# The Inertia of Belief: Hiding in Plain Sight

Robert C. Dennis
Independent Researcher
Leander, Texas 78641
cdenn016@gmail.com

November 25, 2025

**Abstract**

We present a dynamical theory of belief evolution where cognitive states possess momentum and inertia proportional to an agent's prior precision. Building on information geometry, we show that the Fisher information metric naturally provides an inertial mass tensor for beliefs updating and transmission. This yields a second order expansion of the variational free energy such that confident beliefs resist change and carry epistemic momentum, while uncertain beliefs change readily. We show that this endows the information geometry with a symplectic structure entirely analogous to Hamiltonian mechanics. Our framework extends the current paradigm of Bayesean belief updating from pure gradient descent to a completeand rich Hamiltonian dynamics. Our theory predicts and offers explanatory power to phenomena such as belief oscillation, cognitive resonance, and allows the study of momentum transfer via attention in social networks. We derive testable predictions: (1) confident beliefs should present over-shooting of equilibrium and oscillate when confronted with strong counter-evidence, (2) belief evolution times scale with the square root of prior precision, and (3) resonant belief forcing via pursuasion is most effective at the epistemic resonance frequency $\omega = \sqrt{\text{precision} \times \text{evidence strength}}$. These predictions offer new explanations and tools for confirmation bias, belief perseverance, echo-chambers, and opinion polarization as natural consequences of epistemic inertia rather than irrational biases.

## 1 Introduction

Why do some beliefs resist change more than others? Some are stiff and yet others readily sway to and fro. While confident beliefs clearly possess more "inertia" than uncertain ones, a principled mathematical foundation for this intuitive phenomenon remains elusive. Current theories of belief updating, from Bayesian inference (Jaynes, 2003) to predictive coding (Friston, 2010; Clark, 2013), model belief change as gradient descent. This is a purely dissipative process where beliefs flow toward lower free energy without momentum, inertia, or dynamics. Though enormously successful across neuroscience (Friston et al., 2016), psychology (Hohwy, 2013), and machine learning (Millidge et al., 2021), this framework fundamentally remains incomplete.

In this article, we show that beliefs possess an "epistemic" inertia proportional to an agent's prior precision. Just as physical objects with mass resist acceleration, beliefs held with high confidence (precision) resist change and, once moving, tend to continue in their direction. This is not merely metaphor but, rather, it is a mathematical consequence of a second order expansion of the variational free energy. In this view the Fisher information metric (Amari, 2016), which measures statistical distinguishability, simultaneously provides an inertial mass tensor for belief dynamics. The second-order terms in the KL divergence expansion, traditionally neglected due to myriad reasons(Friston, 2008; Bogacz, 2017), generate rich Hamiltonian dynamics with conserved quantities.

Furthermore, beliefs propagate through networks of agents in attention patterns ranging from coordinated consensus to turbulent disagreement, often exhibiting distortion, resonance, and phase transitions (Castellano et al., 2009; Galam, 2012). While numerous models, from opinion dynamics (Hegselmann and Krause, 2002) to quantum-inspired approaches (Busemeyer and Bruza, 2012), capture aspects of collective belief evolution, a principled geometric foundation remains incomplete and wholly absent.

As an intuitive example, consider an agent with strong priors about a political position (high precision). When presented with contradicting evidence, their belief doesn't immediately flip but instead resists change (inertia), may overshoot when it does shift (momentum), and might oscillate before settling (under-damped dynamics). Conversely, an uncertain agent (low prior precision) updates quickly toward new evidence or observation with minimal resistance. These phenomena, typically attributed to cognitive biases (Kahneman, 2011), emerge naturally from belief inertia.

Our framework makes three contributions to the field:

1. **Theoretical**: We derive a second order belief dynamics from first principles, showing the Fisher metric provides a natural inertial mass tensor $M = \Sigma_p^{-1} = \Lambda_p$ (prior precision). Via pullback geometry on informational bundles (Amari, 2016; Nielsen, 2020), we extend variational free energy to multi-agent systems characterized by belief-momentum exchange and gauge-covariant transport between agents in attention patterns where agents can hold identical beliefs yet have distinct perspectives**?** Dennis (2025a)(Kobayashi and Nomizu, 1963).

2. **Phenomenological**: We predict novel cognitive and social phenomena including belief oscillations, overshooting, and resonance emerging from belief inertia. These effects are absent in first order treatments (Parr et al., 2022) yet provide testable predictions distinguishing our framework from purely dissipative models for a variety of informational systems.

3. **Psychological**: We argue that cognitive biases such as confirmation bias (Nickerson, 1998), Dunning-Kruger effect (Kruger and Dunning, 1999), and belief perseverance (Anderson et al., 1980) are natural consequences of belief inertia rather than irrationality, offering a unified geometric explanation for seemingly disparate phenomena.

Our approach unlocks mathematical tools traditionally relegated to physics such as symplectic geometry (Arnold, 1989), perturbation theory (Holmes, 2012), Noether's theorem (Olver, 1993), renormalization group methods (Wilson and Kogut, 1975; Goldenfeld, 1992), topological phenomena (Nakahara, 2003; Bernevig and Hughes, 2013), and critical point analyses (Strogatz, 2015; Sornette, 2006) for understanding cognitive, social, and economic dynamics. By recognizing beliefs as dynamical objects with genuine inertia, we bridge information geometry, cognitive science, and collective behavior within a unified Hamiltonian framework.

# 2 Mathematical Framework

## 2.1 Beliefs as Points on Statistical Manifolds

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

For the remainer of this article we shall consider multio-variate Gaussian (MVG) beliefs and priors

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

where $\mu_\nu$ represents the believed value and $\Sigma_\nu$ represents uncertainty.

The Kullback-Leibler (KL) divergence measures the epistemic distance between an agent's belief $q$ and their prior model $p$

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

## 2.2 Multi-Agent Belief Geometry

We extend our single-agent framework to networks of interacting cognitive agents via attention. Following Dennis (2025a), we model agents as residing on a gauge-theoretic bundle geometry where each agent $i$ maintains beliefs and priors $q_i = \mathcal{N}(\mu_i, \Sigma_i)$ as well as an internal reference frame $\phi_i$ that determine how they interpret information.

Importantly, agents cannot directly compare beliefs. Instead, they must first align their gauge frames via parallel transport operators given by

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

This operator transforms agent $j$'s beliefs into agent $i$'s gauge frame of reference. This captures the fundamental reality that understanding requires translation between different cognitive perspectives (here, represented by $\phi$).

This operator acts by right action as

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

For example, it may be helpful to consider $\Omega_{ij} \in SO(3)$, the group of rotations where $\phi \in \mathfrak{so}3$, the Lie algebra of SO(3).

The transformed belief can then be compared with agent $i$'s own beliefs via KL divergence

$$D_{ij} = D_{\mathrm{KL}}(q_i \| \Omega_{ij} \cdot q_j) \tag{5}$$

Notice that this transport is, in general, asymmetric.

### 2.3 Multi-Agent Free Energy

As we derive in full detail in Dennis (2025a) the total variational free energy for a network of agents balances individual belief maintenance with social consensus pressure as

$$\mathcal{F}[\{q_i\},\{\phi_i\}] = \sum_i \underbrace{D_{\mathrm{KL}}(q_i \| p_i)}_{\text{Prior beliefs}} + \sum_{i,j} \underbrace{\beta_{ij} D_{\mathrm{KL}}(q_i \| \Omega_{ij} \cdot q_j)}_{\text{Social alignment}} \tag{6}$$

$$- \sum_i \underbrace{\mathbb{E}_{q_i}[\log p(o_i \mid \mu_i)]}_{\text{Sensory evidence}} \tag{7}$$

where $\beta_{ij}$ represents the attention agent $i$ places in agent $j$'s beliefs and we take $p_i$ to be quasi-static. The attention naturally emerges as

$$\beta_{ij} = \frac{\exp(-D_{\mathrm{KL}}(q_i \| \Omega_{ij} \cdot q_j)/\tau)}{\sum_k \exp(-D_{\mathrm{KL}}(q_i \| \Omega_{ik} \cdot q_k)/\tau)} \tag{8}$$

with temperature $\tau$ controlling selectivity (recovering transformer attention mechanisms Dennis (2025a)). In previous work we have shown that sensory evidence and/or observations are equivalent to agent-agent attention coupling. Hence, we will no longer utilize $\mathbb{E}_{q_i}[\log p(o_i \mid \mu_i)]$ and instead consider only the first two terms of the variational multi-agent free energy Dennis (2025b)

## 3 Hamiltonian Formulation of Belief Dynamics

The variational free energy principle is typically formulated as gradient descent—a purely dissipative dynamics where beliefs flow downhill toward equilibrium. However, this picture is incomplete. The second-order Taylor expansion of KL divergence reveals a kinetic energy term systematically neglected in standard treatments, extending the free energy principle from gradient flow to fully conservative Hamiltonian mechanics. This extension has profound implications: beliefs acquire inertia, cognitive systems exhibit momentum, and the Fisher information metric emerges as a mass matrix identifying *precision with inertial mass.*

### 3.1 The Adiabatic Approximation

Cognitive systems operate across multiple timescales often hierarchically. Beliefs generally update rapidly in response to sensory evidence when compared to priors which encode stable world-views, personality traits, or cultural assumptions. These evolve slowly through learning and interaction. We formalize this separation via the adiabatic approximation

Let prior parameters $(\bar{\mu}_i, \bar{\Sigma}_i)$ evolve on a slow timescale $T$, while beliefs $(\mu_i, \Sigma_i)$ evolve on a fast timescale $t$, with $\epsilon = t/T \ll 1$.

