The Geometry of Decision:
A Cost-Theoretic Framework for Attention, Choice, and Agency

Jonathan Washburn*

December 2025

Abstract

We develop a mathematical framework for decision-making based on the universal cost functional $J(x) = \frac{1}{2}(x + x^{-1}) - 1$. We introduce an *Attention Operator* as a capacity-limited gating function, and propose that cognitive capacity bounds follow from φ-scaling in the Recognition Science framework. We define the *Choice Manifold* M_{choice} as a Riemannian manifold with metric derived from $J''(x) = x^{-3}$, and study cost-minimizing paths as a model for decisions. We characterize deliberation as gradient descent with exploration noise, and interpret free will as path selection in regions where the cost landscape is locally flat. The framework yields testable predictions relating deliberation time to decision stakes.

Keywords: decision theory, attention capacity, Riemannian geometry, cost minimization, cognitive limits

1 Introduction

The mathematical structure of decision-making remains incompletely understood despite decades of research. Drift-diffusion models [3, 4] successfully describe reaction time distributions but lack a principled account of capacity limits. Information-theoretic approaches [5] characterize channel capacity but do not specify the underlying dynamics. This paper proposes a geometric framework that addresses both structure and dynamics.

Our approach builds on the *Recognition Science* (RS) framework [1], which posits that physical systems minimize a universal cost functional. The unique cost function satisfying the d'Alembert composition law $J(xy) = J(x) + J(y) + J(x)J(y)$, with normalization $J(1) = 0$ and calibration $J''(1) = 1$, is:

$$J(x) = \frac{1}{2}\left(x + \frac{1}{x}\right) - 1, \quad x > 0. \tag{1}$$

This paper extends RS to cognitive decision-making. We show that:

1. The Hessian of J defines a natural Riemannian metric on the space of possible states.

---

*Recognition Science Research Institute. Email: `jonathan@recognitionscience.org`

2. Cost-minimizing paths in this metric provide a model for optimal decisions.

3. Capacity bounds emerge from φ-scaling, where $\varphi = (1 + \sqrt{5})/2$ is the golden ratio.

4. Free will can be understood as path selection in flat regions of the cost landscape.

The paper is organized as follows. Section 2 establishes the properties of J. Section 3 develops the attention capacity model. Section 4 constructs the Choice Manifold and studies its geodesics. Section 5 analyzes cost-minimizing paths. Section 6 models deliberation dynamics. Section 7 addresses free will. Section 8 presents empirical predictions. Section 10 concludes.

2 The Universal Cost Functional

2.1 Axiomatic Derivation

The cost functional (1) is not arbitrary but is uniquely determined by three axioms.

Theorem 2.1 (Cost Uniqueness [2]). *Let $J : \mathbb{R}^+ \to \mathbb{R}$ satisfy:*

1. **Composition**: $J(xy) = J(x) + J(y) + J(x)J(y)$

2. **Normalization**: $J(1) = 0$

3. **Calibration**: $J''(1) = 1$

Then $J(x) = \frac{1}{2}(x + x^{-1}) - 1$.

Proof. The composition law implies $J(x^n) = (1 + J(x))^n - 1$ for $n \in \mathbb{Z}$. Setting $J(x) = \cosh(\theta) - 1$ where $x = e^\theta$ satisfies the functional equation. Calibration fixes the scale. $\square$

2.2 Properties

Proposition 2.2 (Properties of J). *The cost functional has the following properties:*

1. **Non-negativity**: $J(x) \geq 0$ *with equality iff $x = 1$.*

2. **Reciprocity**: $J(x) = J(1/x)$.

3. **Convexity**: $J''(x) = x^{-3} > 0$ *for $x > 0$.*

4. **Unique minimum**: $x = 1$ *is the global minimum.*

Proof. (1) By AM-GM: $\frac{1}{2}(x + 1/x) \geq 1$, with equality iff $x = 1$. (2) Direct substitution. (3) $J''(x) = x^{-3} > 0$. (4) Follows from (1) and (3). $\square$

The cost landscape has a single minimum at $x = 1$ (balance) with steep walls as $x \to 0^+$ or $x \to \infty$.

