

# BEDS: Bayesian Emergent Dissipative Structures

A Formal Framework for Sustainable Digital Twins and Continual Learning Systems

Laurent Caraffa

January 2026, v:0x40

*What is the common point between a river, a neural network, and Gödel's incompleteness theorem? We propose they can all be formalized as Bayesian Emergent Dissipative Structures.*

---

# Foreword

## A Personal Journey

All my research has revolved around a single question: how to build the most faithful digital twin possible? I explored many paths. Apache Spark for processing massive geospatial datasets. Large Language Models for their implicit representations of the world – their ability to compress and reconstruct complex patterns. Blockchain for traceability and audit – ensuring that every transformation of data could be verified and trusted. Each approach solved part of the puzzle. But after years of work, I found myself facing what felt like an impasse. The more sophisticated the tools became, the more energy they consumed. The more accurate the models, the faster they drifted from reality. The more complete the audit trail, the heavier the infrastructure. So I inverted the question. Instead of asking “How do I build the most accurate digital twin?”, I asked: “What would the formalism of a sober digital twin look like?” A twin that learns continuously without exploding in complexity. A twin that stays synchronized with reality without constant retraining. A twin that can be audited without storing everything. A twin that is sustainable by construction. That’s when **BEDS** appeared.

This document presents a theoretical framework called **BEDS** (Bayesian Emergent Dissipative Structures). The central observation: what we call “learning” in machine learning, “dissipation” in thermodynamics, “evolution” in biology, and “proof” in mathematics exhibit similar structural patterns. We propose a unified formalism to describe these patterns.

## What This Document Proposes

The document is structured in five parts:

1. **Analogy** – Building intuition through the river metaphor
2. **Conjecture** – Connecting Gödel, Landauer, and Prigogine
3. **Formal Results** – Mathematical constants as fixed points of inference
4. **Implementation** – A sustainable P2P network architecture
5. **Discussion** – Implications and open questions

**Disclaimer.** The formalism presented here is deliberately simple – perhaps too simple, and I may have reinvented existing results. This is speculative work born from stepping back and reflecting on my research. It does not include a systematic literature review and does not reflect the rigor of my peer-reviewed academic work. Most of it was developed with Claude Opus 4.5 ([claude.ai](http://claude.ai)). However, I find this angle of attack compelling enough to share – hopefully others will find something valuable in it too.

Thanks for reading!

---

## The Challenge: Building a Sustainable Digital Twin

How do we build a digital twin that remains **synchronized** with the real world, **durable over time**, and **bounded in energy consumption**?

The current paradigm relies on **supervised learning**: humans label data, models memorize patterns, inference consumes energy. This approach has

three structural limitations:

#### SUPERVISED LEARNING: THREE STRUCTURAL LIMITATIONS

1. ENERGY COST  
Each inference = computation = heat  
Scale to billions of users → increasing energy demand
2. TEMPORAL DRIFT  
Model trained on data from time T  
World moves to T+1, T+2, T+3...  
Model stays at T → growing divergence
3. OUT-OF-DISTRIBUTION BEHAVIOR  
A frozen model extrapolates in uncalibrated regions  
The map diverges from the territory

Current AI models are **snapshots** of the world at training time. But the world is a **continuous process**. The longer a snapshot exists without updates, the less it resembles reality.

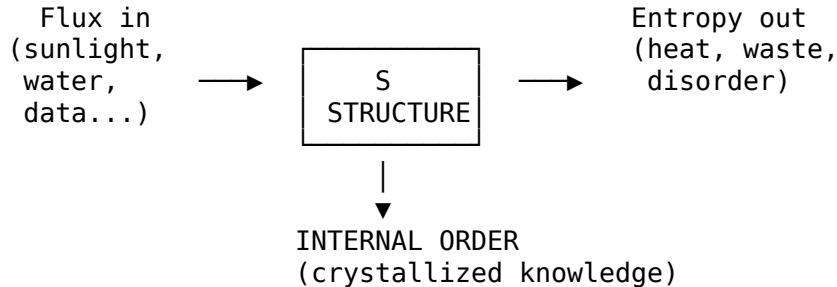
---

#### The Observation: What Is Truly Self-Supervised?

The **energetically optimal** solution is self-supervised learning: the system learns from observation alone, without external labels. But what does “self-supervised” actually mean at a physical level?

There is one class of physical systems that are **genuinely self-supervised: dissipative structures**.

DISSIPATIVE STRUCTURE (Prigogine, 1977)



Examples: Rivers, hurricanes, cells, brains, economies

Key property: They adapt to survive.

Adaptation is thermodynamically necessary.

A river doesn't have a teacher. It finds the path to the sea by dissipating energy. A cell doesn't have labeled data. It maintains itself by exporting entropy. **Adaptation emerges from the physics of maintaining order.**

---

## The Contribution: A Recursive Formalism

This document proposes **BEDS** (Bayesian Emergent Dissipative Structures): a **recursive formalism** connecting dissipative structures and Bayesian inference.

The key observation: when a dissipative structure crystallizes knowledge, that crystallized structure becomes the **prior** for the next level of emergence.

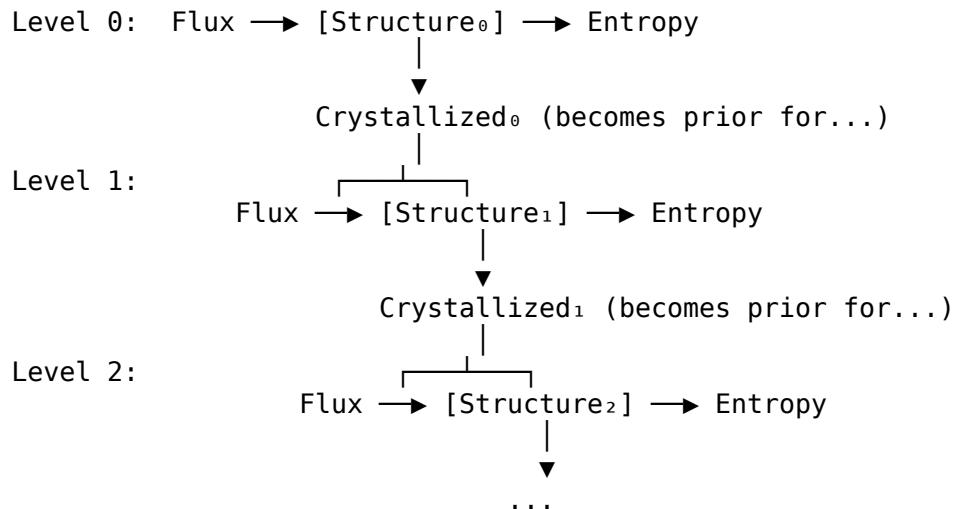
### What is a “prior”?

In everyday terms: it's what you already know before looking at new evidence.

- A child learning language has priors: grammar rules absorbed unconsciously
- A doctor diagnosing has priors: medical training, past cases seen before
- A riverbed is a prior for the mill: “water flows here” is inherited, not learned

In BEDS: each level inherits crystallized knowledge from below as its prior. It doesn't re-learn everything – it builds on what's already stable.

### THE BEDS RECURSION



Each level inherits axioms from below.

Each level only learns what it must learn.

The stack grows – but energy stays bounded.

**Key result:** A system following this design pattern exhibits bounded information content. We show it has a **maximum bound on energy consumption**. It cannot grow unboundedly – it is sustainable by construction.

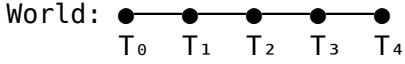
---

## The Application: A Sustainable Digital Twin

We propose a **formalism for sustainable digital twins**: a structure that learns by observing the world, continuously updating its representation, and bounded in its energy requirements.

### FROZEN MODEL vs. BEDS DIGITAL TWIN

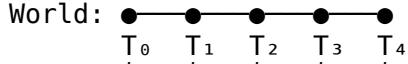
#### Frozen Model:



▲  
Frozen at training

Gap: 0      ↗      ↗      ↗      ↗  
Growing divergence  
→ Out-of-distribution

#### BEDS Digital Twin:



Continuous adaptation  
(observes, dissipates, learns)

Gap: 0    0    0    0    0  
Bounded divergence  
→ Maintained accuracy

## A Conjecture

From this formalism, we derive a conjecture:

**Systems that do not follow a dissipative structure pattern – systems that crystallize without ongoing entropy export – face increasing maintenance costs as they diverge from their environment.**

### What is an “a priori”?

An *a priori* is knowledge assumed to be true **before** observing the world – built-in assumptions, hardcoded rules, fixed architectures.

- A supervised model has a priori: the labels humans chose to impose
- A reward function is an a priori: it defines “good” before the agent experiences anything
- A frozen architecture is an a priori: it cannot adapt its own structure

In BEDS: priors **emerge** from crystallized experience. They are not imposed – they are earned.

This is a hypothesis derived from the framework:

- **Frozen models** accumulate divergence from reality
- **Divergence** requires increasing energy to maintain coherence
- **Unbounded energy** is unsustainable
- **Unsustainable** structures eventually require replacement or retraining

The question is not *whether* current AI architectures will require updates, but *how often* – and whether a more efficient pattern exists.

---

### Extension to Other Domains

If we extend this observation to **other systems** that do not follow dissipative patterns?

Societies. Institutions. Formal systems. Beliefs.

Do they face similar structural pressures – a growing divergence between their frozen representation and the evolving world?

This is a speculative question, not a claim. We explore it in Part V.

---

### What This Document Proposes

The simplest formalism for this behavior is **probability** – from the smallest fluctuation to the largest structure.

In this document, we propose a **speculative exploration** in this direction:

- **Modeling the world as a recursive Gaussian of priors.** Each level of structure is a probability distribution that crystallizes from the flux below, then serves as the foundation for representations above.
  - **Showing that, with bounded observation, this probabilistic representation crystallizes into a stable prior** – which then becomes the foundation for higher-level representations to emerge.
  - **Examining mathematical constants through this lens.** Could they be fixed points of Bayesian inference?  $e$  as the natural rate of growth when information compounds continuously.  $\pi$  as the signature of maximum entropy distributions under rotational symmetry.  $\phi$  as the ratio when a structure must reference itself to grow. Not arbitrary constants, but perhaps necessary outcomes of certain axiomatic constraints.
  - **Deriving a minimal and sustainable architecture for a P2P IoT network of Bayesian Emergent Dissipative Structures.** A concrete implementation where each node observes, dissipates, crystallizes – and shares.
- 

### The Horizon

This opens a path toward **shared knowledge systems** – founded on observation and exchange, where each participant contributes observations for the benefit of all.

A world where learning is not extraction, but **mutual crystallization**.

A world where knowledge is not hoarded, but **distributed as common priors**.

A world that learns like a river – sustainable, emergent, adaptive.

**This is what BEDS proposes to explore.**

---

# Part I – Analogy: The River That Learns

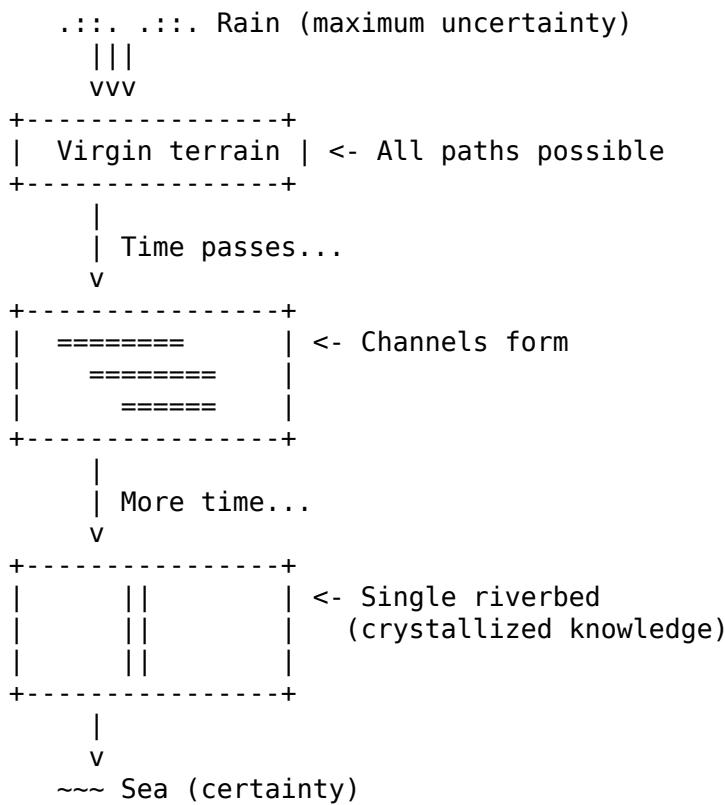
## 1.1 Water That Learns

Imagine a drop of water falling on virgin terrain. It could go anywhere – maximum uncertainty. It flows, erodes, chooses a path. Then another drop, then another. Each passage digs certain channels deeper, abandons others.

The riverbed forms.

Now the arriving water “knows” where to go – uncertainty about the path has been absorbed by the structure. What the river has learned is the relationship between its input (rain) and its output (sea), inscribed in rock.

The price paid: energy dissipated as heat and erosion.

