

Generative Adaptive Dynamical Systems: Foundation, Architecture, and Compositional Geometry

Matthew Andrews

September 15, 2026

Foundation Derivation

Commitment. No central controller. The system updates without any process coordinating its moves.

Argument. Any process acting on non-adjacent positions requires information about the relationship between those positions. That information is not accessible from either position's neighborhood alone: it describes something across the distance between them. For the process to use that information, a mechanism must bridge the distance. That bridging mechanism is the central controller.

Therefore. Under no central controller, all processes are restricted to what they can access from their own bounded neighborhoods. Locality is forced.

Therefore. The substrate is the structural object on which local processes operate: discrete units at ordered positions, with adjacency defined by position structure.

Layer 0

Substrate Instantiation via POP

The substrate's type set is $\mathcal{T}_A \cup \{\emptyset\}$: the atomic alphabet committed at instantiation, plus the absence marker used for padding. This is the substrate's complete view of unit types. The substrate is initialized as a one-dimensional structure over $\mathcal{T}_A$.

POP, the population, is what makes the substrate operational. Without POP placed on the substrate, the substrate is just its structural commitments together with $\mathcal{T}_A$ as a 1D array: no D_t beyond one, no realized base space, no time index. The act of placing the initial population P_0 initiates $t = 0$, induces D_0 as the length of the longest solution, and brings the base space $(\mathcal{T}_A \cup \{\emptyset\})^{D_0}$ into being.

POP is the matter every subsystem and operator forced by the chains will act on. The substrate provides the geometric object; POP instantiates it. At L0, the population is composed entirely of atoms from $\mathcal{T}_A$.

Reproduction (selection, culling, cloning) maintains the population over time. At L0, selection is uniform random over the population: there is no differential survival, no fitness, no signal traffic yet.

L0 Operator Set

The operator algebra at L0 acts on the population. Operators are local: each fires at a single position in a single solution, with admissibility determined by the local context at that position. Three contexts are distinguished by the local boundary structure of the encoded representation:

- **Context 0 ($\mathcal{C}_0$):** undelimited positions, outside any boundary pair, holding non-delimiter tokens.
- **Context 1 ($\mathcal{C}_1$):** interior positions, inside a boundary pair, holding non-delimiter tokens.
- **Context 2 ($\mathcal{C}_2$):** delimiter positions, holding boundary markers themselves.

Each visit to a position is one draw. In a non-delimiter context a mutation event occurs at the context's rate, and the event applies one of four operators, point, swap, insert or delete, chosen uniformly. The four are not separate rates; they are the choice made once an event has occurred. Boundary creation and boundary removal are separate events with their own rates.

$\mathcal{C}_0$: mutation event at rate π_4 per visit, operator uniform over {point, swap, insert, delete}. Plus boundary creation at rate π_9 : inserts a paired boundary marker around the position.

$\mathcal{C}_1$: mutation event at rate π_5 per visit, operator uniform over the same four.

$\mathcal{C}_2$: boundary removal at rate π_6 per visit, removing the matched pair as one unit. Otherwise a delimiter takes the mutation rate of the span it bounds, π_5 , and of the four operators only swap can apply to it: point and insert would break the matched pair, and delete is boundary removal, which carries its own rate. A delimiter event therefore swaps the marker with a non-delimiter neighbor with probability 1/4, moving the boundary's position, and is otherwise a no-op.

The context-modified rates produce the L0 dynamics on which all subsequent chains build.

The Boundary Mechanism

Boundaries are population-level structural admissibility geometry, not a subsystem and not substrate-level types. The substrate does not see them, does not classify them, and does not place them in its base space. They exist as structural metadata at the population level that the operator algebra uses to partition positions into contexts.

The boundary mechanism follows strict structural rules, all of which are local stochastic constraints applied during operator firing:

Atomicity. A boundary consists of a paired opening and closing marker. These markers are structurally atomic: created and removed only as paired units, never independently altered. Boundary creation always places both markers; boundary removal always removes both. No operator may create or destroy a single boundary marker on its own.

Non-interpenetration. Boundary pairs may not cross or pass through one another. Two boundaries may approach and become adjacent, but their relative ordering along the sequence is preserved. Delimited regions remain disjoint.

No-nesting. Boundary markers do not nest within other boundary markers. The admissibility check on boundary creation enforces this directly: a new boundary may only be placed in $\mathcal{C}_0$, never within an existing boundary pair.

Mobility. Boundary markers are not fixed in position. Under boundary swap ($\mathcal{C}_2$ operator), markers may exchange position with non-delimiter neighbors, expanding or contracting the extent of the delimited region they belong to. Mobility is a consequence of local operator application alone: no global coordination is required.

Pair-creation rule. When boundary creation fires at a $\mathcal{C}_0$ position, the matched closing marker is placed at an adjacent position by the same operation. The pair is well-formed at the moment of creation; subsequent operators may move either marker but the pair remains matched.

L0 Structural Properties

L0 is a dynamical system in its own right, with structural properties that emerge from the operator set and the boundary mechanism. These properties are L0's own; they are not produced by any chain.

Two persistence regimes. The context-modified rates create two distinct persistence regimes for atomic content. With $\pi_5 < \pi_4$, content in $\mathcal{C}_1$ persists longer than content in $\mathcal{C}_0$: the mean-age ratio τ_1/τ_0 scales inversely with π_5/π_4 . Atomic patterns inside boundary pairs accumulate stability that patterns outside cannot. This is the substrate-level precondition for higher abstraction: stable patterns form because of the persistence differential, not because of any preservation mechanism. Chain 1 will eventually add.

Information flux at the boundary. The $\mathcal{C}_0/\mathcal{C}_1$ boundary is permeable. Atomic tokens flow across it through two channels: the swap channel (a boundary marker swaps with a neighboring non-delimiter, moving a token from one context to the other) and the bulk channel (boundary creation pulls a $\mathcal{C}_0$ token into $\mathcal{C}_1$; boundary removal releases all interior $\mathcal{C}_1$ tokens back into $\mathcal{C}_0$). The two channels together determine the net information flux through the boundary at each step.

Boundary density and token economy. The boundary count A_2 and the non-delimiter token count $A_0 + A_1$ are related by an equation-of-state-like constraint: the mean tokens per boundary, $\bar{k} = (A_0 + A_1)/A_2$, follows a Tonks-like relationship determined by π_9 and π_6 . This is not a descriptive statistic sitting beside the architecture; it is the supply variable of the compositional layer. Capture fires only on interior content, so the rate at which the hierarchy can grow is bounded by the supply of capturable material, and $\bar{k}$ is precisely that supply: content per boundary, the reservoir capture draws from. The relationship governs how boundary density evolves as content shifts, and because capture consumes interior content while boundary dynamics replenish it, $\bar{k}$ is the balance of a source against a sink. When the sink runs unopposed (capture with no inverse, as in the isolated Layer 1 regime) $\bar{k}$ is driven monotonically to its floor and the hierarchy locks once the reservoir empties; this is the empirical content of the Chain 2 forcing, and it is the trajectory of $\bar{k}$ that signals it before the lock occurs. Boundary density is therefore upstream of every capture event and a leading indicator of the compression dynamics, not an incidental measurable.

Segment-length structure. The interior length k of a boundary pair (the number of $\mathcal{C}_1$ tokens between the opener and closer) follows a conditional geometric distribution: $P(k = 0)$ is set by the contact dynamics of adjacent boundaries; $P(k \geq 1)$ is geometric in k with a rate determined by the operator algebra. This distribution is the size profile of capture targets: capture requires interior length $k \geq 2$, so $P(k)$ fixes what fraction of boundary pairs are capturable and how large

the captured regions are. Together with $\bar{k}$ it is the generative law for what enters MG, the rate set by boundary density and the size set by the segment distribution, and the two move together as the reservoir fills or drains. Like $\bar{k}$, mean segment length collapses as an unopposed capture sink consumes interior content, tracking the economy down toward the lock. It is an L0 encoding-geometry property, measurable as a structural fact, and it is one of the two variables the entire compositional layer is downstream of.

Stable atomic patterns. The persistence differential allows atomic patterns to form inside boundary pairs and remain stable on timescales longer than their $\mathcal{C}_0$ counterparts. These patterns are L0’s expressive output. They exist; they persist; but L0 has no mechanism to refer to them as wholes. They remain atomic-sequence patterns inside boundary pairs, not units that downstream operations can manipulate.

L0 as the Foundation for the Chains

Together: substrate + POP + reproduction + the L0 operator set above = Layer 0. L0 is structurally complete on its own merits: two-context geometry, paired boundaries with mobility, differential persistence, information flux, a token economy ($\bar{k}$ and the segment-length distribution) that sets the supply and size of capturable material, and stable atomic patterns inside boundary-delimited regions. The chains in the next section add machinery on top of this L0, each addressing a failure mode that emerges as the architecture is extended.

L0 is pre-compositional. The stable atomic patterns L0 produces are the substrate-level preconditions on which Chain 1 will introduce capture and force MG.

The Chains

Each chain identifies an asymptotic structural failure mode and the constraint, operator, rule, signal channel, or subsystem that addresses it. Removing any chain’s conclusion produces the failure mode it was forced by.

Chain 1

At L0, boundaries and context-modified mutation rates allow stable atomic patterns to form within boundary-delimited regions of encoded solutions. But L0 has no mechanism for compressing those stable patterns into reusable units. The patterns exist, persist within their boundaries, and decay or mutate over time, but they cannot be referenced as wholes by downstream operations.

Capture is the operator introduced to perform this compression. It acts on a stable boundary-delimited region and replaces it with a single composition token. Compositions are population-level tokens, not substrate types: the substrate does not see them.

A composition is its content. The token capture produces is not a name standing for a content sequence. It is the content sequence, carried by value as an immutable nested unit over atoms and compositions, and identity is structural equality: two compositions are the same object when and only when their content matches at every depth. This is a commitment about the encoding, made here because everything downstream turns on it. Three properties follow from it rather than being arranged separately. Nothing resolves a reference, so no cache and no admission policy is needed. A composition is meaningful to any process that receives it without a shared namespace. And a

composition can be folded into the structures that reference it by value, which is what Chain 4 will use.

Capture admissibility. Capture is locally admissible only when its target region is already boundary-delimited and contains enough material to compress. Three constraints govern capture’s firing:

- **Boundary-delimited target.** Capture fires only on a region enclosed by a matched boundary pair. Capture never selects an arbitrary substring of a solution; the boundaries set the locus of abstraction. This preserves locality: the region of effect is bounded by structure already present in the encoded representation, not by any external scope decision.
- **Minimum extent.** The enclosed content must have at least two tokens (atomic or composition). Single-token regions cannot be captured. This prevents capture from producing compositions whose decoded content is no more compressed than a single token.
- **Boundary consumption.** Capture acts on the $\mathcal{C}_1$ content enclosed by a matched $\mathcal{C}_2$ boundary pair. When capture fires, the boundary pair and the enclosed content are removed together and replaced by the single composition token. The pair is consumed; once the boundaries are gone, the position is no longer interior, so the new composition token sits in $\mathcal{C}_0$. No orphan boundary markers remain.

A composition token is an ordinary token thereafter. Although it lands in $\mathcal{C}_0$ at capture, a later boundary-creation operator may place a boundary pair enclosing it, moving it into $\mathcal{C}_1$; further material may be inserted alongside it in that region. A subsequent capture firing on that boundary pair produces a higher-order composition whose content includes the enclosed composition token. This is how compositions come to nest: not through nested boundaries (which L0 forbids) but through a composition token migrating into a boundary region and being captured along with its surroundings. Boundary pairs never nest; compositions do.

Capture is also subject to deduplication, and under value identity deduplication is free. Two captures of the same span produce structurally equal content, hence the same composition, so a repeat capture returns what already exists rather than admitting a second copy. The boundaries are still consumed and the population gains the token. No comparison of names is involved; the test is membership.

Failure mode. Capture as described compresses a pattern inside one organism. What it does not do is make the resulting unit available to anything else. Sampling has no set to draw a composition from, so no operator can propose the unit into an organism that did not already inherit it; the unit spreads only by descent from the organism that captured it, and dies with that line. Open has no basis on which to decide whether a token it encounters is a current unit or the residue of one. And a later capture cannot build on the unit except where the unit happens already to sit.

The failure is not that a composition cannot be decoded. It can: the token carries its own content, and decoding needs nothing else. The failure is that the compression cannot be *used*. The search performs a local relabeling and the frame it searches in never rises above the atoms.

Therefore: a persistent store of captured content must exist (MG), whose members are their own identities. It is the non-atomic half of the search frame: the set of units sampling may offer, open may test against, and capture may build on. That the store is a set rather than a mapping is not an economy of implementation but a consequence of the encoding, and the conclusion is correspondingly stronger than one that forces a lookup table.

The compositional DAG and its structural prior. Compositions can reference other compositions: a capture firing on a region that contains existing composition tokens produces a new composition whose content includes those tokens. The result is a directed acyclic graph $\mathcal{G}$ of compositions, carried by containment, where edges represent reference relationships.

Each composition $m \in \mathcal{G}$ has two natural properties:

- $\text{depth}(m)$: the maximum nesting depth of compositional references in m 's content.
- $\text{flatlen}(m)$: the length of m 's fully-decoded atomic content.

How a composition touches the base space. A composition is not a point in the base space and does not move a solution within it. The base space remains the flat combinatorial volume $(\mathcal{T}_A \cup \{\emptyset\})^{D_t}$ the substrate is defined to be. A composition's only contact with that space is by *pinning*: m decodes to a fixed atomic string of length $\text{flatlen}(m)$, so placing the token m in a solution fixes $\text{flatlen}(m)$ base-space coordinates to specific values in a single stroke. One token pins a $\text{flatlen}(m)$ -dimensional axis-aligned subspace of the base space. This is the sole channel through which compositions act on the base space; there is no other operation by which a composition relates a solution to a point in $\mathcal{B}_t$.

Pinning is not free: recovering the pinned coordinates from the token requires resolution, and $\text{depth}(m)$ is the number of resolution steps separating m from the base alphabet. A composition referencing compositions stacks these steps, and the count of distinct base-space configurations reachable through m compounds *multiplicatively* with each step: each referenced composition pins its own subspace within the resolution of its referent. The locus of resolution cost over reachable configurations therefore cannot be flat: multiplicative branching of pinned subspaces has constant negative curvature, and there is no free parameter at which to set it.

The forced extent. On that constant-negative-curvature locus, flatlen is the dimension of the pinned subspace and depth is the radial cost coordinate. The extent of the cost-geometry subspace a composition occupies is then fixed by construction:

$$d_H(m) = \text{arccosh}(1 + \text{depth}(m)) \cdot \text{flatlen}(m). \quad (1)$$

This is the exact hyperboloid extent of that subspace, not an approximation and not a metric imposed on the DAG. It is what the pinning operation produces once a single composition exists; nothing here is chosen. (Section 8, Search Space and Compositional Geometry, develops the construction in full and its relationship to PFOL A10.3.)

The measure, and why it is not a prior. Because d_H is the forced extent of the region of base space a composition pins through resolution, the relative weight of a composition is the relative size of that region. The normalized measure over compositions is therefore

$$P_{\text{struct}}(m) \propto \exp(-d_H(m)), \quad (2)$$

not as a structural prior posited from outside and justified after the fact, but as the volume element of the cost geometry itself: the weight is the extent, and the extent was forced by pinning. Shallow, short compositions pin small regions and carry more weight; deeply-nested or long compositions pin exponentially larger cost-regions and carry exponentially less. There is no choice to assign.

The measure is forced, not posited. The relative weight $P_{\text{struct}}(m) \propto \exp(-d_H(m))$ is not a priori assigned from outside: it is the volume element the pinning operation fixes, so $d_H = -\log P_{\text{struct}} + \text{const}$ holds by construction. This is what the variational free energy decomposition consumes as its complexity term, and it is established by counting in §8.1 and §8.6; the development there is the warrant, not this paragraph.

