GENERALIZED RESIDUAL CLOSURE: GENERAL LEARNING DYNAMICS FOR STABILITY–PLASTICITY COMPATIBILITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Learning algorithms typically choose their adaptation coordinates before optimizing them. Generalized Residual Closure (GRC) asks what a future-relevant residual should change. We formulate learning as recursive least-lawful closure: a representation opens into relations that constrain future continuation, and disagreement with realized continuation diagnoses insufficiency without selecting its own write address. At a fixed learner–world boundary, a complete representation–relation description yields two primitive internal modes: Transformation within a representation and revision of the Representation itself. We give a local tangent decomposition and a fixed-representation obstruction criterion. In an affine sector, a necessary-and-sufficient condition and a quadratic construction characterize nonzero learning that preserves registered old responsibilities. Protecting reconstructible semantics weakly enlarges the safe-plasticity operator relative to preserving an exact historical realization. Dynamic sufficiency and future-closure viability extend the analysis beyond current prediction. Growth Learning requires non-regression of the closure-capability frontier, with strict expansion of its domain or reduction of an optimal closure cost. A conditional commit rule preserves this order without guaranteeing discovery of improvement. Optimization, differentiable machine families, and fixed-state control admit restricted correspondences. Under explicit operational and minimal-selection assumptions, the resulting learner class admits a conditional GRC representation, without implying implementation uniqueness or universal performance dominance.

1 INTRODUCTION: WHAT SHOULD A RESIDUAL CHANGE?

Machine learning usually begins after the coordinates of adaptation have been chosen. Neural networks update parameters; reinforcement learning updates values, policies, or models on a declared state space; adapters restrict writes to prescribed subspaces; architecture search explores a structural grammar. Optimization determines how these selected objects should change. Their selection raises a prior question: *what should a residual change?*

A residual does not carry its own write address. A prediction error may indicate a wrong relation within an adequate representation, aliasing of future-distinguishable histories, unavailable evidence, or a family that cannot express the required continuation. Treating every discrepancy as a parameter-update instruction conflates evidence of insufficiency with a decision about where it resides.

Generalized Residual Closure (GRC) describes learning through this distinction. A represented state opens into relations that explain the present and predict continuation. Interaction exposes a contract-relative discrepancy, and a lawful persistent revision establishes the conditions for subsequent predictions, actions, and learning. Closure is recursive: the successor may expose a new residual, or remain sufficient under later interaction. At a fixed boundary, Transformation changes relations within the current abstraction, while Representation Learning changes which states must be distinguished. A joint revision may contain both.

Stability and plasticity then share a constraint geometry. Stability protects reconstructible future-operational responsibilities; plasticity is the useful response that remains lawful under those constraints. Present success remains insufficient for continual learning. Dynamic sufficiency preserves

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

information needed for future updating, and future-closure viability retains lawful continuation paths. Growth Learning additionally compares the domain and optimal cost of those paths.

The paper makes four contributions. First, it specifies the recursive semantics and least-lawful closure formulation. Second, it gives a local two-mode decomposition, a conditional finite classification, and a representation-obstruction criterion. Third, it derives affine stability–plasticity compatibility and reconstructive semantic dominance, then distinguishes dynamic sufficiency from future viability. Fourth, it defines certified growth, develops restricted correspondences with established methods, and supplies a conditional GRC representation theorem. The results organize these requirements jointly; their assumptions do not establish universal algorithmic superiority.

Related work. Prior work has separately defined state through future predictions or behavioral equivalence (Littman et al., 2001; Ferns et al., 2011), constrained continual updates to protect previous capabilities (Lopez-Paz & Ranzato, 2017; Farajtabar et al., 2020), studied the loss and preservation of plasticity in nonstationary learning (Dohare et al., 2024; Han et al., 2026; Qiu et al., 2026), and enlarged the object of learning through meta-learning and architecture search (Finn et al., 2017; Liu et al., 2019). GRC does not claim these components individually as new; its question is upstream: which representational level owns a residual, what must remain invariant during closure, and what future learning must remain possible afterward. Appendix K details the comparisons and scope.

2 Recursive Generalized Residual Closure

Let $M \in \mathcal{M}$ contain all persistent learner variables relevant to future responses and updates. Let $x \in \mathcal{X}$ be a detailed interaction state or history. The contract Ξ registers admissible contexts, actions, evidence, information rights, tolerances, and resources; ∂ fixes the learner–world boundary. GRC imposes three structural requirements: $U^\star$ assigns an operational owner to every required distinction; $I^\star$ identifies states through registered future behavior; and $L^\star$ seeks a least-cost lawful completion. Their satisfaction requires the relevant existence and sufficiency conditions.

2.1 Future-Operational Representation

For a fixed learner M, define

$$x \sim_{\Xi,M} y \iff \mathrm{Resp}_M(c,x) \equiv_\Xi \mathrm{Resp}_M(c,y) \quad \forall c \in \mathrm{Ctx}_\Xi. \tag{1}$$

Here $\equiv_\Xi$ is a response equivalence; quantitative approximation is treated separately. The *Kernel* is the quotient $K_M = \mathcal{X}/\!\sim_{\Xi,M}$, with map $C_M : x \mapsto [x]$ and current state $k = C_M x$. It identifies detailed states under a fixed learner, not different learner implementations.

Defining state through future behavior has close precedents in predictive-state representations and bisimulation-based state abstraction (Littman et al., 2001; Ferns et al., 2011). GRC makes compatibility with registered learner-update events an explicit admissibility requirement, formalized as dynamic sufficiency in Section 5.

The *Ring* $\mathcal{R}_\Xi(K_M)$ comprises admissible relational realizations on that quotient, with active relations $\rho_M \in \mathcal{R}_\Xi(K_M)$. The quotient alone does not uniquely select these relations.

A representation is unchanged when $C_N = \phi \circ C_M$ for an admissible isomorphism ϕ. Write $\varpi_\Xi(M) = [C_M]$ for this abstraction class. Preserving it permits changes to learned relations. Stability separately protects the still-valid responsibilities registered by Ξ_{old}.

2.2 Explanation, Prediction, and Residual Formation

Opening exposes active relations; for an admissible event e, they predict $\widehat{k}^+ = \widehat{U}_{M,e}(C_M x)$. Actual interaction produces $U_e(x)$. Where both continuations are defined on the current quotient, their discrepancy is

$$\mathfrak{R}_{\Xi,e}(M,x) = \mathrm{Def}_\Xi\big(C_M U_e(x), \widehat{U}_{M,e} C_M(x)\big). \tag{2}$$

The prediction $\widehat{U}_{M,e}$ is distinct from an exactly induced update. If the quotient cannot support an adequate continuation, that obstruction must itself be represented rather than evaluated through an undefined map. The residual diagnoses insufficiency without selecting its remedy.

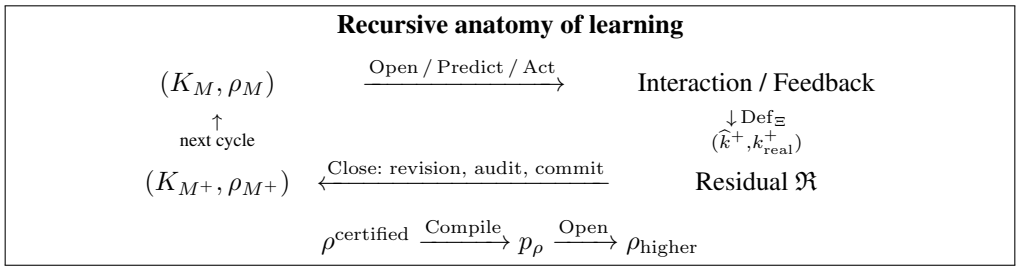


Figure 1: Closure establishes a successor learner for continued interaction. A Transformation retains the quotient while revising relations. Compilation separately turns validated reusable relations into primitives. The next interaction may expose another residual or remain sufficient.

2.3 Least-Lawful Closure

Let $\mathrm{Adm}_\Xi(M, r)$ enforce registered hard obligations and let C_{full} charge the declared search, revision, verification, deployment, and maintenance costs. The completion specification is

$$M^+ \in \operatorname*{arg\,min}_{N \in \mathrm{Adm}_\Xi(M,r)} C_{\mathrm{full}}(N \mid M) \quad \text{subject to} \quad d_\Xi\big(\mathfrak{R}_\Xi(N)\big) \le \varepsilon_\Xi. \tag{3}$$

The notation presupposes an attained lawful minimum; it does not construct a solver. Temporary reasoning or retrieval becomes persistent learning when its consequences change future-relevant learner state. The successor may retain its quotient and change relations, or revise both. Prediction proposes an unrealized continuation; Compile installs a validated reusable primitive. Closure can support either further learning or sufficient ongoing operation with zero new residual.

2.4 Same-Representation Closure Load

Let $\mathcal{C}_T^\Xi(M)$ contain the residual corrections achievable by a single admissible edit within the current representation and unit declared cost. For a nested dilation family D_λ, define

$$q_T(r \mid M) = \inf\{\lambda > 0 : r \in D_\lambda \mathcal{C}_T^\Xi(M)\}. \tag{4}$$

All protected duties constrain that same edit. Under the calibrated threshold-to-feasibility conditions of Appendix E, $q_T > 1$ identifies a limitation of the declared closure family. It does not by itself establish representation insufficiency: missing evidence, transport, and incomplete search or solving must be distinguished. Boundary feasibility depends on attainment of the infimum. The proposed closure-grade extension and its remaining conditions are kept in Appendix E.6.

3 The Two Primitive Modes of Learning

A residual identifies a mismatch, but does not determine which internal object should change. Fix the contract Ξ and learner–world boundary ∂, and write $\varpi_\Xi(M) = [C_M]$ for the representation class defined in Section 2. It records the distinctions made between detailed states, up to admissible relabeling. The current state $k = C_M(x)$, the representation C_M, and the protected old responsibilities $\mathcal{S}_{\mathrm{old}}(M)$ are different objects: a state can change without changing its representation, and relations can change while the representation remains fixed.

Theorem 3.1 (T1: Two-mode learning factorization)**.** *On a finite-dimensional smooth stratum with fixed Ξ and ∂, suppose ϖ_Ξ is smooth and has locally constant rank. Set $V_M = \ker D\varpi_\Xi(M)$. Then*

$$0 \longrightarrow V_M \longrightarrow T_M \mathcal{M}_\Xi^{(\partial)} \xrightarrow{D\varpi_\Xi(M)} \operatorname{im} D\varpi_\Xi(M) \longrightarrow 0 \tag{5}$$

is exact. After choosing a local splitting with complementary space H_M, each direction has a unique decomposition

$$d = d_T + d_R, \qquad d_T \in V_M, \quad d_R \in H_M. \tag{6}$$

The components depend on the splitting; whether $D\varpi_\Xi(M)[d] = 0$ does not.

Proof sketch and finite transitions. The derivative's kernel gives fixed-representation tangent motion. A chosen right splitting supplies d_R, and $d_T = d - d_R$ lies in that kernel. Joint learning can have both components. The full proof is in Appendix C. Finite transitions are classified by whether $C_{M^+} \cong C_M$; their exhaustion of relevant persistent changes uses representation–relation completeness, Assumption A3 in Appendix A. This conditional classification does not give a global smooth splitting across singular strata. A realization change that preserves the abstraction is not automatically Representation learning.

When representation revision is necessary. Let $\mathfrak{R}(C)$ be the exhaustive class of admissible relation laws on a fixed representation, and let F be the required future response. Define

$$\Omega_C(F) = \inf_{f \in \mathfrak{R}(C)} D_\Xi(F, f \circ C). \tag{7}$$

Proposition 3.2 (P2: Representation necessity)**.** *With the boundary, evidence, relation class, and tolerance fixed, if $\Omega_C(F) > \varepsilon_\Xi$, no relation-only update satisfies the contract. Any successful internal closure must leave that representation class.*

The proof and a Hilbert-space decomposition are in Appendix C.4. The obstruction concerns the whole admissible class; optimizer failure alone cannot establish it. It does not guarantee a feasible revised representation, and an infimum at the tolerance boundary is not itself an attained solution.

Querying, retrieval, sensing, and tool use can supply evidence without persistent internal learning. Incorporating their results changes relations, representation, or both. If the interface machinery itself is learned, the enlarged persistent state and boundary must be declared; classification still depends on whether C_M actually changes. Interface modification is not automatically a third mode or a Representation transition.

4 Stability and Plasticity from One Geometry

GRC treats stability as a constraint on lawful change and plasticity as the residual response remaining within that constraint. The protected object is the registered old responsibility, rather than every future response; learning may introduce new behavior while preserving what remains required. The following affine sector makes this compatibility condition explicit.

Semantic-safe response geometry. Let $d \in \mathbb{R}^n$ be an edit and $\mathcal{S}_{\mathrm{old}}(M) \in \mathbb{R}^p$ record the old responsibilities declared by Ξ_{old}. The matrix $A \in \mathbb{R}^{p \times n}$ encodes their change, and $B \in \mathbb{R}^{m \times n}$ maps edits to reductions of a current residual $r \in \mathbb{R}^m$. For this section's exact model,

$$\mathcal{S}_{\mathrm{old}}(M + d) - \mathcal{S}_{\mathrm{old}}(M) = Ad, \qquad r(M + d) = r - Bd. \tag{8}$$

Choose $H \succ 0$ with edit energy $d^\top H d$. The safe space is $\ker A$. Its metric-dependent safe inverse and residual-response operator are

$$P_H = H^{-1} - H^{-1}A^\top (AH^{-1}A^\top)^\dagger AH^{-1}, \tag{9}$$

$$K_{\mathrm{plast}} = BP_H B^\top \succeq 0. \tag{10}$$

Here $P_H = P_H^\top \succeq 0$, $AP_H = 0$, and $\operatorname{im} P_H = \ker A$. It is a safe inverse, not generally an idempotent Euclidean projector.

Theorem 4.1 (T2: Stability–plasticity compatibility)**.** *Under the affine model* (8) *with $H \succ 0$, the following conditions are equivalent:*

$$\exists d \in \ker A : \ \langle r, Bd \rangle > 0, \tag{11}$$

$$P_H B^\top r \neq 0, \tag{12}$$

$$r^\top K_{\mathrm{plast}} r > 0. \tag{13}$$

Thus old-responsibility preservation and a strict first-order reduction of squared residual error are compatible precisely when the residual has an effective safe response.

Proof sketch. If $P_H B^\top r \neq 0$, choose $d = P_H B^\top r$. Then $Ad = 0$ and $\langle r, Bd \rangle = (B^\top r)^\top P_H (B^\top r) > 0$. Conversely, $P_H B^\top r = 0$ places $B^\top r$ in $\operatorname{im} A^\top$, which annihilates every safe direction. The full proof is in Appendix D. The result characterizes Transformation directions only when the edit coordinates fix the representation; otherwise it applies to the joint edit space.

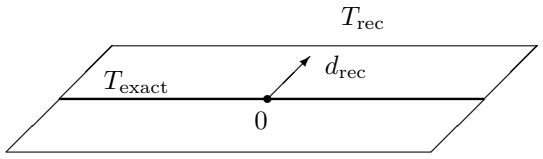


Figure 2: Stability is a geometry of lawful change. A line and a patch of a plane illustrate nested safe subspaces; the drawing is conceptual, not to scale. Reconstructive protection permits additional directions when historical realization constraints are unnecessary. Strict response gain still requires those directions to affect the residual through B.

A constructive affine step. For $\eta > 0$, consider the strictly convex subproblem

$$\min_{Ad=0} \frac{1}{2}\|r - Bd\|^2 + \frac{1}{2\eta} d^\top H d. \tag{14}$$

Its unique solution and resulting residual are

$$d^\star = \eta P_H B^\top (I + \eta K_{\mathrm{plast}})^{-1} r, \tag{15}$$

$$r^+ = (I + \eta K_{\mathrm{plast}})^{-1} r. \tag{16}$$

The step satisfies $Ad^\star = 0$ and $\|r^+\| \le \|r\|$, with strict inequality exactly when $K_{\mathrm{plast}} r \neq 0$. The derivation appears in Appendix D.3. This affine subproblem need not eliminate the residual and is not the complete Close operation.

Reconstructive protection. Preserving old operational responsibilities can allow more edits than freezing their historical implementation. Let $T_{\mathrm{exact}} \subseteq T_{\mathrm{rec}}$ be the corresponding safe spaces. With the same B and H, define $P_T^H = H^{-1/2}\Pi_{H^{1/2}T}H^{-1/2}$ and $K_T = BP_T^H B^\top$.

Corollary 4.2 (Reconstructive semantic stability dominance)**.** *If $T_{\mathrm{exact}} \subseteq T_{\mathrm{rec}}$, then $P^H_{T_{\mathrm{rec}}} - P^H_{T_{\mathrm{exact}}} \succeq 0$ and*

$$K_{\mathrm{rec}} - K_{\mathrm{exact}} = B(P^H_{T_{\mathrm{rec}}} - P^H_{T_{\mathrm{exact}}})B^\top \succeq 0. \tag{17}$$

Whitening by $H^{1/2}$ reduces the claim to nested orthogonal projectors; Appendix D.8 gives the full proof. Strict gain for the current residual requires $r^\top(K_{\mathrm{rec}} - K_{\mathrm{exact}})r > 0$. Operator dominance alone gives neither universal strict benefit nor a uniformly smaller post-update Euclidean residual.

This geometry is related to continual-learning methods that restrict updates to protect previous behavior, including gradient-projection approaches (Farajtabar et al., 2020) and recent explicit treatments of the stability–plasticity tradeoff (Han et al., 2026; Qiu et al., 2026). The distinction here is that the protected object is future-operational semantics rather than a particular historical parameter or gradient subspace.