In the quasi-static limit $\epsilon \to 0$

- Priors $(\bar{\mu}_i, \bar{\Sigma}_i)$ are treated as fixed external parameters
- Only beliefs $(\mu_i, \Sigma_i)$ are dynamical variables
- The configuration space reduces to $\mathcal{Q} = \prod_i [\mathbb{R}^d \times \mathrm{SPD}(d)]$

This approximation captures the phenomenology of rapid belief inference against a stable anchor of learned expectations and behaviors. The slow drift of priors toward equilibrated beliefs constitutes learning.

## 3.2 State Space and Phase Space

Each agent $i$ maintains a Gaussian belief $q_i = \mathcal{N}(\mu_i, \Sigma_i)$ anchored to a fixed prior $p_i = \mathcal{N}(\bar{\mu}_i, \bar{\Sigma}_i)$. The dynamical state vector is then

$$\xi_i = (\mu_i, \Sigma_i) \in \mathbb{R}^d \times \mathrm{SPD}(d) \tag{9}$$

with dimension $d + \frac{d(d+1)}{2} = \frac{d(d+3)}{2}$ per agent.

The full system for $N$ agents is $\xi = (\xi_1, \dots, \xi_N)$, living on the product manifold

$$\mathcal{Q} = \prod_{i=1}^{N} \left[\mathbb{R}^d \times \mathrm{SPD}(d)\right] \tag{10}$$

To formulate Hamiltonian mechanics, we introduce **conjugate momenta**

$$\pi_i^{\mu} \in \mathbb{R}^d \qquad \text{(momentum conjugate to mean)} \tag{11}$$

$$\Pi_i^{\Sigma} \in \mathrm{Sym}(d) \qquad \text{(momentum conjugate to covariance)} \tag{12}$$

The phase space is then the cotangent bundle $T^*\mathcal{Q}$ where $(\xi, \pi) = (\mu, \Sigma, \pi^{\mu}, \Pi^{\Sigma})$.

## 3.3 The Reduced Free Energy

With priors fixed, the free energy becomes a functional of beliefs only

$$\boxed{F[\{q_i\}] = \sum_i \mathrm{KL}(q_i \| p_i) + \sum_{i,j} \beta_{ij}\, \mathrm{KL}(q_i \| \Omega_{ij}[q_j])} \tag{13}$$

The first term pegs each agent's belief to its prior and the second aligns beliefs across the social network through gauge-covariant transport $\Omega_{ij}$ and attention. This unified form subsumes observation terms via the freedom to impose that sensory data simply constitute messages from "environmental agents" with no ontological distinction. Dennis (2025a)

### 3.4 Mass as Precision: The Fisher-Rao Metric

The primary result enabling Hamiltonian mechanics is that the Hessian of free energy serves as the inertia/mass matrix

$$\mathbf{M} = \frac{\partial^2 F}{\partial\xi\partial\xi} = \mathcal{G} \tag{14}$$

where $\mathcal{G}$ is the Fisher-Rao information metric on the statistical manifold.

#### 3.4.1 Mean Sector Mass Matrix

For the $\mu$ parameters, this mass matrix has block structure

$$[\mathbf{M}^\mu]_{ik} = \begin{cases} \bar{\Lambda}_{pi} + \sum_l \beta_{il}\tilde{\Lambda}_{ql} + \sum_j \beta_{ji}\Lambda_{qi} & i = k \\ -\beta_{ik}\Omega_{ik}\Lambda_{qk} - \beta_{ki}\Lambda_{qi}\Omega_{ki}^T & i \neq k \end{cases} \tag{15}$$

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 diagonal block defines the effective mass of agent $i$

$$\boxed{M_i = \bar{\Lambda}_i + \sum_k \beta_{ik}\tilde{\Lambda}_k + \sum_j \beta_{ji}\Lambda_i} \tag{16}$$

This three-part structure has clear physical interpretation:

- $\bar{\Lambda}_i$: Bare mass from prior precision—inertia against deviation from deep expectations
- $\sum_k \beta_{ik}\tilde{\Lambda}_k$: Incoming relational mass—inertia from being "pulled" by neighbors
- $\sum_j \beta_{ji}\Lambda_i$: Outgoing relational mass—recoil from "pulling" neighbors

The off-diagonal blocks $[\mathbf{M}^\mu]_{ik}$ encode a kinetic coupling. This means that when agent $k$ accelerates, agent $i$ experiences a correlated force proportional to coupling strength and relative precision.

#### 3.4.2 Covariance Sector Mass Matrix

The covariance parameters $\Sigma_i \in \mathrm{SPD}(d)$ live on a curved manifold with its own metric structure. The mass matrix in this sector is

$$[\mathbf{M}^\Sigma]_{ii} = \frac{1}{2}(\Lambda_i \otimes \Lambda_i) \cdot \mathcal{S} \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \tag{17}$$

where $\mathcal{S}$ is the symmetrizer and $\otimes$ denotes the Kronecker product.

This gives the shape kinetic energy. This is the cost of changing one's uncertainty structure. Crucially, precision again plays the role of mass as agents with tight beliefs (high $\Lambda_i$) have large covariance inertia, resisting changes to their confidence levels.

#### 3.4.3 Cross Terms

The mean/covariance cross terms $\mathbf{C}^{\mu\Sigma}$ couple position and shape dynamics as

$$[\mathbf{C}^{\mu\Sigma}]_{ik} = -\beta_{ik}\,\Omega_{ik}\Lambda_{qk}(\cdot)\Lambda_{qk}\Omega_{ik}^T\,(\mu_i - \tilde{\mu}_k) \tag{18}$$

Importantly, these vanish when group consensus is achieved. This occurs when $\mu_i = \tilde{\mu}_k$, i.e. the situation where agents share beliefs modulo gauge transformation. Near consensus, mean and covariance dynamics approximately decouple.

### 3.5 The Hamiltonian

The Hamiltonian is the total energy in the standard "kinetic + potential" form

$$\boxed{H = \underbrace{\frac{1}{2}\langle \pi, \mathbf{M}^{-1}\pi\rangle}_{\text{kinetic energy}} + \underbrace{F[\xi]}_{\text{potential energy}}} \tag{19}$$

Expanding in sectors:

$$H = \frac{1}{2}(\pi^\mu)^T(\mathbf{M}^\mu)^{-1}\pi^\mu + \frac{1}{2}\mathrm{tr}\left[(\mathbf{M}^\Sigma)^{-1}[\Pi^\Sigma, \Pi^\Sigma]\right] + F[\mu, \Sigma] \tag{20}$$

The kinetic energy has the form $T = \frac{1}{2}mv^2$ with precision playing the role of mass and belief velocity $\dot{\xi}$ related to momentum through the metric.

### 3.6 Hamilton's Equations: Mean Sector

The equations of motion for mean parameters follow from the canonical structure in physics

#### 3.6.1 Velocity Equation

$$\boxed{\dot{\mu}_i = \sum_k [\mathbf{M}^{-1}]_{ik}^{\mu\mu}\,\pi_k^\mu + \sum_k [\mathbf{M}^{-1}]_{ik}^{\mu\Sigma}\,\Pi_k^\Sigma} \tag{21}$$

This relates belief velocity to momentum through the inverse mass matrix. The off-diagonal terms show that agent $i$'s motion depends on the momenta of all coupled agents. That is to say that beliefs are not independent but kinetically entangled.

#### 3.6.2 Force Equation

$$\boxed{\dot{\pi}_i^\mu = -\frac{\partial F}{\partial \mu_i} - \frac{1}{2}\pi^T\frac{\partial \mathbf{M}^{-1}}{\partial \mu_i}\pi} \tag{22}$$

The first term is the potential force. This is the standard gradient of the variational free energy often considered in the literature but here modified by our multi-agent gauge structure

$$-\frac{\partial F}{\partial \mu_i} = -\bar{\Lambda}_i(\mu_i - \bar{\mu}_i) - \sum_k \beta_{ik}\tilde{\Lambda}_k(\mu_i - \tilde{\mu}_k) - \sum_j \beta_{ji}\Lambda_i\Omega_{ji}^T(\tilde{\mu}_i^{(j)} - \mu_j) \tag{23}$$

This decomposes into

- **Prior restoring force**: Pull toward the prior mean $\bar{\mu}_i$
- **Consensus force**: Pull toward transported neighbor beliefs $\tilde{\mu}_k$
- **Reciprocal force**: Reaction from neighbors being pulled toward agent $i$

The second term is a geodesic force arising from the state-dependent metric

$$f_i^{\text{geo}} = -\frac{1}{2}\sum_{jkl}(\pi_j^\mu)^T \frac{\partial[\mathbf{M}^{-1}]_{jk}^{\mu\mu}}{\partial \mu_i}\pi_k^\mu \tag{24}$$

This encodes the curvature of the statistical manifold: even without potential gradients, beliefs follow curved trajectories known as the geodesics of information geometry Amari (2016).

