positive lower support and midpoint ideal;*
3. *terminal indicators $m_s = m_t$ word by word;*
4. *labelled residual digraphs; and*
5. *deterministic statistics computed from those labelled digraphs.*

Consequently the complete terminal payload is constant on every relatively open face of the realizable arrangement defined by $\mathcal{D}_{\mathcal{A},k}$ and is invariant under positive scaling of all magnitudes.

Proof. The signs of the tagged seed-real and seed-imag occurrences determine the negative-ray and strict-crossing branches on every nonzero PCR. The paired coordinates tagged embedding-zero determine exactly which PCRs use the fixed, weight-independent zero rule. Hence the seed indicators agree.

The lower-lambda signs determine the same set $U_s = U_t$. On that set, the lower-midpoint signs, followed only when zero by the fixed lexicographic rule, determine the same function $a_s = a_t$. Substitution in the endpoint formula (10) gives $m_s = m_t$ on every labelled word. Deleting the same labelled vertices gives the same residual digraph, so any deterministic statistic of it is identical.

The sign covectors are constant on relatively open faces of a finite real hyperplane arrangement [3]. Every catalog form is homogeneous of degree one in the magnitudes, so multiplication by a positive scalar preserves all signs and all tie decisions. □

The five-role trace is sufficient but not irredundant. Theorem 2.2 reduces it to two lower-word families with fixed ties. Equal aggregate zero counts alone still do not determine the terminal.

Remark 2.6 (Resonance strata versus terminal strata)**.** *The resonance stratification records only which selector equalities vanish. The terminal decision arrangement records the full sign of every labelled catalog occurrence. It therefore refines the resonance stratification: equal zero loci need not determine equal seed choices, midpoint ideals, or terminals. The bounded atlas of Proposition D.1 exhibits strict refinement in its declared rows; no all-order irredundancy claim is inferred from that finite witness.*

Corollary 2.7 (Exact one-parameter DNA diagram)**.** *For DNA weights*

$$s(A) = -\rho, \quad s(C) = -1, \quad s(G) = 1, \quad s(T) = \rho, \qquad \rho > 1, \tag{14}$$

every catalog form is $a\rho + b$ with real algebraic coefficients a, b. After structural forms are removed, the exact phase diagram is the ordered finite set of realized critical ratios $-b/a > 1$, the intervening open intervals, and the critical points evaluated with the frozen tie rules. The complete terminal payload is constant on every such cell.

Proof. Every selector form is real-linear in the weight table. Substitution of (14) leaves an affine function of ρ. A nonzero affine form can change sign only at its unique root. Exact algebraic comparison orders the finitely many admissible roots. The conclusion follows from Theorem 2.5; point cells are kept separate because their zero branches need not agree with either adjacent interval. □

This corollary proves a complete diagram only after the entire catalog has been generated. Omitting a branch family can merge cells whose terminals are different.

3 Retiming and reverse-complement sectors

3.1 Integer retiming and reverse-complement components

The ordinary statements in this section assume $q \geq 2$ and $k \geq 2$; the RC statements have the additional hypotheses specified below. This section works on $G = B(q, k-1)$, with lower vertices and length-k arcs, including loops. Write $m = \mathbf{1}_M$ for an ordinary MDS. A complete move at a residual source u exchanges all incoming and outgoing arcs; its indicator change is $\nabla \mathbf{1}_u$. Inverse moves act at residual sinks. Let $\mathcal{C}_F(M)$ denote the component generated by these moves, allowing non-RC intermediate MDS.

Lemma 3.1 (Classical component interface)**.** *Two MDS M, N belong to the same ordinary component if and only if*

$$\mathbf{1}_N = m + \nabla f \quad \textit{for an integer function } f.$$

Thus $m(Q) = \sum_{e \in Q} m(e)$ is constant on each directed cycle throughout the component.

Proof. A move adds a unit gradient, and gradients telescope on cycles. Conversely normalize $\min f = 0$. If its maximum h is positive, choose a source u of the current residual induced on $\{f = h\}$. An incoming residual arc from a lower level would give $\mathbf{1}_N(v \to u) = f(v) - h < 0$. Hence u is a source of the entire residual and a complete move is legal. A new cycle is impossible: after the exchange no residual arc leaves u. For each letter, au and ua lie in the same PCR, so the one-per-PCR condition makes the exchange Boolean and preserves cardinality. A homopolymer loop stays selected and has zero gradient. Replace f by $f - \mathbf{1}_u$. Its sum decreases and stays nonnegative, so iteration reaches N. □

The complete-move and cycle-invariant framework is due to Marçais, DeBlasio and Kingsford [17, Sections 4–5]. The proof above fixes the gradient sign and the constructive scheduler.

Lemma 3.2 (Short separation)**.** *Different ordinary components are separated by a simple directed cycle of length at most* $2k-1$ *whose selected-arc counts differ.*

Proof. Let $d = \mathbf{1}_N - \mathbf{1}_M$ and root the graph at $o = a^{k-1}$. Choose $A_u : o \to u$ by appending u and $B_u : u \to o$ by appending a; both have length at most $k-1$. Put $p(u) = \sum_{A_u} d$. If some closed walk $A_u B_u$ has nonzero sum, its length is at most $2k-2$. Otherwise $\sum_{B_u} d = -p(u)$, and $A_u e B_v$ has sum $p(u) + d(e) - p(v)$ for $e : u \to v$. If all these sums vanished, $d = \nabla(-p)$, contradicting Lemma 3.1. A nonzero closed walk of length at most $2k-1$ therefore exists. Decomposing it into simple cycles leaves at least one with nonzero sum and no greater length. □

3.1.1 The cycle ratio determines the ordinary optimum

Define

$$R(M) = \max_{Q \text{ simple directed cycle}} \frac{|Q|}{m(Q)}.$$

Every denominator is positive because M meets every cycle.

Theorem 3.3 (Unit retiming)**.** *For every integer* $T \geq 1$*, the following are equivalent:*

1. *Some* $N \in \mathcal{C}_F(M)$ *has* $L(N) \leq T-1$.
2. *There is an integer function* H *satisfying*

$$H(v) - H(u) \geq 1 - Tm(e) \qquad (e : u \to v). \tag{15}$$

3. *Every simple directed cycle* Q *satisfies* $|Q| \leq Tm(Q)$.

Consequently

$$\boxed{\min_{N \in \mathcal{C}_F(M)} L(N) = \lceil R(M) \rceil - 1.}$$

Proof. Given N and $\mathbf{1}_N = m + \nabla f$, let h be its longest-path rank, with values between zero and $T-1$. Then $H = h - Tf$ satisfies (15), on both selected and residual arcs.

Conversely divide any integer solution as $H = Ta + h$, with $0 \leq h < T$, and set $f = -a$, $m' = m + \nabla f$. The inequality is

$$h(v) - h(u) + Tm'(e) \geq 1.$$

Since $h(v) - h(u) \leq T-1$, the integer $m'(e)$ is nonnegative. Each PCR has sum one: the gradient telescopes. The PCRs partition all arcs, so m' is Boolean. On each zero arc h strictly increases, giving a DAG of length at most $T-1$ and exactly one selected arc per PCR. Thus m' is an MDS in the required component.

Summing (15) proves necessity of the cycle inequalities. For sufficiency give arcs weight $1 - Tm(e)$ and add a source with zero arcs to every vertex. With no positive cycle, a maximum-weight walk can have its cycles deleted without loss. Maximum walk weights from the source are finite integers and satisfy (15). Closed walks decompose into simple cycles. The smallest feasible integer T is $\lceil R(M) \rceil$. □

This is a specialization of classical retiming [15]. The proof is self-contained; its MDS-specific step is that nonnegative retiming and one unit per PCR force Boolean arc values. Bellman–Ford gives $O(|V||E|)$ integer relaxations per feasibility test; binary search on $1 \le T \le |V|$ finds the optimum. This is polynomial in the explicit graph, whose size is exponential in k, and is distinct from the direct word-query theorem.

A simple cycle with r selected arcs gives the lower bound $\lceil |Q|/r \rceil - 1$ throughout a component: its r residual stretches have total length $|Q| - r$. In particular a one-hit cycle of length $h+1$ and an attainer of length h certify the component optimum.

3.1.2 An integer rank with prescribed mate sum

For the RC results assume q and $k \ge 4$ are even, with fixed-point-free alphabet complement. Then $\theta = \mathrm{RC}$ on lower vertices is fixed-point-free and reverses arcs.

Lemma 3.4 (Bounded autodual rank)**.** *If a finite DAG has such an involutive antiautomorphism and longest path at most an integer $c \ge 0$, it has an integer rank $h : V \to \{0, \dots, c\}$ with*

$$h(\theta v) = c - h(v), \qquad h(v) - h(u) \ge 1 \quad (u \to v).$$

Proof. Let $d(v)$ be the longest path ending at v. The half-integral rank $r(v) = (d(v) + c - d(\theta v))/2$ lies in $[0, c]$, has mate sum c, and increases by at least one on every arc. Keep its integer values; round half-integers below $c/2$ down and those above up. For half-integers exactly at $c/2$, choose opposite rounding in each mate pair. There are no edges between two such central vertices. For two half-integral endpoints on the same side the rounding agrees; across the center it increases the gap, with the worst central choice preserving it. For one integer endpoint the original gap is at least $3/2$, and loses at most $1/2$. Thus all final gaps are at least one. □

If M, N are RC and $\mathbf{1}_N = m + \nabla f$, then $f(v) + f(\theta v) = C$ is constant, by comparing mate arcs. Adding an integer constant to f changes C by an even number. The two parity sectors relative to M are denoted $\mathcal{C}_\epsilon(M)$, $\epsilon \in \{0, 1\}$, normalizing $f + f \circ \theta = \epsilon$.

Theorem 3.5 (RC retiming in a fixed sector)**.** *There is $N \in \mathcal{C}_\epsilon(M)$ with $L(N) \le T - 1$ if and only if an integer solution of (15) satisfies*

$$H(v) + H(\theta v) = c_\epsilon, \qquad c_\epsilon = T - 1 - T\epsilon. \tag{RC}$$

It reconstructs N by $f(v) = -\lfloor H(v)/T \rfloor$ and $\mathbf{1}_N = m + \nabla f$.

Proof. For necessity take the rank from Lemma 3.4 with $c = T - 1$ and set $H = h - Tf$. Conversely write $H = Ta + h$, $0 \le h < T$. Its mate sum is congruent to -1 modulo T, so the sum of the two remainders is $T - 1$, also when $T = 1$. Hence $a + a \circ \theta = -\epsilon$ and $f + f \circ \theta = \epsilon$. The ordinary proof gives a Boolean MDS; this constant mate sum makes its gradient RC invariant. □

Given an ordinary integer solution d of (15), set

$$H_*(v) = \frac{d(v) - d(\theta v) + c_\epsilon}{2}.$$

It is a half-integral solution with the required mate sum. Let K_* consist of tight arcs $H_*(v) - H_*(u) = 1 - Tm(e)$ between half-integral vertices.

Theorem 3.6 (Integral feasibility by tight components)**.** *The system* (15) *and (RC) has an integer solution if and only if no strongly connected component of K_* contains both v and θv.*

Proof. If tight paths go both ways between v and θv, their weights force every feasible solution to have the same difference $H(\theta v) - H(v)$ as H_*. The mate sum then forces $2H(v) = 2H_*(v)$, an odd integer, a contradiction.

Otherwise the strongly connected components of K_* come in distinct mate pairs. Number the condensation topologically by i. Round a half-integer up exactly when $i([v]) > i([\theta v])$, otherwise down, leaving integers fixed. Mate roundings are opposite. A tight edge cannot round its tail up and head down: that would require

$$i([\theta u]) < i([u]) \le i([v]) < i([\theta v]) \le i([\theta u]).$$

A nontight edge with two half-integral endpoints has slack at least one and loses at most one. With one half-integral endpoint it has slack at least $1/2$ and loses at most $1/2$. Thus all inequalities and mate sums survive. □

The tight-component integrality mechanism has a classical UTVPI antecedent in Lahiri–Musuvathi [13, Section 4.2]. Here the prescribed mate sum and PCR reconstruction implement it for Boolean minimum decycling sets.

Corollary 3.7 (At most one unit of parity cost)**.** *Let $T_0 = \lceil R(M) \rceil$. Both RC sectors are nonempty and*

$$\min_{N \in \mathcal{C}_\epsilon(M)} L(N) \in \{T_0 - 1, T_0\}.$$

If $R(M)$ is nonintegral, both optima equal $T_0 - 1$.

Proof. For $T > R(M)$ every cycle has strictly negative clock weight, so the tight graph is acyclic. Its components are singletons, none containing a mate pair. Apply Theorem 3.6 at $T = T_0$ when noncritical, otherwise at $T = T_0 + 1$. The ordinary optimum supplies the lower bound. □

Proposition 4.3 supplies an example attaining the extra unit in one prescribed sector.

3.1.3 Exactly two paired-move components

A paired RC step is a complete move at u and its inverse at θu, with disjoint incident supports. RC is required after the compound step, not after its first half.

Theorem 3.8 (Exactly two RC sectors)**.** *Every ordinary component containing an RC MDS has exactly two components under paired RC steps. They are the two parity sectors. In one sector, if $\mathbf{1}_N - \mathbf{1}_M = \nabla g$ and $g \circ \theta = -g$, the paired-step distance is*

$$d_{\text{pair}}(M, N) = \tfrac{1}{2} \sum_v |g(v)|.$$

Between different sectors the membership of every RC-fixed length-k word is reversed.

Proof. Both sectors exist by Corollary 3.7. In a single sector the difference potential has a unique integer antisymmetric centering g. If its maximum is $h > 0$, choose a residual source u on that level as in Lemma 3.1. Its mate is a sink at level $-h$. An edge between them would have gradient $\pm 2h$, impossible as the difference of two Boolean indicators. Their stars are disjoint and the paired move is legal. It replaces g by $g - \mathbf{1}_u + \mathbf{1}_{\theta u}$, reducing

its norm by two. Iteration gives the asserted length; every paired step changes the norm towards the target by at most two, proving optimality. Paired steps preserve parity. For a fixed arc $e : u \to \theta u$,

$$\mathbf{1}_N(e) - m(e) = 2f(u) - \epsilon.$$

It is zero for $\epsilon = 0$ and ± 1 for $\epsilon = 1$. Such words exist at even order, so the sectors are distinct. □

Within the weight-selector family a fixed periodic word $(a\bar{a})^{k/2}$ always has the same frozen zero-PCR membership. Thus two terminals in the same ordinary component lie in the same paired sector. The second sector need not be realized by weights.

Proposition 3.9 (Components encountered by the DNA family)**.** *At $k = 4, 6, 8$ the fixed labelled DNA family $\mathcal{W}$ with fixed $A < C < G < T$ ties encounters respectively $1, 3, 21$ ordinary components, and the same numbers of paired components. At $k = 8$ its component diagram consists of the eleven intervals and ten distinct points cut out by*

$$(1 + 2\sqrt{2})/3,\ \sqrt{2},\ (5 + 4\sqrt{2})/7,\ (3 + 6\sqrt{2})/7,\ 2\sqrt{2} - 1,$$
$$1 + \sqrt{2},\ 3,\ 1 + 2\sqrt{2},\ 3 + 2\sqrt{2},\ 5 + 4\sqrt{2}.$$

Exactly ten fixed homogeneous linear sign forms suffice and are necessary to identify this component.

Exact certificates. The atlas covers every real ratio $\rho > 1$ in $\mathcal{W}$. Cut-potential normal forms classify its 199 strata; literal paired moves join each to its assigned representative. The replay checks 3857 compound steps. Every pair of distinct classes has a separating cycle: three at $k = 6$ and 210 at $k = 8$, all of length at most $2k - 1$. These are positive and negative certificates, not an assumption that a search found all paths. The component index is $2\#\{r < \rho\} + \mathbf{1}_{\rho \text{ is a listed root}}$. A fixed linear form can have only one normalized root; each listed wall is indispensable and one lambda form per root suffices. The periodic anchor proves that the same classification holds for paired moves inside the family. □

This classifies components encountered by the selector, not all components in the graph. As a small independent boundary check, binary order four has six RC MDS in three ordinary components; all six are isolated under paired moves. Parity cannot be discarded.

4 The certified 141/107 component separation

Throughout the DNA atlas and weighted-component statements in this section, the parameter domain is exactly $\mathcal{W}$, with fixed labels and $A < C < G < T$ ties. The complete atlas does not classify other labelled antisymmetric weight maps.

4.1 The complete fixed-family order-eight atlas

The selector reduction of Theorem 2.2 makes the exact chamber classification finite. The wall-locality and mixed-wall formulas are Theorems D.2 and D.4; their technical arguments and the small-order regression are retained in the supplement (Sections D.1 and D.2). The component coverage is Proposition 3.9. We now give the complete fixed-$\mathcal{W}$ atlas, followed directly by its component optima and the separate 107 construction.

The alphabet is $A < C < G < T$ with weights $(-\rho, -1, 1, \rho)$, $\rho > 1$. Zero-PCR ties are fixed: the smallest RC-fixed word in a stable PCR, and the smallest word in a paired union together with its RC mate.

Here is an explicit finite reduction for the root list. Let $V = \{\pm e_1, \pm e_2\}$ and $D = V + V$, a nine-element alphabet. For a lower word of seven letters write

$$x_1 = v_0 + v_6, \quad x_2 = v_1 + v_5, \quad x_3 = v_2 + v_4, \quad z = v_3.$$

Its vector midpoint form satisfies

$$2r = \sqrt{2}(x_1 - x_3) - 2(x_1 + x_2 + x_3 + 2z).$$

Enumerating $D^3 \times V$ is exhaustive and positionally reversible. For lambda use the same word, with d_i the three paired differences: $\lambda = (d_1 + \sqrt{2}d_2 + d_3)/\sqrt{2}$. An affine scalar form $(a + c\sqrt{2})\rho + b + d\sqrt{2}$ has, when its slope is nonzero, the unique root

$$\rho = \frac{-ab + 2cd + (bc - ad)\sqrt{2}}{a^2 - 2c^2}.$$

Activity must be tested on one and the same positional word; independently choosing pair sums and pair differences would be an invalid relaxation.

Theorem 4.1 (Complete DNA terminal atlas)**.** *At $k = 8$ there are exactly 86 distinct terminal walls in $\rho > 1$: 87 open intervals and 86 point strata. They give 114 different labelled terminal sets. All 86 walls separate different adjacent terminals; at 27 walls the point is a third terminal, and at the other 59 it equals one side. Each terminal has 8230 words. The minimum number of a fixed dictionary of homogeneous linear forms whose ternary signs determine the terminal is exactly 86.*

Exact certificates and real-parameter reduction. The reduction above enumerates every candidate; the exact comparison and activity filters produce 86 midpoint roots, including all ten lambda roots. The certificate stores each root and a realizing word. On each open interval all signs are fixed, and structural ties are fixed, so Theorem 2.2 proves constancy for every real ratio, including transcendental ratios. Rational representatives are chosen by exact dyadic bisection; point roots are evaluated in $\mathbb{Q}(\sqrt{2})$.

The producer and a separate verifier compare all membership bits; the latter reconstructs the original seed in $\mathbb{Q}[X]/\Phi_8$. Integer levels certify acyclicity, and explicit paths certify the reported longest lengths. PCR and RC checks are literal. Terminals are grouped by equality of bit vectors, not by assuming hash injectivity. A separating word at each wall proves necessity. A nonzero homogeneous linear form has at most one root on the normalized line, so at least one form per indispensable wall is necessary. Conversely one midpoint form at each of the 86 roots distinguishes all strata. The files and replay commands are listed in `evidence/extended/README.md`. □

The minimum 86 concerns a fixed linear sign dictionary. It is not a lower bound for adaptive queries, arbitrary nonlinear encodings, or graph isomorphism classes.

4.2 The optimum over the weighted component union

The complete atlas and Proposition 3.9 supply 21 ordinary components at $k = 8$. These are only the components encountered by $\mathcal{W}$, with fixed labels and ties. The next result optimizes their complete union, rather than only the terminals of $\mathcal{W}$. Theorem 4.4 then separates this union from the wider RC MDS space.

