The Inertia of Belief

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

December 5, 2025

1 Hamiltonian Mechanics on Statistical Manifolds

This appendix derives the complete mass matrix structure for multi-agent belief dynamics with explicit sensory evidence, 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).

1.1 Setup and Notation

Each agent i maintains a belief distribution $q_i = \mathcal{N}(\mu_i, \Sigma_i)$ anchored to a fixed prior $p_i = \mathcal{N}(\bar{\mu}_i, \bar{\Sigma}_i)$ and receives observations o_i through a likelihood $p(o_i \mid \theta) = \mathcal{N}(o_i; \theta, \Sigma_{o_i})$. Define:

$$\Lambda_{qi} = \Sigma_i^{-1} \quad \text{(belief precision)} \tag{1}$$
$$\bar{\Lambda}_{pi} = \bar{\Sigma}_i^{-1} \quad \text{(prior precision)} \tag{2}$$
$$\Lambda_{o_i} = \Sigma_{o_i}^{-1} \quad \text{(observation precision)} \tag{3}$$
$$\tilde{\mu}_k = \Omega_{ik}\mu_k \quad \text{(transported mean)} \tag{4}$$
$$\tilde{\Lambda}_{qk} = \Omega_{ik}\Lambda_{qk}\Omega_{ik}^T \quad \text{(transported precision)} \tag{5}$$

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 $\phi_i \in \mathfrak{so}(d)$.

1.2 The Extended Free Energy Functional

The complete variational free energy with explicit sensory evidence is:

$$\boxed{\mathcal{F}[\{q_i\}] = \sum_i D_{\mathrm{KL}}(q_i \| p_i) + \sum_{i,k} \beta_{ik} D_{\mathrm{KL}}(q_i \| \Omega_{ik}[q_k]) - \sum_i \mathbb{E}_{q_i}[\log p(o_i \mid \theta)]} \tag{6}$$

The three terms represent:

1. **Prior anchoring**: Deviation from internal world-model
2. **Social consensus**: Alignment with other agents via gauge-covariant transport
3. **Sensory evidence**: Grounding in observations

1.3 Component Free Energies for Gaussians

1.3.1 KL Divergence Between Gaussians

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

$$D_{\mathrm{KL}}(q\|p) = \frac{1}{2}\left[\mathrm{tr}(\Sigma_p^{-1}\Sigma_q) + (\mu_p - \mu_q)^T\Sigma_p^{-1}(\mu_p - \mu_q) - d + \ln\frac{|\Sigma_p|}{|\Sigma_q|}\right] \tag{7}$$

1.3.2 Expected Log-Likelihood

For the Gaussian likelihood $p(o_i \mid \theta) = \mathcal{N}(o_i; \theta, \Sigma_{o_i})$:

$$\mathbb{E}_{q_i}[\log p(o_i \mid \theta)] = -\frac{d}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_{o_i}| - \frac{1}{2}\mathbb{E}_{q_i}\left[(o_i - \theta)^T\Lambda_{o_i}(o_i - \theta)\right] \tag{8}$$

The quadratic expectation evaluates to:

$$\mathbb{E}_{q_i}\left[(o_i - \theta)^T\Lambda_{o_i}(o_i - \theta)\right] = (o_i - \mu_i)^T\Lambda_{o_i}(o_i - \mu_i) + \mathrm{tr}(\Lambda_{o_i}\Sigma_i) \tag{9}$$

Therefore:

$$\boxed{-\mathbb{E}_{q_i}[\log p(o_i \mid \theta)] = \frac{1}{2}(o_i - \mu_i)^T\Lambda_{o_i}(o_i - \mu_i) + \frac{1}{2}\mathrm{tr}(\Lambda_{o_i}\Sigma_i) + \mathrm{const}} \tag{10}$$

1.4 First Variations (Gradient)

1.4.1 Prior Term: $D_{\mathrm{KL}}(q_i\|p_i)$

$$\frac{\partial D_{\mathrm{KL}}(q_i\|p_i)}{\partial\mu_i} = \bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i) \tag{11}$$

$$\frac{\partial D_{\mathrm{KL}}(q_i\|p_i)}{\partial\Sigma_i} = \frac{1}{2}(\bar{\Lambda}_{pi} - \Lambda_{qi}) \tag{12}$$