## 3.7 Hamilton's Equations: Covariance Sector

The covariance sector requires extremely careful treatment since $\Sigma_i \in \text{SPD}(d)$ is in a curved, hyperbolic manifold.

### 3.7.1 Velocity Equation

$$\boxed{\dot{\Sigma}_i = \sum_k [\mathbf{M}^{-1}]_{ik}^{\Sigma\mu}\,\pi_k^\mu + \sum_k [\mathbf{M}^{-1}]_{ik}^{\Sigma\Sigma}\,\Pi_k^\Sigma} \tag{25}$$

The inverse metric on $\text{SPD}(d)$ takes the form

$$[\mathbf{M}^{-1}]^{\Sigma\Sigma}[\Pi] = 2\Sigma\,\Pi\,\Sigma \cdot (\text{normalization}) \tag{26}$$

This is the natural (affine-invariant) metric on the cone of positive definite matrices.

### 3.7.2 Force Equation

$$\boxed{\dot{\Pi}_i^\Sigma = -\frac{\partial F}{\partial \Sigma_i} - \frac{1}{2}\pi^T\frac{\partial \mathbf{M}^{-1}}{\partial \Sigma_i}\pi} \tag{27}$$

The potential force in the covariance sector is

$$-\frac{\partial F}{\partial \Sigma_{qi}} = -\frac{1}{2}(\bar{\Lambda}_{pi} - \Lambda_{qi}) - \sum_k \frac{\beta_{ik}}{2}(\tilde{\Lambda}_{qk} - \Lambda_{qi}) \tag{28}$$

This drives covariance toward the precision-weighted average of prior and neighbor uncertainties.

### 3.8 Compact Form

The full system can be written compactly as

$$\boxed{\begin{aligned} &\dot{\xi} = \mathbf{M}^{-1}\pi \\ &\dot{\pi} = -\nabla F - \frac{1}{2}\nabla_\xi \langle \pi, \mathbf{M}^{-1}\pi \rangle \end{aligned}} \tag{29}$$

These are the informational Hamilton's equations on the cotangent bundle $T^*\mathcal{Q}$ with symplectic structure. The Hamiltonian $H$ is conserved and gauge covariant along trajectories

$$\frac{dH}{dt} = 0 \tag{30}$$

This conservation law is typically absent in gradient descent formulations but has profound implications for cognitive dynamics, allowing oscillatory and quasi-periodic behavior rather than pure relaxation.

### 3.9 Physical Interpretation

The Hamiltonian formulation reveals that epistemic dynamics are not merely optimization but a genuine informational mechanics on belief space

The identification of mass with precision has several consequences

1. **Confident agents are sluggish**: High precision implies high inertia. Strongly held beliefs resist change even under social pressure.
2. **Uncertain agents are responsive**: Low precision means low inertia and therefore agents with diffuse beliefs readily adopt others' views and follows the crowd.
3. **Social influence is bidirectional**: The recoil term $\sum_j \beta_{ji}\Lambda_i$ shows that influencing others costs inertia.
4. **Momentum enables overshooting**: Unlike gradient descent, Hamiltonian dynamics can overshoot equilibria and oscillate.

This info-mechanical picture transforms our understanding of socio-cognition from a static optimization problem to a bona-fide dynamical system with conserved quantities, characteristic frequencies, gauge covariance, and rich temporal structure.

# 4 Cognitive Phenomena from Belief Momentum

The Hamiltonian formulation introduces a quantity absent from standard treatments of Bayesean belief updating: epistemic momentum. Just as physical momentum allows objects to flow past equilibrium, epistemic momentum allows beliefs to overshoot, oscillate, and resist change in ways that pure gradient descent fundamentally cannot capture.

Table 1: Correspondence between Hamiltonian mechanics and cognitive dynamics. The free energy principle, when extended to second order, reveals that belief updating follows the same mathematical structure as classical mechanics, with precision playing the role of inertial mass.

| **Physical Quantity** | **Cognitive Interpretation** | **Mathematical Form** |
|---|---|---|
| Position $\mu_i$ | What agent $i$ believes | Mean of belief $q_i$ |
| Velocity $\dot{\mu}_i$ | Rate of belief change | $\mathbf{M}^{-1}\pi$ |
| Momentum $\pi_i$ | Belief velocity × precision | $\mathbf{M}\dot{\mu}$ |
| Mass $M_i$ | Epistemic inertia (precision) | $\bar{\Lambda}_i + \sum_k \beta_{ik}\tilde{\Lambda}_k + \sum_j \beta_{ji}\Lambda_i$ |
| 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 | $F[\mu, \Sigma]$ |
| *Covariance (shape) sector:* | | |
| Position $\Sigma_i$ | Agent $i$'s uncertainty | Covariance of belief $q_i$ |
| Momentum $\Pi_i$ | Shape change × precision$^2$ | $\mathbf{K}\dot{\Sigma}$ |
| Shape mass $\mathcal{K}_i$ | Resistance to uncertainty change | $\frac{1}{2}(\Lambda_i \otimes \Lambda_i)(1 + \sum \beta)$ |
| *Conserved quantities:* | | |
| Hamiltonian $H$ | Total cognitive energy | $T + V$ |
| Angular momentum $L$ | Rotational symmetry in belief space | $\mu \wedge \pi + \Sigma \wedge \Pi$ |

## 4.1 Defining Cognitive Momentum

**Definition 1** (Cognitive Momentum)**.** *The cognitive momentum of agent $i$ is the product of epistemic mass and belief velocity*

$$\boxed{\pi_i = M_i\dot{\mu}_i = \left(\bar{\Lambda}_{pi} + \sum_k \beta_{ik}\tilde{\Lambda}_{qk} + \sum_j \beta_{ji}\Lambda_{qi}\right)\dot{\mu}_i} \tag{31}$$

*where $\dot{\mu}_i$ is the rate of belief change.*

For an isolated agent with isotropic uncertainty $\Sigma_i = \sigma_i^2 I$, this simplifies to

$$\pi_i = \frac{1}{\sigma_i^2}\dot{\mu}_i = \Lambda_i\dot{\mu}_i \tag{32}$$

Momentum is not simply the velocity of belief. A confident agent (high $\Lambda$) moving slowly has the same momentum as an uncertain agent (low $\Lambda$) moving quickly. This asymmetry has interesting consequences for belief dynamics.

Table 2: Components of cognitive momentum and their psychological interpretations.

| **Component** | **Formula** | **Psychological Role** |
|---|---|---|
| Bare momentum | $\bar{\Lambda}_{pi}\dot{\mu}_i$ | Inertia from prior expectations |
| Social momentum | $\sum_k \beta_{ik}\tilde{\Lambda}_{qk}\dot{\mu}_i$ | Inertia from social embedding |
| Recoil momentum | $\sum_j \beta_{ji}\Lambda_{qi}\dot{\mu}_i$ | Inertia from influencing others |

## 4.2 Confirmation Bias as Momentum

Presently, research treats confirmation bias as a flaw in evidence evaluation and/or irrationality. Epistemic momentum yields an interesting perspective: confirmation bias is the dynamical consequence of beliefs possessing inertia and the underlying informational geometry holding a Fisher metric.

Therefore, we may predict that confident beliefs possess momentum that causes continued motion in their current direction even against mild opposing evidence. The stopping distance for a belief moving at velocity $\dot{\mu}$ against constant opposing force $f$ is then

$$d_{\text{stop}} = \frac{M_i\|\dot{\mu}_i\|^2}{2\|f\|} = \frac{\|\pi_i\|^2}{2M_i\|f\|} \tag{33}$$

From energy conservation we have that the initial kinetic energy $\frac{1}{2}\pi^T M^{-1}\pi$ must be dissipated by the work done against force $f$ over distance $d$

$$\frac{1}{2}\pi^T M^{-1}\pi = f \cdot d_{\text{stop}} \tag{34}$$

Solving for $d_{\text{stop}}$ gives the result.

This is a distance in "epistemic/informational space".

As an intuitive example, a person with a strong prior (high $\bar{\Lambda}_p$) who has been moving towards a conclusion/equilibrium (nonzero $\dot{\mu}$) doesn't simply stop when opposing evidence appears. Instead, they continue such that the cognitive momentum carries them beyond where the evidence alone would have lead them. Although this appears as confirmation bias, in our view it is actually belief inertia.

This then leads to a quantitative prediction: the ratio of stopping distances for high-precision ($\Lambda_H$) versus low-precision ($\Lambda_L$) agents is

$$\frac{d_H}{d_L} = \frac{\Lambda_H}{\Lambda_L} \tag{35}$$

This implies that a person twice as confident takes twice as long to stop and overshoots twice as far as another. In principle this can be tested by a clever experimentalist in order to falsify or validate the dynamical framework.

### 4.3 Belief Oscillation and Overshooting

Another prediction of our Hamiltonian epistemic dynamics is oscillation phenomena. Unlike gradient descent, which monotonically approaches equilibrium, Hamiltonian systems can overshoot, oscillate, and decay.