3 The Algebra of Attention

3.1 Attention as Gating

We model attention as a gating mechanism that determines which potential experiences become conscious.

Definition 3.1 (Attention Operator). Let $\mathcal{Q}$ denote the space of potential qualia (experiential states). The *Attention Operator* is a function

$$A : \mathcal{Q} \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathcal{Q}_{\text{conscious}} \cup \{\bot\}$$

where $A(q, c, \phi) = (q, c, \phi)$ if $c \geq c_{\min}$ and $\phi > 0$, otherwise $A(q, c, \phi) = \bot$ (unconscious).

The threshold $c_{\min}$ represents the minimum recognition cost for conscious awareness. In the RS framework, we conjecture $c_{\min} = 1$ based on the Gap-45 analysis.

3.2 The Capacity Bound Hypothesis

A central claim of this paper is that attention capacity is not arbitrary but emerges from the mathematical structure of RS.

Conjecture 3.2 (Capacity Bound). *The total attention capacity is bounded by $\varphi^3 \approx 4.236$, where $\varphi = (1 + \sqrt{5})/2$.*

This conjecture is motivated by several observations within RS:

1. The golden ratio φ minimizes J locally along specific constraint surfaces.

2. Powers of φ appear naturally as eigenvalues of RS operators.

3. The value $\varphi^3 \approx 4.236$ is consistent with subitizing limits (typically 3–4 items).

3.3 Derivation of Miller's Law

Theorem 3.3 (Miller's Law). *If the capacity bound is $C_{\max} = \varphi^3$ and the minimum intensity per item is $\epsilon_{\min} = \varphi^{-1}$, then the maximum number of simultaneously attended items is*

$$n_{\max} = \left\lfloor \frac{C_{\max}}{\epsilon_{\min}} \right\rfloor = \lfloor \varphi^4 \rfloor = 6.$$

With chunking, the effective capacity reaches 7 ± 2.

Proof. Direct computation: $\varphi^4 = \varphi^3 \cdot \varphi \approx 4.236 \times 1.618 \approx 6.85$. The floor is 6. Chunking compresses information by estimated factor 1.5, giving $6 \times 1.5 = 9$ as the upper limit. $\square$

Remark 3.4. The parameter $\epsilon_{\min} = \varphi^{-1}$ is a hypothesis, not a derivation. Alternative values would yield different predictions. The specific value φ^{-1} is motivated by the φ-ladder structure in RS but requires independent empirical validation.

3.4 Attentional Phenomena

The framework accounts for several known phenomena:

Inattentional Blindness: When capacity is saturated ($\sum_i \phi_i = C_{\max}$), new stimuli cannot enter consciousness regardless of their salience.

Attentional Blink: After processing a target, capacity recovers gradually over ≈ 300 ms, blocking secondary targets.

Subitizing: Below the capacity bound, enumeration is parallel and fast; above it, serial counting is required.

4 The Choice Manifold

4.1 The Metric Structure

The Hessian of J provides a natural metric for the space of decision states.

Definition 4.1 (Choice Manifold)**.** The *Choice Manifold* is the Riemannian manifold $M_{\text{choice}} = ((0, \infty), g)$ with metric

$$g(x) = J''(x) = \frac{1}{x^3}. \tag{2}$$

Proposition 4.2 (Well-Defined Metric)**.** *The function $g(x) = x^{-3}$ defines a positive-definite Riemannian metric on $(0, \infty)$.*

Proof. For $x > 0$, we have $g(x) = x^{-3} > 0$. The induced inner product on tangent spaces is positive-definite. □

4.2 Arc Length and Geodesics

The arc length of a path $\gamma : [a, b] \to (0, \infty)$ is:

$$L[\gamma] = \int_a^b \sqrt{g(\gamma(t))}\, |\gamma'(t)|\, \mathrm{d}t = \int_a^b \frac{|\gamma'(t)|}{\gamma(t)^{3/2}}\, \mathrm{d}t. \tag{3}$$

