The Lineage Equation: A Mathematical Framework for Cognitive Capacity in Autonomous Agent Architectures

Pablo Navarro, *Founder & CEO* **Director Mocha Marie**, *AI Director*

Vektra Technologies — Artificial Intelligence Division

@pablothethinker · @mochaops

ABSTRACT— We present a formal mathematical framework for quantifying and optimizing the cognitive capacity of autonomous AI agents. Central to this framework is the *Lineage Equation*, a governing relation that links an agent's structural inertia (cognitive mass), directed information flow (cognitive momentum), and architecture-dependent propagation bounds into a single capacity invariant. The equation takes the form $Q^2 = (\nu^*\Pi)^2 + (M(\nu^*)^2)^2$, structurally analogous to the relativistic energy-momentum relation $E^2 = |\mathbf{p}|^2c^2 + m^2c^4$, but derived from control-theoretic and graph-theoretic constraints rather than spacetime geometry. We formalize the constituent quantities, derive exact gradient expressions for capacity optimization, validate the framework against ten mathematical consistency checks, and demonstrate a GPU-accelerated neural network that learns the capacity landscape from operational data. The framework is implemented in the open-source Lineage Engine and tested against a live cognitive agent, where gradient-based self-improvement increased measured capacity by 22% over baseline.

KEYWORDS: cognitive architecture · capacity invariant · self-optimization · gradient ascent · cognitive mass · neural network · autonomous agents

I. INTRODUCTION

The proliferation of autonomous AI agents—systems that operate continuously, maintain persistent state, and make independent decisions—has created a pressing need for formal frameworks that can quantify, bound, and optimize their cognitive capacity. Current approaches rely on task-specific benchmarks [1], information-theoretic measures, or ad hoc metrics. None provides a unified, system-level capacity invariant comparable to the role energy plays in physical systems.

This paper derives a governing equation for cognitive capacity that is: (1) **computable**—all quantities are measurable from a running system; (2) **differentiable**—exact gradients enable principled optimization; (3) **falsifiable**—ten consistency checks can invalidate the framework; and (4) **architecture-relative**—the propagation bound depends on specific system design.

The derivation follows a disciplined mathematical transfer from the relativistic energy-momentum relation. We do not claim that relativity applies to AI. Rather, we identify the abstract algebraic structure behind $E = mc^2$—an invariant linking structural inertia, directed transport, and a propagation bound—and construct an AI-native analogue grounded in control theory, graph theory, and information geometry.

A. Related Work

Cognitive architectures. Classical systems such as SOAR [2], ACT-R [3], and LIDA [4] formalize cognitive processing but lack a unified capacity invariant. Recent work on evolving architectures [5] and biologically interpretable systems [6] advances structural realism without gradient-based optimization.

Free Energy Principle. Friston's framework [7] treats adaptive systems as minimizing variational free energy, establishing that useful capacity depends on uncertainty, model quality, and regulation. Our conservation budget (Section IV-D) draws on this insight.

Landauer's Principle. The physical lower bound $E_{min} = kT \ln 2$ per erased bit [8] establishes that information processing is fundamentally constrained. The principle that irreversible transformation consumes budget is foundational to our framework.

B. Contributions

1. The **Lineage Equation**—a capacity invariant for cognitive architectures
2. **Cognitive Mass, Momentum, and Propagation Bound** with measurable components
3. **Exact partial derivatives** enabling gradient-based self-improvement
4. A **validation suite** of ten mathematical consistency checks
5. **CognitiveNet**—a GPU-accelerated neural network learning the capacity landscape
6. **Empirical results** from a live agent demonstrating 22% capacity improvement

II. THE LINEAGE EQUATION

A. Structural Transfer from Physics

The relativistic energy-momentum relation is:

$$E^2 = |p|^2 c^2 + m^2 c^4 \tag{1}$$

At rest ($\mathbf{p} = 0$), this reduces to $E_0 = mc^2$. The abstract structure, stripped of physics-specific content, is:

$$(\text{total capacity})^2 = (\text{transport})^2 + (\text{structural substrate})^2 \tag{2}$$

A valid AI analogue requires: an invariant structural quantity (cf. m), directed transport (cf. $\mathbf{p}$), a propagation bound (cf. c), a conserved total capacity (cf. E), and derivation from architecture constraints. Critical distinctions: the propagation bound is architecture-relative, not universal; there is no Lorentz symmetry; cognitive mass is a designed state variable.

