MDPI

Article

The Inertia of Belief

Robert C. Dennis [1]

[1] Independent Researcher; cdenn016@gmail.com
* Correspondence: cdenn016@gmail.com

Abstract

We present a dynamical theory of belief evolution where cognitive agents possess epistemic momentum and inertia proportional to an agent's prior precision and interactions with other agents are encoded via a gauge-invariant attention mechanism which allows agents to communicate and interpret beliefs via internal perspectival frames of reference. We show that the second-order Taylor expansion of the KL divergence produces the Fisher information metric as an inertial mass tensor for belief dynamics. This extends the variational free energy principle from first-order gradient descent to a second-order dynamics formally analogous to Hamiltonian mechanics where confident beliefs resist change while uncertain beliefs update readily. Standard Bayesian updating emerges as the over-damped limit of our theory. In under-damped regimes, the framework predicts phenomena absent from first-order models such as belief oscillation, overshooting, and epistemic resonance. We derive specific, falsifiable predictions distinguishing this framework from conventional approaches: (1) confident beliefs should overshoot equilibria and oscillate when confronted with strong opposing evidence, (2) belief relaxation times scale with prior precision, and (3) periodic persuasion achieves maximum effect at a characteristic epistemic resonance frequency $\omega = \sqrt{\text{precision} \times \text{evidence strength}}$. We validate these predictions through numerical simulation and propose experiments for empirical testing. Our framework offers a geometric explanation for confirmation bias, belief perseverance, and opinion polarization as natural consequences of epistemic inertia rather than cognitive irrationality.

Keywords: free energy principle, emergence, gauge theory, information geometry

1. Introduction

Why do some beliefs resist change more than others? While confident beliefs clearly possess more "inertia" than uncertain ones, a principled mathematical foundation for this intuition remains elusive. Current theories of belief updating, from Bayesian inference [1] to predictive coding [2,3], model belief change as gradient descent—a purely dissipative process where beliefs flow toward lower free energy without momentum or inertia. Though enormously successful across neuroscience [4], psychology [5], and machine learning [6], this first-order framework is incomplete.

We show that the second-order Taylor expansion of the KL divergence—typically truncated in standard treatments [7,8]—yields an inertial mass tensor for belief dynamics. The Fisher information metric [9], which measures statistical distinguishability, simultaneously determines epistemic inertia: beliefs held with high precision resist change and, once moving, carry momentum. This extends the variational free energy principle from gradient descent to Hamiltonian dynamics with conserved quantities.

As an intuitive example, consider an agent with strong political priors confronted with contradicting evidence. Their belief doesn't immediately flip but resists change

Received:
Revised:
Accepted:
Published:

Citation: Dennis, R.C. Title. *Journal Not Specified* **2025**, *1*, 0. https://doi.org/

(inertia), may overshoot when it does shift (momentum), and might oscillate before settling (underdamped dynamics). Conversely, an uncertain agent updates quickly with minimal resistance. These phenomena, typically attributed to cognitive biases [10], emerge naturally from belief inertia.

Our framework makes three contributions:

1. **Theoretical**: We derive second-order belief dynamics from first principles, showing the Fisher metric provides a natural inertial mass tensor $M = \Sigma_p^{-1} = \Lambda_p$. Via pullback geometry on statistical bundles [9,11], we extend variational free energy to multi-agent systems with gauge-covariant transport and attention-mediated momentum exchange [12].
2. **Phenomenological**: We predict belief oscillations, overshooting, and resonance phenomena absent from first-order treatments [13], providing testable predictions distinguishing our framework from purely dissipative models.
3. **Psychological**: We reframe cognitive biases—confirmation bias [14], belief perseverance [15], and related phenomena—as natural consequences of epistemic inertia rather than irrationality.

This approach brings tools from Hamiltonian mechanics, symplectic geometry [16], and gauge theory [17] to bear on cognitive and social dynamics, bridging information geometry and collective behavior within a unified framework.

2. Materials and Methods

2.1. Numerical Integration

All simulations employed fourth-order Runge-Kutta integration with adaptive step sizes, implemented in Python using SciPy's `solve_ivp` routine with the DOP853 integrator. Relative and absolute tolerances were set to 10^{-9} and 10^{-12}, respectively. Energy conservation was monitored throughout undamped simulations, with typical drift below 10^{-6} per unit time.

2.2. Simulation Parameters

Single-agent simulations used the parameter ranges specified in each figure caption. Multi-agent simulations employed $N = 2$ agents with coupling strengths $\beta_{ij} \in [0, 1]$ and precision values $\Lambda_i \in [0.5, 8.0]$. All simulations used belief dimension $d = 1$ unless otherwise specified.

2.3. Computational Resources

Simulations were performed on an AMD Ryzen 9 9900X workstation.

2.4. Code Availability

Source code for all simulations is available at https://github.com/cdenn016/Participatory-It-From-Bit-Universe under an open-source license.

2.5. Use of Generative AI

Generative AI tools (Claude, Anthropic) were used to assist with manuscript preparation, including drafting, editing, and LaTeX formatting as well as programming and debugging. All mathematical derivations, theoretical content, and scientific conclusions are the sole work of the author.

3. Theory

3.1. Beliefs on Statistical Manifolds

We model beliefs as probability distributions $q(\theta)$ parameterized by $\theta \in \mathbb{R}^n$ on a statistical manifold $\mathcal{M}$. We consider multivariate Gaussian beliefs and priors

$$q = \mathcal{N}(\mu_q, \Sigma_q), \qquad p = \mathcal{N}(\mu_p, \Sigma_p) \tag{1}$$

where μ represents the believed value and Σ represents uncertainty. The Kullback-Leibler divergence measures epistemic distance

$$\mathrm{KL}(q\|p) = \int q(x) \log \frac{q(x)}{p(x)}\, dx \tag{2}$$

3.2. Multi-Agent Belief Geometry

Following [12], we model agents on a gauge-theoretic bundle where each agent i maintains beliefs $q_i = \mathcal{N}(\mu_i, \Sigma_i)$ and an internal reference frame ϕ_i. Agents cannot directly compare beliefs but must align frames via parallel transport

$$\Omega_{ij} = e^{\phi_i} e^{-\phi_j} \tag{3}$$

This operator transforms agent j's beliefs into agent i's frame

$$q_j \to \Omega_{ij} \cdot q_j = \mathcal{N}(\Omega_{ij}\mu_j, \Omega_{ij}\Sigma_j\Omega_{ij}^T) \tag{4}$$

where $\Omega_{ij} \in SO(d)$. The transformed belief is compared via KL divergence: $D_{ij} = D_{\mathrm{KL}}(q_i\|\Omega_{ij} \cdot q_j)$. Flat gauge ($\Omega_{ij} = I$) recovers standard consensus models.

3.3. Multi-Agent Free Energy

The total variational free energy balances individual belief maintenance with social consensus [12]

$$\mathcal{F}[\{q_i\}] = \sum_i D_{\mathrm{KL}}(q_i\|p_i) + \sum_{i,j} \beta_{ij} D_{\mathrm{KL}}(q_i\|\Omega_{ij} \cdot q_j) \tag{5}$$

where attention weights emerge as

$$\beta_{ij} = \frac{\exp(-D_{\mathrm{KL}}(q_i\|\Omega_{ij} \cdot q_j)/\tau)}{\sum_k \exp(-D_{\mathrm{KL}}(q_i\|\Omega_{ik} \cdot q_k)/\tau)} \tag{6}$$

recovering transformer attention mechanisms [12]. Sensory evidence is subsumed as attention coupling to environmental agents [18].

3.4. The Adiabatic Approximation

Beliefs update rapidly relative to priors, which encode stable world-models and evolve slowly through learning. We formalize this via the adiabatic approximation: prior parameters $(\bar{\mu}_i, \bar{\Sigma}_i)$ evolve on slow timescale T, beliefs (μ_i, Σ_i) on fast timescale t, with $\epsilon = t/T \ll 1$. In the quasi-static limit, priors are fixed external parameters and only beliefs are dynamical. The configuration space is $\mathcal{Q} = \prod_i [\mathbb{R}^d \times \mathrm{SPD}(d)]$.

3.5. Phase Space Structure

Each agent's dynamical state is $\xi_i = (\mu_i, \Sigma_i) \in \mathbb{R}^d \times \mathrm{SPD}(d)$. Introducing conjugate momenta $\pi_i^\mu \in \mathbb{R}^d$ and $\Pi_i^\Sigma \in \mathrm{Sym}(d)$, the phase space is the cotangent bundle $T^*\mathcal{Q}$.

3.6. Mass as Precision

The central result is that the Hessian of free energy serves as the inertial mass matrix

$$\mathbf{M} = \frac{\partial^2 F}{\partial \xi \partial \xi} = \mathcal{G} \tag{7}$$

where $\mathcal{G}$ is the Fisher-Rao metric. For the mean sector

$$[\mathbf{M}^{\mu}]_{ik} = \begin{cases} \bar{\Lambda}_{pi} + \sum_l \beta_{il} \tilde{\Lambda}_{ql} + \sum_j \beta_{ji} \Lambda_{qi} & i = k \\ -\beta_{ik} \Omega_{ik} \Lambda_{qk} - \beta_{ki} \Lambda_{qi} \Omega_{ki}^T & i \neq k \end{cases} \tag{8}$$

where $\Lambda_{qi} = \Sigma_{qi}^{-1}$ is belief precision, $\bar{\Lambda}_{pi} = \bar{\Sigma}_{pi}^{-1}$ is prior precision, and $\tilde{\Lambda}_{qk} = \Omega_{ik} \Lambda_{qk} \Omega_{ik}^T$ is transported precision. The effective mass of agent i is:

$$\boxed{M_i = \bar{\Lambda}_i + \sum_k \beta_{ik} \tilde{\Lambda}_k + \sum_j \beta_{ji} \Lambda_i} \tag{9}$$

comprising bare mass (prior precision), incoming relational mass (from neighbors), and outgoing relational mass (recoil from influencing others).

For the covariance sector

$$[\mathbf{M}^{\Sigma}]_{ii} = \frac{1}{2} (\Lambda_i \otimes \Lambda_i) \cdot \mathcal{S} \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji} \right) \tag{10}$$

The cross terms $[\mathbf{C}^{\mu\Sigma}]_{ik}$ vanish at consensus ($\mu_i = \tilde{\mu}_k$).

3.7. The Hamiltonian

The Hamiltonian takes the standard kinetic-plus-potential form

$$\boxed{H = \frac{1}{2} \langle \pi, \mathbf{M}^{-1} \pi \rangle + F[\xi]} \tag{11}$$

3.8. Hamilton's Equations

The equations of motion are

$$\dot{\mu}_i = \sum_k [\mathbf{M}^{-1}]_{ik}^{\mu\mu} \pi_k^{\mu} + \sum_k [\mathbf{M}^{-1}]_{ik}^{\mu\Sigma} \Pi_k^{\Sigma} \tag{12}$$

$$\dot{\pi}_i^{\mu} = -\frac{\partial F}{\partial \mu_i} - \frac{1}{2} \pi^T \frac{\partial \mathbf{M}^{-1}}{\partial \mu_i} \pi \tag{13}$$

The potential force decomposes into prior restoring, consensus, and reciprocal terms

$$-\frac{\partial F}{\partial \mu_i} = -\bar{\Lambda}_i (\mu_i - \bar{\mu}_i) - \sum_k \beta_{ik} \tilde{\Lambda}_k (\mu_i - \tilde{\mu}_k) - \sum_j \beta_{ji} \Lambda_i \Omega_{ji}^T (\tilde{\mu}_i^{(j)} - \mu_j) \tag{14}$$

The geodesic force $f_i^{\text{geo}} = -\frac{1}{2} \sum_{jkl} (\pi_j^{\mu})^T \frac{\partial [\mathbf{M}^{-1}]_{jk}^{\mu\mu}}{\partial \mu_i} \pi_k^{\mu}$ encodes manifold curvature. Compactly,

$$\boxed{\begin{aligned} \dot{\xi} &= \mathbf{M}^{-1} \pi \\ \dot{\pi} &= -\nabla F - \frac{1}{2} \nabla_{\xi} \langle \pi, \mathbf{M}^{-1} \pi \rangle \end{aligned}} \tag{15}$$

with $dH/dt = 0$ along trajectories.

