

Chapter 9.5 — The Persistence Ratio

Contents

Chapter 9.5: The Persistence Ratio: The Thermodynamics of Existence	1
---	---

Chapter 9.5: The Persistence Ratio: The Thermodynamics of Existence

Up to this point, the argument of **Useful Approximations Framework (UAF)** has been one of *functional necessity*. The **Epistemic Veil** (Chapter 5) forces approximation; **Skin in the Game** (Chapter 6) supplies the imperative; the **Internal Self-Model**, **Qualia**, and **World-Model** (Chapters 7–9) are the necessary computational instruments. But functional necessity, as a mode of argument, only tells us that *some* solution must exist. It does not tell us *what an existing system must continuously do, in physical terms, in order to remain a system at all*.

This chapter closes that gap. We will show that the imperative UAF names “Skin in the Game” is not a colourful metaphor borrowed from Taleb (2018) but a strict thermodynamic accounting identity: the **Persistence Ratio**, here denoted $\mathcal{R}$. A node — be it a particle, a cell, a brain, a corporation, or a civilization — exists at time t if and only if its rate of negentropic production exceeds its rate of entropic dissipation. The moment this ratio falls below unity, the node dissolves into the background noise of its environment. There is no exception, no shelter, no theological reprieve. *Existence is a calculation that must be repaid every instant in the currency of free energy* (Friston, 2010; Schrödinger, 1944).

The Persistence Ratio is the missing engine beneath every other concept in this book. It explains why PEM (Chapter 12) is not merely a learning convenience but a survival requirement; why qualia (Chapter 8) must be causally efficacious rather than epiphenomenal (Chalmers, 1996); why no node can be self-caused, and why the entire fractal of nested systems described in **Chapter 3** is held together by a single, recursive thermodynamic identity.

The Question of Why Anything Persists The Second Law of Thermodynamics states that the entropy of an isolated system never decreases (Clausius, 1865). Read naively, this seems to forbid the existence of structured things: stars, bacteria, brains, civilizations. Yet such structures clearly exist, and some of them persist for billions of years. The resolution, first articulated by Schrödinger (1944) and made formal by Prigogine (1977), is that ordered structures are not isolated. They are **dissipative structures**: open systems that import free energy, perform work to maintain their internal order, and export entropy to the environment. *A whirlpool, a flame, a cell, and a conscious brain are all instances of the same general type — a pattern that exists by continuously paying its entropic tax.* (Prigogine and Stengers, 1984; Lane, 2015.)

This reframing turns ontology into accounting. Instead of asking *what* something is, we ask *what flux of energy and information it must sustain to remain itself*. The Persistence Ratio is the formal answer to that question.

The Persistence Ratio: Definition Let a **node** at level L be any discrete informational pattern with a well-defined boundary (a Markov blanket, in Friston’s vocabulary). Let $\dot{I}_{gen}^{(L)}$ be the rate at which the node generates internal order (negentropy) by performing work against its environment. Let $\dot{S}_{diss}^{(L)}$ be the rate at which entropy is being produced inside the node — by thermal noise, by erroneous predictions, by internal friction. Then the **Persistence Ratio** is simply the ratio of the two:

$$\mathcal{R}^{(L)} = \frac{\dot{I}_{gen}^{(L)}}{\dot{S}_{diss}^{(L)}}.$$

The **First Law of Persistence** is the single inequality on which this entire framework rests — for the node to exist at time t :

$$\mathcal{R}^{(L)} \geq 1.$$

If $\mathcal{R} > 1$, the node accumulates structural reserves; it grows, learns, expands its Markov blanket. If $\mathcal{R} = 1$, the node is in a non-equilibrium steady state (NESS): structurally constant, metabolically alive, but with no surplus. If $\mathcal{R} < 1$, the node is dissolving — the only question is how quickly. *There is no fourth option. There is no “static existence” outside this calculation; thermodynamic equilibrium is the precise technical name for being dead* (England, 2013).