#### 4.3.1 The Damped Epistemic Oscillator

By including dissipation (for example, attention deficits, fatigue, etc), the equation of motion becomes

$$M_i\ddot{\mu}_i + \gamma_i\dot{\mu}_i + \nabla_{\mu_i}F = 0 \tag{36}$$

where $\gamma_i > 0$ is a damping coefficient. This equation, from the physics perspective, is the well-known driven and damped oscillator.

For small displacements from equilibrium $\mu^*$ we have

$$M_i\ddot{\delta\mu} + \gamma_i\dot{\delta\mu} + K_i\delta\mu = 0 \tag{37}$$

where $K_i = \nabla^2 F|_{\mu^*}$ represents the belief's "stiffness" (curvature of free energy at equilibrium, completely analogous to a spring).

Once again we arrive at a quantifiable prediction:

In the sub-critical ($\gamma_i < 2\sqrt{K_i M_i}$) regime, beliefs will oscillate around equilibrium with a frequency and decay time given by

$$\boxed{\omega = \sqrt{\frac{K_i}{M_i} - \frac{\gamma_i^2}{4M_i^2}} \approx \sqrt{\frac{\text{Evidence strength}}{\text{Epistemic mass}}}} \tag{38}$$

$$\tau = \frac{2M_i}{\gamma_i} \tag{39}$$

#### 4.3.2 Three Dynamical Regimes

As the standard physics of oscillators show the discriminant $\Delta = \gamma_i^2 - 4K_iM_i$ manifestly determines different behaviors/evolution

1. **Over-damped** ($\Delta > 0$): Beliefs decay to equilibrium monotonically without oscillation. This resembles standard Bayesian updating in the literature

2. **Critically damped** ($\Delta = 0$): This regime exhibits the fastest approach to equilibrium without oscillation. This suggests it may be optimal for rapid learning.

3. **Under-damped** ($\Delta < 0$): In this regime beliefs oscillate around the equilibrium value, overshooting periodically before equilibrating. This regime produces distinctly non-standard Bayesean dynamics.

As an illuminating and timely example consider an (conspiracy theorist) agent with high precision (strong prior beliefs) and low damping (resistance to evidence). When confronted with strong opposing evidence the agent will exhibit

1. **Initial resistance**: High mass $M = \Lambda$ resists the force of evidence
2. **Acceleration**: Persistent evidence eventually accelerates belief change
3. **Overshoot**: Momentum carries belief past the truth
4. **Oscillation**: Belief swings between acceptance and rejection
5. **Settling**: Damping eventually brings convergence to equilibrium

This pattern (resist, over-correct, oscillate) is frequently observed in attitude change studies but unexplained by standard Bayesian models. Here we find a natural account

## 4.4 Cognitive Resonance

Interestingly, general oscillatory systems exhibit resonance phenomena whereby maximum response occurs when the driving frequency matches the system's natural frequency. In our epistemic view this then has direct implications for persuasion and learning.

A prediction presents itself: periodic evidence driving achieves maximum belief change at the agents belief resonance frequency given by

$$\boxed{\omega_{\text{res}} = \sqrt{\frac{K_i}{M_i}} = \sqrt{\frac{\text{Evidence strength} \times \text{Precision}}{\text{Epistemic mass}}}} \tag{40}$$

### 4.4.1 Amplitude at Resonance

For example, with sinusoidal forcing $f(t) = f_0 \cos(\omega t)$, the steady-state amplitude is shown (in physics/engineering) to be

$$A(\omega) = \frac{f_0/M_i}{\sqrt{(\omega_0^2 - \omega^2)^2 + (\gamma\omega/M_i)^2}} \tag{41}$$

where $\omega_0 = \sqrt{K/M}$ is the system's "natural" frequency.

At resonance ($\omega = \omega_{\text{res}} \approx \omega_0$) then, we have

$$A_{\max} = \frac{f_0 M_i}{\gamma_i K_i^{1/2} M_i^{1/2}} = \frac{f_0}{\gamma_i}\sqrt{\frac{M_i}{K_i}} \tag{42}$$

Curiously this implies that high-mass (confident) agents have larger resonance amplitudes rather than smaller. While they resist off-resonance forcing, properly timed evidence produces dramatic swings. This prediction then offers myriad applications in psychological/sociological fields (education, advertising, negotiating, therapy, etc).

## 4.5 Belief Perseverance

The characteristic time for a belief to relax toward equilibrium in a social setting is given by

$$\boxed{\tau = \frac{M_i}{\gamma_i} = \frac{\bar{\Lambda}_i + \sum_k \beta_{ik}\tilde{\Lambda}_k + \sum_j \beta_{ji}\Lambda_i}{\gamma_i}} \tag{43}$$

High-precision beliefs have long decay times. This suggests phenomena where agents tend to hold onto beliefs even after thorough debunking and evidence to their contrary.

For example, if agent A has precision $\Lambda_A = 10$ and agent B has $\Lambda_B = 1$ (both with equal damping $\gamma$), then

$$\frac{\tau_A}{\tau_B} = \frac{\Lambda_A}{\Lambda_B} = 10 \tag{44}$$

Agent A's false beliefs persist ten times longer than that of B's, despite identical evidence exposure.

### 4.5.1 The Debunking Problem

Typically debunking assumes beliefs respond instantaneously to evidence yet our theory of epistemic momentum predicts that immediate debunking is ineffective. The belief should flow past the correction target. Furthermore, repeated debunking, if not properly timed, can lead to amplification (a well studied phenomenon in debunking studies). A candidate method for debunking, then, is to properly time the belief trajectory before reinforcing the correction. However, predicting that time scale for a given agent may be difficult.

## 4.6 Sociology and Multi-Agent Momentum Transfer

When agents interact through the attention free energy ($\beta_{ij}$ term), momentum can transfer between beliefs, i.e. one agent's beliefs affects another's. This suggests a system of coupled equations of motion given an attention pattern of a multi-agent system.

### 4.6.1 Coupled Equations of Motion

The full multi-agent dynamics with damping are

$$\boxed{M_i\ddot{\mu}_i + \gamma_i\dot{\mu}_i + \nabla_{\mu_i}F = 0} \tag{45}$$

We may expand the gradient as

$$M_i\ddot{\mu}_i = -\gamma_i\dot{\mu}_i - \bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) - \sum_k \beta_{ik}\tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k) - \sum_j \beta_{ji}\Lambda_{qi}\Omega_{ji}^T(\tilde{\mu}_i^{(j)} - \mu_j) \tag{46}$$

Then this can be written as

$$\boxed{\underbrace{M_i\ddot{\mu}_i}_{\text{Inertia}} = -\underbrace{\gamma_i\dot{\mu}_i}_{\text{Damping}} - \underbrace{\nabla_{\mu_i}F_{\text{prior}}}_{\text{Prior force}} - \underbrace{\nabla_{\mu_i}F_{\text{consensus}}}_{\text{Social force}}} \tag{47}$$

### 4.6.2 Momentum Transfer Theorem

**Theorem 2** (Momentum Transfer Between Agents)**.** *When agent $k$ changes belief, it transfers epistemic momentum to agent $i$ according to*

$$\left.\frac{d\pi_i}{dt}\right|_{\text{from } k} = -\beta_{ik}\tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k) - \beta_{ki}\Lambda_{qi}\Omega_{ki}^T(\tilde{\mu}_k^{(i)} - \mu_i) \tag{48}$$

*The total momentum transfer over a given interaction time scale $[0, T]$ is*

$$\boxed{\Delta\pi_i = -\int_0^T \left[\beta_{ik}\tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k) + \beta_{ki}\Lambda_{qi}\Omega_{ki}^T(\tilde{\mu}_k^{(i)} - \mu_i)\right] dt} \tag{49}$$

### 4.6.3 Conservation and Non-Conservation

Without priors and damping, the total momentum is a conserved quantity.

$$\frac{d}{dt}\sum_i \pi_i = 0 \quad \text{(closed system)} \tag{50}$$

In contrast, with priors and damping, momentum is assuredly not conserved. Momentum flows into the environment (the prior) and is then dissipated

$$\frac{d}{dt}\sum_i \pi_i = -\sum_i \gamma_i\dot{\mu}_i - \sum_i \bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) \tag{51}$$

This allows us to define a momentum current from agent $k$ to agent $i$ as

$$J_{k\to i} = \beta_{ik}\tilde{\Lambda}_{qk}(\tilde{\mu}_k - \mu_i) \tag{52}$$

This satisfies the continuity equation

$$\dot{\pi}_i + \gamma_i\dot{\mu}_i + \bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) = \sum_k J_{k\to i} \tag{53}$$

We find that momentum flows from agents with different beliefsvia attention $\beta_{ik}$ and sender precision $\Lambda_{qk}$. High-precision agents are powerful momentum sources as their motion strongly affects coupled neighbors. Howver, their strength is weighted by their relative attentions $\beta_{ij}$

## 4.7 Summary

Our epistemic momentum framework unifies seemingly disparate phenomena such as confirmation bias, belief perseverance, oscillation, and social influence into manifestations of a single underlying epistemic Hamiltonian mechanics. Beliefs are not just updated, they are accelerated. Evidence does not instantly change minds but rather applies an epistemic force. Finally, confident beliefs don't only resist change rather, they possess epistemic inertia that carries them further than evidence alone would have lead them.