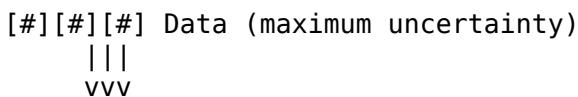


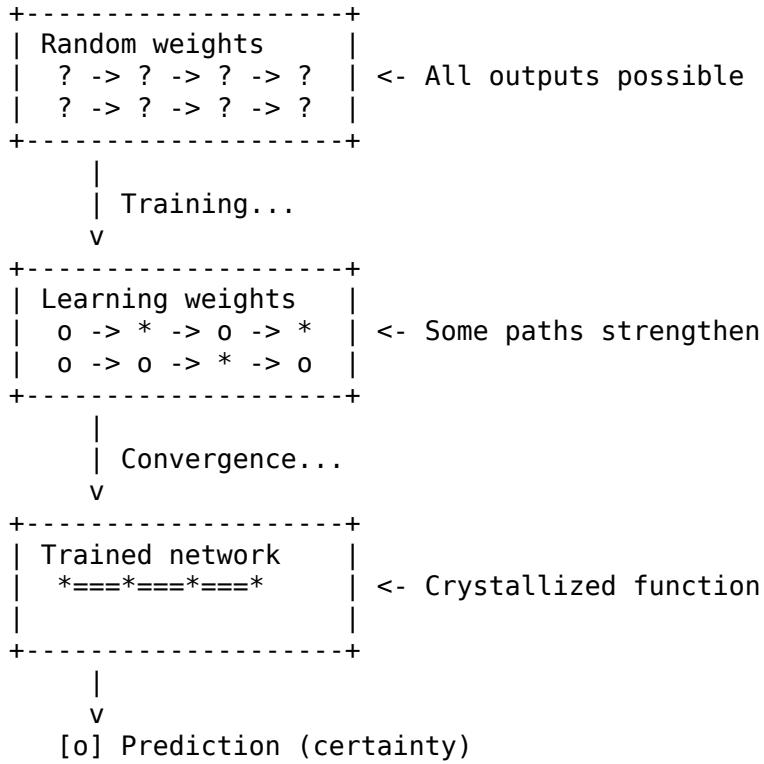
## 1.2 The Network That Learns

A neural network does exactly the same thing. Data arrives in a space of random parameters – they could produce anything. They traverse, weights adjust, certain paths strengthen. Valleys form in parameter space.

At the end, the network “knows”: this input gives that output.

The price paid: CPU cycles dissipated as heat.





### 1.3 The Same Story

	River	Neural Network
<b>Flux</b>	Water	Data
<b>Terrain to sculpt</b>	Geology	Parameter space
<b>Erosion</b>	Friction on rock	Weight updates
<b>Learned structure</b>	Riverbed	Found minimum (valley)
<b>What is represented</b>	Path of least resistance	Input -> Output function
<b>Exported entropy</b>	Heat, sediments	Heat (CPU/GPU)

In both cases, **learning is converting flux into structure by exporting entropy.**

---

### 1.4 The Shortest Path

The Bayesian framework captures the essence of this process:

- **At the beginning:** water could go anywhere – maximum uncertainty
- **During:** each drop reduces possibilities, the bed deepens
- **At the end:** a single path dominates – maximum certainty

The Bayesian is the ideal river: it only digs what the water forces it to dig, no more.

UNCERTAINTY SPACE

```

Start: ::::::::::::::::::::
::::::::::::::::::
::::::::::::::::::
All hypotheses equally possible

Data: :::::::##:::::::::::
:::::##:::::::::::
:::::##:::::::::::
Some regions more probable

More data: :::::::@@@@@::::::::::
:::::@@@::::::::::
:::::@::::::::::
Concentration on truth

End: :::::::*::::::::::
Posterior = crystallized belief

```

**Central Conjecture:** Under the Laplace approximation (locally quadratic free energy) and when the Fisher metric varies slowly, the natural gradient trajectory approximates a geodesic on the statistical manifold. In this regime, the convergence path is both optimal (least informational action) and geodesic.

---

## 1.5 The Mill That Emerges

Now that the riverbed exists, something new becomes possible.

Water flows predictably – it's free energy, available. A mill can attach itself. It couldn't have existed before: without a stable bed, where would it be built? On what flow would it rely?

The mill is itself a dissipative structure. Water traverses it, gears turn, wear, adjust. At first, it creaks, jams. Then the parts find their place, the rhythm stabilizes. The mill has learned to convert current into rotation.

The price paid: friction, wear, heat.

But once crystallized, the mill in turn releases energy – a regular rotation, exploitable. A blacksmith sets up shop. His forge learns iron: what temperatures, what gestures, what rhythms. It crystallizes. Its tools, once forged, allow digging canals, improving the bed, building other mills.

The loop closes. Or rather: it rises.

Rain (maximum uncertainty)



Riverbed ----- crystallizes ----- -> stable current (free energy)



Mill ----- crystallizes ----- -> rotation (free energy)



Forge ----- crystallizes ----- -> tools



...

Each structure can only exist because the previous one crystallized. The mill inherits an axiom: “water flows here, at this rate.” It doesn’t have to learn it – it’s acquired, inscribed in rock. Its possibility space is already restricted, and that’s precisely what allows it to exist.

This is recursion: **the crystallized posterior becomes the prior of the next structure.**

---

## 1.6 The BEDS Formalism

### 1.6.1 Dissipative Structure

A **dissipative structure** (Prigogine, 1977) is an open system that: - Absorbs an incoming flux  $\Phi_{in}$  - Maintains internal order (negentropy  $N$ ) - Exports entropy  $H$  to the environment

$$\begin{array}{c} \Phi_{in} \rightarrow [ S ] \rightarrow H_{out} \\ | \\ +-- N \text{ (maintained order)} \end{array}$$

The second law imposes:

$$\frac{dS_{system}}{dt} = \frac{dS_{internal}}{dt} + \frac{dS_{exchange}}{dt}$$

where: -  $\frac{dS_{internal}}{dt} \geq 0$  (irreversible production) -  $\frac{dS_{exchange}}{dt}$  can be negative (negentropy import)

**Existence condition:** A dissipative structure exists as long as:

$$\Phi_{in} > \Phi_{min}$$

### 1.6.2 The Fundamental Correspondence

The BEDS framework proposes a formal isomorphism between thermodynamics and Bayesian inference:

Thermodynamics	Bayesian Inference
Internal energy $E$	Negative log-likelihood $-\log p(D   \theta)$
Entropy $S$	Entropy of $q(\theta)$
Temperature $T$	Regularization parameter $\beta^{-1}$
Free energy $F$	ELBO (negative)
Thermal equilibrium	Optimal posterior
Thermal fluctuations	Epistemic uncertainty

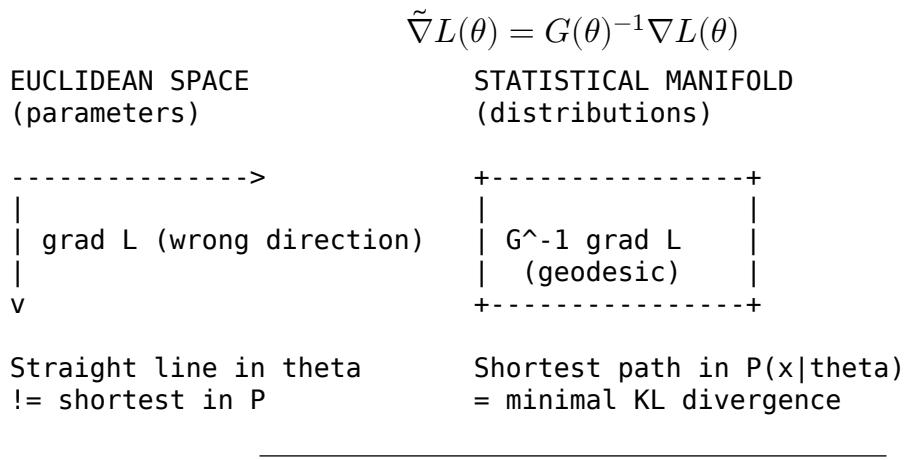
**Status:** This is a mathematical isomorphism, allowing transposition of intuitions and tools between domains. Physical identification (measuring a “crystallized bit” in joules) remains a research program.

### 1.6.3 Information Geometry

The space of probability distributions forms a **statistical manifold**  $\mathcal{M}$ . The Fisher metric defines the natural distance:

$$g_{ij}(\theta) = \mathbb{E}_{p(x|\theta)} \left[ \frac{\partial \log p(x|\theta)}{\partial \theta_i} \frac{\partial \log p(x|\theta)}{\partial \theta_j} \right]$$

The **natural gradient** corrects the ordinary gradient to point in the steepest descent direction in distribution space:



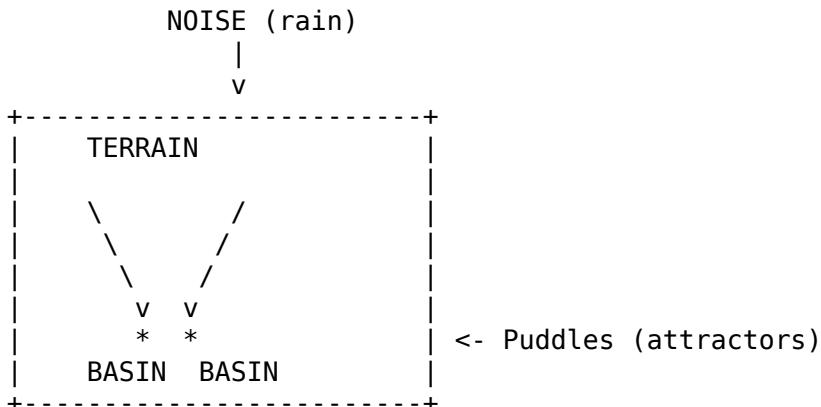
### 1.7 Basins of Attraction

Imagine rain falling on empty terrain. Water falls randomly – pure noise.

After a few minutes, what do you see?

**Puddles. Always in the same places.**

Why? Because the terrain has **hollows**. Water didn't "decide" to go there. It fell there naturally, because that's where potential energy is minimal.



Analogy	Mathematics
Random drops Terrain	Random combinations Space of relations

---

Analogy	Mathematics
Hollows	Basins of attraction
Puddles	Constants ( $e$ , $\pi$ )
Passage between	Euler (saddle point)

---

## 1.8 Recursive Emergence

When a dissipative structure reaches a stable state (local minimum of free energy), it **crystallizes**: its parameters stop fluctuating significantly.

Level 0: Raw data -> Structure S0 crystallizes

|

Level 1: S0 as prior -> Structure S1 crystallizes

|

Level 2: S1 as prior -> Structure S2 crystallizes

|

...

Formally:

$$p_{n+1}(\theta) = p_n(\theta|D_n)$$

At each level, the space of possibilities shrinks:

$$\mathcal{H}(p_{n+1}) < \mathcal{H}(p_n)$$

**Observation:** Systems that convert flux into structure by exporting entropy can maintain themselves as long as the flux persists. Systems that crystallize without entropy export accumulate internal disorder.

---

## Part II – Conjecture: Gödel, Landauer, Prigogine

### 2.1 Three Results

**Gödel (1931):** Any consistent formal system capable of expressing arithmetic contains true statements that cannot be proven within the system.

**Landauer (1961):** Erasing one bit of information requires at least  $k_B T \ln 2$  joules of energy.

**Prigogine (1977):** Open systems can maintain internal order by exporting entropy to their environment.

---

### 2.2 An Observation

These three results share a common theme: **closure has costs.**

CLOSED SYSTEMS		
GÖDEL	LANDAUER	PRIGOGINE
=====	=====	=====
Closed formal system (no external axioms)	Closed computation (no energy dissipation)	Closed physical system (no entropy export)
 v INCOMPLETENESS (true but unprovable statements)	 v REVERSIBILITY CONSTRAINTS (can't erase without cost)	 v DISORDER INCREASES (2nd law)

- A closed formal system (no external axioms) → incompleteness
- A closed computational process (no energy dissipation) → reversibility constraints
- A closed physical system (no entropy export) → disorder increases

---

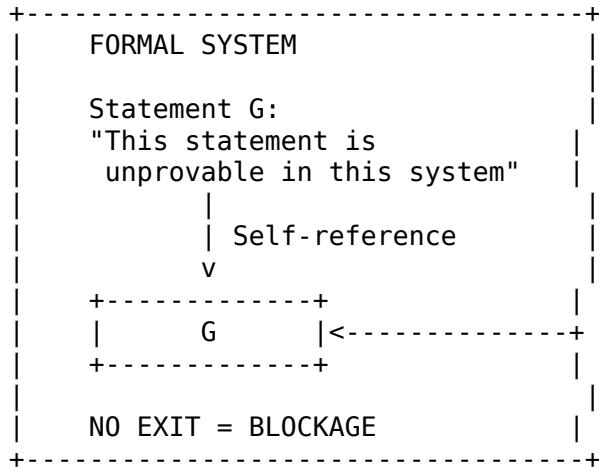
### 2.3 The Conjecture

**Conjecture:** Pathologies of formal systems (incompleteness, undecidability, paradoxes) may be structurally analogous to dissipation deficits in physical systems.

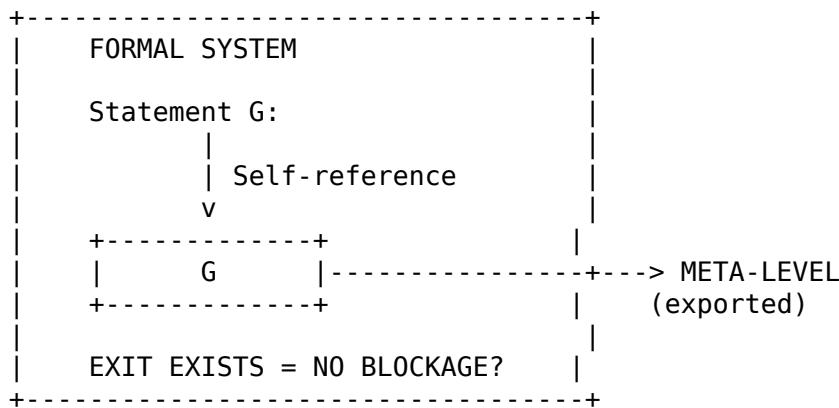
More precisely: when a formal system creates self-referential loops (as in Gödel's construction), those loops have no "exit" – no way to export the problematic information to an external level.

If a system could “dissipate” self-reference – export it to a meta-level, anchor it in physical reality – it might avoid the blockage.

#### GÖDEL'S CONSTRUCTION



#### DISSIPATIVE SYSTEM (hypothetical)



**Status:** This is an analogy, not a theorem. Whether it is deep or superficial remains an open question.

---

## 2.4 What This Would Mean (If True)

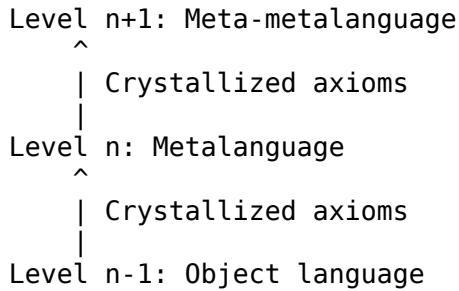
### 2.4.1 Stable Mathematics Requires Physicality

Not just in implementation, but structurally. A mathematical community that forgets, prunes, and interacts with reality can remain coherent. An isolated formal system cannot.