This single inequality is the formal content of “Skin in the Game.” Taleb’s (2018) intuition that a system without genuine downside is not really a system at all is, on this view, a vernacular restatement of the requirement $\mathcal{R} \geq 1$.

The Fractal Persistence Equation To make $\mathcal{R}$ useful, we must decompose the numerator and denominator into observable terms and acknowledge that no node is isolated from the nodes above and below it. **Chapter 3** established the Network Imperative: a node at level L is constituted by nodes at level $L - 1$ and embedded in nodes at level $L + 1$. Persistence therefore couples vertically across the fractal. The full equation — which we will call the **Fractal Persistence Equation (FPE)** — is:

$$\mathcal{R}^{(L)} = \Psi(\mathcal{R}^{(L+1)}) \cdot \left[\frac{P_{in}^{(L)} \eta(I)}{\omega^{(L)} \mathcal{E}_{\Sigma}^{(L)} (1 + \mathcal{D}_{KL}^{(L)} + \Gamma^{(L)})} \right] \cdot \Phi(\mathcal{R}^{(L-1)}).$$

Each term has a precise meaning and a direct interpretation within the UAF framework:

- $P_{in}^{(L)}$ — **Power Input**. The rate at which usable free energy is imported across the node’s boundary. In a cell, this is metabolic ATP throughput; in a brain, glucose and oxygen; in a corporation, revenue; in a conscious agent, the energetic and attentional cost of doing work against the world. *In UAF terms, P_{in} is the metabolic substrate of Skin in the Game; without it, the imperative for Coherence and Agency has nothing to run on.*
- $\eta(I)$ — **Informational Efficiency**. The fraction of imported power that is converted into useful predictive work rather than wasted as heat. By Landauer’s Principle (Landauer, 1961; Bennett, 2003), every erased bit of erroneous internal state costs at least $k_B T \ln 2$ joules. A system with a well-tuned World-Model and ISM erases few bits unnecessarily and therefore has high η .
- $\omega^{(L)}$ — **Structural Complexity**. The number of degrees of freedom that the node’s boundary must continuously stabilise against thermal drift. Larger, more elaborate nodes have larger ω and pay a higher maintenance cost. *This is why simple bacteria can persist on trace metabolism while a human cortex demands $\sim 20\%$ of the body’s total energy budget* (Aiello and Wheeler, 1995).
- $\mathcal{E}_{\Sigma}^{(L)}$ — **Fundamental Noise Floor**. The aggregate of thermal, quantum, and environmental fluctuations that constantly perturb the node. It is non-zero everywhere in spacetime; this is the universe’s irreducible cost of existence.
- $\mathcal{D}_{KL}^{(L)}$ — **Model Divergence (the Honesty Penalty)**. The Kullback–Leibler divergence between the node’s internal generative model Q and the true environmental distribution P :

$$\mathcal{D}_{KL}(P \parallel Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}.$$

This is the exact technical form of the Epistemic Veil’s *cost*. A node whose internal models drift away from reality must perform extra work to suppress the resulting prediction errors (Friston,

2010; Clark, 2013). $\mathcal{D}_{KL}$ thus quantifies what Chapter 12 will call the surcharge on imperfect approximation — and what Chapter 31 will identify as a defining feature of mental illness.

- $\Gamma^{(L)}$ — **Structural Fatigue / Friction**. Accumulated wear: senescence in biology, technical debt in software, resentment in relationships, sclerosis in institutions. Γ grows over time unless actively repaired and contributes additively to the denominator.
- $\Phi(\mathcal{R}^{(L-1)})$ — **Bottom-Up Coupling**. The contribution of sub-nodes to the node’s integrity. If your cells fail, your body fails; if your neurons starve, your conscious stream collapses. Formally: $\mathcal{R}^{(L)} \leq \min_i \mathcal{R}_i^{(L-1)}$ in the limit, so $\lim_{\mathcal{R}^{(L-1)} \rightarrow 0} \mathcal{R}^{(L)} = 0$.
- $\Psi(\mathcal{R}^{(L+1)})$ — **Top-Down Coupling**. The shielding the environment provides against raw cosmic entropy. A healthy biosphere shields its organisms; a stable civilization shields its citizens. When $\mathcal{R}^{(L+1)} \rightarrow 0$, $\Psi \rightarrow 0$ and the noise floor effectively diverges: *no individual node, however internally well-tuned, can persist in a collapsing environment* (Tainter, 1988; Diamond, 2005).