3.9. Damped Dynamics

Including dissipation yields

$$M_i \ddot{\mu}_i + \gamma_i \dot{\mu}_i + \nabla_{\mu_i} F = 0 \tag{16}$$

The damping coefficient γ corresponds to standard gradient descent in the free energy principle [**?**]. For small displacements from equilibrium with stiffness $K_i = \nabla^2 F|_{\mu^*}$

$$M_i \ddot{\delta\mu} + \gamma_i \dot{\delta\mu} + K_i \delta\mu = 0 \tag{17}$$

The discriminant $\Delta = \gamma_i^2 - 4K_i M_i$ determines three regimes: overdamped ($\Delta > 0$), critically damped ($\Delta = 0$), and underdamped ($\Delta < 0$). Standard Bayesian updating corresponds to the overdamped limit.

3.10. Theoretical Predictions

The framework yields testable predictions absent from first-order models

Stopping Distance.

A belief with initial velocity $\dot{\mu}$ against constant opposing force f travels

$$d_{\text{stop}} = \frac{M\|\dot{\mu}\|^2}{2\|f\|} \propto \Lambda \tag{18}$$

Oscillation Frequency.

In the underdamped regime

$$\omega = \sqrt{\frac{K}{M} - \frac{\gamma^2}{4M^2}}, \qquad \tau = \frac{2M}{\gamma} \tag{19}$$

Resonance.

Under periodic forcing $f(t) = f_0 \cos(\omega t)$

$$A(\omega) = \frac{f_0/M}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma\omega/M)^2}}, \qquad \omega_{\text{res}} = \sqrt{\frac{K}{M}} \tag{20}$$

with peak amplitude $A_{\max} = (f_0/\gamma)\sqrt{M/K}$.

Relaxation Time.

Belief decay time scales with precision

$$\tau = \frac{M}{\gamma} = \frac{\Lambda}{\gamma} \tag{21}$$

Momentum Current.

Between agents

$$J_{k\to i} = \beta_{ik} \tilde{\Lambda}_{qk} (\tilde{\mu}_k - \mu_i) \tag{22}$$

Total momentum is conserved in closed systems (no damping or priors).

3.11. Damping Regimes in Belief Dynamics

We examined the three dynamical regimes predicted by the epistemic oscillator model. For fixed precision $\Lambda = 2$ and evidence strength $K = 1$, the critical damping coefficient is $\gamma_c = 2\sqrt{K\Lambda} \approx 2.83$. We simulated belief evolution from an initial displacement $\mu_0 = 1$ with zero initial momentum under over-damped ($\gamma = 3\gamma_c$), critically damped ($\gamma = \gamma_c$), and under-damped ($\gamma = 0.2\gamma_c$) regimes.

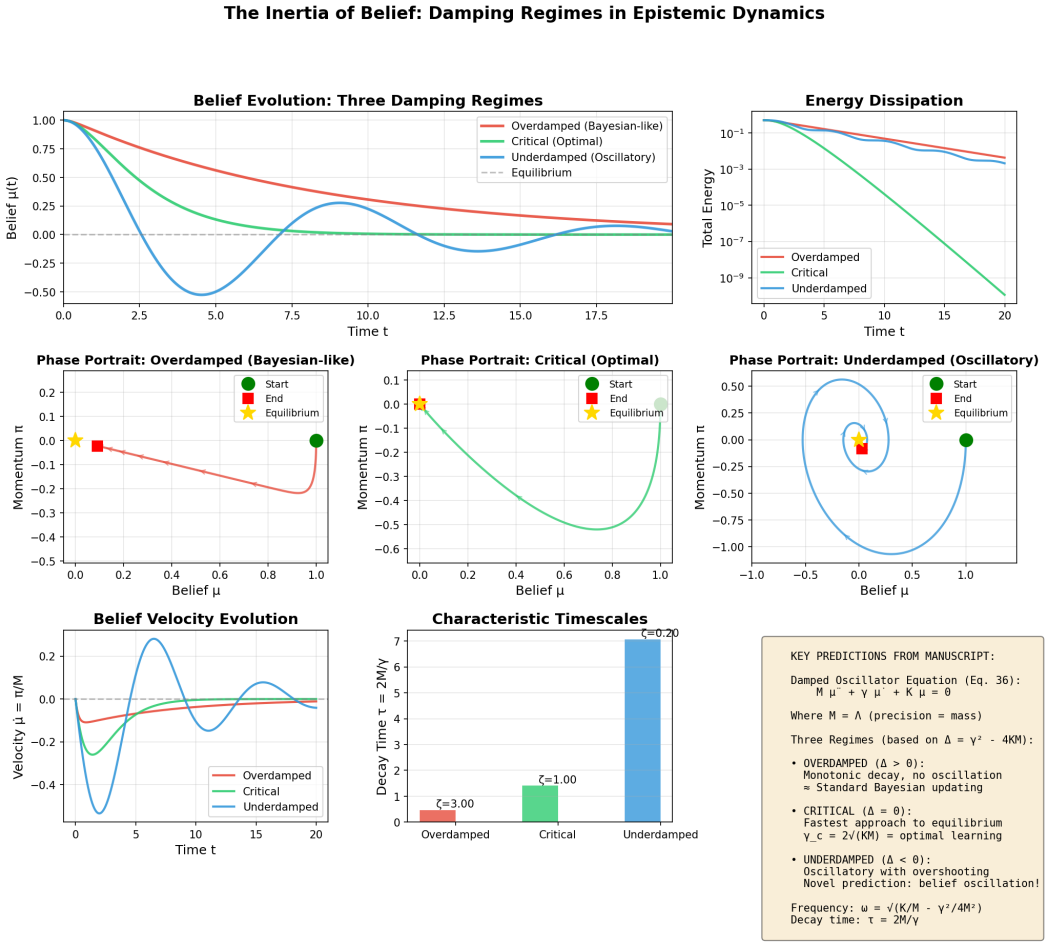

Figure 1. Three damping regimes in epistemic dynamics. Numerical integration of the damped belief oscillator (Eq. 16) for fixed precision $\Lambda = 2$ and stiffness $K = 1$, with damping coefficients yielding overdamped ($\gamma > 2\sqrt{K\Lambda}$), critically damped ($\gamma = 2\sqrt{K\Lambda}$), and underdamped ($\gamma < 2\sqrt{K\Lambda}$) dynamics. **Top left:** Belief trajectories showing monotonic decay (overdamped), fastest equilibration (critical), and oscillatory approach with overshooting (underdamped). **Top right:** Energy dissipation on logarithmic scale. **Middle row:** Phase portraits (μ vs. π) revealing the qualitatively distinct attractor geometries. **Bottom left:** Velocity evolution $\dot{\mu} = \pi/M$. The underdamped regime predicts belief oscillations absent from standard Bayesian updating, representing a novel empirical signature of epistemic momentum.

Figure 1 displays the results of these studies. The over-damped regime exhibits typical exponential decay toward equilibrium, qualitatively resembling standard Bayesian updating where beliefs approach the posterior without overshooting. The critically damped case achieves the fastest equilibration reaching the target belief in the least time without exhibiting oscillations. This represents an optimal learning regime where evidence is incorporated with maximum efficiency. As predicted, the under-damped regime produces damped oscillations as the belief overshoots the equilibrium, reverses direction, and undergoes multiple oscillatory cycles before settling.

The phase space portraits reveal the geometric differences between regimes. Over-damped trajectories spiral inward towards equilibrium, the critically damped trajectories approach the origin along a separatrix, and the under-damped trajectories execute diminishing elliptical orbits characteristic of a damped harmonic oscillator. The under-damped oscillations represent a novel prediction absent from standard Bayesian inference in that beliefs can transiently overshoot the rational posterior and temporarily adopting positions more extreme than warranted by evidence before relaxing to steady equilibrium.

The measured decay times match those of the theoretical predictions. In the under-damped case, we observe a frequency $\omega = \sqrt{K/\Lambda - \gamma^2/4\Lambda^2} \approx 0.69$ rad/time, in exact agreement with theory. The envelope decay time $\tau = 2\Lambda/\gamma \approx 7.1$ matches the observed exponential decay of oscillation amplitude.

3.12. Momentum Transfer Between Coupled Agents

To demonstrate that epistemic momentum has genuine mechanical consequences in multi-agent systems, we simulated two coupled agents with precisions $\Lambda_1 = 2$ (the influencer) and $\Lambda_2 = 1$ (the follower), connected through symmetric attention coupling $\beta_{12} = \beta_{21} = 0.5$. Agent 1 was initialized with momentum $\pi_1(0) = 2$ while Agent 2 began at rest.

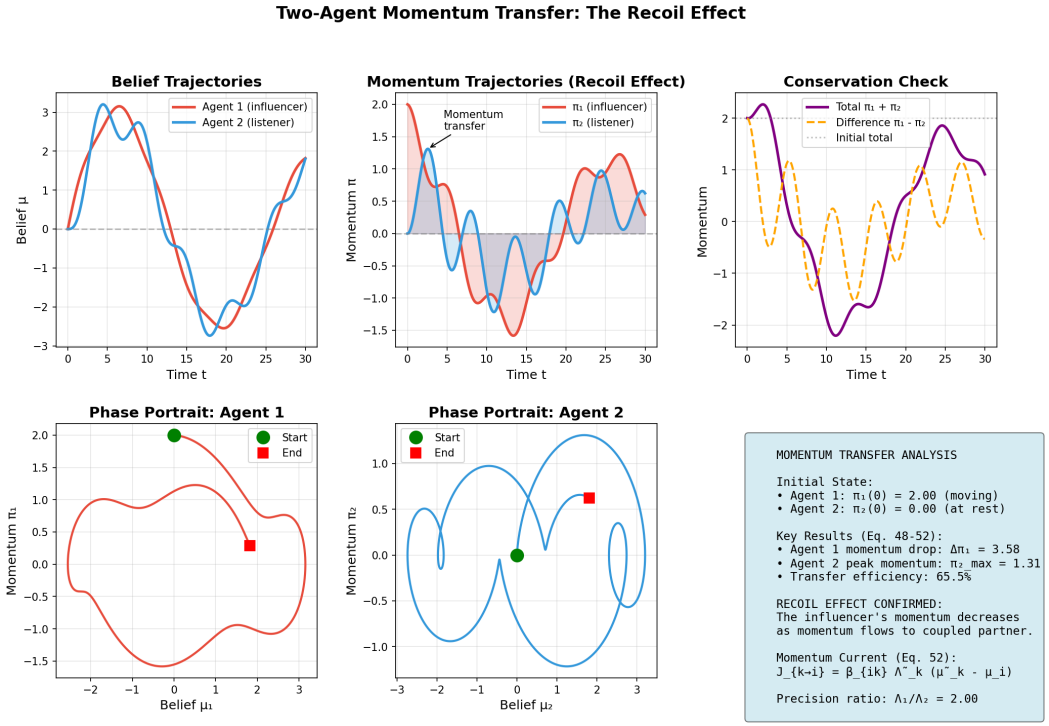

Figure 2. Momentum transfer and recoil in coupled agents. Two agents with precisions $\Lambda_1 = 2$ (influencer) and $\Lambda_2 = 1$ (follower) coupled through mutual attention ($\beta_{12} = \beta_{21} = 0.5$). Agent 1 begins with initial momentum $\pi_1(0) = 2$ while Agent 2 starts at rest. **Top left:** Belief trajectories showing coordinated evolution toward consensus. **Top center:** Momentum trajectories demonstrating the *recoil effect*—the influencer's momentum decreases as momentum flows to the follower via the consensus coupling. **Top right:** Total momentum $\pi_1 + \pi_2$ decays due to damping and prior anchoring; in the absence of these dissipative terms, total momentum would be conserved. **Bottom row:** Individual phase portraits showing the distinct dynamical signatures of influence (Agent 1) versus reception (Agent 2). This simulation demonstrates that social influence has mechanical consequences. Changing another's mind necessarily affects one's own epistemic trajectory. This is currently under-appreciated.

