Logic, Identity, Existence, and the Recognition Composition Law

Jonathan Washburn[1] and Philip Beltracchi[1*]

[1] Recognition Physics Research Institute, Austin, Texas, 78701, USA.

*Corresponding author(s). E-mail(s): philipbeltracchi@gmail.com;
Contributing authors: jon@recognitionphysics.org;

Abstract

There have been various programs attempting to construct axiomatic methods for physics. Recognition Science is such a program, in which physical structure is derived from constraints on a cost functional defined over recognition events, but with an important distinction that it focuses on formalizing results with machine checked theorems in Lean 4. Starting from the *Recognition Composition Law* (RCL) plus normalization and calibration axioms, we derive a unique cost functional $\boldsymbol{J(x)}$. We examine how the equivalence classes in the Recognition Geometry state and event spaces are related to the ground state of this functional ($\boldsymbol{J(1) = 0}$), and examine induced Boolean dynamics of some specifically defined predicates used in the Lean database. We then demonstrate the $\boldsymbol{J \to \infty}$ limit as $\boldsymbol{x \to 0}$ and provide analysis of the implications this has on the state and event spaces from Recognition Geometry. Finally, we show how the Finite Resolution of recognizers induces discreteness in observable structures and explore the connection of ledger balance to conservation laws.

Keywords: Recognition Science, Functional Equations, Lean, Axiomatic Physics

1 Introduction

The search for a fundamental theory of physics has historically been a search for the correct equations of motion governing matter and energy on a pre-existing manifold. Often, the primary criterion for success is empirical. Alternatively, various programs attempt to reconstruct known physics, exactly or approximately, from a smaller axiomatic footprint or simpler concepts.

One simple concept central to modern physics is variational/optimization methods with respect to some "cost". Historically early examples include Fermat's principle in optics, which states that light takes the path that minimizes the time[1], and trajectories in classical geometric mechanics that minimize the action functional through the Lagrangian $S = \int_{t_0}^{t_1} L dt$[2, 3]. Direct analogies to these approaches have carried through to modern times, for instance, field theories are typically phrased in terms of an action minimized with respect to a Lagrangian density $S = \int \mathcal{L} d^n x$[4, 5], and test particles in general relativity follow geodesic trajectories extremizing the spacetime interval[6, 7]. Further examples include statistical mechanics, where equilibrium configurations are those that extremize a thermodynamic potential such as the Helmholtz free energy for canonical systems or entropy for microcanonical systems[8], and ground state configurations in standard quantum mechanics that minimize the energy through the Hamiltonian $\hat{H}$[9]. The foundation of Recognition Science rests on a primitive functional equation, the **Recognition Composition Law (RCL)**. The cost uniqueness theorem (T5) shows that any continuous system satisfying this law, subject to normalization and calibration, is forced into a unique cost structure:

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

It is this J cost that Recognition Science attempts to use to frame problems. In the sections that follow, we trace the consequences of this cost. We show how the law of identity is connected to zero cost (T0), and under what circumstances the boundary divergence of J as $x \to 0$ (T1) can be interpreted as forcing dynamics away from "nonexistence". We then show how finite local resolution in recognition quotients [10] yields discrete observable structure (T2), and summarize how reciprocity together with a balance-preserving update rule yields a double-entry ledger form [11] (T3).

An additional distinguishing feature of Recognition Science is formalizing results as theorems that can be machine checked in the Lean 4 [12] programming language. The theorem numbering here is taken from the Lean database [13] so it does not match the order in which they are discussed in this manuscript[1].

1.1 Related work

Application of the cost function itself (1) and associated structures to physics is relatively new. However, there is a published paper examining fitting of galactic rotation curves under ΛCDM, MOND, and a modified Poisson equation source motivated by the J cost structure [14]. Additional preprints analyze patterns in the masses of standard model particles [15], and in properties of the chemical elements [16]. The focus on machine verified proofs of important theorems, while becoming more commonplace in mathematics research, is still uncommon in approaches to physics.