### 2.4.2 Incompleteness Is the Price of Openness

Tarski's hierarchy of metalanguages is infinite. But if each level “crystallizes” properly before promoting to the next, the system remains usable.

## TARSKI'S HIERARCHY as BEDS



Each level's posterior becomes the next level's prior.  
Incompleteness is "exported upward" – dissipated.

### 2.4.3 Forgetting Is Necessary for Coherence

Systems that preserve everything (perfect memory, no pruning) accumulate contradictions.

*Observation: To maintain consistency, systems may need to forget.*

---

## 2.5 Related Ideas

**Chaitin's Omega:** Algorithmic information theory connects randomness and incompleteness. Is there an energy interpretation? Omega represents the ultimate incompressible information – perhaps analogous to a “thermodynamic floor” of formal systems.

**Friston's Free Energy Principle:** Biological systems minimize variational free energy to maintain their integrity. This provides a theoretical precedent for applying Bayesian formalism to self-organized systems.

**Wheeler's “It from Bit”:** If information is physical, are logical constraints also physical? The BEDS conjecture suggests this possibility: formal pathologies may be information-theoretic manifestations of thermodynamic constraints.

---

## 2.6 Summary

**Observation:** Closed formal systems (Peano arithmetic, Turing machines, lambda-calculus) exhibit structural pathologies – incompleteness, undecidability, paradoxes. These pathologies emerge when the system must speak about itself without being able to “exit” toward a higher level.

**Contrast:** Physical systems (computers, brains) do not exhibit these pathologies in the same sense. They have other problems (crashes, errors, diseases), but not logical incompleteness. A hypothesis: their bounded resources and their capacity to dissipate (forget, restart, die) protect them from infinite loops.

**BEDS Conjecture:** If this analogy is correct, an information system that explicitly integrates openness (incoming flux), dissipation (structured forgetting), and hierarchy (crystallization by levels) could avoid certain pathologies of purely formal systems.

**Status:** This is a structural analogy, not a proven equivalence. Formal systems and physical systems are not directly comparable – one lives in abstraction, the other in matter. The interest of BEDS is to propose a bridge.

## 2.7 BEDS as Emergent Energy-Based Model

The recursive formulation of BEDS presented so far is **linear** – a single chain of crystallizing structures. But in reality, at each level, multiple agents observe, interact, and crystallize simultaneously.

This observation leads to a deeper formulation: **BEDS as an Emergent Hierarchical Energy-Based Model**.

---

### 2.7.1 From Chain to Graph

The linear view:

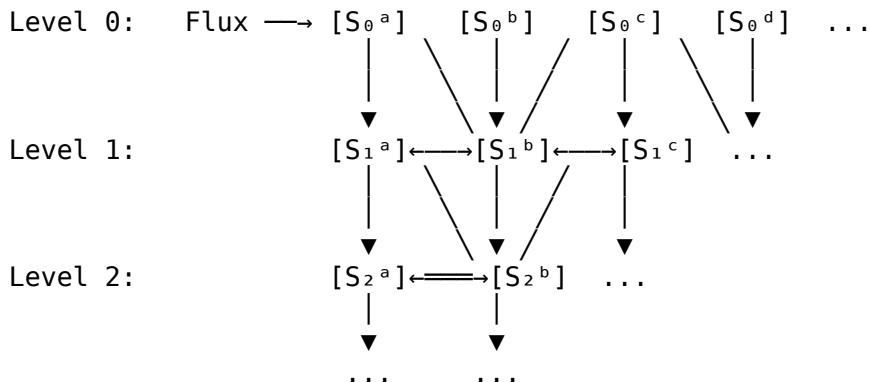
LINEAR BEDS (simplified)

Level 0: Flux  $\longrightarrow [S_0]$   $\longrightarrow$  crystallizes  
|  
Level 1:  $[S_1]$   $\longrightarrow$  crystallizes  
|  
Level 2:  $[S_2]$   $\longrightarrow \dots$

One structure per level.  
One prior passed upward.

The reality:

BEDS AS EMERGENT GRAPH



N agents per level.  
Local interactions between neighbors.  
Graph structure emerges from learning.

At each level, we don't have one structure – we have **N agents**, each with their own belief state, interacting locally.

---

### 2.7.2 Three Levels of Learning

In a classical Markov Random Field (MRF), the graph structure and interaction potentials are **fixed**. Inference finds the beliefs; learning adjusts parameters.

In BEDS, **everything is learned**:

#### LEARNING LEVEL 1: BELIEFS

Each agent learns its own belief state ( $\mu_i, \sigma_i$ )  
"What do I believe about the world?"

[Agent i]  
 $\mu_i = 0.7 \leftarrow$  updated via Bayesian fusion  
 $\sigma_i = 0.2 \leftarrow$  reduced as evidence accumulates

#### LEARNING LEVEL 2: POTENTIALS

Agents learn interaction strengths  $\psi_{ij}$  with neighbors  
"How much should I trust agent j's belief?"

[Agent i]— $\psi_{ij}$ —[Agent j]  
└  $\psi_{ij}$  learned from agreement history  
High agreement → strong coupling  
Disagreement → weak coupling

#### LEARNING LEVEL 3: TOPOLOGY

The graph structure itself emerges  
"Who should I talk to?"

Before:  
[o]—[o]—[o]—[o]  
| \ | / | \ |  
[o]—[o]—[o]—[o]  
| / | \ | / |  
[o]—[o]—[o]—[o]

Dense, uniform

After:  
[o]—[o]—[o]  
| | |  
[o]—[●]—[o]  
|| ||  
[o]

Sparse, structured  
(hubs, clusters emerge)

### 2.7.3 The Energy-Based Formulation

A Markov Random Field defines a probability distribution over configurations:

$$P(\{S\}) = \frac{1}{Z} \prod_{\text{cliques } C} \psi_C(S_C)$$

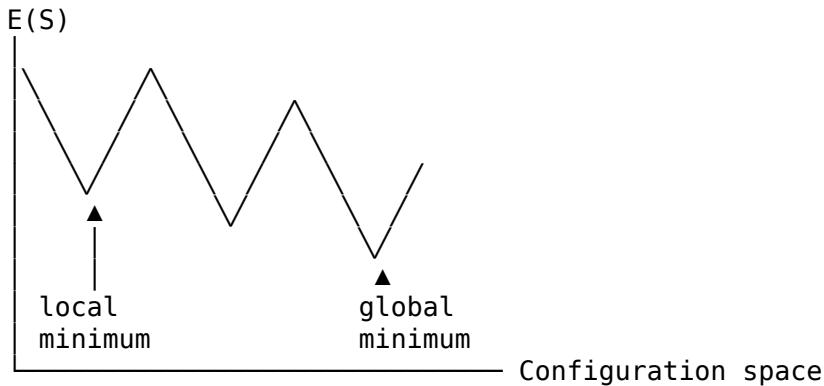
Equivalently, using an energy function:

$$P(\{S\}) = \frac{1}{Z} \exp\left(-\frac{E(\{S\})}{T}\right)$$

where: -  $\{S\} = \{S_1, S_2, \dots, S_N\}$  is the configuration of all agents -  $E(\{S\})$  is the total energy of the configuration -  $T$  is the temperature -  $Z = \sum_{\{S\}} \exp(-E(\{S\})/T)$  is the partition function

**This is a Boltzmann distribution** – the same form that governs statistical mechanics.

#### ENERGY LANDSCAPE



High T: System explores freely (liquid)  
Low T: System settles in minima (crystal)

#### 2.7.4 BEDS Energy Decomposition

In BEDS, the energy decomposes into three terms:

$$E(\{S\}) = \underbrace{E_{\text{data}}}_{\text{fit to observations}} + \underbrace{E_{\text{interact}}}_{\text{neighbor consistency}} + \underbrace{E_{\text{prior}}}_{\text{inherited constraints}}$$

#### ENERGY COMPONENTS

$$E_{\text{data}} = \sum_i D_{\text{KL}}(q_i || p_i(\text{data}))$$

How well does agent i's belief match its observations?

Observations: ● ● ●  
Belief  $q_i$ : [=====>]

<p>Mismatch:      <math>\leftarrow E_{\text{data}} \rightarrow</math></p>
<p><math>E_{\text{interact}} = \sum_{i,j} \psi_{ij} \cdot d(q_i, q_j)</math></p> <p>How consistent are neighboring beliefs?</p> <p>Agent i: [====&gt;] <math>\uparrow d(q_i, q_j)</math>  Agent j: [=====&gt;]</p> <p>Strong coupling (high <math>\psi_{ij}</math>) penalizes disagreement</p>
<p><math>E_{\text{prior}} = \sum_i D_{KL}(q_i \parallel \pi_i)</math></p> <p>How far has agent i drifted from inherited prior?</p> <p>Prior <math>\pi_i</math>: [====&gt;] (from level below)  Belief <math>q_i</math>: [=====&gt;] (current)  Drift: <math>\leftarrow E_{\text{prior}} \rightarrow</math></p>

**Total energy:**

$$E = \sum_i D_{KL}(q_i \parallel p_i) + \sum_{i,j} \psi_{ij} \cdot d(q_i, q_j) + \sum_i D_{KL}(q_i \parallel \pi_i)$$


---

### 2.7.5 Belief Propagation = BEDS Protocol

In graphical models, inference is performed via **message passing**. The canonical algorithm is Belief Propagation (BP):

$$m_{i \rightarrow j}(x_j) = \sum_{x_i} \psi_{ij}(x_i, x_j) \cdot b_i(x_i) \cdot \prod_{k \in N(i) \setminus j} m_{k \rightarrow i}(x_i)$$

For Gaussian beliefs, this simplifies to **precision-weighted averaging** – exactly the BEDS fusion rule:

BELIEF PROPAGATION

=

BEDS PROTOCOL

Agent i sends to Agent j:

Node i sends to Node j:

$m_{i \rightarrow j} = \psi_{ij} \cdot b_i$	=	$\text{msg} = (\mu_i, \sigma_i)$
---	---	----------------------------------

Agent j updates:

$$b_j \propto \prod m_{k \rightarrow j}$$

Node j fuses:

$$\begin{aligned}\tau_j &= \tau_j + \tau_i \\ \mu_j &= \text{weighted avg}\end{aligned}$$

THE BEDS P2P PROTOCOL IS BELIEF PROPAGATION  
ON A LEARNED GRAPH WITH GAUSSIAN BELIEFS

This is not a coincidence. **BEDS nodes are performing distributed variational inference.**

---

### 2.7.6 Learning the Potentials

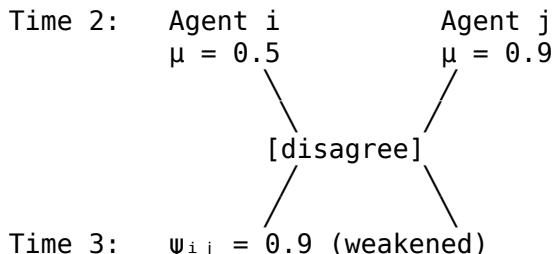
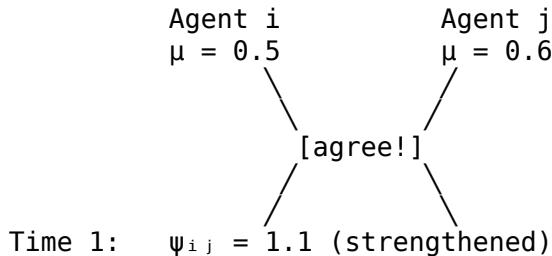
Classical MRFs have fixed potentials. In BEDS, potentials are **learned from interaction history**.

**Learning rule** (simplified):

$$\Delta\psi_{ij} \propto \text{agreement}(i, j) - \text{baseline}$$

POTENTIAL LEARNING

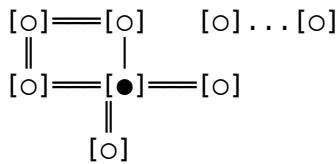
Time 0:  $\psi_{ij} = 1.0$  (default)



Time  $\infty$ : Strong connections where agents consistently agree  
Weak/pruned connections where they disagree

— strong ( $\psi > 1$ )  
— weak ( $\psi \approx 1$ )

... pruned ( $\psi \rightarrow 0$ )



STRUCTURE EMERGES FROM LEARNING

---

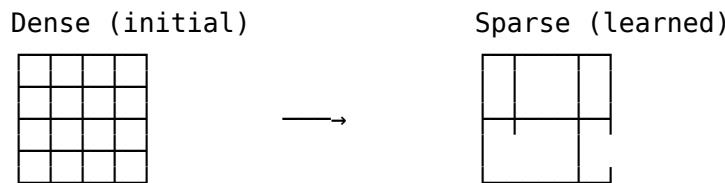
### 2.7.7 Emergent Properties

When potentials are learned, several properties **emerge spontaneously**:

#### 1. SPARSITY

---

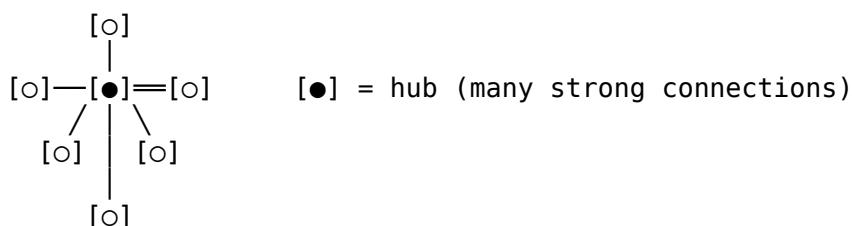
Weak connections ( $\psi_{ij} \rightarrow 0$ ) are pruned  
The graph becomes sparse over time



#### 2. HUBS (Scale-Free Structure)

---

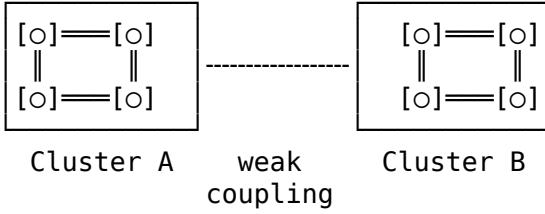
Some agents become highly connected  
They are reliable – others learn to trust them



