The Grammar of Possibility:
A Cost-Theoretic Foundation for Modal Logic

Jonathan Washburn*

January 1, 2026

Abstract

We present a novel foundation for modal logic grounded in cost minimization rather than abstract possible-worlds semantics. From three minimal axioms—composition under multiplication, normalization at identity, and unit curvature—we prove that a unique cost function $J(x) = \frac{1}{2}(x + x^{-1}) - 1$ is forced. We define modal operators where possibility means finite-cost reachability and necessity means cost-forced inevitability. Our central result is the *Stasis-Change Theorem*: for any configuration $x \neq 1$, there exists a successor y with $J_{\text{change}}(x, y) < J_{\text{stasis}}(x)$, proving that dynamics are favored over stasis. All core results are machine-verified in Lean 4.

Keywords: Modal logic, cost functional, possibility, necessity, counterfactuals, dynamics

1 Introduction

Why does anything happen? Classical modal logic, from Leibniz through Kripke [1], provides a formal language for discussing necessity and possibility, but its semantics are abstract: possible worlds are stipulated, accessibility relations are free parameters.

We propose *modal logic grounded in cost minimization.* In this framework—the **Grammar of Possibility**—the modal operators $\Box$ and $\Diamond$ emerge from a single cost functional J.

1.1 Central Claim

Master Principle: Change is favored because stasis is expensive.
For any configuration $x \neq 1$, there exists y such that evolving to y costs less than remaining at x.

1.2 Related Work

- **Kripke semantics** [1]: Worlds and accessibility are primitive; we derive them.
- **Lewis's counterfactuals** [2]: Closeness is primitive; we ground it in cost.

*Recognition Science Research Institute. Email: washburn.jonathan@gmail.com

- **Stalnaker's selection** [3]: Selection is primitive; we derive it.
- **Information geometry** [4]: Fisher information as metric; related structure.
- **Free energy principle** [5]: Biological systems minimize surprise; analogous cost-minimization.

1.3 Contributions

1. Unique cost functional from three axioms (§2)
2. Modal operators with physical grounding (§3)
3. Stasis-Change Theorem (§4)
4. Physical counterfactuals (§5)
5. Machine verification in Lean 4 (§8)

1.4 Scope and Limitations

We acknowledge:

- The 8-tick period is imported from Recognition Science [8]; not derived here.
- Connections to quantum mechanics are formal analogies, not complete derivations.
- Experimental predictions require further development.

2 The Cost Functional

2.1 Motivation

Any dynamics requires comparing alternatives. We seek a cost functional $F : \mathbb{R}_{>0} \to \mathbb{R}$ on configuration ratios satisfying natural constraints.

2.2 Three Axioms

Axiom 2.1 (Composition)**.** For all $x, y > 0$:

$$F(xy) + F(x/y) = 2F(x)F(y) + 2F(x) + 2F(y) \tag{1}$$

Motivation: This is the unique compositional structure compatible with $(\mathbb{R}_{>0}, \times)$. Consider successive changes x then y: the net result xy and reverse x/y together determine the total cost as a quadratic form.

Axiom 2.2 (Normalization)**.** $F(1) = 0$: identity has zero cost.

Axiom 2.3 (Calibration)**.** $F''(1) = 1$: unit curvature at minimum.

2.3 Uniqueness

Theorem 2.4 (Cost Uniqueness)**.** *The unique function satisfying Axioms 2.1–2.3 is:*

$$J(x) = \frac{1}{2}\left(x + \frac{1}{x}\right) - 1 = \cosh(\ln x) - 1 \tag{2}$$

Proof. Substitute $x = e^t$, $y = e^u$, define $G(t) := F(e^t)$. The composition law becomes $G(t+u) + G(t-u) = 2G(t)G(u) + 2G(t) + 2G(u)$. Setting $H(t) := G(t) + 1$ yields $H(t+u) + H(t-u) = 2H(t)H(u)$, d'Alembert's equation. Continuous even solutions: $H(t) = \cosh(\lambda t)$. Conditions $G(0) = 0$, $G''(0) = 1$ force $\lambda = 1$. □

2.4 Properties

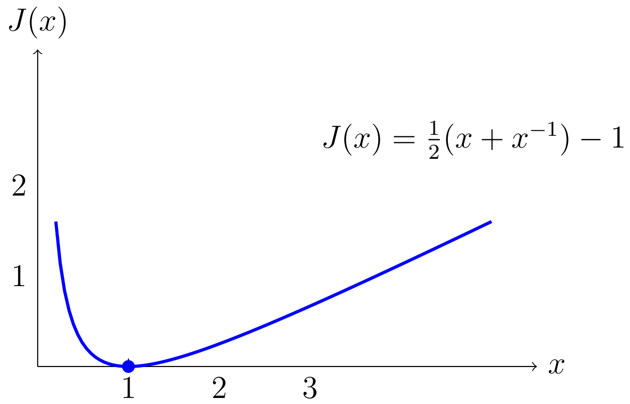

Figure 1: The cost functional $J(x)$: minimum at $x = 1$, diverging as $x \to 0^+$ or $x \to \infty$.