However, there are conceptual parallels to other established lines of research. For instance, ideas formulating quantum mechanics in a relational manner [17] or in terms of fundamental single bit measurements [18] have direct parallels to the concept of recognition. The idea of minimization of recognition cost, in addition to the resemblance to action principles as described above, shows a stronger parallel to the usage of information theory and entropic arguments in deriving fundamental physics [19–22]. The idea and structure of the ledger has parallels to the study of dynamics on graphs [23–25], cellular automata approaches to science [26, 27], and the concept of discretized spacetime [28, 29].

2 The Primitive: The Recognition Composition Law (RCL)

The foundation of Recognition Science is a constraint on how recognition events combine. We posit a single primitive functional equation, the **Recognition Composition Law (RCL)**, which governs the cost of interaction between states.

2.1 Defining the Primitive

Let $J(x)$ be a cost functional defined on positive real numbers $x \in \mathbb{R}_{>0}$. The primitive law is given by:

$$J(xy) + J(x/y) = 2J(x)J(y) + 2J(x) + 2J(y) \tag{2}$$

In the RS framework, this is the **RCL**. This equation describes how the cost of a composite state (xy) and a relative state (x/y) relates to the costs of the individual components. It is a calibrated multiplicative form of the d'Alembert functional equation, see Eqs. (8,9) below.

2.2 Conditions

To select a physical solution from this functional constraint, we impose two natural boundary conditions corresponding to the existence of a neutral identity and a standard scale:

1. **Normalization.** The identity element must have zero cost.

$$J(1) = 0 \tag{3}$$

[1]The uniqueness of the cost is foundational in our presentation, but is numbered as (T5) for this reason. The theorem (T4) about the discrete time evolution from a recognition is not directly relevant here and will be examined in other papers.

2. **Calibration.** We fix the scale of the cost function by defining the curvature at the minimum in logarithmic coordinates. Let $\widetilde{J}(u) := J(e^u)$ for $u \in \mathbb{R}$. We set:

$$\widetilde{J}''(0) = 1 \tag{4}$$

This condition rules out the trivial solution $J \equiv 0$ and fixes the units of the cost landscape.

From the standpoint of this paper, the **RCL**, normalization, and calibration conditions can be considered separate axioms. However, arguments have been made motivating the form of one or more of these axioms from more primitive or weaker conditions. For instance, [30] shows that if $F : (0, \infty) \to \mathbb{R}$ is continuous and nontrivial, $F(1) = 0$, and there exists a *symmetric quadratic polynomial* P such that

$$F(xy) + F(x/y) = P(F(x), F(y)) \qquad (x, y > 0), \qquad P(u, v) = P(v, u), \tag{5}$$

then the law itself is forced into the unique bilinear family

$$P(u, v) = 2u + 2v + cuv \qquad (c \in \mathbb{R}). \tag{6}$$

A natural log-curvature calibration at the identity then fixes $c = 2$, recovering (2) exactly. Similarly, [11] starts with a general equation of the form

$$F(xy) + F(x/y) = \alpha F(x)F(y) + \beta F(x) + \beta F(y) + \gamma \tag{7}$$

and uses normalization to eliminate γ.

Alternatively, it is possible to take (2) as given and derive the normalization from a global nonegativity constraint, which we show below:

Lemma 2.1 (Normalization from the **RCL** + non-negativity)**.** *Let* $F : (0, \infty) \to [0, \infty)$ *satisfy the RCL:*

$$F(xy) + F(x/y) = 2F(x)F(y) + 2F(x) + 2F(y) \qquad (x, y > 0).$$

Then $F(1) = 0$.