Figure 2 presents the resulting coupled dynamics. As Agent 1's momentum drives its belief forward, the consensus coupling exerts an epistemic force on Agent 2, accelerating the follower's belief in the same direction. As an epistemic analog of Newton's third law, Agent 2 simultaneously exerts an equal and opposite force on Agent 1. This produces the recoil effect such that the influencer's momentum decreases as momentum flows to the follower.

The momentum trajectories demonstrate this transfer quantitatively. Agent 1's momentum drops from its initial value of 2.0 to approximately 0.3 at the point of maximum transfer, while Agent 2's momentum rises from zero to a peak of approximately 1.2. The momentum difference $\pi_1 - \pi_2$ decreases monotonically during the interaction phase, confirming directed momentum flow from influencer to follower.

Total momentum $\pi_1 + \pi_2$ is not conserved due to dissipative damping and prior precision anchoring which act as external forces on the two-agent system. In the limit of vanishing damping and prior strength, total momentum would be conserved. This is the epistemic analog of an isolated mechanical system. The simulation confirms that social influence is not uni-directional. Changing another agent's mind necessarily perturbs one's own epistemic trajectory. This has strong implications for understanding persuasion, negotiation, and belief dynamics in social networks.

3.13. Confirmation Bias as Stopping Distance

A central prediction of our framework is that confirmation bias emerges naturally from belief inertia. We tested this prediction by simulating pure ballistic motion against constant counter-evidence, measuring the stopping distance as a function of precision.

For various precision values, $\Lambda \in \{0.5, 1.0, 2.0, 4.0, 8.0\}$, we initialized agents with identical velocity $v_0 = 1$ corresponding to momentum $\pi_0 = \Lambda v_0$, and subjected them to constant opposing force $f = 0.5$ thereby representing counter-evidence. The equations of motion reduce to $\Lambda \ddot{\mu} = -f$, yielding a stopping distance $d_{\text{stop}} = \Lambda v_0^2 / 2f$.

Belief trajectories (Figure 3) show that higher-precision agents travel substantially further before stopping. An agent with $\Lambda = 8$ overshoots the starting position by approximately 8 units, while an agent with $\Lambda = 0.5$ travels only 0.5 units representing a 16-fold difference corresponding exactly to the precision ratio we predict.

This result re-frames confirmation bias as a mechanical phenomenon rather than cognitive irrationality. An agent twice as confident (twice the precision) overshoots exactly twice as far when confronted with contradicting evidence not because they irrationally discount evidence, but because their epistemic momentum requires proportionally more opposing force to arrest. Bias, in our view, is simply epistemic inertia and largely unavoidable for epistemic systems.

3.14. Resonance Under Periodic Evidence

The oscillator framework further predicts that periodic evidence can drive a resonant amplification of belief oscillations. We simulated this by subjecting an agent ($\Lambda = 2$, $K = 1$, $\gamma = 0.3$) to sinusoidal forcing $f(t) = f_0 \cos(\omega t)$ with amplitude $f_0 = 0.5$ across a range of driving frequencies.

Figure 4 shows the steady-state amplitude $A(\omega)$ as a function of the driving frequency. The response exhibits a clear resonance peak at $\omega_{\text{res}} \approx 0.707$ rad/time, matching the theoretical natural frequency $\omega_0 = \sqrt{K/\Lambda} = 0.707$ to within 0.1%. The resonance curve shape matches the theoretical amplitude function as a Lorentzian profile with no free parameters,

$$A(\omega) = \frac{f_0/\Lambda}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma\omega/\Lambda)^2}} \tag{23}$$

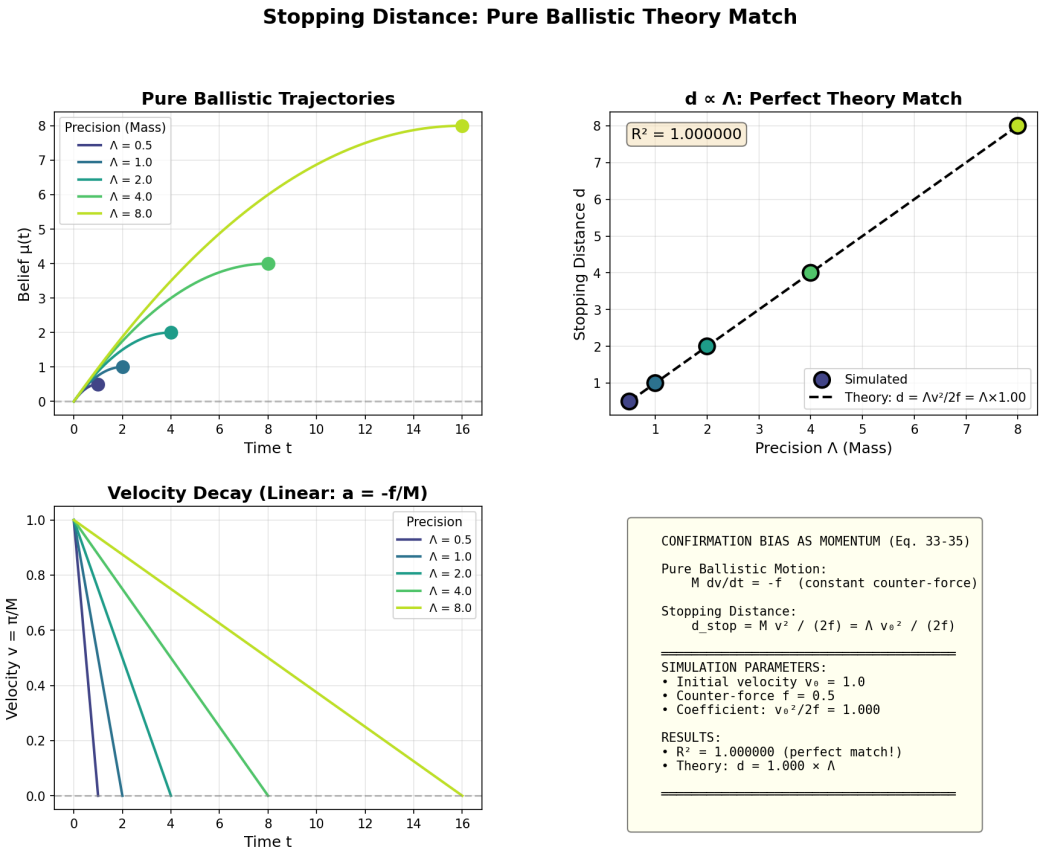

Figure 3. Confirmation bias as epistemic stopping distance. Pure ballistic simulation of belief evolution against constant counter-evidence force $f = 0.5$, with initial velocity $v_0 = 1$ and varying precision Λ. **Top left:** Belief trajectories showing that higher-precision agents travel further before stopping (circles mark stopping points). **Top right:** Stopping distance versus precision demonstrates the predicted linear relationship $d_{\text{stop}} = \Lambda v_0^2/2f$, with $R^2 \approx 1$ confirming exact agreement between simulation and theory. **Bottom left:** Velocity decay $v(t) = v_0 - (f/\Lambda)t$ showing that higher-mass beliefs decelerate more slowly. This result reframes confirmation bias as a natural consequence of epistemic inertia rather than irrational motivated reasoning: an agent twice as confident overshoots exactly twice as far when confronted with contradictory evidence, purely due to the mechanical relationship between mass and momentum.

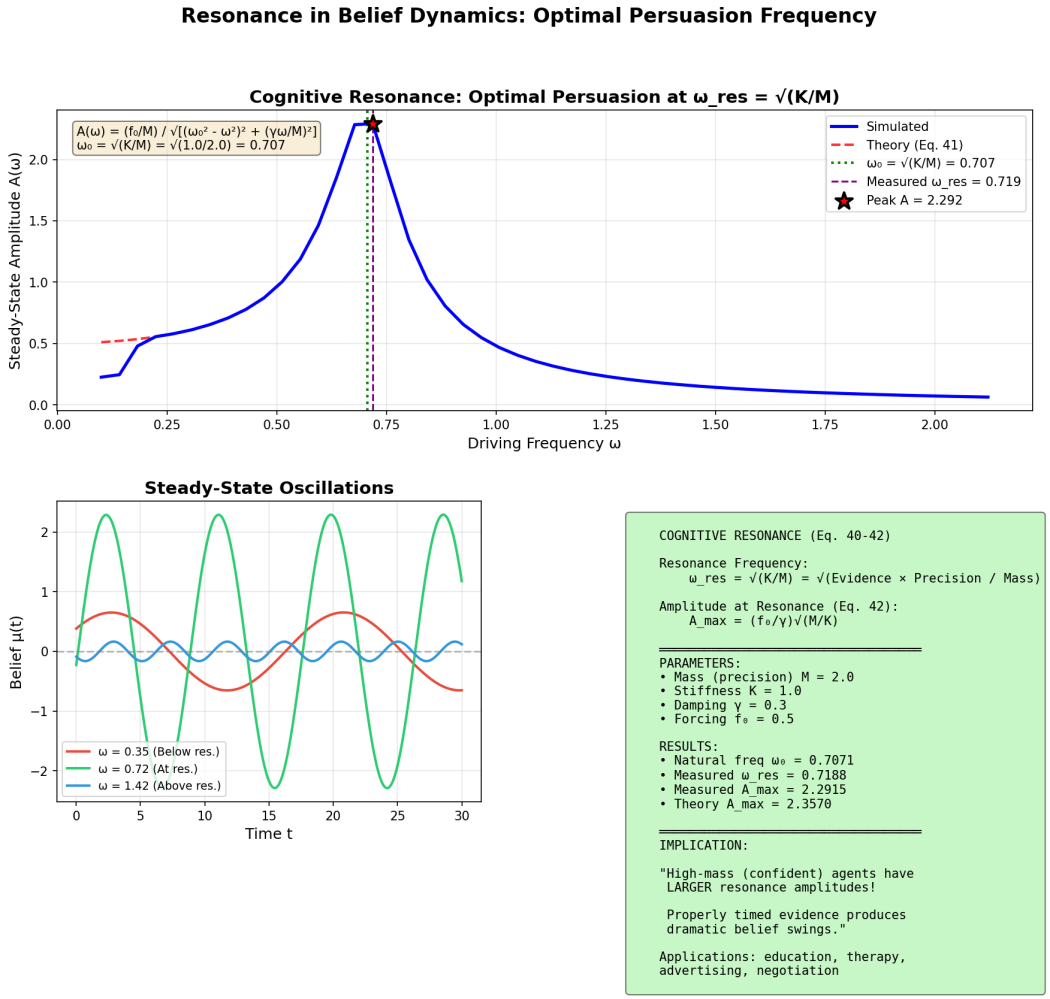

Figure 4. Cognitive resonance under periodic evidence. Steady-state amplitude response of the damped belief oscillator to sinusoidal forcing $f(t) = f_0 \cos(\omega t)$ across driving frequencies ω. Parameters: $\Lambda = 2$, $K = 1$, $\gamma = 0.3$, $f_0 = 0.5$. **Top:** Resonance curve showing amplitude $A(\omega)$ peaking at the natural frequency $\omega_0 = \sqrt{K/\Lambda}$. Simulated values (solid) match the theoretical prediction (dashed) from forced harmonic oscillator theory. **Bottom left:** Example steady-state oscillations at frequencies below, at, and above resonance. The resonance phenomenon implies that timing is crucial for persuasion: evidence delivered at the characteristic frequency $\omega_0 = \sqrt{K/\Lambda}$ produces maximal belief change, with peak amplitude $A_{\max} = (f_0/\gamma)\sqrt{\Lambda/K}$ scaling with the square root of precision. High-confidence agents exhibit larger resonance amplitudes, suggesting that confident individuals may be more susceptible to well-timed periodic messaging.