Nonlinear scope. For a smooth nonlinear learner, $A = D\mathcal{S}_{\mathrm{old}}(M)$ and B is the local residual-response Jacobian. Tangent safety alone does not preserve a finite nonlinear endpoint; actual old responsibilities must be checked before commit. Nor does a missing local response exclude a nonlocal path within the same representation. Global Representation necessity requires the separate obstruction in Proposition 3.2.

5 Dynamic Sufficiency, Continual Learning, and Growth

Preserving the past and learning the present need not preserve the ability to learn again. A compression may support the current prediction yet discard information needed to determine the correct update after future evidence. For a candidate representation $C_t : \mathcal{X}_t \to Z_t$ and registered detailed update $U_e : \mathcal{X}_t \to \mathcal{X}_{t+1}$, update sufficiency requires

$$C_{t+1}U_e \simeq_\Xi \bar{U}_e C_t. \tag{18}$$

Here $\simeq_\Xi$ is exact operational equivalence; approximation requires a separate error contract. The induced updater $\bar{U}_e$ is distinct from the prediction $\widehat{U}_e$. Equal compressed inputs cannot determine different required post-update states, even if their current predictions agree. Appendix F.1 gives the factorization criterion and a two-state counterexample. Dynamic sufficiency additionally requires the corresponding commutation conditions for registered background changes and internal transitions.

5.1 Future-Closure Viability

Let $\Gamma_\Xi(M, r)$ contain the finite lawful paths that close a registered future demand $r \in \mathcal{Y}_\Xi^{\mathrm{fut}}$. Define

$$\mathcal{V}_\Xi(M) = \{r \in \mathcal{Y}_\Xi^{\mathrm{fut}} : \Gamma_\Xi(M, r) \neq \varnothing\}. \tag{19}$$

Viability on a declared class $\mathcal{Y}_0$ means $\mathcal{Y}_0 \subseteq \mathcal{V}_\Xi(M)$. A certificate such as $B_f P_H B_f^\top \succeq \kappa I$, $\kappa > 0$, concerns only the registered local probe family. It neither presumes knowledge of unknown future tasks nor replaces multi-step path viability. Appendix F.4 distinguishes these objects and shows how acquisition can exhaust a particular future response.

5.2 Closure Capability

For finite, nonnegative path costs under a common accounting convention, set

$$\begin{aligned} c_\Xi^\star(r \mid M) &= \inf_{\gamma \in \Gamma_\Xi(M,r)} C_{\mathrm{full}}(\gamma), \qquad \inf \varnothing = +\infty, \\ \mathscr{F}_\Xi(M) &= \{(r, c) \in \mathcal{Y}_\Xi^{\mathrm{fut}} \times \mathbb{R}_{\geq 0} : c_\Xi^\star(r \mid M) \leq c\}. \end{aligned} \tag{20}$$

This frontier is an optimal-cost epigraph; its infimum need not be attained. Write $M' \succeq_\Xi M$ when old registered semantics remain reconstructible and $\mathscr{F}_\Xi(M') \supseteq \mathscr{F}_\Xi(M)$. Strict growth $M' \succ_\Xi M$ additionally expands the viable domain or strictly lowers the optimal cost of a previously closable demand. Thus growth means a broader or cheaper lawful future, with no regression elsewhere in the registered comparison. Information rights, responsibilities, and cost accounting must be aligned.

5.3 Certified Persistent Growth

Let $\widetilde{M}_t = \mathsf{T}_{t \to t+1} M_t$ be a lawful no-learning transport baseline in the next background. A trial may regress; a persistent commit requires both a sound audit and certified dominance over that baseline.

Theorem 5.1 (T3: Certified Monotone Causal Growth)**.** *Assume the transport baseline is lawful and the audit certifies current closure, semantic preservation, dynamic sufficiency, declared future viability, and resource and safety obligations. Commit a candidate only if the audit passes and the candidate dominates $\widetilde{M}_t$; otherwise retain $\widetilde{M}_t$. Then*

$$M_{t+1} \succeq_{\Xi_{t+1}} \widetilde{M}_t.$$

If an accepted candidate also strictly expands the closure domain or lowers an optimal closure cost, then $M_{t+1} \succ_{\Xi_{t+1}} \widetilde{M}_t$.

The two branches of the rule give the result by dominance or reflexivity; Appendix F.9 supplies the full statement and proof. The theorem does not construct an audit, a dominance certificate, or an improving candidate. Rejection retains the lawful baseline without claiming current closure or refunding trial costs. The restricted discovery result in Appendix F needs additional evidence, reachability, and fair terminating search; it does not imply unlimited growth.

Carry can enlarge the domain of lawful future closure; Compile can turn recurrent certified relations into reusable primitives with lower future cost. These are two possible growth effects, not automatic consequences of either operation. Frontier preservation and an actual expansion or optimal-cost improvement remain required. Appendix F.11 states the compilation, reopening, and full-cost conditions. Successful closure can thereby become reusable capacity for subsequent learning.

6 EXISTING PARADIGMS AS RESTRICTED GRC SECTORS

Established methods fix different parts of the closure problem. GRC can retain their solvers while reconsidering those choices. An exact sector embedding must preserve the declared inputs, outputs, information rights, randomness, and updates; a similar functional form alone is insufficient.

6.1 OPTIMIZATION

Fix the representation, boundary, program, and writable coordinates. For $g = \nabla L(\theta)$, $\eta > 0$, and $H \succ 0$,

$$\arg\min_d \left\{ g^\top d + \frac{1}{2\eta} d^\top H d \right\} = -\eta H^{-1} g. \tag{21}$$

Thus $H = I$ yields ordinary gradient descent, while positive-definite Fisher geometry recovers a natural-gradient sector (Amari, 1998); suitable positive-definite local curvature choices produce Newton-like sectors. This algebra determines how chosen coordinates move. GRC additionally asks whether those coordinates can express the required repair.

6.2 DEEP AND RESIDUAL LEARNING

A differentiable family $\mathcal{F}_{\mathcal{P}} = \{M_\theta : \theta \in \Theta\}$ fixes an outer machine grammar. Parameter changes can alter relations and latent representations, so they are not automatically Transformation learning. For a residual block (He et al., 2016),

$$h_{\ell+1} = h_\ell + F_\ell(h_\ell),$$

the forward program fixes the residual owner and the additive correction form. This is a forward-map correspondence, not a proof of lawful closure or an embedding of the full training dynamics. Completion can reopen the owner, correction form, or machine family.

6.3 RL AND CONTROL

Classical reinforcement learning and dynamic programming formulate value, policy, and Bellman updates on a declared state–action representation (Sutton & Barto, 2018). With a future-sufficient state abstraction held fixed, lawful persistent updates that preserve $[C_M]$ are Transformation closures. For a sufficient control abstraction $s = C(x)$, with specified actions, transitions, and costs, $\mathcal{T}V - V$ is a Bellman consistency residual (Bellman, 1952). Updates preserving the registered encoder class are Transformation steps. If a proposed C instead merges detailed histories with inequivalent required continuations, a law depending only on s cannot represent both. This state-aliasing obstruction concerns the abstraction, rather than the optimizer (Givan et al., 2003; Littman et al., 2001). Recovering a particular control or RL algorithm also requires its sampling, update, and feasibility semantics.

6.4 CONTAINMENT AND COMPLETION

A completion preserves lawful incumbent realizations while reopening selected choices: $\mathcal{F}_A \subseteq \mathcal{F}_{A^+}$. Under a common objective, information rights, constraints, and complete cost accounting,

$$J^\star_{A^+} \le J^\star_A, \qquad J^\star_A = \inf_{M \in \mathcal{F}_A} J_\Xi(M).$$

Strict improvement requires a lawful witness outside $\mathcal{F}_A$ with lower objective, including search, audit, compilation, maintenance, and deployment costs. Inclusion does not guarantee discovery or runtime advantage. Appendix G provides the derivations; Appendix G.9 discusses further structural correspondences and their differing proof status.

7 CONDITIONAL GRC REPRESENTATION AND INTELLIGENCE

7.1 A CONDITIONAL GRC REPRESENTATION

Can an abstract learner admit the preceding operational description? Let $\mathcal{X}$ contain detailed interaction states or histories, and let $M \in \mathcal{M}$ be its complete persistent state. Under a fixed contract and

Table 1: Representative restricted choices and candidate GRC revisions. Rows describe conditional correspondences, not universal algorithmic equivalence or guaranteed improvement.

Paradigm	Predeclared structure	Candidate GRC revision
Optimization	Coordinates and metric	Representation or metric
Deep learning	Differentiable family	Outer machine family
ResNet	Interface and additive form	Owner or correction form
LoRA/adapters (Hu et al., 2022)	Factorized edit family	Rank, family, or owner
RAG (Lewis et al., 2020)	Evidence-access boundary	Separate missing evidence from internal failure
RL/control	State/action abstraction	State representation
Meta-learning/NAS (Finn et al., 2017; Liu et al., 2019)	Meta/search family	The family itself

learner–world boundary, registered responses define

$$\begin{aligned} x \sim_{\Xi,M} y \iff \mathrm{Resp}_M(c,x) \equiv_\Xi \mathrm{Resp}_M(c,y) \quad \forall c \in \mathrm{Ctx}_\Xi, \\ K_M = \mathcal{X}/\!\sim_{\Xi,M}, \qquad C_M(x) = [x]_{\sim_{\Xi,M}}. \end{aligned} \tag{22}$$

The response identity must induce an equivalence relation for each fixed M. Thus K_M identifies detailed states, not entire learners. Write $\varpi_\Xi(M) = [C_M]$ for the encoder's class up to admissible relabeling; relation learning can change responses without changing this partition.

Theorem 7.1 (T4: Conditional GRC representation)**.** *Under a fixed (Ξ, ∂), suppose a persistent learner satisfies R1–R8 of Appendix I: state completeness, future identity, conditional dynamic sufficiency, nontrivial discrepancy response, representation revisability, future viability, fail-closed persistence, and minimum complete-cost selection. On the events admitting the required quotient-level laws, its accepted resolving transitions have an operationally equivalent description through K_M, relation laws ρ_M, and generalized residuals. A transition either preserves $[C_M]$ (Transformation) or changes it (Representation), possibly with simultaneous relation changes. For the learner's inherited lawful resolving family,*

$$M^+ \in \operatorname*{arg\,min}_{N \in \mathrm{Law}_\Xi(M,r)} C_{\mathrm{full}}(N \mid M). \tag{23}$$

Proof sketch. Future identity gives the quotient; dynamic sufficiency permits continuation laws to descend. Comparing the current law's prediction with realized continuation gives a residual. The endpoint partition test gives the T/R classification, and the inherited acceptance semantics and R8 give lawful minimum-cost selection. Appendix I supplies the construction. Failure of descent is recorded as an obstruction, not a missing map supplied by the theorem. The prediction can be wrong even when an exact induced continuation exists.

The conclusion preserves registered responses and accepted transition semantics. It does not establish a unique implementation, efficient discovery, or superiority over other sectors. R8 includes an actual selected minimizer; candidate admission alone guarantees neither search success nor an attained optimum.

Independent capitalization property. R9 additionally permits repeatedly certified, recurrent relations to be compiled into reopenable primitives p_ρ, while preserving semantics, dynamic sufficiency, and future viability with lower declared complete future cost. This permits recursive Open/Compile operation when the eligibility conditions hold. A primitive is distinct from the entire quotient K_M; the two operations are not algebraic inverses. Neither R9 nor strict capability growth follows from R1–R8.

7.2 An operational view of intelligence

GRC suggests interpreting intelligence as the recursive capacity to turn unresolved future-relevant distinctions into reusable closure capability without exhausting the ability to learn again. The

capability–cost profile of Section 5 gives this interpretation a declared operational scope: which future demands remain lawfully closable, and at what complete cost.

Reasoning performs temporary closure search; learning installs a certified persistent successor; compiled memory retains a reopenable realization of acquired capability. Repeatedly useful relations can thereby become primitives supporting later relations. These are distinct outcomes: a successful trial need not be a persistent commit, and a commit need not qualify for compilation. The account is a contract-relative interpretation, not a unique or universal scalar definition of intelligence. Appendix J develops its scope and failure tests.

8 Limitations and Conclusion

The theory is conditional. The two-mode classification fixes the learner–world boundary and assumes a complete representation–relation description; its tangent decomposition additionally requires regular smooth structure. The compatibility construction is exact in its affine sector and first-order for nonlinear learners, whose finite steps require endpoint semantic audit. Future viability is contract-relative, and finite probes do not certify arbitrary unknown futures. Certified growth does not imply discovery or liveness. Sector containment does not imply efficient rediscovery or runtime superiority, and the representation theorem does not establish a unique implementation.

These restrictions separate representation failure from missing evidence, search incompleteness, and environmental transport. They also distinguish local response geometry from global guarantees. The counterexamples in Appendix H make those boundaries explicit.

GRC organizes learning around four connected ideas: a residual diagnoses insufficiency without choosing its repair location; internal revision at a fixed boundary changes relations, representation, or both; stability preserves registered future-relevant identity through change, while plasticity is the useful response that remains lawful; and recurrent successful closure can become reusable capacity when its full costs and future obligations justify compilation. In this view, learning recursively constructs, repairs, and recompiles the state space in which further learning becomes possible.

AI/LLM use statement

AI/LLM tools were used only for writing assistance and language polishing. The mathematical source material, theorem arguments, and proofs were supplied by the author. AI/LLMs were not used to generate new mathematical results or supply missing proofs.

Reproducibility statement

This paper develops a theoretical framework without reporting computational experiments. Appendix A collects the semantic and geometric assumptions; Appendices C, D, F, and I give the arguments for T1–T4. Appendix H records counterexamples and impossibility regimes. Conditions still needed for broader results are identified explicitly, including the grade construction in Appendix E and unrestricted certification and discovery.

References

Shun-ichi Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10(2):251–276, 1998.

Marcin Andrychowicz, Misha Denil, Sergio Gómez, Matthew W. Hoffman, David Pfau, Tom Schaul, Brendan Shillingford, and Nando de Freitas. Learning to learn by gradient descent by gradient descent. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Richard Bellman. On the theory of dynamic programming. *Proceedings of the National Academy of Sciences of the United States of America*, 38(8):716–719, 1952.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Shibhansh Dohare, J. Fernando Hernandez-Garcia, Qingfeng Lan, Parash Rahman, A. Rupam Mahmood, and Richard S. Sutton. Loss of plasticity in deep continual learning. *Nature*, 632:768–774, 2024.

Kevin Ellis, Catherine Wong, Maxwell Nye, Mathias Sablé-Meyer, Lucas Morales, Luke Hewitt, Luc Cary, Armando Solar-Lezama, and Joshua B. Tenenbaum. DreamCoder: Bootstrapping inductive program synthesis with wake-sleep library learning. In *Proceedings of the 42nd ACM SIGPLAN International Conference on Programming Language Design and Implementation*, pp. 835–850. ACM, 2021.

Mehrdad Farajtabar, Navid Azizan, Alex Mott, and Ang Li. Orthogonal gradient descent for continual learning. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 3762–3773. PMLR, 2020.

William Fedus, Barret Zoph, and Noam Shazeer. Switch Transformers: Scaling to trillion parameter models with simple and efficient sparsity. *Journal of Machine Learning Research*, 23(120):1–39, 2022.

Norm Ferns, Prakash Panangaden, and Doina Precup. Bisimulation metrics for continuous Markov decision processes. *SIAM Journal on Computing*, 40(6):1662–1714, 2011.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1126–1135. PMLR, 2017.

Karl Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11: 127–138, 2010.

Robert Givan, Thomas Dean, and Matthew Greig. Equivalence notions and model minimization in Markov decision processes. *Artificial Intelligence*, 147(1–2):163–223, 2003.

Isaac Han, Sangyeon Park, Seungwon Oh, Donghu Kim, Hojoon Lee, and KyungJoong Kim. FIRE: Frobenius-isometry reinitialization for balancing the stability–plasticity tradeoff. In *International Conference on Learning Representations*, 2026.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems 33*, pp. 9459–9474, 2020.

Michael L. Littman, Richard S. Sutton, and Satinder Singh. Predictive representations of state. In *Advances in Neural Information Processing Systems*, volume 14, pp. 1555–1561, 2001.

Hanxiao Liu, Karen Simonyan, and Yiming Yang. DARTS: Differentiable architecture search. In *International Conference on Learning Representations*, 2019.

David Lopez-Paz and Marc'Aurelio Ranzato. Gradient episodic memory for continual learning. In *Advances in Neural Information Processing Systems 30*, pp. 6467–6476, 2017.

D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

Haomiao Qiu, Miao Zhang, Ziyue Qiao, Weili Guan, Min Zhang, and Liqiang Nie. SplitLoRA: Balancing stability and plasticity in continual learning through gradient space splitting. In *International Conference on Learning Representations*, 2026.

R. Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.

Cosma Rohilla Shalizi and James P. Crutchfield. Computational mechanics: Pattern and prediction, structure and simplicity. *Journal of Statistical Physics*, 104:817–879, 2001.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2 edition, 2018.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems 30*, pp. 5998–6008, 2017.

A Assumptions and Formal Semantic Setting

The assumptions below are used selectively. Each result requires its own declared subset; they are not simultaneous universal axioms for every learner.

A.1 Contract, state, and boundary

The operational contract Ξ registers queries, actions, feedback, update events, evidence rights, interventions, tolerances, resources, and safety constraints. Equivalence, stability, sufficiency, viability, and growth are relative to this contract. A detailed interaction state or history is $x \in \mathcal{X}$; the persistent learner state is $M \in \mathcal{M}$.

A0: Persistent-state completeness. Every internal variable whose persistent value can affect a registered response or later learning transition belongs to M. Temporary reasoning traces, speculative candidates, and computational workspace need not belong to M unless they become persistent. An exhaustive classification requires this state description to include the relevant update variables.