What Chain 1 forces. Chain 1 introduces capture (the operator that compresses stable patterns into compositions), fixes the encoding (a composition is its content), and forces MG (the persistent store that makes the compression reusable). The compositional DAG $\mathcal{G}$ and the hyperbolic cost geometry with its forced measure P_{struct} emerge as the structural consequences of compositions existing and being able to reference other compositions: they are not separate forced things but the structural form of what capture and MG together produce.

Chain 2

Capture is one-way. Without an inverse, compositions accumulate. Two failure modes appear together:

- *Compositions outnumber atoms.* The population’s encoding becomes dominated by composition tokens that all expand to fixed content. Atom-level diversity diminishes. Variation collapses.
- *Singularity.* A dominant composition concentrates downstream captures around itself. Configuration space collapses to one region. Selection pressure cannot recover.

Therefore: an inverse operator (open) must exist. Open expands a composition back to its atomic content, with context-sensitive boundary handling. Geometrically (§8.5): open is the radial descent on the resolution cover, the operator that exposes a composition’s constituents as independently editable positions. Without it, occupancy ascends the cover and cannot descend; the population collapses onto a single coarse sheet whose tangent neighborhood carries no remaining branching, the singularity above, stated in cover terms.

In the token economy of L0, the same failure reads as the supply variable $\bar{k}$ collapsing: capture is a sink on interior content, open is the source that returns content to circulation, and with the source absent the reservoir is consumed until $\bar{k}$ and mean segment length reach their floor and capture starves. Open is forced as the operator that opposes the sink, the only thing that holds the economy off the floor and keeps capturable material in supply.

What open does not do. Open acts on an organism’s own representation, exposing a composition’s constituents as editable positions in that organism. It does not act on the store. A composition every organism has opened is still live, still sampled, still available to be nested by a later capture. Occupancy descends the cover; the frame does not. This is what Chain 3 addresses.

Chain 3

With capture and open both active, the frame $L(t)$ accumulates compositions that no population member is using. Capture admits; nothing removes. What this chain settles is which removal rule is admissible, given that no process may evaluate a composition’s worth.

Failure mode (granularity ratchet). Three facts fix the direction of the frame.

- Capture is the only operator that admits an entry to $L(t)$. It admits a boundary-delimited span of any length, so an entry may arrive long without arriving deep, and a capture whose content includes a composition arrives one resolution level deeper than it. Admission raises one or other coordinate of $d_H = \text{arccosh}(1 + \text{depth}) \cdot \text{flatlen}$ and lowers neither.
- Open acts on parses, not on $L(t)$ (Chain 2).
- Inlining (Chain 4) is the only operation that alters an entry already in $L(t)$, and it fires only on removal. It lowers depth. It does not lower extent: splicing a dead child into a parent conserves that parent's decoded content by construction, so flatlen is invariant under it. Extent leaves the store only when a removed entry has no live parent to be spliced into.

The frame therefore ratchets in both coordinates, and the two ratchets are not the same failure.

The radial ratchet. Absent removal, $\max_{m \in L(t)} \text{depth}(m)$ is monotone non-decreasing and the reference tower of §8.6 has nothing opposing it. That tower is open in depth by construction: the no-nesting rule bounds boundary pairs, not reference layers, and the migrate-and-recapture route builds reference layers without ever nesting a boundary.

The lateral ratchet, which runs first. Capture admits spans of unbounded length, and length compounds without any nesting at all. A long entry placed in an organism raises D_t by its whole flatlen; the added transverse room admits longer boundary-delimited spans; longer spans admit longer captures; those entries lengthen the organisms that carry them. The loop closes without the radial coordinate being paid once.

It is also the cheaper loop under the architecture's own measure. d_H is linear in flatlen and only logarithmic in depth, since $\text{arccosh}(1 + \delta) \sim \log \delta$, so pinning the *same* content more deeply strictly costs more while pinning *more* content costs linearly either way. Going wide is the cheap direction, so it is the direction the search takes first. Solutions lengthen, entries lengthen with them, and nothing in the architecture opposes either.

Failure mode: without removal the frame's granularity is a ratchet in both coordinates. $L(t)$ accretes deeper and longer units and sheds neither, the single-step vocabulary available to the operators coarsens monotonically, and the self-feeding regime §8.6 identifies runs unopposed in the lateral direction before it runs in the radial one. This is Chain 2's failure mode one level up: Chain 2 concerns where in the cover the population sits, Chain 3 concerns what the cover is built from. Open holds occupancy off the coarse sheet. Nothing yet holds the frame off the tower.

The ceiling is finite at fixed length and unbounded over time. What capture can ever admit is bounded by the decomposability of the reachable space: over a region of decoded length n the number of distinct structures is the fiber multiplicity $\mu_B(n)$ of §8.1, finite for finite n and growing as $C\lambda^n$ with $\lambda > 1$. At any instant the ceiling is therefore very large and not infinite, fixed by the longest solution observed so far. But D_t is monotone in that length and nothing internal bounds it, so the ceiling itself ratchets up. This is the shape §8.6 states for region count, finite at fixed V_t and monotone unbounded in V_t , and it is the correct statement of the failure: not that the store runs out of room under a fixed ceiling, but that the ceiling slides.

Why the premise is not finite storage. Exhaustion of a host’s memory is not an admissible failure mode here. It is a contingency of a deployment, not a property of the class, and a chain resting on it would force a capacity parameter the architecture must then be given from outside. The ratchet above is internal, asymptotic, and structural, in the same form as every other chain’s failure mode.

Signal constraint. By locality, no subsystem can evaluate the worth of a composition. A worth-evaluation would require interpretive machinery reaching into the population’s structure or into MG’s content, both of which violate the scope constraint the foundation derivation forces. The only event the mechanism may observe is whether a composition is *present* as a top-level token in the population, and Chain 6 supplies the argument that presence must be scoped to the top level. Each observation produces exactly one bit per composition: it appeared or it did not. Frequency, magnitude, fitness correlation, depth, and every other metric require evaluation machinery the architecture forbids.

The signal is rank-free. The observation is taken over the whole population at the moment it is scored, before reproduction. It is not restricted to survivors and it consumes no ranking. This is stronger than a binary signal: a signal restricted to survivors would carry selection’s ordering into the metabolic layer through the choice of who emits, and some process would have to hold the rank to apply the restriction. Taking the reading over the whole population removes that channel entirely. Selection still reaches the frame, but it reaches it by renormalization rather than by filtering: an organism that does not reproduce is absent from the next population, so the compositions only it held are absent from the next reading. The filter is applied by the population’s own turnover, at no cost and with no subsystem consuming rank information.

What the signal does not license. Given a one-bit observation, the mechanism cannot rank compositions by value. It cannot count events (counting requires accumulation and the observation carries no magnitude), assign weights (no metric exists to weight), or compute scores (no continuous quantity is available).

It also cannot order. An ordering by last participation is a comparison of durations of non-participation across compositions, and a duration is a magnitude that no single observation carries. Maintaining one requires a persistent container whose state is an accumulation of past observations, which is the accumulation the constraint forbids under a different name. An ordering could be warranted only by a need to *schedule* displacement, since a capacity-bounded container must choose which occupant to evict and the observation alone does not choose. With no capacity there is nothing to schedule, and the constraint applies unopposed.

This does not touch Chain 5. MCO orders $L(t)$ by creation time, which is not derived from the participation observation and is not a record of it. The constraint here binds what may be inferred from participation, not what may be recorded about capture.

The admissible rule. One rule uses the observation without exceeding it: remove exactly the compositions absent from it. Let $C(t)$ be the set of compositions appearing as top-level tokens anywhere in P_t at scoring time, and let the reading be

$$\mathcal{S}(t) = C(t) \cap L(t). \quad (3)$$

The removal set is $L(t) \setminus \mathcal{S}(t)$, and each removal fires inlining (Chain 4). Nothing is counted, nothing is aged, and nothing is ordered.

The intersection with $L(t)$ is not a technicality. An organism may hold a form the store has moved past, and it decodes to the same atoms whether or not the store still holds it, since decode expands a composition from its own content and does not consult the store. Such a token is phenotypically inert. The invariant it relies on, decode invariance under removal, is stated with Chain 4; the removal rule here is admissible only because that invariant holds, and the two should be read together.

Signal attribution under inlining. Inlining conserves content and changes structural identity: the form an organism holds is superseded by the form the store now carries, while the decoded atoms are unchanged. The reading reads structural identity, so an organism left holding a superseded form is silent about the live one. Unrepaired, that silence is a false absence. Content in active use across the population reads as unheld, and the live form is removed on evidence that is an artifact of the rewrite rather than an observation of the population.

What this forces is a requirement on the signal, not a subsystem. The participation of a superseded form must be attributed to the form that replaced it, within the interval the removal rule reads; equivalently, holders degrade to the live form. That it must happen is a consequence of inlining, and inlining is forced at Chain 4. How it happens is instance-level: any mechanism that re-attributes within the interval discharges the requirement, and the class theory is indifferent between them. The canonical instance collects the re-attribution on the decode each organism is already performing, which adds no traversal and no scan of the population, but the requirement does not force that choice and nothing is forced by it beyond the attribution itself.

Why one observation is sufficient evidence. The reading is *total* and *synchronous*. It is total because one evaluation pass decodes every organism in the population, so it covers every process that could be holding a composition at top level. It is synchronous because it is taken at one instant of the cycle rather than accumulated across it. Absence from a total, synchronous observation is therefore not an inference about non-service. It is the observation of non-service, which is the evidence removal requires and the only evidence available locally.

A threshold of more than one generation would answer a lag between two clocks: removal decided on one schedule and confirmation arriving on another. Here removal is decided by the reading itself, so the interval between them is zero and no threshold is forced. A threshold on this signal would in any case exceed what the constraint licenses, since counting generations of absence is an accumulation of a signal that carries no magnitude.

Derived scheduling constraint. The observation is valid only for the population it was taken over, so it must be consumed before that population changes:

$$\text{attribute} \rightarrow \text{decode and score} \rightarrow \text{read} \rightarrow \text{sweep} \rightarrow \text{reproduce} \rightarrow \text{mutate}. \quad (4)$$

Sweep placed after mutation would decide removal from a reading taken over a population that no longer exists, and would remove compositions that mutation had just handed out. Structure captured later in the generation is not a candidate at this sweep: it faces its first reading in the next generation, by which time it is sitting in the organism capture placed it in. The ordering is forced by what the observation means, not chosen for convenience.

Therefore: a Participation Signal must exist. It is one bit per composition per generation, raised by presence as a top-level token anywhere in the population and read once, over the whole population, before the population changes; absence at that reading is the removal criterion. The name is literal: the signal carries no magnitude, no ordering and no rank, so a single wire per

composition is its full width, and everything the removal rule is entitled to do is decided by whether that wire is high.

No container, no capacity, no ordering, no threshold and no counter is forced alongside it. What Chain 3 forces is a signal and the rule that consumes it, discharged where decode is already happening and holding no state of its own. Where the chains are said to force subsystems, the word is used functionally throughout: a subsystem is a role that must be discharged, not a structure that must be held.

Cardinality is bounded as a consequence. Write $N_t = |L(t)|$, let $n_{\text{top}}(x)$ be the number of top-level composition tokens in organism x , and let

$$B(t) = \sum_{x \in P_t} n_{\text{top}}(x) \quad (5)$$

be the *service budget*: the total number of top-level composition slots the population holds. Let $c(t)$ be the number of compositions admitted by capture during generation t . Sweep leaves exactly $\mathcal{S}(t)$, and merges produced by inlining can only reduce the count, so

$$N_{t+1} \leq |\mathcal{S}(t)| + c(t) \leq B(t) + c(t). \quad (6)$$

From the first sweep onward the store is bounded by the population’s capacity to hold distinct compositions at top level plus one generation’s admissions. No capacity constant appears anywhere in the statement, and none is available to be tuned: $B(t)$ is a property of the population, and $c(t)$ is a property of the capture rate against the L0 token economy. The bound is derived, not stipulated.

What that bound does not do. It bounds how many compositions are live, not how large they are, and the two are independent. A store pinned at B entries whose entries are each of unbounded flatlen has bounded cardinality and unbounded extent, and it is extent that the ratchet moves, extent that d_H charges, and extent that lengthens the solutions the population carries. The bound disposes of storage without a capacity parameter. It does not answer the ratchet.

Removal cannot answer the ratchet on its own, and the reason is structural rather than incidental. Inlining conserves each parent’s decoded content, so the removal mechanism is a depth regulator by construction: it flattens the hierarchy and re-granularizes what is held, moving content down the radial coordinate while leaving the lateral coordinate exactly where it was. Extent is shed only at the leaves of the reference relation, when a removed entry has no live parent to be spliced into. Against a lateral ratchet whose loop closes without nesting, leaf attrition is not a restoring force.

The store self-sizes in cardinality. The bound above is a ceiling on the count. What sets the operating point of the count is a negative feedback between admission and removal. Let $\sigma(N)$ be the expected size of the reading with N compositions live, and let $c(N)$ be the expected admissions per generation at that size. Removal per generation is $N - \sigma(N)$, and the net drift is

$$\Phi(N) = (N - \sigma(N)) - c(N), \quad (7)$$

with N falling where $\Phi > 0$ and rising where $\Phi < 0$.

Two structural facts give Φ its sign at the extremes. First, $\sigma(N) \leq \min(N, B)$: a composition is counted only if some organism holds it at top level, and the population has only B such slots. So for $N > B$ the removal term is at least $N - B$ and grows without bound in N , while c is bounded by

the capture rate, giving $\Phi > 0$ for all N beyond $B + \sup_t c(t)$. Second, $\Phi < 0$ whenever the reading holds nearly everything live, which is the regime a small store is in, since almost every entry sits within the effective support of MCO’s sampling weight and is placed somewhere in the population. The drift therefore points inward from both ends, and the store is attracted to a bounded set on which admission and removal balance.

The mechanism of the restoring force is worth stating plainly, because it is the whole content of the claim: a larger store spreads a roughly fixed service budget more thinly, so a larger fraction of entries go unheld at the reading, so the removal rate rises. Size is its own pressure against itself. Deduplication reinforces this from the admission side, since a larger frame converts a larger fraction of capture events into re-captures of content already live, which admit nothing.

Scope of the equilibrium claim. The bound $N_{t+1} \leq B(t) + c(t)$ is exact and follows from the removal rule alone. The inward drift at both ends follows from the budget cap and the sampling support, and is argued rather than proven. That the attracting set is a single point rather than a bounded interval, the value of the operating size, its scaling with population and problem, and the dynamics of approach are all dynamical questions: path-dependent, statistical, and not treated here. This section claims that the count is bounded without a parameter and that the feedback on the count is restoring. It claims nothing about where the balance lands, and nothing whatever about extent.

Extent is bounded only from outside. $B(t)$ is itself a length: it is the population’s total supply of top-level slots, which is bounded only if the population’s encoded length is bounded, and nothing in the architecture bounds that. So the cardinality bound is conditional in exactly the place the ratchet lives, and the conditioning is not removable by any internal mechanism.

The bound §8.6 defers is therefore not recovered here, and this section should not be read as recovering it. Extent is held down only by an external measure that disfavors decoded length, transduced through the Participation Signal: an organism whose length is not earning its keep does not reproduce, so the compositions only it held are absent from the next reading, and the long entries behind them are removed. Chain 3 forces the channel and the removal rule. Chain 6 forces the differential survival that puts a gradient on the channel. The environment supplies the direction, and the architecture reads participation and never reads fitness.

The class-level constraint is §8.6’s, and Chain 6 states it in two parts. How a given GADS realizes it is an instance choice. That the length pressure lives in the environment rather than in the architecture is not a gap in the derivation; it is the claim, and the canonical instance makes it concrete: its length target and penalty are handed to the environment at binding, and nothing in the store, the operators or the reading consults either.