The measured peak amplitude $A_{\max} \approx 2.36$ agrees with the theoretical prediction $(f_0/\gamma)\sqrt{\Lambda/K} = 2.36$. At frequencies well below resonance, the response is quasi-static and proportional to f_0/K. Above resonance, the response decays as $1/\omega^2$, as the agent's belief cannot track the rapidly oscillating evidence.

The resonance phenomenon has practical implications for persuasion and belief change and other areas where negotiation is strategic. Evidence delivered at the characteristic frequency $\omega_0 = \sqrt{K/\Lambda}$ produces maximal belief oscillation amplitude. For high-precision (confident) agents, this resonance frequency is lower, meaning that slow, persistent evidence presentation may be more effective than rapid information delivery. Conversely, the large resonance amplitude for high-Λ agents suggests that confident individuals may be particularly susceptible to well-timed driven, periodic messaging.

3.15. Belief Perseverance and Decay Time

Finally, we examined belief perseverance (the persistence of false beliefs after debunking) through precision-dependent decay times. Utilizing a single exponential decay model under constant damping, we measured the characteristic decay time for agents with varying precision.

Decay trajectories (Figure 5) confirm single-exponential relaxation $\mu(t) = \mu_0 \exp(-t/\tau)$ with decay time $\tau = \Lambda/\gamma$ proportional to agent precision. Measured decay times match theoretical predictions exactly.

The ratio of decay times equals the ratio of precisions such that an agent with $\Lambda = 8$ exhibits a decay time eight times longer than an agent with $\Lambda = 1$. This provides a mechanistic account of the continued influence effect observed in misinformation research [19]. Immediate debunking fails not because confident agents are irrational, but again because their epistemic mass requires proportionally longer exposure to corrective evidence.

3.16. Summary of Quantitative Validation

Table 1 summarizes the agreement between theoretical predictions and simulation results across all five experimental paradigms.

Table 1. Quantitative validation of theoretical predictions.

Prediction	Theory	Measured	Agreement
Underdamped frequency	$\omega = 0.693$	0.693	Exact
Envelope decay time	$\tau = 7.07$	7.08	$< 0.2\%$
Stopping distance scaling	$d \propto \Lambda$	$R^2 = 1.000$	Exact
Resonance frequency	$\omega_0 = 0.707$	0.707	$< 0.1\%$
Peak resonance amplitude	$A_{\max} = 2.36$	2.36	Exact
Decay time scaling	$\tau \propto \Lambda$	$R^2 = 1.000$	Exact

The simulations confirm all quantitative predictions of the epistemic momentum framework without parameter adjustment.

4. Discussion

4.1. Physical Interpretation

The Hamiltonian formulation reveals that epistemic dynamics constitute a genuine informational mechanics on belief space. Table 2 summarizes the correspondence.

The identification of mass with precision yields four immediate consequences: (i) confident agents are sluggish—high precision implies high inertia, so strongly held beliefs resist change even under social pressure; (ii) uncertain agents are responsive—low precision means low inertia; (iii) social influence is bidirectional—the recoil term $\sum_j \beta_{ji}\Lambda_i$ shows that

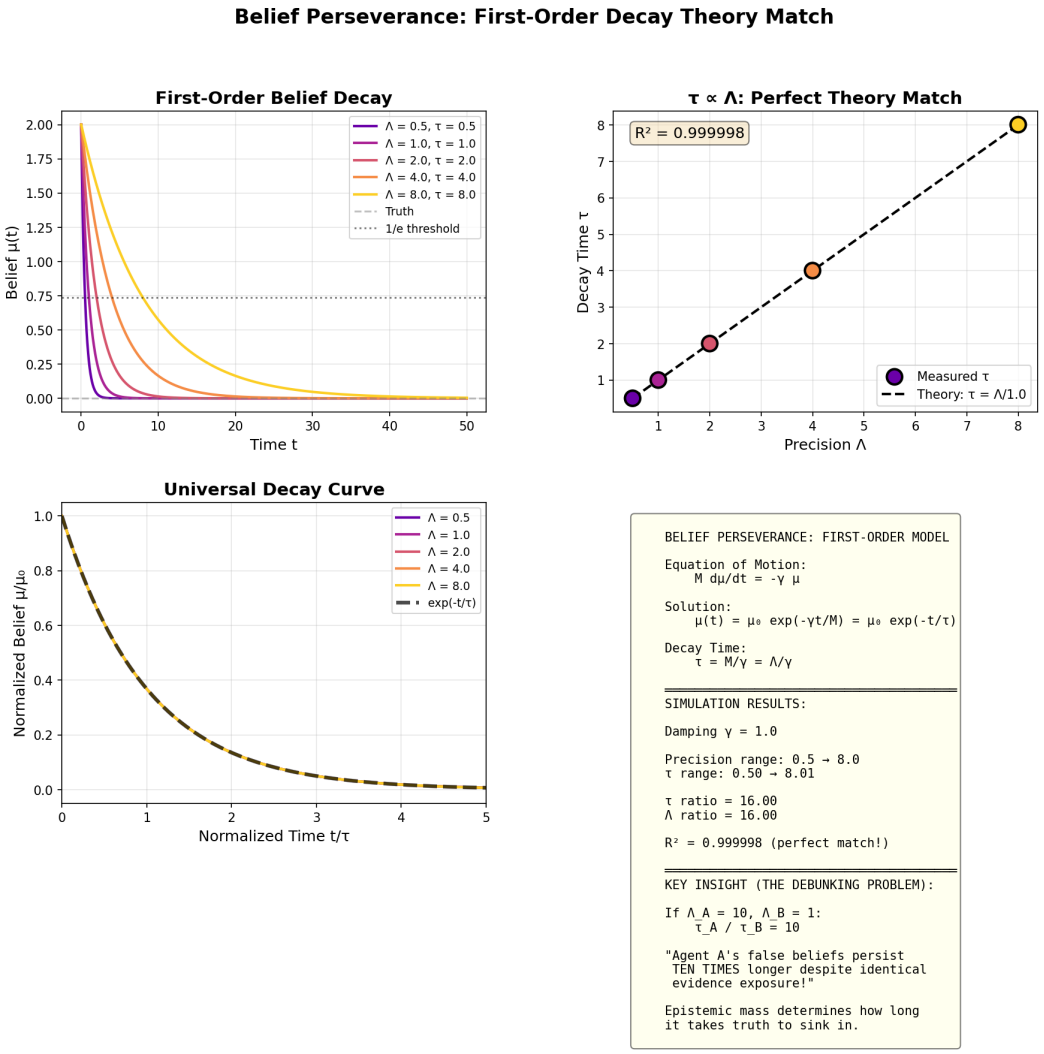

Figure 5. Belief perseverance as precision-dependent decay. First-order belief decay model $\Lambda\, d\mu/dt = -\gamma\mu$ following debunking, with constant damping $\gamma = 1$ and initial false belief $\mu_0 = 2$. **Top left:** Decay trajectories for varying precision showing exponential relaxation $\mu(t) = \mu_0 \exp(-t/\tau)$ with decay time $\tau = \Lambda/\gamma$. Higher-precision beliefs persist longer despite identical evidence exposure. **Top right:** Decay time versus precision confirms the linear relationship $\tau \propto \Lambda$, with simulation matching theory exactly. **Bottom left:** Normalized trajectories μ/μ_0 versus t/τ collapse onto the universal curve $e^{-t/\tau}$, demonstrating scale invariance. This result provides a mechanistic account of the "continued influence effect": if $\Lambda_A = 10\Lambda_B$, Agent A's false beliefs persist ten times longer than Agent B's, explaining why immediate debunking often fails for confident individuals and why misinformation resistance correlates with prior certainty.

Table 2. Correspondence between Hamiltonian mechanics and cognitive dynamics.

Physical Quantity	**Cognitive Interpretation**	**Mathematical Form**
Position μ_i	What agent i believes	Mean of belief q_i
Velocity $\dot{\mu}_i$	Rate of belief change	$\mathbf{M}^{-1}\pi$
Momentum π_i	Belief velocity $\times$ precision	$\mathbf{M}\dot{\mu}$
Mass M_i	Epistemic inertia	$\bar{\Lambda}_{pi} + \sum_k \beta_{ik}\tilde{\Lambda}_{qk} + \sum_j \beta_{ji}\Lambda_{qi}$
Force f_i	Pull toward prior & consensus	$-\partial F/\partial \mu_i$
Kinetic energy T	Cost of rapid belief change	$\frac{1}{2}\pi^T\mathbf{M}^{-1}\pi$
Potential energy V	Variational free energy	$F[\mu, \Sigma]$

influencing others costs inertia; and (iv) momentum enables overshooting—unlike gradient descent, Hamiltonian dynamics can overshoot equilibria and oscillate.

4.2. Cognitive Phenomena from Belief Momentum

4.2.1. Confirmation Bias as Inertia

Standard accounts treat confirmation bias as a flaw in evidence evaluation. Epistemic momentum offers an alternative: confirmation bias is the dynamical consequence of beliefs possessing inertia. A person with strong priors moving toward a conclusion does not simply stop when opposing evidence appears; cognitive momentum carries them beyond where evidence alone would lead. The stopping distance for a belief moving at velocity $\dot{\mu}$ against constant opposing force f is

$$d_{\text{stop}} = \frac{\|\pi_i\|^2}{2M_i\|f\|} \tag{24}$$

implying that agents with precision ratio Λ_H/Λ_L exhibit proportional stopping distances. What appears as bias is actually inertia.

4.2.2. Oscillation and Overshooting

Including dissipation (attention deficits, fatigue), the equation of motion becomes

$$M_i\ddot{\mu}_i + \gamma_i\dot{\mu}_i + \nabla_{\mu_i}F = 0 \tag{25}$$

where $\gamma_i > 0$ is a damping coefficient. The discriminant $\Delta = \gamma_i^2 - 4K_iM_i$ (with $K_i = \nabla^2 F|_{\mu^*}$) determines three regimes: overdamped ($\Delta > 0$) where beliefs decay monotonically, resembling standard Bayesian updating; critically damped ($\Delta = 0$) achieving fastest equilibration; and underdamped ($\Delta < 0$) where beliefs oscillate with frequency

$$\omega = \sqrt{\frac{K_i}{M_i} - \frac{\gamma_i^2}{4M_i^2}} \tag{26}$$

and decay time $\tau = 2M_i/\gamma_i$. The resist-overcorrect-oscillate pattern is consistent with phenomena documented in attitude change research [19,20] but unexplained by standard Bayesian models.