Proposition 4.2 (Exact optima in the 21 components)**.** *In each row of Table 1, the ordinary optimum and both RC-sector optima equal the displayed value. Hence the minimum over the union of these entire components encountered by $\mathcal{W}$ is 141.*

Exact certificates. For each component, a simple cycle with one selected arc gives the lower bound throughout that component. Two explicit MDS, one per RC sector, attain it. The checker verifies the clock inequalities, mate sums, gradients, Boolean membership, every PCR, RC, ranks, paths and cycle intersections. A separate sink-removal dynamic program confirms each length. The 42 sector solutions are also reconstructed from ordinary clock solutions by tight-SCC rounding. The proof does not enumerate each entire component. □

Component	Base row	Best weighted	Component optimum	Cycle/hits
0	c000	161	161	162/1
1	c015	157	157	158/1
2	c016	153	153	154/1
3	c027	153	153	154/1
4	c028	149	149	150/1
5	c037	149	149	150/1
6	c038	149	149	150/1
7	c045	147	143	144/1
8	c046	149	149	150/1
9	c055	161	157	158/1
10	c056	147	147	148/1
11	c077	163	161	162/1
12	c078	145	145	146/1
13	c093	141	141	142/1
14	c094	141	141	142/1
15	c109	155	155	156/1
16	c119	143	143	144/1
17	c129	163	161	162/1
18	c130	143	143	144/1
19	c151	155	155	156/1
20	c152	161	161	162/1

Table 1: DNA $k = 8$. Row identifiers have prefix `k8_`. The optimum is attained in both paired RC sectors.

Thus no sequence of complete moves and inverses starting at a terminal of $\mathcal{W}$ can go below 141, even if RC is abandoned in the intermediate states. Changing component is necessary. The corresponding small regression optima are 17 at $k = 4$ and $63, 65, 57$ for the three $k = 6$ components.

4.3 Sharp parity cost and a shorter DNA construction

The atlas of $\mathcal{W}$ is a starting family within the space of MDS. Other labelled antisymmetric weight maps are not classified here. We now give finite witnesses outside it. All lengths in this and the following sections count residual arcs in $B(q, k-1)$.

Proposition 4.3 (The one-unit sector cost is attained)**.** *For DNA at $k = 8$ there is an RC MDS M_* whose ordinary component optimum is 124, while the optima in $\mathcal{C}_0(M_*)$ and $\mathcal{C}_1(M_*)$ are respectively 125 and 124.*

Proof with finite witnesses. The directory `evidence/advances/r1/witnesses/` supplies `Mstar125.words.txt` and `M124_sector1.words.txt`, each an RC set of 8230 words. Direct DAG reconstruction gives lengths 125 and 124. Their difference is the integer gradient stored in `r1/certificates/retiming_Mstar125_to_M124.json`, whose mate sum is odd, so they are in the indicated sectors of one component.

The parity certificate supplies a residual path $P : A^7 \rightsquigarrow T^7$ of length 71 and a return path Q of length 54 containing exactly one selected arc of M_*. Their concatenation is a simple cycle of length 125 with one selected arc. It gives the ordinary lower bound 124. At period $T = 125$, the two clock inequalities force

$$H(T^7) - H(A^7) = 71.$$

Sector zero requires $H(T^7) + H(A^7) = 124$, hence $2H(A^7) = 53$, impossible for an integer clock. The length-125 endpoint attains the next possible value, whereas the other endpoint attains 124. The independent verifier checks all arc differences, both paths, the cycle and both DAGs. □

This resolves sharpness for a prescribed sector. It does not show that the union of the two RC sectors ever has a worse optimum than the ordinary component.

Theorem 4.4 (Certified DNA upper bound 107)**.** *There is an RC DNA MDS M_{107} of cardinality 8230 with $L(M_{107}) = 107$. Its ordinary component has exact optimum 107. Consequently the global RC MDS optimum at $k = 8$ is at most 107. The upper-graph residual has length 106; a longest avoiding DNA word has 114 letters, and every DNA word of length 115 contains a selected eight-letter window.*

Exact certificates. The canonical witness is `r5/witnesses/M107.words.txt`. Its sorted ASCII lines, including the final newline, have SHA-256

`4e0c44d87ea9782f261cf3ed2c0e918784fa404bbf0f57041b494509ff7d5edc.`

The checker enumerates all 65,536 arcs and 16,384 vertices, verifies RC and one selection in each of the 8230 PCRs, and checks a rank in $[0, 107]$ increasing along each residual arc. A residual path of 107 arcs attains the bound. A simple cycle of 108 arcs has exactly one selected arc, proving the component lower bound by Theorem 3.3. Both path and cycle are included in `r5/certificates/M107.json`. The length conversions follow from overlapping consecutive eight-letter words. □

Earlier component-optimal endpoints and their one-hit cycles are retained in the companion; the first explicit escape is described in Section B.1. A four-witness panel at length 107 certifies two ordinary components with both parity sectors represented in each. The complete stable fibre in Proposition 5.3 supplies further components. Neither finite panel exhausts the global component space.

The 108-cycle forces any improvement to $L \leq 106$ to meet that cycle at least twice, hence to leave this ordinary component. Changing the intersection with one old critical cycle is only a necessary condition: Proposition C.1 includes examples where all old critical cycles are repaired but the result is cyclic or belongs to a component with a worse optimum.

5 Completions, bounded charts, and cardinality heights

5.1 Exact completion of the stable choices

Let a finite digraph have an involution θ on vertices reversing arcs. An individually fixed arc has the form $e : u \to \theta u$. Let D be an RC-invariant DAG from which all such candidate arcs have been removed. Write $a_D(v)$ and $b_D(v)$ for its longest incoming and outgoing path lengths and L_D for its longest path. Reachability below includes paths of length zero.

Lemma 5.1 (At most one fixed arc on a path)**.** *An RC-invariant DAG has no directed path containing two individually fixed arcs.*

Proof. If $e : u \to \theta u$ and $f : v \to \theta v$ occur in that order, let $P\colon \theta u \rightsquigarrow v$ be the intervening path. Then $e, P, f, \mathrm{RC}(P)$ is a nonempty closed directed walk, a contradiction. The argument includes an empty P. □

Theorem 5.2 (Stable completion by 2-SAT)**.** *For a subset R of the removed fixed arcs, $D \cup R$ is acyclic if and only if it satisfies the following prohibitions:*

1. Do not restore $e : u \to \theta u$ if D has a path $\theta u \rightsquigarrow u$.

2. Do not restore distinct $e : u \to \theta u$ and $f : v \to \theta v$ together if D has a path $\theta u \rightsquigarrow v$.

For every acyclic completion,

$$L(D \cup R) = \max\{L_D, \max_{e:u\to\theta u\in R}(2a_D(u) + 1)\}. \tag{16}$$

Thus, after fixing the nonstable PCR selections in an even-order RC de Bruijn graph, existence of an MDS completion with $L \le \ell$ is decided by 2-SAT in polynomial time in the explicit graph size.

Proof. Each prohibition is necessary by closing one fixed arc or applying Lemma 5.1. Conversely a cycle must use a restored arc. If it uses exactly one it violates the first prohibition; otherwise two consecutive restored arcs on the cycle violate the second. For the length formula, every path uses at most one such arc. The longest prefix ending at u and suffix starting at θu can be concatenated with it: any repeated vertex would give a cycle. RC gives $b_D(\theta u) = a_D(u)$, proving equality.

For MDS completion delete the fixed nonstable selections and all individually fixed arcs to form D. If D is cyclic or $L_D > \ell$, reject. Otherwise introduce a variable for restoring each fixed arc, require exactly one per stable pair, and add the two types of prohibition as unary and binary clauses. Also forbid any restored arc with $2a_D(u) + 1 > \ell$. These are exactly the necessary and sufficient constraints. Reachability and DAG distances are polynomial. The implication graph decides 2-SAT by checking that no variable and its negation belong to one strongly connected component. □

The implication-graph algorithm is classical [1]; the complete RC reachability and length reductions are proved above. The algorithm fixes the nonstable selections first; it is not a polynomial algorithm for the global MDS optimization problem or for counting all 2-SAT models. For $\ell = 106$, a restored fixed arc must have $a_D(u) \le 52$.

Proposition 5.3 (The complete stable fibre of M_{107})**.** *Keeping all nonstable selections of the canonical M_{107} fixed leaves exactly 128 RC MDS: 121 of the 128 stable choices are forced and seven are free. The resulting cube has ordinary-component minimum profile*

$$107 : 8, \qquad 109 : 8, \qquad 111 : 16, \qquad 113 : 32, \qquad 115 : 64.$$

The 128 endpoints belong to distinct ordinary components. None of these fixed-exterior completions has $L \le 106$.

Exact certificates. The stable-cube certificate reconstructs all fixed arcs, all reachability prohibitions and all costs in (16). It verifies the 128 DAGs and, for each, a one-hit cycle one arc longer than its length. These certify the component minima, not just the fibre minima. Distinct stable assignments in this cube are neither equal nor global complements, since 121 bits are fixed. Lemma 5.7 separates their components. A 107-arc path using only nonstable PCRs remains after every stable change, and independently obstructs the target 106. □

5.1.1 Allowing one extra selected arc

Let N be the number of PCRs. At cost at most $N+1$, an RC feedback set has exactly one selection in every nonstable PCR: surplus there comes in RC pairs and costs at least two. At cost $N+1$, exactly one stable PCR is double. Its selected pair is either its two individually fixed arcs or a nonfixed RC pair.

Corollary 5.4 (Complete fixed-exterior oracle for excess at most one)**.** *For even q, k and free alphabet complement, fix an RC selection of one arc per nonstable PCR. Existence of an RC feedback set extending it with cost at most $N+1$ and $L \le \ell$ is decided by at most $1 + \sum_{P \text{ stable}} |P|/2$ instances of 2-SAT. This is polynomial in the explicit graph size. For DNA $k = 8$ the number of branches is $1 + 128 + 366 = 495$.*

Proof. The preceding classification is exhaustive because every PCR must be hit. Take one branch with no exceptional PCR, or choose the exceptional PCR and its double selection. A stable PCR of period p has one fixed pair and $(p-2)/2$ nonfixed pairs, giving $p/2$ exceptional branches. Construct D by deleting the fixed nonstable selections, the chosen exceptional nonfixed pair if any, and all individually fixed arcs. Reject cyclic or overlong D. Ordinary stable pairs require exactly one restored fixed arc. For the exceptional pair, force both restoration variables to zero if both fixed arcs were selected, or both to one if a nonfixed pair was selected. All constraints are 2-CNF. Theorem 5.2 applies to any subset of restored fixed arcs, including these forced cases, and gives exact clauses. Their union over the branches is precisely the desired family. The stable PCRs have 988 arcs at DNA order eight, yielding 495. □

This corollary combines the completion theorem with the surplus classification; it does not choose the 4051 nonstable RC-PCR groups of the global problem. Its finite regression includes satisfiable and unsatisfiable 2-SAT branches, not just cases rejected before the SAT step.

5.1.2 Rank projection and cuts with the correct surplus

Proposition 5.5 (Continuous ranks for Boolean endpoints)**.** *For a Boolean arc indicator n and integer $\ell \ge 0$, the residual is a DAG with $L \le \ell$ if and only if there is a real rank $r \in [0, \ell]^V$ such that*

$$r(h(e)) - r(t(e)) \ge 1 - (\ell+1)n(e).$$

If n is RC invariant the rank may have mate sum ℓ.

Proof. An integral longest-path rank proves necessity, including on selected arcs where the right side is $-\ell$. Conversely summing along zero arcs excludes cycles and paths longer than ℓ. Replace $r(v)$ by $(r(v) + \ell - r(\theta v))/2$ for the last claim. □

Integrality of the signed clock in Theorem 5.9 remains essential; this proposition assumes Boolean n.

The rank system can be projected exactly to a boundary B of variable arcs, retaining every variable-arc inequality on its boundary endpoints. Let D be the fixed residual DAG and keep all inequalities $r(v) - r(u) \ge d_D(u, v)$ for reachable boundary pairs, where d_D is the longest path distance. Also keep $a_D(v) \le r(v) \le \ell - b_D(v)$ and check $L_D \le \ell$. To extend feasible boundary values, set the rank of any vertex v to the maximum of $a_D(v)$ and all $r(u) + d_D(u, v)$ with $u \in B$ reaching v. The boundary values are preserved; the suffix bounds keep every extended rank at most ℓ, and all fixed arc inequalities hold. This proves exact elimination. It may create many boundary-pair constraints and is not an automatic reduction in total model size.

Corollary 5.6 (Two-arm inequality with stable surplus)**.** *Let a stable PCR P have fixed arcs $e_i : u_i \to \theta u_i$, $i = 0, 1$. For each i choose a path of h nonfixed arcs ending at u_i, and let U be the union of their arc sets. Every RC feedback set M with $L(M) \le 2h$ satisfies*

$$\sum_{e \in U} \mathbf{1}_M(e) + \big(|M \cap P| - 1\big) \ge 1.$$

Proof. The PCR is hit. If it is hit at least twice the inequality is immediate. Otherwise RC forces its sole selected arc to be fixed, so the other e_i is residual. If M avoided both arms, RC would also leave both reflected arms residual. The arm ending at u_i, followed by e_i and its reflected arm, would be a residual walk of $2h + 1$ arcs. It is either an overlong path or contains a cycle, both impossible. □

The sum uses the full arc universe, including arcs that become selectable when surplus is allowed. The surplus belongs to the *stable* PCR P. At MDS cost it vanishes; at cost $N + 1$ it can legitimately deactivate a clause. Omitting it is false: for binary $k = 4$, the RC feedback set

$$\{0000, 0011, 0100, 0101, 1000, 1100, 1101, 1110, 1111\}$$

has $L = 2$. The stable PCR $\{0011, 0110, 1001, 1100\}$ is hit twice, while neither one-arc arm 0001 or 0110 is selected. The unlifted inequality would read $0 \ge 1$.

5.2 A complete bounded chart of RC endpoints

Assume throughout this section that q and $k \ge 2$ are even, the alphabet complement has no fixed letter, and a base indicator m is RC invariant with exactly one selection per PCR. Lower vertices have odd length and occur in distinct pairs $\{u, \theta u\}$. The base need not be acyclic for the algebraic representation below.

Lemma 5.7 (Stable PCRs and a parity restriction)**.** *Every RC-stable PCR has exactly two individually RC-fixed arcs. An RC selection of one arc in that PCR chooses one of them. Every integral combination of paired transfers*

$$g_{u,a} = \mathbf{1}_{ua} - \mathbf{1}_{au} + \mathbf{1}_{\mathrm{RC}(ua)} - \mathbf{1}_{\mathrm{RC}(au)}$$

preserves the parity of each individually fixed arc coefficient. If two RC MDS differ by an integral gradient, all their fixed-arc bits are either unchanged or simultaneously complemented.

Proof. Write $\mathrm{RC}(w) = \mathrm{rot}^j w$ on an orbit of primitive period p. The induced reflection of letter positions cannot fix a position, since a letter is different from its complement. Thus p is even and j is even. Fixed rotations solve $2i = j \pmod p$, giving two solutions. A one-element RC-invariant selection must be fixed. At a fixed arc the two summands in each RC pair in $g_{u,a}$ coincide, so its coefficient is even. For $n - m = \nabla f$, comparison of mate arcs makes $f(v) + f(\theta v) = C$ constant. On $e : u \to \theta u$, $n(e) - m(e) = 2f(u) - C \equiv C \pmod 2$; Boolean endpoints prove the claim. □

For DNA $k = 8$, the 128 stable choices, modulo simultaneous complementation, yield a 127-bit ambient label for ordinary components. This is neither a complete invariant nor a claim that all labels are realized. Omitting independent stable corrections therefore imposes a genuine restriction on an integral transfer search.

Lemma 5.8 (Reconstruction from a signed cochain)**.** *Let $x : E \to \mathbb{Z}$ be RC invariant and sum to one on every PCR. If $T \geq 1$ and an integral H satisfy*

$$H(h(e)) - H(t(e)) + Tx(e) \geq 1, \qquad H(v) + H(\theta v) = T - 1, \tag{17}$$

then

$$n(e) = x(e) - \lfloor H(t(e))/T \rfloor + \lfloor H(h(e))/T \rfloor$$

is an RC MDS indicator and $L(n) \leq T - 1$.

Proof. Write $H = Ta + r$, $0 \leq r < T$. Then $Tn(e) + r(h(e)) - r(t(e)) \geq 1$, so integrality gives $n(e) \geq 0$. Its sum on each PCR is one by telescoping. Since the PCRs partition the arcs, these conditions force $n(e) \in \{0, 1\}$. Rank r increases along each zero arc. The mate sum gives $r(v) + r(\theta v) = T - 1$ and $a(v) + a(\theta v) = 0$, making $n = x - \nabla a$ RC invariant. Its cardinality attains the PCR lower bound. This includes $T = 1$. □

Choose one representative of each lower-vertex pair and a reference letter a_*. In each stable PCR P, let f_P^0 be the selected fixed arc of m and f_P^1 the other, and set $\eta_P = \mathbf{1}_{f_P^1} - \mathbf{1}_{f_P^0}$. Put

$$x = m + \sum_{u,\, a \neq a_*} z_{u,a} g_{u,a} + \sum_{P \text{ stable}} \gamma_P \eta_P.$$

Theorem 5.9 (Complete bounded endpoint chart)**.** *An RC MDS of residual length at most $T - 1$ exists if and only if (17) has a solution with*

$$z_{u,a} \in \{-2, -1, 0, 1, 2\}, \quad \gamma_P \in \{0, 1\}, \quad -T \leq H(v) \leq 2T - 1,$$

all coordinates integral. There is no restriction on the number of nonzero sites u.

Proof. Each generator preserves RC and each PCR sum, so sufficiency is Lemma 5.8. For necessity, take a required endpoint n. In one PCR of each nonstable RC pair, transport the selection of m to that of n along a rotation path shorter than a full turn, and add its RC companion. After canonicalizing u, this gives transfer coefficients $t_{u,a} \in \{-1, 0, 1\}$: a coordinate can arise only from a rotation step or its reversed RC companion in the other PCR, and different PCR pairs cannot share it. Treat stable PCRs directly with $\gamma_P \eta_P$.

The identities

$$g_{\theta u, \bar{a}} = -g_{u,a}, \qquad \sum_a g_{u,a} = \nabla(\mathbf{1}_u - \mathbf{1}_{\theta u})$$

allow elimination of the reference coordinate. Set $b_u = t_{u,a_*}$ and $z_{u,a} = t_{u,a} - b_u$. Then $n = x + \nabla f$, where $f = \sum_u b_u(\mathbf{1}_u - \mathbf{1}_{\theta u})$ takes values in $\{-1, 0, 1\}$ and is antisymmetric. Thus $|z_{u,a}| \leq 2$. The bounded autodual-rank proof in Lemma 3.4 applies also here at $k = 2$: take an integral rank $r \in [0, T - 1]$ for the residual of n, with mate sum $T - 1$. Setting $H = r - Tf$ gives the stated bounds and all inequalities. □

At DNA order eight the chart uses $8192 \cdot 3 = 24576$ transfer coordinates, 128 stable bits and 8192 independent clock coordinates, for 32,896 integer variables. A signed x need not be Boolean or acyclic. The theorem concerns its reconstructed endpoint, not the legality of intermediate local moves. The three-site exclusion below allows arbitrary integer coefficients on those sites and arbitrary retiming; its local scope is not a limitation of this complete chart.

5.3 Bounded heights and cardinality above the PCR minimum

The MDS objective fixes cardinality first. A complementary question fixes the path bound and minimizes cardinality. Write

$$\tau_{\mathrm{RC}}(\ell) = \min\{|M| : M \text{ is an RC feedback set and } L(M) \le \ell\}.$$

For DNA $k = 8$, every such set has at least 8230 arcs. A set of 8232 arcs is not an MDS, even when it is inclusion-minimal.

Theorem 5.10 (Exact antisymmetric height formulation)**.** *Let G be a finite digraph with a vertex involution θ reversing arcs, and let $H \ge 0$ be an integer. The minimum cardinality of an RC feedback set with residual length at most $2H$ is*

$$\min_{\substack{h:V\to\mathbb{Z},\ |h(v)|\le H \\ h(\theta v)=-h(v)}} |\{e : u \to v : h(v) \le h(u)\}|.$$

The statement allows fixed vertices of θ.