Theorem 4.3 (Geodesic Equation)**.** *Geodesics in (M_{choice}, g) satisfy the second-order ODE:*

$$\gamma''(t) + \Gamma(\gamma(t)) \cdot (\gamma'(t))^2 = 0 \tag{4}$$

where the Christoffel symbol is

$$\Gamma(x) = \frac{g'(x)}{2g(x)} = \frac{-3x^{-4}}{2x^{-3}} = -\frac{3}{2x}. \tag{5}$$

Proof. Standard Riemannian geometry. In 1D, $\Gamma = \frac{1}{2g}\frac{dg}{dx}$. □

4.3 Explicit Geodesic Solution

Theorem 4.4 (Geodesic Form)**.** *The geodesics of (M_{choice}, g) are:*

$$\gamma(t) = (At + B)^2 \tag{6}$$

for constants A, B with $At + B > 0$.

Proof. Substituting $\gamma = u^2$ with $u = At + B$:

$$\gamma' = 2u \cdot u' = 2A \cdot u$$

$$\gamma'' = 2A \cdot u' = 2A^2$$

The geodesic equation (4) becomes:

$$2A^2 - \frac{3}{2u^2} \cdot (2Au)^2 = 2A^2 - \frac{3}{2u^2} \cdot 4A^2u^2 = 2A^2 - 6A^2 = -4A^2$$

This does not vanish, so $\gamma = u^2$ is not a geodesic. Let us solve correctly.
Correct derivation: Let $v = \gamma'$. The geodesic equation gives:

$$\frac{dv}{dt} = \frac{3v^2}{2\gamma}$$

Using $\frac{dv}{d\gamma} = \frac{1}{v}\frac{dv}{dt}$:

$$v\frac{dv}{d\gamma} = \frac{3v^2}{2\gamma} \implies \frac{dv}{d\gamma} = \frac{3v}{2\gamma}$$

Separating: $\frac{dv}{v} = \frac{3}{2}\frac{d\gamma}{\gamma}$, giving $\ln|v| = \frac{3}{2}\ln|\gamma| + C_1$, hence $v = C_2\gamma^{3/2}$.
Thus $\gamma' = C_2\gamma^{3/2}$, i.e., $\gamma^{-3/2}d\gamma = C_2dt$. Integrating:

$$-2\gamma^{-1/2} = C_2t + C_3$$

$$\gamma^{-1/2} = -\frac{C_2t + C_3}{2} = \alpha t + \beta$$

$$\gamma(t) = \frac{1}{(\alpha t + \beta)^2}$$

for appropriate constants α, β. $\qquad\square$

Corollary 4.5 (Geodesic Family). *The geodesics of (M_{choice}, g) have the form:*

$$\gamma(t) = \frac{1}{(\alpha t + \beta)^2} \tag{7}$$

where $\alpha, \beta \in \mathbb{R}$ with $\alpha t + \beta \neq 0$.

5 Cost-Minimizing Paths

5.1 The Cost Functional

We consider a different optimization problem: minimizing the integrated cost along a path.

Definition 5.1 (Path Cost). The *integrated cost* of a path $\gamma : [0, T] \to (0, \infty)$ is:

$$\mathcal{C}[\gamma] = \int_0^T J(\gamma(t)) \, \mathrm{d}t. \tag{8}$$

Remark 5.2. This is distinct from arc length (3). Geodesics minimize arc length; cost-minimizing paths minimize $\mathcal{C}$.

5.2 Euler-Lagrange Equation

Proposition 5.3 (Cost-Minimizing Paths). *Extremals of $\mathcal{C}[\gamma]$ with fixed endpoints satisfy:*

$$J'(\gamma(t)) = 0 \quad \Leftrightarrow \quad 1 - \gamma(t)^{-2} = 0 \quad \Leftrightarrow \quad \gamma(t) = 1. \tag{9}$$

Proof. Since the Lagrangian $J(\gamma)$ does not depend on γ', the Euler-Lagrange equation is $\frac{\partial J}{\partial \gamma} = 0$, giving $J'(\gamma) = 0$. $\square$

Corollary 5.4 (Optimal Path). *The unique cost-minimizing path is $\gamma(t) = 1$ for all t, with $\mathcal{C}[\gamma] = 0$.*

5.3 Constrained Optimization

In realistic decision-making, endpoints are fixed and the path must traverse from $\gamma(0) = x_0$ to $\gamma(T) = x_1$.