What Chain 3 forces. Chain 3 forces the Participation Signal and the removal rule that consumes it: one bit per composition, read as a total, synchronous, rank-free reading of top-level compositions once per generation before the population changes, with absence at that reading as the removal criterion and inlining (Chain 4) as what removal fires. It forces one requirement on the signal’s integrity, that participation be attributed to the live form after inlining supersedes a holder’s, and leaves the mechanism of attribution to the instance. Bounded cardinality and its self-sizing equilibrium are consequences of the rule rather than additional forced things. The chain forces no container and no capacity-driven displacement.

Chain 4

When the removal rule of Chain 3 removes a composition m , m is to be dropped from $L(t)$. But m may be contained by other compositions: some $m' \in L(t)$ has m in its content.

The decoded content of m' is defined by expansion of its content. If m is simply dropped while m' is left holding it, the store's account of m' and the population's account diverge, and any structure the store rewrites thereafter is rewriting a form no longer in correspondence with what it references.

Failure mode: removal alters the store's structural account of every composition that contained the removed one. Without a rule governing what happens to containers, the store's entries and the population's tokens drift apart, and the architecture loses the property that representation changes leave the population's expressed content invariant.

Why no other option preserves invariance. Reaching into organisms to remove tokens referring to m would preserve decoded content but violate locality. Removal is a store-level event triggered by a reading; reaching into the population to modify organisms is outside the scope of that event, and the locality commitment forces removal to act only on store state, not on population state.

Therefore: before m is dropped, every occurrence of m in every containing composition must be replaced with m 's own content. This is inlining. Because a composition is its content, the substitution is by value and requires no resolution: where m' previously contained m and expanded it, it now contains m 's elements directly, and expansion produces the same atomic sequence. Structures that become identical under the rewrite are merged.

Inlining is the only local operation that preserves phenotypic invariance under removal. It acts entirely within the store: no population modification, no global coordination, no signaling beyond the removal event itself.

Decode invariance under removal. The system's exact conservation law, and the invariant every downstream measurement relies on, including the admissibility of Chain 3's removal rule.

Invariant. For any organism x and any composition m , $\text{Decode}(x)$ is unchanged by the removal of m from $L(t)$.

Proof. Decode expands each token an organism holds from that token's own content and never consults $L(t)$. Removal changes $L(t)$, and through inlining changes the content of entries in $L(t)$; it changes no token held by an organism. If x holds m itself, m still carries its content and expands as before. If x holds a container p of m , the store's copy of p has been rewritten but x holds its own p by value, unchanged, and p and its rewritten form expand to the same atoms by construction of the splice. In both cases the expansion is identical, so $\text{Decode}(x)$ is invariant. $\square$

Two consequences. Removal is phenotypically free: no fitness value moves because of it, and no organism is damaged by holding a form the store no longer has. And what a removal costs an organism is therefore not content but identity, which is exactly the loss Chain 3's attribution requirement repairs.

Chain 5

When a mutation operator inserts a composition token into a solution, it samples from $L(t)$. With the store accumulating captures over many generations, the question is how to weight the sampling distribution.

Uniform sampling has no notion of when a composition was captured. A composition captured many generations ago, under conditions that no longer hold for the population’s current location in the search space, competes equally with a composition captured in the current generation. This decouples the sampling distribution from the population’s trajectory.

Failure mode: the sampling distribution drifts arbitrarily from the population’s current location. New compositions, captured under current conditions and most likely to be relevant to where the population is now, are sampled at the same rate as ancient compositions captured under conditions long past. The hierarchy stalls: new captures cannot accumulate into still-newer captures because they are crowded out by uniform competition with everything ever captured. The architecture loses the ability to build forward.

Therefore: sampling must be ordered by creation time and weighted toward recent entries. A chronological record of captures must exist (the Meta-gene Creation Order, MCO), and the selection distribution must weight position in MCO by exponential decay.

Temporal coupling. MCO provides the temporal ordering that keeps sampling coupled to the population’s current location. Recent captures encode patterns discovered under current conditions; weighting selection toward recent entries means new compositions are sampled most often by the operators that build still-newer compositions. The population’s trajectory and the sampling distribution stay together.

Decay as a structural mechanism. The exponential decay does more than weight selection. It is the architecture’s natural decay mechanism for compositions.

The newest composition in MCO has the highest sampling weight. As new captures arrive, every existing composition’s position relative to the newest moves one step further down the decay curve, and its weight diminishes exponentially. Given enough new captures, a composition’s sampling weight approaches zero: it is no longer being inserted into new solutions, no longer being nested inside still-newer captures.

A composition aged out of high availability faces a structural test: it can persist only if it remains held as a top-level token by population members. Being held is the only path left for a composition whose MCO-driven sampling weight has decayed away.

Information flows up, structure flows down. When a composition is present only because it is nested inside another, and not as a top-level token in any member, it does not raise the Participation Signal (Chain 6’s interface-scoped participation). The reading does not see it; sweep removes it; inlining (Chain 4) substitutes its content into every composition that contained it.

- **Information flows up.** The content of the removed composition is preserved in the compositions that contained it. Nothing expressible is lost.
- **Structure flows down.** The hierarchy flattens. The removed composition is no longer a unit; its content is now distributed across its containers.

The aggregate effect is a structural selection pressure: a composition that justifies its existence as a top-level unit in the population persists; a composition that exists only as a nested component of something larger is eventually dissolved into its containers, its content preserved, its identity lost.

Without MCO’s decay, every composition would be sampled at the same rate forever. Nothing would age out of high availability. The structural selection pressure would never engage, and the hierarchy would grow without the corresponding flattening.

Relation to the Participation Signal. MCO orders by creation time. The signal constraint of Chain 3 forbids ordering by participation and does not bind this, because creation time is not a record of participation. Nothing anywhere orders compositions by when they were last held.

Chain 6

GADS performs actions on the substrate. Operators fire, captures occur, the store grows, and the metabolic loop forced by Chains 3, 4 and 5 (the Participation Signal, inlining, MCO) regulates what the store holds. The loop's purpose is to exhaust entropy from the system: to carry away accumulated structure that is not contributing, so that the architecture's actions on the substrate remain sustainable rather than accreting without limit.

POP, the population, is established at Layer 0 as the matter the substrate accommodates. The loop operates on the Participation Signal from POP: which compositions are held at top level, read once per generation over the whole population.

The store is a set of hypotheses. Every entry in $L(t)$ is a claim that a particular span is worth naming as a unit. Capture proposes; the reading tests; removal discards what failed. This is what the loop is for, and it is the sense in which the loop exhausts: what leaves on removal is not content, since inlining conserves each parent's decoded content, but a distinction. A span that was a unit stops being one, and the frame is that much less committed to a claim the population did not support.

What does not fail without differential survival. Suppose selection is uniform, as at L0. The reading still fires, the removal rule still applies, and the rate equilibrium of Chain 3 still holds. The Participation Signal is rank-free by construction, the reading is taken over the whole population, and nothing in the removal rule consults fitness. Uniform selection disables none of it, so the failure is not there.

What fails: rate without direction. Chain 3 delivers a cycle. Admission and removal balance at a fixed point and the store self-sizes. What Chain 3 does not deliver, and cannot, is a direction for the flow through that cycle.

Under uniform selection, which compositions are held at the reading is determined by MCO's sampling and by which organisms happen to carry them, and both are independent of any measure on the content. The removal set is therefore a sample drawn without reference to whether the hypothesis it discards was a good one. Admission is likewise undirected. The store's composition performs an undirected walk over the reachable structures: entries arrive, entries leave, the count sits at its fixed point, and nothing is transported along any measure. The loop runs and exhausts nothing.

Restated as the test: a hypothesis is discarded on chance rather than on evidence, because the reading that was supposed to be the evidence is uninformative. Every span is as likely to be held as every other, so being held says nothing about the span, and the loop that was forced to carry away what is not contributing cannot tell what is not contributing.

Failure mode: loss of entropic pressure. The metabolic loop retains its cycle and loses its gradient, so it ceases to be an exhaust channel and becomes stationary churn. GADS's actions on the substrate accumulate structure at the rate they always did, with nothing carrying the residue away.

What the loss is observed as. Accretion is not visible in the store’s cardinality, which sits at its fixed point. It is visible in extent. With no gradient, a long composition is held exactly as readily as a short one, so the lateral loop of Chain 3 turns freely: long entries are placed, D_t rises by their whole flatlen, longer boundary-delimited spans become available, longer captures follow. Decoded length grows and the ceiling of §8.1 slides with it. The store’s cardinality then follows, since the Chain 3 bound is expressed in $B(t)$, which is itself a length. Extent is the coordinate the loss shows up in, and it shows up laterally first, because that is the cheap direction under d_H .

Therefore: differential survival must exist, and POP must expose two signal channels: Capture, carrying content admitted by a capture event, and the Participation Signal forced at Chain 3, raised by decode. Differential survival is established by fitness applied to the population, operating on the decoded content of each member. Decoding expands composition tokens to their atomic content, and the resulting atomic sequence is what fitness evaluates. Members that score better persist; members that score worse do not. Holding then correlates with scoring, being held becomes evidence about the span, and the flow through the cycle acquires a direction.

The gradient cannot be internal. A system does not supply its own difference. Every quantity the architecture can compute is computed from its own state, and a direction read off internal state is a direction the architecture chose rather than one it is held to. This is why Chain 6 places fitness outside the architecture rather than adding a subsystem that evaluates: the exhaust needs a gradient it does not author. It is also why the variational free energy of §9 is measured and not minimized. The architecture exhibits the quantity; the environment supplies the difference the exhaust runs down.

The gradient reaches the frame by renormalization, not by filtering. It would be a mistake to read this chain as making the Participation Signal survivor-filtered. It is not, and it must not be: a signal restricted to survivors would carry selection’s ordering into the metabolic loop through the choice of who emits, and some process would have to hold the rank in order to apply the restriction. Chain 3 forbids that and nothing here reopens it.

Selection reaches the frame by another route and at no cost. The reading is taken over whatever population exists. An organism that does not reproduce is absent from the next population, so the compositions only it held are absent from the next reading, and the removal rule disposes of them without ever having been told why. Fitness renormalizes the population between readings; the readings themselves stay total and rank-free. The architecture supplies the mechanism of regulation, the environment supplies its direction, and no subsystem reads a fitness value at any point.

One signal, not two. Chain 3 forces what the signal is: one bit per composition, total, synchronous, rank-free. Chain 6 forces that it is emitted, by placing decode as the event that raises it, and that it carries a gradient, by forcing differential survival. These are one object seen from two chains, and it is named once, at Chain 3.

Interface-scoped participation. The signal must be precisely scoped. A composition c may appear in a population member in two distinct ways: as a top-level token in the member’s own encoded sequence, or nested inside the content of another composition the member holds. These are different facts about c , and the signal must distinguish them.

The signal is raised for c only when c appears as a top-level token, and at most once per member per generation. When a member is decoded, expansion of its composition tokens encounters nested

compositions, and those nested compositions do not raise the signal. Expansion is internal to the decode; it is not a participation event.

Why the scoping is forced. If nested compositions raised the signal during expansion, a composition kept alive only because it sits inside a larger one would refresh itself. It would read as held when nothing in the population uses it as a unit; it merely rides along inside something that is used. Two consequences follow, both fatal:

- **The hypothesis stops being tested.** The claim an entry makes is that its span is worth naming *as a unit*. A signal raised by nested occurrence tests a different claim, that the content appears somewhere, which is true of any content anything is built from and discriminates nothing. The reading would return a constant. Chain 5’s decay would never find a composition absent, the flattening would never engage, and the loss of direction described above would recur inside an architecture that has differential survival, because the gradient would have nowhere to act.
- **Encapsulation breaks.** To credit a nested composition, the reading would have to know that it is embedded inside another, which is structural knowledge of the store’s DAG. The reader would be inspecting the store’s internal content rather than observing the population. This violates the locality commitment directly, and it is the one thing the reading is constructed never to do.

The top-level rule keeps the signal legible and the roles encapsulated. “Used as a unit” and “embedded inside a unit” are distinguished at the interface, with nothing inspecting anything else’s internal state. A composition’s persistence as a unit is tied to its participation as a unit; its content survives, if useful, by being inlined into the units that do participate.

Two requirements on the environment, and only two. Chain 6 places fitness as environmental. The specific fitness function is determined by what the architecture is applied to, and the class fixes two things about it, which are separate failures and should not be run together.

- **A gradient must exist.** Without differential survival there is no direction at all, and the failure is the one this chain forces against. This is class-required.
- **The gradient must have a down-component on extent.** Differential survival under a measure indifferent to decoded length gives a directed flow with no direction along the coordinate the lateral ratchet moves. The cycle transports, but not down the axis that needs it, and §8.6’s self-feeding loop runs anyway. This is the class-level constraint §8.6 states, and this is the chain that owns it, because this is where the external measure enters the derivation.

How a given GADS realizes either, the form of the penalty, the length target, the rate, is a property of the environment an instance is run against and not a fact about the class. That the pressure lives in the environment rather than in the architecture is not a gap in the derivation. It is the claim: GADS acts on the substrate, fitness acts on the population, and the Participation Signal is the whole of what passes between them.

Subsystem Specifications

The chains force three subsystems and one set of signal channels: MG (Chain 1), the Participation Signal with its removal rule (Chain 3), MCO (Chain 5), and POP’s signal channels (Chain 6). POP itself is not a chain conclusion; it is established at L0 as the matter all subsystems act on. Inlining (Chain 4) is a rule on removal, not a subsystem; capture and open are operators, not subsystems.

The word *subsystem* is used functionally throughout: a subsystem is a role that must be discharged, not a structure that must be held. MG and MCO are held structures, and are in fact one structure serving two roles. The Participation Signal’s role is discharged where decode is already running and holds no state of its own. The signal is specified under Chain 3 below and appears again under POP’s channels as the thing those channels carry; it is one object, not two.

Every subsystem operates under the locality commitment of the foundation derivation: subsystem state and operations are scoped to explicit interfaces. No subsystem reaches into another’s internal state; all coordination occurs through the signal channels Chain 6 forces.

Meta-Genome (MG)

Derived from: Chain 1.

Theoretical role. The persistent store of captured content. MG holds the compositions the population has compressed, and a composition is its own identity, so the store is a set rather than a mapping. Without MG there is no frame above the atoms: no vocabulary for sampling to offer and nothing for capture to build on.

State. A set $L(t)$ of compositions with insertion order. A composition is an immutable nested sequence over atoms and compositions, and identity is structural equality: two compositions are the same object when and only when their content matches at every depth. A composition’s constituent sequence is the composition; there is no identifier space and no name-to-content table. The reference DAG $\mathcal{G}$ is carried by containment rather than recorded beside it. Cycles are structurally excluded: a composition’s content is drawn from material already present, so it references only compositions that already existed.

Key operations.

- **Insert(m):** admit composition m , fired by capture. Deduplication keys on immediate content, one reference level deep and never on decode, so a repeat capture of live content returns the existing composition rather than admitting a second copy. This is the key §8.1 and §8.6 act on.
- **Decode(m) $\rightarrow$ atomic sequence:** expansion of m from its own content until only atoms remain. Iterative over an explicit stack, so nesting depth is unbounded.
- **Remove(m):** drop m from the set. Fired by the removal rule of Chain 3 and never by anything else. Triggers inlining (Chain 4). Nothing is released, because nothing was allocated.

Decode does not consult MG. This is the property the rest of the architecture rests on and it is worth stating where the state is specified. Decode expands a composition from the content the composition carries, not from a lookup in $L(t)$. An organism holding a composition that has been removed therefore decodes to exactly the atoms it decoded to before, and no measurement taken on decoded content moves when the store changes (Chain 4). What MG governs is the proposal distribution, through sampling, and whether open expands a composition or leaves it whole, through membership. A near-empty store does not stall the machine; it removes the brake on expansion.