1.4.2 Consensus Term: $D_{\mathrm{KL}}(q_i\|\tilde{q}_k)$

With respect to receiver i:

$$\frac{\partial D_{\mathrm{KL}}(q_i\|\tilde{q}_k)}{\partial\mu_i} = \tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k) \tag{13}$$

$$\frac{\partial D_{\mathrm{KL}}(q_i\|\tilde{q}_k)}{\partial\Sigma_i} = \frac{1}{2}(\tilde{\Lambda}_{qk} - \Lambda_{qi}) \tag{14}$$

With respect to sender k:

$$\frac{\partial D_{\mathrm{KL}}(q_i \| \tilde{q}_k)}{\partial \mu_k} = \Lambda_{qk} \Omega_{ik}^T (\tilde{\mu}_k - \mu_i) \tag{15}$$

$$\frac{\partial D_{\mathrm{KL}}(q_i \| \tilde{q}_k)}{\partial \Sigma_k} = \frac{1}{2} \Omega_{ik}^T \left[\tilde{\Lambda}_{qk} - \tilde{\Lambda}_{qk} \Sigma_i \tilde{\Lambda}_{qk} \right] \Omega_{ik} \tag{16}$$

1.4.3 Sensory Term: $-\mathbb{E}_{q_i}[\log p(o_i \mid \theta)]$

$$\frac{\partial}{\partial \mu_i} \left[-\mathbb{E}_{q_i}[\log p(o_i \mid \theta)] \right] = \Lambda_{o_i} (\mu_i - o_i) \tag{17}$$

$$\frac{\partial}{\partial \Sigma_i} \left[-\mathbb{E}_{q_i}[\log p(o_i \mid \theta)] \right] = \frac{1}{2} \Lambda_{o_i} \tag{18}$$

1.4.4 Total Gradient

$$\boxed{\frac{\partial \mathcal{F}}{\partial \mu_i} = \bar{\Lambda}_{pi} (\mu_i - \bar{\mu}_i) + \sum_k \beta_{ik} \tilde{\Lambda}_{qk} (\mu_i - \tilde{\mu}_k) + \sum_j \beta_{ji} \Lambda_{qi} \Omega_{ji}^T (\tilde{\mu}_i^{(j)} - \mu_j) + \Lambda_{o_i} (\mu_i - o_i)} \tag{19}$$

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

$$\frac{\partial \mathcal{F}}{\partial \Sigma_i} = \frac{1}{2} (\bar{\Lambda}_{pi} - \Lambda_{qi}) + \sum_k \frac{\beta_{ik}}{2} (\tilde{\Lambda}_{qk} - \Lambda_{qi}) + \sum_j \frac{\beta_{ji}}{2} \Omega_{ji}^T \left[\tilde{\Lambda}_{qi}^{(j)} - \tilde{\Lambda}_{qi}^{(j)} \Sigma_j \tilde{\Lambda}_{qi}^{(j)} \right] \Omega_{ji} + \frac{1}{2} \Lambda_{o_i} \tag{20}$$

1.5 Second Variations (Hessian = Mass Matrix)

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

1.5.1 Mean Sector: $\partial^2 \mathcal{F} / \partial \mu \partial \mu^T$

Diagonal blocks $(i = k)$**:** From prior:

$$\frac{\partial^2 D_{\mathrm{KL}}(q_i \| p_i)}{\partial \mu_i \partial \mu_i^T} = \bar{\Lambda}_{pi} \tag{21}$$

From consensus (as receiver):

$$\frac{\partial^2 D_{\mathrm{KL}}(q_i \| \tilde{q}_k)}{\partial \mu_i \partial \mu_i^T} = \tilde{\Lambda}_{qk} = \Omega_{ik} \Lambda_{qk} \Omega_{ik}^T \tag{22}$$

From consensus (as sender to agent j):

$$\frac{\partial^2 D_{\mathrm{KL}}(q_j \| \tilde{q}_i)}{\partial \mu_i \partial \mu_i^T} = \Omega_{ji}^T \tilde{\Lambda}_{qi}^{(j)} \Omega_{ji} = \Lambda_{qi} \tag{23}$$

From sensory evidence:

$$\frac{\partial^2}{\partial\mu_i\partial\mu_i^T}\left[-\mathbb{E}_{q_i}[\log p(o_i \mid \theta)]\right] = \Lambda_{o_i} \tag{24}$$