Table 3: Testable predictions from cognitive momentum theory.

| Phenomenon | Prediction | Experimental Test |
|---|---|---|
| Confirmation bias | Stopping distance is $\propto$ precision | Measure belief change latency vs. covariance |
| Belief oscillation | Under-damped agents overshoot truth and oscillate | Track belief trajectories over time |
| Resonance | Optimal persuasion occurs at $\omega_{\text{res}} = \sqrt{K/M}$ | Vary message timing, measure change |
| Perseverance | Decay time $\tau = M/\gamma$ | Measure false belief persistence vs. uncertainty |
| Social momentum | High-$\Lambda$ agents transfer more momentum | Attention vs. source confidence |
| Recoil | Persuaders become harder to persuade | Measure attitude stiffness after persuasion attempts |

# 5 Results

# 6 Discussion

## 6.1 Why Was This Overlooked?

The connection between precision and inertial mass, despite its naturalness, has remained hidden at the intersection of several far-flung disciplines that rarely communicate and interact with meaningful attention patters ($\beta_{ij}$). Psychology has historically focused on static biases and heuristics by cataloging the ways beliefs deviate from normative standards rather than on the temporal dynamics of how and why beliefs change in the manner they do. Questioning "how fast does this belief evolve?" has not been asked by utilizing the proper tools. Neuroscience, meanwhile, focuses on gradient-based descent precisely because neural systems are highly damped. For example, synaptic time constants, metabolic constraints, and homeostatic regulation ensure that neural dynamics are overdamped. This leads researchers to avoid the complexities of oscillatory or momentum-like behavior that might otherwise have been visible. The brain appears to do gradient descent because it operates in a regime where inertial effects are suppressed; not because momentum is absent from the underlying mathematics. It's been there the whole time. We may anticipate its utility in other fields which lean on informational systems.

Information geometry, meanwhile, provides the mathematical language for these ideas. It was developed largely within statistics and machine learning communities, far isolated from psychological or sociological theory. The Fisher metric was studied as an abstract structure on complicated probability spaces, rather than as an inertia tensor governing dynamics. Mostly, the idea that beliefs possess momentum is counter-intuitive at first glance yet obvious in hindsight. We aren't accustomed to thinking of beliefs as probabilistic dynamical variables with velocity and inertia. We "hold" beliefs, we don't "move" them. Yet the mathematics is clear and unambiguous: the second order expansion of the KL divergence simply contains kinetic terms, and ignoring them discards half of reality.

### 6.2 Limitations and Extensions

Our current theory rests on several simplifying assumptions that future work should relax and pursue. For simplicity and tractability, we've restricted our present attentions to Gaussian beliefs in the quasi-static regime where priors do not evolve meaningfully. While analytically tractable, Gaussians simply cannot capture the multi-modal distributions characteristic of the complexities of human informational systems. The extension to general exponential families is straightforward but the computational complexity increases substantially. We've also assumed weak coupling between agents, treating the consensus terms perturbatively. It is unclear at present if strong couplings introduce non-perturbative interaction terms that may qualitatively change the dynamics and potentially enabling a zoo of phase transitions. Additionally, in the present study we have considered belief precision to be quasi-static. While we present simulation results of precision flow we do not consider them analytically here. Furthermore, as we discuss elsewhere (in other contexts) our full theory treats priors (as well as gauge frames) as dynamical variables, yielding complicated coupled systems on $\mathbb{R}^d \times \mathrm{SPD}(d) \otimes \mathfrak{g}$ where beliefs and priors both evolve.

Several directions present exciting areas of future study. Multi-modal distributions representing conflicting hypotheses would connect to models of cognitive dissonance and attitude ambivalence. Social networks with explicit momentum exchange across varied rich and dynamic attention patters could illuminate phenomena like viral belief propagation, where the velocity of a meme matters as much as its content. Our framework presented here provides mathematical rigor and a fountain of possibilities; experiment and implementation are the next frontiers.

## 7 Conclusion

We have shown that beliefs naturally possess inertia in relation to prior precision. The straightforward identification that epistemic mass equals statistical precision transforms our understanding of belief dynamics provides new tools that extend beyond dissipative gradient flow and into rich Hamiltonian dynamics.

Our theory predicts oscillations, over-shooting, resistance, decay, and resonances in belief dynamics. More fundamentally, it re-frames cognitive biases not as irrationality but instead as unavoidable consequences of belief inertia. Just as physical mass resists acceleration, cognitive precision resists belief change.

This shift in perspective offers researchers new tools and methods for understanding persuasion, education, therapy, negotiation, and social dynamics. By recognizing that confident beliefs are massive and uncertain beliefs are light, we chart new frontiers in research and socio-psychological understanding.

The mathematics has been hiding in plain sight for decades due to lack of transitive communication between far-flung fields of differential geometry, physics, and informational geometry. The Fisher information metric has been whispering this entire time that it is actually an inertia tensor for the dynamics of thought.

# 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 $(\mu_i, \Sigma_i)$.

## 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:

$$P_i = \Sigma_i^{-1} \qquad \text{(belief precision)} \tag{54}$$

$$\bar{P}_i = \bar{\Sigma}_i^{-1} \qquad \text{(prior precision)} \tag{55}$$

$$\tilde{\mu}_k = R_{ik}\mu_k + t_{ik} \qquad \text{(transported mean)} \tag{56}$$

$$\tilde{P}_k = R_{ik}P_kR_{ik}^T \qquad \text{(transported precision)} \tag{57}$$

where $R_{ik} \in \mathrm{SO}(d)$ and $t_{ik} \in \mathbb{R}^d$ define the parallel transport from agent $k$ to agent $i$.

The unified free energy functional is:

$$\boxed{F[\{q_i\}] = \sum_i \mathrm{KL}(q_i \| p_i) + \sum_{i,k} \beta_{ik}\, \mathrm{KL}(q_i \| \Omega_{ik}[q_k])} \tag{58}$$

## A.2 KL Divergence for Gaussians

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

$$\mathrm{KL}(q\|p) = \frac{1}{2}\left[\mathrm{tr}(\Sigma_p^{-1}\Sigma_q) + (\mu_p - \mu_q)^T\Sigma_p^{-1}(\mu_p - \mu_q) - d + \ln\frac{|\Sigma_p|}{|\Sigma_q|}\right] \tag{59}$$

## A.3 First Variations (Gradient)

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

$$\frac{\partial \mathrm{KL}(q_i\|p_i)}{\partial \mu_i} = \bar{P}_i(\mu_i - \bar{\mu}_i) \tag{60}$$

$$\frac{\partial \mathrm{KL}(q_i\|p_i)}{\partial \Sigma_i} = \frac{1}{2}(\bar{P}_i - P_i) \tag{61}$$

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

With respect to receiver $i$:

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i} = \tilde{P}_k(\mu_i - \tilde{\mu}_k) \tag{62}$$

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \Sigma_i} = \frac{1}{2}(\tilde{P}_k - P_i) \tag{63}$$

With respect to sender $k$:

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_k} = P_k R_{ik}^T(\tilde{\mu}_k - \mu_i) \tag{64}$$

$$\frac{\partial \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \Sigma_k} = \frac{1}{2} R_{ik}^T \left[\tilde{P}_k - \tilde{P}_k \Sigma_i \tilde{P}_k\right] R_{ik} \tag{65}$$

### A.3.3 Total Gradient

$$\boxed{\frac{\partial F}{\partial \mu_i} = \bar{P}_i(\mu_i - \bar{\mu}_i) + \sum_k \beta_{ik} \tilde{P}_k(\mu_i - \tilde{\mu}_k) + \sum_j \beta_{ji} P_i R_{ji}^T(\tilde{\mu}_i^{(j)} - \mu_j)} \tag{66}$$

## 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.

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

**Diagonal blocks** $(i = k)$**:** From self-consistency:

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial \mu_i \partial \mu_i^T} = \bar{P}_i \tag{67}$$

From consensus (as receiver):

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i \partial \mu_i^T} = \tilde{P}_k = R_{ik} P_k R_{ik}^T \tag{68}$$

From consensus (as sender to agent $j$):

$$\frac{\partial^2 \mathrm{KL}(q_j \| \tilde{q}_i)}{\partial \mu_i \partial \mu_i^T} = R_{ji}^T \tilde{P}_i^{(j)} R_{ji} = P_i \tag{69}$$

$$\boxed{[\mathbf{M}^\mu]_{ii} = \bar{P}_i + \sum_k \beta_{ik} \tilde{P}_k + \sum_j \beta_{ji} P_i} \tag{70}$$

**Off-diagonal blocks** $(i \neq k)$**:** From $\mathrm{KL}(q_i \| \tilde{q}_k)$:

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i \partial \mu_k^T} = -\tilde{P}_k R_{ik} = -R_{ik} P_k \tag{71}$$

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

$$\frac{\partial^2 \mathrm{KL}(q_k \| \tilde{q}_i)}{\partial \mu_i \partial \mu_k^T} = -P_i R_{ki}^T \tag{72}$$

$$\boxed{[\mathbf{M}^\mu]_{ik} = -\beta_{ik} R_{ik} P_k - \beta_{ki} P_i R_{ki}^T \quad (i \neq k)} \tag{73}$$