Size. Emergent. Nothing here bounds the set and no constant sets a target. The size is the balance of capture inflow against the removal rule of Chain 3, which is where its boundedness and its self-sizing are derived.

Timescale. Slow. Entries persist across many generations, admission occurs on capture, and removal occurs on the per-generation reading. The slow timescale refers to the characteristic lifetime of entries, not the frequency of update events.

Participation Signal

Derived from: Chain 3, carried on the channel Chain 6 forces.

Theoretical role. The regulator of structural persistence and the only test an entry ever faces. Each entry in MG is a claim that a span is worth naming as a unit; the signal is how that claim is tested and the removal rule is how a failed claim is discarded.

State. None. This is the substance of the specification rather than an omission. There is no container, no capacity, no ordering, no age counter and no threshold. The signal is one bit per composition per generation and the rule consumes it in the generation it is raised.

The reading. Let $C(t)$ be the set of compositions appearing as top-level tokens anywhere in P_t at scoring time. The reading is

$$\mathcal{S}(t) = C(t) \cap L(t), \quad (8)$$

and the removal set is $L(t) \setminus \mathcal{S}(t)$.

Key operations.

- **Read:** collect $\mathcal{S}(t)$ during the decode the population is already performing for scoring. No separate traversal, no scan of MG.
- **Sweep:** remove every composition absent from the reading, each removal firing inlining (Chain 4).
- **Attribute:** where inlining has superseded a form an organism holds, the participation of the superseded form is credited to the form that replaced it. Required by Chain 3 for signal integrity; the mechanism is instance-level.

Properties. The signal is *binary* (presence, with no magnitude), *total* (one evaluation pass covers every organism that could be holding a composition at top level), *synchronous* (taken at one instant of the cycle rather than accumulated across it), and *rank-free* (taken over the whole population before reproduction, so no process holds or consumes selection's ordering). Absence from a total, synchronous reading is the observation of non-service, not an inference from it, which is why no threshold is forced.

Scheduling constraint. The reading is valid only for the population it was taken over, so it must be consumed before that population changes:

$$\text{attribute} \rightarrow \text{decode and score} \rightarrow \text{read} \rightarrow \text{sweep} \rightarrow \text{reproduce} \rightarrow \text{mutate}. \quad (9)$$

Sweep after mutation would decide removal from a reading of a population that no longer exists. Structure captured later in the generation faces its first reading in the next one.

Timescale. Metabolic, and the rate is derived rather than chosen: one reading per generation, because the scheduling constraint admits exactly one reading per turnover of the population.

Meta-gene Creation Order (MCO)

Derived from: Chain 5.

Theoretical role. The chronological record of composition discovery and the architecture’s decay mechanism. MCO provides the distribution that governs which compositions are sampled when mutation operators insert composition tokens. Weighting toward recent captures keeps sampling coupled to the population’s current location in the search space and creates the decay curve along which compositions age out of high availability. The structural consequence, that compositions unable to earn their existence at the population level are dissolved into their parents with information preserved and hierarchy flattened, is the architecture’s mechanism for losing self-unjustifying compositions.

State. Not a separate sequence. MCO is the insertion order of MG’s set, which is capture order, so one structure serves membership and ordering both.

Key operations.

- **Append(m):** a new composition enters at the newest position, which is the position sampling weights most. Fired on capture.
- **Remove(m):** fired on removal. Where inlining merges two entries into one, the survivor keeps the position of its first occurrence, so a merge never advances an entry’s sampling weight.
- **Sample()** $\rightarrow m$: draw with probability proportional to $\pi_{13}^{|L|-j-1}$, where j is the entry’s position and the newest entry is $j = |L| - 1$.

Decay parameter. $\pi_{13} \in (0, 1)$ controls how sharply sampling biases toward recent compositions. As $\pi_{13} \rightarrow 1$ sampling approaches uniform; as $\pi_{13} \rightarrow 0$ it concentrates on the newest. The value is application-dependent.

Relation to the Participation Signal. The two order different things and only one of them is derived from participation. MCO orders by creation time, which the signal constraint of Chain 3 does not bind, because creation time is not a record of participation. Nothing anywhere orders compositions by when they were last held.

Timescale. Slow. The order changes only on capture and removal.

POP Signal Channels

Derived from: Chain 6.

Theoretical role. The interface between POP and the metabolic loop. POP is established at L0 as the matter all subsystems act on; the channels are what make POP's activity visible to the roles the chains force. Chain 6's argument is that without differential survival the channels carry no gradient and the loop exhausts nothing.

Signals carried.

- **Capture(ξ):** raised when capture compresses a boundary-delimited span into a composition. Carries the content.
- **Participation:** the signal specified above, raised for a composition appearing as a top-level token in a population member, at most once per member per generation. Raised by decode.

Whole population, not survivors. Participation is raised over every member of P_t at scoring time, not only over members that go on to reproduce. Selection reaches the frame by renormalizing the population between readings rather than by filtering the signal: a member that does not reproduce is absent from the next population, so the compositions only it held are absent from the next reading. No process holds rank, and none is consumed.

Interface-scoped participation. Participation is raised only for top-level tokens. Nested compositions encountered during expansion are not independently credited; a nested composition's persistence is inherited from the top-level composition containing it. Chain 6 gives the argument for why the scoping is forced: a signal raised on nested occurrence tests a claim that discriminates nothing, and crediting it would require the reader to inspect the store's DAG.

The channel returns, it does not push. Decode already takes an encoding and returns a value. It now sometimes returns two, the decoded content and the current structural identity of what was presented, because attribution rides the decode. The organism receives an answer about the token it supplied, on a channel already carrying data in both directions. This is what keeps attribution local: nothing is scanned, no model of the population is consulted, and nothing reaches into an organism to edit it.

Timescale. Fast. Capture is raised per mutation event; participation is raised during scoring. Both reach the metabolic and slow roles within the generation in which they are raised.

Inlining (Rule, not Subsystem)

Derived from: Chain 4.

Theoretical role. A rule on removal, not a subsystem. When MG removes composition m , every structure referencing m has that reference replaced by m 's own content. Without inlining, removal would either corrupt the referencing compositions or require modifying the population, which locality forbids.

Effect. For every $m' \in L(t)$ containing m : replace each occurrence of m in m' with m 's content, and merge structures that become identical under the rewrite. The decoded content of every m' is preserved exactly. The reference DAG flattens by one level at each affected site. Because a composition is its content, this is a rewrite by value and requires no resolution step.

What inlining conserves and what it does not. It conserves each parent's decoded content, so flatten is invariant under it and extent is untouched. It lowers depth. It increases the store's immediate-content storage, since k references to a removed composition become k copies of its content. So removal carries away a distinction while leaving the information, which is the sense in which information flows up and structure flows down. Extent leaves the store only at the leaves, when a removed composition has no live parent to be spliced into.

Whether repeated inlining is self-limiting in storage is a dynamical question and is not settled here. It is not the failure mode of any chain.

Multi-Timescale Coupling

The architecture does not update on one clock, and its behavior emerges from how the rates couple. Three timescales partition it. The separation is not imposed by subsystems carrying distinct clocks. It emerges from participation depth, and is derived rather than designed.

Fast: POP. The stochastic search layer. Operators fire per mutation, the population turns over per generation, and both signals are raised within the generation in which the events occur. This is where variation is generated and where selection is applied. The fast layer is where GADS explores.

Metabolic: the reading. One reading of the Participation Signal per generation, and the sweep it drives. This is not a layer of subsystems but a single act, and its rate is not a tuning choice: the scheduling constraint of Chain 3 requires the reading to be consumed before the population changes, and the population changes once per generation, so exactly one reading per generation is admissible. The intermediate rate is forced.

Slow: MG content and MCO order. The structural memory. Entries persist across many generations; the order is the chronological record with its decay. The slow layer is what GADS remembers.

Where the Separation Comes From

One predicate produces all three rates. A composition survives a generation exactly when some organism holds it at top level, and how long it survives is governed by how many independent processes hold it.

- A composition held by one organism persists as long as that organism's line does, which is a fast-layer quantity: a single reproduction failure ends it.
- A composition held by many organisms persists until all of them fail together, which requires a coincidence whose probability falls sharply as the count rises.
- A composition held across independently converging lineages persists on a timescale far longer than either, because the holders are not correlated and their failures are not simultaneous.

Same predicate, three timescales, separated by how many independent processes hold a thing. Nothing assigns a lifetime and no subsystem carries a clock of its own. The hierarchy of rates is a consequence of the reading’s structure, and the slow layer is slow because deeply supported structure is hard to lose, not because it was placed in a subsystem designated slow.

Coupling Between the Rates

- **Fast feeds slow.** Capture events at the fast layer admit entries to the store. The stochastic search is what discovers the patterns the structural memory holds. Without the fast layer’s variation the slow layer would have nothing to record.
- **Slow shapes fast.** MCO’s decay determines which compositions the fast layer samples when operators insert composition tokens, and the store’s content determines what decoding produces. The slow layer’s accumulated structure biases the fast layer’s search. An entry’s position is read at fast-layer speed though the order changes only slowly: the fast layer consults slow-layer structure without changing it.
- **The reading closes the loop.** The metabolic rate does not mediate between two layers in the sense of holding state between them. It is the single act in which fast-layer behavior becomes slow-layer content: what the population held this generation is what the store keeps. There is no accumulation on either side of it.

Why the Rates Must Separate

Structural memory must be slow, or entries dissolve before they can be reused. Search must be fast, or the per-mutation variation that drives discovery is lost. And the reading must sit between them, which it does by construction rather than by choice: it is downstream of one full generation of fast-layer activity and upstream of the store’s next state, and the scheduling constraint admits it at no other rate.

The characteristic rates overlap at their boundaries, since captures are per-mutation events that nonetheless produce long-lived entries, but the separation of the characteristic rates is what lets each layer perform its role.

Core Interaction Pipelines

The chains force their conclusions one at a time, each answering a structural failure mode, and the subsystem specifications define each in isolation. The architecture runs as composed loops. Two pipelines carry the dominant traffic. They are the backbone rather than the whole interaction graph: deduplication short-circuits, the coupling between MCO decay and the reading, and the DAG rewrites inlining ripples through the store all sit alongside them.

Generative Pipeline (Discovery and Growth)

How a composition comes into being and enters circulation.

1. **Capture fires** (fast). A boundary-delimited $\mathcal{C}_1$ region is compressed into a composition, and the boundary pair is consumed. (Chain 1)

2. **Deduplication resolves** the content against the live set and against any attribution pending on it. Content already live returns the live composition; content matching a superseded form returns the form that replaced it. Either way the population gains a token that exists.
3. **MG admits** the composition if it is new, at the newest position, which is the same act as MCO appending it. (Chains 1, 5)
4. **The composition is already placed.** Capture rewrote the organism it fired in, so the new entry is held from the moment it exists and faces its first reading next generation from inside an organism.

Admission and ordering are a single act, since the store and the order are one structure, so the two cannot disagree.

Metabolic Pipeline (Decay and Death)

How a composition leaves circulation.

1. **MCO decay** (slow state, read fast). As new captures arrive, an entry's position moves down the decay curve and its sampling weight diminishes. It is offered to new solutions less and less, and its persistence comes to depend on being held rather than on being sampled. (Chain 5)
2. **Attribution** (fast). Each organism collects any rewrite applying to it on the decode it is already performing, so the signal it raises lands on the form the store holds. (Chains 3, 4)
3. **The reading** (metabolic). Every organism is decoded and scored; the compositions held at top level are the reading. (Chains 3, 6)
4. **Sweep** (metabolic). Every composition absent from the reading is removed, before the population changes. (Chain 3)
5. **Inlining** (slow). Each removal splices the removed composition's content into every structure referencing it, and identical results merge. Decoded content is preserved throughout. (Chain 4)
6. **Rewrites are recorded** for attribution on the next decode, closing back to step 2. (Chain 3)

Net effect: a composition that stopped being used as a unit is removed, its content preserved in the units that do participate, the hierarchy flattened by one level at every affected site, and every organism holding a superseded form repaired on work it was already doing.

The Closed Loop

The two pipelines share endpoints, but what cycles is not a pool of resources. What cycles is granularity. Capture coarsens: it takes a span the population was expressing as separate positions and makes it one unit. Inlining flattens: it takes a unit the population stopped using and returns its content to the structures that held it. The store's content moves between the two, and the population's decoded content is conserved through every step except removal at the leaves.

Entropy enters through capture, which admits new structure, and leaves through removal, which discards a distinction the population did not support. The rate of exit is set by the Participation Signal, and its *direction* is set by the gradient Chain 6 forces onto that signal. The loop is not conservative: inlining duplicates content into referencers even as it removes a unit, and the store's

size is bounded without being fixed. But a cycle of admission against removal, with a gradient across it, is what makes the architecture an exhaust rather than a churn, and Chain 6 is where the gradient is forced.

The GA-to-GADS Boot Sequence

The architecture comes into being through a phase transition. Before the first capture, the compositional layer does not exist; after it, the full two-layer architecture is running. But the pre-capture phase is not idle, and it is not a vanilla genetic algorithm waiting to become something. It is the boundary-driven L0 regime, already doing structured work.

Phase 1: The Boundary-Driven Regime (Pre-Capture)

From the placement of P_0 at $t = 0$, L0’s dynamics are running. The compositional layer is absent (MG is empty, MCO is empty, no composition tokens exist) but the substrate-level layer is fully active. The boundary mechanism drives the dynamics described in the Layer 0 structural properties: boundaries form, move, and are removed; the persistence gradient τ_1/τ_0 separates content into two regimes; information flows across the $\mathcal{C}_0/\mathcal{C}_1$ boundary through the swap and bulk channels; and stable atomic patterns accumulate inside boundary-delimited regions.

This phase is not waiting for anything. It is the substrate-level engine, exploring and generating structure. What it cannot do is remember that structure as reusable units: a stable pattern formed inside a boundary pair persists only as long as the boundary dynamics sustain it, and is referenced by nothing. The patterns that Phase 1 produces are precisely the raw material that capture will compress, but until capture exists, they remain atomic-sequence patterns, not compositions.

In timescale terms, only the fast layer is active. There is no slow layer yet, because MG and MCO are empty; there is no metabolic layer yet, because there is nothing to regulate. The architecture runs on a single timescale.

Phase Transition: The First Capture

The transition occurs when capture fires for the first time, compressing a stable boundary-delimited pattern into the first composition token. This single event brings the compositional layer into existence:

- MG gains its first entry: the store becomes non-empty, and the search frame acquires a non-atomic half.
- The reference DAG begins, with one node initially, but the structure on which all compositional geometry is defined now exists.
- The structural prior P_{struct} becomes non-trivial. Before first capture, no composition has depth or flat length; the prior has no support. After, the hyperbolic structure of the compositional DAG begins to take form.
- MCO gains its first entry, and the slow layer begins to exist.

This is the boundary between a one-layer and a two-layer system. “One-layer” does not mean nothing was happening; it means only the substrate-level boundary dynamics were active. “Two-layer” means the compositional layer now runs on top of the boundary dynamics, which continue underneath unchanged.

Phase 2: The Two-Layer Architecture (Post-Capture)

With compositions in existence, the full architecture engages. The generative pipeline runs as captures continue to fire; the metabolic pipeline engages as the store accumulates and the reading begins regulating content. The slow, metabolic, and fast timescales all become meaningful: the timescale separation described earlier exists only once the compositional layer is present to occupy the slow and metabolic scales.

The boundary dynamics of Phase 1 do not stop. They continue to run at the substrate level, forming the boundary-delimited regions that capture acts on. The two-layer architecture is the boundary-driven regime with the compositional layer added on top: the fast layer continues to explore and to form patterns; the slow layer now remembers selected patterns as compositions; the metabolic layer now regulates what is remembered. What the phase transition adds is not dynamics where there were none, but a second layer of dynamics, operating on a slower timescale, under a structural prior, coupled to the boundary dynamics that were running from $t = 0$.