A1: Fixed learner–world boundary. The internal-transition classification fixes the declared interface $\partial_{t+1} = \partial_t = \partial$. The world can still change and the learner can observe, act, and receive evidence through that interface. A query or measurement that changes available evidence without changing persistent learner state is an external resolution action. If sensing or action machinery becomes a persistent learning variable, it must be included in an enlarged state description. Its transition is classified there by its actual effect on the representation and relations; changing interface machinery alone does not establish a changed state partition.

A.2 Future identity and representation

For each fixed learner M, registered contexts define the response equivalence of Equation (1).

A2: Future-operational congruence. The relation $\sim_{\Xi,M}$ is a well-defined equivalence on $\mathcal{X}$. Each operation claimed to descend to the quotient additionally respects that equivalence on its declared domain. Thus

$$K_M = \mathcal{X}/{\sim_{\Xi,M}}, \qquad C_M : \mathcal{X} \to K_M, \qquad k = C_M x.$$

The quotient space, the quotient map, and a represented state are distinct. A tolerance discrepancy is not automatically a transitive equivalence relation.

Encoders describe the same abstraction when $C_N = \phi \circ C_M$ for an admissible isomorphism $\phi : K_M \to K_N$. Write $C_N \cong C_M$ and $\varpi_\Xi(M) = [C_M]$. This removes mere relabeling of the same partition; it does not identify entire learners by equality of every response. Old responsibilities are separately registered by Ξ_{old} and, in the affine analysis, encoded by $\mathcal{S}_{\mathrm{old}}(M)$.

A3: Representation–relation completeness. For a fixed encoder, let $\mathrm{Rel}_\Xi(C_M)$ contain its admissible relation laws and let ρ_M be the active law. This is the encoder-indexed description of the Ring used in Section 2, with the same admissibility conditions. Every persistent internal change relevant to the claimed complete description must be represented by a relation-law change at fixed abstraction, a change of abstraction, or both. This is an exhaustiveness assumption about the chosen description. It is not a consequence of the elementary kernel–image exact sequence and is not an empirical assertion about an arbitrary parameterization.

A.3 Regular strata and lawful transitions

Discrete structural changes may be organized by a stratified learner space

$$\mathcal{M}_\Xi^{(\partial)} = \bigsqcup_{\sigma \in \Sigma_\Xi} \mathcal{S}_{\Xi,\sigma}.$$

A stratum can fix a discrete ownership structure, memory topology, program support, or grammar. A change of stratum need not change the quotient partition.

A4: Local regularity. Where the differential factorization is invoked, M belongs to a finite-dimensional regular smooth stratum and the smooth map $\varpi_\Xi : \mathcal{S}_{\Xi,\sigma} \to \mathfrak{B}_\Xi$ has locally constant rank. The horizontal/vertical decomposition is unique only after a complementary space or connection is chosen. This structure is additional to the existence of a future-behavior quotient. Cross-stratum events are compared through their endpoint representations; no global tangent decomposition is implied.

For current residual r, $\mathrm{Adm}_\Xi(M, r)$ enforces the hard obligations relevant to a transition claim: old-semantic reconstruction, dynamic sufficiency, future viability, information rights, resource and safety conditions, and current closure where it is required. In Equation (3), closure is displayed separately. The resolving set $\mathrm{Law}_\Xi(M, r)$ of T4 includes closure to tolerance. These obligations are not mutually tradable soft penalties.

A5: A lawful baseline when invoked. A comparison that relies on fallback must contain a fallback satisfying that comparison's declared obligations, often the incumbent or a lawful transport. An available incumbent is not automatically lawful. In particular, one that leaves the present residual unresolved need not belong to a resolving successor set. T3's rejection branch retains a transported baseline without asserting current closure. Nonemptiness of a resolving set, attainment of a least-cost solution, and availability of a fallback are separate requirements.

A.4 Costs, closure load, and diagnosis

The declared complete cost $C_{\mathrm{full}}(N \mid M)$ may charge editing, search, audit, deployment, maintenance, and compilation. Its resource scope and accounting origin must be fixed, including how unsuccessful trials are charged.

A6: Comparable costs and information. Feasible-family comparisons use common information rights, output semantics, initialization conventions, objectives, and resource accounting. Additional search or audit cannot be treated as free overhead. A comparison of future continuation costs need not be a comparison of total lifecycle costs.

A7: Identifiable closure-load scale. A scalar closure-load comparison requires a declared family with

$$\lambda_1 \le \lambda_2 \implies D_{\lambda_1}\mathcal{C} \subseteq D_{\lambda_2}\mathcal{C}.$$

The scale and accessible set must be identified sufficiently for the claimed comparison. A Minkowski-gauge interpretation additionally requires the corresponding radial scaling and origin/regularity conditions. Nonconvexity does not prevent writing an infimum, but it prevents importing convex-gauge properties without their hypotheses. If only bounds or set-valued information are available, the output can be an interval or a scoped certificate such as INSIDE, CROSSES, OUTSIDE, or UNIDENTIFIED. The meaning and reliability of these certificates must be specified; their names are not proofs of reachability. Appendix E records the relevant distinctions.

A8: No direct overflow-to-Carry inference. Excessive same-representation load does not identify a representation obstruction. Diagnosis must distinguish saturation of the declared response family, transport mismatch, missing evidence or boundary limitations, incomplete search, and a certified expressive obstruction. This is a logical distinction, not a mandatory fixed routing algorithm. Representation necessity requires a class-wide certificate such as Proposition 3.2; it does not follow from a numerical solver failure.

A.5 Result-specific assumptions

The following map identifies the additional conditions used in the present results.

- T1 uses a complete persistent state and a fixed boundary for its learning interpretation, with A4 for the regular differential setting. Its stronger global interpretation invokes A3; the exact sequence alone is a local linear-algebra construction.
- P2 fixes the representation, evidence, exhaustive admissible relation class, discrepancy, and tolerance. A strict infimal obstruction implies necessity; equality at the tolerance boundary requires an attainment argument for a converse.

- T2 uses the exact affine response model, a common edit for all old responsibilities and the new residual, and $H \succ 0$. With nonlinear Jacobians it describes first-order behavior; finite-endpoint preservation is an additional requirement.
- Finite future-probe certificates concern only their registered probe family and information rights. Unavailable future-task labels cannot be used to certify current viability.
- T3 assumes sound audit, a declared capability order and dominance gate, consistent full-cost accounting, and a lawful comparison transport when the background changes. It does not establish search liveness.
- T4 uses its separately enumerated R1–R8 and the stipulated existence of quotient-level relations on their applicability domain. Compilation requires the additional R9 property. The extensional construction alone supplies neither the regularity of A4 nor an executable full-state encoding.

B Explain–Predict Semantics

B.1 Opening and Explanation

For a fixed learner M, the Kernel $K_M = \mathcal{X}/{\sim_{\Xi,M}}$ is a quotient state space, C_M is its encoding map, and $k = C_M x$ is a current state. An admissible opening is a correspondence

$$\mathrm{Open}_{\Xi,M}(k) \subseteq \mathcal{R}_{\Xi}(K_M), \qquad \rho \in \mathrm{Open}_{\Xi,M}(k). \tag{24}$$

Here the Ring comprises relational realizations admitted for the fixed encoding and its operational contract. The source's notation $\mathrm{Rel}_{\Xi}(C_M)$ emphasizes that same dependence; identifying the two descriptions requires the same admissibility semantics. The quotient alone does not select a unique relation law, and the correspondence does not assert that a suitable opening always exists.

Explanation means operationally reopening the represented state into relations through which it acts, predicts, controls, or returns. It need not be an explanation in natural language. A selected opening supplies the active realization used for the following prediction.

B.2 Prediction and Its Comparison Space

For event e, the selected realization predicts a continuation state,

$$\widehat{k}_{t+1} = \mathrm{Pred}_{M_t,e}(\rho_t). \tag{25}$$

This is the constraint that the active relation places on an unrealized future. It is broader than scalar-target or next-token prediction. Where a selected opening and prediction rule compose into a state-level map, we write $\widehat{k}_{t+1} = \widehat{U}_{M_t,e}(k_t)$. The hat denotes the learner's prediction, not an exactly induced update.

Prediction and evaluation must share a declared comparison space, denoted here by $Z_{t,e}$. Its declaration includes any admissible comparison or transport needed after the event. The notation $\widehat{k}_{t+1}$ refers to an element of this space; it does not predict a new quotient $K_{M_{t+1}}$ or presuppose that a successor learner has already been accepted.

B.3 Realized Continuation Before Learning

Let

$$U_e : \mathcal{X}_t \longrightarrow \mathcal{X}^-_{t+1}, \qquad x^-_{t+1} = U_e(x_t), \tag{26}$$

be the detailed continuation generated by the event before a persistent learning update. If the current representation has a lawful evaluation on that domain, declare

$$C^e_{M_t} : \mathcal{X}^-_{t+1} \longrightarrow Z_{t,e}, \qquad k^{\mathrm{real}}_{t+1} = C^e_{M_t}(x^-_{t+1}). \tag{27}$$

The superscript e records evaluation under the event's declared domain or transport; it does not denote a newly learned encoder. In the fixed-domain case, $Z_{t,e} = K_{M_t}$ and $C^e_{M_t} = C_{M_t}$ recover the expression in Section 2.

Such an evaluation is not automatic. If a missing future distinction prevents a lawful quotient-level continuation, record a representation obstruction. An undefined evaluation is neither a numerical defect nor evidence of zero residual. The contract must specify how that obstruction is witnessed.

B.4 The Pre-learning Residual

When both values are well typed in the comparison space, define

$$\mathfrak{R}^{\Xi}_{t,e} = \mathrm{Def}_{\Xi}\big(k^{\mathrm{real}}_{t+1}, \widehat{k}_{t+1}\big). \tag{28}$$

The defect object need not be a scalar loss. Its operational meaning and closure tolerance are part of the contract. This sequence separates explanation, prediction, realized continuation, and the observed residual.

A residual diagnoses insufficiency without prescribing its own write address. In particular, $\mathfrak{R} \neq 0$ does not by itself require a parameter update. Resolution may involve relation revision, representation revision, acquiring admissible evidence, lawful transport, or further search. Boundary actions and temporary search must still be distinguished from persistent learner-internal changes.

B.5 Update Commutation Is a Separate Requirement

For declared source and target encoders $C_t : \mathcal{X}_t \to K_t$ and $C_{t+1} : \mathcal{X}_{t+1} \to K_{t+1}$, update sufficiency asks whether an induced updater $\bar{U}_e : K_t \to K_{t+1}$ exists with

$$C_{t+1} \circ U_e \simeq_{\Xi} \bar{U}_e \circ C_t. \tag{29}$$

Here K_t and K_{t+1} denote the quotients associated with those declared encoders. Their detailed update domain must be specified; it is not silently identified with the pre-learning domain in Equation (26). For a typed update diagram, the source writes its operator-level diagnostic as

$$\mathfrak{R}^{\Xi,\mathrm{upd}}_{e} = \mathrm{Def}_{\Xi}\big(C_{t+1} \circ U_e, \bar{U}_e \circ C_t\big). \tag{30}$$

This diagnostic additionally requires a declared comparison of maps. If $\bar{U}_e$ is already the exactly induced updater, the diagram commutes and its exact defect is zero. If no such updater exists, that failure is an obstruction to update sufficiency, not a defined zero value of the expression. Approximate commutation requires its own error contract.

Thus $\widehat{U}_{M,e}$ makes a prediction that may fail, whereas $\bar{U}_e$ denotes a compatible induced update when one exists. The pre-learning residual diagnoses an observed failure using the current representation. Commutation tests whether declared representations support correct updating; it does not define today's learning signal through an unaccepted future representation.

B.6 Prediction and Compilation Have Different Roles

Prediction constrains an unrealized continuation. Compilation instead makes a repeatedly validated, reopenable relation available as a persistent primitive after successful use. Writing p_ρ for that primitive,

$$\rho^{\mathrm{certified}} \xrightarrow{\mathrm{Compile}} p_\rho. \tag{31}$$

The symbol denotes the installed primitive, not a predicted world state or an entire quotient space. Its installation and subsequent reopening must satisfy the relevant representation, sufficiency, and cost contract. No identity between Predict and Compile is assumed, and an accepted closure need not already qualify for compilation.

B.7 Recursion and Continued Sufficient Operation

The cycle begins with $k_t = C_{M_t}(x_t)$ and active relations ρ_t. Prediction and realized continuation supply the comparison in Equation (28). Diagnosis and lawful closure may then produce a persistent successor M_{t+1} with its own encoding and active relations. For a declared successor detailed state,

$$k_{t+1} = C_{M_{t+1}}(x_{t+1}). \tag{32}$$

The encoding acts on x_{t+1}, not on the learner M_{t+1}. Any writeback, state transport, or domain change from x^{-}_{t+1} to x_{t+1} must be declared separately.

The successor supplies the conditions for subsequent opening and prediction. A Transformation may preserve the quotient up to an admissible isomorphism while changing the active relation law. Closure therefore need not change the state partition, and repeated operation need not expose another defect: the next residual can remain zero. Compilation is a separate qualified transition, not a mandatory stage after every closure.

Scope of the semantic construction. P1 specifies typed operational semantics. It does not establish general opening existence, background transport, obstruction detection, or quantitative multi-step approximate sufficiency. Stronger claims require additional definitions and proofs.

C Two-Mode Factorization and Representation Necessity

C.1 Local representation projection

Fix the contract Ξ and learner–world boundary ∂. The persistent learner-state space may be stratified as

$$\mathcal{M}_{\Xi}^{(\partial)} = \bigsqcup_{\sigma \in \Sigma_{\Xi}} \mathcal{S}_{\Xi,\sigma}, \qquad \varpi_{\Xi}(M) = [C_M].$$

The representation projection records the abstraction class, rather than the current value $k = C_M(x)$ or a particular engineering realization. For the local statement, M lies in a finite-dimensional regular stratum on which ϖ_{Ξ} is smooth and has locally constant rank. Tangent spaces below refer to that stratum. In particular, $V_M = \ker D\varpi_{\Xi}(M)$ consists of directions that preserve the representation class to first order. Completeness of the persistent-state description and the representation–relation description are the declared assumptions, including Assumption A3 in Appendix A.

C.2 Proof of Theorem 3.1

Proof. Write $L = D\varpi_{\Xi}(M)$ and $V_M = \ker L$. The inclusion of V_M into the tangent space is injective, and the quotient by V_M is naturally isomorphic to $\operatorname{im} L$. This gives the exact sequence (5).

Choose a right splitting $s_M : \operatorname{im} L \to T_M\mathcal{M}_{\Xi}^{(\partial)}$ with $L \circ s_M = I$, and let $H_M = \operatorname{im} s_M$. For an arbitrary tangent direction d, set

$$d_R = s_M(Ld), \qquad d_T = d - d_R.$$

Then $Ld_T = Ld - Ls_M(Ld) = 0$, so $d_T \in V_M$, while $d_R \in H_M$ by construction. Since $V_M \cap H_M = \{0\}$, the decomposition is unique for the chosen splitting. The horizontal representative depends on that choice; whether $Ld = 0$ does not. □

Local constant rank supports a regular family of vertical spaces; the pointwise exact sequence itself is the standard kernel–image construction. Neither the argument nor the finite-jump classification selects a canonical connection or establishes a global smooth decomposition across strata.

Interpretation and joint learning. A Transformation component satisfies $D\varpi_{\Xi}(M)[d_T] = 0$. It may change weights, policies, transitions, routing, controllers, or other relation variables, provided that it preserves the abstraction to first order. A Representation component changes the abstraction when $D\varpi_{\Xi}(M)[d_R] \neq 0$. Examples can include state splitting, memory-state creation, or a change of the representational grammar; their classification depends on their effect on C_M, not on the implementation name. An event can have both $d_T \neq 0$ and $d_R \neq 0$. This is a joint event, not a third primitive component or a partition of algorithms into mutually exclusive families. The chosen horizontal vector is not canonical, whereas the test $D\varpi_{\Xi}(M)[d] = 0$ is independent of the splitting.

C.3 A fixed-representation Hilbert-space sector

Suppose future responses lie in a Hilbert space $\mathcal{H}$ and the responses expressible under C form a closed linear subspace $\mathcal{A}_C$. Let Π_C be its orthogonal projector and $Q_C = I - \Pi_C$. For a target u and a current response $\widehat{u} \in \mathcal{A}_C$, the decomposition is

$$u - \widehat{u} = \underbrace{\Pi_C u - \widehat{u}}_{r_{\mathrm{rel}}} + \underbrace{Q_C u}_{r_{\mathrm{rep}}}.$$

For any $\widetilde{u} \in \mathcal{A}_C$, orthogonality gives

$$\|u - \widetilde{u}\|^2 = \|Q_C u\|^2 + \|\Pi_C u - \widetilde{u}\|^2, \qquad \inf_{\widetilde{u} \in \mathcal{A}_C} \|u - \widetilde{u}\| = \|Q_C u\|.$$

Thus r_{rep} gives an irreducible error floor in this sector. The closed-linear-subspace assumption is substantive: the calculation is not an orthogonal decomposition of arbitrary nonlinear program classes. Nor does representability certify that an admissible finite-budget search finds the projected response.

C.4 Proof of Proposition 3.2

Proof. By the definition of the infimum, every $f \in \mathfrak{R}(C)$ satisfies

$$D_{\Xi}(F, f \circ C) \geq \Omega_C(F).$$

If $\Omega_C(F) > \varepsilon_{\Xi}$, each admissible relation law therefore violates the declared tolerance. Under the fixed-boundary, fixed-evidence, and exhaustive-relation-class assumptions, a successful internal successor cannot remain in that representation class. □

This is a necessary condition on a successful repair, not an existence theorem for a revised representation. A converse at $\Omega_C(F) = \varepsilon_{\Xi}$ would require attainment; it is not asserted here. A numerical search failure also does not certify the obstruction.