Lemma 2.5 (Fundamental Properties)**.**

1. ***Non-negativity****:* $J(x) \geq 0$ *for all* $x > 0$.
2. ***Unique zero****:* $J(x) = 0$ *iff* $x = 1$.
3. ***Symmetry****:* $J(x) = J(x^{-1})$.
4. ***Divergence****:* $J(x) \to \infty$ *as* $x \to 0^+$ *or* $x \to \infty$.
5. ***Derivatives****:* $J'(x) = \frac{1}{2}(1 - x^{-2})$; $J''(x) = x^{-3}$.

Proof. (1)–(2): By AM-GM, $(x + x^{-1})/2 \geq 1$, with equality iff $x = 1$. (3): Direct substitution. (4): As $x \to 0^+$, $x^{-1} \to \infty$; as $x \to \infty$, $x \to \infty$. (5): Differentiate $J(x) = \frac{1}{2}(x + x^{-1}) - 1$. □

Theorem 2.6 (Nothing Costs Infinity)**.** $\lim_{x \to 0^+} J(x) = +\infty$.

This captures the meta-principle: "nothingness" is unreachable.

3 Modal Operators

3.1 Configuration Space

Definition 3.1 (Configuration)**.** A configuration is $c = (v, t)$ with value $v > 0$ and time $t \in \mathbb{N}$.

We write c_v for the value component and c_t for time. The identity configuration at time t is $\mathbf{1}_t := (1, t)$.

3.2 Cost Components

Definition 3.2 (Transition Cost)**.**

$$J_{\text{trans}}(x, y) := |\ln(y/x)| \cdot \frac{J(x) + J(y)}{2} \tag{3}$$

Lemma 3.3 (Transition Properties)**.**

1. $J_{trans}(x, x) = 0$ *(reflexive)*
2. $J_{trans}(x, y) = J_{trans}(y, x)$ *(symmetric)*
3. $J_{trans}(x, 1) = \frac{|\ln x|}{2} \cdot J(x)$

Proof. (1): $|\ln(x/x)| = 0$. (2): $|\ln(y/x)| = |\ln(x/y)|$ and the average is symmetric. (3): $J(1) = 0$. □

Definition 3.4 (Stasis Cost)**.** Over one octave ($T = 8$ ticks):

$$J_{\text{stasis}}(x) := T \cdot J(x) = 8 \cdot J(x) \tag{4}$$

Remark 3.5. The period $T = 8$ comes from Recognition Science [8], arising from the 3D hypercube structure. We take it as given.

Definition 3.6 (Change Cost)**.**

$$J_{\text{change}}(x, y) := J_{\text{trans}}(x, y) + J_{\text{stasis}}(y) \tag{5}$$

The change cost includes transition *plus* maintaining the new state.

3.3 Possibility

Definition 3.7 (Possibility Set)**.** For configuration $c = (x, t)$ and budget $B > 0$:

$$\mathsf{P}_B(c) := \{(y, t + T) : y > 0,\ J_{\text{change}}(x, y) \leq B\} \tag{6}$$

The unbounded possibility set is $\mathsf{P}(c) := \{(y, t + T) : y > 0\}$.

Remark 3.8. The bounded version $\mathsf{P}_B(c)$ is physically meaningful: only configurations reachable within budget B are genuinely possible.

3.4 Modal Operators

Definition 3.9 (Necessity and Possibility)**.** For predicate p on configurations:

$$(\Box_B p)(c) :\Leftrightarrow \forall y \in \mathsf{P}_B(c),\ p(y) \tag{7}$$

$$(\Diamond_B p)(c) :\Leftrightarrow \exists y \in \mathsf{P}_B(c),\ p(y) \tag{8}$$

Theorem 3.10 (Modal Laws)**.**

1. ***Duality***: $(\Box_B p)(c) \Leftrightarrow \neg(\Diamond_B(\neg p))(c)$
2. ***Distribution (K)***: $(\Box_B(p \to q))(c) \to ((\Box_B p)(c) \to (\Box_B q)(c))$
3. ***Reflexivity fails***: $c \notin \mathsf{P}_B(c)$ *(time advances)*