Total diagonal mass:

$$\boxed{[\mathbf{M}^\mu]_{ii} = \underbrace{\bar{\Lambda}_{pi}}_{\text{prior}} + \underbrace{\sum_k \beta_{ik}\tilde{\Lambda}_{qk}}_{\text{incoming social}} + \underbrace{\sum_j \beta_{ji}\Lambda_{qi}}_{\text{outgoing recoil}} + \underbrace{\Lambda_{o_i}}_{\text{sensory}}} \tag{25}$$

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

$$\frac{\partial^2 D_{\mathrm{KL}}(q_i\|\tilde{q}_k)}{\partial\mu_i\partial\mu_k^T} = -\tilde{\Lambda}_{qk}\Omega_{ik} = -\Omega_{ik}\Lambda_{qk} \tag{26}$$

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

$$\frac{\partial^2 D_{\mathrm{KL}}(q_k\|\tilde{q}_i)}{\partial\mu_i\partial\mu_k^T} = -\Lambda_{qi}\Omega_{ki}^T \tag{27}$$

The sensory term does not couple different agents. Therefore:

$$\boxed{[\mathbf{M}^\mu]_{ik} = -\beta_{ik}\Omega_{ik}\Lambda_{qk} - \beta_{ki}\Lambda_{qi}\Omega_{ki}^T \quad (i \neq k)} \tag{28}$$

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

1.5.2 Covariance Sector: $\partial^2\mathcal{F}/\partial\Sigma\partial\Sigma$

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

$$\frac{\partial^2 f}{\partial\Sigma\partial\Sigma}[V,W] = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\left.\frac{\partial f}{\partial\Sigma}\right|_{\Sigma+\epsilon W} - \left.\frac{\partial f}{\partial\Sigma}\right|_{\Sigma}\right)[V] \tag{29}$$

Key identity:

$$\frac{\partial}{\partial\Sigma}(\Sigma^{-1}) = -\Sigma^{-1}\otimes\Sigma^{-1} \tag{30}$$

Diagonal blocks $(i = k)$**:** From prior:

$$\frac{\partial^2 D_{\mathrm{KL}}(q_i\|p_i)}{\partial\Sigma_i\partial\Sigma_i}[V,W] = \frac{1}{2}\mathrm{tr}\left[\Lambda_{qi}V\Lambda_{qi}W\right] \tag{31}$$

In tensor notation:

$$\frac{\partial^2 D_{\mathrm{KL}}(q_i\|p_i)}{\partial\Sigma_i\partial\Sigma_i} = \frac{1}{2}(\Lambda_{qi}\otimes\Lambda_{qi}) \tag{32}$$

From consensus (as receiver and sender), identical contributions arise.

Critical observation: The sensory term $\frac{1}{2}\text{tr}(\Lambda_{o_i}\Sigma_i)$ is *linear* in Σ_i, so its second derivative **vanishes**:

$$\frac{\partial^2}{\partial\Sigma_i\partial\Sigma_i}[\text{tr}(\Lambda_{o_i}\Sigma_i)] = 0 \tag{33}$$

Therefore:

$$\boxed{[\mathbf{M}^\Sigma]_{ii} = \frac{1}{2}(\Lambda_{qi} \otimes \Lambda_{qi}) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right)} \tag{34}$$

The sensory precision Λ_{o_i} does **not** contribute to the covariance-sector mass.

1.5.3 Mean-Covariance Cross Blocks

Prior term:

$$\frac{\partial^2 D_{\text{KL}}(q_i \| p_i)}{\partial\mu_i\partial\Sigma_i} = 0 \tag{35}$$

Sensory term: The sensory free energy decomposes as:

- Quadratic in μ_i: $(o_i - \mu_i)^T\Lambda_{o_i}(o_i - \mu_i)$
- Linear in Σ_i: $\text{tr}(\Lambda_{o_i}\Sigma_i)$

These are independent, so:

$$[\mathbf{C}^{\mu\Sigma}]_{ii}^{\text{sensory}} = 0 \tag{36}$$

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

$$\frac{\partial^2 D_{\text{KL}}(q_i\|\tilde{q}_k)}{\partial\mu_i\partial\Sigma_k}[V] = -\Omega_{ik}\Lambda_{qk}V\Lambda_{qk}\Omega_{ik}^T(\mu_i - \tilde{\mu}_k) \tag{37}$$