C.5 Finite transitions and conditional global classification

A finite transition $M^- \to M^+$ need not admit a tangent description. Compare its representation classes on the same declared detailed-state domain. If $C_{M^+} \cong C_{M^-}$, the abstraction is unchanged: the event can be a fixed-representation relation change or a realization change. Otherwise it contains a Representation revision. This comparison also applies across engineering strata, so a discrete recompilation or reparameterization is not automatically Representation learning.

Under Assumption A3 in Appendix A, these cases exhaust the relevant persistent internal changes at the fixed boundary.

Argument under Assumption A3. For each persistent transition, exactly one of

$$[C_{M^+}] = [C_M] \qquad \text{or} \qquad [C_{M^+}] \neq [C_M]$$

holds. In the first case, A3 places the represented learning content in relation or realization changes within the current abstraction. In the second, the representation class changes and relation changes may accompany it. This gives the asserted conditional classification. □

The equality test is set-theoretic; the assertion that this description captures every relevant persistent change is the completeness assumption A3. The argument does not establish that assumption for arbitrary learners. It also does not supply a global connection, a canonical finite Transformation/Representation decomposition, or smooth coordinates across singular strata. The exact-sequence theorem adds a local linear factorization where its separate regularity assumptions hold. A pure realization change or a no-op need not acquire any new capability, and this classification alone does not establish preservation of old responsibilities.

C.6 Boundary actions and persistent incorporation

A one-time query, retrieval, sensor use, or intervention can resolve an external information need without changing persistent learner state. An already permitted call need not change the boundary itself. If new evidence is incorporated persistently, its internal effect is classified by changes to the relation law, the representation class, or both under A3.

If stateful interface machinery becomes part of the learner, include it in the complete persistent state as required by Assumption A0 in Appendix A, and restate the boundary and comparison domain. Its modification is then classified by whether C_M actually changes. It is not automatically Representation learning: the modification may preserve the abstraction and change a relation or realization instead. Actual changes to information rights or the available action set require that enlarged setting to be declared before applying a fixed-boundary classification. External resolution is therefore distinct from the conditional classification of persistent internal changes.

D Affine Stability–Plasticity Compatibility

All statements in this appendix use the finite-dimensional affine model of Section 4: an edit $d \in \mathbb{R}^n$ satisfies $Ad = 0$, the updated residual is $r - Bd \in \mathbb{R}^m$, and the edit energy is $d^\top H d$. Here $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{m\times n}$, and $H \in \mathbb{R}^{n\times n}$ is positive definite. These relations are exact in the stated affine sector. When A and B are nonlinear learner Jacobians, they describe first-order behavior only.

D.1 The semantic-safe inverse

Let P_H be the matrix in Equation (9). Introduce

$$\widetilde{A} = AH^{-1/2}, \qquad Q_A = I - \widetilde{A}^\top(\widetilde{A}\widetilde{A}^\top)^\dagger \widetilde{A}.$$

Then Q_A is the Euclidean orthogonal projector onto $\ker \widetilde{A}$, and $P_H = H^{-1/2} Q_A H^{-1/2}$. In general, P_H itself is not an idempotent Euclidean projector.

Lemma D.1 (Properties of the semantic-safe inverse)**.** *The matrix P_H satisfies*

$$P_H = P_H^\top \succeq 0, \qquad AP_H = 0, \qquad P_H A^\top = 0,$$
$$\operatorname{im} P_H = \ker A, \qquad \ker P_H = \operatorname{im} A^\top,$$
$$P_H H d = d \quad \textit{for every } d \in \ker A.$$

Proof. Because Q_A is an orthogonal projector, congruence by $H^{-1/2}$ gives symmetry and positive semidefiniteness of P_H. The identity $\widetilde{A} Q_A = 0$ gives $AP_H = 0$; symmetry gives $P_H A^\top = 0$. Moreover,

$$\operatorname{im} P_H = H^{-1/2} \ker(AH^{-1/2}) \subseteq \ker A.$$

For the reverse inclusion, let $z \in \ker A$ and put $u = H^{1/2} z$. Then $AH^{-1/2}u = Az = 0$, so $Q_A u = u$. Consequently,

$$P_H H z = H^{-1/2} Q_A H^{1/2} z = z,$$

which proves the range identity and the stated action on safe directions. Finally, $P_H x = 0$ is equivalent to $Q_A H^{-1/2} x = 0$, hence to

$$H^{-1/2} x \in \operatorname{im} \widetilde{A}^\top = \operatorname{im}(H^{-1/2} A^\top).$$

Since $H^{-1/2}$ is invertible, this is equivalent to $x \in \operatorname{im} A^\top$. □

The residual-response operator $K_{\text{plast}} = BP_H B^\top$ is therefore positive semidefinite, and $r^\top K_{\text{plast}} r = (B^\top r)^\top P_H (B^\top r)$ measures the residual's coupling to safe edit directions.

D.2 Proof of Theorem 4.1

Proof. Set $x = B^\top r$. If $P_H x \neq 0$, choose $d = P_H x$. Then $Ad = 0$ and

$$\langle r, Bd \rangle = x^\top P_H x > 0,$$

where strict positivity follows from positive semidefiniteness and $P_H x \neq 0$. This proves that (12) implies (11).

Conversely, if $P_H x = 0$, symmetry and $\operatorname{im} P_H = \ker A$ imply that x is orthogonal to $\ker A$. Hence $\langle r, Bd \rangle = x^\top d = 0$ for every $d \in \ker A$, excluding (11). Finally,

$$r^\top K_{\text{plast}} r = (B^\top r)^\top P_H (B^\top r),$$

which is positive exactly when $P_H B^\top r \neq 0$. Thus all three conditions are equivalent. Since $K_{\text{plast}} \succeq 0$, the same conditions are also equivalent to $K_{\text{plast}} r \neq 0$. □

D.3 THE PROXIMAL AFFINE STEP

For $\eta > 0$, the strictly convex quadratic objective in (14) has a unique minimizer. Its KKT conditions are

$$\eta^{-1} H d - B^\top (r - Bd) + A^\top \lambda = 0, \qquad Ad = 0.$$

Multiplying the stationarity equation by P_H, and using $P_H A^\top = 0$ and $P_H H d = d$ on $\ker A$, gives

$$d = \eta P_H B^\top (r - Bd).$$

Writing $r^+ = r - Bd$ gives $d = \eta P_H B^\top r^+$ and consequently

$$(I + \eta K_{\text{plast}}) r^+ = r.$$

Because $K_{\text{plast}} \succeq 0$, the matrix $I + \eta K_{\text{plast}}$ is positive definite. Substitution therefore yields (15) and (16). The identity $AP_H = 0$ gives $Ad^\star = 0$.

D.4 WEAK, STRICT, AND UNIFORM RESIDUAL CONTRACTION

Write $K_{\text{plast}} = U \Lambda U^\top$ with orthogonal U and $\Lambda = \operatorname{diag}(\lambda_i)$, $\lambda_i \geq 0$, and let $r = Uc$. The residual formula gives

$$\|r^+\|_2^2 = \sum_i \frac{c_i^2}{(1 + \eta \lambda_i)^2} \leq \sum_i c_i^2 = \|r\|_2^2. \tag{33}$$

The inequality is strict exactly when r has a nonzero component in a positive-eigenvalue eigenspace, equivalently $K_{\text{plast}} r \neq 0$. If $r \in \ker K_{\text{plast}}$, then $r^+ = r$. Thus positive semidefiniteness gives weak contraction, not strict improvement for every residual.

If the stronger bound $K_{\text{plast}} \succeq \kappa I$ holds on the entire registered residual space $\mathbb{R}^m$, with $\kappa > 0$, the same calculation gives

$$\|r^+\|_2 \leq (1 + \eta\kappa)^{-1} \|r\|_2.$$

This uniform statement uses a lower bound on every registered direction. A quadratic-form lower bound only on an arbitrary proper subspace is not used here as a certificate for the full resolvent's contraction on that subspace.

D.5 MINIMUM ENERGY FOR EXACT AFFINE CLOSURE

Proposition D.2 (Minimum semantic-safe edit energy)**.** *The constraints $Ad = 0$ and $Bd = r$ have a solution if and only if $r \in \operatorname{im} K_{\text{plast}}$. When they are feasible,*

$$\min_{Ad=0,\, Bd=r} d^\top H d = r^\top K_{\text{plast}}^\dagger r. \tag{34}$$

Proof. Let the columns of Z form a basis of $\ker A$. Every safe edit is $d = Zz$. Define $H_T = Z^\top H Z \succ 0$ and $G = BZ$. The problem becomes

$$\min_z z^\top H_T z \quad \text{subject to } Gz = r.$$

Its feasible residual space is $\operatorname{im} G$, and the standard minimum-energy formula gives the value $r^\top (G H_T^{-1} G^\top)^\dagger r$ there. The constrained inverse identity

$$Z (Z^\top H Z)^{-1} Z^\top = P_H$$

implies $G H_T^{-1} G^\top = B P_H B^\top = K_{\text{plast}}$. The feasible residual space is consequently $\operatorname{im} K_{\text{plast}}$, and the minimum energy is Equation (34). □

The basis reduction uses the usual zero-dimensional convention if $\ker A = \{0\}$; only the zero edit is then available.

D.6 A GEOMETRIC LOAD FOR THE STATED EDIT BUDGET

For a positive energy radius $R_0 > 0$ and budget $d^\top H d \le R_0^2$, define

$$q_{\mathrm{geo}}(r) = \begin{cases} \sqrt{r^\top K_{\mathrm{plast}}^\dagger r}/R_0, & r \in \operatorname{im} K_{\mathrm{plast}}, \\ +\infty, & r \notin \operatorname{im} K_{\mathrm{plast}}. \end{cases} \tag{35}$$

Proposition D.2 makes $q_{\mathrm{geo}}(r) \le 1$ equivalent to an exact affine semantic-safe closure within this energy budget. The range condition cannot be dropped: the pseudoinverse quadratic form alone does not certify feasibility.

This is an energy-normalized local load. Identifying it with the general operational closure load requires a matching edit family, resource scale, and full-cost convention. Calling an infeasible affine response a Transformation obstruction additionally requires the coordinates to be restricted to fixed-representation edits. Even then it does not establish global Representation necessity or exclude a nonlocal same-representation path.

D.7 A FIXED LINEAR CONTINUOUS-TIME FLOW

Keep A, B, and H fixed, start with $Ad(0) = 0$, and normalize time so that the safe edit velocity is $\dot{d} = P_H B^\top r$. With $r = r_0 - Bd$, the induced residual equation and its energy derivative are

$$\dot{r} = -K_{\mathrm{plast}} r, \qquad \frac{d}{dt}\frac{1}{2}\|r\|_2^2 = -r^\top K_{\mathrm{plast}} r \le 0. \tag{36}$$

If $K_{\mathrm{plast}} \succeq \kappa I$ on the entire registered residual space, the displayed inequality gives the stated bound

$$\|r(t)\|_2 \le e^{-\kappa t}\|r(0)\|_2.$$

These statements concern fixed affine operators. They do not provide a state-dependent nonlinear flow theorem, finite nonlinear endpoint preservation, or a derivation of the full discrete GRC procedure as a continuous-time limit. A rate for an arbitrary proper residual subspace is not asserted without further subspace conditions.

D.8 RECONSTRUCTIVE SEMANTIC STABILITY DOMINANCE

Let $T_{\mathrm{exact}} \subseteq \mathbb{R}^n$ be the linear safe space obtained by preserving a historical realization exactly, and let T_{rec} be the linear safe space obtained by preserving the registered future-operational semantics while permitting a different realization. The comparison assumes

$$T_{\mathrm{exact}} \subseteq T_{\mathrm{rec}}. \tag{37}$$

This is a condition on the declared preservation rules: exact preservation must impose the semantic requirements together with additional realization constraints. It need not hold for arbitrary pairs of constraint systems. Throughout the comparison, the positive-definite metric H and the residual-response map B are fixed.

For a linear subspace T, define its safe inverse by

$$P_T^H = H^{-1/2}\Pi_{H^{1/2}T}H^{-1/2}, \tag{38}$$

where $\Pi_{H^{1/2}T}$ is the Euclidean orthogonal projector onto the whitened subspace $H^{1/2}T$. For $T = \ker A$, this is the operator P_H of Lemma D.1. Write $P_j^H = P_{T_j}^H$ and $K_{\mathrm{plast}}^j = BP_j^H B^\top$ for $j \in \{\mathrm{exact}, \mathrm{rec}\}$.

Theorem D.3 (Reconstructive semantic stability dominance)**.** *Under* (37)*,*

$$\begin{aligned} P_{\mathrm{rec}}^H - P_{\mathrm{exact}}^H &\succeq 0, \\ K_{\mathrm{plast}}^{\mathrm{rec}} - K_{\mathrm{plast}}^{\mathrm{exact}} &= B(P_{\mathrm{rec}}^H - P_{\mathrm{exact}}^H)B^\top \succeq 0. \end{aligned} \tag{39}$$

In particular, this establishes the response-operator ordering in (17)*.*

Proof. Since $H^{1/2}$ is invertible, the whitened spaces satisfy $H^{1/2}T_{\mathrm{exact}} \subseteq H^{1/2}T_{\mathrm{rec}}$. For nested Euclidean subspaces $S_1 \subseteq S_2$, the difference $\Pi_{S_2} - \Pi_{S_1}$ is the orthogonal projector onto $S_2 \cap S_1^\perp$ and is therefore positive semidefinite. Consequently,

$$\Pi_{H^{1/2}T_{\mathrm{rec}}} - \Pi_{H^{1/2}T_{\mathrm{exact}}} \succeq 0.$$

Congruence by $H^{-1/2}$ gives the first ordering in (39); congruence by B gives the second. □

D.9 When the safe response improves strictly

The ordering is weak in general. A strict gain exists for at least one residual direction if and only if

$$B(P^H_{\mathrm{rec}} - P^H_{\mathrm{exact}})B^\top \neq 0. \tag{40}$$

For a specified residual r, the response quadratic form improves strictly exactly when

$$r^\top B(P^H_{\mathrm{rec}} - P^H_{\mathrm{exact}})B^\top r > 0. \tag{41}$$

Thus a strict enlargement of the safe space alone is insufficient: the newly available directions must couple to B and, for a given residual, to that residual direction. These are statements about the local response quadratic form, not a general ordering of finite-step residual norms.

D.10 Closure-load comparison on a common nondegenerate support

Let $\mathcal{Y} \subseteq \mathbb{R}^m$ be a common invariant residual subspace of $K^{\mathrm{exact}}_{\mathrm{plast}}$ and $K^{\mathrm{rec}}_{\mathrm{plast}}$ on which the former is positive definite. In particular, the comparison applies on a shared image space when both operators have that image. Invariance makes the restrictions operators on $\mathcal{Y}$; an arbitrary compression to a subspace is not being identified with the original closure problem. Set

$$K^{\mathcal{Y}}_j = K^j_{\mathrm{plast}}\big|_{\mathcal{Y}}, \qquad j \in \{\mathrm{exact}, \mathrm{rec}\}.$$

The PSD comparison gives $K^{\mathcal{Y}}_{\mathrm{rec}} \succeq K^{\mathcal{Y}}_{\mathrm{exact}} \succ 0$. Positive-definite inversion reverses Löwner order, hence

$$(K^{\mathcal{Y}}_{\mathrm{rec}})^{-1} \preceq (K^{\mathcal{Y}}_{\mathrm{exact}})^{-1}. \tag{42}$$

With the same edit-budget radius $R_0 > 0$, the geometric load of (35) therefore satisfies, for $r \in \mathcal{Y}$,

$$\begin{aligned} q^{\mathrm{rec}}_{\mathrm{geo}}(r) &= \frac{\sqrt{r^\top (K^{\mathcal{Y}}_{\mathrm{rec}})^{-1} r}}{R_0} \\ &\leq \frac{\sqrt{r^\top (K^{\mathcal{Y}}_{\mathrm{exact}})^{-1} r}}{R_0} = q^{\mathrm{exact}}_{\mathrm{geo}}(r). \end{aligned} \tag{43}$$

The inequality is strict whenever

$$r^\top \big[(K^{\mathcal{Y}}_{\mathrm{exact}})^{-1} - (K^{\mathcal{Y}}_{\mathrm{rec}})^{-1}\big] r > 0.$$

Every inverse here is taken on $\mathcal{Y}$. No ambient pseudoinverse-order claim is made outside a common nondegenerate support.

D.11 Interpretation and scope

Under the stated inclusion, imposing additional realization constraints that the registered semantics do not require cannot increase safe plasticity in the response-operator sense above. Forgetting violates a registered preservation requirement; rigidity preserves more than that contract requires. The distinction does not make all historical constraints unnecessary or establish that a reconstructive realization is inexpensive to discover or execute.

The comparison holds H, B, and the residual contract fixed; the load comparison also holds the edit budget fixed. It does not compare different metrics, response maps, total search or reconstruction costs, or arbitrary nonlinear finite updates. In particular, PSD ordering alone is not used to claim that every residual has a smaller Euclidean norm after a finite proximal step.

E Representation Obstruction, Closure Load, and Structural Carry

E.1 Solver Failure and Representation Obstruction

A failed solve does not identify its cause. A solver may miss an available relation, explore too narrow a family, or lack evidence available only through another declared interface. These possibilities differ from the absence of any admissible relation law on the current representation.

For consistency with Section 3, write $\mathfrak{R}(C)$ for the exhaustive admissible relation family on representation C. Let $\mathcal{A}_{\mathrm{search}}(C) \subseteq \mathfrak{R}(C)$ be the family actually explored, and let F denote the required future response. Its search-restricted obstruction is

$$\Omega_C^{\mathrm{search}}(F) = \inf_{f \in \mathcal{A}_{\mathrm{search}}(C)} D_\Xi(F, f \circ C). \tag{44}$$

A certified value above ε_Ξ rules out that searched family, not the entire representation. The loss of one returned candidate does not by itself certify this infimum either.