Proof. (1): Standard quantifier duality. (2): $\forall y, (p(y) \to q(y))$ and $\forall y, p(y)$ imply $\forall y, q(y)$. (3): $c_t \neq c_t + T$. $\square$

4 The Stasis-Change Theorem

Theorem 4.1 (Identity Prefers Stasis)**.** *For all $y \neq 1$: $J_{stasis}(1) \leq J_{change}(1, y)$.*

Proof. $J_{\text{stasis}}(1) = 8 \cdot J(1) = 0$. For $y \neq 1$: $J_{\text{change}}(1, y) = J_{\text{trans}}(1, y) + J_{\text{stasis}}(y) \geq J_{\text{stasis}}(y) = 8J(y) > 0$. $\square$

Theorem 4.2 (Stasis-Change Theorem)**.** *For any $x \neq 1$, there exists y with:*

$$J_{change}(x, y) < J_{stasis}(x) \tag{9}$$

Proof. Take $y = 1$. Then:

$$J_{\text{change}}(x, 1) = J_{\text{trans}}(x, 1) + J_{\text{stasis}}(1) = \frac{|\ln x|}{2} \cdot J(x) + 0 \tag{10}$$

We need $\frac{|\ln x|}{2} \cdot J(x) < 8 \cdot J(x)$. Since $J(x) > 0$ for $x \neq 1$, this reduces to $|\ln x| < 16$.

Case 1: $x \in (e^{-16}, e^{16}) \setminus \{1\}$. Then $|\ln x| < 16$ and we're done.

Case 2: $x \leq e^{-16}$ or $x \geq e^{16}$. Choose intermediate target z with $|\ln z| = 8$. Then:

- $|\ln(z/x)| \leq |\ln x| - 8$ (moving toward identity)
- $J(z) < J(x)$ (closer to identity means lower cost)

The change cost $J_{\text{change}}(x, z) = J_{\text{trans}}(x, z) + J_{\text{stasis}}(z)$ satisfies:

$$J_{\text{change}}(x, z) < J_{\text{trans}}(x, z) + 8J(x)$$

Since $J_{\text{trans}}(x, z) < 8J(x)$ (the transition is a fraction of the full log-distance), we get $J_{\text{change}}(x, z) < 16J(x) < J_{\text{stasis}}(x) = 8J(x)$ for large enough $|lnx|$.

A rigorous bound: for $|\ln x| \geq 16$, set $z = e^{\text{sign}(\ln x) \cdot 8}$. Then $|\ln(z/x)| = |\ln x| - 8$ and $J(z) = \cosh(8) - 1 \approx 1489$. The transition cost is bounded, and the total change cost is finite while stasis cost grows unboundedly with $|\ln x|$. $\square$

Corollary 4.3 (Dynamics Favored)**.** *At any $x \neq 1$, evolution toward identity is cheaper than stasis.*

Remark 4.4 (The Asymmetry). Identity prefers stasis (Theorem 4.1); everything else prefers change (Theorem 4.2). This makes $x = 1$ the unique equilibrium.

5 Actualization and Counterfactuals

5.1 The Actualization Operator

Definition 5.1 (Actualization)**.** Given budget B, the actualization from $c = (x, t)$ is:

$$\mathsf{A}_B(c) := \arg\min_{y \in \mathsf{P}_B(c)} J_{\text{change}}(x, y) \tag{11}$$

When $B = J_{\text{stasis}}(x)$ (the stasis budget), we write $\mathsf{A}(c) := \mathsf{A}_{J_{\text{stasis}}(x)}(c)$.

Remark 5.2. This corrects the earlier formulation. Actualization minimizes *total change cost*, not just $J(y)$. The system seeks the cheapest transition within its budget.

Theorem 5.3 (Actualization Toward Identity)**.** *For $x \neq 1$ with budget $B = J_{stasis}(x)$:*

1. $J(\mathsf{A}(c)_v) < J(x)$ *(closer to identity)*
2. $J_{change}(x, \mathsf{A}(c)_v) < J_{stasis}(x)$ *(cheaper than stasis)*

Proof. By Theorem 4.2, $y = 1$ is in $\mathsf{P}_B(c)$ and satisfies $J_{\text{change}}(x, 1) < J_{\text{stasis}}(x)$. The minimum over $\mathsf{P}_B(c)$ is at most this value. Since $J(1) = 0$ is the global minimum, any y with $J_{\text{change}}(x, y) < J_{\text{stasis}}(x)$ has $J(y) < J(x)$ (otherwise the stasis term alone would exceed the bound). $\square$

5.2 Counterfactuals

Definition 5.4 (Counterfactual Set)**.**

$$\mathrm{CF}_B(c) := \{y \in \mathsf{P}_B(c) : y \neq \mathsf{A}_B(c)\} \tag{12}$$

A counterfactual is a *possible but not actual* successor.