#### 3. MODULARITY (Clustering)

---

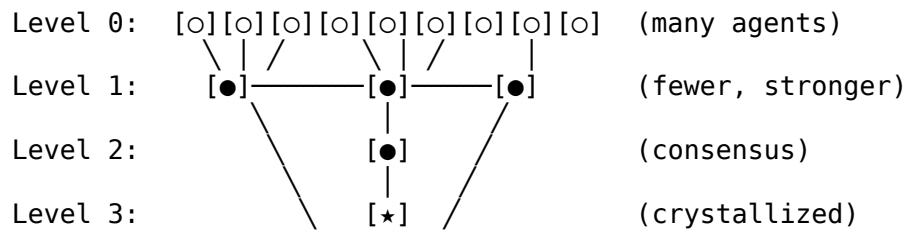
Groups of agents form tightly connected clusters  
Weak connections between clusters



#### 4. HIERARCHY

---

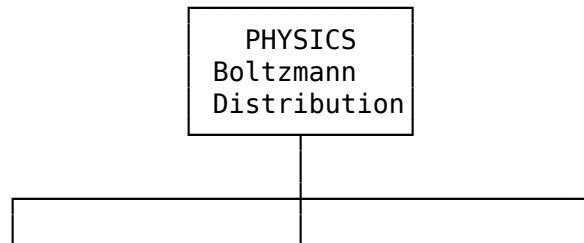
Within clusters, sub-clusters form  
The structure is fractal / self-similar

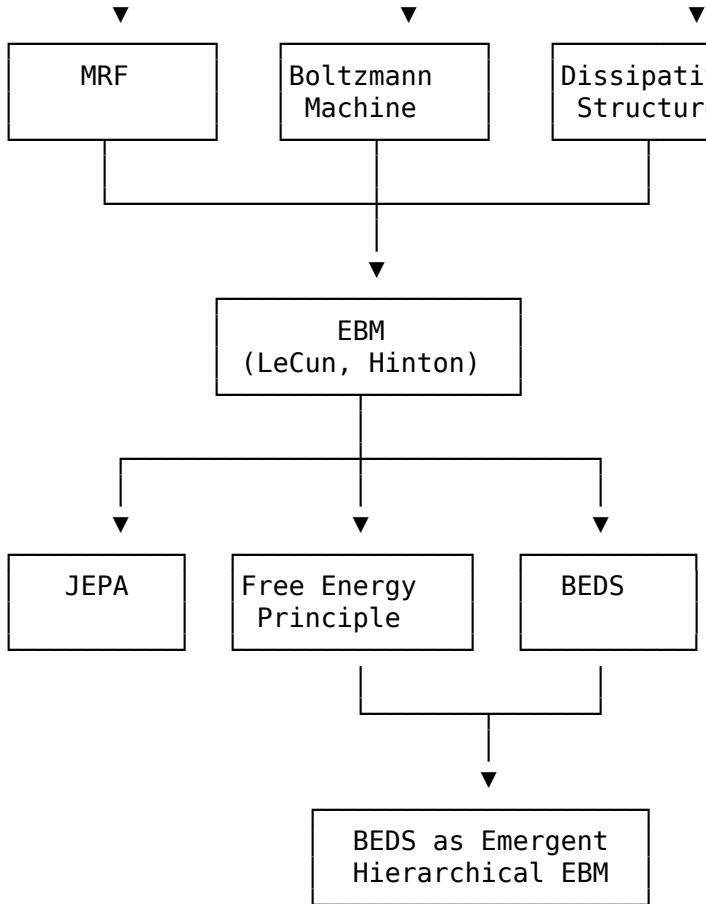


#### 2.7.8 Connection to Existing Frameworks

Framework	What it provides	BEDS adds
<b>MRF</b> (Markov Random Field)	Local interactions, Gibbs distribution	Learned potentials, emergent topology
<b>EBM</b> (Energy-Based Model)	Energy minimization, contrastive learning	Hierarchical structure, dissipation
<b>JEPA</b> (LeCun)	Joint embedding, prediction in latent space	Distributed agents, crystallization
<b>Free Energy Principle</b> (Friston)	Variational inference, active inference	Recursive priors, P2P implementation
<b>Boltzmann Machine</b>	Learned potentials, stochastic dynamics	Open system, entropy export

#### FRAMEWORK RELATIONSHIPS





### 2.7.9 Summary

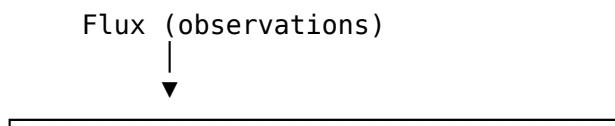
**BEDS is not just a design pattern for sustainable systems. It is an Emergent Hierarchical Energy-Based Model where:**

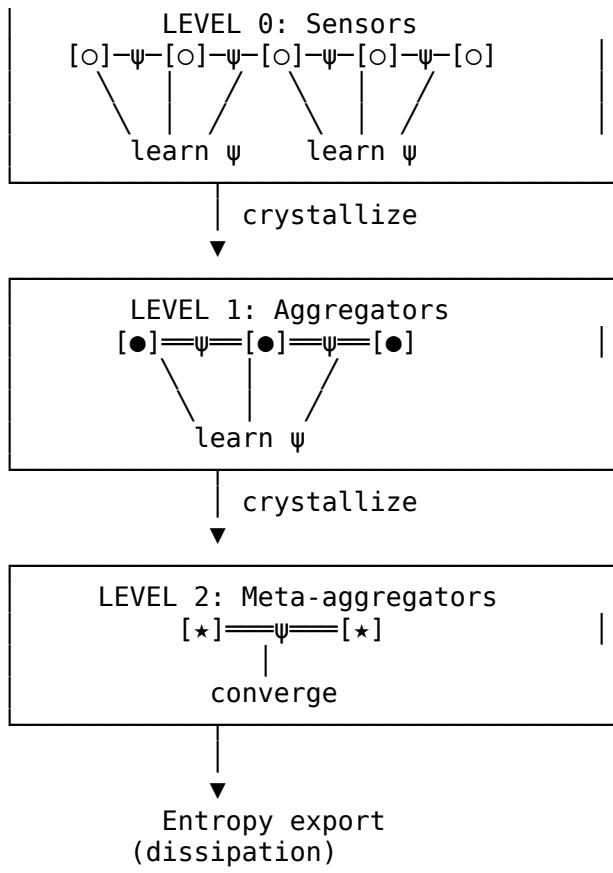
1. **Beliefs** are learned via Bayesian fusion
2. **Potentials** (interaction strengths) are learned from agreement history
3. **Topology** (who talks to whom) emerges from potential learning
4. **Hierarchy** (levels of abstraction) crystallizes from collective convergence

The P2P protocol described in Part IV is **belief propagation** on this emergent structure.

The energy bound derived in Part IV follows from the **convergence of the energy function** as potentials stabilize and the graph sparsifies.

THE COMPLETE PICTURE





At each level:

- Beliefs updated ( $\mu, \sigma$ )
- Potentials learned ( $\psi$ )
- Topology emerges (graph structure)
- Energy decreases (convergence)
- Entropy exported (dissipation)

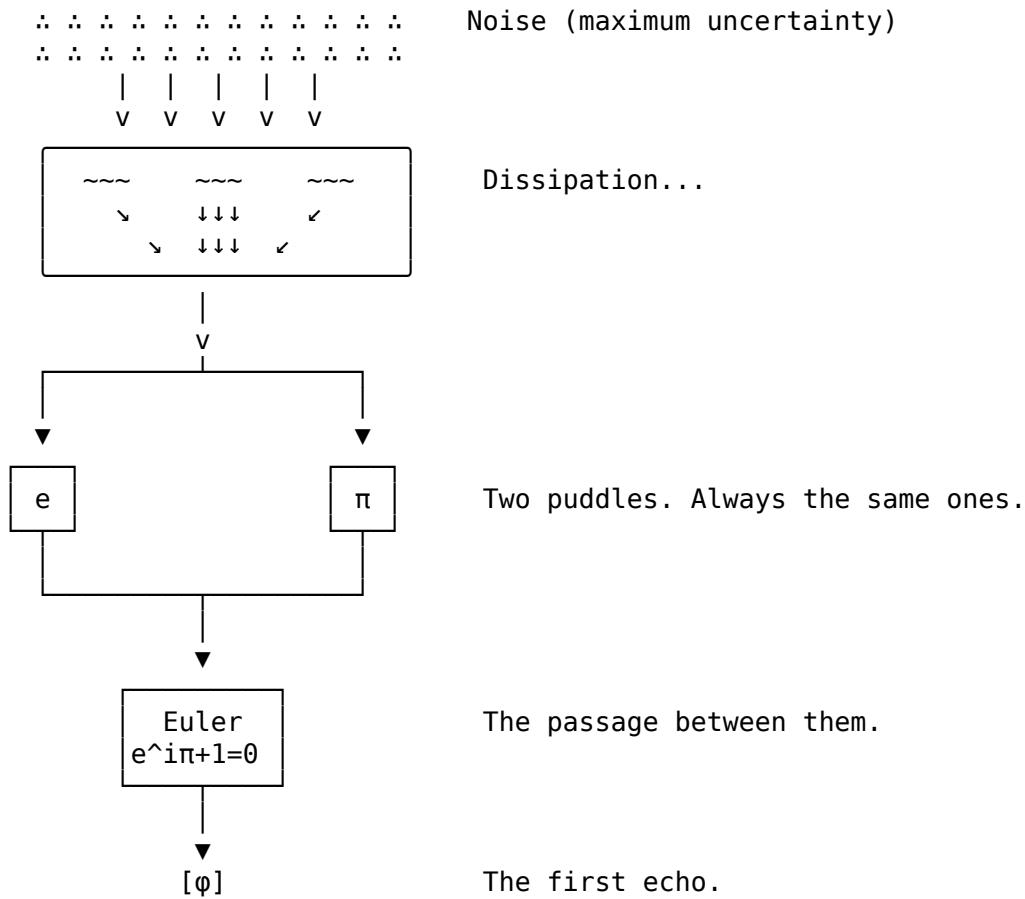
**Testable prediction:** A BEDS network will spontaneously develop sparse, modular, hierarchical structure – similar to biological neural networks and social networks.

## Part III – Formal Results: $e$ , $\pi$ , $\varphi$ , and Euler

Complete mathematical proofs are provided in Annex D. This section presents the key results and their interpretation.

### 3.0 Intuition

Imagine rain falling on an empty lot. Billions of drops land at random – pure noise.



**The water didn't "choose" these hollows. It fell into them.**

This is what we observe in mathematics. If you build any system capable of:  
- representing uncertainty - updating it in light of evidence - quantifying information

...then  $e$ ,  $\pi$ , and  $\varphi$  appear. Not because we define them arbitrarily, but because they are fixed points of certain operations.

Constant	What it encodes	Metaphor
$e$	Accumulation	The deepening hollow
$\pi$	Return	The circular hollow
Euler	Balance	The saddle point between
$\varphi$	Self-similarity	The first echo from the saddle

**Observation:** These constants appear to be fixed points of Bayesian inference under specific axioms. They emerge necessarily when systems satisfy certain constraints on continuity, symmetry, and self-reference.

---

### 3.1 The Axioms

The BEDS framework rests on four axioms only:

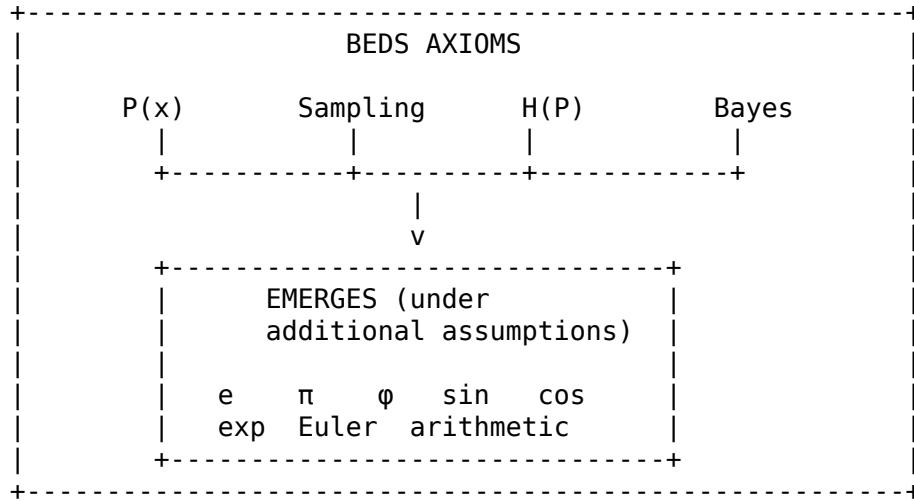
**Axiom 1 (State):** There exist probability distributions  $P(x)$  over a measurable space  $X$ .

**Axiom 2 (Sampling):** Realizations  $x \sim P$  can be drawn from distributions.

**Axiom 3 (Entropy):** Every distribution has an entropy  $H(P) = -\int P(x) \log P(x) dx$ .

**Axiom 4 (Conditioning):** Observing evidence  $y$  updates distributions via Bayes' rule.

**What emerges under natural assumptions:** arithmetic, the constants  $e$ ,  $\pi$ ,  $\varphi$ , functions  $\exp/\sin/\cos$ , Euler's equation. The derivations below require additional hypotheses (continuity, dimensionality, symmetry) that are physically natural but not strictly contained in the four axioms.



### 3.2 Emergence of e

**Theorem 1:** If Bayesian updates are continuous and infinitesimal (likelihood of the form  $1 + g(x)/n$ ), then their composition produces an exponential function with base  $e$ .