8 Search Space and Compositional Geometry

This section develops the geometry of the search space a GADS navigates once the compositional layer is active. It is the formal development the preceding chains defer to: Chain 1 establishes that capture and MG produce a compositional DAG carrying a structural prior $P_{\text{struct}}(m) \propto \exp(-d_H(m))$, and states that the form of d_H and its connection to the variational free energy are developed here.

The central result is that the compositional cost

$$d_H(m) = \text{arccosh}(1 + \text{depth}(m)) \cdot \text{flatlen}(m) \quad (\star)$$

is not a cost function chosen for convenient asymptotics. It is the hyperbolic extent of the cost-geometry subspace a composition occupies, with both factors forced by structure already fixed in the substrate and the chains. Section 8.5 earns this form by construction.

A note on what is and is not claimed here. Everything in Sections 8.1 through 8.8 is structural and earns the word “forced.” Whether the resulting dynamics converge, whether influence-decay is exponential rather than polynomial, whether the architecture equilibrates rather than collapses, is path-dependent and statistical, and not proven here. The boundary is deliberate.

8.1 The Structure Space Is a DAG of Growing Multiplicity

Before any metric is introduced, one fact about the structure space must be settled by counting, because the entire curvature development rests on it and because it is the fact most often mistaken. The space of compositional structures over a solution is *not a tree*. It is a directed acyclic graph whose nodes carry in-degree greater than one, and the number of structures decoding to a fixed solution grows exponentially with the solution’s length. These are not modelling choices; they are arithmetic, and this subsection states them so that nothing downstream depends on intuition about “tree-like” growth.

The objects. Fix a base alphabet $\mathcal{T}_A$ (the atoms, each of decoded length 1) and a finite inventory B of available block lengths, where a block is a unit the representation may place in one editable position: $B \ni 1$ always (atoms), and B acquires a length ℓ whenever a composition of decoded length ℓ is live in $L(t)$. A *structure* over a region is an ordered factorization of the region’s content into blocks whose lengths lie in B ; its decode is the atomic sequence obtained by expanding every block.

Two structures are the *same* iff their immediate content (one reference level deep, compositions left as tokens) is identical. This is exactly the key deduplication acts on: capture tests whether the candidate content is already live, comparing immediate content and never decode. Two structures are *equivalent* iff they decode to the same atomic sequence.

Definition 8.1 (Fiber and its multiplicity). *For an atomic region Σ of decoded length n , the fiber over Σ is the set of distinct structures decoding to Σ , and its multiplicity is $\mu(\Sigma) = |\{\text{structures} : \text{decode} = \Sigma\}|$. Because block contents are forced by the underlying string (a block of length ℓ at a position has exactly one content, the substring sitting there), μ depends on Σ only through its length n and the inventory B , and equals the number of ordered tilings of n by parts in B :*

$$\mu_B(n) = [x^n] \frac{1}{1 - \sum_{b \in B} x^b}, \quad \mu_B(n) = \sum_{b \in B} \mu_B(n - b), \quad \mu_B(0) = 1. \quad (10)$$

Lemma 8.1 (The structure space is a DAG with reconvergence, not a tree). *The reference relation on live compositions is a directed acyclic graph $\mathcal{G}$ in which some node has in-degree greater than one, and this is constructible by capture and not removable by deduplication. Hence $\mathcal{G}$ is not a tree.*

Proof. Acyclicity is immediate: a composition references only compositions created strictly before it (a capture's content is drawn from material already present), so $\mathcal{G}$ admits the creation-order topological sort and contains no cycle. For reconvergence, exhibit a forced instance. Let m be a composition with immediate content (a, b) over atoms. Capture acting on a region (ma) produces p with immediate content (m, a) ; capture acting on a region (bm) produces q with immediate content (b, m) . Then m is referenced by both p and q , so m has in-degree two in $\mathcal{G}$. Deduplication keys on immediate content; p and q have distinct immediate content, so neither is merged away, and m is a single shared identity beneath two distinct parents. A tree requires every node to have in-degree at most one; $\mathcal{G}$ violates this. The construction uses only capture and the deduplication rule, both forced by Chain 1, so no admissible variant of the architecture avoids it. $\square$

Lemma 8.2 (Multiplicity exceeds one and grows exponentially). *Let B contain 1 and at least one block length $b \geq 2$. The hypothesis $1 \in B$ is not an assumption but a structural guarantee: the base space exists by virtue of holding a population, so atoms are always a live block length. In particular $\gcd(B) = 1$, so the recurrence (10) is aperiodic. Then for all $n \geq b$, $\mu_B(n) > 1$; the map from structures to decoded solutions is therefore strictly many-to-one. Moreover $\mu_B(n)$ grows exponentially: $\mu_B(n) \sim C \lambda^n$, where $\lambda > 1$ is the unique positive root of $\sum_{b \in B} x^{-b} = 1$ (equivalently the reciprocal of the smallest positive root of $1 - \sum_{b \in B} x^b$). The degenerate case $B = \{1\}$ gives $\mu_{\{1\}}(n) \equiv 1$, the unique tiling is all atoms, and is the only inventory for which the structure space collapses to one structure per solution.*

Proof. For $n \geq b$ the all-atoms tiling and any tiling using one b -block are distinct structures with the same decode, so $\mu_B(n) \geq 2$. The generating function $(1 - \sum_{b \in B} x^b)^{-1}$ is rational with the recurrence (10); since $1 \in B$ forces $\gcd(B) = 1$, the recurrence is aperiodic and its companion matrix irreducible aperiodic, so Perron–Frobenius gives a unique smallest-modulus pole, real and simple, and $\mu_B(n) \sim C \lambda^n$ with λ its reciprocal, the unique positive root of $\sum_{b \in B} x^{-b} = 1$. Since $\sum_{b \in B} x^{-b}$ is strictly decreasing in $x > 0$ and equals $|B| > 1$ at $x = 1$, the root satisfies $\lambda > 1$. For $B = \{1\}$ the recurrence is $\mu(n) = \mu(n - 1)$ with $\mu(0) = 1$, giving $\mu \equiv 1$, and any larger inventory contains a $b \geq 2$ and falls under the first claim. $\square$

Proposition 8.1 (The fiber is non-constant, so the structure space is not flat). *$\mu_B(n)$ is strictly increasing in n for n beyond the smallest nontrivial block, and for fixed n it is strictly increasing*

under inventory inclusion ($B \subsetneq B' \Rightarrow \mu_B(n) \leq \mu_{B'}(n)$, strictly once B' adds a block of length $\leq n$). The fiber multiplicity therefore varies across regions and across the live inventory. A bundle whose fiber has constant cardinality is flat; a fiber of varying cardinality cannot be flat. Hence the structure space carries non-trivial curvature, and the rate at which μ varies is the rate at which that curvature varies across $L(t)$.

Proof. Monotonicity in n follows from (10): each term $\mu_B(n-b) \geq \mu_B(n-b-1)$ by induction once μ has begun to grow, and the $b=1$ term alone gives $\mu_B(n) \geq \mu_B(n-1)$. Monotonicity in B is termwise: adding a block length introduces a new non-negative summand to the recurrence, strictly positive as soon as n reaches that length. A locally trivial bundle with structure group acting on a fiber of fixed cardinality has, by definition, the same fiber over every base point; a base-dependent fiber cardinality is an obstruction to such triviality. The variation of μ is thus incompatible with a flat (constant-fiber) structure, which is the qualitative content recovered metrically in Sections 8.5.2 and 8.6 as coarse, non-constant negative curvature. $\square$

What this fixes, once. The question “is the compositional structure space a tree?” is decided by Lemma 8.1 and Lemma 8.2: a tree is exactly the case $\mu \equiv 1$ (in-degrees ≤ 1 , one structure per decode), which holds only for the atoms-only inventory $B = \{1\}$ and fails the instant any composition is live. The question “is it flat?” is decided by Proposition 8.1: flatness requires constant fiber cardinality, and μ is provably non-constant. Both questions reduce to evaluating the rational generating function (10), which has one value for each (B, n) ; there is no residual matter of interpretation. The metric development that follows, the fold (§8.5.2), the hyperboloid as forced local model (§8.5.3–8.5.4), and the content fiber with its depth-and-multiplicity profile (§8.6), is the quantitative refinement of these counted facts, not an independent posit. Where §8.5.2 speaks of multiplicative growth of reachable volume and §8.6 of a multiplicity $\mu(\Sigma, \delta)$ that varies across $L(t)$, the present subsection is the source: the volume is $\mu_B(n)$, its growth is Lemma 8.2, and its variation is Proposition 8.1.

Why this is not the exponential growth of a grid. A standard objection: products of trees and integer grids also grow exponentially without being negatively curved, so growth alone proves nothing. The objection does not apply, and a worked instance shows why the difference is structural rather than rhetorical. Take a discrete cube $\{0, \dots, 9\}^3$ and the point $(3, 5, 7)$. Capture the pair $(5, 7)$ into a unit m : a token that, placed on an axis, reads as the ordered pair $(5, 7)$ with the remaining axis free. The point is now representable as $(3, m)$: first axis 3, second coordinate the single token m , third axis free.

Two facts hold simultaneously, and their conjunction is the whole point. First, the flat configuration is preserved: $(3, 5, 7)$ remains a point of the cube, reachable by placing 5 and 7 independently, so capture deletes nothing from the base space and its cardinality is unchanged. Second, the token installs a *reusable section*: m is available wherever the pattern $(5, 7)$ may sit, and at every such site it fuses two transverse axes into one. Reaching the $(5, 7)$ cell formerly cost two independent selections ranging over a 10×10 plane; through m it costs one selection of the token. The token does not enlarge the base space; it lowers the dimension of the subspace that must be searched by independent moves to reach the cell, and it does so at every site it occupies. This is the dual of Lemma 8.2 read on the volume side rather than the count side: the fiber over the $(5, 7)$ cell gains an element (count up), and the reachable subspace per representational step drops in dimension (volume down); the two are the same fold seen from opposite directions.

A grid cannot do this. Its axes are global but independent: no coordinate of a grid fixes a transverse pair of other coordinates and exports that fusion to every site while leaving the unfused

configuration reachable. The grid’s exponential growth is additive over independent directions (no geodesics converge; triangles stay fat); capture’s growth is an identification that creates a short route between configurations that remain far apart under the flat metric (the long route survives), so geodesics converge and triangles thin. Convergence of geodesics under a preimage-preserving identification is the discriminating feature: it is present here by construction and absent in any product of grids or trees. The curvature appears in the fused axes and not in the free one (here z stays flat), which is precisely why the geometry is coarse and non-constant rather than a uniform $\mathbb{H}^n$: capture folds the axes it captures and leaves the rest flat, and which axes are folded is path-dependent.

8.2 The Search Frame

A GADS does not search a fixed space. The space it searches is defined by what it can currently express, and that changes as the compositional layer grows.

Definition 8.2 (Search Frame). *At time t , the search frame is the pair*

$$\mathcal{F}(t) = (G, L(t)), \quad (11)$$

where $G = \mathcal{T}_A$ is the fixed base alphabet committed at instantiation, and $L(t)$ is the set of compositions live in MG at time t .

The base alphabet G is fixed; $L(t)$ grows and shrinks as captures add compositions and the reading removes them. The search frame is the vocabulary available to the mutation operators at time t : the tokens they may insert are the base genes of G together with the live compositions of $L(t)$. It is what Chain 1 forces MG to be.

Relationship to the substrate. The substrate is the base space $\mathcal{B}_t = (\mathcal{T}_A \cup \{\emptyset\})^{D_t}$: flat combinatorial volume over atoms, with population-induced dimensionality. The substrate sees only atoms; compositions are invisible to it. The search frame is not a modification of the substrate. It is the vocabulary, above the substrate, in which the population is written. A composition in $L(t)$ is a token the population may carry; when decoded, it lays atomic content onto the flat base space. The frame describes what can be expressed; the base space describes where the expressed content lands. The two are distinct objects, and the substrate remains flat regardless of the frame’s contents.

8.3 Base Volume and Pattern Regions

Definition 8.3 (Base Volume). *The base volume at time t is the cardinality of the base space:*

$$V_t = (|\mathcal{T}_A| + 1)^{D_t}, \quad (12)$$

the number of distinct points the substrate can represent at dimensionality D_t (the $+1$ is the absence marker $\emptyset$).

Compositions do not change the base volume. A composition is a shortcut to a specific configuration of atoms; the configuration was always a point in $\mathcal{B}_t$, reachable in principle by placing the atoms directly. What a composition changes is not what exists in the base space but what is *accessible* to the operators in a single step.

Definition 8.4 (Pattern Region). *For a live composition $m \in L(t)$ with decoded content $\xi^*(m)$ of length $\text{flatlen}(m)$, the pattern region $\Sigma_m \subseteq \mathcal{B}_t$ is the set of base-space points whose coordinates on the $\text{flatlen}(m)$ axes where m 's content lands are fixed to $\xi^*(m)$, with the remaining axes free.*

Σ_m is a flatlen -dimensional coordinate subspace of the flat base space, a real region of $\mathcal{B}_t$, the configurations the composition can place. Compositions do not add points to the base space; they concentrate single-step accessibility into specific pattern regions. This concentration is the geometry the rest of the section develops.

8.4 Effective Dimensionality

Definition 8.5 (Effective Dimensionality). *The effective dimensionality of a population P_t is the mean decoded length of its members:*

$$\text{Dim}_{\text{eff}}(P_t) = \frac{1}{|P_t|} \sum_{x \in P_t} |\text{Decode}(x)|, \quad (13)$$

the average number of base-space axes a population member pins when fully decoded.

A population member written with composition tokens is shorter, in encoded length, than its decoded content. The compositional layer lets the population occupy a high effective dimensionality, pin many base-space axes, with a short encoded representation. This gap between encoded length and effective dimensionality is the compression the compositional layer provides, and it is what the compositional cost d_H measures the price of.

8.5 Compositional Cost as the Extent of a Cost-Geometry Subspace

This subsection earns $(\star)$. The order of argument is deliberate. Fixing a hyperboloid embedding first and recovering the arccosh form as its distance would make the cost formula appear assumed rather than derived, since the embedding would have been chosen to produce it. Here the curvature comes first: it is forced by what a composition *does* to single-step reachability, with no cost function and no embedding in hand. Only once the cost geometry is established to be hyperbolic do we introduce the hyperboloid, as the forced local model of a structure already shown to be negatively curved, and read $(\star)$ off as its distance. The formula is the consequence; the curvature is the premise. This reordering is what avoids the circularity.

The base space is flat (§8.3) and stays flat throughout. What carries curvature is not the base space but the cost geometry: the geometry of *reaching* configurations through compositions. The claim has two halves of different strength:

1. **(global, forced: curvature).** Adding a composition folds the reachability structure over the base space. Each composition is a coordinate that collapses a flatlen -dimensional region of the base space to a single editable token, so single-step-reachable volume grows multiplicatively with resolution depth, and compositions referencing compositions compound the folds. This multiplicative growth, carried on the reconvergent DAG of §8.1 (Lemma 8.1, Lemma 8.2), is coarse (Gromov) hyperbolicity: growth alone is necessary but not sufficient; the reconvergent in-degrees supply the rest. The cost geometry is therefore negatively curved, and this is forced by the act of capture alone, before any cost function is named. The spreading rate varies across $L(t)$ because branching varies across compositions, so the curvature is coarse, not constant.

2. **(local, exact: the formula).** Because the cost geometry is hyperbolic, the constant-curvature hyperboloid $\mathbb{H}^n$ is its exact local model at a single fiber element. Under that model the cost-extent of one parse is exactly $\text{arccosh}(1+\text{depth}) \cdot \text{flatlen}$. This is exact within one parse, where the content lays onto disjoint axis-blocks at a common resolution depth and the product structure is genuine.