The numerator therefore captures what aggressive, active work the node does to maintain itself; the denominator captures what the universe charges it for the privilege. Persistence is the difference between the two — and that difference, by the inequality above, must be non-negative.

The fractal graph: scale limits, substrate, and overlapping shelter Chapter 3 traced the network imperative from quarks to cosmos. This section names the bookkeeping behind that picture — three points readers often import implicitly but the formalism must keep separate.

Scale limits and the global graph. The level index L orders persisting patterns from fine to coarse. At the **lower bound** L_0 , decomposition stops at constituents we do not resolve as separate blankets on the timescales we care about — quarks and hadrons in physics, organelles and cells in biology. Fundamental physics may extend L_0 further; the accounting is unchanged. At the **upper bound**, the closed universe is the ultimate phase space: every node $\Sigma \subset \Gamma$ has environment $\Gamma \setminus \Sigma$. Cosmic structure (galaxies, the expanding background) is the highest *persisting* IPS we model in practice; Γ itself is not a driven subsystem paying $\mathcal{R}$ in the same way a cell does — when nothing encloses you, shelter Ψ is undefined and we treat it as unity by convention. **Between** L_0 and cosmic scale, reality is not a single linear chain but a **fractal graph**: each node has substrate edges to $\{\Sigma_i^{(L-1)}\}$ and shelter edges to enclosing $\Sigma_j^{(L+1)}$, with atoms, organisms, firms, and polities as intermediate nodes — branching, skipping levels, and cross-linking where blankets overlap.

Composition (Φ) vs shelter (Ψ). These are different relations and must not be conflated.

- **Made of** is *substrate*: a person from organs and immune networks; a polity such as Finland from firms, households, infrastructure, and territory; a cell from molecular machines. Φ is the integrity of that heterogeneous graph — not one part, but which constituents still satisfy $\mathcal{R}_i \geq 1$ and how critical each is. Lose a critical sub-node and $\Phi \rightarrow 0$ regardless of how good the weather is.
- **Belongs to** is *shelter*: attenuation of environmental noise by whatever actually buffers shocks for the node — body and household for a person; climate, regional cooperation, and supranational institutions for a state. A label without thermodynamic buffering is not a blanket.

Failure modes differ empirically. Losing one shelter (unemployment, trade isolation) lowers one channel; losing critical substrate (organ failure, banking-system collapse) attacks Φ directly. A node can have rich substrate *and* several overlapping shelters without having only one parent.

Multiple shelters. When several enclosures buffer *distinct* noise spectra, effective shelter combines conservatively — to leading order $\Psi_{\text{eff}} \approx \prod_j \Psi_j$ when channels are independent, or the minimum for a worst-case bound. Nested physical containment (galaxy $\rightarrow$ planet $\rightarrow$ biosphere) is the special case of one dominant shelter per scale, not the general case.

Part IV-B will show how aggression, honesty, community, and respect are behavioural faces of the same numerator and denominator at the social layer. The full formal treatment, with proofs, is in [papers/information_persisting_systems.md](#) (Sections 2.4–2.5, 6.7).

Axioms and the Proof of Necessary Convergence The full FPE looks elaborate, but its claim — that $\mathcal{R} \geq 1$ is mandatory — follows from three modest axioms that any reader who accepts standard physics and information theory already grants.