The exhaustive quantity is $\Omega_C(F)$ from Equation (7). Under the fixed boundary, evidence, tolerance, and exhaustive-family assumptions, $\Omega_C(F) > \varepsilon_\Xi$ excludes relation-only closure by Proposition 3.2. Its existing proof is in Appendix C.4; no separate proposition is needed here. Diagnosing this inequality from a computational failure requires ruling out incomplete search at the level of the claim.

The inequality $\Omega_C(F) \leq \varepsilon_\Xi$ does not automatically certify an attained feasible relation at the tolerance boundary. It also does not guarantee that a chosen solver finds a feasible relation, or exclude a revised representation with lower cost or better future capacity.

E.2 A Residual-Space Correction Body

Closure load compares a residual with attainable corrections in a common residual space. Let $\mathcal{F}^{\Xi}_{[C_M]}$ be the lawful realization fibre that preserves the representation class $[C_M]$. In particular, members preserve $\varpi_\Xi(M)$, not every acquired future response. Denote by $\Delta r(M, N)$ the correction induced by an admissible successor N in that common residual space. The unit-cost body is

$$\mathcal{C}_T^{\Xi}(M) = \left\{ \Delta r(M, N) : \begin{array}{l} N \in \mathcal{F}^{\Xi}_{[C_M]}, \\ C_{\mathrm{full}}(N \mid M) \leq 1 \end{array} \right\}. \tag{45}$$

The unit is a declared cost normalization. Each member must be realized by one lawful successor satisfying the protected duties together. Separately feasible corrections for different duties need not certify one joint edit. The body is not a set of machine states and need not be convex, connected, or star-shaped.

E.3 Calibrated Load and Its Boundary

Declare a family $D_\lambda \mathcal{C}_T^{\Xi}(M)$ of reachable corrections, nested in $\lambda > 0$. Interpreting $\lambda = 1$ as the original unit budget requires $D_1 \mathcal{C}_T^{\Xi}(M) = \mathcal{C}_T^{\Xi}(M)$. The load is Equation (4), with the convention $\inf \varnothing = +\infty$. This definition does not make the reachable sets identifiable or computable. The scalar-calibration requirements in Appendix A are additional obligations.

Where the declared scaling is calibrated to actual feasible edits, a load below one indicates a correction within the unit budget, and a load above one rules out a correction in that declared unit-cost family. This is a budget-relative statement. "Below one" is not a general assertion of topological interior. At $q_T = 1$, the value of an infimum alone does not certify a feasible boundary correction: the correspondence between a load threshold and membership must include the necessary attainment or set regularity conditions. No such conditions are presumed for an arbitrary nonlinear family.

Radial gauge sector. For standard radial dilation $D_\lambda \mathcal{C} = \lambda \mathcal{C}$, an absorbing body star-shaped about the origin gives the corresponding Minkowski gauge. Convex gauge properties are invoked only in a convex sector, with the origin, scaling, and any boundary regularity required by the particular property specified. This specialization is not a claim that all reachable correction sets have gauge geometry.

Several protected constraints. If a common calibrated radial description genuinely has joint correction body $\mathcal{C} = \bigcap_{i=1}^m \mathcal{C}_i$, the proposed scalar consolidation is

$$q(r) = \max_{1 \leq i \leq m} q_i(r). \tag{46}$$

Using this identity requires compatible gauge assumptions and a dilation that represents the same cost scale for every component. In particular, intersecting componentwise reachable images must

not be substituted for certifying a single edit that meets all constraints. The identity is not asserted for arbitrary response sets or differently calibrated diagnostics; those cases retain a vector or certificate description. A complete P3 argument must establish the joint-set and gauge conditions before applying the scalar formula to a learner family.

E.4 Bounds and Certificate-Valued Diagnosis

Disconnected, non-star-shaped, or partially identified correction sets need not admit a useful scalar summary. An interval $[\underline{q}, \overline{q}]$ can record certified load bounds when a common scalar comparison is justified. Otherwise retain

$$\mathrm{Cert}_T(r \mid M) \in \{\mathrm{INSIDE}, \mathrm{CROSSES}, \mathrm{OUTSIDE}, \mathrm{UNIDENTIFIED}\}. \tag{47}$$

INSIDE. A lawful same-representation closure within the declared budget is certified.

OUTSIDE. The absence of such a closure in the declared family is certified, with the comparison scope fixed.

CROSSES. Available certified bounds straddle the threshold and do not settle feasibility, including the applicable boundary qualification.

UNIDENTIFIED. Calibration or evidence is insufficient for a valid decision.

An implementation using inner and outer approximations must state which reachable set they bound and certify the corresponding inclusions. An inner witness must realize the joint obligations; an outer exclusion must cover the full declared family. No construction of such bounds, or general certificate-completeness theorem, is supplied by the load definition. Unknown calibration is not evidence for structural growth.

E.5 When Structural Carry Is Justified

The implication $q_T > 1 \Rightarrow$ Representation Learning is invalid. An overflow can reflect an exhausted budget, a changed background, unavailable evidence, or incomplete search. The diagnostic ordering

$$\begin{aligned} &\text{same-grade saturation} \longrightarrow \text{transport} \longrightarrow \text{boundary/evidence} \\ &\longrightarrow \text{search/solver} \longrightarrow \text{representation obstruction} \end{aligned}$$

expresses logical distinctions, not a mandatory execution schedule.

For a persistent, future-relevant defect, a structural Carry requires the following claims to be established under the declared information and cost contract:

1. Current same-representation closure is infeasible at the stated scope, and lawful transport does not remove the defect.
2. Missing admissible evidence, an inadequate comparison budget, and search or solver incompleteness do not explain the asserted representation obstruction.
3. The exhaustive obstruction satisfies $\Omega_C(F) > \varepsilon_\Xi$, and a lawful revised representation exists.

These conditions distinguish the necessity of leaving the present class from the existence of a successful successor. The notation

$$M^+ = \mathrm{LeastLawfulCarry}_\Xi(M, r)$$

specifies a least-lawful structural selection; it does not construct the successor, prove an attained minimum, or provide a solver. Carry may create, split, merge, or revise a state, owner, memory variable, or program component. It is not synonymous with adding parameters.

E.6 Filtration, Closure Valuation, and Learning Modes

Let $F_0(M) \subseteq F_1(M) \subseteq \cdots$ be lawful residual-response families on a common domain, where $F_s(M)$ contains distinctions closable at grade s. The costs, evidence rights, protected duties, and tolerances must be comparable. A calibrated description may identify

$$F_s(M) = \{r : q^M_{\Xi,s}(r) \le 1\}, \tag{48}$$

only after the threshold-to-feasibility and boundary conditions are established. Under that identification, define the closure valuation

$$\nu_{\Xi}(r \mid M) = \inf\{s : q^{M}_{\Xi,s}(r) \leq 1\}, \tag{49}$$

using $+\infty$ when no grade qualifies.

The load q_s asks about closure within a grade; ν_{Ξ} asks which grade can first qualify. Transformation/Representation classification asks instead whether the representation structure changes. Same-grade closure is pure Transformation only when that grade's edit family fixes the representation schema. Likewise, a larger grade must not be equated with representation revision merely because it permits a larger budget or search space. The implication from $\nu_{\Xi}(r) > s$ to required structural revision additionally needs the structural meaning of the grades and the exclusions in the Carry diagnosis.

E.7 Associated Grades Require Additional Structure

For an actual filtration of vector spaces or modules, with the appropriate subspace or submodule inclusions, one may form

$$V_s = F_s/F_{s-1}. \tag{50}$$

A class of $r \in F_s \setminus F_{s-1}$ then records a distinction not absorbed at the preceding grade. Cost-bounded reachable sets are not automatically such spaces. This notation therefore does not define a quotient for general nonlinear closure families.

An alternative notation $F_{s+1}/\sim_s$ also requires a genuine equivalence relation on the stated domain. Saying that the difference between two residuals is absorbable at grade s does not by itself establish reflexivity, symmetry, or transitivity in an arbitrary reachable set; even the meaning of a difference must be declared. No nonlinear equivalence theorem or general associated-graded algebra is established here.

Status of P3. The correction bodies, loads, diagnostic certificates, and filtration provide a conditional organization of closure. A complete P3 result still requires the claimed calibration, attainability, joint-feasibility, and structural-grade assumptions, together with any bound or Carry construction. The representation-necessity implication remains P2 and its existing proof. An OUTSIDE certificate alone is not a CARRY certificate.

F Dynamic Sufficiency, Growth, and Compilation

F.1 Update-Sufficiency Factorization

Let $\widetilde{C}_t : \mathcal{X}_t \to \widetilde{Z}_t$ be a candidate compressed representation, $U_e : \mathcal{X}_t \to \mathcal{X}_{t+1}$ a registered detailed update, and $C_{t+1} : \mathcal{X}_{t+1} \to Z_{t+1}$ the required post-update semantic encoding. We take $\widetilde{Z}_t = \widetilde{C}_t(\mathcal{X}_t)$, so the claimed updater is defined on attained compressed states.

Proposition F.1 (Update-sufficiency factorization criterion)**.** *There is a map $\bar{U}_e : \widetilde{Z}_t \to Z_{t+1}$ satisfying*

$$C_{t+1} \circ U_e = \bar{U}_e \circ \widetilde{C}_t \tag{51}$$

if and only if the required continuation is constant on the fibres of $\widetilde{C}_t$:

$$\widetilde{C}_t(x_1) = \widetilde{C}_t(x_2) \implies C_{t+1}U_e(x_1) = C_{t+1}U_e(x_2) \tag{52}$$

for all $x_1, x_2 \in \mathcal{X}_t$.

Proof. If the updater exists, equal compressed inputs give

$$C_{t+1}U_e(x_1) = \bar{U}_e\widetilde{C}_t(x_1) = \bar{U}_e\widetilde{C}_t(x_2) = C_{t+1}U_e(x_2).$$

Conversely, for $k = \widetilde{C}_t(x)$ define $\bar{U}_e(k) = C_{t+1}U_e(x)$. Fibre constancy makes this independent of the choice of representative x. The map is therefore well defined on the stated image and satisfies Equation (51). □

For contract-relative identity, the equality is taken in the appropriate semantic quotient. This mathematical factorization criterion concerns retained information; admissible implementation and resource requirements are separate obligations. It does not establish a quantitative approximate factorization result.

F.2 Current Prediction Need Not Determine a Future Update

The supplied example has two detailed histories $\mathcal{X}_t = \{x_1, x_2\}$ requiring the same current output. The encoding $\widetilde{C}(x_1) = \widetilde{C}(x_2) = k$ is sufficient for that output. After a registered feedback event, let

$$U_e(x_1) = a, \qquad U_e(x_2) = b, \qquad C^+(a) = k_a, \quad C^+(b) = k_b, \quad k_a \neq k_b.$$

An updater based only on the shared compressed state would have to satisfy both $\bar{U}_e(k) = k_a$ and $\bar{U}_e(k) = k_b$, a contradiction. Current-prediction sufficiency therefore does not imply update sufficiency.

The example uses a candidate prediction compression, not an already sufficient future-operational quotient. If the contract registers the differing update behavior, its full quotient cannot silently merge the two histories. Updating contexts must therefore contribute to the declared future equivalence.

F.3 A Family of Lawful Continuations

Update sufficiency requires a lawful induced updater for every registered event, with the appropriate source and target encodings as in Equation (18). Dynamic sufficiency extends this requirement to all registered evolution classes, including evidence updates, background transport, and internal control transitions. The relevant commuting diagrams must each be declared; a single-event factorization does not supply arbitrary background transports or their composition. The interpretation that knowledge retains both what is known and how it should change describes this information requirement, rather than adding another mathematical assumption.

F.4 Future-closure Viability and Lawful Successor Paths

Let $\mathcal{Y}_\Xi^{\mathrm{fut}}$ be the registered class of future residual demands. A lawful successor closure path for r is a finite sequence $\gamma = (M_0, \dots, M_L)$, starting at $M_0 = M$, whose persistent transitions satisfy the declared semantic, dynamic-sufficiency, resource, and safety requirements, and whose final state closes r to the contract tolerance. Thus the path-based definition of Section 5 is

$$\mathcal{V}_\Xi(M) = \big\{ r \in \mathcal{Y}_\Xi^{\mathrm{fut}} : \Gamma_\Xi(M, r) \neq \varnothing \big\}. \tag{53}$$

Future viability on a declared target class $\mathcal{Y}_0$ requires $\mathcal{Y}_0 \subseteq \mathcal{V}_\Xi(M)$. It is an existence condition for lawful continuation paths, with no claim about unregistered futures.

F.5 Sequential Safe Inverse

Work in the finite-dimensional affine geometry of Section 4. Let $P = P^\top \succeq 0$ denote the current semantic-safe inverse for the declared edit metric and previous responsibilities. Suppose a newly acquired relation is registered as an additional hard local responsibility $Dd = 0$. The remaining safe inverse is

$$P_{\mathrm{after}} = P - PD^\top (DPD^\top)^\dagger DP. \tag{54}$$

The following supplied linear-algebra result permits rank-deficient P and D; the dagger denotes the Moore–Penrose pseudoinverse.

Lemma F.2 (Sequential safe inverse)**.** *Let P be a finite-dimensional symmetric positive-semidefinite matrix, let D be a compatible real matrix, and set $G = DP^{1/2}$. Then*

$$P_{\mathrm{after}} = P^{1/2}\big[I - G^\top (GG^\top)^\dagger G\big] P^{1/2}, \qquad 0 \preceq P_{\mathrm{after}} \preceq P, \tag{55}$$

and

$$\operatorname{im} P_{\mathrm{after}} = \operatorname{im} P \cap \ker D. \tag{56}$$

Proof. The identities $PD^\top = P^{1/2}G^\top$, $DPD^\top = GG^\top$, and $DP = GP^{1/2}$ give the factorization by substitution into Equation (54). The bracketed matrix is the Euclidean orthogonal projector onto $\ker G$. It is positive semidefinite and bounded above by I, so congruence by $P^{1/2}$ gives $0 \preceq P_{\text{after}} \preceq P$. The range of P_{after} consists precisely of the vectors in the previous safe space whose image under D vanishes, giving Equation (56). □

Registering an additional responsibility can therefore consume local safe directions. Successful current closure alone does not establish that a useful response remains available for the next registered demand.

F.6 A Certificate for Registered Local Future Probes

Let B_f be the response Jacobian of a registered local probe family. Its post-update response operator is

$$K_f^{\text{after}} = B_f P_{\text{after}} B_f^\top. \tag{57}$$

The probe family and its construction must respect the declared information rights. The notation B_f neither assumes nor authorizes access to unavailable future-task identities or target labels; it does not encode all unknown future tasks.

Proposition F.3 (Registered local future-plasticity certificate)**.** *Suppose the declared probe operator satisfies*

$$B_f P_{\text{after}} B_f^\top \succeq \kappa_f I, \qquad \kappa_f > 0. \tag{58}$$

Then every nonzero r_f in the registered probe output space has $P_{\text{after}} B_f^\top r_f \neq 0$. In the corresponding frozen affine proximal model, for $\eta > 0$, its residual update satisfies

$$r_f^+ = (I + \eta K_f^{\text{after}})^{-1} r_f, \qquad \|r_f^+\| \leq \frac{1}{1+\eta\kappa_f}\|r_f\|. \tag{59}$$

Proof. For $r_f \neq 0$,

$$r_f^\top B_f P_{\text{after}} B_f^\top r_f \geq \kappa_f \|r_f\|^2 > 0.$$

The compatibility criterion of Theorem 4.1, applied with the additional responsibility, gives a nonzero semantic-safe response. The corresponding proximal residual operator is $(I + \eta K_f^{\text{after}})^{-1}$, as in Equation (16). Every eigenvalue of K_f^{after} is at least κ_f, so

$$\left\|(I + \eta K_f^{\text{after}})^{-1}\right\|_2 \leq (1+\eta\kappa_f)^{-1}.$$

The residual contraction bound follows. □

The certificate concerns the declared linear probe directions. For a nonlinear learner, this calculation is local and does not certify its finite endpoint or a sequence of later updates. It also does not establish $\mathcal{V}_\Xi(M) = \mathcal{Y}_\Xi^{\text{fut}}$: closure paths can require nonlocal changes, evidence, or representation revision absent from B_f. Conversely, a zero local response need not imply global non-viability when a lawful representation revision remains available.

F.7 A Minimal Collapse Example and Continual Admissibility

Example F.4 (Loss of a registered future response)**.** Take a two-dimensional current safe edit space with $P = I_2$, and register $D = [1 \quad 0]$ as a new responsibility. Then $DPD^\top = 1$ and

$$P_{\text{after}} = I_2 - \begin{bmatrix}1\\0\end{bmatrix}[1 \quad 0] = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}. \tag{60}$$

For the registered future probe $B_f = [1 \quad 0]$, this gives

$$B_f P_{\text{after}} B_f^\top = 0. \tag{61}$$

The current responsibility can thus be installed while leaving no same-representation first-order response for that future probe, even though the previous safe requirements remain in force.

This example distinguishes retaining old responsibilities and acquiring the present one from retaining a particular future local response. It does not exclude a nonlocal or representation-changing repair. At the general level, continual admissibility requires the declared future-closure condition $\mathcal{Y}_0 \subseteq \mathcal{V}_\Xi(M^+)$, along with semantic preservation, current closure, and correct updating. Where local probes are registered, Equation (58) provides an additional finite certificate. It does not replace the multi-step viability requirement.