Proof Substitute $x = y = 1$ into the RCL. The left side is

$$F(1 \cdot 1) + F(1/1) = F(1) + F(1) = 2F(1).$$

The right side is

$$2F(1) \cdot F(1) + 2F(1) + 2F(1) = 2F(1)^2 + 4F(1).$$

Equating:

$$2F(1) = 2F(1)^2 + 4F(1).$$

Rearranging:

$$0 = 2F(1)^2 + 2F(1),$$

which factors as

$$0 = 2F(1)\big(F(1) + 1\big).$$

This gives $F(1) = 0$ or $F(1) = -1$. But F maps into $[0, \infty)$, so $F(1) \geq 0$. This eliminates $F(1) = -1$, leaving $F(1) = 0$. □

2.3 Theorem T5: Cost Uniqueness

Theorem T5 shows that the Recognition Composition Law, together with the normalization and calibration conditions, selects a unique element from the complete family of continuous solutions to the d'Alembert functional equation classified by Aczél[31]. A detailed proof and analysis of the cost function (1), and description of what happens when any one of the cost axioms is not enforced, is given in [32], and a machine checked proof is in the Lean database [13]. However, we sketch the proof here to make this manuscript self-contained.

Theorem (T5: Uniqueness of the calibrated cost). Let $J : (0, \infty) \to \mathbb{R}$ be continuous, satisfy the Recognition Composition Law (2), and obey $J(1) = 0$. Assume moreover that $\widetilde{J}(u) := J(e^u)$ is twice differentiable at $u = 0$ with $\widetilde{J}''(0) = 1$. Then the cost function is uniquely determined on $\mathbb{R}_{>0}$ as the form in (1) $J(x) = \frac{1}{2}\left(x + \frac{1}{x}\right) - 1$.

Proof Define K by $K(x) := 1 + J(x)$. Expanding the right-hand side of (2) shows that (2) is equivalent to the generalized d'Alembert equation for groups [33, 34]

$$K(xy) + K(xy^{-1}) = 2K(x)K(y) \qquad (x, y > 0), \tag{8}$$

here the group in question is the positive reals under multiplication. Now define f by $f(u) := K(e^u)$. Continuity of J implies continuity of f, and substituting $x = e^u$ and $y = e^v$ into (8) yields the standard d'Alembert equation [31]

$$f(u + v) + f(u - v) = 2f(u)f(v) \qquad (u, v \in \mathbb{R}), \tag{9}$$

with $f(0) = K(1) = 1$ from normalization and $f''(0) = \widetilde{J}''(0) = 1$ from calibration. The classical classification of continuous solutions to (9) admits $f \equiv 1$, $f \equiv 0$, $f(u) = \cos(au)$, or $f(u) = \cosh(au)$ for some $a > 0$ [35]. Notice that $f(0) = 1$ excludes $f \equiv 0$ and $f''(0) = 1$ excludes $f(u) \equiv 1$ (because then $f''(0)$ would be 0) and $f(u) = \cos(au)$ (because then $f''(0)$ would be $-a^2 \cos(0) = -a^2 < 0$ for real a). Hence $f(u) = \cosh(au)$ for some $a > 0$, so

$$K(x) = \cosh(a \log x) \quad \text{and} \quad J(x) = \cosh(a \log x) - 1 = \tfrac{1}{2}(x^a + x^{-a}) - 1. \tag{10}$$

Finally, $\widetilde{J}(u) = J(e^u) = \cosh(au) - 1$ has $\widetilde{J}''(0) = a^2$, so the calibration $\widetilde{J}''(0) = 1$ forces $a = 1$. Thus $J(x) = \cosh(\log x) - 1 = \frac{1}{2}(x + x^{-1}) - 1$, as claimed.

□

Alternatively, it is possible to construct the solution as an ODE initial value problem once we have obtained (9). Setting $u = 0$ gives $f(v) + f(-v) = 2f(0)f(v)$. Because of the normalization $f(0) = 1$, we obtain $f(v) + f(-v) = 2f(v)$, which means f is even. Even functions differentiable at 0 must have $f'(0) = 0$, and the calibration axiom implies f is at least twice differentiable at 0; in fact, a theorem from [31] gives that solutions to (9) with $f(0) = 1$ are C^∞. Taking partial derivatives of (9) with respect to v gives

$$\frac{\partial^2}{\partial v^2}(f(u + v) + f(u - v)) = \frac{\partial^2}{\partial v^2}(2f(u)f(v)) \tag{11}$$

$$f''(u + v) + f''(u - v) = 2f(u)f''(v) \tag{12}$$

Setting $v = 0$ and recalling the curvature condition $f''(0) = 1$ from A_3 gives

$$2f''(u) = 2f(u) \tag{13}$$

From which the initial conditions of $f'(0) = 0$ (evenness) and $f(0) = 1$ (consequence of normalization) gives $f(u) = \cosh(u)$, after which we can recover $K(x)$ and $J(x)$.