Axiom 1 (Universal Noise). The fundamental noise floor $\mathcal{E}_\Sigma$ — comprising thermal, quantum-vacuum, and gravitational fluctuations — is strictly positive in all regions of spacetime. *No node anywhere gets noise for free.*

Axiom 2 (Landauer Efficiency). The erasure of any erroneous bit requires a minimum energetic dissipation of $k_B T \ln 2$ (Landauer, 1961; Bérut et al., 2012). *Correcting one’s errors is never thermodynamically free.*

Axiom 3 (Finite Information Density). No node within the observable universe possesses infinite internal information I or infinite power throughput P_{in} . *Resources are always bounded.*

From these three axioms, three lemmas follow.

Lemma 1 — The Dissipation Constraint. If $P_{in}\eta < \omega\mathcal{E}_\Sigma$, the node must compensate by consuming its own internal structural information. Since that store is finite (Axiom 3), it is exhausted in finite time and the node disintegrates. *A house cannot heat itself indefinitely by burning its own walls.*

Lemma 2 — The Delusion Penalty. Any $\mathcal{D}_{KL} > 0$ generates persistent prediction errors. By the Free Energy Principle, surprise is mathematically equivalent to $\mathcal{D}_{KL}$ (Friston, 2010). By Axiom 2, every correction costs energy. Drift between model and reality therefore acts as an exponential multiplier on the entropic tax: a small, ignored delusion grows into a catastrophic energy sink. *This is the mathematical core of Chapter 31’s account of mental illness, and of the “Honesty Penalty” we will encounter in Chapter 31.2.*

Lemma 3 — Fractal Dependency. No node is self-caused. If $\mathcal{R}^{(L-1)} \rightarrow 0$, the node’s substrate dissolves and $\Phi \rightarrow 0$. If $\mathcal{R}^{(L+1)} \rightarrow 0$, the shielding $\Psi \rightarrow 0$ and the effective noise floor diverges. Therefore $\mathcal{R}^{(L)}$ is bounded above by the persistence of both the nodes that compose it and the node that contains it.

These three lemmas yield the **Proof of Necessary Convergence**, conducted as a reductio ad absurdum:

1. Assume a node persists indefinitely with $\mathcal{R} < 1$.
2. By Lemma 1, the node produces more entropy than it exports.
3. By the Second Law, this excess entropy must either heat the node or break its internal bonds.
4. If the bonds break, the structural complexity ω dissolves.
5. Once ω dissolves, the Markov blanket vanishes.
6. A pattern without a Markov blanket is, by definition, background noise — not a distinct node.
7. **Contradiction.** The “persisting” node was assumed to be distinct from its environment, but the analysis shows it has merged with it.

Therefore $\mathcal{R} \geq 1$ is a necessary condition for the existence of any node at any level. The inequality is not a heuristic; it is an accounting identity that the universe enforces continuously.

Empirical Witness Across Scales The FPE is not idle formalism. The same inequality is visible — with different units, but identical structure — across every domain in which something either lasts or fails to last.

- **Radioactive decay.** A heavy isotope is a configuration in which the strong-force binding energy only marginally exceeds the weak-force perturbation. When local fluctuations push the internal $\mathcal{E}$ above the binding term, $\mathcal{R}$ transiently drops below 1 and the nucleus decays into a more stable configuration with $\mathcal{R} \geq 1$. The half-life of every isotope is, in effect, an empirical estimate of how often its $\mathcal{R}$ falls below unity (Krane, 1988).
- **Biological death.** As an organism ages, Γ (cellular wear, mitochondrial damage, accumulated misfolded proteins) grows. The bottom-up factor Φ degrades as sub-cellular nodes fail (López-Otín et al., 2013). When the integrated denominator outpaces $P_{in}\eta$, the organism crosses the threshold. *Death, on this view, is not an event but a continuous calculation that finally evaluates to $\mathcal{R} < 1$.*