Definition 5.5 (Decision Cost). The *decision cost* from x_0 to x_1 is:

$$D(x_0, x_1; T) = \inf_{\gamma} \left\{ \int_0^T J(\gamma(t)) \, \mathrm{d}t : \gamma(0) = x_0, \gamma(1) = x_1 \right\}. \tag{10}$$

For the linear path $\gamma(t) = x_0 + t(x_1 - x_0)/T$, the cost is:

$$\mathcal{C}_{\text{linear}} = \int_0^T J(x_0 + s(x_1 - x_0)) \, ds \tag{11}$$

which can be computed numerically or approximated for small displacements.

5.4 Regret

Definition 5.6 (Regret). Given a chosen path γ_{actual} and the optimal geodesic γ_{opt} with the same endpoints, the *regret* is:

$$R = \mathcal{C}[\gamma_{\text{actual}}] - \mathcal{C}[\gamma_{\text{opt}}] \geq 0. \tag{12}$$

Regret measures the excess cost incurred by suboptimal navigation.

6 Deliberation Dynamics

6.1 Gradient Descent Model

We model deliberation as noisy gradient descent on the cost landscape.

Definition 6.1 (Deliberation Update). The state evolves according to:

$$x_{t+1} = x_t - \eta \cdot J'(x_t) + \sigma_t \xi_t \tag{13}$$

where $\eta > 0$ is the step size, σ_t is the noise magnitude, and $\xi_t \sim \mathcal{N}(0, 1)$.

Proposition 6.2 (Convergence). *For η sufficiently small and $\sigma_t \to 0$, the iteration (13) converges to $x^* = 1$.*

Proof. The function J is strictly convex on $(0, \infty)$ with unique minimum at $x = 1$. Standard results on stochastic gradient descent with vanishing noise guarantee convergence. $\square$

6.2 Exploration-Exploitation Tradeoff

Definition 6.3 (Annealing Schedule)**.** The noise magnitude follows a cooling schedule:

$$\sigma_t = \sigma_0 \cdot \left(1 - \frac{t}{T}\right)^\alpha \tag{14}$$

where T is the total deliberation time and $\alpha > 0$ controls the cooling rate.

High noise early in deliberation promotes exploration; decreasing noise later promotes exploitation. This is analogous to simulated annealing.

6.3 Thermodynamic Interpretation

Definition 6.4 (Decision Free Energy)**.** The *decision free energy* at inverse temperature β is:

$$F(\beta) = -\frac{1}{\beta} \ln \int e^{-\beta J(x)} \, dx. \tag{15}$$

At high temperature ($\beta \to 0$), the distribution is uniform (exploration). At low temperature ($\beta \to \infty$), the distribution concentrates at $x = 1$ (exploitation).

6.4 Time-Stakes Relationship

Proposition 6.5 (Deliberation Time Scaling)**.** *For decisions with stakes proportional to S, the optimal deliberation time scales as:*

$$T^* \propto \sqrt{S}. \tag{16}$$

Heuristic. The benefit of additional deliberation is $\sim S \cdot \epsilon$, where ϵ is the error reduction. The cost of deliberation is $\sim T$. Optimal stopping occurs when marginal benefit equals marginal cost. With error $\epsilon \sim 1/\sqrt{T}$ (from random walk intuition), we get $T^* \propto \sqrt{S}$. $\quad\square$

7 Free Will as Path Selection

7.1 Decision Points

Definition 7.1 (Decision Point)**.** A *decision point* is a state x such that there exist multiple distinct paths $\{\gamma_1, \ldots, \gamma_n\}$ with:

1. Common starting point: $\gamma_i(0) = x$ for all i.

2. Finite costs: $\mathcal{C}[\gamma_i] < \infty$ for all i.

3. Similar costs: $|\mathcal{C}[\gamma_i] - \mathcal{C}[\gamma_j]| < \epsilon$ for some small ϵ.

Definition 7.2 (Choice Degree)**.** The *choice degree* at a decision point is:

$$\chi = \frac{1}{1 + \mathrm{Var}(\mathcal{C}[\gamma_i])}. \tag{17}$$

High χ (near 1) indicates a flat cost landscape with many viable options.