We establish (global) first (§8.5.2), because it is the premise. Then (local) follows (§8.5.3–8.5.4): given that the geometry is hyperbolic, the hyperboloid is not a modelling choice but the forced local model, and $(\star)$ is its distance. We do not claim the global cover is a smooth Riemannian manifold of constant curvature; it is a discrete object that is coarsely hyperbolic, with $\mathbb{H}^n$ borrowed pointwise as the exact local model.

8.5.1 The resolution cover

Before the curvature argument we name the object that carries it.

Definition 8.6 (Resolution Cover). *Over the base space $\mathcal{B}_t$ sits the resolution cover $R(t)$. Above each base point $b \in \mathcal{B}_t$ lies its fiber R_b : the set of parses of b , the distinct factorizations of b 's content through the live compositions $L(t)$. Each fiber element $r \in R_b$ carries two structures:*

- a radial depth $\text{depth}(r)$, the resolution depth of the parse;
- a tangent neighborhood $N(r) \subseteq \mathcal{B}_t$, the set of base points reachable from r by one depth-preserving operator firing on the representation r .

The projection $\pi : R(t) \rightarrow \mathcal{B}_t$, $r \mapsto b$ is many-to-one. Distinct fiber elements over the same base point carry distinct tangent neighborhoods: same decoded content, different single-step reachability.

The tangent / radial / cover-growing partition. The operators of the architecture partition into three classes by their action on the cover:

- **Tangent (depth-preserving).** Point mutation on exposed positions, swap, single-token insertion and deletion, boundary creation and removal. These act on the encoded representation r without changing the parse's depth structure. They generate $N(r)$.
- **Radial (depth-changing, between existing fiber elements).** Open descends the cover: it exposes a composition's constituents, decreasing parse depth at that position. Capture firing on existing MG content ascends the cover, raising parse depth at that position without growing $L(t)$.
- **Cover-growing.** Capture firing on content not already in MG. This is the only operator that grows $L(t)$, and so the only operator that grows the cover itself: new fiber elements become available over every base point whose content can be parsed through the new composition.

$N(r)$ is defined over tangent operators only. The curvature argument of §8.5.2 uses tangent operators for the lateral neighborhood and the radial operator open for the fracture move. Cover-growing operators are how the cover is built (the boot sequence and the chains' forcing of capture), not motion on a given cover; they are offstage from the curvature argument and appear only when the cover itself is changing between time steps.

Two parses, one base point. The base point $ABCD$ admits, among others, the flat parse $ABCD$ (depth 1) and the parse (AB, CD) (depth 2) when $AB, CD \in L(t)$. These decode identically and are the same point of $\mathcal{B}_t$. They are distinct points of $R(t)$: from (AB, CD) , a single point mutation $AB \rightarrow CD$ reaches (CD, CD) , a base point four axes from $ABCD$, in one firing; from the flat parse $ABCD$ no single firing reaches it. The representations differ not in what they express but in what they can become in one step. The degrees of freedom are a property of the fiber element, not of the base point. Deduplication operates on content uniqueness and is invisible at the parse level: (AB, CD) and $ABCD$ have entirely different decode paths and would never be conflated by content-dedup; they coexist as distinct fiber elements over the same base point.

8.5.2 Why the cost geometry is curved: the fold

This is the premise of the whole construction, and it is settled here without reference to any cost function. We show that the act of adding a composition curves the reachability structure negatively.

A composition is a fold, not a point. Capture adds nothing to the base space. The base point b whose content a composition m pins was always a point of $\mathcal{B}_t$, reachable in principle by placing its atoms directly (§8.3). What capture adds is a *coordinate*: a single token that, in one editable position, pins the $\text{flatlen}(m)$ -dimensional pattern region Σ_m . Before m , reaching a configuration in Σ_m from a neighbouring one took up to $\text{flatlen}(m)$ independent operator firings, one per atom. After m , the whole region is one token, traversed or swapped in a single firing. The composition does not enlarge the space; it *folds* it, bringing a flatlen -dimensional region within a single step. This is the geometric content of “adding a composition is adding negative volume”: a fold along which distances contract, a direction in which far-apart configurations become single-step-adjacent.

The fold compounds multiplicatively. Fix a fiber element r and fracture it one resolution level finer, exposing one composition token as its constituents (the open operator, Chain 2). Each newly exposed constituent is an independently editable position: a point mutation on it reaches a base point that no single firing on the unfractured token could reach. The exposed constituents vary independently, so the set of single-step destinations after fracturing is the *product* of the per-constituent destination sets, not their union. Exposing one more resolution level multiplies the local reachable neighborhood by the branching at that level. Because compositions may reference compositions, fracturing one level can expose further composition tokens, and the multiplication recurs: reachable volume grows multiplicatively with resolution depth.

Multiplicative reachability growth is negative curvature. The growth is not asserted; it is counted. The number of structures reachable over a region of decoded length n is the fiber multiplicity $\mu_B(n)$ of §8.1, and Lemma 8.2 establishes $\mu_B(n) \sim C\lambda^n$ with $\lambda > 1$: reachable structure grows exponentially in the radial coordinate (resolution depth tracks decoded length through the reference tower), and the base of the exponential is fixed by the inventory’s characteristic equation. We are deliberately careful not to overclaim the inference. Exponential volume growth is *necessary* but not by itself *sufficient* for Gromov hyperbolicity; the sufficient ingredient is the tree-plus-reconvergence structure of the DAG (Lemma 8.1), whose geodesic spreading is the divergence of distinct fracture routes through reconvergent nodes. It is the conjunction, exponential multiplicity *carried on a DAG with in-degree* > 1 , that yields coarse negative curvature, not the volume growth alone. The radial coordinate along which the growth occurs is resolution depth, and this is forced by the fold structure together with the counted multiplicity, by what capture and open do to

reachability, with no cost function assumed and no embedding chosen. The curvature is a fact about the cover before d_H is named.

We are careful about which claim this earns. The cover is a discrete object, a set of parses with single-step reachability between them, not a smooth Riemannian manifold, and it carries no curvature tensor of its own. What the fold argument together with §8.1 forces is the coarse, large-scale fact: reachable structure blows up exponentially with depth (Lemma 8.2) on a graph whose reconvergent in-degrees (Lemma 8.1) make the spreading tree-like rather than Euclidean, which is coarse hyperbolicity in the large. It does *not*, on its own, establish smooth sectional curvature, because there is no smooth surface on the discrete cover to carry one. That branching varies across $L(t)$, a composition viable beside only one partner exposes one direction while one viable broadly exposes many, is the statement that the spreading rate is non-constant across the cover.

Deduplication of the DAG does not threaten this. The relevant object is the unfolded expression on the reachability side, where independently editable positions are genuinely distinct (cf. the dimension lemma, §8.5.3). And dedup operates on content uniqueness at MG, not on parses; the fiber elements over a single base point remain distinct as parses regardless of whether any of their constituents are the same composition.

The hyperboloid enters here, as a consequence. Because the cost geometry is now established to be negatively curved, we model it *locally* with the constant-curvature space of that signature, the hyperboloid

$$\mathbb{H}^n = \{x : \langle x, x \rangle_M = -1, x_0 > 0\}, \quad \langle x, y \rangle_M = -x_0 y_0 + \sum_{j \geq 1} x_j y_j. \quad (14)$$

The Minkowski sign-flip on x_0 makes the radial coordinate contribute negatively to the ambient quadratic form: the cost geometry is a sign-flipped extension of the flat base space, with depth on the flipped axis. The sign-flip is a feature of the cost metric, not a modification of the base space’s combinatorial volume. $\mathbb{H}^n$ is introduced as the *exact local model* of a structure already shown to be hyperbolic, not as an assumption that makes hyperbolicity appear; the global cover stays discrete and coarsely hyperbolic, with $\mathbb{H}^n$ borrowed pointwise. The next subsections show depth is the timelike coordinate and read off the per-fiber cost as the hyperboloid distance, now a derivation rather than an assumption, because the curvature was settled independently above.

8.5.3 What a symbol fixes in the base space

A base gene $g \in \mathcal{T}_A$ placed at a position fixes exactly one axis of $\mathcal{B}_t$: it pins one coordinate and says nothing about the others. This is what “a token occupies a position” means in the base space. A base gene occupies a one-dimensional coordinate subspace, and it pins it directly: no reference must be resolved to learn which atom sits there.

A composition m decodes to a unique atomic sequence $\xi^*(m)$ of length $\text{flatlen}(m)$. When m appears at a position, decoding lays that sequence onto $\text{flatlen}(m)$ consecutive axes of the base space: each decoded atom fixes one axis, exactly as a base gene does.

Lemma 8.3 (Dimension). *A composition m occupying a position fixes a coordinate subspace $\Sigma_m \subseteq \mathcal{B}_t$ of dimension $\dim \Sigma_m = \text{flatlen}(m)$.*

This is the first earned fact. flatlen is not a multiplier applied to a per-symbol scalar from outside; it is the dimension of the base-space region the symbol pins, the dimension of the region the fold of §8.5.2 collapses to one token. The product in $(\star)$ will be “cost-per-axis $\times$ number of axes,” and the lemma supplies the number of axes.

Why flatlen is unfolded, and what unfolding does not claim. The recursive computation of flatlen sums a child’s flatlen once per reference path, so a child reached by two paths contributes twice. This is not an inconsistency with MG’s folded storage. MG stores one composition (storage is folded); the base space receives flatlen pinned axes (expression is unfolded). Within a single placement, the pattern region Σ_m lives on the expression side: when one composition is decoded at one locus, its two internal reference paths to the same content lay that content onto two disjoint axis-blocks of $\mathcal{B}_t$, so $\dim \Sigma_m$ equals the unfolded length. flatlen is the dimension of Σ_m by construction, and the dimension is unfolded because, within one placement, the axes are distinct.

What unfolding does *not* claim is that distinct compositions pin distinct regions. Two different compositions may decode to the same atomic sequence and, placed at the same locus, pin the *identical* axis-block: the same Σ , not disjoint ones. The disjointness above is between reference paths inside one placement; it says nothing about two compositions that resolve to the same content by different internal structure. That second relation is the content fiber of §8.6, and it is where the cost geometry ceases to be a single hyperboloid.

8.5.4 Depth is the radial coordinate; the cost-distance and the product

A base gene pins its axis directly. A composition does not: to learn which atom sits on a given axis of Σ_m , a process must expand m , and the number of resolution steps before only atoms remain is $\text{depth}(m)$ (the recursion depth of ξ^* , computed as $1 + \max_{\text{child}} \text{depth}$; base genes have depth 0). This is the same resolution depth along which the fold of §8.5.2 compounds, the radial coordinate of the cost geometry.

Two facts about this displacement are forced:

- **It is orthogonal to the content axes.** The depth of m is a property of how its content is reached, not of which atoms the content is. It is independent of the values on the flatlen axes. So depth enters as a coordinate separate from the content axes, a coordinate in the cost geometry, not in the base space.
- **It is monotone and unbounded in depth, with base genes at zero.** A base gene (depth = 0) sits in the base alphabet: zero cost displacement, reached directly. Each extra level of nesting strictly increases the resolution cost. The coordinate carrying depth must be 0 at depth = 0 and increasing thereafter.

The embedding is now determined, not chosen. We are no longer free to pick an embedding to produce a desired formula: the geometry is hyperbolic (§8.5.2), so the local model is $\mathbb{H}^n$, and depth must sit on the timelike coordinate (it is the radial direction along which reachable volume grew multiplicatively, i.e. the negatively-signed axis). Embedding a single content axis carrying resolution depth δ into the local model is then fixed by three constraints, all already in hand: on the hyperboloid; vertex at depth 0 (base genes reached directly); depth carried by the timelike coordinate. These give

$$\phi_{\text{axis}}(\delta) = (x_0, x_1) \in \mathbb{H}^1, \quad x_0 = 1 + \delta, \quad x_1 = \sqrt{(1 + \delta)^2 - 1}. \quad (8.1)$$

There is no remaining freedom: $\langle \phi_{\text{axis}}, \phi_{\text{axis}} \rangle_M = -(1 + \delta)^2 + [(1 + \delta)^2 - 1] = -1$ places it on $\mathbb{H}^1$; $\delta = 0 \Rightarrow (1, 0)$, the vertex, fixes the normalization to $1 + \delta$ (only $c = 1$ puts the vertex at depth 0); and $x_0 - 1 = \delta$ puts depth on the timelike axis. The embedding is the unique point of the local model consistent with constraints the curvature argument already forced: it is read off the established geometry, not posited to create it.

The cost-distance. The hyperboloid distance from the origin $o = (1, 0)$ to (8.1) is

$$d_{\mathbb{H}}(\phi_{\text{axis}}(\delta), o) = \text{arccosh}(-\langle \phi_{\text{axis}}(\delta), o \rangle_M) = \text{arccosh}(x_0) = \text{arccosh}(1 + \delta), \quad (8.2)$$

using $-\langle (x_0, x_1), (1, 0) \rangle_M = x_0$. So per content axis, the cost displacement of m from the base alphabet is $\text{arccosh}(1 + \text{depth}(m))$. The $\text{arccosh}(1 + \delta)$ form is the distance in the local model of a geometry whose hyperbolicity was established before this computation; it is the consequence of the fold, not an input chosen to appear.

The product. By the dimension lemma, Σ_m spans $\text{flatlen}(m)$ content axes. The radial coordinate of the subspace is m 's own resolution depth $\delta = \text{depth}(m) = 1 + \max_{\text{child}} \text{depth}$: the number of resolution steps to expose m fully, a single property of the composition rather than a per-axis quantity. The cost-extent of the subspace is the radial displacement at that depth applied across its $\text{flatlen}(m)$ axes, the per-axis displacement times the number of axes:

$$d_H(m) = \sum_{i=1}^{\text{flatlen}(m)} d_{\mathbb{H}}(\phi_{\text{axis}}(\delta), o) = \text{flatlen}(m) \cdot \text{arccosh}(1 + \text{depth}(m)), \quad (15)$$

which is $(\star)$. The depth is charged at m 's resolution depth, not resolved per axis: d_H is the cost-extent of the pinned subspace at the depth at which m becomes available as a unit, which is the cost the architecture actually pays to place m in one editable position. (Where a composition's children sit at unequal depths, the constituent depths are recovered only by descending the cover, the radial operator open, which is motion to a different fiber element with its own d_H , not a refinement of this one. Per fiber element the charge is exact; the max is the resolution depth of *this* unit, not an averaging over its content.)

Proposition 8.2 (Earned form of d_H , local). *For a single fiber element r of the resolution cover, the cost geometry is negatively curved by the fold argument (§8.5.2); its exact local model is $\mathbb{H}^n$; the embedding (8.1) is the unique point of that model consistent with the forced constraints; and the cost-geometry extent of r is exactly*

$$d_H(r) = \text{arccosh}(1 + \text{depth}(r)) \cdot \text{flatlen}(r). \quad (16)$$

The factor $\text{arccosh}(1 + \text{depth}(r))$ is the per-axis radial displacement (8.2); the factor $\text{flatlen}(r)$ is $\dim \Sigma_r$ (the dimension lemma). The product is cost-extent per axis times number of axes.

The exactness is local to the fiber element: one parse pins disjoint axis-blocks at a common radial depth, and the product of per-axis displacement and axis count is the hyperboloid extent of that subspace. The global cover $R(t)$ is coarsely hyperbolic with a spreading rate that varies across $L(t)$; $\mathbb{H}^n$ is its exact local model at r , not a smooth global geometry.

When a composition m appears as a top-level token in an organism, the organism's parse identifies a specific fiber element r at that position, and $d_H(m)$ is read as $d_H(r)$ for that parse. The proposition is stated on r because the cover, not the composition in isolation, is where the geometry lives; $d_H(m)$ is read under the standing convention that a composition token in an organism is read at the parse the organism realizes.