Proof. For a feasible height h, retain only strictly increasing arcs. This is an RC DAG of path length at most $2H$, proving one inequality. Conversely take a feasible feedback set and its residual longest incoming rank d. The half-integral height $c(v) = (d(v) - d(\theta v))/2$ lies in $[-H, H]$, is antisymmetric, and increases by at least one on residual arcs. Round nonintegral values away from zero, keeping integers fixed. For endpoints on one side of zero their rounding increments agree if both are half-integral; across zero the gap grows. With exactly one half-integral endpoint an integral lower gap bound of one means the original gap is at least $3/2$, so losing at most $1/2$ is harmless. Thus the resulting integral h still increases on all residual arcs. It stays in $[-H, H]$ and is zero on fixed vertices. Its nonincreasing arcs form a subset of the given feedback set, proving the reverse inequality. □

Without antisymmetry, maximizing the increasing arcs is the classical Max-k-Ordering formulation, here with $2H + 1$ labels [10]. The statement above supplies the exact RC restriction for an even path bound, with a proof rather than a continuous-relaxation assumption.

Proposition 5.11 (Cost 8232 at length 105)**.** *There is an RC feedback set S with $|S| = 8232$ and $L(S) = 105$. It is inclusion-minimal even among feedback sets without RC. It hits 8228 PCRs once and two nonstable companion PCRs twice. In particular*

$$8230 \le \tau_{\mathrm{RC}}(106) \le 8232.$$

The same upper bound holds for $\tau_{\mathrm{RC}}(105)$.

Exact certificates. The word list `r11/certificates/S8232.words.txt`, height certificate, rank and attaining path verify the first two numbers. The checker reconstructs the PCR partition and the stated multiplicities. For each singly hit PCR, its rotation cycle certifies necessity of its selected arc. For the four selected arcs in the double PCRs, four explicitly stored one-hit cycles of lengths six or nine provide the missing necessity witnesses. Removing any selected arc therefore restores a directed cycle. The PCR count proves the lower bound 8230 independently of this construction. Neither inclusion-minimality nor these short cycles prove that 8232 is a minimum at the required length. □

There are two further cost-8232 witnesses S_A, S_B of length 105 in `r13/certificates/`. They arise by adding the RC pair $\{\texttt{GAAAAAAA}, \texttt{TTTTTTTC}\}$ to the two length-107 MDS `M107_panel_02.words.txt` and `M107_panel_03.words.txt`. Those MDS are in one ordinary component and opposite sectors. In this component, two is the minimum number of RC additions needed to reach length at most 106: the supplied simple 108-cycle has one selected arc and no individually fixed arcs. Any single RC addition would have to be fixed, so it cannot repair that cycle. Gradients preserve its intersection count. The two displayed additions attain length 105. For the original canonical M_{107}, a separate exact census instead gives minimum RC augmentation cost four. These are fixed-base or fixed-component augmentation statements, not global cardinality lower bounds.

Adding either word of the displayed pair alone gives a set of 8231 arcs with $L = 106$, but that set is not RC invariant. This separates an ordinary construction from the still-open RC target.

Proposition 5.12 (Affine pruning is exhausted for three bases)**.** *For each of S, S_A, S_B, no smaller Boolean RC feedback indicator of the form*

$$n = \mathbf{1}_S - a + \nabla f, \qquad a : E \to \mathbb{Z}_{\geq 0}, \quad a \neq 0, \quad a \text{ RC invariant}, \quad f : V \to \mathbb{R},$$

has residual length at most 106 (replace S by the chosen base). For the first base no smaller feedback endpoint of this form exists at any length. For S_A, S_B the smallest attainable length in this family is 107.

Reduction and exact cycle certificates. On each PCR, gradients telescope and a feedback indicator has positive integral sum. Hence a vanishes on every singly hit PCR, including the stable ones. Each base has just one doubled nonstable PCR pair. In each of its two PCRs, a has sum at most one. RC and $a \neq 0$ force the sums to be one, so a is one RC pair of arcs chosen from this orbit pair. There are exactly eight possibilities, including arcs not selected in the base. The whole de Bruijn graph is balanced, so $\sum_e \nabla f(e) = 0$; any endpoint then has cardinality 8230.

For the first base all eight choices have a certified closed walk of nonpositive $\mathbf{1}_S - a$ weight. Its weight cannot change under a gradient; a Boolean feedback indicator must have positive weight on every nonempty closed walk. For each of the other two bases there is one negative and six zero-weight obstructions. The final case has a 108-cycle of weight one, forcing $L \geq 107$. In that case subtracting the pair actually gives the corresponding length-107 MDS, so the bound is attained. All 24 walks and their integer weights are checked in `r13/certificates/affine_pruning.json`. □

Allowing real f does not weaken the cycle obstructions. Between Boolean endpoints an integral gradient also has an integral potential after a constant normalization, by propagation along the connected graph. This exhaustion concerns subtraction and retiming. It does not allow arbitrary new inserted arcs, which are a separate research direction.

6 Balanced loads and finite-pool closure

6.1 Balanced closed-walk loads and the exact closure of a finite pool

The complete endpoint charts above retain Boolean and acyclicity constraints. We now isolate a weaker necessary test: nonnegative combinations of a fixed pool of closed-walk inequalities whose exterior incidence is invariant under real gradients. This test has an exact finite linear-programming description. Its scope is a real-gradient class on the nonstable arcs, rather than an ordinary complete-move component on the full graph.

Let $k \geq 2$ be even and let the alphabet have a fixed-point-free complement. In $G = B(q, k-1)$ write S for the union of RC-stable PCRs, $E_0 = E \setminus S$, and N_S for their number. Fix an integer $T \geq k$. For $b \geq 0$ define the aggregate stable budget

$$\mathcal{B}_b = \{c \in \mathbb{R}^{N_S} : c_P \geq 1,\ \sum_P c_P \leq N_S + b\}.$$

This relaxation has no upper bounds $c_P \leq |P|$, Boolean constraints or acyclicity requirement. All projections below refer to this stated budget. For an exterior $y : E_0 \to \mathbb{R}$, its real-gradient class is $y + \nabla f|_{E_0}$ with $f : V \to \mathbb{R}$ and $\nabla f(e) = f(\operatorname{tail} e) - f(\operatorname{head} e)$, as in Section 2.1.

Lemma 6.1 (Closed-walk counting inequality)**.** *If M is an RC feedback set with $L(M) \leq T - 1$ and $n = \mathbf{1}_M$, then every nonempty closed walk W, with multiplicity vector χ_W, satisfies*

$$\langle \chi_W, n \rangle \geq \lceil |W|/T \rceil . \tag{18}$$

Proof. The walk meets M, since a nonempty residual closed walk contains a cycle. Cut its circular sequence at the selected occurrences. Each piece has one selected arc followed by at most $T-1$ residual arcs: a residual stretch cannot repeat a vertex and is bounded by $L(M)$. Thus r selected occurrences give $|W| \leq Tr$. Repetitions are counted with multiplicity. □

Proposition 6.2 (Stable-load criterion and aggregate normal form)**.** *The stable subgraph consists of vertex-disjoint directed PCR cycles. Let $F = \sum_j \lambda_j \chi_{W_j}$ with $\lambda_j \geq 0$, and put $A = F|_{E_0}$. Then $A(\nabla f) = 0$ for every real potential f if and only if F has a constant load m_P on the arcs of each stable PCR P. For such a combination set*

$$K = \sum_j \lambda_j \lceil |W_j|/T \rceil , \qquad M_* = \max_P m_P.$$

Its exact aggregate projection is

$$A(y) \geq K - \sum_P m_P - bM_*. \tag{19}$$

If $M_* > 0$, adjoining nonnegative weighted stable PCR cycles makes all positive stable loads equal to M_* without changing this projected inequality. Rational weights remain rational and can be cleared to integers.

Proof. A word in a stable PCR has equal counts in every complementary letter pair, because RC is a rotation of that word. A fixed $(k-1)$-letter prefix can be completed to balanced counts by at most one letter: distinct letters add distinct signed unit vectors to its imbalance. The same argument applies to suffixes. Hence each vertex has stable indegree and outdegree at most one; rotation closure gives vertex-disjoint cycles. This argument uses necessity of balanced counts, not a converse characterization of RC stability.

An arc vector annihilates all gradients exactly when its divergence is zero, as testing the potentials supported at individual vertices shows. The full vector F is a circulation. Thus A is a circulation exactly when $F|_S$ is one; on the disjoint stable cycles this means constant loads. Summing (18) yields $A(n|_{E_0}) + \sum_P m_P c_P \geq K$. Over $\mathcal{B}_b$ the maximum stable term is $\sum_P m_P + bM_*$: place all excess on a largest-load coordinate. This proves the exact projection of this single combined row and the budget; it is not a simultaneous projection of all original walk rows.

For $U = \{P : m_P > 0\}$ adjoin weight $M_* - m_P$ of the PCR cycle P. Its length is at most $k \le T$, so its counting right side is one. The new right side is $K' = K + \sum_{P \in U}(M_* - m_P)$ and its stable maximum is $M_*(|U| + b)$; subtraction gives precisely (19). If every load vanishes no normalization is needed. Clearing denominators proves the last assertion. □

The circulation and normalization argument works on any finite graph whose distinguished subgraph is a union of vertex-disjoint cycles of length at most T. The RC de Bruijn geometry supplies this hypothesis.

Theorem 6.3 (Exact finite-pool closure)**.** *Fix a finite pool of mixed closed walks $W_1, \dots, W_J$, each containing a stable arc. Write $s_{ej} = \chi_{W_j}(e)$ for $e \in S$, $A_j(y) = \sum_{e \in E_0} \chi_{W_j}(e) y_e$ and $d_j = \lceil |W_j|/T \rceil$. Let $F_*(y)$ be the maximum of*

$$\sum_j \lambda_j (d_j - A_j(y)) - \sum_P m_P \tag{20}$$

subject to $\lambda_j \ge 0$, $0 \le m_P \le 1$ and $\sum_j s_{ej}\lambda_j = m_P$ for every $e \in P$, every stable P. For every $b \ge 0$, the following are equivalent:

1. *No nonnegative combination from this pool with gradient-invariant exterior incidence excludes y by (19).*
2. *$F_*(y) \le b$.*
3. *There are signed real coefficients γ_e, $e \in S$, such that*

$$\begin{aligned} A_j(y) + \sum_{e \in S} s_{ej}\gamma_e &\ge d_j && (1 \le j \le J), \\ \Gamma_P := \sum_{e \in P} \gamma_e &\ge 1 && (P \text{ stable}), \\ \sum_{e \in S} \gamma_e &\le N_S + b. \end{aligned} \tag{21}$$

If y is RC invariant and the pool is RC closed, the lift γ can be required to be RC invariant. Individual lift coordinates have no sign, integrality, Boolean or upper-bound restrictions.

Proof. The primal is nonempty because zero is feasible. Every $\lambda_j \le 1$: choose a stable arc in W_j, whose positive integer multiplicity contributes to a total load at most one. Thus the feasible set is compact and the maximum is finite and attained. Its equalities impose exactly the circulation criterion of Proposition 6.2.

A feasible objective larger than b is larger than $b \max_P m_P$ because $b \ge 0$ and $m_P \le 1$, and hence separates y. Conversely a separating combination has $M_* > 0$: in a mixed pool, a nonzero nonnegative combination cannot have all stable loads zero. Divide its weights and loads by M_*. The normalized feasible primal point has objective larger than b. This proves the first equivalence.

We explicitly derive the dual, using classical finite LP strong duality [4, Chapter 5]. Put free multipliers γ_e on $m_P - \sum_j s_{ej}\lambda_j = 0$ and nonnegative t_P on $m_P \le 1$. The Lagrangian is

$$\sum_P t_P + \sum_j \lambda_j \left[d_j - A_j(y) - \sum_e s_{ej}\gamma_e \right] + \sum_P m_P[-1 + \Gamma_P - t_P].$$

Its supremum over nonnegative λ, m is finite precisely when

$$\sum_e s_{ej}\gamma_e \ge d_j - A_j(y), \qquad t_P \ge \Gamma_P - 1, \quad t_P \ge 0.$$

The dual objective is $\sum_P t_P$, and a feasible primal of finite value has an attaining dual of equal value.

Formally add every stable PCR cycle as a column. This leaves the primal maximum unchanged. If its coefficient is η_P, remove it and replace m_P by $m_P - \eta_P$; the remaining loads are nonnegative and at most one, and both right-side contribution and subtracted mass drop by η_P. Conversely add zero coefficients. These columns impose $\Gamma_P \geq 1$ in the dual because their right sides are one. Its minimum is at most b exactly when $\Gamma_P \geq 1$ and $\sum_P \Gamma_P \leq N_S + b$, taking $t_P = \Gamma_P - 1$ for sufficiency. This is (21). Finally RC permutes the mixed rows and preserves each stable PCR set. Averaging γ with $\gamma \circ \theta$ preserves feasibility and budget. □

The convention can also be checked directly. If an endpoint satisfies $n|_{E_0} = y + \nabla f|_{E_0}$, take $\gamma = n|_S - \nabla f|_S$. The full cochain $y + \gamma = n - \nabla f$ has the same closed-walk and PCR evaluations as n. It satisfies the extension whenever the endpoint satisfies its stable budget. Data using $\delta f = f(\text{head}\, e) - f(\text{tail}\, e)$ require $\delta = -\nabla$; transport $y = x + \delta a$ then transports the lift by $\gamma_y = \gamma_x + \delta a = \gamma_x - \nabla a$.

Corollary 6.4 (The MDS budget)**.** *At $b = 0$, (21) forces $\Gamma_P = 1$ on every stable PCR. Every balanced row reduces to $A(y) \geq K - \sum_P m_P$.*

Proof. Each of the N_S masses is at least one and their total is at most N_S. Substitute $c_P = 1$ in the combined row. □

This specialization remains a necessary test on the exterior. Pure E_0 walks must be checked separately, and feasibility of a signed lift does not imply a Boolean stable completion, a clock or a full endpoint. Adding mixed walks can strictly strengthen the test, as the following single example shows.

Proposition 6.5 (An exterior passing one pool and excluded by new returns)**.** *There is an RC exterior Y of 8102 nonstable DNA8 arcs, one per nonstable PCR, whose residual core $(V, E_0 \setminus Y)$ is a DAG of length exactly 106. At $T = 107$ a specified pool of 1804 mixed walks admits an RC signed rational lift of total stable mass 128 for $y = \mathbf{1}_Y$. Thus this entire balanced-pool closure passes at both $b = 0$ and $b = 1$. Nevertheless every RC feedback set M with $L(M) \leq 106$ and*

$$\mathbf{1}_M|_{E_0} = \mathbf{1}_Y + \nabla f|_{E_0}, \qquad f : V \to \mathbb{R},$$

has $|M| \geq 8235$. Attainment of 8235 in this class is not asserted.

Finite certificate reduction. Encode A, C, G, T by $0, 1, 2, 3$. A length-eight word has ID $e < 65536$, tail $\lfloor e/4 \rfloor$ and head e mod 16384; RC reverses the digits and replaces d by $3 - d$. Rotation partitions the arcs into 8230 PCRs. Exactly 128 PCRs are stable, containing 988 arcs. The other 8102 PCRs form 4051 RC pairs.

All paths in this proof are relative to `evidence/balanced_loads/`. The literal witness and exact payload are `results/joint_core106.words.txt` and `results/joint_core106.json.gz`. The companion verifier `verify_case.py` uses the declared shared exact kernel to reconstruct the carrier and check all selections, RC, nonstable PCR counts, and a rank equal to the longest incoming core distance. Every residual core arc increases this rank, whose maximum is 106; a listed 106-arc residual path proves equality. The core omits every stable arc freely and is not a complete feedback-set witness.

The mixed pool is reconstructed from the literal sources and declared alphabet images, joined walks and their RC images, deduplicating by arc multiplicity. Its 1804 walks

compress to 902 mixed rows on 622 stable RC arc orbits. The exported rational lift has denominators one or two. Exact row evaluation checks all mixed lowers, each of the 128 PCR masses $\Gamma_P = 1$, and total mass 128. By Theorem 6.3, no balanced combination of this pool and the stable PCR cycles separates Y.

For the stronger bound, forty further walks supply a return Q_e entirely in E_0 from head e to tail e for every arc e of five stable PCRs. Each eQ_e is closed. Put

$$K_P = \sum_{e \in P} \lceil (1 + |Q_e|)/107 \rceil , \qquad A_P(z) = \sum_{e \in P} \sum_{a \in Q_e} z(a).$$

Their stable restriction is the incidence of the PCR cycle, so their external incidence is a circulation. Hence $A_P(\nabla f) = 0$ and (18) gives $c_P \geq K_P - A_P(\mathbf{1}_Y)$. The checked closed-walk lengths range from 53 to 111, with the following values:

PCR representative	K_P	$A_P(\mathbf{1}_Y)$	lower on c_P
`ATCCCGGG`	10	8	2
`ATCGCGCG`	9	7	2
`CCCGGGCG`	9	7	2
`CCGCGCGG`	9	7	2
`CCGCGGCG`	9	7	2

The verifier checks adjacency, closure, E_0 support, complete PCR coverage, ceilings, multiplicities, circulation and evaluation. The remaining 123 stable PCRs require one hit each, so stable cost is at least 133. A gradient telescopes on each nonstable PCR, preserving its selected count one. Nonstable cost is therefore 8102, proving $|M| \geq 8102 + 133 = 8235$ for arbitrary real f. □

Summing the five return rows gives $A(y) + \sum_{P \in U} c_P \geq 46$. For an RC endpoint of cost at most 8231, nonstable surplus is even and must vanish. Its stable budget is at most 129; after the other 123 mandatory hits, $\sum_{P \in U} c_P \leq 6$. Thus $A(y) \geq 40$, or $A(y) \geq 41$ in the MDS branch. On Y the value is 36. These are weighted occurrence thresholds, not counts of distinct newly selected words. The example proves strictness relative to the stated pool, without asserting independence from every retained historical row. It leaves $8230 \leq \tau_{\mathrm{RC}}(106) \leq 8232$ unchanged and gives no global lower bound or attained optimum in the new real-gradient class.

7 Main arithmetic results

The arithmetic theory explains which positional cancellations can occur before one asks whether they change a terminal or its component. Its complete proofs, including all ancillary structural and orbit counts, are in the technical supplement, Sections A.1 to A.6.

For even $k \geq 4$ with t distinct odd prime divisors, the accidental DNA embedding spectrum in $\mathcal{W}$ is exactly

$$\mathcal{E}_t = \left\{ \frac{v+u}{v-u} : 0 < u < v \leq 2^t,\ (u,v) = 1 \right\}.$$

This is Theorem A.20: prime-power splitting and the rational tensor kernel reduce necessity to a bounded mixed difference, and the signed-one-hot realization supplies actual labelled words for sufficiency. The result depends on the number of distinct odd primes, not their values or multiplicities. It is an embedding spectrum, not a terminal-wall classification.

A real algebraic ratio $\rho > 1$ has an accidental positional DNA embedding at some even order if and only if it belongs to a cyclotomic field (Theorem A.28). The constructive sufficiency allows the order to grow. Field membership at a prescribed order is only a necessary gate (Theorem A.1 and corollary A.2); minimal realizing orders and active algebraic midpoint realization remain open.

At twice-prime orders the structural formula and rational lower-resonance classification are Theorems A.7 and A.15. Exact certificate transport, the two-odd-prime counts, rational seed and midpoint classifications, and direct-membership bit complexity are retained in Theorems A.12, A.13, A.17 and A.24 to A.26. The proofs keep the positional alphabet restrictions separate from unrestricted root-of-unity relations. The residual three-block interface in Theorem A.30 connects this arithmetic to the wall geometry of Section 4.

8 Limits, open questions, and evidence availability

The all-order rational embedding spectrum, rational twice-prime classifications, complete fixed-labelled-$\mathcal{W}$ selector $k = 8$ atlas, ordinary component optimization and existence of exactly two RC sectors are established, together with sharp fixed-sector cost, a complete bounded chart, conditional stable completion and exact balanced finite-pool closure. The latter remains a relaxation with signed real lifts. The following questions remain.