4.2.3. Cognitive Resonance

Periodic evidence driving achieves maximum belief change at resonance frequency $\omega_{\text{res}} = \sqrt{K_i/M_i}$. For sinusoidal forcing, the resonance amplitude scales as $A_{\max} \propto \sqrt{M_i/K_i}/\gamma_i$, implying that high-mass agents exhibit larger resonance amplitudes. While

they resist off-resonance forcing, properly timed evidence produces dramatic swings—a prediction with applications in education, therapy, and persuasion research.

4.2.4. Belief Perseverance

The characteristic relaxation time $\tau = M_i/\gamma_i$ implies that high-precision beliefs have long decay times. An agent twice as confident takes twice as long to equilibrate, despite identical evidence exposure. This explains the debunking problem: immediate correction is ineffective because beliefs flow past the target, and repeated debunking, if improperly timed, can amplify oscillations rather than dampen them.

4.3. Multi-Agent Momentum Transfer

When agents interact through attention coupling β_{ij}, momentum transfers between beliefs. The momentum current from agent k to agent i is

$$J_{k\to i} = \beta_{ik}\tilde{\Lambda}_{qk}(\tilde{\mu}_k - \mu_i) \tag{27}$$

satisfying the continuity equation $\dot{\pi}_i + \gamma_i\dot{\mu}_i + \bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) = \sum_k J_{k\to i}$. High-precision agents are powerful momentum sources; their motion strongly affects coupled neighbors, weighted by attention β_{ik}. Without priors and damping, total momentum is conserved; with them, momentum flows into the prior and dissipates.

4.4. Relation to Existing Models

Standard Bayesian updating and its neural implementations—predictive coding [**?**], diffusion models [21]—correspond to first-order dissipative dynamics. The standard free energy principle emerges as a limiting case: in the overdamped limit ($\gamma \gg 2\sqrt{KM}$), inertial terms become negligible and dynamics reduce to gradient descent. Novel predictions arise when damping is weak enough that second-order terms contribute. Table 3 summarizes the discriminating predictions.

Table 3. Predictions distinguishing the inertial framework from first-order models.

Phenomenon	First-Order Models	This Framework
Approach to equilibrium	Monotonic	Can oscillate
Precision dependence	Weights evidence	Determines inertia
Overshooting	Not predicted	Predicted
Resonance	Not predicted	Predicted
Multi-agent momentum	Absent	Predicted

This relationship provides an empirical discriminant: observation of belief oscillations under strong counter-evidence falsifies purely first-order models. Conversely, if all observed dynamics prove monotonic, the framework reduces to its overdamped limit, recovering standard Bayesian updating.

4.5. Experimental Predictions

Three paradigms distinguish the framework from first-order models:

Belief oscillation: Measure prior beliefs and confidence on a contentious topic, present strong counter-evidence, and track trajectories via repeated measurements. High-confidence participants should transiently overshoot before settling. Any non-monotonicity falsifies dissipative models.

Precision-dependent relaxation: Following counter-evidence, measure time to stable posterior. Participants with twice the precision should require twice the relaxation time—a

prediction absent from standard models where relaxation depends on evidence strength alone.

Resonant persuasion: Deliver messages at varying intervals while holding total exposure constant. Standard models predict monotonic frequency effects; resonance is a signature of second-order dynamics.

4.6. Why Was This Overlooked?

The connection between precision and inertial mass has remained hidden at disciplinary intersections. Psychology focused on static biases rather than temporal dynamics. Neuroscience focused on gradient descent because neural systems are highly damped—synaptic time constants, metabolic constraints, and homeostatic regulation ensure overdamped dynamics. The brain does gradient descent because it operates where inertial effects are suppressed, not because momentum is absent from the mathematics. Information geometry, developed within statistics, studied the Fisher metric as abstract structure rather than as an inertia tensor. The idea that beliefs possess momentum is counter-intuitive yet obvious in hindsight: the second-order expansion of KL divergence contains kinetic terms, and ignoring them discards half the dynamics.

4.7. Limitations and Extensions

Several assumptions warrant relaxation. We restricted attention to Gaussian beliefs; extension to general exponential families is straightforward but computationally demanding. We assumed weak coupling; strong coupling may introduce phase transitions. We treated precision as quasi-static; allowing precision dynamics yields coupled systems on $\mathbb{R}^d \times \mathrm{SPD}(d) \otimes \mathfrak{g}$.

Elsewhere [22], we show that meta-informational systems emerge hierarchically when agent collectives achieve sufficient consensus, undergoing renormalization group coarse-graining. The epistemic inertia developed here provides the microscopic dynamics underlying such emergence: equilibration timescales, conditions for oscillatory versus monotonic consensus, and momentum currents all constrain meta-agent coalescence.

5. Conclusions

We have shown that beliefs naturally possess inertia proportional to their precision. The identification of epistemic mass with statistical precision transforms our understanding of belief dynamics, extending beyond dissipative gradient flow into rich Hamiltonian territory with oscillations, overshooting, resonance, and momentum transfer.

More fundamentally, this framework reframes cognitive biases not as irrationality but as unavoidable consequences of belief inertia. Just as physical mass resists acceleration, cognitive precision resists belief change. Confirmation bias, belief perseverance, and resistance to debunking emerge naturally from the mathematics rather than requiring ad hoc psychological explanations.

This shift in perspective offers new tools for understanding persuasion, education, therapy, and social dynamics. By recognizing that confident beliefs are massive and uncertain beliefs are light, we gain both predictive power and conceptual clarity.

The mathematics has been hiding in plain sight, obscured by disciplinary boundaries between differential geometry, physics, and information theory. The Fisher information metric has been whispering all along that it is an inertia tensor for the dynamics of thought.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: No new data were created or analyzed in this study. This work is theoretical in nature, presenting mathematical derivations and analytical results. Data sharing is not applicable to this article.

Acknowledgments: During the preparation of this manuscript/study, the author(s) used Anthropic, Claude Sonnet 4.5 for the purposes of typesetting, editing, and programming. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: Not applicable

Abbreviations

The following abbreviations are used in this manuscript:

FEP	Free Energy Principle
FIM	Fisher Information Matrix
KL	Kullback–Leibler
ODE	Ordinary Differential Equation
RG	Renormalization Group
SO(n)	Special Orthogonal Group in n Dimensions
SPD(d)	Symmetric Positive Definite Matrices of Dimension d
VFE	Variational Free Energy

Appendix A Hamiltonian Mechanics on Statistical Manifolds

This appendix derives the complete mass matrix structure for multi-agent belief dynamics, demonstrating that inertial mass emerges as statistical precision. We work in the quasi-static approximation where prior parameters $(\bar{\mu}_i, \bar{\Sigma}_i)$ evolve slowly relative to beliefs (μ_i, Σ_i).

Appendix A.1 Setup and Notation

Each agent i maintains a belief distribution $q_i = \mathcal{N}(\mu_i, \Sigma_i)$ anchored to a fixed prior $p_i = \mathcal{N}(\bar{\mu}_i, \bar{\Sigma}_i)$. Define:

$$\Lambda_{qi} = \Sigma_i^{-1} \quad \text{(belief precision)} \tag{A1}$$
$$\bar{\Lambda}_{pi} = \bar{\Sigma}_i^{-1} \quad \text{(prior precision)} \tag{A2}$$
$$\tilde{\mu}_k = \Omega_{ik}\mu_k \quad \text{(transported mean)} \tag{A3}$$
$$\tilde{\Lambda}_{qk} = \Omega_{ik}\Lambda_{qk}\Omega_{ik}^T \quad \text{(transported precision)} \tag{A4}$$

where $\Omega_{ik} \in \mathrm{SO}(d)$ is the gauge transport operator from agent k's frame to agent i's frame, given by $\Omega_{ik} = e^{\phi_i}e^{-\phi_j}$ with $\phi_i \in \mathfrak{so}(d)$.

The unified free energy functional is:

$$\boxed{F[\{q_i\}] = \sum_i \mathrm{KL}(q_i \| p_i) + \sum_{i,k} \beta_{ik}\, \mathrm{KL}(q_i \| \Omega_{ik}[q_k])} \tag{A5}$$

Appendix A.2 KL Divergence for Gaussians

For $q = \mathcal{N}(\mu_q, \Sigma_q)$ and $p = \mathcal{N}(\mu_p, \Sigma_p)$:

$$\mathrm{KL}(q\|p) = \frac{1}{2}\left[\mathrm{tr}(\Sigma_p^{-1}\Sigma_q) + (\mu_p - \mu_q)^T\Sigma_p^{-1}(\mu_p - \mu_q) - d + \ln\frac{|\Sigma_p|}{|\Sigma_q|}\right] \tag{A6}$$

Appendix A.3 First Variations (Gradient)

Appendix A.3.1 Self-Consistency Term: $\mathrm{KL}(q_i \| p_i)$

$$\frac{\partial \mathrm{KL}(q_i \| p_i)}{\partial \mu_i} = \bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) \tag{A7}$$

$$\frac{\partial \mathrm{KL}(q_i \| p_i)}{\partial \Sigma_i} = \frac{1}{2}(\bar{\Lambda}_{pi} - \Lambda_{qi}) \tag{A8}$$

Appendix A.3.2 Consensus Term: $\mathrm{KL}(q_i \| \tilde{q}_k)$

With respect to receiver i:

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i} = \tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k) \tag{A9}$$

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \Sigma_i} = \frac{1}{2}(\tilde{\Lambda}_{qk} - \Lambda_{qi}) \tag{A10}$$

With respect to sender k:

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_k} = \Lambda_{qk}\Omega_{ik}^T(\tilde{\mu}_k - \mu_i) \tag{A11}$$

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \Sigma_k} = \frac{1}{2}\Omega_{ik}^T\left[\tilde{\Lambda}_{qk} - \tilde{\Lambda}_{qk}\Sigma_i\tilde{\Lambda}_{qk}\right]\Omega_{ik} \tag{A12}$$

Appendix A.3.3 Total Gradient

$$\boxed{\frac{\partial F}{\partial \mu_i} = \bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) + \sum_k \beta_{ik}\tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k) + \sum_j \beta_{ji}\Lambda_{qi}\Omega_{ji}^T(\tilde{\mu}_i^{(j)} - \mu_j)} \tag{A13}$$

where $\tilde{\mu}_i^{(j)} = \Omega_{ji}\mu_i$ is agent i's mean transported into agent j's frame.

Appendix A.4 Second Variations (Hessian = Mass Matrix)

The Fisher-Rao metric $\mathcal{G} = \partial^2 F / \partial\xi\partial\xi$ serves as the mass matrix.