### A.4.2 Covariance Sector: $\partial^2 F / \partial\Sigma\partial\Sigma$

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

$$\frac{\partial^2 f}{\partial\Sigma\partial\Sigma}[V, W] = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left( \left.\frac{\partial f}{\partial\Sigma}\right|_{\Sigma+\epsilon W} - \left.\frac{\partial f}{\partial\Sigma}\right|_{\Sigma} \right) [V] \tag{74}$$

**Key identity:**

$$\frac{\partial}{\partial\Sigma}(\Sigma^{-1}) = -\Sigma^{-1} \otimes \Sigma^{-1} \tag{75}$$

**Diagonal blocks** $(i = k)$**:** From self-consistency:

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial\Sigma_i \partial\Sigma_i}[V, W] = \frac{1}{2} \left[ P_i V P_i W P_i + P_i W P_i V P_i \right] \tag{76}$$

In tensor notation with symmetrizer $\mathcal{S}$:

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial\Sigma_i \partial\Sigma_i} = \frac{1}{2}(P_i \otimes P_i) \cdot \mathcal{S} \tag{77}$$

From consensus (as receiver):

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial\Sigma_i \partial\Sigma_i} = \frac{1}{2}(P_i \otimes P_i) \cdot \mathcal{S} \tag{78}$$

From consensus (as sender):

$$\frac{\partial^2 \mathrm{KL}(q_j \| \tilde{q}_i)}{\partial\Sigma_i \partial\Sigma_i} = \frac{1}{2}(R_{ji}^T \otimes R_{ji}^T)(\tilde{P}_i \otimes \tilde{P}_i)(R_{ji} \otimes R_{ji}) \cdot \mathcal{S} + (\text{cross terms}) \tag{79}$$

Simplifying using $R_{ji}^T \tilde{P}_i R_{ji} = P_i$:

$$\boxed{[\mathbf{M}^\Sigma]_{ii} = \frac{1}{2}(P_i \otimes P_i) \cdot \mathcal{S} \cdot \left( 1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}(1 + \cdots) \right)} \tag{80}$$

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

$$\boxed{[\mathbf{M}^\Sigma]_{ik} = -\frac{1}{2}\beta_{ik}(R_{ik}^T \otimes R_{ik}^T)(\tilde{P}_k\Sigma_i\tilde{P}_k \otimes \tilde{P}_k + \tilde{P}_k \otimes \tilde{P}_k\Sigma_i\tilde{P}_k)(R_{ik} \otimes R_{ik})} \tag{81}$$

### A.4.3 Mean-Covariance Cross Blocks

**Self-consistency:**

$$\frac{\partial^2 \mathrm{KL}(q_i \| p_i)}{\partial \mu_i \partial \Sigma_i} = 0 \tag{82}$$

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

**Consensus (cross-agent):** From $\partial \mathrm{KL}(q_i \| \tilde{q}_k)/\partial \mu_i = \tilde{P}_k(\mu_i - \tilde{\mu}_k)$, varying $\Sigma_k$:

$$\frac{\partial^2 \mathrm{KL}(q_i \| \tilde{q}_k)}{\partial \mu_i \partial \Sigma_k}[V] = -R_{ik}P_k V P_k R_{ik}^T(\mu_i - \tilde{\mu}_k) \tag{83}$$

In components:

$$\boxed{\frac{\partial^2 \mathrm{KL}}{\partial(\mu_i)_a \partial(\Sigma_k)_{bc}} = -[R_{ik}P_k]_{ab}[P_k R_{ik}^T(\mu_i - \tilde{\mu}_k)]_c} \tag{84}$$

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

## A.5 Complete Mass Matrix Assembly

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

### A.5.1 Block Structure

$$\boxed{\mathbf{M} = \begin{pmatrix} \mathbf{M}^\mu & \mathbf{C}^{\mu\Sigma} \\ (\mathbf{C}^{\mu\Sigma})^T & \mathbf{M}^\Sigma \end{pmatrix}} \tag{85}$$

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

$$\mathbf{M}^\mu = \begin{pmatrix} [\mathbf{M}^\mu]_{11} & [\mathbf{M}^\mu]_{12} & \cdots \\ [\mathbf{M}^\mu]_{21} & [\mathbf{M}^\mu]_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad \mathbf{M}^\Sigma = \begin{pmatrix} [\mathbf{M}^\Sigma]_{11} & [\mathbf{M}^\Sigma]_{12} & \cdots \\ [\mathbf{M}^\Sigma]_{21} & [\mathbf{M}^\Sigma]_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{86}$$

### A.5.2 Explicit Formulae

**Mean sector diagonal:**

$$[\mathbf{M}^\mu]_{ii} = \underbrace{\bar{P}_i}_{\text{prior anchoring}} + \underbrace{\sum_k \beta_{ik} R_{ik} P_k R_{ik}^T}_{\text{incoming consensus}} + \underbrace{\sum_j \beta_{ji} P_i}_{\text{outgoing recoil}} \tag{87}$$

**Mean sector off-diagonal:**

$$[\mathbf{M}^{\mu}]_{ik} = -\beta_{ik} R_{ik} P_k - \beta_{ki} P_i R_{ki}^T \quad (i \neq k) \tag{88}$$

**Covariance sector diagonal:**

$$[\mathbf{M}^{\Sigma}]_{ii} = \frac{1}{2}(P_i \otimes P_i) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \tag{89}$$

**Cross mean-covariance (at consensus):**

$$[\mathbf{C}^{\mu\Sigma}]_{ik} = 0 \quad \text{when } \mu_i = \tilde{\mu}_k \tag{90}$$

## A.6 Physical Interpretation

### A.6.1 Mass as Precision

The mean-sector mass for agent $i$ is:

$$\boxed{M_i = \bar{P}_i + \sum_k \beta_{ik} \tilde{P}_k + \sum_j \beta_{ji} P_i} \tag{91}$$

- $\bar{P}_i$: **Bare mass** — inertia against deviation from prior
- $\sum_k \beta_{ik} \tilde{P}_k$: **Incoming relational mass** — inertia from being “pulled” by neighbors
- $\sum_j \beta_{ji} P_i$: **Outgoing relational mass** — inertia from “pulling” neighbors (recoil)

### A.6.2 Kinetic Energy

$$\boxed{T = \frac{1}{2}\dot{\mu}^T \mathbf{M}^{\mu} \dot{\mu} + \frac{1}{2}\mathrm{tr}\left[\mathbf{M}^{\Sigma}[\dot{\Sigma}, \dot{\Sigma}]\right]} \tag{92}$$

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

### A.6.3 Inter-Agent Kinetic Coupling

$$T_{\text{couple}} = \sum_{i<j} \beta_{ij} \left[-\dot{\mu}_i^T R_{ij} P_j \dot{\mu}_j + (\text{cov. terms})\right] \tag{93}$$

This represents **kinetic correlation**: when agent $j$ accelerates, agent $i$ feels a “drag” proportional to coupling strength and relative precision.

## A.7 Hamilton’s Equations

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

$$\boxed{H = \frac{1}{2}\langle \pi, \mathbf{M}^{-1}\pi \rangle + F[\xi]} \tag{94}$$

### A.7.1 Equations of Motion

$$\dot{\mu}_i = \sum_k [\mathbf{M}^{-1}]^{\mu\mu}_{ik} \pi^\mu_k + \sum_k [\mathbf{M}^{-1}]^{\mu\Sigma}_{ik} \Pi^\Sigma_k \tag{95}$$

$$\dot{\Sigma}_i = \sum_k [\mathbf{M}^{-1}]^{\Sigma\mu}_{ik} \pi^\mu_k + \sum_k [\mathbf{M}^{-1}]^{\Sigma\Sigma}_{ik} \Pi^\Sigma_k \tag{96}$$

$$\dot{\pi}^\mu_i = -\frac{\partial F}{\partial \mu_i} - \frac{1}{2}\pi^T \frac{\partial \mathbf{M}^{-1}}{\partial \mu_i} \pi \tag{97}$$

$$\dot{\Pi}^\Sigma_i = -\frac{\partial F}{\partial \Sigma_i} - \frac{1}{2}\pi^T \frac{\partial \mathbf{M}^{-1}}{\partial \Sigma_i} \pi \tag{98}$$

### A.7.2 Force Terms

The potential forces are:

$$-\frac{\partial F}{\partial \mu_i} = -\bar{P}_i(\mu_i - \bar{\mu}_i) - \sum_k \beta_{ik} \tilde{P}_k (\mu_i - \tilde{\mu}_k) - \sum_j \beta_{ji} P_i R^T_{ji} (\tilde{\mu}^{(j)}_i - \mu_j) \tag{99}$$

$$-\frac{\partial F}{\partial \Sigma_i} = -\frac{1}{2}[\bar{P}_i - P_i] - \sum_k \frac{\beta_{ik}}{2} [\tilde{P}_k - P_i] \tag{100}$$

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

## A.8 Isotropic Simplification

With $\Sigma_i = \sigma_i^2 I$, $\bar{\Sigma}_i = \bar{\sigma}_i^2 I$, and $R_{ik} \in \mathrm{SO}(d)$:

### A.8.1 Reduced Variables

$$\xi_i = (\mu_i, \sigma_i) \in \mathbb{R}^d \times \mathbb{R}^+ \tag{101}$$

### A.8.2 Scalar Mass

$$\boxed{m_i = \frac{1}{\bar{\sigma}_i^2} + \sum_k \frac{\beta_{ik}}{\sigma_k^2} + \sum_j \frac{\beta_{ji}}{\sigma_i^2}} \tag{102}$$

**Mass = Total precision** (prior + consensus partners).

### A.8.3 Simplified Hessian

$$[\mathbf{M}^{\mu}]_{ii} = \frac{1}{\bar{\sigma}_i^2} I + \sum_k \frac{\beta_{ik}}{\sigma_k^2} R_{ik} R_{ik}^T \tag{103}$$

$$[\mathbf{M}^{\mu}]_{ik} = -\frac{\beta_{ik} + \beta_{ki}}{\sigma_k^2} R_{ik} \quad (i \neq k, \text{ symmetric case}) \tag{104}$$

$$[\mathbf{M}^{\sigma}]_{ii} = \frac{d}{\sigma_i^2} \quad \text{(hyperbolic geometry in log-variance)} \tag{105}$$

### A.8.4 Force

$$f_i = -\frac{\mu_i - \bar{\mu}_i}{\bar{\sigma}_i^2} - \sum_k \frac{\beta_{ik}}{\sigma_k^2}(\mu_i - \tilde{\mu}_k) - \sum_j \frac{\beta_{ji}}{\sigma_i^2} R_{ji}^T (\tilde{\mu}_i^{(j)} - \mu_j) \tag{106}$$

## A.9 Summary

[title=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:**

$$F = \sum_i \mathrm{KL}(q_i \| p_i) + \sum_{i,k} \beta_{ik} \mathrm{KL}(q_i \| \Omega_{ik}[q_k]) \tag{107}$$

**Mass Matrix:**

$$\mathbf{M} = \frac{\partial^2 F}{\partial \xi \partial \xi} = \text{Fisher information} = \text{Precision} \tag{108}$$

**Dynamics:**

$$\dot{\xi} = \mathbf{M}^{-1}\pi, \qquad \dot{\pi} = -\nabla F - \frac{1}{2}\nabla_{\xi}\langle \pi, \mathbf{M}^{-1}\pi \rangle \tag{109}$$

**Physical Meaning:**

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

# B Gauge Frame Variations and Pullback Geometry

The Hamiltonian formulation of belief dynamics is not merely a mathematical convenience but 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.

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

### B.1.1 The Principal Bundle

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 $k$'s frame to agent $i$'s frame.

### B.1.2 Gauge Transformations

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

$$g : \{1, \ldots, N\} \to \mathrm{SO}(d), \quad i \mapsto g_i \tag{110}$$

Under this transformation, belief parameters transform as:

$$\mu_i \mapsto \mu_i' = g_i \mu_i \tag{111}$$

$$\Sigma_i \mapsto \Sigma_i' = g_i \Sigma_i g_i^T \tag{112}$$

$$\Lambda_i \mapsto \Lambda_i' = g_i \Lambda_i g_i^T \tag{113}$$

The transport operators transform as:

$$\Omega_{ik} \mapsto \Omega_{ik}' = g_i \Omega_{ik} g_k^{-1} \tag{114}$$

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

$$\tilde{\mu}_k' = g_i \tilde{\mu}_k, \quad \tilde{\Lambda}_k' = g_i \tilde{\Lambda}_k g_i^T \tag{115}$$

## B.2 Transformation of the Mass Matrix

### B.2.1 Mean Sector

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

**Diagonal blocks:**

$$[\mathbf{M}^\mu]_{ii}' = \bar{\Lambda}_i' + \sum_k \beta_{ik} \tilde{\Lambda}_k' + \sum_j \beta_{ji} \Lambda_i' \tag{116}$$

$$= g_i \bar{\Lambda}_i g_i^T + \sum_k \beta_{ik} g_i \tilde{\Lambda}_k g_i^T + \sum_j \beta_{ji} g_i \Lambda_i g_i^T \tag{117}$$

$$= g_i \left[ \bar{\Lambda}_i + \sum_k \beta_{ik} \tilde{\Lambda}_k + \sum_j \beta_{ji} \Lambda_i \right] g_i^T \tag{118}$$

$$= g_i \, [\mathbf{M}^\mu]_{ii} \, g_i^T \tag{119}$$

**Off-diagonal blocks:**

$$
\begin{aligned}
[\mathbf{M}^\mu]'_{ik} &= -\beta_{ik}\Omega'_{ik}\Lambda'_k - \beta_{ki}\Lambda'_i(\Omega'_{ki})^T && (120)\\
&= -\beta_{ik}(g_i\Omega_{ik}g_k^{-1})(g_k\Lambda_k g_k^T) - \beta_{ki}(g_i\Lambda_i g_i^T)(g_k\Omega_{ki}g_i^{-1})^T && (121)\\
&= -\beta_{ik}g_i\Omega_{ik}\Lambda_k g_k^T - \beta_{ki}g_i\Lambda_i\Omega_{ki}^T g_k^T && (122)\\
&= g_i\,[\mathbf{M}^\mu]_{ik}\,g_k^T && (123)
\end{aligned}
$$

**Block matrix form:** Define the block-diagonal gauge matrix:

$$
\mathbf{G} = \mathrm{diag}(g_1, g_2, \ldots, g_N) \in \mathrm{SO}(d)^N \tag{124}
$$

Then the full mean-sector mass matrix transforms as:

$$
\boxed{(\mathbf{M}^\mu)' = \mathbf{G}\,\mathbf{M}^\mu\,\mathbf{G}^T} \tag{125}
$$

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

### B.2.2 Covariance Sector

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

$$
\begin{aligned}
[\mathbf{M}^\Sigma]'_{ii} &= \frac{1}{2}(\Lambda'_i \otimes \Lambda'_i)\cdot\mathcal{S}\cdot(\ldots) && (126)\\
&= \frac{1}{2}(g_i\Lambda_i g_i^T \otimes g_i\Lambda_i g_i^T)\cdot\mathcal{S}\cdot(\ldots) && (127)\\
&= \frac{1}{2}(g_i \otimes g_i)(\Lambda_i \otimes \Lambda_i)(g_i^T \otimes g_i^T)\cdot\mathcal{S}\cdot(\ldots) && (128)
\end{aligned}
$$

The transformation law is:

$$
\boxed{(\mathbf{M}^\Sigma)' = (\mathbf{G}\otimes\mathbf{G})\,\mathbf{M}^\Sigma\,(\mathbf{G}^T\otimes\mathbf{G}^T)} \tag{129}
$$

### B.2.3 Cross Blocks

The mean-covariance cross blocks transform as:

$$
(\mathbf{C}^{\mu\Sigma})' = \mathbf{G}\,\mathbf{C}^{\mu\Sigma}\,(\mathbf{G}^T\otimes\mathbf{G}^T) \tag{130}
$$

## B.3 Transformation of Momenta

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

### B.3.1 Mean Momentum

The mean momentum transforms as a covector:

$$
\boxed{(\pi_i^\mu)' = g_i\,\pi_i^\mu} \tag{131}
$$

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

$$
\langle (\pi^\mu)', \dot{\mu}'\rangle = (g_i\pi_i^\mu)^T(g_i\dot{\mu}_i) = (\pi_i^\mu)^T g_i^T g_i\dot{\mu}_i = (\pi_i^\mu)^T\dot{\mu}_i = \langle \pi^\mu, \dot{\mu}\rangle \tag{132}
$$

### B.3.2 Covariance Momentum

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

$$\boxed{(\Pi_i^\Sigma)' = g_i\,\Pi_i^\Sigma\,g_i^T} \tag{133}$$

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

$$\mathrm{tr}[(\Pi^\Sigma)'\dot{\Sigma}'] = \mathrm{tr}[(g_i\Pi_i^\Sigma g_i^T)(g_i\dot{\Sigma}_i g_i^T)] = \mathrm{tr}[\Pi_i^\Sigma\dot{\Sigma}_i] \tag{134}$$

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

## B.4 Covariance of Hamilton's Equations

### B.4.1 Velocity Equation

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

$$\begin{aligned}
\dot{\mu}' &= ((\mathbf{M}^\mu)')^{-1}(\pi^\mu)' && (135)\\
&= (\mathbf{G}\mathbf{M}^\mu\mathbf{G}^T)^{-1}\mathbf{G}\pi^\mu && (136)\\
&= \mathbf{G}^{-T}(\mathbf{M}^\mu)^{-1}\mathbf{G}^{-1}\mathbf{G}\pi^\mu && (137)\\
&= \mathbf{G}(\mathbf{M}^\mu)^{-1}\pi^\mu \quad (\text{since } \mathbf{G}^{-T} = \mathbf{G} \text{ for } \mathrm{SO}(d)) && (138)\\
&= \mathbf{G}\dot{\mu} && (139)
\end{aligned}$$

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

### B.4.2 Force Equation

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

$$\left(\frac{\partial F}{\partial \mu_i}\right)' = g_i \frac{\partial F}{\partial \mu_i} \tag{140}$$

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

The geodesic force transforms similarly, ensuring full covariance:

$$\boxed{\dot{\pi}' = \mathbf{G}\dot{\pi}} \tag{141}$$

## B.5 The Connection and Its Variation

### B.5.1 Connection 1-Form

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

$$A_e = \Omega_{ik} \in \mathrm{SO}(d) \tag{142}$$

Under gauge transformation:

$$A_e \mapsto A'_e = g_i A_e g_k^{-1} \tag{143}$$

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

### B.5.2 Curvature

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

$$F_\gamma = \Omega_{ij}\Omega_{jk}\Omega_{ki} \in \mathrm{SO}(d) \tag{144}$$

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.