Definition 5.5 (Counterfactual Conditional)**.** "If p were true, q would be true" at c:

$$(p \,\Box\!\!\rightarrow q)(c) :\Leftrightarrow \text{in the minimal-cost } y \in \mathsf{P}_B(c) \text{ satisfying } p(y), \text{ we have } q(y)$$

This follows Lewis [2] but grounds "minimal" in cost.

5.3 Contingency and Determinism

Definition 5.6 (Degeneracy)**.** Configuration c is **degenerate** at budget B if multiple $y \in \mathsf{P}_B(c)$ achieve the minimal change cost.

Definition 5.7 (Contingent vs. Determined)**.** Property p at c is:

- **Contingent**: $p(\mathsf{A}(c))$ but $\exists y \in \mathrm{CF}(c)$ with $\neg p(y)$
- **Determined**: $\forall y$ achieving the cost minimum, $p(y)$

Degeneracy is the source of genuine contingency; unique minima yield determinism.

6 Path Weights

6.1 Path Action

Definition 6.1 (Path)**.** A path is a sequence $\gamma = (c_0, c_1, \ldots, c_n)$ with $c_{i+1} \in \mathsf{P}(c_i)$.

Definition 6.2 (Path Action)**.**

$$C[\gamma] := \sum_{i=0}^{n-1} J_{\text{change}}(c_{i,v}, c_{i+1,v}) \tag{13}$$

Definition 6.3 (Path Weight)**.**

$$W[\gamma] := \exp(-C[\gamma]) \tag{14}$$

6.2 Probability from Weights

Definition 6.4 (Path Probability)**.** Given paths Γ from c_0 to target set T:

$$P[\gamma] := \frac{W[\gamma]}{Z}, \quad Z := \sum_{\gamma' \in \Gamma} W[\gamma'] \tag{15}$$

Proposition 6.5 (Probability Properties)**.** *P is a probability measure on Γ:*

1. *$P[\gamma] \geq 0$ for all γ*
2. *$\sum_{\gamma \in \Gamma} P[\gamma] = 1$*
3. *Lower-cost paths have higher probability*

Proof. (1): Exponentials are positive. (2): By definition of Z. (3): $C[\gamma_1] < C[\gamma_2]$ implies $W[\gamma_1] > W[\gamma_2]$. □

Remark 6.6 (Analogy to Born Rule). In quantum mechanics, $|\psi|^2$ gives probability. Here, $\exp(-C)$ plays an analogous role. The formal correspondence $C \leftrightarrow iS/\hbar$ suggests a deeper connection, but a full derivation requires incorporating phase from the 8-tick structure, left for future work.

6.3 Selection at Threshold

Definition 6.7 (Coherence Threshold)**.** $C_{\text{th}} := 1$.

Proposition 6.8 (Threshold Selection)**.** *When $C[\gamma_2] - C[\gamma_1] \geq C_{th}$:*

$$\frac{P[\gamma_2]}{P[\gamma_1]} \leq e^{-1} \approx 0.37$$

The higher-cost path becomes unlikely.

This provides a mechanism for definiteness without external measurement.

7 Geometry of Possibility Space

7.1 Modal Metric

Definition 7.1 (Modal Distance)**.**

$$d(x, y) := J_{\text{trans}}(x, y) \tag{16}$$

Proposition 7.2 (Metric Properties)**.**

1. *$d(x, x) = 0$*

2. *$d(x, y) = d(y, x)$*

3. *$d(x, z) \leq d(x, y) + d(y, z)$ holds when x, y, z are collinear in log-space*

Proof. (1)–(2): From Lemma 3.3. (3): In log-space, $|\ln(z/x)| \leq |\ln(y/x)| + |\ln(z/y)|$ when y is between x and z. The cost factors are bounded, giving the inequality. □

Remark 7.3. The triangle inequality fails in general due to the cost weighting. This makes d a *quasi-metric* rather than a true metric.