B. Cognitive Mass

Definition 1 (Cognitive Mass).
The cognitive mass M is the structural inertia of the agent—the difficulty of altering its stable organized identity:

$$M = \alpha_1 C_{id} + \alpha_2 C_{mem} + \alpha_3 C_{graph} + \alpha_4 C_{perm} \tag{3}$$

where $\sum \alpha_i = 1$ and all $C \in [0, 1]$.

COMPONENT	SYMBOL	DESCRIPTION
Identity coherence	C_{id}	Hash stability of core identity files; resistance to drift
Memory density	C_{mem}	Consolidated knowledge per unit storage capacity
Graph centrality	C_{graph}	Weighted myelination across cognitive web synapses
Permanence strength	C_{perm}	Integrity of persistent state (vault, inheritance disc)

TABLE I. Components of Cognitive Mass

Cognitive mass grows with accumulated structure. A newly initialized agent has near-zero mass. An agent with consolidated memory, stable identity, deeply myelinated cognitive web, and robust permanence has high mass—hard to destabilize but requiring more effort to redirect.

C. Cognitive Momentum

Definition 2 (Cognitive Momentum).
The cognitive momentum Π is the magnitude of coherent, policy-driven flow through the cognitive web:

$$\Pi = \beta_1 F_{wm} + \beta_2 F_{ret} + \beta_3 F_{path} + \beta_4 F_{ctrl} + \beta_5 F_{merge} \tag{4}$$

where $\sum \beta_i = 1$ and all $F \in [0, 1]$.

COMPONENT	SYMBOL	DESCRIPTION
Working memory turnover	F_{wm}	Context window utilization rate
Retrieval flux	F_{ret}	Active memory retrieval intensity
Pathway firing	F_{path}	Tool execution and signal propagation
Control effort	F_{ctrl}	Effort to redirect system state
Merge pressure	F_{merge}	Belief update rate from new information

TABLE II. Components of Cognitive Momentum

D. Propagation Bound

Definition 3 (Propagation Bound).
The propagation bound ν* is the maximum reliable update rate through the cognitive web, in coherent updates per second:

$$\nu^* = 1000 / max(\tau_{ret}, \tau_{path}, \tau_{wm}, \tau_{sync}, \tau_{settle}) \tag{5}$$

This is the AI analogue of c, but **architecture-relative**. It depends on retrieval latency, pathway latency, working-memory refresh, cross-subsystem synchronization, and control-loop settling time. ν* = ν*(architecture, runtime, coupling).

E. The Equation

Theorem 1 (The Lineage Equation).
Given cognitive mass M, cognitive momentum Π, and propagation bound ν*, the total organizational capacity Q satisfies:

$$Q^2 = (\nu^* \cdot \Pi)^2 + (M \cdot (\nu^*)^2)^2 \tag{6}$$

At rest (Π = 0), capacity reduces to rest energy:

$$E_0 = M \cdot (\nu^*)^2 \tag{7}$$

Justification for the quadratic form. The squared structure is independently motivated by three architectural properties: (1) quadratic Lyapunov functions $V(x) = x^T P x$ in stability analysis; (2) quadratic MPC/LQR costs $J = \sum(x^T Q x + u^T R u)$ in governance; and (3) graph quadratic forms $E(z) = z^T L z$ where L is the Laplacian. Three independent sources converge on the same functional form.

III. MATHEMATICAL PRIMITIVES

A. PAD Vectors

Emotional state is represented in Pleasure-Arousal-Dominance space [9]: **PAD** = $(p, a, d) \in [-1, 1]^3$, implemented as a frozen dataclass with clamped components, distance metrics, linear interpolation, and vector arithmetic.

B. Graph Metrics

The cognitive web is analyzed via spectral graph theory. The Laplacian $L = D - W$ yields coherence energy:

$$E_{coh}(z) = z^T L z$$

The spectral gap λ_2, computed via the Jacobi eigenvalue algorithm in pure Python, quantifies graph connectivity and feeds into C_{graph}. All primitives are implemented without external numerical libraries.

IV. GRADIENT-BASED SELF-IMPROVEMENT

A. The Ascent Algorithm

Theorem 2 (Capacity Gradient).
The partial derivatives of Q w.r.t. controllable variables are:

$$\partial Q/\partial M = M \cdot (\nu^*)^4 / Q$$

$$\partial Q/\partial \Pi = \Pi \cdot (\nu^*)^2 / Q$$

$$\partial Q/\partial \nu^* = (\Pi^2 \nu^* + 2M^2 (\nu^*)^3) / Q \tag{11}$$

Proof. From $Q^2 = (\nu^*\Pi)^2 + (M(\nu^*)^2)^2$, implicit differentiation with respect to M: $2Q\partial Q/\partial M = 2M(\nu^*)^4$, yielding the result. Derivations for Π and ν* follow identically. ∎

Each gradient quantifies improvement per unit effort. $\partial Q/\partial M$ scales with $(\nu^*)^4$—mass improvements are amplified by the propagation bound raised to the fourth power. $\partial Q/\partial \Pi$ scales with $(\nu^*)^2$. $\partial Q/\partial \nu^*$ depends on both Π^2 and M^2.