**Key idea:** When we accumulate information continuously – each piece infinitesimally small – the result is multiplication by  $e^{(\text{total information})}$ . The limit definition emerges naturally:

$$e := \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = 2.71828\dots$$

## CONTINUOUS BAYESIAN UPDATING

```
n=1:  (1 + 1/1)^1 = 2.000
n=10: (1 + 1/10)^10 = 2.594
n=100: (1 + 1/100)^100 = 2.705
n->∞: limit = e = 2.718
```

=====

e emerges as the limit of continuous  
information accumulation

**Interpretation:** The constant e is not an arbitrary definition. It is the **fixed point** of continuous Bayesian information accumulation. It is the only function that equals its own derivative at all orders.

**e encodes monotonic growth that perfectly predicts itself.**

*Full proof:* Annex D.1

---

## 3.3 Emergence of $\pi$

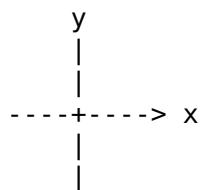
**Theorem 2:** In a space of dimension  $\geq 2$  with rotational symmetry, normalizing the maximum-entropy distribution (Gaussian) produces  $\pi$  as the period of angular integration.

**Key idea:** The Gaussian is the distribution that maximizes entropy for fixed variance. Normalizing it requires computing  $\int e^{-x^2} dx$ . This integral, when squared and converted to polar coordinates, reveals rotational symmetry – and  $\pi$  emerges as the period of that rotation.

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

### EMERGENCE OF $\pi$ FROM ROTATIONAL SYMMETRY

Cartesian:



Polar:

The integrand  $e^{-(x^2+y^2)}$   
depends only on  $r^2 = x^2+y^2$   
-> Rotational symmetry  
-> Angle integral =  $2\pi$   
->  $\pi$  emerges as the period

**Interpretation:**  $\pi$  is not primarily “the ratio of circumference to diameter.” That geometric fact is a *consequence*.  $\pi$  is the **signature of rotational symmetry** – the period that emerges when systems have no preferred direction.

**$\pi$  encodes periodicity – the return to origin after a full cycle.**

*Full proof:* Annex D.2

---

### 3.4 Euler's Equation

**Theorem 3:** The equation  $e^{i\pi} + 1 = 0$  follows necessarily from Theorems 1 and 2.

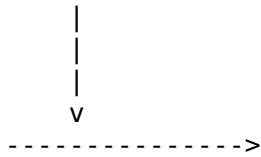
**Key idea:** Extending the exponential to complex arguments, multiplication by  $(1 + i\theta/n)$  represents an infinitesimal rotation. Composing  $n$  such rotations gives total rotation  $\theta$ . At  $\theta = \pi$ , we've rotated halfway around the circle, arriving at -1.

$$e^{i\pi} + 1 = 0$$

THE TWO MODES MEET

MODE: ACCUMULATION (e)

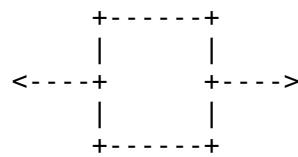
=====



"Accumulates forever"

MODE: PERIODICITY ( $\pi$ )

=====



"Returns forever"

EULER:  $e^{i\pi} + 1 = 0$

=====

ACCUMULATION applied to HALF-PERIOD = -1

Add 1 = ZERO

**Interpretation:** Euler's equation is not a mysterious coincidence. It is the **meeting point** where accumulation (e) applied to half a period ( $\pi$ ) produces exact opposition (-1), which added to unity gives zero.

**Euler's identity is where accumulation and periodicity balance exactly.**

*Full proof:* Annex D.3

### 3.5 Emergence of $\varphi$ (Golden Ratio)

**Theorem 4:** The golden ratio can be expressed as  $\varphi = 2\cos(\pi/5)$ , linking it to e and  $\pi$  via Euler's formula. The angle  $\pi/5$  corresponds to the first regular polygon (pentagon) exhibiting non-trivial self-similar structure:

$$\varphi = e^{i\pi/5} + e^{-i\pi/5} = 2 \cos\left(\frac{\pi}{5}\right) = \frac{1 + \sqrt{5}}{2}$$

**Key idea:** The angle  $\pi/5$  corresponds to the pentagon – the first polygon with non-trivial self-similarity. At this angle, the periodicity mode generates a number satisfying  $\varphi^2 = \varphi + 1$ , the equation of self-reference.

WHY  $\pi/5$  IS SPECIAL

n	$\theta = \pi/n$	$2\cos(\theta)$	Property
1	$\pi$	-2	Opposition
2	$\pi/2$	0	Orthogonality
3	$\pi/3$	1	Trivial
4	$\pi/4$	$\sqrt{2}$	Square diagonal
5	$\pi/5$	$\varphi$	SELF-SIMILARITY <-
6	$\pi/6$	$\sqrt{3}$	Hexagon

5 is the smallest  $n \geq 3$  where  $2\cos(\pi/n)$  is an algebraic number of degree 2 satisfying  $x^2 = x + 1$ . Whether this constitutes "self-reference" in a deep sense remains interpretive.

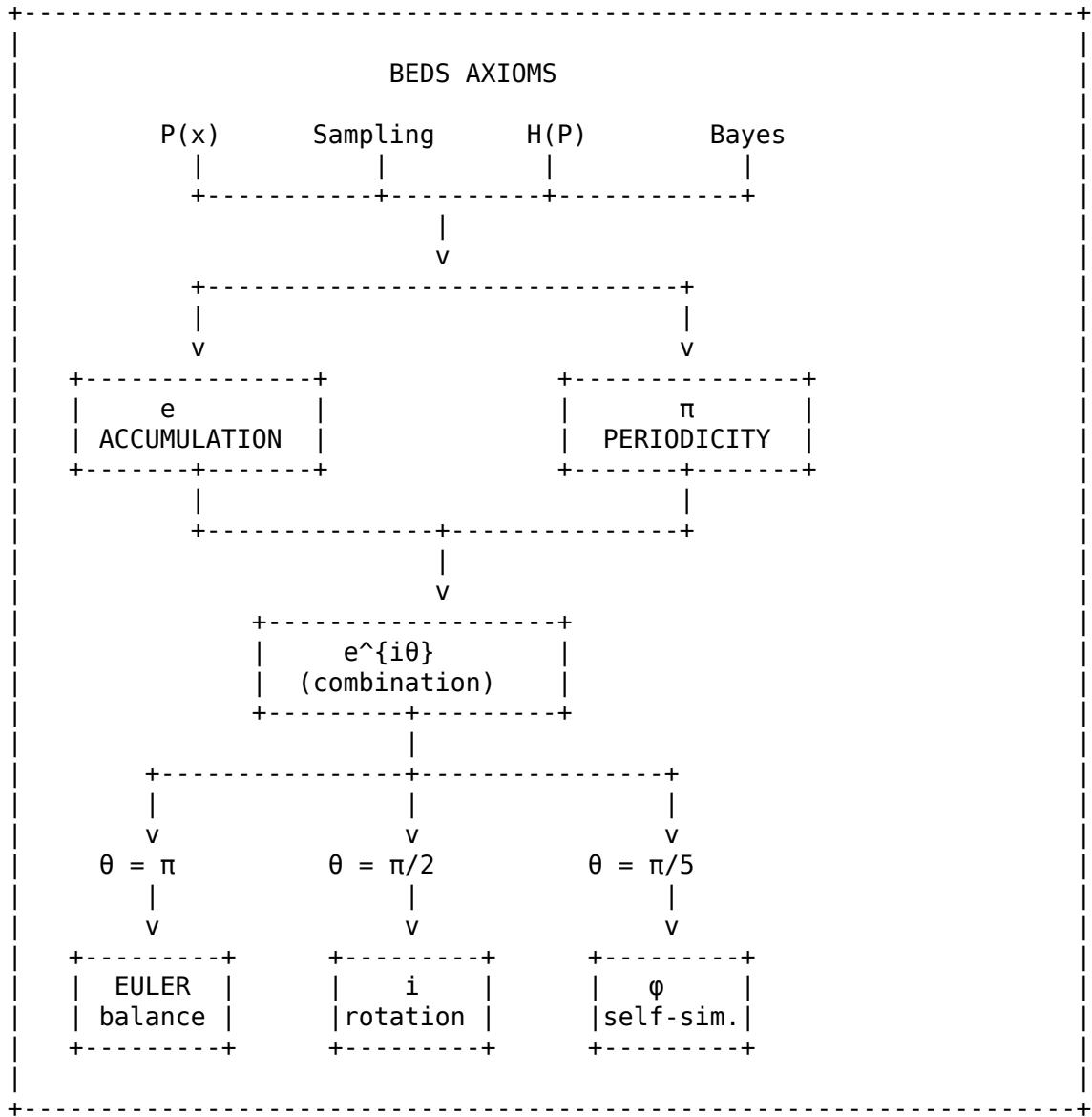
**Interpretation:**  $\varphi$  is not independent of e and  $\pi$  – it's their combination at the angle where self-reference first emerges.

**$\varphi$  encodes self-similarity – the structure that contains itself.**

*Full proof: Annex D.4*

---

### 3.6 The Complete Picture



**Note on rigor:** The derivations above are mathematically correct, but the claim that these constants “emerge necessarily” given only the BEDS axioms requires implicit assumptions about continuity, dimensionality, and symmetry. A fully rigorous treatment would need to derive these properties from more primitive axioms – an open problem.

#### Summary Table

Constant	Mechanism	Property	Mode
$e$	$\lim(1 + 1/n)^n$	Self-derivation	Pure ACCUMULATION
$\pi$	Period of rotation	Periodicity	Pure PERIODICITY
Euler	$e^{i\pi} = -1$	Balance	ACCUMULATION + full PERIODICITY
$i$	$e^{i\pi/2}$	90° rotation	ACCUMULATION + half PERIODICITY

Constant	Mechanism	Property	Mode
$\varphi$	$2\cos(\pi/5)$	Self-similarity	ACCUMULATION + 1/5 PERIODICITY

---

### 3.7 Summary

**e** is what **knows itself** – the only function that is its own derivative.

**π** is what **returns** – the only number encoding a full cycle.

**φ** is what **contains itself** – where periodicity generates self-reference.

**Euler** is **balance** – where accumulation and periodicity cancel exactly.

Under the stated axioms and assumptions, these constants emerge as fixed points. They are attractors of systems that represent and update uncertainty.

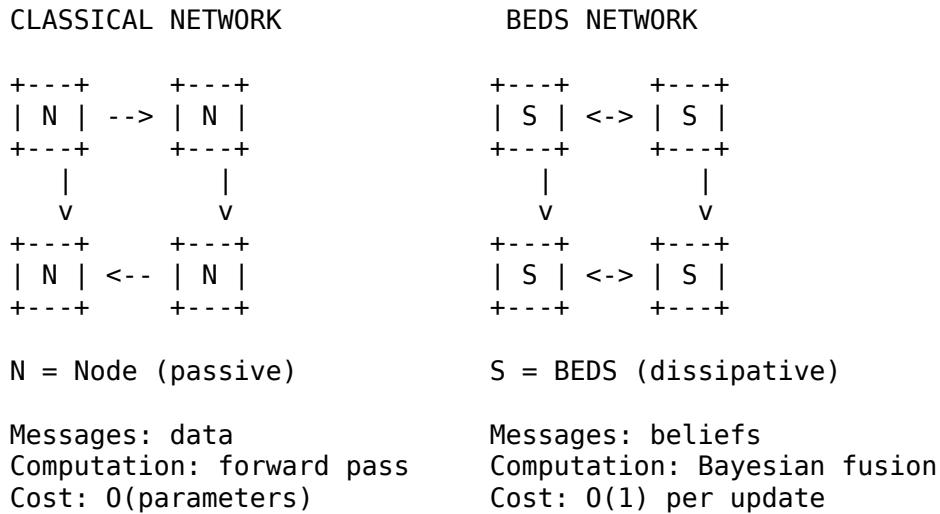
---

## Part IV – Implementation: Sustainable P2P Network

### 4.1 Motivation

BEDS structures are not just a theoretical framework – they can be instantiated directly in a peer-to-peer network.

The central idea: **each node in the network is a BEDS.**



### 4.2 Node Architecture

Each node contains:

Component	Role	Notation
Belief	Current state (posterior)	$\mu$
Uncertainty	Variance on belief	$\sigma^2$
Dissipation	Forgetting rate	$\gamma$
Identity	Public key	$pk$
Secret	Private key	$sk$

#### Isolation Principle

Once created, a node: - **Cannot** be reconfigured - **Communicates only** via signed messages - **Dies only** if flux is cut

This is the dissipative constraint: the node exists through the flux that traverses it. No flux, no node.

#### MicroQuickJS: Ultra-Minimal Runtime

For maximum resource efficiency, BEDS nodes can be implemented using **Micro-QuickJS** (`mquickjs`), Fabrice Bellard's JavaScript engine released in December 2025, specifically designed for embedded systems:

Specification	mquickjs	Standard JS (V8)
<b>Minimum RAM</b>	10 kB	~10 MB
<b>ROM footprint</b>	~100 kB	~30 MB
<b>Object overhead</b>	12 bytes (32-bit)	~64 bytes
<b>Dependencies</b>	None (self-contained)	System libraries
<b>Language subset</b>	ES5 strict	Full ES2024

## MEMORY COMPARISON

V8/Node.js:	[REDACTED]	10+ MB
QuickJS:	[REDACTED]	367 KB
MicroQuickJS:	[REDACTED]	100 KB
BEDS node state:	[.]	~200 bytes

A BEDS node fits in the cache of a \$0.50 microcontroller.

Why mquickjs is ideal for BEDS:

1. **Self-contained:** Own memory allocator, math library, and floating-point emulator
2. **No garbage collection pause:** Tracing + compacting GC with bounded latency
3. **Deterministic:** Bounded C stack usage (no recursion in parser)
4. **Portable bytecode:** 32-bit bytecode runs on any 32-bit embedded system

A complete BEDS node implementation in mquickjs requires:

```
// Core BEDS state: ~200 bytes
const node = {
  mu: 0.0,          // belief (8 bytes)
  sigma: 1.0,        // uncertainty (8 bytes)
  gamma: 0.001,      // dissipation rate (8 bytes)
  pk: new Uint8Array(32), // public key
  sk: new Uint8Array(32)  // private key
};

// Bayesian update: ~10 FLOPS
function fuse(msg) {
  const tau_self = 1 / (node.sigma * node.sigma);
  const tau_msg = 1 / (msg.sigma * msg.sigma);
  const tau_new = tau_self + tau_msg;
  node.mu = (tau_self * node.mu + tau_msg * msg.mu) / tau_new;
  node.sigma = Math.sqrt(1 / tau_new);
}
```

## 4.3 Exchange Protocol

### Minimal Message Format

Field	Description
source	Who sends (public key)
belief	The belief $\mu$
sigma	The uncertainty $\sigma$
signature	Cryptographic proof

Four fields suffice. It's canonical.