1. **Global DNA optimum at $k = 8$.** The construction has $L = 107$, and 107 is its exact component optimum. The minimum over all RC MDS is unknown. Any $L \leq 106$ construction must leave that component and hit the displayed 108-cycle at least twice. This necessary condition is not a sufficient construction or a global lower bound of 107. The 128-member stable fibre and the three-site chart domain are already closed at this target. Searching those same domains again cannot improve the result.

2. **Exact cardinality at path bound 106.** We have $8230 \leq \tau_{\mathrm{RC}}(106) \leq 8232$, with cost 8232 already attaining length 105. Cost 8231 would require exactly one double stable PCR; the 495-branch completion theorem solves its stable subproblem only after the nonstable choices are fixed. The eight regional exclusions and three affine-pruning exhaustions are not global exclusions. The real-gradient-class lower in Proposition 6.5 is likewise local to its stated affine class. A new insertion or a change of nonstable component may escape these domains.

3. **Uniform changes of component.** The four patches are certified at $k = 8$. No theorem supplies analogous improvements at arbitrary even order, jointly preserving Boolean membership, RC, PCR counts and acyclicity.

4. **Fixed-order algebraic realization and active branches.** The cyclotomic-field condition is sufficient when the order may grow. At a prescribed order the full positional image, minimal realizing order and efficient conductor bounds remain unclassified. Eventual raw midpoint witnesses can have lambda zero; general algebraic active midpoint realization is open. The resolved quadratic power-of-two slice is not the full higher-order domain.

5. **Higher-order terminal and component arrangements.** At $k = 8$, 86 fixed linear signs determine the terminal and ten determine its encountered component, both minimally. General orders and higher-dimensional weight spaces lack such a minimal classification here. Adaptive or nonlinear encodings are distinct questions. Other labelled sign/order assignments already at $k = 8$ are outside $\mathcal{W}$ and their terminal and

component arrangements are not classified here; no exclusion of all such weight maps follows from the 21-component atlas.

6. **Length control at mixed walls.** Regular walls have width at most two. The mixed-wall indicator formula does not imply a comparable uniform path-length bound; a width-20 example is certified. The max-plus formula is exact, but equality in $L \leq 2L_P + L_N + 2$ is verified only in the finite atlas and still needs a uniform alignment argument.

7. **Simultaneous cost in both RC sectors.** Each sector optimum is $\lceil R \rceil - 1$ or $\lceil R \rceil$. The extra unit is now attained in one sector, while the other sector in that example attains the ordinary optimum. Whether both sectors of some ordinary component require the extra unit remains unresolved here. All 42 sectors from the fixed-labelled-$\mathcal{W}$ atlas still attain their ordinary optima.

8. **Component range and qualified symmetries.** The minimum does not determine the maximum in a component or settle the component-range questions of [17]. Naive Galois descent to terminals is false. A useful restricted transport theorem must explicitly preserve positivity, positional data and ties.

None of these questions is an unproved premise of the established results. Biological and software benefits require comparative benchmarks. Odd-order constrained decycling and packing identities are separate problems. The general height theorem permits fixed vertices, but this does not remove the even-order hypotheses of the PCR chart, paired-sector and stable-completion applications.

8.1 Code and data availability

The accompanying companion has repository home

https://github.com/aconsciousfractal/
FCIG-Cyclotomic-Weight-Chambers-for-RC-Invariant-MDS

This manuscript accompanies version 3.2.0. Its distributed evidence is identified by the immutable SHA-256 digest

`1c9666c6fa78651514c260d1da8569be4c530469e32b94241dfbcd474efdd034`

The companion documents the digest's scope and replay procedure. Paths abbreviated as `rN/...` in this article are relative to `evidence/advances/` at that repository; `evidence/finite_assurance/` contains the original regression capsule. Literal witnesses, ranks, paths, cycles and rational certificates support the finite statements. Computation does not replace the uniform proofs, certify global optimality, or establish biological performance. This version identifier and digest do not assert an existing tag, release, DOI or journal publication.

AI assistance and responsibility

AI systems assisted proof exploration, counterexample search, implementation of exact checks, drafting and editorial revision. The author is responsible for the mathematical claims and the artifact. Separate implementations provide computational cross-checks; their AI-assisted origin is not a claim of independent human peer review or formal proof verification. No complete historical model/version inventory is asserted.

A Technical supplement: arithmetic proofs and exact counts

A.1 Component fields, structural zeros, and algebraic designs

A zero of a selector form is called *structural* when it holds componentwise, and therefore for every choice of positive magnitudes. A zero that is not structural is called *accidental*. This terminology refers to the weight parameter, not to numerical error.

Theorem A.1 (Component-field zero criterion and structural extremality)**.** *Fix an even $k \geq 4$.*

1. *If $\beta_1, \ldots, \beta_m$ are linearly independent over K_k, then*

$$P_s(w) = 0 \quad \Longleftrightarrow \quad X_i(w) = 0 \text{ for every } i.$$

2. *If $\beta_1, \ldots, \beta_m$ are linearly independent over K_k^+, then*

$$\lambda_s(u) = 0 \quad \Longleftrightarrow \quad L_i(u) = 0 \text{ for every } i.$$

3. *For arbitrary admissible magnitudes, every componentwise zero remains zero. Hence a field-generic table attains the universal structural floor for each of these two zero sectors.*

Proof. The decompositions (13) are linear relations in the declared coefficient fields. Under the corresponding field-linear independence, the relation can vanish only when every coefficient vanishes. The converse is immediate. Componentwise vanishing does not involve the β_i, so those states remain zero for every admissible table. Every other zero is an additional cancellation, proving the extremal assertion. □

The lower statement would be false as written without dividing out the common factor $\sin(2\pi/k)$. For example, at $k = 6$ that sine is $\sqrt{3}/2$; treating it as a rational coefficient confuses the coefficient field with a common nonzero scalar.

Corollary A.2 (DNA field gate)**.** *For DNA magnitudes $(\rho, 1)$, an accidental embedding zero forces $\rho \in K_k$, and an accidental lower zero forces $\rho \in K_k^+$. Field membership is necessary, not sufficient, for a realized DNA cancellation.*

Proof. An embedding relation has the form $\rho X_H + X_L = 0$. If it is accidental, $X_H \neq 0$ and $\rho = -X_L/X_H \in K_k$. The lower proof is identical with $L_H, L_L \in K_k^+$. Membership alone supplies no labelled word whose two component coefficients realize the required quotient. □

The gate gives exact schedules for probing or avoiding algebraic resonance. We record three useful regimes. All field-theoretic implications below are standard consequences of conductor and ramification theory for cyclotomic fields [23].

Corollary A.3 (Three exact weight-design regimes)**.** *Let $\rho > 1$ be the high DNA magnitude.*

1. *Put $\varphi = (1+\sqrt{5})/2$ and $\rho_{\text{gold}} = \varphi^4 = (7+3\sqrt{5})/2$. Then $\rho_{\text{gold}} \in K_k$ exactly when $5 \mid k$. Thus the family is a conductor-five resonance probe.*

2. *If ρ generates a non-Galois real extension of $\mathbb{Q}$, then $\rho \notin K_k$ for every k. In particular $\rho = \sqrt[3]{2}$ gives a uniform all-order structural-only design for the two sectors of Theorem A.1.*

3. *Given a finite horizon K, choose a prime $\ell > K$ with $\ell \equiv 1 \pmod 4$ and put $\rho = \sqrt{\ell}$. Then $\rho \notin K_k$ for every $4 \leq k \leq K$, so the same two sectors are structural-only throughout the declared horizon.*

Proof. The quadratic field $\mathbb{Q}(\sqrt{5})$ has conductor 5, which gives the first equivalence. Every subfield of the abelian extension $K_k/\mathbb{Q}$ is Galois over $\mathbb{Q}$; a non-Galois real field therefore cannot be contained in K_k. For the last statement, the prime ℓ ramifies in $\mathbb{Q}(\sqrt{\ell})$, whereas all finite ramification in K_k is supported on primes dividing k. □

These regimes do not assert that the golden family resonates whenever $5 \mid k$. They separate where resonance is algebraically impossible from where a certificate-level realization must still be proved.

Corollary A.4 (Uniform radical magnitude family)**.** *Fix $m \geq 1$, choose an odd prime $d \geq \max\{m, 3\}$, put $\alpha = 2^{1/d}$, and define*

$$\beta_i = \alpha^{i-1}, \qquad 1 \leq i \leq m. \tag{22}$$

Then the magnitudes are positive and distinct, and they are linearly independent over K_k for every k. Consequently, at every even $k \geq 4$, the embedding and lower-lambda zero sectors contain only their componentwise structural zeros.

Proof. The polynomial $x^d - 2$ is Eisenstein at 2, so $E = \mathbb{Q}(\alpha)$ has prime degree d. For any cyclotomic field K_k, the intersection $E \cap K_k$ therefore has degree either 1 or d over $\mathbb{Q}$. The second case would embed E as an intermediate field of the abelian Galois extension $K_k/\mathbb{Q}$, forcing $E/\mathbb{Q}$ to be Galois. This is impossible: $E \subset \mathbb{R}$, whereas the irreducible polynomial $x^d - 2$ has nonreal roots $\alpha\zeta_d^j$ and does not split in E.

Thus $E \cap K_k = \mathbb{Q}$. Since $K_k/\mathbb{Q}$ is Galois, the two fields are linearly disjoint, and the degree of α over K_k remains d. No nonzero polynomial over K_k of degree below d can vanish at α; because $m \leq d$, the powers $1, \alpha, \ldots, \alpha^{m-1}$ are independent over K_k. The conclusion follows from Theorem A.1, with independence over K_k also implying independence over K_k^+. □

This family is an exact existence design. It makes no claim about numerical conditioning, magnitude range, downstream path quality, or practical superiority to a DNA weight table.

A.2 The twice-prime structural formula

Let $k = 2p$ with p an odd prime and put $h = (p-1)/2$. Identify the letters with the vertices

$$V_m = \{\pm e_1, \ldots, \pm e_m\} \subset \mathbb{Z}^m$$

through the component map $x \mapsto (\varepsilon_1(x), \ldots, \varepsilon_m(x))$.

Lemma A.5 (Prime sine basis)**.** *The positive numbers*

$$\sin(\pi/p), \ldots, \sin(h\pi/p)$$

are linearly independent over $\mathbb{Q}$.

Proof. Suppose $\sum_{r=1}^{h} c_r \sin(r\pi/p) = 0$ with $c_r \in \mathbb{Q}$, clear denominators, and write $\xi = e^{\pi i/p}$. Then

$$g(x) = \sum_{r=1}^{h} c_r (x^{h+r} - x^{h-r})$$

satisfies $g(\xi) = 0$ and has degree at most $p-1$. The minimal polynomial $\Phi_{2p}(x) = x^{p-1} - x^{p-2} + \cdots - x + 1$ also has degree $p-1$. Therefore g is zero or a scalar multiple of Φ_{2p}. The former polynomial is anti-palindromic and the latter palindromic, so comparison of leading and constant coefficients forces the scalar to vanish. The disjoint coefficient pairs of g then give $c_r = 0$ for every r. □

For a lower word of length $2p-1$, sine indices $1,\dots,2p-1$ split into the free central index p and, for $1 \le r \le h$, the four-position block

$$\{r, p-r, p+r, 2p-r\}.$$

The coefficient of $\sin(r\pi/p)$ is therefore

$$v(u_r) + v(u_{p-r}) - v(u_{p+r}) - v(u_{2p-r}) \in \mathbb{Z}^m. \tag{23}$$

Lemma A.6 (Cross-polytope additive energy)**.**
The number of ordered quadruples $(x_1,x_2,x_3,x_4) \in V_m^4$ satisfying $x_1+x_2=x_3+x_4$ is

$$E(V_m) = 6m(2m-1). \tag{24}$$

Proof. Count ordered pairs by their sum. The zero sum has multiplicity $2m$. Each sum $\pm 2e_i$ has multiplicity one, contributing $2m$. For each unordered pair of distinct coordinate axes and each independent choice of signs, a sum $\pm e_i \pm e_j$ has multiplicity two. There are $4\binom{m}{2}$ such sums. Thus

$$E(V_m) = (2m)^2 + 2m + 4\binom{m}{2} 2^2 = 6m(2m-1).$$

□

Theorem A.7 (Twice-prime structural zeros and firings)**.** *Let $q=2m$, $k=2p$, p an odd prime, and $h=(p-1)/2$. The number of structural lower zeros is*

$$Z_{\mathrm{str}}(q,2p) = q\,[6m(2m-1)]^h. \tag{25}$$

For a table whose magnitudes are linearly independent over K_{2p}^+, this is the exact lower-zero count. The direct descent then has

$$F_{\mathrm{lower}}(q,2p) = \frac{q^{2p-1} - Z_{\mathrm{str}}(q,2p)}{4} \tag{26}$$

physical lower firings.

Proof. By Lemma A.5, componentwise lower vanishing is equivalent to the independent vanishing of (23) in every one of the h blocks. Each block has $E(V_m) = 6m(2m-1)$ ordered assignments by Lemma A.6; the central letter has sine coefficient zero and is free, giving the factor q. This proves (25). Field genericity turns structural into total lower vanishing by Theorem A.1.

Letterwise complementation changes the sign of every lower value. Thus the nonzero states split equally between the two signs. Reverse complement acts freely on the positive states and the self-dual ideal chooses exactly one state from each pair. Hence the firing count is one quarter of the nonzero lower-state count, proving (26). □

The factor q in (25) is load-bearing: it is the free central coordinate. The count is for ordered letter assignments, and the firing denominator is four rather than two.

Corollary A.8 (Closed structural-zero density)**.** *Under the hypotheses of Theorem A.7,*

$$Z_{\mathrm{str}}(q,2p) = q\,[3q(q-1)]^h, \qquad \frac{Z_{\mathrm{str}}(q,2p)}{q^{2p-1}} = \left(\frac{3(q-1)}{q^3}\right)^h. \tag{27}$$

Proof. Substitute $q = 2m$ into $6m(2m-1) = 3q(q-1)$. Since $p = 2h+1$, the lower-word length is $2p-1 = 4h+1$; division by q^{4h+1} gives the second identity. □

Theorem A.9 (Single-odd-prime structural embedding words and PCRs)**.** *Let $q = 2m$, let p be an odd prime, and let*

$$k = 2^b p^a, \qquad r = 2^{b-1}p^{a-1}, \qquad a, b \geq 1.$$

Put

$$M_p(q) = q^p + q + 2^{p-1}q(q-2).$$

The number of componentwise structural embedding-zero words is

$$Z_{\mathrm{emb,str}}(q,k) = M_p(q)^r. \tag{28}$$

The number of their rotation orbits is

$$N_{\mathrm{emb,str}}(q,k) = \frac{1}{k}\left(M_p(q)^r + \sum_{\substack{d \mid k \\ d<k}} \varphi(k/d) q^d \right). \tag{29}$$

In particular, when $a = b = 1$, this becomes

$$N_{\mathrm{emb,str}}(q,2p) = \frac{2q^p + pq + (p-1)q^2 + 2^{p-1}q(q-2)}{2p}. \tag{30}$$

If the magnitudes are independent over K_k, the formulas count every scalar embedding-zero word and PCR.

Proof. Set $\zeta = \zeta_k$ and $\xi = \zeta^r$, so that ξ is a primitive $2p$th root. The standard prime-power identity and the degree calculation give

$$\Phi_k(X) = \Phi_{2p}(X^r), \qquad [\mathbb{Q}(\zeta) : \mathbb{Q}(\xi)] = r.$$

Thus $1, \zeta, \ldots, \zeta^{r-1}$ is a basis of $\mathbb{Q}(\zeta)$ over $\mathbb{Q}(\xi)$. Writing the component vector at position $s + rt$ as $x_{s,t}$, we obtain

$$\sum_{j=0}^{k-1} x_j \zeta^j = \sum_{s=0}^{r-1} \zeta^s \left(\sum_{t=0}^{2p-1} x_{s,t} \xi^t \right).$$

Consequently the full embedding vanishes componentwise exactly when each of the r disjoint length-$2p$ blocks vanishes componentwise.

For one block, $\xi^p = -1$, so vanishing is equivalent to

$$\sum_{t=0}^{p-1} (x_{s,t} - x_{s,t+p}) \xi^t = 0.$$

In every coordinate this is a rational polynomial of degree at most $p-1$. The minimal polynomial $\Phi_{2p}(z) = 1 - z + \cdots + z^{p-1}$ has the same degree, so there is a vector v such that

$$x_{s,t} - x_{s,t+p} = (-1)^t v, \qquad 0 \leq t < p. \tag{31}$$

If $n(v)$ counts ordered pairs $(c,d) \in V_m^2$ with $c - d = v$, the block count is $\sum_v n(v)^p$. The difference histogram has multiplicity q at zero, multiplicity one at each of the q values

$\pm 2e_i$, and multiplicity two at each of the $q(q-2)/2$ signed two-axis differences. This gives the block count $M_p(q)$. Independence of the r blocks proves (28).

For the orbit count, apply Burnside to this rotation-invariant zero set. Every word fixed by a nonidentity shift is periodic with period dividing $d = \gcd(s,k) < k$; its embedding contains the vanishing geometric factor $1+\zeta^d+\cdots+\zeta^{k-d}$. There are $\varphi(k/d)$ shifts with this gcd and each fixes q^d words. Adding the identity contribution and dividing by k proves (29). At $k = 2p$, the proper divisors $1, 2, p$ give (30). The final assertion is Theorem A.1. □

Corollary A.10 (Rational lower energy and the structural floor)**.** *Under the same twice-prime notation, let the distinct positive magnitudes β_i be rational. Put*

$$S_\beta = \{\pm\beta_1, \ldots, \pm\beta_m\}.$$

If

$$E(S_\beta) = \#\{(a,b,c,d) \in S_\beta^4 : a+b = c+d\},$$

then the scalar lower-zero count is

$$Z_{\text{lower}}(\beta, 2p) = q\, E(S_\beta)^h. \tag{32}$$

It attains the structural floor $q[3q(q-1)]^h$ exactly when the projection $\pm e_i \mapsto \pm\beta_i$ creates no new two-sum equality, equivalently when it is a Freiman 2-isomorphism on V_m. In particular, $\beta_i = B^{i-1}$ attains the floor for every integer $B \geq 4$. The threshold cannot be replaced uniformly by $B = 3$.

Proof. For rational magnitudes, Lemma A.5 makes lower vanishing equivalent to one scalar two-sum equality in each of the h independent blocks; the central letter is free. This proves (32). The projection from V_m to S_β is a bijection and preserves every structural two-sum relation, so equality with the structural energy $E(V_m) = 3q(q-1)$ holds exactly when it creates no additional relation.

Now let $\beta_i = B^{i-1}$ with $B \geq 4$. After cancelling identical signed terms from a scalar two-sum equality, a largest uncancelled magnitude on only one side exceeds the sum of the at most three smaller terms. Equal largest terms with the same sign would already have cancelled, while opposite largest signs require the smaller terms to bridge at least twice that magnitude. The only remaining cancellations are the structural opposite-pair relations. Thus the projection is a Freiman 2-isomorphism. For $B = 3$ and $m = 2$, the nonstructural equality $3 + (-1) = 1 + 1$ shows that no uniform base-three statement is possible. □

The corollary is deliberately finite-additive. Classification of all energy hyperplanes, chambers, or optimal magnitude systems lies outside this article.

A.3 Certificate-preserving DNA resonance

Abstract divisibility by a cyclotomic polynomial is not a DNA word witness. To obtain a reversible statement, we retain position, complement pair, sign, and selector-family information throughout the translation.