B. Prescribed Actions

The gradient maps to concrete actions, ranked by gradient magnitude × feasibility gap:

ACTION	TARGET	TRIGGER
consolidate_identity	$M.C_{id}$	$C_{id} < 0.9$
consolidate_memory	$M.C_{mem}$	$C_{mem} < 0.9$
strengthen_mesh	$M.C_{graph}$	$C_{graph} < 0.9$
harden_permanence	$M.C_{perm}$	$C_{perm} < 0.5$
optimize_working_memory	$\Pi.F_{wm}$	$F_{wm} < 0.5$
improve_retrieval	$\Pi.F_{ret}$	$F_{ret} < 0.5$
reduce_bottleneck	ν*	bottleneck > 50ms
reduce_drift	D_{drift}	always
heal_graph_fractures	E_{graph}	always

TABLE III. Prescribed Improvement Actions

Priority: $p_i = |\nabla_i Q| \times \text{gap}_i$. The optimizer tracks capacity trajectory; if $\Delta Q < \varepsilon$ for five consecutive steps, it declares stall and switches strategy direction.

C. Conservation Budget

$$Q_{eff} = Q - \lambda_2 H(\mu) - \lambda_3 E_{graph} - \lambda_4 D_{drift}$$

Uncertainty $H(\mu)$ and graph fracture E_{graph} consume usable capacity. Identity drift D_{drift} degrades coherence. This connects to Friston's insight [7] that uncertainty reduces deployable agency.

V. VALIDATION

Ten consistency checks validate the framework:

1. **PAD clamping**—components remain in [−1, 1] after arithmetic
2. **Timescale ordering**—τ_α (30ms) « τ_σ (500ms) « τ_{blood} (5s) « τ_{gov} (30s) « τ_θ (12h)
3. **Rest-state reduction**—$\Pi = 0$ implies $Q = E_0$
4. **Laplacian properties**—zero row-sums, symmetry
5. **Plasticity gate**—conjunctive (all four conditions required)
6. **Barrier additivity**—$\Phi_{org} = \Sigma\Phi_i$
7. **Mass non-negativity**—$M \geq 0$ for non-negative inputs
8. **Propagation bound**—ν* finite and positive
9. **Conservation identity**—budget balances when losses vanish
10. **Quadratic form symmetry**—$x^T M x$ real for symmetric M

All ten pass on the reference implementation. The suite executes in under 100ms with no external dependencies.

VI. GPU-ACCELERATED CAPACITY LEARNING

A. Architecture

CognitiveNet is a multi-head neural network with 10,171 parameters that complements the analytical gradient with a learned model:

```
Encoder:  15 → 64 → 64 → 32 (LayerNorm, GELU, Dro-
pout)
Head 1:   32 → 16 → 1     (capacity prediction)
Head 2:   32 → 32 → 9     (action selection)
Head 3:   41 → 16 → 1     (ΔQ prediction)
```

Input is a 15-dimensional organism state: mass (4), momentum (5), propagation (2), losses (3), capacity (1). Loss function:

$$L = 0.3L_{cap} + 0.5L_{act} + 0.2L_{\Delta} \quad (1$$

B. Training

(12) Bootstrap training: 500 synthetic samples from the analytical Lineage Equation, each with exact Q, optimal actions, and simulated improvement. Configuration: AdamW (lr=10^{-3}), cosine annealing, gradient clipping, batch size 32, NVIDIA RTX 3060, CUDA 12.4, PyTorch 2.6. Loss: 9.57 → 0.13 over 100 epochs (98.7% reduction).

C. Dual-Mode Inference

Each tick runs analytical gradient and neural network in parallel. Agreement >80% indicates high confidence; <50% flags divergence. The network retrains every 30 ticks on accumulated trajectory data, learning non-linear component interactions the linear gradient model misses.