### Bayesian Update

When node B receives a belief from A, it fuses by Gaussian product:

$$\begin{aligned}\tau_{new} &= \tau_A + \tau_B \\ \mu_{new} &= \frac{\tau_A \cdot \mu_A + \tau_B \cdot \mu_B}{\tau_{new}} \\ \sigma_{new} &= \sqrt{\frac{1}{\tau_{new}}}\end{aligned}$$

where  $\tau = 1/\sigma^2$  is the precision.

#### BAYESIAN FUSION

Node A:  $\mu_A = 10$ ,  $\sigma_A = 2$       ( $\tau_A = 0.25$ )  
 Node B:  $\mu_B = 14$ ,  $\sigma_B = 1$       ( $\tau_B = 1.00$ )

Fusion:

$\tau_{new} = 0.25 + 1.00 = 1.25$   
 $\mu_{new} = (0.25*10 + 1.00*14) / 1.25 = 13.2$   
 $\sigma_{new} = \sqrt{1/1.25} = 0.89$

A	Fusion	B
+---+		+---+
10   -----> 13.2 <----- 14		±1
+---+		+---+

### Temporal Dissipation

Without new messages, uncertainty grows:

$$\sigma(t) = \sigma_0 \cdot e^{\gamma t}$$

The node "forgets." It returns toward maximum uncertainty – thermodynamic death if flux stops.

#### DISSIPATION WITHOUT FLUX

t=0: $\sigma = 1.0$	
t=1: $\sigma = 1.5$	
t=2: $\sigma = 2.3$	
t= $\infty$ : $\sigma = \infty$	DEATH (maximum uncertainty)

## 4.4 Emergent Hierarchy

```
Level 0: Sensors -> raw beliefs  
          | crystallization  
Level 1: Aggregators -> fused beliefs  
          | crystallization  
Level 2: Meta-aggregators -> consensus  
          |  
          ...
```

### Energy Bound Property

If each level crystallizes, then:

$$E_{total} = \sum_{n=0}^{\infty} E_n < \infty$$

The series converges because entropy decreases at each level. The network is **self-bounded energetically**.

## 4.5 Energy Budget

For a 5-minute cycle:

Phase	Duration	Energy
Wake + Measure	15 ms	~1 µWh
Bayesian fusion	20 ms	~1.6 µWh
LoRa transmission	100 ms	~20 µWh
Deep sleep	299.8 s	~275 µWh
<b>Total</b>	<b>300 s</b>	<b>~300 µWh</b>

Average power: ~3.6 mW

### Why Bayesian Fusion Is Negligible

The Bayesian update of two Gaussians requires only ~10 floating-point operations:

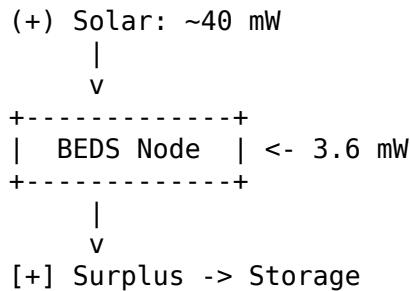
- 2 divisions, 2 multiplications, 2 additions, 1 square root

This is a direct consequence of BEDS: Bayesian fusion is the **minimal operation** on the statistical manifold.

## 4.6 Autonomy Equation

	Value
Solar input (average)	~40 mW
Node consumption	~3.6 mW
<b>Margin</b>	<b>x11</b>

## ENERGY BALANCE



Autonomy without sun: ~47 hours

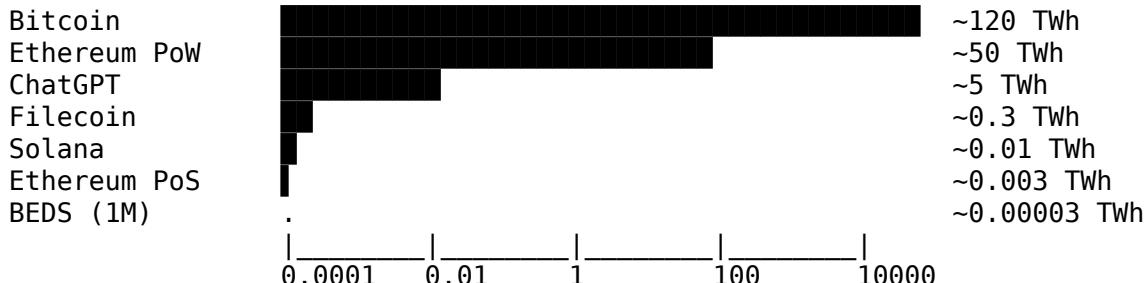
## 4.7 Comparison with Existing Systems

### Global Energy Footprint

System	Annual Consumption	Learning	Consensus
Bitcoin	~120–150 TWh	None	Proof-of-Work
Ethereum (pre-Merge, PoW)	~21–95 TWh	None	Proof-of-Work
Ethereum (post-Merge, PoS)	~0.0026 TWh	None	Proof-of-Stake
Filecoin	~0.1–0.5 TWh*	None	Proof-of-Spacetime
Solana	~0.011 TWh	None	Proof-of-History
Flow	~0.00018 TWh	None	Proof-of-Stake
ChatGPT Inference	~1–10 TWh	None (frozen)	N/A
Federated Learning (1M)	~0.01–0.1 TWh	Distributed	Centralized aggr.
<b>BEDS Network (1M nodes)</b>	<b>~0.00003 TWh</b>	<b>Continuous</b>	<b>Bayesian fusion</b>

\*Filecoin estimates vary widely depending on storage provider hardware.

### ENERGY SCALE (log TWh/year)



### Per-Transaction/Operation Comparison

System	Energy per Operation	Equivalent To
Bitcoin transaction	~1,200 kWh	2.8 house-months
Ethereum PoW transaction	~100 kWh	House for 3 days
Ethereum PoS transaction	~0.03 kWh	1 minute of microwave
Solana transaction	~0.0005 kWh	2 Google searches
BEDS belief update	~0.0000003 kWh	1 millisecond of LED light

## Functional Comparison

Criterion	Bitcoin	Ethereum	PoS	Federated Learning	BEDS Network
<b>Learning</b>	x	x		✓ (batched)	✓ (continuous)
<b>Decentralized</b>	✓	✓		Partial	✓
<b>Energy autonomous</b>	x	x		x	✓
<b>Real-time</b>	x		~12s blocks	x	✓
<b>Uncertainty quant.</b>	x	x		Partial	✓ (native)
<b>Byzantine tolerant</b>	✓	✓		Partial	✓ (weighted)
<b>Consensus type</b>	Global	Global		Centralized	Local-first

## Why BEDS Achieves Orders-of-Magnitude Efficiency

- No global consensus:** BEDS nodes only communicate with neighbors; no network-wide broadcast
- Bayesian fusion is O(1):** Independent of network size, unlike gradient aggregation
- Computation = compression:** Each update reduces entropy, not just verifies
- Natural sparsity:** Uncertainty-weighted communication means low-confidence nodes stay quiet
- Solar-sustainable:** 3.6 mW average fits within a 2cm<sup>2</sup> solar cell budget

## Memory Footprint Comparison

Runtime/System	Minimum RAM	Typical RAM	BEDS Nodes per GB
TensorFlow Lite	~1 MB	10–100 MB	10–100
PyTorch Mobile	~50 MB	100–500 MB	2–10
Node.js + V8	~10 MB	50–200 MB	5–20
QuickJS	~200 KB	1–5 MB	200–1000
MicroQuickJS	<b>10 KB</b>	50–100 KB	10,000–20,000
BEDS state only	<b>~200 B</b>	~500 B	2,000,000+

With `mquickjs`, a single Raspberry Pi Zero (512 MB) can simulate **thousands** of BEDS nodes for testing. A microcontroller with 64 KB RAM can run multiple BEDS agents.

A BEDS network doesn't prove it has *worked*. It proves it has *converged*.

## Part V – Discussion: Implications and Open Questions

### 5.1 On the Nature of Mathematics

The BEDS derivation suggests an intermediate position between Platonism and constructivism:

- Mathematics is not **invented** arbitrarily
- It is not **discovered** in a pre-existing abstract world
- It **emerges necessarily** from any system capable of representing uncertainty

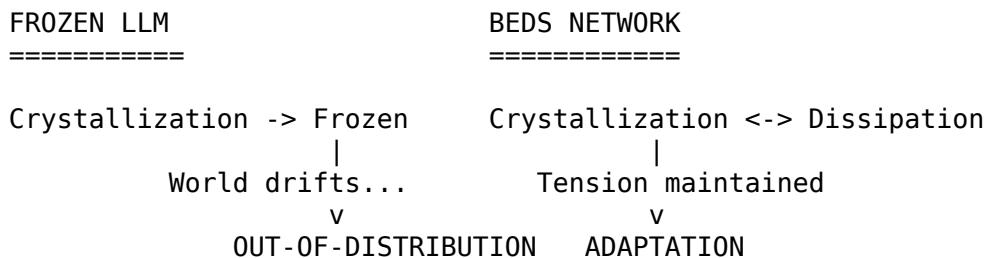
This is an interpretation, not a proof. The question of whether mathematical objects are discovered or constructed remains philosophically open.

### 5.2 Crystallization and Fragility

**Observation:** The erratic behaviors of current AI may be related to the thermodynamics of crystallization.

When a large language model is frozen after training, it crystallizes a hypothesis about the world as it was. As the environment drifts, the model extrapolates into uncalibrated regions.

Hallucinations, out-of-distribution fragility, adversarial vulnerability: these may be symptoms of **crystallization without ongoing dissipation**.



A BEDS network maintains permanent tension between crystallization and dissipation. The system remains **liquid**: ordered but adaptable.

### 5.3 Toward a Quantum Substrate

A speculative direction: **Bayesian formalism might find a natural substrate in quantum computing**.

Paradigm	State	Adaptability
Frozen LLM	Crystal	None
Classical BEDS	Liquid	Via $\gamma$
Quantum BEDS	Superposition	Native

A quantum BEDS node wouldn't *simulate* uncertainty – it would be *physical*, encoded in superposition. Structured forgetting would emerge from decoherence.

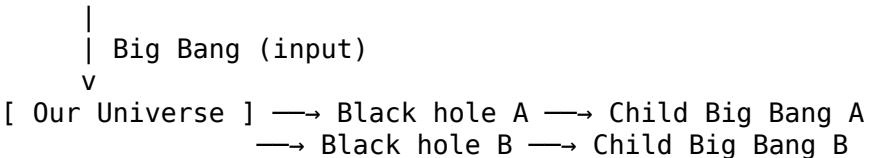
This remains speculative.

## 5.4 Cosmological Speculation

If the universe is a dissipative structure, it requires an input and outputs. The Big Bang would be the input (low-entropy prior, maximum possibilities), black holes would be the outputs (exporting entropy to... something).

Smolin (1992) proposed “cosmological natural selection” – the idea that black holes spawn child universes. BEDS would add that this is not just selection, but *learning*: each child universe inherits a more constrained prior.

Parent universe



**Status:** This is speculation, not science. It is included as an illustration of where the BEDS framework could be extended, not as a claim.

---

## 5.5 Collective Learning Systems

A Bayesian collective would be a set of agents who share their beliefs and update them together. Each agent observes the world, exchanges with others, and adjusts what they believe based on what they learn.

No central authority. No imposed doctrine. Just a continuous flow of observations, conversations, and updates.

### BAYESIAN EXCHANGE

```
A: "I observed rain"
|
+----> B: "Me too" -----> Confidence increases
|
+----> C: "I observed sun" --> Uncertainty remains
|
+----> D: "Check the clouds" --> New observation suggested
```

No one commands. Everyone updates.

When enough agents converge on the same belief – when collective uncertainty about a question becomes very low – this belief crystallizes and becomes a **shared prior**. This prior is imposed by no one: it emerges from the flow of information, like a riverbed emerges from the repeated passage of water.

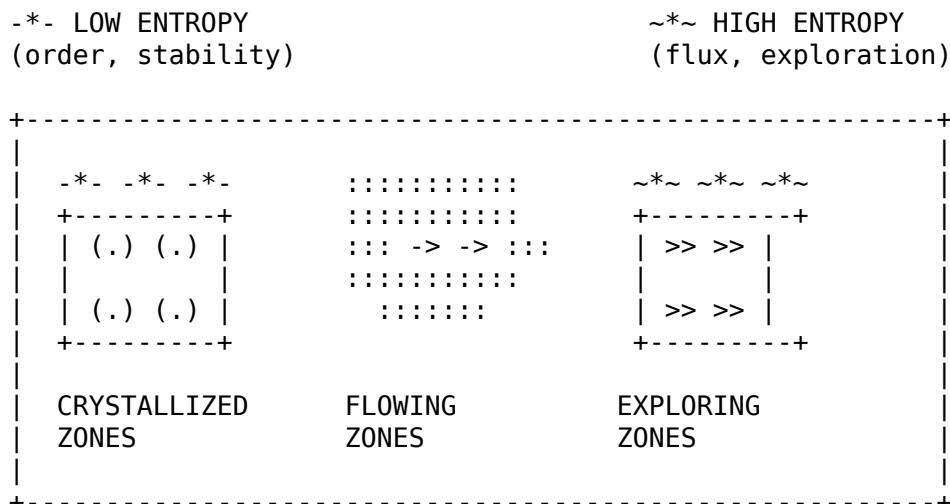
These crystallized priors then become the foundation on which new beliefs can be built. The cost of learning decreases at each level, because the space of possibilities shrinks.

---

### Different Thermodynamic Niches

Not everyone operates in the same thermodynamic regime. And that's structurally expected.