For a DNA word $w = w_0 \cdots w_{n-1}$ define signed one-hot coordinates (h_j, ℓ_j) by

$$A \mapsto (-1, 0), \quad C \mapsto (0, -1), \quad G \mapsto (0, 1), \quad T \mapsto (1, 0).$$

Thus $|h_j| + |\ell_j| = 1$ and $s(w_j) = \rho h_j + \ell_j$. Put

$$H_w(z) = \sum_{j=0}^{n-1} h_j z^{j+1}, \qquad L_w(z) = \sum_{j=0}^{n-1} \ell_j z^{j+1}. \tag{33}$$

Reverse complement acts transparently:

$$H_{\mathrm{RC}\,w}(z) = -z^{n+1}H_w(z^{-1}), \qquad L_{\mathrm{RC}\,w}(z) = -z^{n+1}L_w(z^{-1}). \tag{34}$$

Each equality role uses a frozen Laurent coefficient template. Up to a nonzero scalar, the templates are

Role	coefficient at position j
embedding zero	z^{j+1}
seed real	$z^{j+1} + z^{-(j+1)}$
seed imaginary / lower lambda	$z^{j+1} - z^{-(j+1)}$
lower midpoint	$z^{j+1} + z^{-(j+1)} - 2$

The lower roles use $0 \le j \le k-2$; the seed roles use the relevant length-k rotation. Multiplication by a monomial or a fixed nonzero scalar does not alter the equality.

Definition A.11 (Realized selector certificate)**.** *A certificate is the tuple consisting of the labelled word, its equality role and rotation, all signed one-hot coordinates, and the resulting Laurent polynomials $A(z), B(z)$ obtained by inserting respectively h_j and ℓ_j into the role template. Its scalar equality is*

$$\rho A(\zeta_k) + B(\zeta_k) = 0. \tag{35}$$

Theorem A.12 (Certificate-preserving DNA translation)**.** *For every frozen DNA selector equality, the map above has the following properties.*

1. *The original equality holds exactly when (35) holds.*
2. *It is structural exactly when $A(\zeta_k) = B(\zeta_k) = 0$.*
3. *If $A(\zeta_k) = 0$ and $B(\zeta_k) \neq 0$, no ratio satisfies it. Otherwise every nonstructural equality has the unique candidate*
$$\rho = -\frac{B(\zeta_k)}{A(\zeta_k)},$$
which must additionally be real and greater than one.
4. *Suppose this candidate belongs to K_k and is represented in a fixed canonical cyclotomic coordinate as $\rho = U(\zeta_k)/d$, with $U \in \mathbb{Z}[z, z^{-1}]$ and $d \in \mathbb{Z}_{>0}$. Then, after multiplication by a sufficient monomial to remove negative exponents,*
$$\Phi_k(z) \mid z^N\big(U(z)A(z) + dB(z)\big) \quad \text{in } \mathbb{Z}[z]. \tag{36}$$
Conversely, this divisibility implies the scalar equality for the same retained certificate.

Proof. Substitution $s(w_j) = \rho h_j + \ell_j$ into the role's linear equality gives (35) term by term, proving the first assertion. The next two assertions are the exhaustive alternatives for the two evaluated coefficients. When $A(\zeta_k) \neq 0$, solving a linear equation in ρ gives the unique candidate.

For the last assertion, multiply $\rho A(\zeta_k) + B(\zeta_k) = 0$ by d and substitute the represented value of ρ. Clearing Laurent exponents gives an integer polynomial vanishing at the primitive root ζ_k. Its minimal polynomial over $\mathbb{Q}$ is Φ_k, so divisibility follows. Conversely, divisibility implies vanishing at ζ_k and hence the original scalar equality. Because the certificate is retained, the converse returns to the same labelled word and equality role; it does not turn an arbitrary divisible polynomial into a DNA witness. □

Theorem A.13 (Typed Galois covariance of equality certificates)**.** *Let $\sigma_t \in \mathrm{Gal}(K_k/\mathbb{Q})$ be defined by $\sigma_t(\zeta_k) = \zeta_k^t$, where $\gcd(t,k) = 1$. Transport a typed certificate C by carrying every signed coefficient, role tag, and rotation tag at exponent e to the residue $te \bmod k$; denote the resulting typed certificate by D_tC. Then*

$$A_{D_tC}(\zeta_k) = \sigma_t(A_C(\zeta_k)), \qquad B_{D_tC}(\zeta_k) = \sigma_t(B_C(\zeta_k)). \tag{37}$$

Whenever $A_C(\zeta_k) \neq 0$, its equality candidate transforms as

$$-\frac{B_{D_tC}(\zeta_k)}{A_{D_tC}(\zeta_k)} = \sigma_t\left(-\frac{B_C(\zeta_k)}{A_C(\zeta_k)}\right). \tag{38}$$

Thus the unfiltered set of typed realized equality candidates is closed under the Galois action.

Proof. For an exponent residue e, the four role templates are

$$z^e, \qquad z^e + z^{-e}, \qquad z^e - z^{-e}, \qquad z^e + z^{-e} - 2.$$

Evaluation of the corresponding template at exponent te is exactly the σ_t-image of its evaluation at e; the rational constant -2 is fixed. Multiplication by a unit permutes all residues for embedding and seed certificates and all nonzero residues for lower certificates, so the transport is again typed in the same role. Integer linearity gives (37), and division in K_k gives (38). □

This theorem is intentionally prior to the selector's order tests. Galois automorphisms need not preserve positivity: at $k = 10$, $1 + 2\cos(2\pi/5) > 1$ is carried by σ_3 to $1 + 2\cos(6\pi/5) < 1$. Therefore no Galois invariance of the admissible $\rho > 1$ slice, complete sign covector, or terminal MDS follows.

General divisibility criteria and classifications of abstract root-of-unity relations [16, 7, 14, 12, 8] can now be applied as filters. They become word theorems only after their coefficient hypotheses are checked against this certificate image.

A.4 Rational twice-prime lower resonances

We now specialize to DNA, $k = 2p$ with p an odd prime, and rational $\rho > 1$. The statement concerns only the lower form λ_s.

In one prime-sine block, the two letters on the positive side and the two on the negative side contribute sums of two vertices of $V_2 = \{\pm e_H, \pm e_L\}$. Under scalar evaluation

$$\omega_\rho(x_H, x_L) = \rho x_H + x_L,$$

the possible nonzero absolute pair-sum values are

$$2\rho, \qquad \rho + 1, \qquad \rho - 1, \qquad 2. \tag{39}$$

Lemma A.14 (Unique rational block collision)**.** *For rational $\rho > 1$, two distinct vector sums in $V_2 + V_2$ have the same scalar value under ω_ρ only when $\rho = 3$. At $\rho = 3$ the only new equality in absolute value is $\rho - 1 = 2$.*

Proof. Apart from zero and sign, distinct scalar pair sums are the four expressions in (39). Pairwise equality among them has no solution $\rho > 1$ except $\rho - 1 = 2$, which gives $\rho = 3$. Equality to the negative of another positive value is impossible. Direct inspection of the signed pair sums proves the claim. □

At $\rho = 3$, the exceptional ordered blocks, including exchange of the two sides, are generated by

$$(T, C), (C, T) \longleftrightarrow (G, G), \qquad (A, G), (G, A) \longleftrightarrow (C, C). \tag{40}$$

There are eight ordered side-aware patterns in total.

Theorem A.15 (Rational twice-prime lower-resonance classification)**.** *Let p be an odd prime, $k = 2p$, $h = (p-1)/2$, and let $\rho > 1$ be rational. Then every realized lower zero is structural unless $\rho = 3$. At $\rho = 3$, a lower word is zero exactly when every prime-sine block is either structural or one of the eight exceptional blocks in (40). It is accidental exactly when at least one block is exceptional. Consequently*

$$Z_{\mathrm{str}}(4, 2p) = 4 \cdot 36^h, \tag{41}$$
$$Z_{\mathrm{tot}}(3; 4, 2p) = 4 \cdot 44^h, \tag{42}$$
$$Z_{\mathrm{acc}}(3; 4, 2p) = 4(44^h - 36^h). \tag{43}$$

The corresponding physical firing counts are

$$F_{\rho \neq 3}(4, 2p) = 4^{2p-2} - 36^h, \tag{44}$$
$$F_{\rho = 3}(4, 2p) = 4^{2p-2} - 44^h. \tag{45}$$

Proof. By Lemma A.5, the full lower form vanishes exactly when the scalar coefficient of every prime-sine block vanishes. If the two vector sums in a block are equal, the block is structural. If they differ but have equal scalar values, Lemma A.14 forces $\rho = 3$, and direct use of the signed one-hot certificate gives exactly (40). This proves the classification without forgetting the realizing letters.

For DNA, Lemma A.6 gives 36 structural ordered assignments per block. At $\rho = 3$ there are eight additional assignments, hence 44. The h blocks are independent and the central letter is free, giving (41) and (42); subtraction gives (43). Finally apply the quartering argument of Theorem A.7 to the total zero count, yielding (44) and (45). □

For $p = 3$, the theorem gives 144 structural zeros, 176 total zeros at $\rho = 3$, and 32 accidental zeros. The firing counts are respectively 220 and 212. These numbers are checks of the theorem, not its proof. The tempting ratio $\rho = 2$ is not exceptional: the four positive quantities in (39) are then $4, 3, 1, 2$, all distinct.

Corollary A.16 (Scoped twice-prime DNA trichotomy)**.** *Fix $k = 2p$ with p an odd prime and consider only the DNA* lower-lambda *equality family. Then:*

1. *if $\rho \notin K_{2p}^{+}$, every lower zero is structural;*
2. *if $\rho > 1$ is rational, an accidental lower zero exists if and only if $\rho = 3$; and*
3. *among ratios not decided by these two clauses, only $(K_{2p}^{+} \setminus \mathbb{Q}) \cap (1, \infty)$ remains.*

The third clause is an exact residual domain, not an existence assertion.

Proof. The first clause is the necessary field gate of Corollary A.2. The second is Theorem A.15. Partitioning the admissible real ratios $\rho > 1$ by field membership and rationality leaves precisely the set in the third clause; field membership alone remains only necessary. □

A.5 Exact direct-membership complexity

The direct formula (10) avoids materializing the MDS, but an exact complexity statement requires a literal representation of every real algebraic input. We use the following model.

Let a real algebraic number α be supplied by a square-free polynomial $f \in \mathbb{Z}[x]$, normalized to have positive leading coefficient, and a rational isolating interval selecting α. Put $e = \deg f$. Each positive magnitude is represented by a dense coefficient list as

$$\beta_i = \frac{b_i(\alpha)}{d_i}, \qquad b_i \in \mathbb{Z}[x], \quad \deg b_i < e, \qquad d_i \in \mathbb{Z}_{>0}.$$

The complement table, a total alphabet order, even $k \geq 4$, and the queried word are explicit. Let L denote the total bit length of this input. Degrees and every coefficient are encoded explicitly; in particular $e = O(L)$.

Theorem A.17 (Dense variable-degree polynomial-bit membership)**.** *Preprocessing time, per-word query time, and working memory for membership in the direct MDS of Theorem 2.1 are polynomial in k, q, and L, even though e is part of the input. The algorithm is exact and stores no table of size q^{k-1} or q^k.*

Proof. Put $c = 2\cos(2\pi/k)$. Its defining polynomial and an isolating interval can be computed from Φ_k in polynomial bit complexity in k: the degree is at most $\varphi(k)/2 \leq k/2$, and the required coefficient bit sizes are polynomial in k. Write $\mu_k \in \mathbb{Z}[y]$ for the square-free minimal polynomial selecting c and $\delta = \deg \mu_k$; thus $\delta = \varphi(k)/2$ and μ_k is monic. The recurrences

$$2\cos(j\theta) = C_j(c), \qquad \frac{\sin(j\theta)}{\sin\theta} = U_{j-1}(c/2), \qquad \theta = 2\pi/k,$$

where $C_0(x) = 2$, $C_1(x) = x$, and $C_{j+1}(x) = xC_j(x) - C_{j-1}(x)$; equivalently, $C_j(x) = 2T_j(x/2)$ for the standard Chebyshev polynomial T_j. Express every real, normalized imaginary, lower-lambda, and midpoint branch as

$$\operatorname{sgn} G(\alpha, c) \tag{46}$$

after multiplication by known positive denominators. Here $G \in \mathbb{Z}[x, y]$ has degree less than e in x, degree $O(k)$ in y, and coefficient bit length polynomial in $k + L$. These bounds follow by the Chebyshev recurrences and by summing at most k represented weights.

Before invoking the sign routine, reduce G in the two algebraic coordinates. Pseudo-division in x by f replaces G by an integral remainder of x-degree less than $e = \deg f$, after multiplication by a known positive power of the leading coefficient of f. Division in y by the monic polynomial μ_k then gives an integral representative $\widetilde{G}$ of y-degree less than δ. These operations preserve the value sign at (α, c), including zero. The two reductions use a number of dense integer-polynomial operations polynomial in $e + k$. Standard pseudo-division height bounds, resultant construction of μ_k, square-free factorization, and real-root isolation keep all degrees and coefficient bit lengths polynomial in $e + k + L$, hence in $k + L$.

Put $d_2 = \max(e, \delta)$ and $d_1 = \max(\deg_x \widetilde{G}, \deg_y \widetilde{G})$; then $d_1 < d_2$ unless $\widetilde{G}$ is constant, whose sign is immediate. The exact equal-degree format of the sign theorem can be met without changing the selected value. If one variable degree of $\widetilde{G}$ is below d_1, multiply by the required power of a rational linear factor that is positive throughout that variable's isolating interval, and clear its positive denominator. The resulting polynomial has bidegree exactly (d_1, d_1) and the same sign at (α, c). If $e < d_2$ or $\delta < d_2$, multiply the corresponding defining

polynomial by enough distinct integral linear factors whose roots lie outside its isolating interval. This preserves square-freeness and the selected algebraic number while making both defining degrees exactly d_2. The added integers can be chosen with polynomial bit length.

Use a common coefficient-height bound τ for the padded defining polynomials, their isolating intervals, and the sign polynomial. Theorem 37 of Emiris and Tsigaridas [9] now applies literally; its bit bound $\widetilde{O}(d_1^3 d_2^3 \tau)$ is polynomial in $k + L$ because d_1, d_2, and the common height bound τ are polynomially bounded by the dense explicit input.

A nonzero seed query uses a constant number of such branch signs per tested rotation. In the zero-PCR branch, at most k distinct rotations and their reverse complements are inspected to apply the deterministic tie rule. The two endpoint values $a_s(\mathrm{tail}(w))$ and $a_s(\mathrm{head}(w))$ each require one lower-lambda sign, one midpoint sign, and, only at equality, a lexicographic comparison. Thus only polynomially many polynomial-bit operations are used per query, and all stored words and polynomials have polynomial size. □

The result is a membership theorem in the dense explicit model. A sparse syntax that hides an exponentially large degree, or an unpriced transcendental/comparison oracle, is outside the theorem. Listing all selected words necessarily has exponential output size in general, so no polynomial-time full materialization or practical-throughput claim follows.

A.6 Signed cyclotomic blocks

The tensor kernel in this section is classical; compare Martin–Reiner [19] and the finite-difference criterion of De Carli–Laporta [8]. The bounded one-hot realization developed from it is stated separately.

Write

$$k = 2^b \prod_{i=1}^{t} p_i^{a_i}, \quad n = 2 \prod_{i=1}^{t} p_i, \quad r = k/n,$$

where the p_i are distinct odd primes. Thus $\Phi_k(X) = \Phi_n(X^r)$ and $[\mathbb{Q}(\zeta_k) : \mathbb{Q}(\zeta_n)] = r$. Every rational-coefficient polynomial decomposes uniquely as

$$f(X) = \sum_{a=0}^{r-1} X^a f_a(X^r), \qquad f(\zeta_k) = 0 \iff f_a(\zeta_n) = 0 \quad \text{for all } a.$$

For a squarefree even block $n = 2N$, antipodal pairing gives coefficients in

$$D_m = V_m - V_m = V_m + V_m.$$

Indeed $\zeta_{2N}^j = (-1)^j \zeta_N^{cj}$ for a unit c modulo odd N; signs and a coordinate permutation preserve the difference alphabet and its multiplicities. After CRT, the coefficient array is indexed by $\prod_i \{0, \ldots, p_i - 1\}$.

Lemma A.18 (Anchored mixed differences)**.** *Let $c_{j_1,\ldots,j_t}$ have rational scalar or rational vector coefficients. Its evaluation at the product of primitive p_ith roots is zero exactly when every full anchored mixed difference vanishes. Each such difference is an alternating sum of 2^t entries, choosing one nonanchor index and the anchor $p_i - 1$ in each coordinate.*

Proof. For prime p, the only relation among $1, \zeta_p, \ldots, \zeta_p^{p-1}$ is their sum. Eliminating the last root replaces coefficients by $c_j - c_{p-1}$. The products of the remaining powers over the distinct primes are a rational basis: their number is $\prod_i (p_i - 1) = \varphi(N)$, by the cyclotomic degree and CRT. Repeating elimination yields precisely the full mixed differences. Vector evaluation is componentwise. □

Lemma A.19 (One-hot realization of a bounded defect)**.** *Every integer vector v with even $\|v\|_1 \le 2^{t+1}$ is a full mixed difference of a D_m-valued array for which every full anchored difference is the same v. Every entry of the array can be assigned to an actual ordered pair of letters. If a projection kills v, that scalar block vanishes; if $v \neq 0$, it is not componentwise zero.*

Proof. Express v as a sum of exactly 2^{t+1} vectors in V_m: use $\|v\|_1$ signed units and pad by opposite pairs. Group these into 2^t pairs and call their sums $d_\epsilon \in D_m$, indexed by $\epsilon \in \{0,1\}^t$. Put $c_\epsilon = (-1)^{|\epsilon|} d_\epsilon$, with the sign convention adjusted uniformly to the chosen difference orientation. Extend to the full grid by mapping each nonanchor coordinate to 0 and its anchor to 1. Every mixed difference is v. Since D_m is centrally symmetric and consists of ordered letter differences, assign letters to each antipodal pair. No grid entry or word position is omitted. □

For $t = 0$, this is the single antipodal difference. A block lifts to higher prime powers by putting it in one residue class modulo r and filling each other class by a constant word, which vanishes componentwise.

A.7 All rational DNA embedding resonances

Theorem A.20 (Farey spectrum)**.** *Let $k \ge 4$ be even, with t distinct odd prime divisors. For a reduced rational $\rho = a/b > 1$, an accidental DNA embedding zero exists if and only if*

$$\begin{cases} a + b \le 2^{t+1}, & a, b \text{ both odd,} \\ a + b \le 2^t, & a, b \text{ of opposite parity.} \end{cases} \tag{F}$$

Equivalently,

$$\boxed{\mathcal{E}_t = \left\{ \frac{v+u}{v-u} : 0 < u < v \le 2^t,\ (u,v) = 1 \right\}.}$$

In particular the spectrum depends only on the number of distinct odd primes, not their values or multiplicities.

Proof. Reduce to squarefree blocks. An accidental zero has a nonzero vector mixed difference v killed by (a, b), hence $v = j(b, -a)$ for a nonzero integer j. Each mixed difference is a signed sum of 2^t elements of D, so its ℓ_1 norm is even and at most 2^{t+1}. If $a + b$ is even, the least allowed $|j|$ is 1; otherwise it is 2. This gives (F). Conversely choose this least j and apply Lemma A.19, then the prime-power lift. The resulting one-hot word has scalar embedding zero and a nonzero component embedding. Finally reduce $(a - b)/(a + b) = u/v$. Its denominator is $(a + b)/2$ when both are odd and $a + b$ otherwise. Condition (F) is precisely $v \le 2^t$. □

Corollary A.21.

$$|\mathcal{E}_t| = \sum_{v=2}^{2^t} \varphi(v), \quad \min \mathcal{E}_t = \frac{2^t + 1}{2^t - 1}, \quad \max \mathcal{E}_t = 2^{t+1} - 1 \quad (t \ge 1).$$

The first spectra are

$$\mathcal{E}_0 = \varnothing, \qquad \mathcal{E}_1 = \{3\}, \qquad \mathcal{E}_2 = \{5/3, 2, 3, 5, 7\}; \qquad |\mathcal{E}_3| = 21.$$

Every rational DNA ratio is eventually an embedding resonance at some even order. Its minimum required number of distinct odd primes is $\lceil \log_2(a + b) \rceil - 1$ for two odd entries, and $\lceil \log_2(a + b) \rceil$ otherwise.