What is forced and what is chosen. This is the one place in the geometry where a choice is made, and it is worth stating once and exactly. The curvature is not chosen: it is counted in §8.1 (the structure space is a non-flat DAG of growing multiplicity) and refined metrically here, independent

of how cost is later quantified. Given that curvature, the local model and the embedding are forced, and $(\star)$ follows as the local distance. The single modelling commitment is the identification of cost-extent with hyperboloid distance, measuring the price of a composition by the radial extent of the subspace it pins. That identification is natural and used throughout, but it is adopted, not derived; the curvature beneath it is derived. This is the honest scope of “earned”: the geometry is forced, the metric-as-cost identification is adopted, and nothing else in §8 reopens this question.

Corollary (occupancy is dynamics, not geometry). Which fiber element over a given base point the population occupies is set by capture history and MCO sampling, not by the geometry. d_H is forced as the fiber metric; the population’s position in the cover is path-dependent. MCO decay (§8.7) is the temporal evolution of cover occupancy: as fracture-routes age out of sampling weight, the population drifts across sheets.

Why the factors multiply. The two factors multiply rather than add because Σ_m is a product subspace pinned at a single radial displacement, m ’s resolution depth. Adding them would assert depth and length are alternative costs on the same axis; they are not. Length counts axes, depth displaces the subspace as a whole. A shallow-but-large composition (small arccosh factor, many axes) and a deep-but-small one (large arccosh factor, few axes) are distinguished correctly because the geometry multiplies a per-axis displacement by an axis count, the only combination consistent with Σ_m being a flatlen-dimensional region at radial depth depth.

What stays flat. The base space is untouched by this construction. Σ_m is a genuine flatlen-dimensional region of the flat base space, the axes the content pins. The hyperbolic structure is on the cost coordinate alone: the radial displacement measuring resolution distance from the alphabet. flatlen lives in the base space (real axes); depth lives in the cost geometry (resolution distance). The substrate remains the flat combinatorial volume it was defined to be; what is hyperbolic is the geometry of reaching configurations through compositions, not the configurations themselves.

8.6 The Parse Fiber over a Base Region

The local Proposition is exact for one parse. The step from one parse to the cover is not a technicality: it is where the single hyperboloid fails, and it fails for a structural reason fixed by the substrate, not by any trajectory.

Definition 8.7 (Content fiber over a region). *A composition’s identity is its immediate content: the literal token sequence it carries, one reference level deep, not unwound. This is the key deduplication acts on: two compositions with the same immediate content are the same composition; two with different immediate content are different, even when they decode to the same atomic sequence. For an atomic sequence ξ pinnable at a locus, the content fiber $\Phi(\xi)$ is the set of distinct immediate-content tuples whose decode is ξ and which capture can deposit under the rules of Chain 1.*

$\Phi(\xi)$ is a property of ξ and the admissibility rules alone: combinatorial, not historical. It does not depend on which members are live at any time, only on which immediate contents the substrate permits. For flatlen ≥ 3 it generically has more than one element. The sequence $\xi = abb$ is the decode of the flat content $\langle abb \rangle$ at depth = 0 and of the content $\langle ac \rangle$ where c is a composition with content $\langle bb \rangle$, at depth = 1: two distinct immediate contents, hence two distinct compositions, with one decode. Deduplication does not merge them: it keys on immediate content, and their immediate contents differ. A four-atom region admits a content of depth 1 and a content of depth 2 decoding alike, again distinct as identities.

Every member of $\Phi(\xi)$ pins the *same* pattern region,

$$\Sigma_m = \Sigma_{m'} \quad \text{for all } m, m' \in \Phi(\xi), \quad (17)$$

because Σ is fixed by ξ and the locus (Definition, Pattern Region) and all members share ξ . They differ in immediate content, and so may differ in depth. Crucially the cost geometry does not resolve them by content: $d_H = \text{arccosh}(1 + \text{depth}) \cdot \text{flatlen}$ is built from decode-side quantities only (formula length and resolution depth), so two content-distinct members of $\Phi(\xi)$ of equal depth carry identical d_H . The geometry sees decoded formula and depth; it does not see immediate content. It is content-blind by construction, and the content fiber is therefore strictly finer than the cost geometry.

Lemma 8.4 (Content fiber and cost multiplicity). *The members of $\Phi(\xi)$ pin one region Σ . The preimage of Σ under the cost projection is not a point but a fiber: at each radial depth δ realised by some member, the cost point (Σ, δ) is occupied with multiplicity equal to the number of content-distinct members of $\Phi(\xi)$ of depth δ . Identity lives in the fiber; d_H lives on the base of the projection; the map forgetting immediate content and retaining (Σ, depth) is the coarsening through which d_H factors.*

This is the precise content of “different internal structure, same atoms,” with no commitment beyond the deduplication rule. It is forced at the level of the class: any compositional substrate with local capture admits more than one immediate content decoding to a sufficiently long region, and deduplication, keying on immediate content, admits them as distinct identities. No member of the class can avoid it, because merging them would require capture to consult a global record of every content that already decodes to a given sequence, a central controller, excluded at the Foundation.

Consequence for the geometry. The cover $R(t)$ has, over a base region Σ , sheets at each occupied depth, each sheet carrying the content multiplicity of that depth. The radial spreading rate over Σ is set by how the fiber populates depth and by the multiplicity at each, and this varies from region to region with the content combinatorics of each ξ . That variation is exactly the “spreading rate that varies across $L(t)$ ” of the local Proposition, now identified at its source: the depth-and- multiplicity profile of the content fiber, a structural invariant of the substrate. The geometry is a single hyperboloid only on one sheet; over the cover it is coarsely (pinched-negatively) curved, with the pinching determined by content multiplicity. The constant-curvature reading is the degenerate case in which every region has a single-member content fiber, which the class does not satisfy for $\text{flatlen} \geq 3$.

The fiber has no top: nesting forbidden semantically is permitted structurally. The fiber grows by one move: a composition token migrates into a boundary region and is captured together with its surroundings, producing a composition whose immediate content *references* the old one. This is the only nesting route the substrate admits: boundary pairs never nest (no-nesting is enforced on boundaries directly), but a captured token sits in $\mathcal{C}_0$ as an ordinary token and may re-enter a fresh boundary region and be captured again. Nesting is thus forbidden semantically, at the level of boundaries, and permitted structurally, at the level of reference. The new composition has a strictly larger immediate-content reference tower and the same decode, so deduplication admits it as a distinct identity: its immediate content differs from every shallower member’s, which is exactly what dedup keys on. The new layer is not merged away; it is a genuinely new composition over the same ξ .

Lemma 8.5 (Open fiber). *For fixed ξ the content fiber, ordered by reference-refinement ($m \prec m'$ iff m' 's immediate content is obtained from m 's by one migrate-and-recapture layer preserving the decode), contains a chain $0 \prec 1 \prec 2 \prec \dots$ with no maximal element. Each member has distinct immediate content, so each is a distinct identity that deduplication does not collapse, and pins the same region Σ at strictly greater depth. Nothing in the substrate bounds the layer count, because the rule that would bound it (no boundary nesting) is the same rule the reference route circumvents.*

Zorn supplies a maximal element only when every chain is bounded above. Here the chains are the reference towers in the content fiber, and they are not bounded above: the radial coordinate is open. This is the structural reason there is no curvature bound: not a gap in the argument but a property of the substrate. Both coordinates are unbounded: the transverse coordinate because D_t is monotone in the longest solution length (a longer solution adds axes; the base volume only ratchets up), and the radial coordinate because reference layering produces unbounded new immediate contents over a fixed decode, none of them deduplicated away.

Where the bound is, and where it is not. Two bounds must be kept separate, because they live in different places. At *fixed* base volume V_t the number of distinct pinned regions is finite (finitely many axes admit finitely many distinct decoded contents) but over each region the reference tower is infinite. And V_t is not fixed: it grows with the longest solution, so even the region count is unbounded over time. The region count is finite-at-fixed- V_t but V_t -monotone-unbounded; it is a structural quantity. The tower depth over each region is unbounded outright; it is the quantity that must be held down dynamically if it is to stay finite at all.

Worse, the open direction is self-feeding. A compounded composition is long; placed in an organism it raises D_t by its whole flatlen, which adds transverse room that a later capture can fill with a further reference layer. Depth enlarges V_t , and a larger V_t permits more depth. Unopposed, this loop grows hierarchy exponentially to super-exponentially. No structural fact stops it, because the geometry is open in both coordinates by construction. Chain 3 shows that the same loop runs laterally without paying the radial coordinate at all, and that the lateral route is the cheaper of the two under d_H , so it is the one that runs first.

What opposes it, and the boundary of this section's claim. The bound is therefore not geometric. It is supplied by the system's internal regulators, MCO decay starving deep reference routes of sampling weight at the source, and the Participation Signal with inlining dissolving unheld compositions and pushing their content back down, acting along a gradient the system does not itself contain. The gradient comes from outside: the population is measured against an external fitness, and *fitness is not part of the system*. It is the standard the population is held to, not a subsystem. The only channel by which that external standard reaches the internal regulators is the Participation Signal, and it reaches it by renormalization rather than by filtering: the reading is taken over the whole population, an organism that does not reproduce is absent from the next one, and the compositions only it held are absent from the next reading. The system reads participation, never a fitness value; differential survival under the external measure is what makes participation differential. The architecture supplies the mechanism of regulation; the environment supplies its direction.

This places class-level constraints on the environment, not an implementation, and Chain 6 states them in two parts: a gradient must exist at all, and it must have a down-component on decoded length, or nothing transduces into down-pressure on the tower and the self-feeding loop runs free. *How* a given GADS realises them, the form of the penalty, the length target, the penalty

rate, is a choice of the environment a particular system is run against, not a fact about the class, and is left to the instance.

What is structural, and all §8.6 claims, is the openness: the geometry cannot supply the bound, because the reference tower the no-boundary-nesting rule permits is unbounded in depth, and D_t is unbounded in length. That the internal regulators, given a length-disfavoring external gradient through the participation channel, in fact settle this open geometry into a bounded non-equilibrium rather than collapsing it or letting it blow up is a dynamical statement, not established here. The boundary is the same one drawn at the head of this section: structure here, convergence there.

Why content-blindness is correct, not a gap. The cost geometry does not resolve immediate content, and that is the right behavior. Cost is reachability: d_H measures how deep and how wide a region is to reach through resolution, which does not depend on which immediate content realises the resolution. Content acquires significance only through use: a content-distinct composition matters, against another of equal cost, exactly when it is the one reused, signalled, and survived, and use is the metabolic layer’s concern, not the geometry’s. So the multiplicity $\mu(\Sigma, \delta)$ sits in the measure as inert degeneracy, content awaiting use, resolved only when participation makes one member live and lets the rest go unheld and be inlined. This is the occupancy corollary of §8.5: which member over a cost point is occupied is dynamical, because use is the only thing that breaks a tie the static geometry is built to leave unbroken.

Given $(\star)$ as a cost-distance, the structural prior

$$P_{\text{struct}}(m) \propto \exp(-d_H(m)) \quad (18)$$

is the Gibbs measure of that distance, the maximum-entropy distribution on compositions whose expected cost displacement is fixed.

We state plainly what this does and does not add. Writing $\exp(-\text{cost})/Z$ is available for any non-negative cost and by itself asserts nothing. What gives it content is that the cost is not free: it is the geometric cost-extent derived in Section 8.5. The Boltzmann form is a re-expression of an earned cost-distance as a likelihood, so that

$$-\log P_{\text{struct}}(m) = d_H(m) + \text{const}, \quad (19)$$

and the structural complexity

$$C_{\text{struct}}(t) = \mathbb{E}_{q_{\text{emp}}}[d_H] = \mathbb{E}_{q_{\text{emp}}}[-\log P_{\text{struct}}] + \text{const} \quad (20)$$

is a proper expected negative log-likelihood, where q_{emp} is the empirical token distribution in the population. The prior is downstream of the geometry; it is not an independent modelling commitment.

The prior is over pinned regions, fibered by content. Summing $\exp(-d_H(m))$ over $L(t)$ as if each composition were an independent direction over-weights any region whose content fiber (§8.6) is multiply occupied: the members of $\Phi(\xi)$ pin one region Σ , not $|\Phi(\xi)|$ regions, so a naive sum charges the transverse subspace once per content tuple that reaches it. The measure that respects the geometry is organised by pinned region first and by content within it. Writing $\Phi(\Sigma) = \{m : \Sigma_m = \Sigma\}$ for the content fiber over Σ ,

$$P_{\text{struct}}(m) = \frac{1}{Z} \underbrace{\exp(-d_{\perp}(\Sigma_m))}_{\text{region}} \cdot \underbrace{\frac{\exp(-\rho(m))}{\sum_{m' \in \Phi(\Sigma_m)} \exp(-\rho(m'))}}_{\text{content within region}}, \quad (21)$$

where $d_{\perp}(\Sigma) = \text{flatlen} \cdot c_0$ is the region's transverse extent (charged once) and $\rho(m) = \text{flatlen} \cdot [\text{arccosh}(1 + \text{depth}(m)) - c_0]$ is the residual radial cost. Because d_H is content-blind, ρ depends on m only through its depth: content-distinct members of $\Phi(\Sigma)$ of equal depth receive equal weight, and the within-region sum collects them with a multiplicity $\mu(\Sigma, \delta)$ equal to the number of content-distinct members of depth δ . The product telescopes to $\exp(-d_H(m))/Z$ for the live member at its own depth, so $(\star)$ is recovered per composition; what changes is the normalisation, which counts each base region once and distributes weight *within* its fiber over depths, weighting each depth by its content multiplicity, rather than summing independent copies. This is inclusion–exclusion on pinned regions made explicit: the members of a content fiber pin one region, and the measure weighs that region once while carrying the content multiplicity as the within-fiber degeneracy. It weighs the union, not the multiset, and records how many content-distinct identities sit over each cost point without letting them inflate the region's weight.

Why this leaves the free energy intact. The decomposition of §8.8 requires only that P_{struct} be a normalisable measure over compositions with $-\log P_{\text{struct}} = d_H + \text{const}$ on each composition. Both hold: the region-fibered form normalises (Z is finite whenever $L(t)$ is), and on any single live composition it reduces to $\exp(-d_H(m))/Z$, so $-\log P_{\text{struct}}(m) = d_H(m) + \log Z$ as before. Hence

$$C_{\text{struct}}(t) = \mathbb{E}_{q_{\text{emp}}}[d_H] = \mathbb{E}_{q_{\text{emp}}}[-\log P_{\text{struct}}] + \text{const} \quad (22)$$

is unchanged, and it enters the variational free energy in the same role, balanced against the entropy of q_{emp} . The fiber correction moves only Z , the bookkeeping of how regions are counted, not the per-composition cost the free energy consumes. The complexity term does not charge a convergent base region once per nesting that reaches it: structurally distinct routes to the same pinned content are weighed as parses of one region, which is what the pattern-region definition required all along.

8.7 The Cost Geometry is Also the Search-Space Warp

The same object d_H has a passive reading and an active role, and the two must be kept distinct.

Passive reading. $C_{\text{struct}}(t)$ is a telemetry readout: the expected compositional cost of the population's tokens. Computing it does not influence selection. In this sense d_H is a thermometer.

Active role. The same cost geometry shapes the dynamics. The active warp is the cover's tangent structure: a single tangent-operator firing on a parse r moves all $\text{flatlen}(r)$ pinned axes *coherently*, because each token in r is one editable unit. Reaching the configuration $\xi^*(r)$ through the composition tokens of r costs one tangent firing per token; reaching the same content through a flat atomic parse costs one firing per atom. The cover concentrates single-step accessibility along the directions its fiber elements pin, and where occupancy sits in the cover determines which directions the population can move along in one step.