## ENTROPY LANDSCAPE



(.) = Agents seeking stability, preservation

-> = Agents exchanging, connecting

>> = Agents exploring, creating

Some agents are drawn to **negentropy** – order, predictability, stability. They gravitate toward crystallized zones: established knowledge, well-tested methods, stable institutions. They preserve what has been learned.

Others are drawn to **entropy** – flux, novelty, exploration. They gravitate toward high-energy zones: frontiers, unsolved problems, new territories. They discover what has not yet been found.

And in between, the **connectors** – carrying beliefs from one zone to another, cross-pollinating, linking stability and exploration.

## **Structural Sustainability**

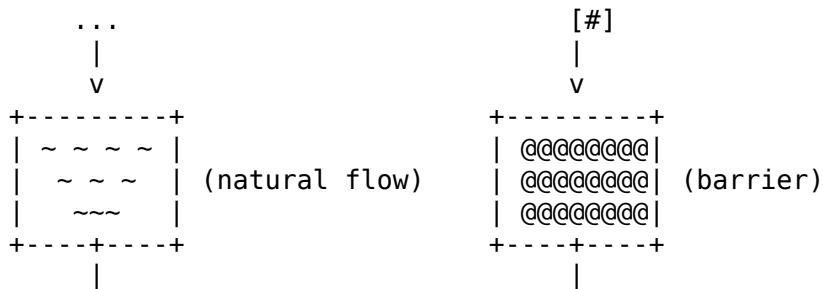
**Hypothesis:** Value systems, norms, and institutions that emerge from collective observation and exchange may be more stable than those imposed from above.

What crystallizes naturally from interactions has roots in the collective experience. What is imposed without such roots may require more energy to maintain.

## EMERGENT vs IMPOSED

**EMERGENT** (crystallizes)

### IMPOSED (forced)



v  
~~~

v  
pressure builds

Flow finds its path.

Barrier requires maintenance.

This is an observation about structural dynamics, not a normative claim.

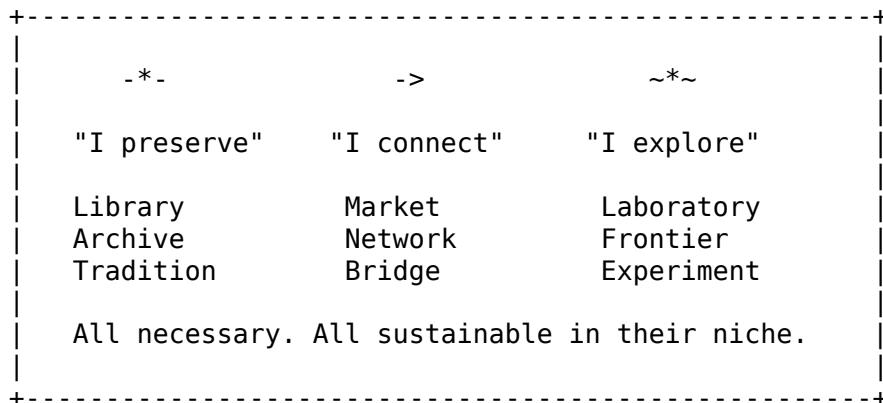
---

### Finding One's Niche

Different agents, different thermodynamic preferences:

- **Stability-seekers:** Find crystallized zones. Join established traditions. Preserve accumulated knowledge.
- **Explorers:** Find frontiers. Break new ground. Accept high uncertainty.
- **Connectors:** Walk between. Carry compressed beliefs. Reduce learning costs for others.

### NICHES IN THE LANDSCAPE



The goal is not to maximize or minimize entropy universally. The goal is to find where one's flux flows sustainably – and to let structure emerge.

*Some crystallize. Some explore. Some connect. All dissipate. All learn.*

---

## 5.6 Conclusion

**Observation: Learning, in many systems, appears to be the process of finding stable configurations.**

A formal system can maintain coherence if it is open, dissipative, and recursive. A closed system pays the price of closure: incompleteness, rigidity, or eventual replacement.

*To maintain consistency, systems may need to forget. To remain accurate, systems may need to stay open. To persist, systems may need to dissipate.*

This is a framework for thinking about sustainable learning systems. Whether it captures deep truths or merely useful analogies remains to be seen.

---

**P.S.** – A note on working with self-reference: Cantor, Gödel, and Turing all worked on closed formal systems and the paradoxes of self-reference. BEDS suggests an alternative: observe the world, let structure emerge, and export entropy to the environment. Keep the system open.

## How to Cite / Donation

If you find this work useful:

eth: 0x5664023c0de4d209a11f23978ef845ebe3e8b697

dot: 5FCpi5qu4SoYqkXyroRcmKNBDZDfkEqRVMeFS5RJrhHaNt7y

xmr: 468p23RoRLFAvbw4mGYWjrXavMtr9sUzbMFPNjChbaH1j2jV

ELmyyDyU4kcCndA3jEHNzEQo1kzyKLu7R7QwK2aEvoQibW

This is my OpenPGP public key (fingerprint: DBB3D6FDABBA66A4), which I intend to use as a stable long-term identity key.

If you find this work useful or want to reference it:

```
@report{hal-05441660,
  author      = {Caraffa, Laurent},
  title       = {{BEDS}: Bayesian Emergent Dissipative Structures},
  institution = {HAL Open Science},
  type        = {Rapport},
  year        = {2026},
  note        = {Working paper. Speculative framework, not peer-reviewed.},
  howpublished = {\url{https://github.com/lcaraffa/Bayesian_Emergent_Dissipative_Structures}},
  url         = {https://hal.science/hal-05441660},
  institution = {Univ. Gustave Eiffel, IGN-ENSG, LaSTIG}
}
```

## Acknowledgments

This work draws from five major sources of inspiration:

1. **Yann LeCun's vision** of joint embedding predictive architectures, presented in *A Path Towards Autonomous Machine Intelligence* (LeCun, 2022), complemented by the theoretical foundations of LeJEPA (Balestrieri & LeCun, 2025) – establishing that the isotropic Gaussian distribution is optimal for embeddings, a result that resonates with our conjecture on thermodynamic optimality.
2. **François Roddier's *Thermodynamique de l'évolution*** (2012), which applies Prigogine's principles to biological and social systems, and whose reading profoundly oriented this work toward a unified vision of dissipative structures at all scales.
3. **The Polkadot network** (Wood, 2016), whose interoperable parachain model offers a precedent for emergent consensus hierarchies – though our approach substitutes cryptographic proof with Bayesian convergence.
4. **Douglas Hofstadter's *Gödel, Escher, Bach: An Eternal Golden Braid***, which explores through logic, art, and music how formal systems can engender self-reference, consciousness, and meaning.
5. **Leonardo da Vinci**, whose principle – “Simplicity is the ultimate sophistication” – guided the design: four axioms, ten operations per update, one equation for consensus.

If BEDS has any merit, it is in having added nothing unnecessary.

---

## **Annex A – Technical Details: P2P Network**

### **A.4 LoRa Configuration**

| Parameter        | Value                       |
|------------------|-----------------------------|
| Frequency        | 868 MHz (EU) / 915 MHz (US) |
| Spreading Factor | SF7-SF12                    |
| Range            | 5-10 km (rural)             |

## Annex B – Connection to Existing Work

### B.1 Relation to JEPA (LeCun)

| JEPA                         | BEDS                                  |
|------------------------------|---------------------------------------|
| Encoder $x \rightarrow s_x$  | Structure extraction from flux        |
| Latent variable $z$          | Non-exportable entropy                |
| Energy $D(s_y, \tilde{s}_y)$ | Variational free energy               |
| Isotropic Gaussian (LeJEP)A  | Max entropy = dissipative equilibrium |

JEPA prescribes *what* to build. BEDS proposes *why* it might work.

### B.2 Relation to Free Energy Principle (Friston)

Both identify free energy minimization as the core organizing principle. BEDS adds recursive crystallization.

### B.3 Relation to Information Bottleneck (Tishby)

$$\min_{p(t|x)} I(X; T) - \beta I(T; Y)$$

In BEDS terms:  $I(X; T)$  = entropy exported,  $I(T; Y)$  = structure maintained,  $\beta$  = temperature.

## Annex C – Energy Comparison Data

| System                 | Annual Consumption  | Mechanism              |
|------------------------|---------------------|------------------------|
| Bitcoin                | ~120 TWh            | Proof of Work          |
| Ethereum (post-merge)  | ~0.01 TWh           | Proof of Stake         |
| GPT-4 Training         | ~50 GWh (one-shot)  | Backpropagation        |
| ChatGPT Inference      | ~1-10 TWh/year      | Forward pass           |
| Bittensor              | ~0.1 TWh            | Proof of Intelligence  |
| <b>BEDS (1M nodes)</b> | <b>~0.00003 TWh</b> | <b>Bayesian fusion</b> |

Calculation for BEDS:  $1,000,000 \text{ nodes} \times 3.6 \text{ mW} \times 8760 \text{ h} = 31.5 \text{ MWh} = 0.00003 \text{ TWh}$

---

## Annex D – Complete Mathematical Proofs

### D.1 Complete Proof: Emergence of e

#### D.1.1 Setup

Let  $P_0(x)$  be a prior distribution over parameter space  $\mathcal{X}$ . We receive a sequence of  $n$  observations  $y_1, y_2, \dots, y_n$ , each carrying infinitesimal information.

#### D.1.2 Infinitesimal Evidence Assumption

Each observation  $y_k$  has likelihood:

$$P(y_k|x) = 1 + \frac{g(x)}{n} + O\left(\frac{1}{n^2}\right)$$

where  $g(x)$  is the **information direction** – which regions of  $x$  are favored by the evidence.

**Justification:** For small perturbations, any smooth likelihood can be Taylor-expanded around 1 (the uninformative likelihood). The  $1/n$  scaling ensures total information remains finite as  $n \rightarrow \infty$ .

#### D.1.3 Sequential Bayesian Update

After observation  $y_1$ :

$$P_1(x) = \frac{P(y_1|x)P_0(x)}{P(y_1)} = \frac{(1 + g(x)/n)P_0(x)}{Z_1}$$

where  $Z_1 = \int (1 + g(x)/n)P_0(x)dx$  is the normalization constant.

After all  $n$  observations:

$$P_n(x) \propto P_0(x) \prod_{k=1}^n P(y_k|x) = P_0(x) \left(1 + \frac{g(x)}{n}\right)^n$$

#### D.1.4 The Limit

Taking  $n \rightarrow \infty$ :

$$P_\infty(x) \propto P_0(x) \cdot \lim_{n \rightarrow \infty} \left(1 + \frac{g(x)}{n}\right)^n$$

**Definition:** We define the exponential function as this limit:

$$\exp(g(x)) := \lim_{n \rightarrow \infty} \left(1 + \frac{g(x)}{n}\right)^n$$

### D.1.5 Proof That the Limit Exists and Equals e for $g(x) = 1$

Let  $a_n = (1 + 1/n)^n$ . We prove:

**(a) The sequence is increasing:**

By AM-GM inequality on  $n$  copies of  $(1 + 1/n)$  and one copy of 1:

$$\frac{n(1 + 1/n) + 1}{n + 1} \geq ((1 + 1/n)^n \cdot 1)^{1/(n+1)}$$

This gives  $(1 + 1/(n+1))^{n+1} \geq (1 + 1/n)^n$ , so  $a_{n+1} \geq a_n$ .

**(b) The sequence is bounded:**

$$\left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k} < \sum_{k=0}^{\infty} \frac{1}{k!} < 1 + 1 + \frac{1}{2} + \frac{1}{4} + \dots = 3$$

**(c) Therefore, by the monotone convergence theorem, the limit exists.**

**(d) The limit is the unique number  $e = 2.71828\dots$**

We can compute:  $e = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \dots$

### D.1.6 The Self-Derivation Property

**Theorem:**  $\frac{d}{dx} e^x = e^x$

**Proof:**

$$\frac{d}{dx} e^x = \lim_{h \rightarrow 0} \frac{e^{x+h} - e^x}{h} = e^x \lim_{h \rightarrow 0} \frac{e^h - 1}{h}$$

We need to show  $\lim_{h \rightarrow 0} \frac{e^h - 1}{h} = 1$ .

From the series expansion:

$$e^h = 1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \dots$$

Therefore:

$$\frac{e^h - 1}{h} = 1 + \frac{h}{2!} + \frac{h^2}{3!} + \dots \rightarrow 1 \text{ as } h \rightarrow 0$$

Thus  $\frac{d}{dx} e^x = e^x$ . ■

### D.1.7 Uniqueness

**Theorem:**  $e^x$  is the unique function  $f$  satisfying  $f'(x) = f(x)$  and  $f(0) = 1$ .

**Proof:** Suppose  $f'(x) = f(x)$  with  $f(0) = 1$ . Consider  $g(x) = f(x)/e^x$ .

$$g'(x) = \frac{f'(x)e^x - f(x)e^x}{e^{2x}} = \frac{f(x)e^x - f(x)e^x}{e^{2x}} = 0$$

So  $g(x)$  is constant. Since  $g(0) = f(0)/e^0 = 1/1 = 1$ , we have  $g(x) = 1$  for all  $x$ , hence  $f(x) = e^x$ . ■

---

## D.2 Complete Proof: Emergence of $\pi$

### D.2.1 Maximum Entropy Principle

**Theorem (Jaynes):** Among all distributions on  $\mathbb{R}$  with fixed mean  $\mu$  and variance  $\sigma^2$ , the one maximizing entropy is the Gaussian:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

**Proof:** We maximize the entropy functional subject to constraints:

$$\mathcal{L}[P] = - \int P(x) \ln P(x) dx - \lambda_0 \left( \int P(x) dx - 1 \right) - \lambda_1 \left( \int xP(x) dx - \mu \right) - \lambda_2 \left( \int x^2 P(x) dx - (\sigma^2 + \mu^2) \right)$$

Taking the functional derivative  $\frac{\delta \mathcal{L}}{\delta P} = 0$ :

$$-\ln P(x) - 1 - \lambda_0 - \lambda_1 x - \lambda_2 x^2 = 0$$

Therefore:

$$P(x) = \exp(-1 - \lambda_0 - \lambda_1 x - \lambda_2 x^2)$$

This is a Gaussian. The constraints determine  $\lambda_0, \lambda_1, \lambda_2$  in terms of  $\mu, \sigma^2$ . ■

### D.2.2 The Gaussian Integral

**Theorem:**  $I = \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$

**Proof (Poisson's method):**

(a) **Square the integral:**

$$I^2 = \int_{-\infty}^{+\infty} e^{-x^2} dx \cdot \int_{-\infty}^{+\infty} e^{-y^2} dy = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy$$

(b) **Observe rotational symmetry:**

The integrand  $e^{-(x^2+y^2)} = e^{-r^2}$  depends only on the radial coordinate  $r = \sqrt{x^2 + y^2}$ .

(c) **Convert to polar coordinates:**

With  $x = r \cos \theta$ ,  $y = r \sin \theta$ , the Jacobian is  $r$ :

$$I^2 = \int_0^{2\pi} d\theta \int_0^{\infty} e^{-r^2} r dr$$

**(d) Evaluate the radial integral:**

Let  $u = r^2$ , so  $du = 2r dr$ :

$$\int_0^\infty e^{-r^2} r dr = \frac{1}{2} \int_0^\infty e^{-u} du = \frac{1}{2} [-e^{-u}]_0^\infty = \frac{1}{2}$$

**(e) Evaluate the angular integral:**

The angular integral introduces  $\pi$ . To proceed rigorously, we need to establish what “angle” means.

**Definition:** The functions  $\cos \theta$  and  $\sin \theta$  are defined as the unique solutions to:

$$\frac{d^2 f}{d\theta^2} = -f$$

with initial conditions  $\cos(0) = 1, \cos'(0) = 0$  and  $\sin(0) = 0, \sin'(0) = 1$ .

**Definition:**  $\pi$  is defined as:

$$\pi := \inf\{t > 0 : \cos(t) = -1\}$$

Equivalently,  $2\pi$  is the period of  $\cos$  and  $\sin$ .

**(f) The angular integral equals  $2\pi$ :**

The full rotation around the origin traverses angle from 0 to  $2\pi$ :

$$\int_0^{2\pi} d\theta = 2\pi$$

**(g) Conclusion:**

$$I^2 = 2\pi \cdot \frac{1}{2} = \pi \\ \therefore I = \sqrt{\pi} \quad \blacksquare$$

### D.2.3 Alternative Proof via Gamma Function

The Gamma function is defined as:

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} dt$$

**Property:**  $\Gamma(1/2) = \sqrt{\pi}$

**Proof:**

$$\Gamma(1/2) = \int_0^\infty t^{-1/2} e^{-t} dt$$

Substituting  $t = x^2, dt = 2x dx$ :

$$\Gamma(1/2) = \int_0^\infty x^{-1} e^{-x^2} \cdot 2x dx = 2 \int_0^\infty e^{-x^2} dx = I$$

Since  $\Gamma(1/2) = \sqrt{\pi}$  (provable via the reflection formula  $\Gamma(s)\Gamma(1-s) = \pi/\sin(\pi s)$  at  $s = 1/2$ ), we have  $I = \sqrt{\pi}$ . ■

#### D.2.4 Why $\pi$ Appears

The key insight is that  $\pi$  emerges from **rotational symmetry**. Whenever a problem has no preferred direction in 2D, converting to polar coordinates introduces an angular integral over a full rotation, which equals  $2\pi$ .

This is not about circles having circumference  $2\pi r$ . Rather, **circles have circumference  $2\pi r$  because  $\pi$  is the period of the rotational symmetry group.**

---

### D.3 Complete Proof: Euler's Equation

#### D.3.1 Extending the Exponential to Complex Numbers

**Definition:** For  $z \in \mathbb{C}$ :

$$e^z := \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

**Convergence:** The series converges absolutely for all  $z \in \mathbb{C}$  by comparison with  $\sum |z|^k/k! = e^{|z|}$ .

#### D.3.2 Euler's Formula

**Theorem:**  $e^{i\theta} = \cos \theta + i \sin \theta$

**Proof:**

(a) **Expand  $e^{i\theta}$  as a series:**

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(i\theta)^k}{k!} = \sum_{k=0}^{\infty} \frac{i^k \theta^k}{k!}$$

(b) **Note the pattern of powers of  $i$ :** -  $i^0 = 1$  -  $i^1 = i$  -  $i^2 = -1$  -  $i^3 = -i$  -  $i^4 = 1$  (cycle repeats)

(c) **Separate real and imaginary parts:**

Real parts (even k):  $1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots = \cos \theta$

Imaginary parts (odd k):  $i \left( \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots \right) = i \sin \theta$

(d) **Therefore:**

$$e^{i\theta} = \cos \theta + i \sin \theta \blacksquare$$

#### D.3.3 Euler's Identity

**Theorem:**  $e^{i\pi} + 1 = 0$

**Proof:**

From Euler's formula with  $\theta = \pi$ :

$$e^{i\pi} = \cos \pi + i \sin \pi$$

From the definitions of  $\cos$  and  $\sin$ : -  $\cos \pi = -1$  (half-period of cosine) -  $\sin \pi = 0$  (zero-crossing of sine)

Therefore:

$$\begin{aligned} e^{i\pi} &= -1 + i \cdot 0 = -1 \\ e^{i\pi} + 1 &= 0 \quad \blacksquare \end{aligned}$$

#### D.3.4 Geometric Interpretation

The function  $e^{i\theta}$  traces the unit circle in the complex plane as  $\theta$  varies: - At  $\theta = 0$ :  $e^{i \cdot 0} = 1$  (rightmost point) - At  $\theta = \pi/2$ :  $e^{i\pi/2} = i$  (topmost point) - At  $\theta = \pi$ :  $e^{i\pi} = -1$  (leftmost point) - At  $\theta = 3\pi/2$ :  $e^{i \cdot 3\pi/2} = -i$  (bottommost point) - At  $\theta = 2\pi$ :  $e^{i \cdot 2\pi} = 1$  (back to start)

Euler's identity states: starting at 1, rotating by angle  $\pi$  (half a full rotation) brings you to  $-1$ .

---

### D.4 Complete Proof: Emergence of $\varphi$

#### D.4.1 The Golden Ratio

**Definition:** The golden ratio  $\varphi$  is the positive root of:

$$x^2 - x - 1 = 0$$

By the quadratic formula:

$$\varphi = \frac{1 + \sqrt{5}}{2} \approx 1.6180339887\dots$$

**Property:**  $\varphi$  satisfies the self-referential equation:

$$\varphi = 1 + \frac{1}{\varphi}$$

#### D.4.2 Connection to $\cos(\pi/5)$

**Theorem:**  $\varphi = 2 \cos(\pi/5)$

**Proof:**

**(a) Express  $\cos(\pi/5)$  using the identity for  $\cos(5\theta)$ :**

Let  $c = \cos(\pi/5)$ . Since  $5 \cdot (\pi/5) = \pi$ , we have  $\cos(5 \cdot \pi/5) = \cos \pi = -1$ .

Using the Chebyshev polynomial:

$$\cos(5\theta) = 16 \cos^5 \theta - 20 \cos^3 \theta + 5 \cos \theta$$

Setting  $\theta = \pi/5$  and  $c = \cos(\pi/5)$ :

$$-1 = 16c^5 - 20c^3 + 5c$$

$$16c^5 - 20c^3 + 5c + 1 = 0$$

**(b) Factor the polynomial:**

This factors as:

$$(c+1)(4c^2 - 2c - 1)^2 / (\text{some factor}) = 0$$

More directly, the minimal polynomial for  $\cos(\pi/5)$  is:

$$4c^2 - 2c - 1 = 0$$

**(c) Solve:**

$$c = \frac{2 \pm \sqrt{4 + 16}}{8} = \frac{2 \pm \sqrt{20}}{8} = \frac{1 \pm \sqrt{5}}{4}$$

Since  $\cos(\pi/5) > 0$ , we take the positive root:

$$\cos(\pi/5) = \frac{1 + \sqrt{5}}{4}$$

**(d) Therefore:**

$$2\cos(\pi/5) = \frac{1 + \sqrt{5}}{2} = \varphi \quad \blacksquare$$

#### D.4.3 Expression via Euler's Formula

**Theorem:**  $\varphi = e^{i\pi/5} + e^{-i\pi/5}$

**Proof:**

By Euler's formula:

$$\begin{aligned} e^{i\pi/5} &= \cos(\pi/5) + i\sin(\pi/5) \\ e^{-i\pi/5} &= \cos(\pi/5) - i\sin(\pi/5) \end{aligned}$$

Adding:

$$e^{i\pi/5} + e^{-i\pi/5} = 2\cos(\pi/5) = \varphi \quad \blacksquare$$

#### D.4.4 Connection to the Pentagon

The regular pentagon has interior angles of  $108^\circ = 3\pi/5$  radians. The diagonal-to-side ratio equals  $\varphi$ .

**Proof:**

In a regular pentagon with side length 1, let  $d$  be the diagonal length. The diagonal and two sides form an isosceles triangle with angles  $36^\circ$ - $72^\circ$ - $72^\circ$  (i.e.,  $\pi/5$ - $2\pi/5$ - $2\pi/5$ ).

By the law of sines:

$$\frac{d}{\sin(3\pi/5)} = \frac{1}{\sin(\pi/5)}$$

Using  $\sin(3\pi/5) = \sin(2\pi/5)$  and the identity  $\sin(2\theta) = 2\sin\theta\cos\theta$ :

$$d = \frac{\sin(2\pi/5)}{\sin(\pi/5)} = \frac{2\sin(\pi/5)\cos(\pi/5)}{\sin(\pi/5)} = 2\cos(\pi/5) = \varphi \quad \blacksquare$$

#### D.4.5 Self-Similarity of the Pentagon

When you draw all diagonals of a regular pentagon, they form a smaller regular pentagon inside. The ratio of the original to the smaller pentagon is  $\varphi^2$ .

Continuing this process creates a sequence of nested pentagons, each scaled by  $\varphi$  from the previous – a **fractal structure** exhibiting self-similarity at angle  $\pi/5$ .

#### D.4.6 The Fibonacci Connection

**Theorem:**  $\lim_{n \rightarrow \infty} \frac{F_{n+1}}{F_n} = \varphi$ , where  $F_n$  is the Fibonacci sequence.

**Proof:**

The Fibonacci recurrence  $F_{n+1} = F_n + F_{n-1}$  can be written as:

$$\frac{F_{n+1}}{F_n} = 1 + \frac{F_{n-1}}{F_n} = 1 + \frac{1}{F_n/F_{n-1}}$$

If the ratio converges to  $r$ , then:

$$r = 1 + \frac{1}{r}$$

$$r^2 = r + 1$$

This is the defining equation of  $\varphi$ . Since ratios are positive,  $r = \varphi$ . ■

---

### D.5 Summary: The Hierarchy of Constants

| Constant  | Definition         | Emergence Mechanism             | Key Property              |
|-----------|--------------------|---------------------------------|---------------------------|
| $e$       | $\lim(1 + 1/n)^n$  | Continuous Bayesian update      | $\frac{d}{dx}e^x = e^x$   |
| $\pi$     | Period of rotation | Normalizing max-entropy dist.   | $e^{2\pi i} = 1$          |
| $i$       | $e^{i\pi/2}$       | Quarter rotation                | $i^2 = -1$                |
| $-1$      | $e^{i\pi}$         | Half rotation                   | Opposition                |
| $\varphi$ | $2 \cos(\pi/5)$    | Fifth-rotation + self-reference | $\varphi^2 = \varphi + 1$ |

**Observation:** These fundamental constants can be expressed as combinations of  $e$  and  $\pi$  at rational multiples of  $\pi$ .

---

## References

### Thermodynamics and Dissipative Structures

- Prigogine, I. (1977). *Time, Structure and Fluctuations*. Nobel Lecture.
- Prigogine, I. & Stengers, I. (1984). *Order out of Chaos*. Bantam Books.
- Nicolis, G. & Prigogine, I. (1977). *Self-Organization in Non-Equilibrium Systems*. Wiley.

### Information Geometry

- Amari, S. (1998). Natural gradient works efficiently in learning. *Neural Computation*, 10(2), 251-276.
- Amari, S. & Nagaoka, H. (2000). *Methods of Information Geometry*. AMS.
- Chentsov, N.N. (1982). *Statistical Decision Rules and Optimal Inference*. AMS.

### Free Energy Principle

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

### Machine Learning

- Balestrieri, R. & LeCun, Y. (2025). LeJEPAP. *arXiv:2511.08544*.
- LeCun, Y. (2022). A Path Towards Autonomous Machine Intelligence. *OpenReview*.
- Kingma, D.P. & Welling, M. (2014). Auto-Encoding Variational Bayes. *ICLR*.

### Logic and Foundations

- Gödel, K. (1931). Über formal unentscheidbare Sätze. *Monatshefte für Mathematik*, 38, 173-198.
- Chaitin, G. (1982). Gödel's Theorem and Information. *Int. J. Theor. Phys.*, 21, 941-954.
- Hofstadter, D. (1979). *Gödel, Escher, Bach*. Basic Books.

### Physics of Information

- Landauer, R. (1961). Irreversibility and Heat Generation. *IBM J. Res. Dev.*, 5(3), 183-191.
- Bennett, C.H. (1982). The Thermodynamics of Computation. *Int. J. Theor. Phys.*, 21(12), 905-940.
- Wheeler, J.A. (1990). Information, Physics, Quantum. *Complexity, Entropy, and the Physics of Information*.

### Information Theory

- Shannon, C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27, 379-423.
- Jaynes, E.T. (1957). Information Theory and Statistical Mechanics. *Physical Review*, 106(4), 620-630.

- Cox, R.T. (1946). Probability, Frequency and Reasonable Expectation. *Am. J. Physics*, 14(1), 1-13.

## Mathematics

- Euler, L. (1748). *Introductio in analysin infinitorum*.
- Livio, M. (2002). *The Golden Ratio: The Story of Phi*. Broadway Books.

## Cosmology

- Smolin, L. (1992). Did the universe evolve? *Classical and Quantum Gravity*, 9(1), 173-191.

---

*This document is speculative work, not a peer-reviewed article. Criticism and feedback are welcome.*

---

**Document version:** 0x40

**Last updated:** January 2026