This closes the rational *embedding* existence question at all even orders. It does not classify the full terminal chamber diagram at these orders.

A.8 The finite-additive criterion and an exact small arrangement

Theorem A.22 (General rational embedding floor)**.** *For rational magnitudes and t distinct odd prime factors, the embedding sector attains its structural floor exactly when $V_m \to \{\pm\beta_1, \dots, \pm\beta_m\}$ is a Freiman 2^t-isomorphism: no new equality between two sums of 2^t signed units is created. Integer geometric magnitudes $\beta_i = B^{i-1}$ attain the floor whenever $B \geq 2^{t+1}$. This uniform integer-base threshold is sharp for $m \geq 2$.*

Proof. By the mixed-difference lemma, the relevant coefficient defects are exactly the even-norm integer vectors with norm at most $M = 2^{t+1}$. They are precisely differences of two 2^t-term sums from V_m; central symmetry and opposite-pair padding prove both directions. Lemma A.19 proves necessity as well as sufficiency on actual words. For a nonzero relation among powers of B, take its largest nonzero index j. The lower terms have absolute sum at most $(M - |c_j|)B^{j-1}$, whereas the leading term is at least $|c_j|B^j$. The latter is strictly larger for $B \geq M$. At $B = M - 1$, the nonzero relation $Be_1 - e_2$ has norm M and even parity, so realization produces an accidental zero. □

Proposition A.23 (Irredundant additive wall list)**.** *Put $M = 2^{t+1}$. Up to sign, the complete distinct additive walls are $c \cdot \beta = 0$, where $c \in \mathbb{Z}^m$ is primitive and nonzero, with*

$$\|c\|_1 \leq M \text{ if } \|c\|_1 \text{ is even}; \qquad \|c\|_1 \leq M/2 \text{ if } \|c\|_1 \text{ is odd}.$$

A wall meets $0 < \beta_1 < \cdots < \beta_m$ exactly when the suffix sums $\sum_{i=j}^m c_i$ take both positive and negative values. After this filter the distinct wall list is irredundant for the additive failure locus in that cone.

Proof. The least even-norm multiple of a primitive normal is itself or twice itself according to parity. Write $\beta_i = \delta_1 + \cdots + \delta_i$, with all $\delta_j > 0$. Then $c \cdot \beta = \sum_j \delta_j \sum_{i=j}^m c_i$. This has a zero in the open positive cone precisely under the stated mixed-sign condition. A distinct intersecting hyperplane has relatively open points not covered by the finitely many other hyperplanes. The rational structure also supplies rational points there. □

The equivalence to the cyclotomic embedding floor here assumes rational magnitudes. The additive arrangement itself is defined for real magnitudes, but it must not be substituted for all algebraic cyclotomic equalities.

A complete three-magnitude instance. For Freiman order 2 and $0 < x < y < z$, the walls are exactly

$$y = 3x, \quad z = 3x, \quad z = 3y, \quad 2y = x + z, \quad z = 2x + y, \quad z = x + 2y.$$

Normalize $x = 1$, $u = y/x > 1$, $v = z/x > u$. Their interior intersections are $(2, 3), (3, 5), (3, 7), (3, 9)$, of multiplicities $2, 3, 2, 2$. Adding lines inside the convex cone gives

$$1 + 6 + \big[(2 - 1) + (3 - 1) + (2 - 1) + (2 - 1)\big] = 12$$

open chambers. There are $6 + (2 + 3 + 2 + 2) = 15$ one-dimensional faces and four vertices. This is an exact additive-floor arrangement (and, by the preceding theorem, the rational lower-floor arrangement at $k = 2p, q = 6$), not the full terminal arrangement.

A.9 Exact embedding counts with two odd primes

Let $n(d)$ be the number of ordered pairs in V_m^2 with difference d. For structural counting retain vector d. For rational weights aggregate multiplicities after scalar projection.

Theorem A.24 (Rectangular counting formula)**.** *At $k = 2p\ell$, with two distinct odd primes, the corresponding number of zero words is*

$$Z_{p,\ell}(n) = \sum_{(b_1,\ldots,b_\ell)\in D_m^\ell} \left(\prod_{j=1}^{\ell} n(b_j)\right) \left(\sum_{a\in D_m - D_m} \prod_{j=1}^{\ell} n(a+b_j)\right)^{p-1}, \tag{C}$$

where n is extended by zero outside its support. For scalar counting the alphabet and difference set in (C) are their projected versions. At $k = 2^b p^a \ell^c$ the count is the $(k/(2p\ell))$th power of (C).

Proof. The two-dimensional anchored-difference kernel consists exactly of arrays $c_{ij} = a_i + b_j$. Fix the gauge $a_1 = 0$, so b is the actual first row and is unique. Each other row independently chooses its unique translating value a_i; its number of letter-pair realizations is the product of entry multiplicities. Any feasible shift lies in $D_m - D_m$. Multiply row weights and sum. The rational prime-power block decomposition is independent in each of the $k/(2p\ell)$ position classes, giving the power. □

For DNA at $k = 30$ this gives:

Ratio	Total embedding-zero words	Accidental words
Structural floor	21,566,097,316	0
5/3	21,573,498,016	7,400,700
2	27,518,030,956	5,951,933,640
3	288,112,407,796	266,546,310,480
5	22,778,390,836	1,212,293,520
7	21,566,157,256	59,940

These values were independently evaluated by first-row enumeration and first-row histograms; no enumeration of 4^{30} words was used.

One odd prime. For DNA at $k = 2p$, structural coefficients must be constant on the prime block. The multiplicities of D are $n(0) = 4$, four axis differences with multiplicity 1, and four diagonals with multiplicity 2. Thus

$$Z_{2p}^{\mathrm{str}} = 4^p + 4 + 4\,2^p, \qquad Z_{2p}(3) = 4^p + 2 + 2\,2^p + 2\,3^p.$$

At $2^b p^a$ raise these to the $2^{b-1}p^{a-1}$th power. At other rational ratios the structural count applies.

PCRs at every even order. Every word of proper period $d \mid k$ has componentwise zero embedding, because the geometric sum across repeats vanishes. Burnside therefore gives, for either structural or fixed-ratio total zero-word count Z_k,

$$\#\{\text{zero PCRs}\} = \frac{1}{k}\left(Z_k + \sum_{d\mid k,\ d<k} \varphi(k/d) q^d\right).$$

The single-odd-prime formula in Theorem A.9 does not apply at $k = 30$; the rectangular formula is its distinct two-prime replacement.

A.10 Twice-prime rational seed equalities

Throughout this and the next section, $k = 2p$ with odd prime p, and $h = (p-1)/2$. The $h+1$ numbers $1, \cos(\pi/p), \ldots, \cos(h\pi/p)$ have precisely the relation

$$1 + 2\sum_{j=1}^{h} (-1)^j \cos(j\pi/p) = 0. \tag{R}$$

Indeed $2\cos(\pi/p)$ has degree h; its powers through degree $h-1$ are a basis, and the cosine polynomials are triangular by degree. Relation (R) follows by pairing terms of $\Phi_{2p}(\zeta_{2p}) = 0$.

Theorem A.25 (Rational raw seed classification)**.** *The accidental ratios for the raw seed-imaginary form are exactly* $\{3\}$*. For the raw seed-real form they are exactly*

$$\{5/3, 2, 3, 5, 7\}.$$

The latter statement concerns raw real equalities, not only negative-ray branch transitions.

Proof. The imaginary form has the same independent prime-sine blocks as the preceding lower-lambda formula, and the two real-axis positions are free. Thus the preceding rational lower classification applies. For the real form, pair the two positions on each cosine and on its negative. With a harmless alternating swap, their coefficient differences C_j are sums of four signed units. Let d be the appropriately signed difference of the letters at exponents $0, p$. Relation (R) makes scalar vanishing equivalent to $\pi(C_j) = 2\pi(d)$ for all j. Each vector defect $C_j - 2d$ has even norm at most 8. A nonzero scalar-killed such vector forces exactly the five ratios of Theorem A.20 with $t = 2$. Conversely the following choices have $C - 2d \neq 0$ and $\pi_\rho(C - 2d) = 0$:

ρ	d	C
$5/3$	$(0,2)$	$(3,-1)$
2	$(0,2)$	$(2,0)$
3	$(0,0)$	$(1,-3)$
5	$(0,2)$	$(1,-1)$
7	$(0,2)$	$(1,-3)$

Every d is a two-letter difference and every C a four-letter signed sum; these choices can be repeated in all h independent blocks. This gives actual words at every odd prime p. □

For reference the exact structural real-form count is

$$Z_{\Re}^{\mathrm{str}} = 4\,36^h + 8\,6^h + 4.$$

The total real-form counts at the five resonances, in increasing order, are

$5/3$	$4\,36^h + 4\,10^h + 4\,6^h + 2\,5^h + 2$
2	$4\,36^h + 4\,22^h + 4\,6^h + 2\,17^h + 2$
3	$4\,44^h + 6\,31^h + 4\,10^h + 2$
5	$4\,36^h + 4\,10^h + 4\,6^h + 2\,25^h + 2$
7	$4\,36^h + 8\,6^h + 2\,5^h + 2.$

A finite-sum proof, avoiding reliance on these expansions, is as follows. Let $N_2(d)$ count ordered letter differences and $N_4(c)$ ordered four-letter signed sums. Then

$$Z_{\Re} = \sum_{d} N_2(d) N_4(2d)^h.$$

Use vector multiplicities for the structural count and scalar multiplicities for a fixed rational ratio. Direct evaluation of the two finite alphabets gives the displayed expansions. Independent positional remainder enumeration checked them at $k = 6, 10$. The imaginary structural count is $16\,36^h$, and its total at 3 is $16\,44^h$; lower counts are four times smaller.

A.11 Complete rational midpoint classification at $2p$

For each $j = 1, \dots, h$, let A_j and B_j be the sums of the two letters at the symmetric positions belonging to $\cos(j\pi/p)$ and $-\cos(j\pi/p)$. Swap the two pairs for odd j, and put

$$C_j = A_j - B_j, \qquad D_j = A_j + B_j.$$

Let $s \in V$ be the central letter at exponent p. Relation (R) gives an exact scalar criterion:

$$\pi(C_j) = c \text{ independent of } j, \qquad c + 2\sum_{j=1}^{h} \pi(D_j) + 4\pi(s) = 0. \tag{M}$$

Up to signed coordinate permutations, the possible common vector C and its D values are the following. Multiplicities count ordered four-letter assignments.

C	D values and multiplicities
$(0,0)$	$(0,0)$:16; $(\pm4,0),(0,\pm4)$:1; $(\pm2,\pm2)$:4.
$(2,0)$	$(\pm2,0),(0,\pm2)$:4.
$(1,1)$	$(1,1),(-1,-1)$:8; $(3,-1),(-1,3),(1,-3),(-3,1)$:2.
$(4,0)$	$(0,0)$:1.
$(3,1)$	$(1,-1),(-1,1)$:2.
$(2,2)$	$(0,0)$:4; $(2,-2),(-2,2)$:1.

This table follows simply by writing $A = (C + D)/2$ and $B = (D - C)/2$ and counting the two signed-unit sums. It is a small exact algebraic identity, not a large-order computational premise.

Define $\mathrm{Rat}(x, y) = \max(|x|, |y|)/\min(|x|, |y|)$ when $xy \neq 0$ and $|x| \neq |y|$, and omit it otherwise. For $h = (p - 1)/2$, define five finite sets:

$$\begin{aligned}
\mathcal{R}_0 &= \{\mathrm{Rat}(a,b) : |a| + |b| \le p,\ |a| + |b| \text{ odd}\},\\
\mathcal{R}_{20} &= \{\mathrm{Rat}(1 + 2a, 2b) : |a| + |b| \le h + 1,\ a + b \equiv h + 1 \pmod 2\},\\
\mathcal{R}_{11} &= \{\mathrm{Rat}(1 + 2u + 4v + 4\epsilon, 1 + 2u - 4v) :\\
&\qquad\qquad u = -h, -h + 2, \dots, h,\ -h \le v \le h,\ \epsilon = \pm1\},\\
\mathcal{R}_{31} &= \{\mathrm{Rat}(7 + 2t, 1 - 2t), \mathrm{Rat}(3 + 2t, 5 - 2t) : t = -h, -h + 2, \dots, h\},\\
\mathcal{R}_{22} &= \{\mathrm{Rat}(3 + 2t, 1 - 2t) : -h \le t \le h\}.
\end{aligned}$$

All parameters are integers; signs and interchange of coordinates are absorbed by Rat.

Theorem A.26 (Midpoint spectrum and active realization)**.** *The complete set of rational accidental midpoint ratios at $k = 2p$ is*

$$\mathcal{R}_p = \mathcal{R}_0 \cup \mathcal{R}_{20} \cup \mathcal{R}_{11} \cup \mathcal{R}_{31} \cup \mathcal{R}_{22}.$$

Every ratio in this set is also realized with $\lambda_s(u) > 0$. The largest ratio is

$$\max \mathcal{R}_p = 2p + 4 - (-1)^{(p-1)/2}.$$

There are exactly four structural midpoint words, and all four have $\lambda = 0$.

Proof. First assume $\rho \notin \mathcal{E}_2$. Equality of the scalar values of two C_j forces equality of their vectors: a nonzero difference is an even-norm vector of norm at most 8, whose possible rational annihilators are exactly $\mathcal{E}_2$. Thus all $C_j = C$. For $C = 0$, the vector $\sum D_j/2 + s$ is any sum of $2h + 1 = p$ signed units, giving $\mathcal{R}_0$. For $C = (2, 0)$, divide the defect in (M) by 2; it becomes $(1 + 2a, 2b)$ where (a, b) is a sum of $h + 1$ units, giving $\mathcal{R}_{20}$. For $C = (1, 1)$, write

$$\sum D_j = u(1, 1) + 2v(1, -1), \quad u = -h, -h + 2, \ldots, h, \quad -h \le v \le h.$$

The choices of u, v are independent by the table. Taking the central letter to be ϵe_1, and using interchange symmetry for ϵe_2, gives $\mathcal{R}_{11}$. For $C = (3, 1)$, the sum of the D values is $t(1, -1)$, with $t \equiv h \pmod 2$ and $|t| \le h$. Two central signs give ratio 1 and are excluded; the other two give $\mathcal{R}_{31}$. For $C = (2, 2)$, the sum is $2t(1, -1)$ for any integer $|t| \le h$, giving $\mathcal{R}_{22}$. The type $(4, 0)$ gives either a structural zero or no ratio greater than one. Every listed choice is reversible into its constituent ordered letter pairs, proving sufficiency on actual words.

The excluded exceptional set $\mathcal{E}_2$ adds nothing. Ratio 2 is in $\mathcal{R}_{20}$: use $(a, b) = (0, 1)$ for even h, and $(-1, 1)$ for odd h. Ratios $3, 5$ are in $\mathcal{R}_{22}$ using $t = 0, 1$. Ratios $5/3, 7$ are in $\mathcal{R}_{31}$ using $t = 0$ for even h and $t = -1$ for odd h. Thus every possible exceptional ratio is already in the proposed union.

For structural vanishing, the vector analogue of (M) holds with common C. Since each coordinate of $C - D_j$ is even, $C + 2\sum D_j + 4s \equiv pC \pmod 4$. Hence each coordinate of C is divisible by four. The type $C = 0$ cannot work: it would make an odd number p of signed units sum to zero. Only $C = \pm 4e_i$ remain, with all $D_j = 0$ and $s = \mp e_i$. They yield exactly four words and each symmetric pair has identical letters, so its lambda contribution is zero.

To make a nonstructural midpoint equality active, it suffices that some symmetric pair has distinct letters. Swapping those two positions leaves C, D, and hence midpoint zero, unchanged; it changes lambda by a nonzero real multiple of their weight difference. At least one of the two words has nonzero lambda; reversal changes its sign while preserving midpoint, yielding positive lambda. Types $20, 11, 31$ have such a pair in every realization. The set $\mathcal{R}_0$ is contained in $\mathcal{R}_{20}$: orient the odd coordinate as $1 + 2a$, and use its two possible signs to obtain the required parity with $|a| + |b| \le h + 1$. For type 22, if $|t| < h$, use at least one $D = 0$ block, whose pairs are distinct. At $t = \pm h$, its ratio is already in $\mathcal{R}_{11}$ with $u = \mp h, v = 0, \epsilon = -1$. Thus all ratios have active witnesses.

Finally, $\mathcal{R}_{11}$ supplies the maximum. Both coordinates are congruent to $1 + 2h$ modulo four, and $x + y = 2 + 4u + 4\epsilon$. If the smaller absolute coordinate is at least 3, the ratio is at most $1 + (4h + 6)/3$. If it is 1, that coordinate must be $+1$ for even h and -1 for odd h. The other coordinate is then bounded in absolute value by $4h + 5$ or $4h + 7$, respectively. These extrema are attained at $u = h$, $\epsilon = 1$, and respectively $v = h/2$ or $(h + 1)/2$. The other four families have bounds at most $p + 6 \le 2p + 3$ (or the smaller $p - 1, p + 1, p + 2$ bounds). This proves the stated maximum. □

An exact counting transfer. For reduced $\rho = a/b$, use integer scalar letters $S = \{\pm a, \pm b\}$. Put

$$T_c(z) = \sum_{\substack{x_1, x_2, y_1, y_2 \in S \\ x_1 + x_2 - y_1 - y_2 = c}} z^{2(x_1 + x_2 + y_1 + y_2)}.$$

Then the total midpoint count is exactly

$$Z_r(p, \rho) = \sum_c \sum_{s \in S} [z^{-c-4s}] T_c(z)^h. \tag{T}$$

For the joint midpoint/lambda-zero count, restrict each summand to $x_1 - x_2 + y_1 - y_2 = 0$. Prime-sine independence makes this restriction necessary and sufficient blockwise for rational weights. Reversal pairs the nonzero lambda values, so the active midpoint count is half the difference between the total and joint counts. RC pairs active midpoint ties and selects exactly half of them in the lexicographic ideal.

At $p = 3$, the spectrum is

$$7/5,\ 3/2,\ 5/3,\ 2,\ 7/3,\ 3,\ 4,\ 5,\ 7,\ 9,\ 11,$$

matching the complete $k = 6$ critical list. The numbers of midpoint ratios at $p = 5, 7, 11, 13, 17, 19$ are $20, 30, 62, 77, 125, 147$. These numbers are finite evaluations of the proven parametrization, not an asserted closed cardinality formula.

Remark A.27 (Do not remove structural midpoint ties at general order)**.** *At $k = 8$, the lower word $TTTAATA$ has componentwise midpoint zero but $\lambda_s = 2\sqrt{2}\rho > 0$. Therefore the twice-prime structural-tie inactivity statement must not be extended to all even orders. The general two-family theorem retains these ties.*

A.12 Every real cyclotomic ratio is eventually realizable

Theorem A.28 (Eventual positional realization)**.** *A real algebraic number $\rho > 1$ is realized as an accidental DNA embedding ratio at some even order if and only if it belongs to a cyclotomic field. For every representation $\rho = U(\zeta_n)/d$, with even $n \geq 4$, $d \in \mathbb{Z}_{>0}$ and $U \in \mathbb{Z}[X]$ reduced to degree less than n, there is an explicit one-hot construction at*

$$k = n\ell_1 \cdots \ell_t,$$

where the ℓ_i are fresh distinct odd primes and $2^t \geq \max_e \|(2d\,\mathbf{1}_{e=0}, -2U_e)\|_1$.