2.4 A connection to the Itakura-Saito and general Bregman divergences

It is interesting that $J(x)$ is related to the Itakura-Saito divergence[36], which is influential in signal processing [37]. The Itakura-Saito divergence is defined in general as

$$d(x|y) = \frac{x}{y} - \log(\frac{x}{y}) - 1 \tag{14}$$

Then

$$\frac{1}{2}(d(1|x) + d(x|1)) = \frac{1}{2}(\frac{1}{x} - \log(\frac{1}{x}) - 1 + x - \log(x) - 1) = J(x) \tag{15}$$

because the logarithmic terms cancel. This connects the cost function to information geometry (see e.g. [38]).

The Itakura-Saito divergence is a Bregman divergence

$$D_\phi(x, y) = \phi(x) - \phi(y) - (x - y) \cdot \nabla\phi(y) \tag{16}$$

for the function $\phi(x) = -\log(x)$ [39]. The cost function in logarithmic coordinates $\widetilde{J}(u) = \cosh(u) - 1$ has been shown to be the Bregman divergence of $\phi(u) = \cosh u$ evaluated at $v = 0$ in [32],

$$D_{\cosh}(u, v) = \cosh(u) - \cosh(v) - \sinh(v)(u - v) \tag{17}$$

$$D_{\cosh}(u,0) = \cosh(u) - 1. \tag{18}$$

However, J and $\widetilde{J}$ can also be shown to be an appropriately defined Bregman divergence with respect to themselves. We need the following lemma:

Lemma 2.2. *If a convex function of a single scalar variable $G(x)$ has a local minimum at y such that $G(y) = G'(y) = 0$, then $D_G(x,y) = G(x)$*

Proof By the definition of Bregman divergence (16), $D_G(x,y) = G(x) - G(y) - (x-y)G'(y)$. Substituting in the given conditions associated with the local minimum $G(y) = 0$ and $G'(y) = 0$ gives $D_G(x,y) = G(x)$. □

The associated theorems are:

Theorem 2.3 (J is its own Bregman divergence at 1)**.**

$$D_J(x,1) = J(x) \tag{19}$$

Proof From the normalization axiom, $J(1) = 0$ and since $J'(z) = \frac{1}{2}(1 - \frac{1}{z^2})$, we get $J'(1) = 0$. $J''(x) = \frac{1}{x^3}$ is positive over the domain $x \in \mathbb{R}_{>0}$ so J is convex. Straightforward application of lemma 2.2 gives $D_J(x,1) = J(x)$. □

Theorem 2.4 ($\widetilde{J}$ is its own Bregman divergence at 0)**.**

$$D_{\widetilde{J}}(u,0) = \widetilde{J}(u) \tag{20}$$

Proof Recall $\widetilde{J}(u) = \cosh(u) - 1$. By direct evaluation,$\widetilde{J}(0) = \cosh(0) - 1 = 0$ and $\widetilde{J}'(0) = \sinh(0) = 0$. $\widetilde{J}$ is convex over $u \in \mathbb{R}$ because $\widetilde{J}''(u) = \cosh(u) > 0$. We can again apply lemma 2.2 to obtain $D_{\widetilde{J}}(u,0) = \widetilde{J}(u)$. □

This information-geometric interpretation motivates treating J not merely as an algebraic solution of the RCL but as a convex divergence measuring reciprocal mismatch.

3 Zero Cost, Recognizer Equivalence Classes, and Identity Conditions (T0)

We begin with a statement and proof of **T0**.