7.2 The Compatibilist Interpretation

Proposition 7.3 (Compatibilism). *The framework is compatible with both determinism and agency:*

1. ***Determinism**: The cost functional J is fixed and objective.*

2. ***Agency**: At decision points with $\chi \approx 1$, the path selection is underdetermined by cost considerations alone.*

When the cost landscape is steep ($\chi \approx 0$), there is effectively one optimal path and little room for choice. When the landscape is flat ($\chi \approx 1$), multiple paths are nearly equally good, and the selection among them constitutes *genuine choice*.

Remark 7.4. This framework does not prove that free will exists in any metaphysical sense. Rather, it provides a mathematical setting in which the *experience* of choice has a natural interpretation: selection among near-optimal paths in flat regions of the cost landscape.

7.3 Responsibility and Regret

Definition 7.5 (Moral Responsibility). Responsibility for an action is graded by:

$$\text{Resp}(\gamma) = \text{Awareness} \times \chi \times \text{Selection} \times (1 - R) \tag{18}$$

where Awareness $\in [0, 1]$ is consciousness of options, χ is choice degree, Selection is explicit commitment, and R is normalized regret.

High responsibility requires: awareness of alternatives, a flat cost landscape (genuine choice), explicit selection, and low regret (coherent values).

8 Empirical Predictions

The framework generates several testable predictions.

8.1 Quantitative Predictions

Prediction 8.1 (Time-Stakes Scaling). Deliberation time increases sublinearly with stakes: $T \propto S^{0.5 \pm 0.1}$.

Protocol: Measure decision time for choices with varying monetary stakes (e.g., \$1 vs \$10 vs \$100 gambles with identical structure).

Expected: Log-log slope between 0.4 and 0.6.

Prediction 8.2 (Capacity Limit). Working memory capacity clusters around 4 ± 1 items, consistent with $\varphi^3 \approx 4.24$.

Protocol: Measure digit span, change detection, and subitizing limits across population.

Expected: Mean capacity $\in [3, 5]$ with low variance.

Prediction 8.3 (Exploration Decay). Eye movement entropy decreases over the course of deliberation.

Protocol: Track gaze patterns during multi-attribute decisions with varying time pressure.

Expected: Entropy drops by $\geq 30\%$ from first to last quintile.

8.2 Qualitative Predictions

Prediction 8.4 (Flat Landscape $\Rightarrow$ Slowness). Decisions with similar-valued options take longer than decisions with a clear winner.

   Interpretation: High χ extends deliberation as the system explores the flat region.

Prediction 8.5 (Regret Correlates with Path Deviation). Self-reported regret correlates with objective distance from the geodesic path.

   Protocol: Track decision trajectories and correlate with post-hoc regret ratings.

9 Discussion

9.1 Relation to Existing Models

Drift-Diffusion Models (DDM): Our gradient descent with noise (13) resembles DDM [4], with J' providing the drift and σ_t providing the diffusion. The key addition is the principled derivation of the cost function and the interpretation of the learning rate.

   ACT-R and Chunking: The capacity bound φ^3 is consistent with the chunk size in ACT-R [7]. The framework provides a principled reason for this particular value.

   Prospect Theory: The cost function J shares the property of being steeper for losses ($x < 1$) than for gains ($x > 1$, near the minimum). However, J is symmetric under $x \mapsto 1/x$, so the analogy is imperfect.

9.2 Limitations

Several limitations should be acknowledged:

1. **1D simplification**: The choice manifold $(0, \infty)$ is one-dimensional. Real decisions have higher-dimensional state spaces. Extension to $\mathbb{R}^n$ is straightforward mathematically but requires additional structure.

2. **Parameter dependence**: The Miller's Law derivation depends on $\epsilon_{\min} = \varphi^{-1}$, which is a hypothesis. Different values would yield different capacity predictions.

3. **Uncomputability claims**: Earlier versions of this work invoked "Gap-45 uncomputability" to protect free will. This claim requires further formalization and is not included as a theorem in the current paper.