Proof. Necessity is the preceding field gate. For sufficiency take $t \geq 1$ satisfying the bound. For each $e = 0, \ldots, n-1$, the vector $c_e = (2d\mathbf{1}_{e=0}, -2U_e)$ has even norm at most 2^t, so express it as a sum of exactly 2^t signed units, padding with opposite pairs. Place these units, with alternating signs, at the corners of a binary t-cube. Extend to the full $\ell_1 \times \cdots \times \ell_t$ grid by collapsing all nonanchor values to one corner index, exactly as in Lemma A.19, but using individual letters rather than letter pairs. The prime-root sum in each new coordinate reduces this array to a nonzero factor Γ times c_e. CRT identifies all $(e, j_1, \ldots, j_t)$ with distinct positions in the length-k word; every position receives one letter. The resulting component embeddings are

$$H = 2d\Gamma \neq 0, \qquad L = -2\Gamma U(\zeta_n),$$

so $\rho H + L = 0$ and the zero is accidental. Explicitly one may take the exponent $\frac{k}{n}e + \sum_i \frac{k}{\ell_i} j_i \pmod{k}$; all coefficients of this CRT map are units in their own factors. No abstract coefficient is left without a positional realization. □

The cyclotomic-field gate is therefore sufficient *after allowing the order to grow*, although it remains insufficient at a fixed order. Examples verified by exact polynomial division are $\rho = \sqrt{2}$, base $n = 8$, fresh prime 3, giving $k = 24$; and $\rho = (7 + 3\sqrt{5})/2$, base $n = 10$, fresh primes $3, 7, 11$, giving $k = 2310$. These are explicit upper bounds, not minimum realizing orders.

Corollary A.29 (Eventual raw realization for the other roles)**.** *Over all even orders, the same necessary and sufficient cyclotomic-field condition holds for the raw real, imaginary, lower-lambda and midpoint equality families.*

Proof. For the real and imaginary roles, rotate an embedding witness until the relevant component of its nonzero H is nonzero. Such rotations exist because $k \geq 4$. Delete the exponent-zero letter for a lower-lambda witness; its imaginary contribution was zero. For midpoint, interleave the embedding witness with its letterwise complement at order $2k$. This gives component weight sum zero and new component embedding $(1 - \zeta_{2k})H \neq 0$, while its scalar embedding is zero. Rotate to make its real high component nonzero and delete the exponent-zero letter; the midpoint form does not see that position. Necessity follows by solving the corresponding nonstructural linear relation in the real cyclotomic coefficient field, after the usual positive sine normalization for lambda. □

The midpoint witnesses in this corollary have scalar lambda zero. It does *not* establish eventual active midpoint realization for every algebraic ratio. Fixed-order realization, efficient/minimal conductors, and active branch classification remain separate problems.

Uniform radical designs. The preceding uniform magnitude family is linearly independent over every cyclotomic coefficient field. Since all catalogue forms, not just embedding and normalized lambda, have coefficients in such fields, the same componentwise argument makes every zero in the entire catalogue structural. This does not remove structurally active midpoint ties such as the $k = 8$ example.

A.13 Residual geometry as a three-block problem

The imported construction proves strict increase of the midpoint potential along arcs between positive-lambda states [2]. The following block interface and max-plus formula are deductions, not a new claim to that monotonicity lemma. Set

$$N = \{u : \lambda_s(u) \leq 0\},$$
$$P_0 = \{u : \lambda_s(u) > 0, a_s(u) = 0\}, \qquad P_1 = \{u : \lambda_s(u) > 0, a_s(u) = 1\}.$$

Theorem A.30 (Three-block residual interface)**.** *There is no original lower-graph arc from P_0 to P_1. Every residual arc respects the order*

$$P_0 \longrightarrow N \longrightarrow P_1.$$

The positive blocks contain all their original internal arcs. The core is the original graph induced on N with only the zero-PCR seed arcs deleted. RC swaps the two positive blocks, reverses residual arcs and preserves the core. Let L_P be either positive-block longest path, L_N the core longest path, and let $d_P(u)$ be the longest P_0-internal path ending at u. Define

$$\eta(n) = \max\big(\{0\} \cup \{d_P(u) + 1 : u \in P_0,\ u \to n\}\big).$$

Then the exact residual longest path is

$$\boxed{L(D_s) = \max\left\{L_P,\ \max_{n_0 \to \cdots \to n_j \text{ in } N}[\eta(n_0) + j + \eta(\mathrm{RC}(n_j))]\right\}.} \tag{P}$$

In particular $L_N \leq L(D_s) \leq 2L_P + L_N + 2$.

Proof. For any lower arc $u \to v$,

$$r_s(v) - r_s(u) = \tan(\pi/k)(\lambda_s(u) + \lambda_s(v)).$$

This follows by writing $P = x + iy$ in $(\zeta_k^{-1} - 1)P$. If both lambdas are positive, the increment is strictly positive. A P_0 state has $r \geq 0$, and a P_1 state has $r \leq 0$, so $P_0 \to P_1$ is impossible. The two-family formula now checks all block types: $N \to P_0$, $P_1 \to N$ and $P_1 \to P_0$ have terminal indicator one; $P_0 \to N$ and $N \to P_1$ have indicator zero. On each positive block it is zero, and on N it is the zero-PCR seed indicator. The imported RC invariance gives the stated anti-isomorphism. A residual path either lies inside a positive block or decomposes into a P_0 prefix, one core path, and a P_1 suffix; either end may be empty. For a fixed core path the longest possible prefix is $\eta(n_0)$. The RC image of a longest incoming prefix at $\mathrm{RC}(n_j)$ is a longest outgoing suffix at n_j, with length $\eta(\mathrm{RC}(n_j))$. These paths lie in disjoint blocks, so they concatenate without collision. Maximizing proves (P). The inequalities follow from $\eta \leq L_P + 1$. □

Formula (P) can be evaluated by a weighted DAG dynamic program on the core after computing the two boundary functions; it does not claim subexponential enumeration of all lower states.

k	Weight or interval	L_P	L_N	$L(D_s)$
4	any atlas cell	2	11	17
6	first atlas interval	9	43	63
6	$\rho = 3$	8	47	65
6	$3 < \rho < 4$	9	39	59
6	$4 < \rho < 5$	8	39	57
8	$\rho = 2$	33	79	147
8	$\rho = (7 + 3\sqrt{5})/2$	28	85	143
8	$\rho = \sqrt{13}$	27	85	141
8	$\rho = \sqrt[3]{2}$	33	93	161

All 199 certified strata at $k = 4, 6, 8$ satisfy equality in the upper bound, but we assert only the exact formula (P) and the inequality at arbitrary order. An all-order equality claim would need an additional argument aligning extremal prefixes, core paths, and suffixes.

A.14 Minimal root locations, positivity, and finite optimization

Proposition A.31 (Minimal one-parameter cut set)**.** *Let a complete finite affine decision catalogue on an open interval be given, and evaluate the deterministic output on every point and interval cell. A root is indispensable to a point/interval partition exactly when the outputs on its left interval, at the point, and on its right interval are not all equal. Keeping exactly these roots gives the unique minimum set of root locations for such a partition. With E indispensable roots it has $2E + 1$ cells; the number of distinct outputs can be smaller.*

Proof. If the three outputs agree, removing that root merges cells with identical output. Conversely, without an indispensable root the output would have to be constant in a neighborhood containing all three local cells, a contradiction. Finite repeated deletion cannot change this local necessity. The count is E points and $E + 1$ open intervals. A point-only anomaly must therefore be retained even when its adjacent outputs agree. □

Applied to the complete small atlases: $k = 4$ has one indispensable root, three cells and two distinct terminals; $k = 6$ has eleven indispensable roots, twenty-three cells and seventeen terminals. This is not a general solution for minimum forms in higher-dimensional weight space. The complete $k = 8$ atlas is given in Theorem 4.1.

Proposition A.32 (Two barriers to terminal Galois transport)**.** *A field automorphism of a real number field preserving positivity of every element is the identity. A fixed coordinate permutation on words over an alphabet with at least two letters preserves all directed de Bruijn arcs only when it is the identity; it reverses all arcs only when it is position reversal.*

Proof. If $x \neq \sigma(x)$, a rational strictly between them supplies an element whose sign changes under σ. For the graph statement, write the image of a word as $(w_{\pi(0)}, \dots, w_{\pi(k-1)})$. In an arc $w \to (w_1, \dots, w_{k-1}, a)$, the free appended letter cannot enter an overlap constraint. Those constraints then force $\pi(j+1) = \pi(j) + 1$ for all j, hence the identity. Compose an arc-reversing permutation with position reversal to obtain the first case. □

These barriers do not exclude a finite set of certificates retaining positivity, accidental terminal coincidences, or alphabet symmetries. They explain why equality-level Galois covariance alone cannot be promoted to a terminal orbit classification.

An explicit, bounded optimization result. Choose the projective angular coordinate $\theta = \arctan \rho \in (\pi/4, \pi/2)$, and define robustness as distance to all critical ratios and the two domain boundaries in this coordinate. First minimize residual longest path in the complete atlas, then maximize robustness inside an optimal open cell. At $k = 4$, the widest optimal cell is $(1, 3)$; its angular midpoint is

$$\rho = (1 + \sqrt{5})/2.$$

At $k = 6$, the minimum is $L(D) = 57$ for $\rho \geq 4$. Among the optimal open cells, $(11, \infty)$ is angularly widest, giving

$$\rho = 11 + \sqrt{122}.$$

For example its width is $\arctan(1/11)$; the other finite optimal-cell widths are $\arctan(1/21)$, $\arctan(1/18)$, $\arctan(1/32)$, and $\arctan(1/50)$, all smaller. This is an exact answer for a declared metric and two complete small atlases, not a universal or biologically optimal weight. The first ratio is the ordinary golden ratio, not the earlier landmark φ^4.

B Technical supplement: the earlier length-131 escape

B.1 An earlier length-131 escape construction

This retained construction documents one explicit route out of the weighted components. Its upper bound is superseded by Theorem 4.4.

For a lower word u and a nonempty proper letter set B, consider the simultaneous cochain change

$$\Delta_{u,B} = \sum_{a \in B} (\mathbf{1}_{ua} - \mathbf{1}_{au}) + J\left(\sum_{a \in B} (\mathbf{1}_{ua} - \mathbf{1}_{au}) \right).$$

It preserves RC and the sum on every PCR. It does not by itself guarantee Boolean indicators or acyclicity; both must be verified.

Proposition B.1 (Four changes of component)**.** *Start with the weighted terminal M_0 for $\rho = 49/16$ (row `k8_c094`). Apply the four changes in Table 2 in order. All five endpoints are RC MDS of size 8230, with*

$$\boxed{L(M_0), \dots, L(M_4) = 141, 135, 135, 133, 131.}$$

They occupy five distinct ordinary components and each attains the optimum in its own component. The last component is distinct from all 21 in Table 1.

Exact certificates. Direct reconstruction checks all 4^8 membership bits after each patch, and checks RC, one word per PCR and acyclicity. Integer ranks and explicit longest paths give the upper and attained bounds. One-hit simple cycles of lengths $142, 136, 136, 134, 132$ give the matching component lower bounds. Ten cycle separators distinguish all pairs of endpoints; another 21 distinguish the last endpoint from the weighted components. Each separator has at most 15 arcs. The separate checker tests these claims without importing the discovery search. □

Step	u	B	Changed words	L after
1	`ATCGGCG`	$\{T\}$	4	135
2	`ATTACTG`	$\{G, T\}$	8	135
3	`CCGCCCG`	$\{C\}$	4	133
4	`GTTCTTG`	$\{T\}$	4	131

Table 2: Explicit simultaneous changes; the neutral second step changes component. Its necessity or minimal number of moves is not claimed.

The four changed supports are disjoint. A compact final membership rule is

$$\mathbf{1}_{M_4}(w) = m_{49/16}(w) - \mathbf{1}_D(w) + \mathbf{1}_A(w),$$

where the following columns list D and A; rows are not canonical pairings.

Remove (D)	Add (A)
`CAAGAACA`	`ACAAGAAC`
`CAGTAATA`	`ACAGTAAT`
`CAGTAATC`	`ACGCCGAT`
`CCCGCCCG`	`ATCGGCGT`
`CGCCGATA`	`ATTACTGG`
`CGGGCGGG`	`ATTACTGT`
`GATTACTG`	`CCAGTAAT`
`TATCGGCG`	`CCGCCCGC`
`TATTACTG`	`GCGGGCGG`
`TGTTCTTG`	`GTTCTTGT`

Corollary B.2 (Global upper bound, exact component optimum)**.** *For RC DNA MDS at $k = 8$,*

$$\min_{M \text{ RC MDS}} L(M) \le 131, \qquad \min_{N \in \mathcal{C}_F(M_4)} L(N) = 131.$$

The final upper residual has length 130 in edges. Its longest avoiding word has 138 letters, so every DNA word of length 139 contains a selected eight-letter window.

The final one-hit cycle is the cyclic concatenation of these lines:

```
CGCGGAGGAGCCGACCCCACCACACAACAAAAAAACAACACAAGAAGACC
AGCAGCCGAGGAGGATAGGCCTATCCTCCTCGGCTGCTGGTCTTCTTGTG
TTGTTTTTTTGTTCTTGGGTGCTGGGCTGATG
```

Its length is 132, its seven-letter prefixes are distinct, and its only selected cyclic eight-letter window is `GCGCGGAG` at zero-based index 131. Deleting that arc leaves 131 residual arcs. The machine-readable string in `cycle_132_compact.json` is checked against the cycle certificate. Any improvement to $L \le 130$ must hit this cycle at least twice, and therefore leave this component.

B.1.1 Reproducibility and evidence boundary

The public package contains the original 30-row regression capsule, the complete 199-row atlas, mixed-wall and component certificates, and the retiming construction. The atlas checks 11,432,704 word decisions. The retiming checker verifies 72 certificate occurrences (not 72 distinct sets), 42 retimings, 30 lower-bound cycles and 31 short separators. Its 24 semantic mutations include missing lower bounds and a duplicated parity sector. The atlas and wall checkers have 18 and 15 further mutation controls respectively. Small exhaustive tests cover $(q, k) = (2, 2), (2, 4), (4, 2)$ and 64 RC-feasibility queries, plus 260 bounded autodual ranks.

All finite outputs have a declared scope. A rank certificate proves a universal path bound for its finite DAG; a completed list proves a complete atlas only because the affine-root reduction is exhaustive. None of these checks proves an all-order theorem by sampling. Distinct implementations remain AI-assisted checks, not independent human acceptance. No timeout or incomplete search is used to claim that a shorter global construction is impossible.

C Technical supplement: local exclusion domains

C.1 What the finite exclusions do and do not prove

A PCR group below is either a stable PCR or a pair of nonstable PCRs interchanged by RC. A chart site is a lower-vertex pair $\{u, \theta u\}$ in Theorem 5.9; it is not a PCR group or a legal move. Changing these definitions changes a neighbourhood and can invalidate a purported exclusion.

Proposition C.1 (Exact local exclusions)**.** *The following finite statements hold for the exact packaged bases.*

1. *In the 99-PCR region generated by the displayed critical cycle of M_{121} and its RC image, with the exterior fixed, the minimum RC MDS length is 121.*
2. *Changing at most four RC-PCR groups of M_{111} cannot produce an RC MDS of length at most 110.*
3. *Changing at most three RC-PCR groups of M_{109} and then applying arbitrary integral retiming cannot attain length at most 108. The minimum in that family is 109. Retiming the base before measuring the radius is not included.*
4. *In the chart based at the canonical M_{107}, allowing at most three sites, arbitrary integral site coefficients, all 128 stable bits and an arbitrary integral gradient gives minimum RC MDS length 107.*
5. *In each of eight specified regions around S of Proposition 5.11, with respectively $1, 13, 139, 64, 128, 256, 512, 1024$ nonstable groups free and all stable choices unrestricted, the minimum cardinality at path bound 106 is 8232. These are eight specified regions, not all regions of those sizes.*

Exact finite proof objects. For the first statement the verifier reconstructs 51 RC-compatible option domains with 382 initial options and checks 53 walk cuts, 202 option eliminations and a terminal contradiction. No numerical solver status enters this implication chain.

For the four-group radius, 2486 path or cycle cuts support an exhaustive branching certificate with 533,138 nodes. Each node has its declared domain and cut bound recomputed, so complete branch coverage supplies the exclusion. A two-group change repairs the

old critical cycles but creates a seven-cycle, showing why PCR and critical-cycle conditions alone are insufficient. For the three-group radius around M_{109}, the 1069 cuts are all closed walks; their weights are unchanged by subsequent retiming. The 45,845-node proof thus excludes the full retiming closure of the stated radius. A separate acyclic three-group change hits each of the 6,931,456 old critical cycles three times but has component optimum 111. The exhaustive old-cycle count uses a tight-clock DAG and is independently replayed.

For three chart sites, globally checked support conditions reduce the $\binom{8192}{3} = 91,592,417,280$ triples to 24,592 necessary ones. Lifting a rational certificate to an omitted site requires both vanishing of its combined coefficients and the divisibility conditions used in row normalization. The verifier checks these conditions before using a lifted support clause. It reconstructs the complete support recursion, then checks every surviving triple by rational Farkas combinations, integer strips, modular equalities or rounded valid inequalities. There are no unresolved triples. The 1323 closed walks and all rational coefficients are included; the checker uses integer and rational arithmetic only.

For the eight regions, deleting all 988 stable-PCR arcs is a valid relaxation. The verifier checks 844 sequential eliminations of nonstable options using residual cycles or 107-arc paths. An option is removed only using already proved residual arcs or fixed exterior arcs. Every region ends with an empty domain if all its nonstable PCRs are singly hit. Cost at most 8231 has this property; the base S attains 8232 in each region. □

The finite cut relaxation for the four-group radius has exact fractional minimum $1/2$, certified by a feasible rational point and matching dual multipliers. The corresponding integer cut system requires at least five changed groups, hence its gap is at least ten. This concerns that particular relaxation. Adding more valid inequalities or changing the objective changes the statement; it is not a universal integrality-gap theorem for the global problem.

C.1.1 A boundary clause needs its assumptions

Each regional elimination proof supplies a set B of exterior RC literals and a set $\mathcal{T}$ of nonstable PCR representatives whose one-selection property the proof used. Writing $c_P = |M \cap P|$, its valid extension to any RC feedback set at the target length is

$$\sum_{b \in B} \mathbf{1}_M(b) + \sum_{P \in \mathcal{T}} (c_P - 1) \geq 1. \tag{47}$$

Indeed, if the nonnegative integral left side were zero, all boundary arcs would remain residual and all used PCRs would be singly hit, so the identical inductive contradiction would apply. At cost at most 8231 the nonstable surplus vanishes, and the bare boundary clause is valid. At higher cost it need not vanish. This surplus is different from the stable-PCR surplus in Corollary 5.6.

The shortest supplied boundary clause has 21 RC literals and uses 12 group assumptions and eight cycles. Both S_A, S_B satisfy all eight bare clauses while moving the double-PCR surplus to a different nonstable group. Thus these clauses are useful necessary constraints, not global refutations. The companion also checks 214 globally valid closed-walk cuts at length 106, of which 187 require one hit and 27 require two. Their validity does not imply their sufficiency.

C.1.2 Reproducibility and scope

The companion retains the literal inputs and exact proof objects for these local exclusions. Its separate base-four checks recompute the positive endpoints and sharp parity example;

small-instance tests exercise the conditional stable-completion and lifted two-arm implementations. Such checks support implementations; the uniform results depend on their proofs. Each exclusion is tied to the stated base, exterior, neighbourhood or affine family. None certifies the missing global optimum.

D Technical supplement: wall proofs, regressions, and finite evidence

D.1 An exact bounded DNA chamber atlas

We retain the original thirty-row regression result alongside the complete order-eight classification in Theorem 4.1. It uses the DNA family (14), the complete catalog of Definition 2.4, exact algebraic sign evaluation, and the deterministic selector of Theorem 2.1. The range is exactly thirty rows: every interval and critical point for $k = 4, 6$, and four named weights for $k = 8$.

The landmarks are

$rat2$	2
$gold5$	$(7 + 3\sqrt{5})/2$
$quad13$	$\sqrt{13}$
$cubic2$	$\sqrt[3]{2}.$

Proposition D.1 (Exact small DNA chamber atlas)**.** *Under the conventions above:*

1. *The complete $k = 4$ diagram has three cells and two terminal/residual classes. Its sole critical ratio is $\rho = 3$; the class for $1 < \rho \leq 3$ differs in sixteen selected words from the class for $\rho > 3$. Both have longest residual path* 17.

2. *The complete $k = 6$ diagram has the eleven ordered critical ratios*

$$\frac{7}{5}, \frac{3}{2}, \frac{5}{3}, 2, \frac{7}{3}, 3, 4, 5, 7, 9, 11,$$

   hence twenty-three point/interval cells and seventeen terminal/residual classes. Its longest-path phase is

$1 < \rho \leq 7/3$	63
$7/3 < \rho \leq 3$	65
$3 < \rho < 4$	59
$\rho \geq 4$	57.