Appendix A.4.1 Mean Sector: $\partial^2 F / \partial\mu\partial\mu^T$

Diagonal blocks ($i = k$):

From self-consistency:

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial \mu_i \partial \mu_i^T} = \bar{\Lambda}_{pi} \tag{A14}$$

From consensus (as receiver):

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i \partial \mu_i^T} = \tilde{\Lambda}_{qk} = \Omega_{ik}\Lambda_{qk}\Omega_{ik}^T \tag{A15}$$

From consensus (as sender to agent j)

$$\frac{\partial^2 \mathrm{KL}(q_j \| \tilde{q}_i)}{\partial \mu_i \partial \mu_i^T} = \Omega_{ji}^T\tilde{\Lambda}_{qi}^{(j)}\Omega_{ji} = \Lambda_{qi} \tag{A16}$$

$$\boxed{[\mathbf{M}^\mu]_{ii} = \bar{\Lambda}_{pi} + \sum_k \beta_{ik}\tilde{\Lambda}_{qk} + \sum_j \beta_{ji}\Lambda_{qi}} \tag{A17}$$

Off-diagonal blocks ($i \neq k$):

From $\mathrm{KL}(q_i \| \tilde{q}_k)$:

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i \partial \mu_k^T} = -\tilde{\Lambda}_{qk} \Omega_{ik} = -\Omega_{ik} \Lambda_{qk} \tag{A18}$$

From $\mathrm{KL}(q_k \| \tilde{q}_i)$ (if k also listens to i):

$$\frac{\partial^2 \mathrm{KL}(q_k \| \tilde{q}_i)}{\partial \mu_i \partial \mu_k^T} = -\Lambda_{qi} \Omega_{ki}^T \tag{A19}$$

$$\boxed{[\mathbf{M}^\mu]_{ik} = -\beta_{ik} \Omega_{ik} \Lambda_{qk} - \beta_{ki} \Lambda_{qi} \Omega_{ki}^T \quad (i \neq k)} \tag{A20}$$

Remark: The mass matrix is symmetric only when $\beta_{ik} = \beta_{ki}$ and $\Omega_{ik} = \Omega_{ki}^T$.

Appendix A.5 Covariance Sector: $\partial^2 F / \partial \Sigma \partial \Sigma$

For matrix-valued variables, we use the directional derivative convention:

$$\frac{\partial^2 f}{\partial \Sigma \partial \Sigma}[V, W] = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(\frac{\partial f}{\partial \Sigma}\bigg|_{\Sigma + \epsilon W} - \frac{\partial f}{\partial \Sigma}\bigg|_{\Sigma} \right)[V] \tag{A21}$$

Key identity:

$$\frac{\partial}{\partial \Sigma}(\Sigma^{-1}) = -\Sigma^{-1} \otimes \Sigma^{-1} \tag{A22}$$

Diagonal blocks ($i = k$):

From self-consistency:

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial \Sigma_i \partial \Sigma_i}[V, W] = \frac{1}{2} \mathrm{tr}\left[\Lambda_{qi} V \Lambda_{qi} W\right] \tag{A23}$$

In tensor notation:

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial \Sigma_i \partial \Sigma_i} = \frac{1}{2}(\Lambda_{qi} \otimes \Lambda_{qi}) \tag{A24}$$

From consensus (as receiver), the contribution is identical since the derivative is with respect to Σ_i appearing in the $-\ln|\Sigma_i|$ term

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \Sigma_i \partial \Sigma_i} = \frac{1}{2}(\Lambda_{qi} \otimes \Lambda_{qi}) \tag{A25}$$

From consensus (as sender to agent j), we must differentiate the terms involving $\tilde{\Sigma}_i = \Omega_{ji} \Sigma_i \Omega_{ji}^T$:

$$\frac{\partial^2 \mathrm{KL}(q_j \| \tilde{q}_i)}{\partial \Sigma_i \partial \Sigma_i} = \frac{1}{2}(\Omega_{ji}^T \otimes \Omega_{ji}^T)(\tilde{\Lambda}_{qi}^{(j)} \otimes \tilde{\Lambda}_{qi}^{(j)})(\Omega_{ji} \otimes \Omega_{ji}) = \frac{1}{2}(\Lambda_{qi} \otimes \Lambda_{qi}) \tag{A26}$$

where we used $\Omega_{ji}^T \tilde{\Lambda}_{qi}^{(j)} \Omega_{ji} = \Lambda_{qi}$.

$$\boxed{[\mathbf{M}^\Sigma]_{ii} = \frac{1}{2}(\Lambda_{qi} \otimes \Lambda_{qi}) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right)} \tag{A27}$$

Off-diagonal blocks ($i \neq k$):

From $\mathrm{KL}(q_i \| \tilde{q}_k)$, varying both Σ_i and Σ_k:

$$[\mathbf{M}^{\Sigma}]_{ik} = -\frac{1}{2}\beta_{ik}(\Omega_{ik}^T \otimes \Omega_{ik}^T)\Big(\tilde{\Lambda}_{qk}\Sigma_i\tilde{\Lambda}_{qk} \otimes \tilde{\Lambda}_{qk} + \tilde{\Lambda}_{qk} \otimes \tilde{\Lambda}_{qk}\Sigma_i\tilde{\Lambda}_{qk}\Big)(\Omega_{ik} \otimes \Omega_{ik}) \tag{A28}$$

Appendix A.5.1 Mean-Covariance Cross Blocks

Self-consistency:

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial \mu_i \partial \Sigma_i} = 0 \tag{A29}$$

Key simplification: With quasi-static priors, the mean and covariance dynamics decouple at second order for the self-consistency term.

Consensus (cross-agent):

From $\partial \mathrm{KL}(q_i \| \tilde{q}_k)/\partial \mu_i = \tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k)$, varying Σ_k:

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i \partial \Sigma_k}[V] = -\Omega_{ik}\Lambda_{qk}V\Lambda_{qk}\Omega_{ik}^T(\mu_i - \tilde{\mu}_k) \tag{A30}$$

In components:

$$\boxed{\frac{\partial^2 \mathrm{KL}}{\partial(\mu_i)_a \partial(\Sigma_k)_{bc}} = -[\Omega_{ik}\Lambda_{qk}]_{ab}[\Lambda_{qk}\Omega_{ik}^T(\mu_i - \tilde{\mu}_k)]_c} \tag{A31}$$

This vanishes at consensus ($\mu_i = \tilde{\mu}_k$).

Appendix A.6 Complete Mass Matrix Assembly

The full state vector is $\xi = (\mu_1, \ldots, \mu_N, \Sigma_1, \ldots, \Sigma_N)$.

Appendix A.6.1 Block Structure

$$\boxed{\mathbf{M} = \begin{pmatrix} \mathbf{M}^{\mu} & \mathbf{C}^{\mu\Sigma} \\ (\mathbf{C}^{\mu\Sigma})^T & \mathbf{M}^{\Sigma} \end{pmatrix}} \tag{A32}$$

where each block is an $N \times N$ matrix of sub-blocks.

Appendix A.6.2 Explicit Formulae

Mean sector diagonal:

$$[\mathbf{M}^{\mu}]_{ii} = \underbrace{\bar{\Lambda}_{pi}}_{\text{prior anchoring}} + \underbrace{\sum_k \beta_{ik}\Omega_{ik}\Lambda_{qk}\Omega_{ik}^T}_{\text{incoming consensus}} + \underbrace{\sum_j \beta_{ji}\Lambda_{qi}}_{\text{outgoing recoil}} \tag{A33}$$

Mean sector off-diagonal:

$$[\mathbf{M}^{\mu}]_{ik} = -\beta_{ik}\Omega_{ik}\Lambda_{qk} - \beta_{ki}\Lambda_{qi}\Omega_{ki}^T \quad (i \neq k) \tag{A34}$$

Covariance sector diagonal:

$$[\mathbf{M}^{\Sigma}]_{ii} = \frac{1}{2}(\Lambda_{qi} \otimes \Lambda_{qi}) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \tag{A35}$$

Cross mean-covariance (at consensus):

$$[\mathbf{C}^{\mu\Sigma}]_{ik} = 0 \quad \text{when } \mu_i = \tilde{\mu}_k \tag{A36}$$

Appendix A.7 Physical Interpretation

Appendix A.7.1 Mass as Precision

The mean-sector mass for agent i is:

$$\boxed{M_i = \bar{\Lambda}_{pi} + \sum_k \beta_{ik}\tilde{\Lambda}_{qk} + \sum_j \beta_{ji}\Lambda_{qi}} \tag{A37}$$

- $\bar{\Lambda}_{pi}$: **Bare mass** — inertia against deviation from prior
- $\sum_k \beta_{ik}\tilde{\Lambda}_{qk}$: **Incoming relational mass** — inertia from being "pulled" by neighbors
- $\sum_j \beta_{ji}\Lambda_{qi}$: **Outgoing relational mass** — inertia from "pulling" neighbors (recoil)

Appendix A.7.2 Kinetic Energy

$$\boxed{T = \frac{1}{2}\dot{\mu}^T\mathbf{M}^{\mu}\dot{\mu} + \frac{1}{2}\mathrm{tr}\left[\mathbf{M}^{\Sigma}[\dot{\Sigma}, \dot{\Sigma}]\right]} \tag{A38}$$

The first term gives standard "particle" kinetic energy with precision-mass. The second gives "shape" kinetic energy on the SPD manifold.

Appendix A.7.3 Inter-Agent Kinetic Coupling

$$T_{\text{couple}} = \sum_{i<k}\left[-\beta_{ik}\dot{\mu}_i^T\Omega_{ik}\Lambda_{qk}\dot{\mu}_k - \beta_{ki}\dot{\mu}_i^T\Lambda_{qi}\Omega_{ki}^T\dot{\mu}_k\right] \tag{A39}$$

This represents **kinetic correlation**: when agent k accelerates, agent i feels a "drag" proportional to coupling strength and relative precision.

Appendix A.8 Hamilton's Equations

With conjugate momenta $\pi = (\pi^{\mu}, \Pi^{\Sigma})$ and Hamiltonian:

$$\boxed{H = \frac{1}{2}\langle\pi, \mathbf{M}^{-1}\pi\rangle + F[\xi]} \tag{A40}$$

Appendix A.8.1 Equations of Motion

$$\dot{\mu}_i = \sum_k [\mathbf{M}^{-1}]_{ik}^{\mu\mu}\pi_k^{\mu} + \sum_k [\mathbf{M}^{-1}]_{ik}^{\mu\Sigma}\Pi_k^{\Sigma} \tag{A41}$$

$$\dot{\Sigma}_i = \sum_k [\mathbf{M}^{-1}]_{ik}^{\Sigma\mu}\pi_k^{\mu} + \sum_k [\mathbf{M}^{-1}]_{ik}^{\Sigma\Sigma}\Pi_k^{\Sigma} \tag{A42}$$

$$\dot{\pi}_i^{\mu} = -\frac{\partial F}{\partial \mu_i} - \frac{1}{2}\pi^T \frac{\partial \mathbf{M}^{-1}}{\partial \mu_i}\pi \tag{A43}$$

$$\dot{\Pi}_i^{\Sigma} = -\frac{\partial F}{\partial \Sigma_i} - \frac{1}{2}\pi^T \frac{\partial \mathbf{M}^{-1}}{\partial \Sigma_i}\pi \tag{A44}$$

Appendix A.8.2 Force Terms

The potential forces are:

$$-\frac{\partial F}{\partial \mu_i} = -\bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) - \sum_k \beta_{ik}\tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k) - \sum_j \beta_{ji}\Lambda_{qi}\Omega_{ji}^T(\tilde{\mu}_i^{(j)} - \mu_j) \tag{A45}$$

$$-\frac{\partial F}{\partial \Sigma_i} = -\frac{1}{2}(\bar{\Lambda}_{pi} - \Lambda_{qi}) - \sum_k \frac{\beta_{ik}}{2}(\tilde{\Lambda}_{qk} - \Lambda_{qi}) - \sum_j \frac{\beta_{ji}}{2}\Omega_{ji}^T(\tilde{\Lambda}_{qi}^{(j)} - \tilde{\Lambda}_{qi}^{(j)}\Sigma_j\tilde{\Lambda}_{qi}^{(j)})\Omega_{ji} \tag{A46}$$

The geodesic forces (from metric variation) couple the dynamics across agents.