(T0: Cost of Identity). The global minimum of the cost function (1) over its domain is $J(x) = 0$ and occurs uniquely at $x = 1$.

Proof By direct evaluation,

$$J(1) = \frac{1}{2}(1+1) - 1 = 0. \tag{21}$$

Further,

$$J'(1) = \frac{1}{2}(1 - \frac{1}{x^2})|_{x=1} = 0 \tag{22}$$

and

$$J''(1) = \frac{1}{x^3}|_{x=1} = 1 \tag{23}$$

so $x = 1$ is a local minimum. Recall that J is defined over the domain $x \in (0, \infty)$. The second derivative is positive over the domain

$$J''(x) = \frac{1}{x^3} > 0 \ \forall x \in (0, \infty) \tag{24}$$

so J is convex over its domain. It is a standard real analysis fact that a strictly convex function with a critical point has that critical point as its unique global minimum, so the local minimum $J = 0$ at $x = 1$ is a unique global minimum. □

A machine checked proof of **T0** is in the Lean database [13].

In this framework, the state $x = 1$ represents perfect ratio match and is "free": It is the ground state. This is effectively a consequence of the normalization axiom, or alternatively from Lemma 2.1. However, **T0** goes further because it shows that $J(x) = 0, x = 1$ is the unique minimum over the domain.

3.1 Cost and identity in state and event spaces

What does **T0** imply for configurations and events? In standard formulations of logic, the Law of Identity ($A = A$) is posited as an axiom. In a recognition-first setting, identity at the observable level is supplied operationally: in Recognition Geometry [10], a recognizer induces an equivalence relation of observational indistinguishability and an observable quotient, where identity is equality of equivalence classes.

For concreteness, consider configurations $a, b, c...$ in the state space $\mathcal{C}$, and a family of recognizers $R_1, R_2, ...$ mapping the state space to the event spaces $E_1, E_2, ...$ For certain types of state spaces, there may only be a finite number of independent recognizers. One physical example is electron states in atoms, which are determined by energy, squared angular momentum, z component of angular momentum, and spin. Another physical example is stationary isolated black hole spacetimes, which are classically fully determined by mass, charge, and angular momentum according to the no hair theorem. In these cases, a meaningful "all properties" composite recognizer[10] $R_{all} = R_1 \otimes R_2 \otimes ...R_{max}$ exists.

As previously mentioned, event spaces could consist of a variety of sorts of elements, such as tuples of real numbers (for instance, if the recognizer corresponds to Cartesian coordinates), integers (for instance, if the recognizer corresponds to electric charge), Boolean values (indicating presence or absence of an attribute) etc. However, the canonical cost function $J(x)$ (1) is defined for $x \in \mathbb{R}_{>0}$. As such, we need to define positive scale functions ι_n to map the results from $E_{na} = R_n(a)$ to $\mathbb{R}_{>0}$ such that the ratio

$$x_{nab} = \frac{\iota_n(R_n(a))}{\iota_n(R_n(b))} \tag{25}$$

is valid as an input to J. This is similar to the state maps ι_s, ι_o from [40], but here we are comparing two items from the same event space rather than an "object" and a "symbol".

3.1.1 From zero cost to identity

Formally, we have

$$J(x_{nab}) = 0 \Leftrightarrow x_{nab} = 1 \Leftrightarrow \iota_n(R_n(a)) = \iota_n(R_n(b)). \tag{26}$$

where the first $\Leftrightarrow$ is because J has a unique minimum at 1 and the second is from (25). Define two structural assumptions:

1. **S1: ι injectivity.** The scale function for the recognizer in this process is injective
2. **S2: All Properties recognizer** For this system, $\exists R_{all}$ that returns all independent properties of a state.

If **S1** holds, we can state

$$R_n(a) = R_n(b) \Leftrightarrow a \sim_n b \tag{27}$$

for the definition of the $\sim$ equivalence class from [10]. Note also that from our labeling of the event space, (27) gives equality of the events $E_{na} = E_{nb}$.