F.8 Path Costs and Capability Comparison

Fix a registered future residual class $\mathcal{Y}_\Xi^{\mathrm{fut}}$. The learner state M includes the persistent variables relevant to future closure; it is distinct from a detailed interaction state x or the represented state $C_M(x)$. Previously acquired duties are registered under Ξ_{old} and retained within the current contract. Every element of $\Gamma_\Xi(M, r)$ must be a finite lawful closure path under the same registered information rights, semantic responsibilities, resource constraints, and closure tolerance. We use finite nonnegative values of C_{full} on these paths. An empty path set has optimal cost $+\infty$. Thus the optimal complete future-closure cost and its finite domain are

$$c_\Xi^\star(r \mid M) = \inf_{\gamma \in \Gamma_\Xi(M,r)} C_{\mathrm{full}}(\gamma), \qquad \mathcal{V}_\Xi(M) = \{r : c_\Xi^\star(r \mid M) < +\infty\}.$$

Unknown costs do not establish feasibility or a dominance certificate. The epigraph in Equation (20) uses finite cost coordinates, so demands with no lawful path contribute no points. It compares infimal costs, not necessarily attained budgets; a statement that a path exists at cost exactly $c_\Xi^\star$ would additionally require attainment.

Write $M' \succeq_\Xi^{\mathrm{sem}} M$ when every previously registered future-operational capability remains lawfully reconstructible from M'. Assume lawful self-reconstruction and composability of these reconstruction relations. For a common contract, define

$$M' \succeq_\Xi M \quad \Longleftrightarrow \quad M' \succeq_\Xi^{\mathrm{sem}} M \text{ and } \mathscr{F}_\Xi(M') \supseteq \mathscr{F}_\Xi(M). \tag{62}$$

The frontier clause preserves both the viable domain and its optimal complete costs. In particular, it requires

$$c_\Xi^\star(r \mid M') \le c_\Xi^\star(r \mid M) \quad \text{for every } r \in \mathcal{V}_\Xi(M),$$

in addition to preservation of registered semantics. Unknown costs cannot be treated as zero or as a certificate of this universal comparison.

Proposition F.5 (Closure-capability preorder)**.** *Under the stated reconstruction assumptions, $\succeq_\Xi$ is reflexive and transitive.*

Proof. Self-reconstruction gives $M \succeq_\Xi^{\mathrm{sem}} M$, and the frontier contains itself. Thus $M \succeq_\Xi M$. If $M_2 \succeq_\Xi M_1$ and $M_1 \succeq_\Xi M_0$, composability gives $M_2 \succeq_\Xi^{\mathrm{sem}} M_0$, while

$$\mathscr{F}_\Xi(M_2) \supseteq \mathscr{F}_\Xi(M_1) \supseteq \mathscr{F}_\Xi(M_0).$$

Together these imply $M_2 \succeq_\Xi M_0$. □

Strict growth. Strict growth, denoted $M' \succ_\Xi M$, requires the weak comparison together with

$$\mathcal{V}_\Xi(M') \supsetneq \mathcal{V}_\Xi(M) \quad \text{or} \quad c_\Xi^\star(r \mid M') < c_\Xi^\star(r \mid M) \text{ for some } r \in \mathcal{V}_\Xi(M).$$

A decrease of the present residual alone does not establish either strict condition. These comparisons are explicitly relative to the declared future class. They make no guarantee about unregistered demands. Parameter count, current accuracy, and the number of stored tasks are not substitutes for these domain or optimal-cost conditions.

F.9 Transport and Proof of T3

If the background changes, we compare a successor with $\widetilde{M}_t = \mathsf{T}_{t \to t+1} M_t$, rather than directly with M_t. Transport must identify the old responsibilities and capability comparisons in the new contract. It is the declared lawful no-learning baseline; environmental transport is not counted as learned growth.

Speculative candidates can fail an audit, regress, or be discarded before commit. Denote the candidate proposed for persistence by $M_t^\star$. The fail-closed rule is

$$M_{t+1} = \begin{cases} M_t^\star, & \mathrm{Audit}_{\Xi_{t+1}}(M_t^\star) = \mathrm{PASS}, \\ & M_t^\star \succeq_{\Xi_{t+1}} \widetilde{M}_t, \\ \widetilde{M}_t, & \text{otherwise.} \end{cases} \tag{63}$$

Audit soundness means that acceptance certifies current closure, preservation of registered semantics, dynamic sufficiency, declared future viability, and the resource and safety contract. The dominance gate is a further assumption, not a consequence of current closure.

Proof of Theorem 5.1. There are two cases. If the candidate fails the audit or the dominance condition, Equation (63) sets $M_{t+1} = \widetilde{M}_t$. By reflexivity, $M_{t+1} \succeq_{\Xi_{t+1}} \widetilde{M}_t$. If the candidate is accepted, the commit condition requires $M_t^\star \succeq_{\Xi_{t+1}} \widetilde{M}_t$, and the rule sets $M_{t+1} = M_t^\star$. Thus the weak comparison holds in both cases. If an accepted candidate additionally expands the lawful closure domain or strictly reduces the optimal cost of a previously closable demand, the definition of $\succ_{\Xi_{t+1}}$ gives the strict comparison. □

This proof concerns certified persistent states. Trial candidates can fail the audit and can be discarded. The fallback branch does not prove that the current residual is closed, and transport-relative non-regression does not establish non-regression across arbitrarily changing environments. The theorem also provides no method for checking optimal-cost dominance over an infinite future class.

Ideal comparison and executable certification. The frontier comparison quantifies over every registered future residual and over optimal lawful path costs. It is an ideal semantic condition, not automatically a terminating audit procedure. A practical commit requires certificates sound for that comparison, or a declared finite sector where the required decisions and costs can actually be evaluated. Finite task or candidate counts alone do not make an unbounded path search or undecidable validation step executable. This appendix does not construct a complete certificate system for the unrestricted comparison.

The baseline and candidate must also be compared at a common accounting point. Candidate search and rejected trials can consume resources even when the persistent model is rolled back. Equation (63) does not restore those resources or prove that the entire search trajectory is cost-free or non-regressive from an earlier accounting origin.

F.10 Conditional Liveness in an Enumerable Sector

Weak non-regression is a safety property. A learner can retain its incumbent forever by rejecting every proposal, so Theorem 5.1 does not imply discovery. Consider a registered candidate family $\mathcal{N}(M, r) = \{N_1, N_2, \ldots\}$ for a persistent residual. For the following result, a completed fair evaluation means that every finite-index candidate is eventually evaluated to completion with its required admissible evidence available. A provisional evaluation before that evidence arrives does not discharge this condition.

Proposition F.6 (Conditional discovery of a lawful closure)**.** *Assume that the residual remains identifiable and future relevant; the required admissible evidence eventually becomes available; a finite-index candidate $N_{j^\star}$ is a lawful closure when evaluated; evaluations are fair in the completed sense just specified; the audit accepts every genuinely lawful candidate in this family; and sufficient resources remain through evaluation of $N_{j^\star}$. Then a lawful closure is eventually found and passes that audit, becoming available for commit subject to the full commit conditions.*

Proof. Fairness gives a completed evaluation of $N_{j^\star}$. The candidate is lawful by assumption, and audit completeness for lawful candidates makes it pass. It is therefore found and available for commit subject to the stated commit conditions. □

Audit soundness, used for safety, and audit completeness on the lawful family, used here for discovery, are separate assumptions. In particular, passing a current-closure audit does not establish the additional dominance gate in Equation (63). That gate and the persistence policy must also be

covered before claiming eventual acceptance under T3. The proposition asserts no strict frontier improvement unless the found candidate additionally has a strict growth witness.

An infinite enumeration with fairness supplies no uniform finite-time or fixed-budget bound. A bounded enumeration claim would require a finite candidate bound, terminating evaluation and certification, appropriate evidence availability, and a corresponding resource bound. No such quantitative bound is derived here. The result also does not decide that no closure exists or imply infinite strict growth. Failure of its assumptions can leave the learner open or inert; it does not automatically produce a verifiable non-closure certificate.

F.11 Compilation, Core, and Full Cost Accounting

Accumulating low-level responsibilities $A_1 d = 0, A_2 d = 0, \ldots$ can reduce the available same-representation directions. GRC proposes compilation as a complementary route to structural expansion. A relation becomes a compilation candidate when it is repeatedly successful, reopenable, dynamically sufficient, and actually reused, and its replacement by a primitive respects the contract:

$$\rho^{\text{certified}} \xrightarrow{\text{Compile}} p_\rho.$$

Here p_ρ is the installed reusable primitive, distinct from an entire quotient state space. When lawful, Core can then factor, merge, prune, retire, or remove and reoptimize redundant realization structure. The intended benefit is to move repeated causal work into a reusable primitive, while preserving the information needed for later reopening and revision.

The two proposed growth mechanisms have different success conditions. Carry supports a growth claim only when a certified successor enlarges the lawful future domain without other regression. Compile supports a growth claim only when a certified successor preserves the frontier and reduces optimal cost for a future demand that reuses the compiled relation. These roles do not by themselves prove either outcome for a specified learner family. Each application must supply a construction or certificate establishing its claimed improvement.

Carry as domain growth. If a registered residual satisfies $c^\star_\Xi(r \mid M) = +\infty$ but $c^\star_\Xi(r \mid M^+) < +\infty$, and the successor also preserves the old semantics and entire frontier, then

$$\mathcal{V}_\Xi(M^+) \supsetneq \mathcal{V}_\Xi(M), \qquad M^+ \succ_\Xi M.$$

This is the domain-growth clause of the definition. Enabling one new demand while damaging existing capabilities does not satisfy it.

Compile as cost growth. Let a recurrent relation be replaced by p_ρ, with registered semantics, dynamic sufficiency, and future viability preserved. To certify growth in the order above, the successor must also preserve every old optimal-cost bound and strictly reduce $c^\star_\Xi(r \mid M)$ for some previously viable registered residual. Mere preservation of feasibility does not establish the cost bounds. Likewise, making one particular reused path cheaper does not by itself show that the infimum over all paths decreases. The path-level benefit is therefore a candidate mechanism whose frontier and optimal-cost conditions still require a certificate.

Core and closure capital. A Core simplification is lawful for this comparison only if $M_{\text{core}} \succeq_\Xi M_{\text{pre-core}}$. Deletion alone is not the defining property. Merge, factorization, reoptimization, pruning, or retirement may simplify an implementation while retaining that comparison; a strict optimal-cost improvement supplies cost growth. Carry and Compile/Core thus address different coordinates of closure capability: the viable domain and its costs. The interpretation of past stable closure as future learning capital is conditional on these certified improvements. Neither the existence of an eligible compilation nor its discovery follows from stable–plastic learning alone; the capitalization property in Appendix I remains an additional requirement.

The cost comparison must fix its accounting origin and resource scope. Search, rejected trials, compilation, maintenance, and reopening must be recorded wherever the declared full-cost contract assigns them. A reduction in continuation cost from the successor is not by itself a reduction of total lifecycle cost from before the search. Rollback of a candidate does not refund resources already consumed. A claim that compilation releases resources or degrees of freedom must be checked under this same accounting and semantic contract.

G Restricted-Sector Correspondences and Comparison Details

G.1 Scope of a restricted-sector correspondence

Let $\mathcal{F}^{\Xi}_{\mathrm{GRC}}$ denote the admissible family of learner states and transition mechanisms under contract Ξ. A restricted method A is contained only when an embedding

$$\iota_A : \mathcal{F}_A \hookrightarrow \mathcal{F}^{\Xi}_{\mathrm{GRC}} \tag{64}$$

preserves the quantities relevant to the comparison: initialization, admissible information, randomness, update behavior, and observable output. Thus representation of an output function is weaker than recovery of the declared learning dynamics. The correspondences below specify restricted choices; they are not a construction of arbitrary solvers from primitive operations.

The proof status is specific to each correspondence. The quadratic optimization calculation below is an exact algebraic recovery under its freezing assumptions. Decision and control correspondences additionally require the declared state sufficiency and transition semantics. The extended map in Section G.9 contains explanatory structural correspondences, not uniformly proved embeddings. None of these categories alone establishes superiority over the methods being compared.

G.2 Fixed-representation quadratic closure

Restrict the persistent parameter state to $M = \theta \in \Theta$, holding the representation, boundary, program, and writable coordinates fixed. At the current state let $g = \nabla_\theta L(\theta)$ and choose

$$\Psi_\theta(d) = \frac{1}{2\eta} d^\top H d, \qquad \eta > 0, \qquad H \succ 0. \tag{65}$$

Proposition G.1 (Quadratic local optimization correspondence)**.** *Under these restrictions, the quadratic local closure problem*

$$d^\star = \arg\min_d \left\{ g^\top d + \frac{1}{2\eta} d^\top H d \right\} \tag{66}$$

has the unique solution $d^\star = -\eta H^{-1} g$.

Proof. The positive-definite quadratic objective is strictly convex. Its first-order condition is $g + \eta^{-1} H d = 0$, which gives $d^\star = -\eta H^{-1} g$. □

The choice $H = I$ recovers a gradient step. Choosing an appropriate positive-definite Fisher metric gives a natural-gradient direction (Boyd & Vandenberghe, 2004; Amari, 1998). A positive-definite local Hessian or a suitable regularized approximation motivates a Newton- or Gauss–Newton-type interpretation; the statement above does not identify arbitrary Hessian choices with a well-posed minimum. Proximal and mirror formulations require their respective objectives, geometries, and regularity assumptions. In particular, the proximal point method is not recovered simply by renaming the quadratic linearization (Rockafellar, 1976). These are additional comparison directions, not further recoveries proved here.

The structural distinction is between choosing a change within a declared coordinate family and deciding whether that family can express the required repair. The quadratic calculation addresses the former. It supplies neither a representation-obstruction test nor a procedure for constructing a new representation.

G.3 Constrained writable coordinates

For a declared local edit set $\mathcal{D}(\theta)$, the corresponding restricted problem is

$$d^\star \in \arg\min_{d \in \mathcal{D}(\theta)} \left\{ g^\top d + \frac{1}{2\eta} d^\top H d \right\}. \tag{67}$$

This specifies constrained optimization in the chosen coordinates when the minimum exists. The quadratic assumptions alone do not supply existence for an arbitrary edit set. Identifying a particular

projected method also requires its projection geometry and update rule. An accepted update is Transformation learning only when it preserves the representation class, $[C_{\theta+d^\star}] = [C_\theta]$, under the fixed contract and complete persistent-state description. Numerical optimization inside this set cannot remove a certified obstruction applying to the whole retained representation family.

G.4 Differentiable families and additive residual blocks

Fix an architecture $\mathcal{P}$, a boundary, and writable parameters $\theta \in \Theta$. The resulting family is

$$\mathcal{F}_{\mathcal{P}} = \{M_\theta : \theta \in \Theta\}. \tag{68}$$

Changes in θ may alter both relations and latent representations $z_\theta(x)$. A restriction to this family therefore fixes the outer machine grammar, not the learned feature values. Its interpretation as a GRC sector is conditional on preserving the declared initialization, execution, and learning updates within the admissible family under Ξ. The inclusion $\mathcal{F}_{\mathcal{P}} \subseteq \mathcal{F}^{\Xi}_{\mathrm{GRC}}$ restricts the candidate family; recovering a particular training method further requires the embedding to intertwine its declared updates. A fixed architecture alone does not imply a fixed quotient K_{M_θ} or make all parameter learning Transformation.

The identity-shortcut additive core

$$h_{\ell+1} = h_\ell + F_\ell(h_\ell; \theta_\ell) \tag{69}$$

prescribes the block interface, reference input, and form of the correction (He et al., 2016). Here a fixed owner means a specified interface at which the correction is applied; it does not mean a constant hidden state or an absence of feature learning. The displayed expression concerns the forward block. A complete ResNet realization also includes its actual shortcut choices, nonlinearities, and training state. Consequently this expression alone is not an embedding of the complete training dynamics.

If a required distinction cannot be expressed anywhere in the declared family, further optimization confined to that family cannot supply it. Revising the representation, owner structure, memory, or program is then a candidate completion. This observation does not assert that every residual reveals such an obstruction, or that a legal revision is always available.

G.5 Fixed-state sequential decision mechanisms

Suppose a declared control abstraction $s = C(x)$ is sufficient for the registered control future, with specified actions, transition semantics, rewards or costs, and control boundary; here x may be an interaction history. The Bellman consistency relation and its residual are

$$V = \mathcal{T}V, \qquad r_{\mathrm{Bellman}} = \mathcal{T}V - V. \tag{70}$$

This residual concerns a relation within an already declared decision model (Bellman, 1952). Value, policy, and controller changes on that fixed state space provide the restricted correspondence. Recovering any particular learning algorithm additionally requires its sampling, estimation, and update semantics; the fixed-point equation alone does not provide them. Value, policy, or controller changes count as Transformation when their persistent transitions preserve $[C_M]$ under the fixed contract and complete learner-state description. A fixed state-vector interface by itself does not establish that quotient condition.

The representation issue arises for a proposed abstraction C when

$$C(x_1) = C(x_2), \qquad \mathrm{Future}_\Xi(x_1) \not\simeq_\Xi \mathrm{Future}_\Xi(x_2). \tag{71}$$

If the contract requires the distinguishing future quantity to be represented, that distinction cannot be expressed as a function of the shared state alone. This is an obstruction of the stated abstraction, consistent with the role of future-relevant state equivalence in predictive representations and MDP reduction (Littman et al., 2001; Givan et al., 2003). Different continuations need not imply different values or optimal actions for every objective, so the claim concerns the quantity actually registered in Ξ.

Proposition G.2 (State-aliasing obstruction)**.** *If two detailed states have the same represented state but require inequivalent registered continuations, no relation law receiving only that represented state can represent both required continuations.*

Proof. Suppose such a law ρ existed. Writing $C(x_1) = C(x_2) = s$, it would satisfy

$$\rho(C(x_1)) = \rho(s) = \rho(C(x_2)).$$

The contract instead requires inequivalent future continuations for x_1 and x_2, a contradiction. □

This is an obstruction of the proposed abstraction: it means that the claimed encoder fails the required sufficiency contract. It does not contradict the definition of a quotient already known to preserve those same futures. Optimizing the relation more accurately while retaining the aliased input cannot supply the missing distinction.