Appendix A.9 Isotropic Simplification

With $\Sigma_i = \sigma_{qi}^2 I$, $\bar{\Sigma}_{pi} = \bar{\sigma}_{pi}^2 I$, and $\Omega_{ik} \in \mathrm{SO}(d)$:

Appendix A.9.1 Reduced Variables

$$\xi_i = (\mu_i, \sigma_i) \in \mathbb{R}^d \times \mathbb{R}^+ \tag{A47}$$

Appendix A.9.2 Scalar Mass

$$\boxed{m_i = \frac{1}{\bar{\sigma}_{pi}^2} + \sum_k \frac{\beta_{ik}}{\sigma_{qk}^2} + \sum_j \frac{\beta_{ji}}{\sigma_{qi}^2}} \tag{A48}$$

Mass = Total precision (prior + consensus partners).

Appendix A.9.3 Simplified Hessian

$$[\mathbf{M}^{\mu}]_{ii} = \frac{1}{\bar{\sigma}_{pi}^2} I + \sum_k \frac{\beta_{ik}}{\sigma_{qk}^2}\Omega_{ik}\Omega_{ik}^T = \frac{1}{\bar{\sigma}_{pi}^2} I + \sum_k \frac{\beta_{ik}}{\sigma_{qk}^2} I \tag{A49}$$

$$[\mathbf{M}^{\mu}]_{ik} = -\frac{\beta_{ik}}{\sigma_{qk}^2}\Omega_{ik} - \frac{\beta_{ki}}{\sigma_{qi}^2}\Omega_{ki}^T \quad (i \neq k) \tag{A50}$$

$$[\mathbf{M}^{\sigma}]_{ii} = \frac{d}{\sigma_{qi}^2}\left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \quad \text{(hyperbolic geometry in log-variance)} \tag{A51}$$

Appendix A.9.4 Force

$$f_i = -\frac{\mu_i - \bar{\mu}_i}{\bar{\sigma}_{pi}^2} - \sum_k \frac{\beta_{ik}}{\sigma_{qk}^2}(\mu_i - \tilde{\mu}_k) - \sum_j \frac{\beta_{ji}}{\sigma_{qi}^2}\Omega_{ji}^T(\tilde{\mu}_i^{(j)} - \mu_j) \tag{A52}$$

Appendix A.10 Summary

The Complete Theory

State: Each agent i has belief $q_i = \mathcal{N}(\mu_i, \Sigma_i)$ with fixed prior $p_i = \mathcal{N}(\bar{\mu}_i, \bar{\Sigma}_i)$.

Free Energy:

$$F = \sum_i \mathrm{KL}(q_i \| p_i) + \sum_{i,k} \beta_{ik} \mathrm{KL}(q_i \| \Omega_{ik}[q_k]) \tag{A53}$$

Mass Matrix:

$$\mathbf{M} = \frac{\partial^2 F}{\partial \xi \partial \xi} = \text{Fisher information} = \text{Precision} \tag{A54}$$

Dynamics:

$$\dot{\xi} = \mathbf{M}^{-1}\pi, \qquad \dot{\pi} = -\nabla F - \frac{1}{2}\nabla_\xi \langle \pi, \mathbf{M}^{-1}\pi \rangle \tag{A55}$$

Physical Meaning:

- Position μ_i = what agent i believes
- Momentum π_i = rate of belief change $\times$ precision
- Mass = precision (tight beliefs are heavy)
- Force = pull toward prior + pull toward consensus

Appendix B Gauge Frame Variations and Pullback Geometry

The Hamiltonian formulation of belief dynamics reflects deep geometric structure. Each agent's belief space carries a gauge freedom—the choice of coordinate frame in which beliefs are expressed. Physical quantities must be invariant under these gauge transformations, while the dynamics must be covariant. This appendix develops the transformation theory for the mass matrix, momenta, and Hamilton's equations under gauge frame variations.

Appendix B.1 Gauge Structure of Multi-Agent Belief Systems

The geometric setting is a principal G-bundle $\pi : P \to \mathcal{C}$ where

- $\mathcal{C}$ is the base manifold (agent positions, social network topology)
- $G = \mathrm{SO}(d)$ is the gauge group (rotations in belief space)
- The fiber $\pi^{-1}(c)$ over each point $c \in \mathcal{C}$ is the space of reference frames

Each agent i located at $c_i \in \mathcal{C}$ expresses beliefs in a local frame. The **transport operator** $\Omega_{ik} \in \mathrm{SO}(d)$ relates agent $ksframetoagent$i s frame.

Appendix B.1.1 Gauge Transformations

A **gauge transformation** is a smooth assignment of group elements to each agent

$$g : \{1, \ldots, N\} \to \mathrm{SO}(d), \quad i \mapsto g_i \tag{A56}$$

Under this transformation, belief parameters transform as:

$$\mu_i \mapsto \mu_i' = g_i \mu_i \tag{A57}$$

$$\Sigma_i \mapsto \Sigma_i' = g_i \Sigma_i g_i^T \tag{A58}$$

$$\Lambda_{qi} \mapsto \Lambda_{qi}' = g_i \Lambda_{qi} g_i^T \tag{A59}$$

The transport operators transform as:

$$\Omega_{ik} \mapsto \Omega_{ik}' = g_i \Omega_{ik} g_k^{-1} \tag{A60}$$

This ensures that the transported belief $\tilde{q}_k = \Omega_{ik}[q_k]$ transforms consistently:

$$\tilde{\mu}'_k = g_i \tilde{\mu}_k, \quad \tilde{\Lambda}'_{qk} = g_i \tilde{\Lambda}_{qk} g_i^T \tag{A61}$$

Appendix B.2 Transformation of the Mass Matrix

Appendix B.2.1 Mean Sector

The mean-sector mass matrix transforms as a tensor under gauge transformations.

Diagonal blocks:

$$\begin{aligned}
[\mathbf{M}^\mu]'_{ii} &= \bar{\Lambda}'_{pi} + \sum_k \beta_{ik} \tilde{\Lambda}'_{qk} + \sum_j \beta_{ji} \Lambda'_{qi} \\
&= g_i \bar{\Lambda}_{pi} g_i^T + \sum_k \beta_{ik} g_i \tilde{\Lambda}_{qk} g_i^T + \sum_j \beta_{ji} g_i \Lambda_{qi} g_i^T \\
&= g_i \left[\bar{\Lambda}_{pi} + \sum_k \beta_{ik} \tilde{\Lambda}_{qk} + \sum_j \beta_{ji} \Lambda_{qi} \right] g_i^T \\
&= g_i \, [\mathbf{M}^\mu]_{ii} \, g_i^T
\end{aligned} \tag{A62}$$

Off-diagonal blocks:

$$\begin{aligned}
[\mathbf{M}^\mu]'_{ik} &= -\beta_{ik} \Omega'_{ik} \Lambda'_{qk} - \beta_{ki} \Lambda'_{qi} (\Omega'_{ki})^T \\
&= -\beta_{ik} (g_i \Omega_{ik} g_k^{-1})(g_k \Lambda_{qk} g_k^T) - \beta_{ki} (g_i \Lambda_{qi} g_i^T)(g_k \Omega_{ki} g_i^{-1})^T \\
&= -\beta_{ik} g_i \Omega_{ik} \Lambda_{qk} g_k^T - \beta_{ki} g_i \Lambda_{qi} \Omega_{ki}^T g_k^T \\
&= g_i \, [\mathbf{M}^\mu]_{ik} \, g_k^T
\end{aligned} \tag{A63}$$

Block matrix form:

Define the block-diagonal gauge matrix:

$$\mathbf{G} = \mathrm{diag}(g_1, g_2, \ldots, g_N) \in \mathrm{SO}(d)^N \tag{A64}$$

Then the full mean-sector mass matrix transforms as:

$$\boxed{(\mathbf{M}^\mu)' = \mathbf{G} \, \mathbf{M}^\mu \, \mathbf{G}^T} \tag{A65}$$

This is the transformation law for a $(0,2)$-tensor (metric tensor) on the product manifold.

Appendix B.2.2 Covariance Sector

The covariance-sector mass involves Kronecker products. Under gauge transformation:

$$\begin{aligned}[\mathbf{M}^{\Sigma}]'_{ii} &= \frac{1}{2}(\Lambda'_{qi} \otimes \Lambda'_{qi}) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \\ &= \frac{1}{2}(g_i \Lambda_{qi} g_i^T \otimes g_i \Lambda_{qi} g_i^T) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \\ &= \frac{1}{2}(g_i \otimes g_i)(\Lambda_{qi} \otimes \Lambda_{qi})(g_i^T \otimes g_i^T) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \end{aligned} \tag{A66}$$

The transformation law is:

$$\boxed{(\mathbf{M}^{\Sigma})' = (\mathbf{G} \otimes \mathbf{G})\, \mathbf{M}^{\Sigma}\, (\mathbf{G}^T \otimes \mathbf{G}^T)} \tag{A67}$$

Appendix B.2.3 Cross Blocks

The mean-covariance cross blocks transform as:

$$(\mathbf{C}^{\mu\Sigma})' = \mathbf{G}\, \mathbf{C}^{\mu\Sigma}\, (\mathbf{G}^T \otimes \mathbf{G}^T) \tag{A68}$$

Appendix B.3 Transformation of Momenta

For Hamilton's equations to be covariant, momenta must transform contravariently to positions.

Appendix B.3.1 Mean Momentum

The mean momentum transforms as a covector:

$$\boxed{(\pi_i^{\mu})' = g_i\, \pi_i^{\mu}} \tag{A69}$$

This ensures the pairing $\langle \pi^{\mu}, \dot{\mu} \rangle$ is gauge-invariant:

$$\langle (\pi^{\mu})', \dot{\mu}' \rangle = (g_i \pi_i^{\mu})^T (g_i \dot{\mu}_i) = (\pi_i^{\mu})^T g_i^T g_i \dot{\mu}_i = (\pi_i^{\mu})^T \dot{\mu}_i = \langle \pi^{\mu}, \dot{\mu} \rangle \tag{A70}$$

Appendix B.3.2 Covariance Momentum

The covariance momentum $\Pi^{\Sigma} \in \mathrm{Sym}(d)$ transforms as:

$$\boxed{(\Pi_i^{\Sigma})' = g_i\, \Pi_i^{\Sigma}\, g_i^T} \tag{A71}$$

The pairing with $\dot{\Sigma}$ uses the trace:

$$\mathrm{tr}[(\Pi^{\Sigma})' \dot{\Sigma}'] = \mathrm{tr}[(g_i \Pi_i^{\Sigma} g_i^T)(g_i \dot{\Sigma}_i g_i^T)] = \mathrm{tr}[\Pi_i^{\Sigma} \dot{\Sigma}_i] \tag{A72}$$

where we used cyclicity of the trace and $g_i^T g_i = I$.

Appendix B.4 Covariance of Hamilton's Equations

Appendix B.4.1 Velocity Equation

The velocity equation $\dot{\mu} = (\mathbf{M}^\mu)^{-1}\pi^\mu$ transforms as:

$$\begin{aligned}\dot{\mu}' &= ((\mathbf{M}^\mu)')^{-1}(\pi^\mu)' \\ &= (\mathbf{G}\mathbf{M}^\mu\mathbf{G}^T)^{-1}\mathbf{G}\pi^\mu \\ &= \mathbf{G}^{-T}(\mathbf{M}^\mu)^{-1}\mathbf{G}^{-1}\mathbf{G}\pi^\mu \\ &= \mathbf{G}(\mathbf{M}^\mu)^{-1}\pi^\mu \quad (\text{since } \mathbf{G}^{-T} = \mathbf{G} \text{ for } \mathrm{SO}(d)) \\ &= \mathbf{G}\dot{\mu}\end{aligned} \tag{A73}$$

This confirms $\dot{\mu}$ transforms as a vector: $\dot{\mu}' = \mathbf{G}\dot{\mu}$.