### B.5.3 Variation of Connection

Consider an infinitesimal variation of the connection:

$$\delta\Omega_{ik} = \omega_{ik}\,\Omega_{ik}, \quad \omega_{ik} \in \mathfrak{so}(d) \tag{145}$$

The variation of transported precision is:

$$\delta\tilde{\Lambda}_k = \omega_{ik}\tilde{\Lambda}_k + \tilde{\Lambda}_k\omega_{ik}^T = [\omega_{ik}, \tilde{\Lambda}_k]_+ \tag{146}$$

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

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

How does the mass matrix change when we vary the connection?

### B.6.1 Diagonal Block Variation

$$\delta[\mathbf{M}^\mu]_{ii} = \sum_k \beta_{ik}\,\delta\tilde{\Lambda}_k \tag{147}$$

$$= \sum_k \beta_{ik}\,[\omega_{ik}, \tilde{\Lambda}_k]_+ \tag{148}$$

### B.6.2 Off-Diagonal Block Variation

$$\delta[\mathbf{M}^\mu]_{ik} = -\beta_{ik}\,\delta(\Omega_{ik}\Lambda_k) - \beta_{ki}\,\delta(\Lambda_i\Omega_{ki}^T) \tag{149}$$

$$= -\beta_{ik}\omega_{ik}\Omega_{ik}\Lambda_k - \beta_{ki}\Lambda_i(\omega_{ki}\Omega_{ki})^T \tag{150}$$

$$= -\beta_{ik}\omega_{ik}\Omega_{ik}\Lambda_k + \beta_{ki}\Lambda_i\Omega_{ki}^T\omega_{ki}^T \tag{151}$$

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

$$\boxed{\delta[\mathbf{M}^\mu]_{ik} = -\beta_{ik}\omega_{ik}\Omega_{ik}\Lambda_k - \beta_{ki}\Lambda_i\Omega_{ki}^T\omega_{ki}} \tag{152}$$

## 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.

### B.7.1 Pullback of the Fisher-Rao Metric

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

$$\phi_g(\mu, \Sigma) = (g\mu, g\Sigma g^T) \tag{153}$$

The pullback metric is:

$$(\phi_g^* \mathcal{G})_{(\mu,\Sigma)}(v, w) = \mathcal{G}_{\phi_g(\mu,\Sigma)}(d\phi_g \cdot v, d\phi_g \cdot w) \tag{154}$$

For the Fisher-Rao metric, gauge invariance implies:

$$\boxed{\phi_g^* \mathcal{G} = \mathcal{G}} \tag{155}$$

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

### B.7.2 Horizontal and Vertical Decomposition

The tangent space at each point decomposes as:

$$T_{(\mu,\Sigma)}\mathcal{Q} = H_{(\mu,\Sigma)} \oplus V_{(\mu,\Sigma)} \tag{156}$$

- **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:

$$\mathcal{G}(v, \xi_X) = 0 \quad \forall X \in \mathfrak{so}(d) \tag{157}$$

where $\xi_X$ is the vector field generated by $X$.

### B.7.3 Physical (Gauge-Invariant) Quantities

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

1. **Consensus divergence**: $\|\mu_i - \tilde{\mu}_k\|^2_{\tilde{\Lambda}_k} = (\mu_i - \tilde{\mu}_k)^T \tilde{\Lambda}_k (\mu_i - \tilde{\mu}_k)$
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

## B.8 Gauge-Fixed Dynamics

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

### B.8.1 Identity Gauge

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

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

### B.8.2 Consensus-Aligned Gauge

Choose gauges so that at equilibrium:

$$\Omega_{ik}^* = I \quad \text{(parallel frames at consensus)} \tag{158}$$

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

### B.8.3 Principal Axis Gauge

For each agent, choose $g_i$ to diagonalize $\Sigma_i$:

$$\Sigma_i' = g_i \Sigma_i g_i^T = \mathrm{diag}(\lambda_1^{(i)}, \ldots, \lambda_d^{(i)}) \tag{159}$$

This separates dynamics along principal axes of uncertainty.

## B.9 Infinitesimal Gauge Transformations and Noether Currents

Continuous gauge symmetry implies conserved quantities via Noether's theorem.

### B.9.1 Infinitesimal Transformation

For $g_i = \exp(\epsilon X_i)$ with $X_i \in \mathfrak{so}(d)$ and $\epsilon \ll 1$:

$$\delta\mu_i = \epsilon X_i \mu_i \tag{160}$$

$$\delta\Sigma_i = \epsilon(X_i\Sigma_i + \Sigma_i X_i^T) = \epsilon[X_i, \Sigma_i]_+ \tag{161}$$

### B.9.2 Noether Current

The conserved current associated with gauge symmetry is the **angular momentum** in belief space:

$$\boxed{L_i = \mu_i \wedge \pi_i^\mu + \Sigma_i \wedge \Pi_i^\Sigma} \tag{162}$$

where $\wedge$ denotes the appropriate antisymmetric product.

For a global gauge transformation ($X_i = X$ for all $i$), total angular momentum is conserved:

$$\frac{d}{dt} \sum_i L_i = 0 \tag{163}$$

This reflects the rotational symmetry of the free energy in belief space.

## B.10 Summary: Gauge-Covariant Hamiltonian Mechanics

[title=Gauge Transformation Laws] **Positions:**

$$\mu_i' = g_i \mu_i \qquad \Sigma_i' = g_i \Sigma_i g_i^T \tag{164}$$

**Momenta:**

$$(\pi_i^\mu)' = g_i \pi_i^\mu \qquad (\Pi_i^\Sigma)' = g_i \Pi_i^\Sigma g_i^T \tag{165}$$

**Mass Matrix:**

$$\mathbf{M}' = \mathbf{G}\mathbf{M}\mathbf{G}^T \tag{166}$$

**Transport Operators:**

$$\Omega_{ik}' = g_i \Omega_{ik} g_k^{-1} \tag{167}$$

**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


Amari, S.-i. (2016). *Information geometry and its applications.* Springer.

Anderson, C. A., Lepper, M. R., and Ross, L. (1980). Perseverance of social theories. *Journal of Personality and Social Psychology*, 39(6):1037.

Arnold, V. I. (1989). *Mathematical methods of classical mechanics.* Springer.

Bernevig, B. A. and Hughes, T. L. (2013). *Topological insulators and topological superconductors*. Princeton University Press.

Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology*, 76:198–211.

Busemeyer, J. R. and Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge University Press.

Castellano, C., Fortunato, S., and Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, 81(2):591.

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204.

Dennis, R. C. (2025a). Attention, transformers, and backpropagation are degenerate limits of the variational free energy principle. Submitted to Journal of Machine Learning Research.

Dennis, R. C. (2025b). A theoretical and computational implementation of a participatory "it from bit" universe. Submitted to Foundations of Physics.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, 4(11):e1000211.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., and Pezzulo, G. (2016). Active inference and learning. *Neuroscience & Biobehavioral Reviews*, 68:862–879.

Galam, S. (2012). *Sociophysics: A physicist's modeling of psycho-political phenomena*. Springer.

Goldenfeld, N. (1992). *Lectures on phase transitions and the renormalization group*. Addison-Wesley.

Hegselmann, R. and Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. In *Journal of Artificial Societies and Social Simulation*, volume 5.

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Holmes, M. H. (2012). *Introduction to perturbation methods*. Springer.

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kobayashi, S. and Nomizu, K. (1963). *Foundations of differential geometry*, volume 1. Interscience Publishers.

Kruger, J. and Dunning, D. (1999). Unskilled and unaware of it. *Journal of Personality and Social Psychology*, 77(6):1121.

Millidge, B., Seth, A., and Buckley, C. L. (2021). Predictive coding: a theoretical and experimental review. *arXiv preprint arXiv:2107.12979*.

Nakahara, M. (2003). *Geometry, topology and physics*. CRC Press.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2):175–220.

Nielsen, F. (2020). *Elementary differential geometry*. Springer.

Olver, P. J. (1993). *Applications of Lie groups to differential equations*. Springer.

Parr, T., Pezzulo, G., and Friston, K. J. (2022). *Active inference: The free energy principle in mind, brain, and behavior*. MIT Press.

Sornette, D. (2006). *Critical phenomena in natural sciences*. Springer, 2nd edition.

Strogatz, S. H. (2015). *Nonlinear dynamics and chaos*. Westview Press, 2nd edition.

Wilson, K. G. and Kogut, J. (1975). The renormalization group and the  expansion. *Physics Reports*, 12(2):75–199.