State splitting, memory, belief states, recurrent states, or a revised model state are possible responses, subject to the same admissibility contract. For constrained control, established stability and feasibility assumptions remain necessary to the claimed correspondence (Mayne et al., 2000). No separate HJB, MPC, or complete RL algorithm embedding is derived from the Bellman residual here.

G.6 Completion and preservation of a lawful incumbent

Write

$$A^+ = \mathrm{GRCComplete}_{\Xi}(A), \qquad \mathcal{F}_A \subseteq \mathcal{F}_{A^+}, \tag{72}$$

for a completion that preserves the lawful realizations of A while reopening selected representation, state, boundary, owner, program, geometry, or update-rule choices. These are permitted variables, not a requirement to revise all of them at every step. If a comparison changes the learner–world boundary or contract, its information rights, transport, and baseline must be restated. Such a change is not silently covered by the fixed-boundary internal two-mode claim.

Let a_{inc} be the incumbent mechanism. Only candidates meeting the current residual, old-semantic, dynamic-sufficiency, future-viability, resource, and safety requirements belong to $\mathcal{A}_t^{\mathrm{legal}}$. The selection specification is

$$a^\star \in \operatorname*{arg\,min}_{a \in \mathcal{A}_t^{\mathrm{legal}}} C_{\mathrm{full}}(a). \tag{73}$$

The incumbent is included whenever it remains lawful. A merely available incumbent is not automatically a feasible fallback if it fails the current contract. The displayed specification also does not establish that a minimizer exists or that an executable search returns one.

G.7 Feasible-set comparison and a strict-improvement witness

Let J_Ξ be a common complete objective. Define

$$J_A^\star = \inf_{M \in \mathcal{F}_A} J_\Xi(M), \qquad J_{A^+}^\star = \inf_{M \in \mathcal{F}_{A^+}} J_\Xi(M). \tag{74}$$

Proposition G.3 (Incumbent-preserving feasible-set comparison)**.** *If $\mathcal{F}_A \subseteq \mathcal{F}_{A^+}$ under the same information rights, constraints, objective, and complete cost accounting, then*

$$J_{A^+}^\star \le J_A^\star. \tag{75}$$

Proof. Every realization feasible for A is feasible for A^+. The infimum over a superset cannot exceed the infimum over its subset. □

Corollary G.4 (Conditional strict extension)**.** *Under the same comparison conditions, if there exists*

$$M' \in \mathcal{F}_{A^+} \setminus \mathcal{F}_A, \qquad J_\Xi(M') < J_A^\star, \tag{76}$$

then $J_{A^+}^\star < J_A^\star$.

The witness condition is not supplied by set enlargement alone. Operational costs include search, audit, compilation, maintenance, and deployment. Inclusion must therefore concern the feasible realizations of the actual comparison, after its resource restrictions and costs are specified. This idealized comparison of infima is not a guarantee that a particular numerical solver or structural search attains either value.

G.8 Containment and discovery are distinct claims

Containment recovers a declared method. Rediscovery would require a learner not given that method to construct an operationally equivalent mechanism from primitive operations and feedback. Retaining the ability to revise a useful mechanism when it later becomes insufficient is a further property. The restricted correspondences and feasible-set comparison do not solve these search and computability questions. A recurrent successful mechanism can be considered for compilation only under the dynamic-sufficiency and cost conditions imposed elsewhere in the framework.

G.9 Extended Structural Correspondences

The following correspondences identify a mechanism's predeclared choices and the choices a GRC completion could reopen. Except for the stated read allocation calculation, they are structural interpretations requiring further specifications for an exact algorithmic embedding. Changing an implementation component is classified by its effect on $[C_M]$ and the relation law, not by its architectural name. The two-mode classification applies to persistent internal transitions under a fixed contract and boundary, with the completeness assumption A3.

Attention as content-addressed opening. The following calculation gives a restricted correspondence with attention (Vaswani et al., 2017). Consider read allocations $a \in \Delta_m$ over a fixed candidate set, with finite costs c_j, a full-support prior allocation $\omega \in \Delta_m$, and temperature $\tau > 0$. The source considers

$$\min_{a \in \Delta_m} \sum_{j=1}^{m} a_j c_j + \tau D_{\mathrm{KL}}(a \| \omega). \tag{77}$$

Its Lagrangian first-order condition is

$$c_j + \tau \left(\log \frac{a_j}{\omega_j} + 1 \right) + \lambda = 0,$$

giving the normalized allocation

$$a_j^\star = \frac{\omega_j \exp(-c_j/\tau)}{\sum_{i=1}^{m} \omega_i \exp(-c_i/\tau)}. \tag{78}$$

For a numerical query $u \in \mathbb{R}^{d_k}$ and numerical keys $z_j \in \mathbb{R}^{d_k}$, choosing $c_j = -u^\top z_j / \sqrt{d_k}$ gives soft content-addressed read weights. The numerical keys are distinct from the quotient states $k = C_M x$; their representation must be declared. This recovers the allocation rule for the chosen objective, not every value-combination, positional, normalization, or training operation of a Transformer. The interpretation as Open also requires those reads to expose the declared relational realizations. A completion may reconsider the candidate owners, read representation, or read grammar.

Recurrent state, state-space models, and filtering. If two detailed histories share a candidate compression $\widetilde{C}(x)$ but require different represented successors under the same event, that compression is not update sufficient. A persistent recurrence

$$s_{t+1} = \Psi(s_t, o_{t+1}) \tag{79}$$

can serve as a remedy when it actually carries the missing distinction and supports the required updates. Recurrence alone does not establish that property. RNN, state-space, or filtering mechanisms provide possible realizations inside a declared recurrent-state grammar; a completion may split, merge, or revise that grammar. Neither recurrence's necessity in every task nor sufficiency of every recurrent model is asserted.

Sparse experts and ownership. When a residual can lawfully be assigned to a small subset of existing owners, a sparse expert mechanism provides a possible realization using selected local relation laws. This is the source's structural analogy with mixtures of experts under a fixed expert/router vocabulary (Fedus et al., 2022). It is not a unique derivation of an MoE architecture or proof that an arbitrary routing rule performs the required diagnosis. A completion may consider owner creation, merging, retirement, and certified Compile/Core operations.

Low-rank writable families. A low-rank matrix update restricts edits to

$$\Delta W = UV, \qquad U \in \mathbb{R}^{m\times r}, \quad V \in \mathbb{R}^{r\times n}, \tag{80}$$

for a prescribed factor size r. This defines a constrained writable family. Calling it a same-representation fibre additionally requires preservation of $[C_M]$; low rank by itself does not supply that condition. The correspondence concerns this factorized update form, not all adapter designs. Exhaustion of a numerical search or edit budget does not by itself certify a representation obstruction or justify automatic rank expansion. Evidence, transport, search, and genuine expressivity limitations must be distinguished before claiming that representation change is necessary.

Retrieval and tools. A query, retrieval, sensing action, or tool call may change available evidence without changing persistent learner state. In that case it is an external resolution action, not a third primitive internal learning mode. Missing evidence can explain an apparent inability to close a residual within the current internal response family. If the evidence is subsequently incorporated persistently, that incorporation is subject to the internal two-mode classification. A persistent revision of the retrieval or tool mechanism must likewise be classified by its actual effect on the representation and relation law; revising an interface does not alone prove that the quotient changed.

Predeclared structural expansion. Width, node, module, or feature expansion rules specify candidate structural edits in advance. A GRC interpretation can consider such candidates when diagnosis supports them, while keeping a representation-obstruction certificate distinct from a failed optimization attempt. A claim that representation revision is necessary requires that certificate; an architectural expansion is not automatically a change of quotient. No unique expansion algorithm or general tractability result follows.

Architecture search and program synthesis. Searching over architectures or programs can reopen structural choices inside a fixed search grammar. It is representation search when those choices change $[C_M]$; some implementation changes may instead preserve that abstraction. A completion can place the grammar, solver, or compiler among the persistent objects available for revision. This extends the candidate description, without establishing that unrestricted search or grammar revision is computationally tractable.

Meta-learning and learned update rules. An initialization, adaptation rule, optimizer, or other meta-state can be included in M. A meta-learning method then operates within its declared meta-family. The recursive question is whether that family remains sufficient: an updater can itself become an object of learning. This interpretation does not identify every change to an updater with Representation learning or supply a method for discovering a better one.

Probabilistic and variational representations. A probabilistic learner can fix latent variables, a generative graph, a probability family, and a variational family, then perform inference or learning within that ontology. Such a method occupies a restricted sector when a semantics-preserving embedding is supplied. A completion may reconsider the latent variables, factorization, or representation language. This is a conditional correspondence; it is not a complete embedding of the Free Energy Principle or active inference (Friston, 2010).

Table 2: Status of the sector map. Reopened choices are candidate freedoms, not guarantees of discovery, strict improvement, or lower complete cost.

Mechanism	Correspondence and status	Choices that may reopen
Quadratic optimization	Exact stated quadratic-step algebra with frozen representation	Edit geometry, representation
Deep networks	Conditional containment of a declared differentiable family	Machine/program family
Residual blocks	Additive forward-map structure	Correction owner and form
Attention	Exact allocation for the stated entropic read objective	Read grammar and candidates
RNN/SSM/filtering	Conditional realization of update-sufficient state	Recurrent-state representation
Sparse experts	Structural ownership interpretation	Owner creation, merge, retirement
Low-rank updates	Constrained edits; Transformation only if $[C_M]$ is preserved	Rank and writable family
Retrieval/tools	External resolution when persistent state is unchanged	Evidence access and mechanism
NAS/synthesis	Structural search in a declared grammar	Grammar, solver, compiler
Meta-learning	Conditional learned-updater sector	Updater/meta-family
Bayes/VI/FEP	Conditional probabilistic representation sector	Latent/model ontology

H Counterexamples and Impossibility Regimes

These supplied examples separate current prediction, lawful acquisition, future response, representation adequacy, and the scope of the declared contract. All comparisons retain the assumptions specified below.

H.1 Prediction Sufficiency Does Not Imply Update Sufficiency

Let $\mathcal{X} = \{x_1, x_2\}$ and suppose both histories require the same current prediction. A compression with $\widetilde{C}(x_1) = \widetilde{C}(x_2)$ is sufficient for that prediction. If the same feedback event requires

$$C^+U_e(x_1) = a, \qquad C^+U_e(x_2) = b, \qquad a \neq b,$$

no map $\bar{U}_e$ on the shared compressed state can satisfy $C^+U_e = \bar{U}_e\widetilde{C}$. This is the two-state obstruction of Appendix F.2; the general factorization criterion is Proposition F.1. The candidate compression is sufficient for the current prediction, not for the full registered future-update contract.

H.2 Loss Reduction Does Not Imply Lawful Learning

Take a scalar learner state $\theta \in \mathbb{R}$, an old registered responsibility $S_{\mathrm{old}}(\theta) = \theta = 0$, and a current loss $L_{\mathrm{new}}(\theta) = (\theta - 1)^2$. Starting at $\theta_0 = 0$, the edit $\theta_1 = 1/2$ gives

$$L_{\mathrm{new}}(\theta_1) = \tfrac{1}{4} < 1 = L_{\mathrm{new}}(\theta_0), \qquad S_{\mathrm{old}}(\theta_1) = \tfrac{1}{2} \neq 0.$$

The current loss improves while the old responsibility fails. Loss reduction by itself therefore certifies neither semantic stability nor the legality of a persistent transition.

H.3 Present Acquisition Can Exhaust a Future Local Response

As in Example F.4, let $P = I_2$ and register the new hard responsibility $D = [1 \quad 0]$. Then

$$P_{\mathrm{after}} = P - PD^\top(DPD^\top)^\dagger DP = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

For $B_f = [1 \quad 0]$, $B_f P_{\mathrm{after}} B_f^\top = 0$. Previous responsibilities and successful current acquisition need not leave a nonzero same-representation first-order response for this future probe. This example does not rule out a lawful nonlocal or representation-changing continuation, and is therefore distinct from an impossibility result for all future closure paths.

H.4 Excessive Closure Load Need Not Diagnose Representation Failure

Suppose the correct relation depends on an environmental bit $z \in \{0, 1\}$ and the existing relation family can already represent both branches $\rho(s, z)$. If the current evidence does not reveal z, an internal edit based only on that evidence cannot identify which branch to install. A registered diagnostic that treats this unresolved evidence requirement as infeasible may report $q_T(r) > 1$.

An admissible external query can reveal z, after which the same representation selects an already expressible relation. Thus an overflow report alone does not establish representation insufficiency. The example concerns an evidence-limited diagnostic; it does not assign a numerical value to the geometric load without its required calibration. Evidence diagnosis must precede an inference that Carry is necessary, as specified in Appendix E.

H.5 A Trivial Safe Space Eliminates Local Plasticity

If $\ker A = \{0\}$, then the only semantic-safe edit is $d = 0$. Lemma D.1 gives

$$\operatorname{im} P_H = \ker A = \{0\}, \qquad P_H = 0, \qquad K_{\mathrm{plast}} = BP_HB^\top = 0.$$

Hence $P_H B^\top r = 0$ for every residual: there is no nonzero safe first-order update in the declared edit space. This is an impossibility regime of the compatibility theorem. Remaining open, obtaining admissible evidence, revising the contract, or undertaking a justified representation revision are different possibilities, none guaranteed to restore feasibility. A contract revision also changes the comparison being made.

H.6 Protecting Historical Realization Can Remove Useful Plasticity

Take $H = I$ on $\mathbb{R}^2$. Suppose exact historical preservation allows only $T_{\text{exact}} = \text{span}(e_1)$, whereas the same registered future semantics permits $T_{\text{rec}} = \text{span}(e_1, e_2)$. For $B = [0 \quad 1]$, the corresponding response operators are

$$K_{\text{plast}}^{\text{exact}} = 0, \qquad K_{\text{plast}}^{\text{rec}} = 1.$$

Exact realization preservation admits no safe response in this direction; semantic preservation admits one. This supplied example realizes the strict-coupling case of Appendix D.8. The extra rigidity comes from protecting implementation details absent from the declared semantic contract.

H.7 The Two-Mode Classification Fixes the Learner–World Boundary

Suppose two environmental states cannot be distinguished because a sensor channel z is unavailable. Adding an external sensor that reveals z changes the boundary ∂. Relative to the original learner-state description, this event is outside the fixed-boundary domain of the internal two-mode classification.

One may instead include a persistent, learner-controlled sensor interface in an enlarged state description. In that enlarged description the classification applies under its stated completeness assumptions: a quotient-changing revision is Representation Learning, and a change within the same quotient is Transformation Learning. Merely naming an interface change does not determine its mode. The boundary and state description must be fixed before applying the classification.

H.8 Semantic Identity Is Contract-Relative

Let x and y have identical responses under every context in Ξ_1, so $x \sim_{\Xi_1, M} y$. Enlarge the contract to Ξ_2 by adding a query $c^\star$ such that

$$\text{Resp}_M(c^\star, x) \not\simeq \text{Resp}_M(c^\star, y).$$

Then $x \not\sim_{\Xi_2, M} y$. The semantic quotient is relative to the declared future-context class. If that class is unidentified, the intended quotient is likewise unidentified; selecting an arbitrary equivalence relation cannot certify sufficiency for it.

These examples identify separate limitations. None licenses replacing an information obstruction with an optimization claim, or extending a local response calculation into an unrestricted guarantee of continual learning.

I Conditional Representation: Assumptions and Construction

I.1 Abstract learner and operational assumptions

Fix a contract Ξ and a learner–world boundary. An abstract learner is described by

$$\mathcal{L} = (\mathcal{X}, \mathcal{M}, \mathcal{E}, \text{Resp}, U, \text{Acc}, C_{\text{full}}).$$

Here $\mathcal{X}$ contains detailed interaction states or histories, $\mathcal{M}$ contains persistent learner states, and $\mathcal{E}$ is the registered event family. The response $\text{Resp}_M(c, x)$ is registered for context $c \in \text{Ctx}_\Xi$ and detailed state x. An event continuation U_e acts on the detailed state, whereas a persistent learning transition changes M to M^+. Their domains and any event-specific evaluation or transport are declared separately, as in Appendix B. The learner's own acceptance specification determines $\text{Acc}_\Xi(M, r)$; $C_{\text{full}}(N \mid M)$ is the complete declared transition cost. No neural, parametric, probabilistic, or differentiable realization is imposed.

The assumptions of Theorem 7.1 concern the process being represented. They are not consequences of stability and plasticity alone.

R1: Persistent-state completeness. Every persistent learner-internal variable affecting registered future responses or subsequent accepted updates belongs to M.

R2: Extensional future distinguishability. For each fixed M, the registered response identity in Equation (22) induces a well-defined equivalence relation $\sim_{\Xi,M}$ on $\mathcal{X}$. Distinctions irrelevant to every registered future context may be collapsed; distinctions that affect such a context may not be silently identified. An approximate discrepancy does not automatically define a transitive equivalence.

R3: Conditional dynamic sufficiency. For events claimed to operate adequately within the current abstraction, equivalent detailed states have equivalent represented continuations. The declared continuation is constant on the relevant quotient fibres up to contract equivalence and therefore admits descent on that domain. Failure caused by a merged future-relevant distinction may instead trigger representation revision under R5. This exception does not construct a missing continuation map.

R4: Nontrivial discrepancy response. If an identifiable future-relevant discrepancy is nonzero and a lawful corrective successor exists, the learner admits a nontrivial corrective candidate. A candidate must still pass the persistence requirements below. This assumption supplies neither a search-time bound nor a guarantee of finding a fully resolving successor.