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

$$[\mathbf{C}^{\mu\Sigma}]_{ik} = 0 \quad \text{when } \mu_i = \tilde{\mu}_k \tag{38}$$

1.6 Complete Mass Matrix Assembly

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

1.6.1 Block Structure

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}^\mu & \mathbf{C}^{\mu\Sigma} \\ (\mathbf{C}^{\mu\Sigma})^T & \mathbf{M}^\Sigma \end{pmatrix} \tag{39}$$

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

1.6.2 Explicit Formulae

Mean sector diagonal:

$$[\mathbf{M}^\mu]_{ii} = \underbrace{\bar{\Lambda}_{pi}}_{\text{prior anchoring}} + \underbrace{\sum_k \beta_{ik}\Omega_{ik}\Lambda_{qk}\Omega_{ik}^T}_{\text{incoming consensus}} + \underbrace{\sum_j \beta_{ji}\Lambda_{qi}}_{\text{outgoing recoil}} + \underbrace{\Lambda_{o_i}}_{\text{sensory grounding}} \tag{40}$$

Mean sector off-diagonal:

$$[\mathbf{M}^\mu]_{ik} = -\beta_{ik}\Omega_{ik}\Lambda_{qk} - \beta_{ki}\Lambda_{qi}\Omega_{ki}^T \quad (i \neq k) \tag{41}$$

Covariance sector diagonal:

$$[\mathbf{M}^\Sigma]_{ii} = \frac{1}{2}(\Lambda_{qi} \otimes \Lambda_{qi}) \cdot \left(1 + \sum_k \beta_{ik} + \sum_j \beta_{ji}\right) \tag{42}$$

Cross mean-covariance (at consensus):

$$[\mathbf{C}^{\mu\Sigma}]_{ik} = 0 \quad \text{when } \mu_i = \tilde{\mu}_k \tag{43}$$

1.7 Physical Interpretation

1.7.1 Mass as Precision

The mean-sector effective mass for agent i is:

$$\boxed{M_i = \bar{\Lambda}_{pi} + \sum_k \beta_{ik}\tilde{\Lambda}_{qk} + \sum_j \beta_{ji}\Lambda_{qi} + \Lambda_{o_i}} \tag{44}$$

- $\bar{\Lambda}_{pi}$: **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)
- Λ_{o_i}: **Sensory mass** — inertia from grounding in observations

1.7.2 Asymmetry of Sensory Contribution

The sensory precision Λ_{o_i} contributes to:

1. The mean-sector mass (Eq. 40)
2. The mean-sector force (Eq. 19)

But **not** to:

1. The covariance-sector mass (Eq. 34)

This asymmetry arises because the sensory term is quadratic in μ but only linear in Σ.

1.7.3 Kinetic Energy

$$T = \frac{1}{2}\dot{\mu}^T \mathbf{M}^\mu \dot{\mu} + \frac{1}{2}\mathrm{tr}\left[\mathbf{M}^\Sigma[\dot{\Sigma}, \dot{\Sigma}]\right] \tag{45}$$

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

1.8 The Hamiltonian

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

$$\boxed{H = \frac{1}{2}\langle \pi, \mathbf{M}^{-1}\pi\rangle + \mathcal{F}[\xi]} \tag{46}$$

1.9 Hamilton's Equations

1.9.1 Equations of Motion

$$\dot{\mu}_i = \sum_k [\mathbf{M}^{-1}]^{\mu\mu}_{ik}\pi^\mu_k + \sum_k [\mathbf{M}^{-1}]^{\mu\Sigma}_{ik}\Pi^\Sigma_k \tag{47}$$

$$\dot{\Sigma}_i = \sum_k [\mathbf{M}^{-1}]^{\Sigma\mu}_{ik}\pi^\mu_k + \sum_k [\mathbf{M}^{-1}]^{\Sigma\Sigma}_{ik}\Pi^\Sigma_k \tag{48}$$

$$\dot{\pi}^\mu_i = -\frac{\partial \mathcal{F}}{\partial \mu_i} - \frac{1}{2}\pi^T \frac{\partial \mathbf{M}^{-1}}{\partial \mu_i}\pi \tag{49}$$

$$\dot{\Pi}^\Sigma_i = -\frac{\partial \mathcal{F}}{\partial \Sigma_i} - \frac{1}{2}\pi^T \frac{\partial \mathbf{M}^{-1}}{\partial \Sigma_i}\pi \tag{50}$$