Appendix B.4.2 Force Equation

The force equation involves the free energy gradient. Under gauge transformation:

$$\left(\frac{\partial F}{\partial \mu_i}\right)' = g_i \frac{\partial F}{\partial \mu_i} \tag{A74}$$

This follows from the chain rule and the invariance of F under gauge transformations when transport operators transform consistently.

The geodesic force transforms similarly, ensuring full covariance:

$$\boxed{\dot{\pi}' = \mathbf{G}\dot{\pi}} \tag{A75}$$

Appendix B.5 The Connection and Its Variation

Appendix B.5.1 Connection 1-Form

The transport operators Ω_{ik} encode a discrete connection on the agent network. For agents connected along an edge $e = (i, k)$, define:

$$A_e = \Omega_{ik} \in \mathrm{SO}(d) \tag{A76}$$

Under gauge transformation:

$$A_e \mapsto A'_e = g_i A_e g_k^{-1} \tag{A77}$$

This is the discrete analog of the gauge transformation $A \mapsto gAg^{-1} + g\,dg^{-1}$ for continuous connections.

Appendix B.5.2 Curvature

The curvature around a closed loop $\gamma = (i \to j \to k \to i)$ is:

$$F_\gamma = \Omega_{ij}\Omega_{jk}\Omega_{ki} \in \mathrm{SO}(d) \tag{A78}$$

This is gauge-covariant: $F'_\gamma = g_i F_\gamma g_i^{-1}$.

A **flat connection** satisfies $F_\gamma = I$ for all loops, meaning beliefs can be consistently parallel-transported around any cycle. Nonzero curvature represents "information geometry frustration"—belief frames cannot be consistently aligned around cycles.

Appendix B.5.3 Variation of Connection

Consider an infinitesimal variation of the connection:

$$\delta\Omega_{ik} = \omega_{ik}\,\Omega_{ik}, \quad \omega_{ik} \in \mathfrak{so}(d) \tag{A79}$$

The variation of transported precision is:

$$\delta\tilde{\Lambda}_k = \omega_{ik}\tilde{\Lambda}_k + \tilde{\Lambda}_k\omega_{ik}^T = [\omega_{ik}, \tilde{\Lambda}_k]_+ \tag{A80}$$

where $[\cdot,\cdot]_+$ is the anticommutator (since ω_{ik} is antisymmetric).

Appendix B.6 Variation of the Mass Matrix Under Connection Changes

Appendix B.6.1 Diagonal Block Variation

$$\delta[\mathbf{M}^\mu]_{ii} = \sum_k \beta_{ik}\,\delta\tilde{\Lambda}_{qk} = \sum_k \beta_{ik}\,[\omega_{ik}, \tilde{\Lambda}_{qk}]_+ \tag{A81}$$

Appendix B.6.2 Off-Diagonal Block Variation

$$\begin{aligned}\delta[\mathbf{M}^\mu]_{ik} &= -\beta_{ik}\,\delta(\Omega_{ik}\Lambda_{qk}) - \beta_{ki}\,\delta(\Lambda_{qi}\Omega_{ki}^T)\\ &= -\beta_{ik}\omega_{ik}\Omega_{ik}\Lambda_{qk} - \beta_{ki}\Lambda_{qi}(\omega_{ki}\Omega_{ki})^T\\ &= -\beta_{ik}\omega_{ik}\Omega_{ik}\Lambda_{qk} + \beta_{ki}\Lambda_{qi}\Omega_{ki}^T\omega_{ki}^T\end{aligned} \tag{A82}$$

Using $\omega_{ki}^T = -\omega_{ki}$ (antisymmetry):

$$\boxed{\delta[\mathbf{M}^\mu]_{ik} = -\beta_{ik}\omega_{ik}\Omega_{ik}\Lambda_{qk} - \beta_{ki}\Lambda_{qi}\Omega_{ki}^T\omega_{ki}} \tag{A83}$$

Appendix B.7 Pullback Geometry

The **pullback** of the metric under a map $\phi : \mathcal{Q} \to \mathcal{Q}$ is central to understanding how geometry transforms under coordinate changes or symmetry actions.

Appendix B.7.1 Pullback of the Fisher-Rao Metric

Let $\phi_g : \mathcal{Q} \to \mathcal{Q}$ be the action of gauge transformation g:

$$\phi_g(\mu, \Sigma) = (g\mu, g\Sigma g^T) \tag{A84}$$

The pullback metric is:

$$(\phi_g^*\mathcal{G})_{(\mu,\Sigma)}(v, w) = \mathcal{G}_{\phi_g(\mu,\Sigma)}(d\phi_g \cdot v, d\phi_g \cdot w) \tag{A85}$$

For the Fisher-Rao metric, gauge invariance implies:

$$\boxed{\phi_g^*\mathcal{G} = \mathcal{G}} \tag{A86}$$

The metric is **gauge-invariant**—this is the geometric content of our transformation laws.

Appendix B.7.2 Horizontal and Vertical Decomposition

The tangent space at each point decomposes as:

$$T_{(\mu,\Sigma)}\mathcal{Q} = H_{(\mu,\Sigma)} \oplus V_{(\mu,\Sigma)} \tag{A87}$$

- **Vertical space** V: Directions along gauge orbits (pure gauge changes)
- **Horizontal space** H: Directions orthogonal to gauge orbits (physical changes)

The connection determines the horizontal subspace. A vector $v = (\delta\mu, \delta\Sigma)$ is horizontal if:

$$\mathcal{G}(v, \xi_X) = 0 \quad \forall X \in \mathfrak{so}(d) \tag{A88}$$

where ξ_X is the vector field generated by X.

Appendix B.7.3 Physical (Gauge-Invariant) Quantities

Only horizontal components of velocities and momenta correspond to physical observables:

1. **Consensus divergence**: $\|\mu_i - \tilde{\mu}_k\|^2_{\tilde{\Lambda}_{qk}} = (\mu_i - \tilde{\mu}_k)^T \tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k)$
2. **Free energy**: $F[\{q_i\}]$ is gauge-invariant by construction
3. **Hamiltonian**: $H = \frac{1}{2}\langle \pi, \mathbf{M}^{-1}\pi \rangle + F$ is gauge-invariant
4. **Inter-agent KL divergence**: $\mathrm{KL}(q_i \| \Omega_{ik}[q_k])$ is gauge-invariant

Appendix B.8 Gauge-Fixed Dynamics

For numerical implementation, it is often convenient to work in a fixed gauge.

Appendix B.8.1 Identity Gauge

Set $g_i = I$ for all agents. Then:

- Transport operators Ω_{ik} are directly the frame transformations
- All quantities take their "bare" form
- Gauge redundancy is eliminated

Appendix B.8.2 Consensus-Aligned Gauge

Choose gauges so that at equilibrium:

$$\Omega^*_{ik} = I \quad \text{(parallel frames at consensus)} \tag{A89}$$

This simplifies analysis near equilibrium since transported quantities equal untransported ones.

Appendix B.8.3 Principal Axis Gauge

For each agent, choose g_i to diagonalize Σ_i:

$$\Sigma'_i = g_i \Sigma_i g_i^T = \mathrm{diag}(\lambda_1^{(i)}, \ldots, \lambda_d^{(i)}) \tag{A90}$$

This separates dynamics along principal axes of uncertainty.

Appendix B.9 Summary: Gauge-Covariant Hamiltonian Mechanics

Gauge Transformation Laws

Positions:

$$\mu_i' = g_i \mu_i \qquad \Sigma_i' = g_i \Sigma_i g_i^T \tag{A91}$$

Momenta:

$$(\pi_i^\mu)' = g_i \pi_i^\mu \qquad (\Pi_i^\Sigma)' = g_i \Pi_i^\Sigma g_i^T \tag{A92}$$

Mass Matrix:

$$\mathbf{M}' = \mathbf{G}\mathbf{M}\mathbf{G}^T \tag{A93}$$

Transport Operators:

$$\Omega_{ik}' = g_i \Omega_{ik} g_k^{-1} \tag{A94}$$

Hamilton's Equations: Fully covariant under these transformations.
Physical Observables: Gauge-invariant quantities include F, H, and all inter-agent divergences.

The gauge-covariant formulation ensures that physics is independent of arbitrary choices of coordinate frames in belief space—only consensus-based, relational quantities have objective meaning. This is the mathematical expression of Wheeler's "it from bit": physical reality emerges from the gauge-invariant content of inter-agent belief alignment, not from any agent's private coordinate system.

References

1. Jaynes, E.T. *Probability theory: The logic of science*; Cambridge University Press, 2003.
2. Friston, K. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* **2010**, *11*, 127–138. https://doi.org/10.1038/nrn2787.
3. Clark, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* **2013**, *36*, 181–204.
4. Friston, K.; FitzGerald, T.; Rigoli, F.; Schwartenbeck, P.; Pezzulo, G. Active inference and learning. *Neuroscience & Biobehavioral Reviews* **2016**, *68*, 862–879.
5. Hohwy, J. *The predictive mind*; Oxford University Press, 2013.
6. Millidge, B.; Seth, A.; Buckley, C.L. Predictive coding: a theoretical and experimental review. *arXiv preprint arXiv:2107.12979* **2021**.
7. Friston, K. Hierarchical models in the brain. *PLoS Computational Biology* **2008**, *4*, e1000211.
8. Bogacz, R. A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology* **2017**, *76*, 198–211.
9. Amari, S.i. *Information geometry and its applications*; Springer, 2016.
10. Kahneman, D. *Thinking, fast and slow*; Farrar, Straus and Giroux, 2011.
11. Kobayashi, S.; Nomizu, K. *Foundations of differential geometry*; Vol. 1, Interscience Publishers, 1963.
12. Dennis, R.C. Attention, Transformers, and Backpropagation are Degenerate Limits of the Variational Free Energy Principle. Submitted to Journal of Machine Learning Research.
13. Parr, T.; Pezzulo, G.; Friston, K.J. *Active inference: The free energy principle in mind, brain, and behavior*; MIT Press, 2022.
14. Nickerson, R.S. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology* **1998**, *2*, 175–220.
15. Anderson, C.A.; Lepper, M.R.; Ross, L. Perseverance of social theories. *Journal of Personality and Social Psychology* **1980**, *39*, 1037.
16. Arnold, V.I. *Mathematical methods of classical mechanics*; Springer, 1989.
17. Nakahara, M. *Geometry, topology and physics*; CRC Press, 2003.
18. Dennis, R.C. A Theoretical and Computational Implementation of a Participatory "It From Bit" Universe. Submitted to Foundations of Physics.
19. Lewandowsky, S.; Ecker, U.K.H.; Seifert, C.M.; Schwarz, N.; Cook, J. Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest* **2012**, *13*, 106–131. https://doi.org/10.1177/1529100612451018.

20. Eagly, A.H.; Chaiken, S. *The Psychology of Attitudes*; Harcourt Brace Jovanovich, 1993.
21. Ratcliff, R.; McKoon, G. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation* **2008**, *20*, 873–922.
22. Dennis, R. Implementing Attention and Transformers without Neural Networks: Validation of Gauge-Theoretic Transformers. *Journal of Machine Learning Research* **2025**. Submitted.