   Except at $\rho = 3$, every cell has 144 *lower zeros and* 220 *physical firings; at $\rho = 3$ the counts are* 176 *and* 212.

3. *At $k = 8$, the four declared landmarks give four distinct terminal sets, midpoint ideals, and residual digraphs, although all four have* 256 *zero-embedding words,* 70 *zero PCRs,* 576 *lower-zero states,* 3952 *firings, and minimum cardinality* 8230*. Their longest paths are respectively* $147, 143, 141, 161$ *in the landmark order above.*

4. *Every unequal pair in the frozen landmark comparisons differs outside zero-PCR tie representatives. No equal complete covector produced two different terminal payloads.*

Computer-assisted proof. All thirty cells were generated with exact arithmetic over rational and represented real algebraic numbers. For each row, the producer enumerated all labelled words, constructed the seed and midpoint ideal, formed the direct terminal, and checked literal reverse-complement invariance, one representative per PCR, and acyclicity. It then computed the labelled residual digraph and longest path.

An independent checker reconstructed the complete decision covector, terminal, legal firing replay, residual arcs, topological order, path histogram, and all pairwise comparisons. Two sequential producer runs and normal/optimized checker runs had identical canonical scientific payloads. The assurance layer rejected twenty-nine typed semantic mutations, including reversed gradients, omitted decision forms, floating zero tests, changed tie rules, incomplete cell sets, and altered path statistics. The canonical producer and validator payload hashes are recorded in Section D.5. □

This regression proposition alone does not prove the complete $k = 8$ diagram; Theorem 4.1 supplies that extension. The incompatibility of optimal weights is resolved in Proposition D.6. In fact the $k = 6$ landmark ranking favors gold5, whereas among the four $k = 8$ controls quad13 is shorter. Equal zero counts can coexist with thousands of selected-word changes, showing why zero-sector extremality and residual geometry are distinct objectives.

D.2 Wall geometry and the complete DNA atlas

Write $\theta = \mathrm{RC}$ on lower vertices, $J = \mathrm{RC}$ on arcs, and $\alpha = 2\pi/k$. We use the gradient convention (2). The basic identity, also used in Theorem A.30, is

$$r_s(v) - r_s(u) = \tan(\pi/k)\big(\lambda_s(u) + \lambda_s(v)\big) \quad (u \to v). \tag{48}$$

D.2.1 Regular midpoint walls

Consider an admissible affine weight path $s_t = s_0 + t\dot{s}$. Call its wall at zero *regular* if every lambda form either vanishes identically on this path or is nonzero at zero. Restrict to a sufficiently small neighborhood with no other roots. Let

$$S = \{u : \lambda_{s_0}(u) > 0,\ r_{s_0}(u) = 0,\ \dot{r}(u) \neq 0\}.$$

Subscripts $-, 0, +$ denote the two exact side germs and the point.

Theorem D.2 (Regular wall locality)**.** *The set S is RC-stable and no two of its vertices are adjacent. The seed is constant and*

$$m_+ - m_- = \nabla f, \qquad f(u) = -\operatorname{sgn}\dot{r}(u) \quad (u \in S), \quad f = 0 \text{ otherwise.}$$

Thus $|M_+ \triangle M_-| = 2q|S|$. Between any two of the three germs the same formula holds using only the vertices whose midpoint indicator changes. Their three lower residual lengths lie in an interval of width at most two.

Proof. The two-family formula makes the seed constant. Only vertices of S can change a_s, with the stated signs. Identity (48) rules out adjacent vertices of S; it also rules out loops there. The incident stars are disjoint. On each star both endpoint indicators are Boolean, so the change exchanges all q incoming and q outgoing arcs. In each residual a changing vertex is a source or a sink. A path meets these vertices only at its endpoints. Deleting them leaves a path in the same induced residual outside S; hence every length lies between the longest path there and that length plus two. Finally $\lambda(\theta u) = \lambda(u)$, $r(\theta u) = -r(u)$ and $a(u) + a(\theta u) = 1$ on S, including the frozen point tie. This proves RC stability and all point comparisons. □

The bound two is attained at $k = 8$, $\rho = 4\sqrt{2} - 1$, where the lengths are $(145, 143, 143)$. It does not apply to a mixed wall.

D.2.2 Singular vertices and mixed walls

Put $Z = \{z : \lambda_{s_0}(z) = 0\}$, including identically zero forms. For a letter c define

$$x_c(z) = \Re P_{s_0}(zc).$$

If y_c is the tail of cz and v_c the head of zc, rotation gives

$$\lambda(y_c) = x_c(z)\sin\alpha, \qquad \lambda(v_c) = -x_c(z)\sin\alpha. \tag{W1}$$

Lemma D.3 (Singular subgraph)**.** *The original graph induced on Z is a disjoint union of isolated vertices and zero-embedding PCR cycles. Each cycle has length dividing k, and its midpoint potential is constant, equal to $-W_s(w)$. A zero PCR with nonzero embedding derivative along the weight path has period exactly k.*

Proof. Since $x_c(z)$ is a fixed real number plus $s_0(c)$, injectivity allows at most one letter with $x_c(z) = 0$. Equation (W1) shows that this letter gives both the unique incoming and the unique outgoing internal arc. Thus internal indegree and outdegree are either both zero or both one. Rotation of a zero embedding supplies the successor, so the cycles are precisely the zero PCRs, without tails. Their total letter weight is rotation invariant, proving the potential formula. A word of proper period $d \mid k$ has embedding zero for every weight, by the geometric sum across repeats; its derivative cannot be nonzero. □

For any two germs $\sigma, \tau \in \{-, 0, +\}$, put $\Delta a = a_\tau - a_\sigma$ and $\Delta h = \mathbf{1}_{\lambda_\tau > 0} - \mathbf{1}_{\lambda_\sigma > 0}$. For $x_c(z) \neq 0$ define

$$\gamma_c(z) = \Delta a(z) - \Delta h(z)\mathbf{1}_{x_c(z)<0}.$$

Let $e_\tau(C)$ be the unique arc of a singular PCR C selected by terminal τ. Keep the definition of S from the regular case.

Theorem D.4 (Disjoint mixed-wall formula)**.** *One has the exact cochain identity*

$$m_\tau - m_\sigma = \nabla(\mathbf{1}_S \Delta a) + \sum_{\substack{z\in Z,\ c\in\mathcal{A} \\ x_c(z)\neq 0}} \gamma_c(z)(\mathbf{1}_{zc} - \mathbf{1}_{cz}) + \sum_{C\subset G[Z]} (\mathbf{1}_{e_\tau(C)} - \mathbf{1}_{e_\sigma(C)}). \tag{49}$$

All nonzero summands have disjoint supports, except for the incoming/outgoing arcs already grouped into one star. In particular,

$$|M_\tau \triangle M_\sigma| = 2q\, s_{\sigma\tau} + 2e_{\sigma\tau} + 2c_{\sigma\tau},$$

where s counts changed midpoint vertices in S, $e = \sum_{z,c} |\gamma_c(z)|$, and c counts singular PCRs whose selected representative changes.

Proof. Outside Z, only midpoint indicators at S change. On zc with $x_c(z) \neq 0$, the seed is $(1 - h_\tau(z))h(v_c)$. After separating a possible midpoint change at v_c, its difference is $\Delta a(z) - \Delta h(z)h(v_c) = \gamma_c(z)$ by (W1). The companion cz has the negative difference. The inequalities $0 \le a_\tau \le h_\tau \le 1$ give $\gamma_c(z) \in \{-1, 0, 1\}$.

For the boundary check, if $u \in S$ and $u \to z$ with $z \in Z$, (48) gives $r(z) > 0$, so $a_\tau(z) = 0$ in all germs and the singular contribution on that arc vanishes. If $z \to u$, then $r(z) < 0$, so $a_\tau(z) = h_\tau(z)$ and the same conclusion holds. Thus nonzero singular contributions do not overlap midpoint stars. Each external arc has only one endpoint in Z; a companion coincidence is a homopolymer loop and has $x_c(z) = 0$, hence is excluded. Internal arcs belong to the disjoint PCRs of Lemma D.3, each with one selected arc. All other arcs are unchanged. Summing absolute values proves the count. □

Freezing a_τ to a_0 on S performs commuting complete source/sink moves and gives an RC MDS within two length units of each original germ. This does not bound the difference between distinct singular germs. Nor does (49) assert that its partial companion exchanges are individually legal moves: Boolean membership and acyclicity must be checked on any proposed intermediate state.

D.3 Ancillary atlas properties and wall examples

Proposition D.5 (Embedding spectrum in the quadratic slice)**.** *At $k = 8$ the accidental DNA embedding ratios are*

$$\sqrt{2}, \quad 2\sqrt{2}-1, \quad 1+\sqrt{2}, \quad 1+2\sqrt{2}, \quad 3+2\sqrt{2}.$$

Their total zero-word counts are respectively $288, 352, 648, 352, 576$*, including 256 structural words. At $k = 2^b$, $b \geq 3$, restricted to $\rho \in \mathbb{Q}(\sqrt{2})$, the same list is complete and*

$$Z_{2^b}(\rho) = Z_8(\rho)^{2^{b-3}}.$$

Proof. Antipodal pairing reduces the embedding to four entries of D. The exact enumeration of 9^4 tuples, with ordered-letter multiplicities, checks both real and imaginary coordinates and agrees with all 4^8 positional words. If a zero is accidental, at least one nonstructural coordinate is nonzero as a form, so its ratio lies in $\mathbb{Q}(\sqrt{2})$; no additional real ratio is omitted. For the lift, $\{1, \zeta_{2^b}, \ldots, \zeta_{2^b}^{r-1}\}$ is a basis over $\mathbb{Q}(\zeta_8)$, where $r = 2^{b-3}$. Separating exponents modulo r gives r independent eight-position words. Scalar and componentwise vanishing hold exactly block by block, proving the spectrum and count. Other subfields at higher order are not classified by this argument. □

Proposition D.6 (Exact family optima and a failed Galois quotient)**.** *For the fixed selector,*

$$\min_{\rho>1} L_6(\rho) = 57, \quad \operatorname{argmin} L_6 = [4, \infty),$$

$$\min_{\rho>1} L_8(\rho) = 141, \quad \operatorname{argmin} L_8 = [3, 1+2\sqrt{2}).$$

Consequently no fixed ratio minimizes this family at every even order. Moreover, Galois conjugation does not define an operation on terminal equivalence classes, even when all four ratios remain admissible.

Proof. Upper ranks and attaining paths give the exact length on each stratum. The ten optimal $k = 6$ strata and sixteen optimal $k = 8$ strata join to the displayed intervals; the excluded right endpoint at $k = 8$ has length 155. Since $1 + 2\sqrt{2} < 4$, these loci are disjoint.

For the last assertion let $x = 5 - \sqrt{2}$ and $y = (17 + 6\sqrt{2})/7$. Exact evaluation gives the same terminal at x and y, but their conjugates under $\sqrt{2} \mapsto -\sqrt{2}$ have residual lengths 143 and 161 and hence different terminals. All ratios exceed one. This disproves descent to terminal classes without contradicting the typed equality-certificate covariance. □

The length bound in Theorem A.30 is attained on all 199 strata at $k = 4, 6, 8$; no all-order equality is inferred. The mixed-wall formula is checked on all 33 comparisons at eleven walls. For example at $k = 8$, $\rho = 3 + 2\sqrt{2}$, the side difference is

$$2056 = 8 \cdot 112 + 2 \cdot 540 + 2 \cdot 40,$$

while the lengths are $(143, 163, 143)$. After freezing regular midpoint changes they are $(145, 163, 143)$, so a width-20 excursion persists. At $(1 + 2\sqrt{2})/3$ the embedding-zero count

remains 256 but the component changes: new zero PCRs alone cannot detect all component changes. At $(5+4\sqrt{2})/7$ the three lengths are all 149 while the components are distinct. Equal length is therefore not a complete component invariant.

D.4 Technical normalization details

The prime-sine independence argument is proved once in Lemma A.5; that proof supplies the normalization used here.

D.4.1 Decision traces and sufficient forms

The five-role catalogue retains a transparent branch trace. Structural forms are registered separately from nonzero hyperplanes. The two-family theorem (Theorem 2.2) proves that the lambda and midpoint signs suffice; retaining all five roles is an audit choice, not a minimality claim. A point stratum is retained separately from its adjacent intervals.

D.4.2 Certificate height

Raw positional polynomials have coefficients in $\{-1, 0, 1\}$, with one signed one-hot letter per position. The role-transformed Laurent polynomials A, B need not have these coefficients. For example the midpoint polynomial of TTT at $k = 4$ contains the constant coefficient -6, since $A(z) = \sum_{e=1}^{3}(z^e + z^{-e} - 2)$. For a word of length n, all five role templates satisfy the safe bounds

$$\|A\|_1, \|B\|_1 \le 4n, \qquad \|UA + dB\|_1 \le 4n(\|U\|_1 + d) \quad (d > 0).$$

A common Laurent shift preserves this bound. The cleared degree is $O(k + \deg U)$. These polynomial height bounds suffice for the bit complexity statement, without asserting that every bounded polynomial has a positional realization.

D.5 Finite evidence and reproducibility boundary

The original regression proposition uses a canonical thirty-row payload. Its frozen identities are

$$\text{producer payload prefix: } aca5d5fc81c839c218b6eefaaf00acdc\ldots,$$
$$\text{validator payload prefix: } b79051b4190d6f13b9bb183fe1299fb6\ldots.$$

The full SHA-256 values, schemas, exact row table, independent reconstruction code, and replay instructions are included in the public finite-assurance capsule distributed with this article.

Theorem A.9 has a separate finite arithmetic checker with payload prefix `9ed015120913ea30`. It verifies selected composition identities, word counts and the $k = 12$ orbit count, and uses $k = 30$ only as an out-of-domain negative control. It is not a premise of the all-order proof.

The finite range is:

$$3 \text{ rows at } k = 4, \qquad 23 \text{ rows at } k = 6, \qquad 4 \text{ named landmarks at } k = 8.$$

No floating-point decision enters the scientific payload. Every row checks the direct formula against a legal firing replay, literal RC invariance, one-per-PCR cardinality, and acyclicity

by both Kahn elimination and an independent depth-first search. Longest paths are counted in edges; the corresponding vertex length is one larger for a nonempty path.

This regression capsule is exact inside its declared range. The extended capsules supply the complete $k = 8$ atlas, component classification, retiming certificates and the 131 construction, as described in Section B.1. Neither capsule claims a terminal atlas at $k \geq 10$, biological performance or a deduction of uniform theorems from finite rows.

References

[1] Bengt Aspvall, Michael F. Plass, and Robert Endre Tarjan. A linear-time algorithm for testing the truth of certain quantified Boolean formulas. *Information Processing Letters*, 8(3):121–123, 1979. Erratum: 14(4), 195 (1982). Used for the classical 2-SAT implication-graph criterion.

[2] Oleksiy Babanskyy. A parity theorem for reverse-complement-invariant minimum decycling sets in q-ary de bruijn graphs. Release v2.0.0, commit 89a29a90962d02add6690d6411db927c4788371a; https://github.com/aconsciousfractal/FCIG-RC-Invariant-Minimum-Decycling-Sets, 2026.

[3] Anders Björner, Michel Las Vergnas, Bernd Sturmfels, Neil White, and Günter M. Ziegler. *Oriented Matroids*, volume 46 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2 edition, 1999. https://doi.org/10.1017/CBO9780511586507.

[4] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. https://web.stanford.edu/~boyd/cvxbook/.

[5] Jean-Marc Champarnaud, Georges Hansel, and Dominique Perrin. Unavoidable sets of constant length. *International Journal of Algebra and Computation*, 14(2):241–251, 2004. https://doi.org/10.1142/S0218196704001700.

[6] John Christie, Ken Dykema, and Igor Klep. Classifying minimal vanishing sums of roots of unity, 2020. https://arxiv.org/abs/2008.11268.

[7] John H. Conway and A. J. Jones. Trigonometric diophantine equations. *Acta Arithmetica*, 30:229–240, 1976. https://doi.org/10.4064/aa-30-3-229-240.

[8] Laura De Carli and Maurizio Laporta. Divisibility criteria and coefficient formulas for cyclotomic polynomials. *Ramanujan Journal*, 70:44, 2026. https://doi.org/10.1007/s11139-026-01421-6.

[9] Ioannis Z. Emiris and Elias P. Tsigaridas. Computations with one and two real algebraic numbers, 2005. arXiv preprint; https://arxiv.org/abs/cs/0512072.

[10] Sreyash Kenkre, Vinayaka Pandit, Manish Purohit, and Rishi Saket. On the approximability of digraph ordering. arXiv:1507.00662v1, 2015.

[11] Gergely Kiss, Izabella Łaba, Caleb Marshall, and Gábor Somlai. Lower bounds for mask polynomials with many cyclotomic divisors. *Advances in Mathematics*, 494:110932, 2026. https://doi.org/10.1016/j.aim.2026.110932.

[12] Biswajit Koley and Satyanarayana Reddy Arikatla. Cyclotomic factors of borwein polynomials. *Bulletin of the Australian Mathematical Society*, 100(1):41–47, 2019. https://doi.org/10.1017/S0004972719000339.

[13] Shuvendu K. Lahiri and Madanlal Musuvathi. An efficient decision procedure for UTVPI constraints. Technical Report MSR-TR-2005-67, Microsoft Research, 2005. https://www.microsoft.com/en-us/research/publication/an-efficient-decision-procedure-for-utvpi-constraints/.

[14] T. Y. Lam and K. H. Leung. On vanishing sums of roots of unity. *Journal of Algebra*, 224(1):91–109, 2000. https://doi.org/10.1006/jabr.1999.8089.

[15] Charles E. Leiserson and James B. Saxe. Retiming synchronous circuitry. *Algorithmica*, 6:5–35, 1991. https://doi.org/10.1007/BF01759032.

[16] Henry B. Mann. On linear relations between roots of unity. *Mathematika*, 12(2):107–117, 1965. https://doi.org/10.1112/S0025579300005210.

[17] Guillaume Marçais, Dan DeBlasio, and Carl Kingsford. Sketching methods with small window guarantee using minimum decycling sets. *Journal of Computational Biology*, 31(7):597–615, 2024. https://doi.org/10.1089/cmb.2024.0544.

[18] Guillaume Marçais, Charles S. Elder, and Carl Kingsford. k-nonical space: Sketching with reverse complements. *Bioinformatics*, 40(11):btae629, 2024. https://doi.org/10.1093/bioinformatics/btae629.

[19] Jeremy Martin and Victor Reiner. Cyclotomic and simplicial matroids. *Israel Journal of Mathematics*, 150:229–240, 2005. https://arxiv.org/abs/math/0402206.

[20] Johannes Mykkeltveit. A proof of golomb's conjecture for the de bruijn graph. *Journal of Combinatorial Theory, Series B*, 13(1):40–45, 1972. https://doi.org/10.1016/0095-8956(72)90006-8.

[21] David Pellow, Lianrong Pu, Bariş Ekim, Lior Kotlar, Bonnie Berger, Ron Shamir, and Yaron Orenstein. Efficient minimizer orders for large values of k using minimum decycling sets. *Genome Research*, 33(7):1154–1161, 2023. https://doi.org/10.1101/gr.277644.123.

[22] J. P. Steinberger. Minimal vanishing sums of roots of unity with large coefficients. *Proceedings of the London Mathematical Society*, 97(3):689–717, 2008. https://doi.org/10.1112/plms/pdn006.

[23] Lawrence C. Washington. *Introduction to Cyclotomic Fields*, volume 83 of *Graduate Texts in Mathematics*. Springer, 2 edition, 1997. https://doi.org/10.1007/978-1-4612-1934-7.

[24] Hongyu Zheng, Carl Kingsford, and Guillaume Marçais. Lower density selection schemes via small universal hitting sets with short remaining path length. *Journal of Computational Biology*, 28(4):395–409, 2021. https://doi.org/10.1089/cmb.2020.0432.