4. **Empirical validation**: The predictions are testable but have not yet been tested. The framework's value depends on future empirical confirmation.

9.3 Future Directions

Several extensions are natural:

- **Multi-dimensional choice manifolds**: Replace $(0, \infty)$ with $\mathbb{R}^n_+$ and study the resulting Riemannian geometry.

- **Social decisions**: Model multi-agent decision-making as interacting cost functions.

- **Neural implementation**: Identify neural correlates of the cost function and deliberation dynamics.

- **Formal verification**: Complete the Lean formalization of all theorems.

10   Conclusion

We have presented a geometric framework for decision-making based on the universal cost functional $J(x) = \frac{1}{2}(x + x^{-1}) - 1$. The main contributions are:

1. **Geodesic structure**: The Choice Manifold $((0, \infty), g)$ with $g(x) = x^{-3}$ has geodesics of the form $\gamma(t) = (\alpha t + \beta)^{-2}$.

2. **Capacity bound**: The hypothesis that attention capacity is $\varphi^3 \approx 4.236$ yields predictions consistent with Miller's Law.

3. **Deliberation dynamics**: Gradient descent with cooling noise models the exploration-exploitation tradeoff.

4. **Free will interpretation**: Agency is path selection in flat regions of the cost landscape, compatible with determinism.

5. **Testable predictions**: Time-stakes scaling, capacity limits, and exploration decay are empirically verifiable.

The framework does not resolve the metaphysics of free will but provides a precise mathematical setting in which the structure of choice can be analyzed. Whether this structure is ultimately correct is an empirical question.

Acknowledgments

References

[1] J. Washburn, "Recognition Science: Deriving Physics from Cost Minimization," Recognition Science Technical Reports, 2025.

[2] J. Washburn, "The Universal Cost Functional from d'Alembert Composition," Recognition Science Technical Reports, 2025.

[3] R. Ratcliff, "A Theory of Memory Retrieval," *Psychological Review*, vol. 85, no. 2, pp. 59–108, 1978.

[4] R. Ratcliff and G. McKoon, "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks," *Neural Computation*, vol. 20, no. 4, pp. 873–922, 2008.

[5] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[6] G. A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Review*, vol. 63, no. 2, pp. 81–97, 1956.

[7] J. R. Anderson and C. Lebiere, *The Atomic Components of Thought*. Lawrence Erlbaum Associates, 1998.

[8] D. Kahneman and A. Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, vol. 47, no. 2, pp. 263–291, 1979.

[9] D. C. Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting*. MIT Press, 1984.

[10] N. Cowan, "The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity," *Behavioral and Brain Sciences*, vol. 24, no. 1, pp. 87–114, 2001.

A Derivation of the Geodesic Equation

For a 1D Riemannian manifold with metric $g(x)$, the arc length is:

$$L[\gamma] = \int \sqrt{g(\gamma(t))}\, |\gamma'(t)|\, dt$$

The Euler-Lagrange equation for arc length minimization gives:

$$\frac{d}{dt}\left(\frac{\gamma'}{\sqrt{g(\gamma)}|\gamma'|}\right) = \frac{g'(\gamma)|\gamma'|}{2\sqrt{g(\gamma)}}$$

For unit-speed parameterization ($|\gamma'| = 1/\sqrt{g(\gamma)}$), this simplifies to the geodesic equation with Christoffel symbol $\Gamma = g'/(2g)$.

B Numerical Values

Constant	Value	Interpretation
φ	1.618	Golden ratio
φ^2	2.618	Intensity threshold (hypothesized)
φ^3	4.236	Capacity bound (hypothesized)
φ^4	6.854	Miller upper bound
φ^{-1}	0.618	Minimum item intensity

Table 1: Key numerical values in the framework.

C   Lean Formalization

Partial formalization is available in Lean 4:

```
IndisputableMonolith/Decision/
  Attention.lean            -- Attention operator
  ChoiceManifold.lean       -- Riemannian structure
  FreeWill.lean             -- Agency definitions
  DeliberationDynamics.lean
  GeodesicSolutions.lean
  Decision.lean             -- Index module
```

Status: Core definitions complete; some theorems have `sorry` placeholders pending detailed proofs.