1.9.2 Force Decomposition

The potential forces decompose into four physically distinct contributions:

$$\boxed{-\frac{\partial \mathcal{F}}{\partial \mu_i} = \underbrace{-\bar{\Lambda}_{pi}(\mu_i - \bar{\mu}_i)}_{\text{prior restoring}} \underbrace{- \sum_k \beta_{ik}\tilde{\Lambda}_{qk}(\mu_i - \tilde{\mu}_k)}_{\text{consensus}} \underbrace{- \sum_j \beta_{ji}\Lambda_{qi}\Omega^T_{ji}(\tilde{\mu}^{(j)}_i - \mu_j)}_{\text{reciprocal}} \underbrace{-\Lambda_{o_i}(\mu_i - o_i)}_{\text{sensory evidence}}} \tag{51}$$

The geodesic force $f^{\text{geo}}_i = -\frac{1}{2}\sum_{jkl}(\pi^\mu_j)^T \frac{\partial [\mathbf{M}^{-1}]^{\mu\mu}_{jk}}{\partial \mu_i}\pi^\mu_k$ encodes manifold curvature.

1.9.3 Compact Form

$$\boxed{\begin{aligned} \dot{\xi} &= \mathbf{M}^{-1}\pi \\ \dot{\pi} &= -\nabla \mathcal{F} - \frac{1}{2}\nabla_\xi \langle \pi, \mathbf{M}^{-1}\pi\rangle \end{aligned}} \tag{52}$$

with $dH/dt = 0$ along trajectories.

1.10 Damped Dynamics

Including dissipation yields:

$$M_i \ddot{\mu}_i + \gamma_i \dot{\mu}_i + \nabla_{\mu_i} \mathcal{F} = 0 \tag{53}$$

For small displacements from equilibrium with stiffness $K_i = \nabla^2 \mathcal{F}|_{\mu^*}$:

$$M_i \ddot{\delta\mu} + \gamma_i \dot{\delta\mu} + K_i \delta\mu = 0 \tag{54}$$

The discriminant $\Delta = \gamma_i^2 - 4K_i M_i$ determines three regimes:

- **Overdamped** $(\Delta > 0)$: Monotonic decay, standard Bayesian updating
- **Critically damped** $(\Delta = 0)$: Fastest equilibration
- **Underdamped** $(\Delta < 0)$: Oscillatory approach with overshooting

1.11 Momentum Current with Sensory Coupling

Between agents, the momentum current is:

$$J_{k \to i} = \beta_{ik} \tilde{\Lambda}_{qk} (\tilde{\mu}_k - \mu_i) \tag{55}$$

The continuity equation becomes:

$$\dot{\pi}_i + \gamma_i \dot{\mu}_i + \bar{\Lambda}_{pi} (\mu_i - \bar{\mu}_i) + \Lambda_{o_i} (\mu_i - o_i) = \sum_k J_{k \to i} \tag{56}$$

The sensory term $\Lambda_{o_i}(\mu_i - o_i)$ acts as an additional "anchoring force" that grounds the agent in observations, distinct from the social momentum currents.

1.12 Summary

The Complete Theory with Sensory Evidence

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

Free Energy:

$$\mathcal{F} = \sum_i D_{\mathrm{KL}}(q_i \| p_i) + \sum_{i,k} \beta_{ik} D_{\mathrm{KL}}(q_i \| \Omega_{ik}[q_k]) - \sum_i \mathbb{E}_{q_i}[\log p(o_i \mid \theta)] \tag{57}$$

Mass Matrix:

$$\mathbf{M} = \frac{\partial^2 \mathcal{F}}{\partial \xi \partial \xi} = \text{Fisher information} = \text{Precision} \tag{58}$$

Effective Mass:

$$M_i = \bar{\Lambda}_{pi} + \sum_k \beta_{ik} \tilde{\Lambda}_{qk} + \sum_j \beta_{ji} \Lambda_{qi} + \Lambda_{o_i} \tag{59}$$

Dynamics:

$$\dot{\xi} = \mathbf{M}^{-1}\pi, \qquad \dot{\pi} = -\nabla \mathcal{F} - \frac{1}{2} \nabla_\xi \langle \pi, \mathbf{M}^{-1}\pi \rangle \tag{60}$$

Physical Meaning:

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