- **Corporate bankruptcy.** A firm whose World-Model has diverged from the market ($\mathcal{D}_{KL}$ high), whose internal friction has compounded (Γ high — bureaucracy, technical debt, internal politics), and whose revenue (P_{in}) no longer covers its structural overhead ($\omega\mathcal{E}_\Sigma$) is — exactly and not metaphorically — running an $\mathcal{R} < 1$. Schumpeter’s (1942) “creative destruction” is the macro-statistic of millions of such ratios falling below unity.
- **Civilizational collapse.** Tainter (1988) and Diamond (2005) describe collapses as failures of marginal returns on complexity: ω grows faster than $P_{in}\eta$, Γ accumulates faster than it can be discharged, and the bottom-up integrity Φ of constituent communities erodes. *No invading horde is needed; the inequality does its own work.*
- **Mental health.** A mind with persistently elevated $\mathcal{D}_{KL}$ — a stuck delusion, a refused grief, an unchallenged anxiety — pays a continuous, growing energetic surcharge. Chapter 31 will return to this in detail: the brain’s inability to update the offending model is precisely an inability to lower the denominator of its own $\mathcal{R}$.

In every case, the formula is the same; only the units differ. *That is what it means to call something a law of nature.*

Persistence and the Free Energy Principle For readers familiar with Friston’s Free Energy Principle (FEP), the FPE will look like a close relative — and it is. Friston (2010) shows that any self-organising system that maintains a Markov blanket must, in the long run, behave as if it minimises a quantity called variational free energy, which is mathematically an upper bound on $\mathcal{D}_{KL}(P \parallel Q)$. The FEP can therefore be read as the operational consequence of the FPE: *systems that survive are systems that behave so as to keep $\mathcal{D}_{KL}$ small, because doing so keeps $\mathcal{R} \geq 1$.*

The FPE adds three things the bare FEP does not make explicit:

1. **An energetic numerator.** FEP describes what surviving systems minimise; FPE makes plain what they must simultaneously *generate* — namely $P_{in}\eta$, the active work term — and why no amount of accurate prediction will save a node whose power budget collapses.
2. **Vertical coupling.** Φ and Ψ make explicit the fractal dependency that the bare FEP, formulated for a single Markov blanket, leaves implicit.
3. **A pass-through to ethics and behaviour.** Because $\mathcal{D}_{KL}$ and Γ are both empirically distinguishable — one is mismatch with reality, one is friction with one’s substrate — the FPE permits a richer behavioural account of what conscious systems must do to persist. *That account is the subject of Part IV-B.*

The Persistence Ratio and the Useful Approximations Framework We can now make explicit what has been implicit since Chapter 6. UAF / UAF claim that consciousness is the **asymptotic best simplified approximation a system makes of itself and its interaction with the universe** is, in thermodynamic terms, the claim that *any sufficiently complex finite system must develop an approximation engine because it is the only available strategy for keeping its $\mathcal{R} \geq 1$ in a complex environment.*

- The Epistemic Veil (Chapter 5) is the reason $\mathcal{D}_{KL}$ can never be zero in practice.
- Approximation (Chapter 4) is the strategy for keeping $\mathcal{D}_{KL}$ as small as the system’s resources allow.
- The ISM, WM, and Qualia (Chapters 7–9) are the *form* the approximations take.
- PEM (Chapter 12) is the *algorithm* by which $\mathcal{D}_{KL}$ is continuously reduced.
- Free Will (Chapter 10) is the ISM’s necessary fiction of agency, useful precisely because acting *as if* it controlled P_{in} and η feeds back into actually controlling them.
- Mental illness (Chapter 31) is what happens when the approximation engine fails to lower $\mathcal{D}_{KL}$ and the node runs persistently with $\mathcal{R} < 1$.

In other words, the entire UAF framework can be re-stated as the answer to a single question: *what must a finite system do to keep $\mathcal{R} \geq 1$ in an environment too complex to model exactly?* The answer is: build the cheapest useful fiction, defend it actively, update it relentlessly, and stay coupled to viable nodes above and below you.

The rest of this book consists, in various idioms, of working out the implications of that one sentence.