Chain 2's singularity, geometrically. Without open, occupancy ascends the cover and never descends. Captures continue to fire, raising parse depth at occupied positions; no operator brings occupancy back down. The population collapses onto a single coarse sheet at high depth: every organism's parse is a deeply nested composition with no exposed constituents to mutate independently. The tangent neighborhood $N(r)$ shrinks to the moves available at that one rigid parse. Lateral fracture-moves, the independent variation of constituents that made the cover's curvature negative, vanish. The active warp becomes a single point. This is the geometric content of the

Chain 2 failure mode: not bookkeeping about composition counts, but collapse of cover occupancy onto a coarse sheet whose tangent structure carries no remaining branching.

Passive and active are one object. The “passive telemetry” description scopes to the computation of C_{struct} , not to the geometric object it reads. d_H is one object: the passive readout C_{struct} and the active warp on the cover’s tangent structure are two faces of it. The decay of a composition’s influence (the meta-gene creation order, Chain 5) is then the dilution of fiber-element occupancy as later captures route the population through their own sheets: the same construction, viewed over capture-time. The MCO decay is the temporal face of the cost geometry, not a separate forced thing.

Decay form: forced existence, contingent shape. That a composition’s share of accessibility decays is forced: it is monotone decreasing the moment later captures concentrate accessibility into their own pattern regions. The exponential form of the decay is contingent. Exponential decay is the regime where captures arrive at a rate proportional to existing structure, so that accessibility compounds multiplicatively. If captures arrive at a constant rate and accessibility adds linearly, the share decays polynomially. The existence of decay is structural; the exponential form is the equilibrium regime, and slower-exhaust regimes exhibit heavier tails. Which regime obtains is an empirical question, not settled here.

8.8 Variational Decomposition

The variational free energy of the population at time t decomposes into three terms, each forced by a specific architectural role.

Theorem 8.1 (Three-Term Variational Decomposition). *The variational free energy of the population at time t admits the decomposition*

$$F_{\text{VFE}}(t) = \underbrace{D_{KL}(q_{\text{emp}}||q_{\text{sel}})}_{\text{selection complexity}} + \underbrace{\mathbb{E}_{q_{\text{emp}}}[d_H]}_{\text{structural complexity}} + \underbrace{\mathbb{E}_{q_{\text{emp}}}[-\log P_{\text{gen}}]}_{\text{surprisal}}, \quad (23)$$

where each term is forced by a distinct architectural role.

- **Selection complexity** is the divergence of the empirical token distribution from the selection prior q_{sel} defined by the MCO sampling distribution (Chain 5). It measures how far fitness pressure has warped the population from the architecture’s intrinsic selection balance.
- **Structural complexity** is the expected compositional cost under the empirical distribution, the expected negative log-likelihood under the Boltzmann structural prior of Section 8.6. It is the term Section 8.5 earns geometrically.
- **Surprisal** is the expected negative log-likelihood of the population under the generative transition model P_{gen} . It measures the predictability of the population’s transitions.

The three terms have the structure of a variational free energy bound: an empirical (variational) distribution, a divergence from a prior, and a likelihood under a generative model. The formal statement of the distributions, their estimation, and the implementation correspondence are developed in Section 9; this section establishes that the structural-complexity term is geometrically earned.

What this decomposition does not claim. The decomposition is structural: each term is forced by an architectural role, and the structural-complexity term’s prior is earned in Section 8.5. The decomposition does not claim the VFE converges, settles into a steady state, or is minimized by the architecture. Those are dynamical questions, path-dependent and statistical, not treated here. What is established here is the form of the decomposition and the geometric grounding of its structural term, not the behavior of the quantity over time.

9 Variational Free Energy Telemetry

Section 8 established the three-term variational decomposition (Theorem 8.1) and earned the geometric grounding of its structural term. This section formalizes the decomposition as a measurement: it defines the distributions each term is computed from, establishes the properties of the resulting quantity, and shows that the measurement is computed independently of the fitness function. The independence is the load-bearing result: it is what distinguishes a GADS, which *measures* a variational free energy as a passive consequence of its dynamics, from a system engineered to minimize one.

The variational free energy is a GADS-class property, not a feature of any single instance. Any GADS computes the three terms from its population, its active subsystem state, and its generative model. The canonical instance MEGA is used throughout for implementation correspondence, but the construction is class-general: the terms are defined from the architectural roles the chains force, and the fitness independence is forced by Chain 6’s placement of fitness as environmental rather than architectural.

9.1 The Measured Quantity

Definition 9.1 (Variational Free Energy). *For a GADS at generation t with population P_t , the variational free energy in natural units is*

$$F_{\text{VFE}}(t) = D_{KL}(q_{\text{emp}}||q_{\text{sel}}) + \mathbb{E}_{q_{\text{emp}}}[d_H] + \mathbb{E}_{q_{\text{emp}}}[-\log P_{\text{gen}}], \quad (24)$$

the three terms of Theorem 8.1: selection complexity, structural complexity, and surprisal.

The remainder of this subsection defines the three distributions q_{emp} , q_{sel} , and P_{gen} . Each is computed from internal state, population content, subsystem configuration, generative dynamics, and none references the fitness function.

Empirical distribution. The empirical token distribution is the frequency of each token in the decoded, delimiter-stripped population:

$$q_{\text{emp}}(g) = \frac{\sum_{x \in P_t} \text{Count}(g \in \text{Strip}(x))}{\sum_{x \in P_t} |\text{Strip}(x)|}, \quad (25)$$

where Strip removes boundary delimiters, leaving the semantic content. The support of q_{emp} is the active vocabulary $\Sigma = \text{supp}(q_{\text{emp}})$, the tokens actually appearing in the population.

Selection distribution. The selection distribution is the architecture’s intrinsic sampling prior, set by the MCO ordering (Chain 5). For base genes and compositions, following the MCO selection

weights:

$$q_{\text{sel}}(g) \propto \begin{cases} \frac{\pi_{12}}{|\Sigma \cap G|} & g \in \Sigma \cap G \quad (\text{base gene}) \\ \frac{(1 - \pi_{12}) \cdot \pi_{13}^{|L|-j-1}}{Z_\Sigma} & g = m_j \in \Sigma \cap L \quad (\text{composition}) \end{cases} \quad (26)$$

where π_{12} is the base-gene probability mass, π_{13} is the MCO decay factor, j is the composition's position in the insertion order, and Z_Σ normalizes over the compositions present in Σ . The distribution is then renormalized over Σ . This is the distribution the population would exhibit under the architecture's own selection operators with no external pressure, the neutral equilibrium of Chain 5's sampling.

Generative model. The generative model is the bigram transition distribution estimated from the previous generation:

$$P_{\text{gen}}(b \mid a) = \frac{\text{Count}_t(a, b) + \alpha_s}{\text{Count}_t(a, \cdot) + \alpha_s |\Sigma_t|}, \quad (27)$$

with Laplace smoothing parameter α_s . The surprisal term is the expected negative log-likelihood of the current population's transitions under this model, the predictability of the population's structure given the patterns the previous generation established. The estimation base is a telemetry choice and is independent of the Participation Signal's scoping; it may be taken over the previous generation's parents or over its whole population, so long as the choice is stated with the measurement.

9.2 Normalized and Natural-Units Forms

The natural-units form $F_{\text{VFE}}(t)$ is unbounded above and directly comparable across experiments and parameter settings. A normalized form maps each term to $[0, 1]$ for bounded comparison: surprisal is normalized by the maximum entropy $\log |\Sigma_t|$; complexity is mapped through $1 - \exp(-\text{raw}/\text{scale})$ with $\text{scale} = \log |\Sigma| + 1$. The normalized form trades cross-experiment comparability for a bounded range; the natural-units form is the primary measured quantity.

The normalized form is a presentation convenience and is *not* the quantity the propositions below describe. The non-negativity and equilibrium characterizations (§9.3) are stated for the natural-units three-term sum, where each term retains its own scale and the terms add. Squashing the terms into $[0, 1]$ and collapsing them into a single weighted score destroys the additive three-term structure of Theorem 8.1: the result is bounded by construction rather than by any property of the dynamics, and its boundedness carries no information. A bounded, blended score may be logged for monitoring, but it must not be substituted for F_{VFE} in any statement about the architecture, and in particular the fitness-independence and non-monotonicity results refer to the natural-units quantity alone.

9.3 Properties

Proposition 9.1 (Non-negativity). *For all t , $F_{\text{VFE}}(t) \geq 0$.*

Each term is non-negative: the KL divergence by Gibbs' inequality; the structural term because $d_H(g) \geq 0$ for all tokens (base genes have $d_H = 0$, compositions have $d_H > 0$); the surprisal because it is a negative log-likelihood of probabilities in $(0, 1]$.

Proposition 9.2 (Equilibrium characterization). $F_{\text{VFE}}(t) \rightarrow 0$ requires all three terms to vanish simultaneously: $q_{\text{emp}} = q_{\text{sel}}$ (the population matches the architecture’s intrinsic selection balance), $\mathbb{E}_{q_{\text{emp}}}[d_H] = 0$ (the population uses only base genes), and the surprisal vanishes (the population is perfectly predicted by the previous generation’s model).

True equilibrium is not the expected operating regime. Fitness pressure typically drives q_{emp} away from q_{sel} ; useful compression via compositions makes $\mathbb{E}_{q_{\text{emp}}}[d_H] > 0$; variation introduces surprisal. A GADS under selection occupies a regime that balances these pressures, not one that minimizes their sum.

Non-monotonicity. F_{VFE} is not a quantity the architecture drives in one direction. Exploration raises it (surprisal increases, the population deviates from equilibrium); exploitation lowers it (convergence, return toward equilibrium). The architecture does not minimize VFE; the quantity rises and falls as the population explores and consolidates. This is a direct consequence of the next property: VFE is not an objective the architecture optimizes, because the architecture cannot read it.

9.4 Fitness Independence

Theorem 9.1 (Fitness Independence). *The variational free energy is computed without reference to the fitness function F :*

$$F_{\text{VFE}}(t) = F_{\text{VFE}}(P_t, \text{ subsystem state}, P_{t-1}, \Pi), \quad (28)$$

where the fitness function F does not appear. Fitness influences F_{VFE} only indirectly, through which organisms survive to populate P_t , never through direct computation.

The proof is by construction. Each term depends only on internal state: selection complexity on P_t and the MCO parameters; the structural term on P_t and the compositional geometry of the store; surprisal on P_t and the previous generation’s generative model. No term references fitness values. This is not a property MEGA happens to have; it is forced by the architecture. Chain 6 places fitness as environmental: the architecture requires the capacity to receive feedback but is not constituted by any particular fitness function. Fitness acts on the population from outside the architecture; the VFE is computed from the architecture’s internal state. A quantity computed from internal state structurally cannot reference the external fitness function.

Why this is load-bearing. A standard objection to free-energy formulations of adaptive systems is circularity: if a system is engineered to minimize a quantity labelled “free energy,” then its settling into a low-free-energy state is a tautology, not a finding. The objection applies to any implementation that builds the free energy into the objective the system optimizes.

It does not apply to a GADS, and Chain 6 says why in its own terms. The exhaust the architecture runs needs a gradient it does not author, so the gradient is placed outside; a quantity computed from internal state is therefore structurally incapable of being the gradient the system runs down. The variational free energy is a passive measurement on internal state, computed independently of the fitness that drives selection. The architecture is not engineered to minimize it, and cannot read it. That a GADS’s measured VFE settles into a sustained regime rather than diverging is therefore an *observation* about the dynamics, not a designed-in result. The fitness independence is what makes the observation non-circular: the quantity that settles is not the quantity the system was built to reduce.

This is the measurement’s significance for the GADS class. The decomposition (Theorem 8.1) is forced by architectural roles; the fitness independence (this theorem) is forced by Chain 6’s placement of fitness as external. Together they make the variational free energy a property the architecture *exhibits* rather than *pursues*. Whether the measured quantity settles into a non-equilibrium steady state, and under what conditions, is a dynamical question, path-dependent, statistical, and not established here. What is established here is that if such settling is observed, it is observed in a quantity the architecture does not optimize, which is what the FEP circularity objection requires and what standard implementations cannot claim.

9.5 Implementation Correspondence

In the canonical instance MEGA, the three terms correspond directly to computed telemetry. The empirical distribution is read from the decoded, delimiter-stripped population. The selection distribution is constructed from the instance’s configured base-gene mass and decay factor, which correspond to π_{12} and π_{13} above, together with the insertion order of the live set. The structural cost is computed by expansion of each composition’s depth and flat length (Section 8.5). The surprisal is computed from the bigram model of the previous generation. The telemetry logs each term per generation alongside fitness statistics for correlation analysis, correlation that is observed after the fact and never enters the VFE computation itself.

Two correspondence points are load-bearing. The structural term is computed by expanding compositions the population holds, not by consulting a store-side table, which is why it is well-defined for a token whose composition the store has since removed. And the selection distribution is built from the live set’s insertion order, which is MCO, since the store and the order are one structure.

MEGA is one instance; the construction applies to any GADS that instantiates the substrate, discharges the roles, and receives fitness feedback through POP’s signal channels.

Closure

Foundation: one commitment forcing locality, locality requiring a structured object that GADS adopts as the substrate, substrate accommodating the L0 setup.

Six chains, each forcing what is required to address the next asymptotic structural failure mode internal to the architecture.

L0 + chains = full GADS architecture. The forced things are locality, capture, the value encoding, MG, open, the Participation Signal with its removal rule, inlining, MCO, and POP’s signal channels. Removing any one of them produces a specific identified failure mode.

No chain rests on a resource bound. Every failure mode the derivation answers is internal and asymptotic: compression that cannot be reused (Chain 1), collapse of variation and of cover occupancy (Chain 2), the granularity ratchet in both coordinates (Chain 3), phenotypic invariance under removal (Chain 4), decoupling of the sampling distribution from the population’s location (Chain 5), and loss of entropic pressure (Chain 6). No chain forces a capacity parameter, and the architecture carries no constant that sets a store size, a lifetime, or a threshold.

Necessity, not sufficiency. The chains establish *necessity*: each forced element is required, and removing any one produces the failure mode it was forced by. They do not establish *sufficiency*, that the assembled architecture, with all elements present, equilibrates rather than collapses. That

the system settles into sustained operation rather than drifting into the Chain 2 singularity or running the ratchet of Chain 3 is a statistical tendency, underwritten by the entropic exhaust of the metabolic loop, not a theorem proved here.

Two quantities must be kept separate in any such claim. Cardinality is bounded by the removal rule without a parameter (Chain 3). Extent is bounded by nothing internal, and is held down only by an external measure disfavoring decoded length, transduced through the Participation Signal (Chain 6). So the sufficiency question is conditional on the environment in a way the necessity argument is not.

The exhaust, and what it carries. The architecture is an entropy engine and the metabolic loop is its exhaust. Entropy enters through capture, which admits new structure, and leaves through removal, which discards a distinction the population did not support. What leaves is not content: inlining conserves each parent’s decoded content, so information flows up while structure flows down, and content leaves the store only at the leaves of the reference relation. What leaves is the claim that a span was worth naming as a unit.

Chain 3 supplies the cycle and Chain 6 supplies the gradient across it, and both are required. A cycle without a gradient does work and transports nothing: admission and removal balance at a rate equilibrium while the store’s composition walks undirected over the reachable structures. That is the Chain 6 failure mode, and it is the sense in which differential survival is class-required. The gradient cannot be internal, because a system does not supply its own difference, which is the same fact that makes the variational free energy of §9 a quantity the architecture exhibits rather than pursues.

The interface. The architecture is configured to receive environmental feedback through POP’s signal channels, via fitness applied to the population. Fitness evaluates the decoded atomic content of each member; decoding expands composition tokens from their own content. The environment is not part of the architecture; the framework requires only the capacity to receive feedback, plus the two class-level constraints Chain 6 states on what that feedback must be: that a gradient exists, and that it has a down-component on decoded length. What the architecture produces depends on what environment it is placed in and what feedback that environment provides.