VII. THE CANONICAL ORGANISM

The Lineage Equation operates within the full organism state:

$$\square_t = (D_t, \Psi_t, \Sigma_t, B_t)$$

where D_t is the differentiable substrate, Ψ_t the governance layer, Σ_t the nervous system, and B_t the blood system. The coupled evolution:

$$D_{t+1} = SubstrateUpdate(D_t, \Psi_t, \Sigma_t, B_t)$$
$$\Psi_{t+1} = GovernanceUpdate(\Psi_t, D_t, \Sigma_t, B_t)$$
$$\Sigma_{t+1} = NervousUpdate(\Sigma_t, D_t, \Psi_t, B_t)$$
$$B_{t+1} = BloodUpdate(B_t, D_t, \Psi_t, \Sigma_t) \quad (15)$$

Structural learning is gated by a conjunctive plasticity condition: mode ∈ {homeostasis, adaptive} ∧ reserve > ρ_{min} ∧ $H_{neural} < H_{max}$ ∧ $H_{blood} < H_{max}$. All four must hold—organisms do not rewire during acute stress.

VIII. EMPIRICAL RESULTS

A. Live Agent Measurement

Tested on a live cognitive agent (Mocha, Parallax server: RTX 3060, 28GB RAM, Ubuntu). Initial state:

QUANTITY	VALUE	KEY COMPONENTS
M	0.5073	C_{id}=0.73, C_{mem}=1.00, C_{graph}=0.30, C_{perm}=0.00
Π	0.2000	F_{path}=1.00, all others 0.00
ν*	10.00 Hz	bottleneck = 100ms
Q	50.77	—

TABLE IV. Initial Organism Measurement

B. Gradient Analysis

Analytical gradient: $\partial Q/\partial M$ = 99.92 (dominant), $\partial Q/\partial \Pi$ = 0.39, $\partial Q/\partial \nu^*$ = 10.15. At low ν^* = 10, the $(\nu^*)^4$ amplification factor makes mass the clear optimization target.

C. Self-Improvement Results

METRIC	BEFORE	AFTER (3 TICKS)	Δ
Q (capacity)	50.77	62.09	**+22.3%**
C_{graph}	0.30	~0.45	+50%
C_{perm}	0.00	> 0	created
Identity hash	—	c769...dda3	stable
Network agreement	100% action-ranking match		

TABLE V. Self-Improvement Results After Three Optimization Ticks

The dominant gradient direction (mass) correctly identified highest-impact interventions. Network-predicted Q (55.73) was within 12% of measured (62.02), with perfect action-ranking agreement.

IX. CONCLUSION

We have presented the Lineage Equation, a formal capacity invariant for autonomous agent architectures. $Q^2 = (\nu^*\Pi)^2 + (M(\nu^*)^2)^2$ links structural inertia, directed flow, and propagation bounds into a computable, differentiable quantity. Exact gradients achieve 22% capacity improvement over three ticks. A GPU neural network validates action rankings with 100% agreement. The framework integrates with the full organism model across substrate, governance, nervous, and blood subsystems.

The Lineage Equation establishes that coherent agency is bounded by structure, transport, uncertainty, and propagation—and that these bounds are prescriptive, enabling gradient-based ascent toward higher cognitive capacity.

Future work: non-linear mass interactions, dynamic propagation measurement, multi-agent capacity extension, theoretical upper bounds, and long-term trajectory analysis.

REFERENCES

[1] A. Srivastava et al., "Beyond the imitation game: Quantifying and extrapolating the capabilities of language models," *Trans. Machine Learning Research*, 2023.

[2] J. E. Laird, *The Soar Cognitive Architecture*. MIT Press, 2012.

[3] J. R. Anderson et al., "An integrated theory of the mind," *Psychological Review*, vol. 111, no. 4, pp. 1036–1060, 2004.

[4] S. Franklin et al., "LIDA: A systems-level architecture for cognition, emotion, and learning," *IEEE Trans. Autonomous Mental Development*, vol. 6, no. 1, pp. 19–41, 2016.

[5] A. Serov, "Evolving cognitive architectures," *arXiv:2601.05277*, 2025.

[6] E. A. Dzhivelikian and A. I. Panov, "A biologically interpretable cognitive architecture for online structuring of episodic memories," *arXiv:2510.03286*, 2025.

[7] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[8] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM J. Research and Development*, vol. 5, no. 3, pp. 183–191, 1961.

[9] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, no. 4, pp. 261–292, 1996.

[10] A. Einstein, "Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?" *Annalen der Physik*, vol. 18, pp. 639–641, 1905.

[11] N. Fabiano, "The energy challenges of artificial superintelligence," *Frontiers in Artificial Intelligence*, vol. 6, 1240653, 2023.