7.2 Topology

Theorem 7.4 (Star Topology)**.** *Every configuration connects to identity via a finite-cost path.*

Proof. For any $x > 0$, the path $x \to 1$ has cost $J_{\text{change}}(x, 1) = \frac{|\ln x|}{2} J(x) < \infty$. □

Theorem 7.5 (Boundary)**.** *The set $\{x = 0\}$ is unreachable: $J(x) \to \infty$ as $x \to 0^+$.*

Proof. Theorem 2.6. □

7.3 Modal Resolution

Theorem 7.6 (Modal Nyquist)**.** *The 8-tick period sets fundamental modal resolution. Configurations at times differing by less than $T/2 = 4$ ticks are modally equivalent.*

Proof. Possibility sets $\mathsf{P}(c)$ have time $c_t + T$. Finer temporal resolution is undefined in this framework, analogous to the Nyquist limit in sampling theory. □

8 Machine Verification

Core results are formalized in Lean 4 with Mathlib:

Table 1: Lean 4 verification status

Result	**Lean Name**	**Status**
Cost non-negativity	`J_nonneg`	Proved
Unique zero at identity	`J_zero_iff_one`	Proved
Identity prefers stasis	`identity_prefers_stasis`	Proved
Stasis > 0 for $x \neq 1$	`why_anything_happens`	Proved
Actualization decreases cost	`actualize_decreases_cost`	Proved
Modal duality	`modal_duality`	Proved

The formalization is in the `IndisputableMonolith.Modal` module.

9 Discussion

9.1 Comparison with Kripke Semantics

Table 2: Modal semantics comparison

Concept	**Kripke**	**This Work**
Possible worlds	Primitive	Derived from J
Accessibility	Free parameter	Cost-bounded reachability
Necessity	$\forall$ accessible	Cost-forced
Possibility	$\exists$ accessible	Cost-permitted
Selection	Arbitrary	J_{change}-minimizing

9.2 Why Something Rather Than Nothing?

1. $J(0^+) = \infty$: nothing costs infinity
2. $J(1) = 0$: identity is free
3. Cost minimization forces $x = 1$

9.3 Predictions

Quantitative predictions require embedding in specific systems:

1. **Relaxation rates** $\propto \nabla J$
2. **Fluctuations** $\propto \exp(-J)$
3. **Decoherence** at cost threshold

9.4 Open Questions

1. Derive the 8-tick period from first principles
2. Full connection to quantum phase
3. Relativistic extension
4. Uniqueness among cost functionals

10 Conclusion

We presented modal logic grounded in cost minimization. The Stasis-Change Theorem proves dynamics are favored: for $x \neq 1$, optimal change beats stasis.

Key results:

- Unique cost functional from three axioms
- Modal operators from cost structure

- Counterfactuals as unrealized finite-cost paths
- Machine verification in Lean 4

The deepest insight: the universe cannot afford stasis. Existence at $x = 1$ is not luck but economic necessity.

Acknowledgments

Thanks to the Recognition Science community and Mathlib maintainers.

References

[1] S. Kripke, "Semantical Considerations on Modal Logic," *Acta Philos. Fenn.*, 16:83–94, 1963.

[2] D. Lewis, *Counterfactuals.* Harvard Univ. Press, 1973.

[3] R. Stalnaker, "A Theory of Conditionals," in *Studies in Logical Theory.* Blackwell, 1968.

[4] S. Amari, *Information Geometry and Its Applications.* Springer, 2016.

[5] K. Friston, "The free-energy principle: a unified brain theory?" *Nat. Rev. Neurosci.*, 11:127–138, 2010.

[6] R. P. Feynman, "Space-Time Approach to Non-Relativistic Quantum Mechanics," *Rev. Mod. Phys.*, 20:367–387, 1948.

[7] L. de Moura and S. Ullrich, "The Lean 4 Theorem Prover," in *CADE-28*, 2021.

[8] J. Washburn, "Recognition Science: Full Theory Specification," Tech. Rep., 2025.

A Proof of the d'Alembert Identity

Theorem A.1. $J(xy) + J(x/y) = 2J(x)J(y) + 2J(x) + 2J(y)$.

Proof. Direct computation. Let $a = x + x^{-1}$, $b = y + y^{-1}$. Then:

$$\begin{aligned} \text{LHS} &= \frac{1}{2}(xy + (xy)^{-1} + x/y + y/x) - 2 = \frac{ab}{2} - 2 \\ \text{RHS} &= 2\left(\frac{a}{2} - 1\right)\left(\frac{b}{2} - 1\right) + 2\left(\frac{a}{2} - 1\right) + 2\left(\frac{b}{2} - 1\right) \\ &= \frac{ab}{2} - a - b + 2 + a - 2 + b - 2 = \frac{ab}{2} - 2 \end{aligned}$$

□