R5: Representation revisability. If no admissible relation law on the current abstraction can satisfy the registered contract, refinement or replacement of the representation is an admissible candidate operation. Permission to consider that operation does not establish that a successful replacement is constructible.

R6: Operational future viability. Accepted transitions preserve the declared future adaptive possibilities: the registered future demands remain supported by lawful successor paths.

R7: Fail-closed persistence. Speculative candidates need not already satisfy all persistence conditions. Only candidates satisfying the declared semantic, dynamic, viability, information, resource, and safety requirements may become persistent states.

R8: Minimal admissible selection. For the accepted resolving transitions under consideration, the learner selects a complete-cost minimizer among its accepted persistent successors that resolve the discrepancy to tolerance. This requires a selected minimizer where the rule is invoked. An empty resolving set or an unattained infimum does not provide a successor. No deterministic tie-breaking rule or unique minimizer is assumed.

I.2 Future-operational quotient and relation descent

R2 gives the fixed-M quotient and its class map,

$$K_M = \mathcal{X}/\!\sim_{\Xi,M}, \qquad C_M : \mathcal{X} \longrightarrow K_M.$$

The quotient is determined up to admissible isomorphism by the registered future responses. Its elements are classes of detailed states or histories; it is not a quotient of entire machines. No historical parameter identity enters this construction.

For an event on the applicability domain of R3, let C_M^e evaluate the detailed event successor in the declared comparison space $Z_{M,e}$. The superscript denotes event-specific evaluation or transport, not a newly accepted learner. Dynamic sufficiency gives

$$C_M x = C_M y \quad \Longrightarrow \quad C_M^e(U_e x) \simeq_\Xi C_M^e(U_e y).$$

Hence there is a descended continuation, in the declared response type,

$$\bar{U}_{M,e} : K_M \rightsquigarrow Z_{M,e}, \qquad \bar{U}_{M,e}(C_M x) \simeq_\Xi C_M^e(U_e x). \tag{81}$$

The represented result is independent of the representative x, up to the registered equivalence. For a fixed unchanged evaluation domain, $Z_{M,e} = K_M$ and $C_M^e = C_M$ recover the same-quotient case.

Collect the descended laws and the remaining persistent operational degrees of freedom into the relation/realization description ρ_M. In this construction that description retains what is needed to reproduce the learner's registered operations, giving the pair (C_M, ρ_M). This use of a complete description must be distinguished from restricting ρ_M to a predetermined small grammar or parametric family. R1 ensures that the variables belong to M; it does not prove that such a restricted relation family can express all of them. Identifying the complete description with a chosen Ring requires the coverage condition stated in Assumption A3 in Appendix A.

I.3 Prediction and the generalized residual

The current relation law imposes a predicted represented continuation,

$$\widehat{k}^{+} = \mathrm{Pred}_{M,e}(\rho_M, k) = \widehat{U}_{M,e}(k), \qquad k = C_M x,$$

where the last notation is used when the selected law defines a state-level map. Prediction need not be probabilistic or differentiable. It must have the same declared comparison space as the realized value $k^{+}_{\mathrm{real}} = C^{e}_{M}(U_e x)$. Where both values are defined, set

$$\mathfrak{R}_{\Xi,e}(M, x) = \mathrm{Def}_{\Xi}\big(C^{e}_{M}(U_e x), \widehat{U}_{M,e}(C_M x)\big). \tag{82}$$

This includes the fixed-domain formula with $C^{e}_{M} = C_M$.

The prediction $\widehat{U}_{M,e}$ may disagree with reality; it is not the exactly descended continuation $\bar{U}_{M,e}$. Comparing reality with an already exact induced continuation would give zero commutation defect, rather than a potentially nonzero prediction residual. Dynamic sufficiency specifies when a represented continuation is well defined; it does not assert that the learner's current prediction is correct.

If no lawful continuation descends because a needed distinction has been merged, record that failure as a representation obstruction. An undefined evaluation is not assigned a numerical residual or interpreted as zero. Thus a relation defect on an adequate representation and a failure of representation sufficiency are distinct diagnoses. R4 applies to an identifiable relevant defect when a lawful corrective successor exists; the construction supplies no general obstruction-detection procedure.

I.4 Representation classes and persistent transitions

Representations C_M and C_N have the same abstraction when an admissible isomorphism satisfies

$$C_N = \phi \circ C_M, \qquad \phi : K_M \longrightarrow K_N. \tag{83}$$

Write $C_N \cong C_M$ and $\varpi_{\Xi}(M) = [C_M]$. The brackets denote the encoder's class up to such relabeling, not a machine-equivalence class. Learners can differ internally and in their relation laws while inducing the same partition of detailed states.

Every accepted persistent transition either preserves $[C_M]$ or changes it. When it preserves the class, its remaining operational changes are recorded in the complete relation/realization component: this is Transformation. When it changes the class, it revises the represented state distinctions: this is Representation learning. Relation changes may accompany that revision, giving a joint event rather than a third primitive class. A no-op requires no nonzero learning component. This endpoint classification is distinct from the connection-dependent local splitting of Theorem 3.1.

If the current abstraction supports no lawful relation satisfying the contract, R5 permits a candidate N with $[C_N] \neq [C_M]$. This supplies the Representation/Carry option within the description; it guarantees neither a successful candidate nor completion of the search.

I.5 Inherited legality, selection, and operational equivalence

On the resolving domain of R8, let $\mathrm{Acc}_{\Xi}(M, r)$ denote the learner's own accepted successor family, including closure to the declared tolerance. Define

$$\mathrm{Law}^{\mathcal{L}}_{\Xi}(M, r) := \mathrm{Acc}_{\Xi}(M, r). \tag{84}$$

This is the family denoted by $\mathrm{Law}_{\Xi}(M, r)$ in Theorem 7.1. Its preservation, dynamic sufficiency, future viability, information rights, resource, and safety requirements are inherited from the learner's registered acceptance semantics. No additional GRC legality filter is inserted to alter that family. R6–R7 constrain accepted successors; they do not imply that every conceivable lawful candidate is accepted or discovered.

The least-lawful selection correspondence is

$$\mathrm{LeastLawfulClose}_{\Xi}(M, r) := \operatorname*{arg\,min}_{N \in \mathrm{Law}^{\mathcal{L}}_{\Xi}(M,r)} C_{\mathrm{full}}(N \mid M).$$

R8 places the learner's selected successor in this correspondence, giving Equation (23). The representation retains the learner's actual selection among minimizers; the correspondence does not license additional observed behavior by replacing that selection with every minimizer. Minimum cost is an assumption about these transitions, not a consequence of R1–R7.

Two descriptions are operationally equivalent under Ξ when, for each registered initial state, context, and event sequence on the declared applicability domain, they agree up to contract equivalence on registered responses, accepted successor semantics, information rights, and transition legality. The constructed description retains the original detailed continuation, response equivalence, acceptance semantics, and selection. Thus it changes their representation, not those operational observables. When descent is unavailable, it records the obstruction instead of claiming a complete quotient-level continuation for that event.

I.6 Proof of Theorem 7.1

Proof. 1. Quotient existence. R2 makes $\sim_{\Xi,M}$ an equivalence relation, yielding K_M and C_M.

2. Relation descent. On events for which the current representation is adequate, R3 makes the represented continuation constant on quotient fibres. This gives Equation (81). The remaining persistent operational degrees of freedom are retained in ρ_M.

3. Residual construction. The represented realized continuation and the current law's prediction give Equation (82). Failure of descent is recorded as a representation obstruction, within the stated typing and applicability restrictions.

4. Nontrivial response. R4 supplies a nontrivial corrective candidate for an identifiable nonzero relevant defect whenever a lawful corrective successor exists.

5. Transition classification. An accepted successor either preserves $[C_M]$ or changes it. Retaining the other persistent operational degrees of freedom in ρ_M yields the Transformation case in the former event and a Representation revision, possibly accompanied by relation changes, in the latter.

6. Revisability. R5 permits representation revision when no relation law on the current abstraction satisfies the contract. This step provides an admissible option, not a constructive success guarantee.

7. Persistent legality. R6–R7 impose the learner's declared future and persistence requirements. Equation (84) retains that learner's accepted resolving family.

8. Minimal closure. R8 places the selected successor among the complete-cost minimizers of that family, giving Equation (23). The construction preserves the declared responses and accepted transition semantics, establishing the stated conditional operational representation. □

This argument constructs descriptive objects from the stipulated process. It does not infer relation descent beyond its applicability domain, an effective encoding of the full quotient, or a search procedure that realizes the assumed selection.

I.7 Additional growth, compilation, and Core properties

The stronger growth description additionally assumes the capability comparison of Section 5 and Appendix F. Relative to the declared background transport T, non-regression requires $M^+ \succeq_\Xi \mathsf{T}M$; a transition designated as strict Growth also requires $M^+ \succ_\Xi \mathsf{T}M$. The representation preserves these properties when imposed. It does not derive them from R1–R8 or imply that strictly improving transitions are always available.

R9: Recurrent capitalization. Theorem 7.1 does not imply compilation. A stable, plastic learner could keep repeating an expensive successful relation. The additional R9 property concerns a relation ρ that recurs across future closure paths, has been repeatedly certified, and is reopenable from a persistent primitive. Its compilation must preserve registered semantics, dynamic sufficiency, and the future-viability frontier while strictly reducing the declared complete future closure cost. When these conditions hold, R9 permits a persistent replacement

$$\rho^{\text{recurrent,certified}} \xrightarrow{\text{Compile}} p_\rho.$$

The primitive p_ρ is neither the quotient space K_M nor a prediction of the next world state. Reopening makes the stored primitive operational; compilation makes certified recurrent activity reusable. These operations have complementary temporal roles and are not asserted to be algebraic inverses. Together they permit a recursive semiclosed description when the stated opportunities occur, without guaranteeing their discovery or recurrence. A compilation transition counts as Growth only under the declared capability comparison and complete-cost accounting.

After a successful compilation, a proposed Core operation changes M_{comp} to M_{core} only under the stated comparison

$$M_{\mathrm{core}} \succeq_{\Xi} M_{\mathrm{comp}}.$$

Strict inequality would make this Core transition a further strict gain. Merge, factorization, pruning, reoptimization, and retirement are possible realization changes, not mechanisms required or constructed by the representation theorem. Neither R9 nor the displayed acceptance condition ensures that a qualifying Core operation can be found.

I.8 Scope

The result is an existence-of-operational-representation statement for a stipulated class of processes under a fixed contract and boundary. It does not show that every learner satisfies R1–R8, identify a unique implementation, or establish efficient discovery, affordable representation revision, strict growth on every event, or empirical superiority. Recursive compilation requires the independent R9 eligibility and admission property; Core retains its own comparison condition.

The quotient's isomorphism class does not determine a unique solver, metric, cost model, or learning trajectory. Retaining complete persistent operational information in a descriptive relation layer does not prove that a restricted executable Ring is complete. A different contract can change the quotient, lawful successor family, and capability order. The conditional structural description therefore supplies no computational universality or algorithmic completeness beyond its stated assumptions.

J Operational Intelligence, Recursive Learning, and Falsifiability

J.1 A Profile of Future Capability and Cost

The proposed intelligence interpretation reuses the objects of Section 5. Its capability profile combines a registered future domain with the optimal complete costs of lawful closure. The contract fixes which distinctions, information rights, semantic responsibilities, resources, and interventions matter. Consequently, more parameters or better performance on one fixed benchmark need not establish dominance of the whole profile. The interpretation is that intelligence can turn unresolved, future-relevant distinctions into reusable capacity while preserving the ability to learn again. It does not posit a unique metaphysical definition or collapse that profile into an unsupported universal score.

We retain the optimal-cost epigraph already defined in Equation (20). Its boundary may represent an unattained infimum. A claim of closure within an exact budget must separately provide an admissible path at that budget; the epigraph alone does not provide one.

J.2 Temporary, Persistent, and Compiled Outcomes

Reasoning is temporary closure search over relations, trajectories, decompositions, or programs. It need not change persistent knowledge. Learning makes a residual-driven lawful transition persistent, subject to the distinctions between learning and effective learning in Section 5. Compilation gives repeatedly useful closures a reusable realization. The operational role of such memory is to preserve conditions for reopening a capability and continuing its revision, rather than merely storing a historical snapshot. These are different outcomes: a successful trial is not automatically a persistent commit, and a commit is not automatically a mature compiled primitive.

Transformation revises relations within a representation; representation revision changes which distinctions the learner can express. A certified relation may subsequently become a primitive supporting further relations. Thus learning can construct and repair the state space of subsequent learning.

A successor need not change the quotient state space: Transformation can preserve it up to isomorphism while changing its relation law. This account is structural: it does not assert that every stored object is compiled or that every recurrence creates a new residual. When continued operation respects the registered obligations, the next residual can remain zero.

J.3 Failure Tests and Their Scope

Formal counterexamples must satisfy the stated hypotheses. For the two-mode classification, the proposed test is a persistent internal transition within the fixed boundary and complete state description that admits neither a relation change nor a representation change nor their combination. For the local affine compatibility claim, the proposed test has $P_H B^\top r \neq 0$ but no semantic-safe direction producing positive first-order residual reduction. For reconstructive dominance, the response operators must be compared under the same metric and residual map; a violation despite the stated nested safe spaces would contradict that claim.

The other checks concern the declared certificates and operational representation. A successor worse than its transported baseline requires examining audit soundness and the dominance gate. A counterexample to the conditional representation result must satisfy its assumptions while lacking the claimed operational representation. These are proof and implementation obligations, not additional theorems established by this discussion.

Empirical failures can also expose incorrect contracts, unsound audits, missing evidence, or excessive complete costs. Such failures need to be distinguished from counterexamples satisfying a theorem's hypotheses. Neither preservation nor containment alone supplies discovery, affordable search, or automatic compilation. The substantive question is whether a specified learner can certify and reuse closures so that acquired stability supports its next lawful adaptation.

K Related Work and Positioning

Predictive state and behavioral abstraction. Predictive state representations use action-conditioned future tests and provide recursive state updates (Littman et al., 2001). Computational mechanics constructs causal states from predictive equivalence (Shalizi & Crutchfield, 2001); MDP equivalence supports behavior-preserving state reduction (Givan et al., 2003). Consequently, a future-operational quotient, minimal predictive state, or commuting update equation alone cannot establish GRC's novelty. Here the registered operations include changes to the learner's persistent knowledge and update state. A candidate encoding must be checked against those operations and the declared background. The distinction between a current prediction and sufficient information for a subsequent learning update must not be confused with claiming that predictive-state methods lack dynamics.

Continual learning and retained plasticity. EWC penalizes changes to important parameters (Kirkpatrick et al., 2017), whereas OGD uses gradient projections to reduce interference with previous predictions (Farajtabar et al., 2020). These are precedents for seeking useful updates under preservation requirements. Loss of plasticity is also an established problem, with continual backpropagation evaluated as a means of sustaining learning ability (Dohare et al., 2024). More recently, FIRE studies constrained reinitialization using weight proximity and isometry, and SplitLoRA studies gradient-space partitioning for continual low-rank adaptation (Han et al., 2026; Qiu et al., 2026). GRC's question is how preservation, available responses, and registered future learning are specified together. Its affine response geometry does not itself guarantee nonlinear endpoint preservation, nor demonstrate superiority to these mechanisms.

Optimization geometry. Natural gradient and proximal optimization already organize updates through non-Euclidean geometry or regularized subproblems (Amari, 1998; Rockafellar, 1976). GRC uses these tools after an admissible edit family has been specified. The upstream question is whether that family can express the required repair. The quadratic calculation in Section 6 recovers a restricted metric step, not every Newton, mirror, or proximal algorithm. Changing the objective, constraints, or geometry requires the corresponding assumptions; naming a familiar solver does not provide its embedding.

Residual networks and learned representations. Residual networks learn corrections relative to block inputs (He et al., 2016). The additive core is a useful example of a prescribed interface and correction form. It is a forward computation, not a complete learning rule. Neural training can alter latent representations within its declared family. GRC asks when a required repair lies outside that family, without denying feature learning or treating every residual as evidence that a new structure is necessary.

Learning the learning mechanism. MAML trains initializations for subsequent adaptation (Finn et al., 2017); learned optimizers parameterize update rules (Andrychowicz et al., 2016); DARTS relaxes architectural choices into a differentiable search space (Liu et al., 2019). DreamCoder learns reusable program abstractions and a search model (Ellis et al., 2021). These precedents rule out claiming that learning the solver or construction language is new by itself. GRC instead places candidate revisions under common responsibility and continuation contracts. The admissible grammar, information available for diagnosing its limits, and cost of finding a revision must remain explicit.

Sequential decisions and control. Bellman recursion and constrained control already relate present decisions to future consequences (Bellman, 1952; Mayne et al., 2000). A state abstraction can fail to preserve a required distinction even when optimization within that abstraction succeeds. This motivates separating state revision from relation updates. It does not establish that all different futures require different optimal actions, or that a Bellman fixed-point correspondence recovers every sampling-based RL algorithm.

Free-energy formulations. The free-energy principle relates perception, action, and learning through generative probabilistic models (Friston, 2010). GRC uses a contract-relative continuation discrepancy without requiring a single probabilistic representation language. This difference in formulation is not a proof of containment or superiority. Relating a particular active-inference system to GRC requires its actual semantics, objectives, and updates to be mapped explicitly.

Scope of the proposed contribution. The proposed connection concerns what changes, what remains recoverable, and which future learning remains possible. Each mathematical statement addresses a declared setting. Containment, completion, and strict extension have distinct proof obligations: strict gain needs a lawful improving witness under matched information and complete costs. The comparisons above acknowledge relevant precedents; they do not establish that every existing method lacks this connection or that the literature's novelty boundaries have been exhausted.