What Follows From This Chapter The remainder of Part I (Chapters 10–17) will treat the cognitive machinery of persistence: free will as the ISM’s fiction of authorship over P_{in} (Chapter 10), the sub-conscious as the pre-trained operator of $A_{reflexive}$ (Chapter 11), PEM as the engine of $\mathcal{D}_{KL}$ -reduction (Chapters 12–14), and the formal mathematics of the system (Chapters 15–17). Parts II and III re-examine the philosophical literature in light of these mechanisms. Part IV (Chapters 28–31) shows how each FPE term has a biological substrate. **Part IV-B (Chapters 31.1–31.4) — the immediate sequel in this book’s argumentative arc — will then treat the *behavioural* implementation of the Persistence Ratio in conscious agents and social systems:** how aggression operationalises P_{in} and η at the social layer, how honesty and friction reduction operationalise $\mathcal{D}_{KL}$ and Γ , how communities and institutions implement Φ and Ψ , and what emergent property — namely *respect* — appears in the limit $t \rightarrow \infty$.

Part V will then carry the same framework into digital substrates, where every term of the FPE has an analogue and AI alignment can be re-stated as the design of systems whose persistence is coupled to ours.

For now, the reader needs only to internalise one image. A conscious system is not a thing. It is a continuous calculation. The calculation has a numerator (what you do) and a denominator (what the universe charges you). When the ratio is at least one, you exist. When it isn’t, you don’t. *To exist is to calculate. To persist is to keep the ratio above one.*

Key References Cited (*Harvard Style, Alphabetical*)

- **Aiello, L.C. and Wheeler, P.** (1995) ‘The Expensive-Tissue Hypothesis: The Brain and the Digestive System in Human and Primate Evolution’, *Current Anthropology*, 36(2), pp. 199–221.
- **Bennett, C.H.** (2003) ‘Notes on Landauer’s Principle, Reversible Computation, and Maxwell’s Demon’, *Studies in History and Philosophy of Modern Physics*, 34(3), pp. 501–510.
- **Bérut, A. et al.** (2012) ‘Experimental Verification of Landauer’s Principle Linking Information and Thermodynamics’, *Nature*, 483, pp. 187–189.
- **Chalmers, D.J.** (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- **Clark, A.** (2013) ‘Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science’, *Behavioral and Brain Sciences*, 36(3), pp. 181–204.
- **Clausius, R.** (1865) ‘Über verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie’, *Annalen der Physik*, 125(7), pp. 353–400.
- **Diamond, J.** (2005) *Collapse: How Societies Choose to Fail or Succeed*. Viking.
- **England, J.L.** (2013) ‘Statistical Physics of Self-Replication’, *Journal of Chemical Physics*, 139(12), 121923.
- **Friston, K.** (2010) ‘The Free-Energy Principle: A Unified Brain Theory?’, *Nature Reviews Neuroscience*, 11(2), pp. 127–138.
- **Krane, K.S.** (1988) *Introductory Nuclear Physics*. Wiley.
- **Landauer, R.** (1961) ‘Irreversibility and Heat Generation in the Computing Process’, *IBM Journal of Research and Development*, 5(3), pp. 183–191.
- **Lane, N.** (2015) *The Vital Question: Energy, Evolution, and the Origins of Complex Life*. W.W. Norton.
- **López-Otín, C. et al.** (2013) ‘The Hallmarks of Aging’, *Cell*, 153(6), pp. 1194–1217.
- **Prigogine, I.** (1977) ‘Time, Structure, and Fluctuations’, Nobel Lecture in Chemistry.
- **Prigogine, I. and Stengers, I.** (1984) *Order Out of Chaos: Man’s New Dialogue with Nature*. Bantam.
- **Schrödinger, E.** (1944) *What Is Life? The Physical Aspect of the Living Cell*. Cambridge University Press.
- **Schumpeter, J.A.** (1942) *Capitalism, Socialism and Democracy*. Harper & Brothers.
- **Tainter, J.A.** (1988) *The Collapse of Complex Societies*. Cambridge University Press.
- **Taleb, N.N.** (2018) *Skin in the Game: Hidden Asymmetries in Daily Life*. Random House.