If additionally we have **S2** and that $a \sim_{all} b$, then it is meaningful to say

$$a = b, \tag{28}$$

in the configuration space itself because every possible property of a and b is the same (Leibniz's law).

Notice that knowing $a \sim_m b$ for all m in a proper subset of $M \subset [R_1..R_{max}]$, does not necessarily mean that $a = b$ because some $R_q \in [R_1..R_{max}]$ with $R_q \notin M$ may be able to distinguish them. Further, if ι_k is not injective for some $k \in [1, 2, ...n]$, then $x_{kab} = 1$ does not require that $a \sim_k b$.

In summary, we can label the following as the "Event space and equivalence class" reasoning pipeline

$$(J(x_{nab}) = 0 \Leftrightarrow x_{nab} = 1) \wedge S1 \implies R_n(a) = R_n(b) \Leftrightarrow E_{na} = E_{nb} \Leftrightarrow a \sim_n b \tag{29}$$

so zero cost plus injective ι_n implies an equivalence class on the states and equality in the event space. A stronger "state space equivalence" reasoning pipeline is

$$\begin{aligned} &S2 \wedge (J(x_{all,ab}) = 0 \Leftrightarrow x_{all,ab} = 1) \wedge S1 \implies \\ &R_{all}(a) = R_{all}(b) \Leftrightarrow E_{all,a} = E_{all,b} \Leftrightarrow a \sim_{all} b \Leftrightarrow a = b \end{aligned} \tag{30}$$

which requires the usage of an all properties recognizer, in addition to an injective ι_{all} and zero cost.

3.1.2 From identity to zero cost

Suppose instead we start with $a = b$. Then it is obvious that $R_n(a) = R_n(b)$ for any recognizer (because R_n is a function) and that $\iota_n(R_n(a)) = \iota_n(R_n(b))$ (because ι_n is a function) from which we can apply (26) and get that the cost of the comparison is zero.

Formally, this can be written

$$a = b \implies a \sim_n b \Leftrightarrow E_{na} = E_{nb} \Leftrightarrow R_n(a) = R_n(b) \implies \frac{\iota_n(R_n(a))}{\iota_n(R_n(b))} = x_{nab} = 1 \Leftrightarrow J(x_{nab}) = 0 \tag{31}$$

Notice that the reasoning in this direction does not require the **S1** property that ι_n is injective, and it also requires only the usage of any single recognizer rather than the composite all properties recognizer, so **S2** is not necessary for this direction either.

3.2 The Cost of Deviation

Consider a state of inconsistency, represented by any deviation $x \neq 1$. Recall that **T0** showed that $J(1) = 0$ is the unique global minimum of J over the domain $R_{>0}$. Then,

$$\forall x \neq 1, \quad J(x) > 0. \tag{32}$$

Any deviation from identity incurs a positive cost penalty. For small deviations $x = 1 + \epsilon$, the cost rises quadratically ($J(x) \approx \epsilon^2/2$). For large deviations the cost grows linearly ($x >> 1$) or as $O(x^{-1})$ for ($0 < x << 1$). In logarithmic coordinates $u = \ln(x)$, the divergence as $u \to \pm\infty$ in $\widetilde{J}(u)$ is exponential for $|u| >> 1$, and still quadratic under small deviations $u = 0 + \delta$, $\widetilde{J}(u) \approx \delta^2/2$.

Therefore, within this model, identity-consistent recognition corresponds to the unique zero-cost equilibrium of the cost landscape.

3.3 RS ontological predicates and induced Boolean Operations

The reasoning pipelines (29), (30), (31) provide bridges between the standard logic notion of identity $A = A$ and the cost condition $J = 0$. In this subsection, we examine some additional predicates involving the $J = 0$ cost condition used in the Lean database and their connection to standard logic. Properties of the predicates are examined in the examples of Section 5.

The predicates used in the database are labeled "RSExists", "RSTrue", and "RSReal". These *should not* be confused with the English words "exists", "true" and "real"; they have very specific definitions. In this paper, any time we are using the specific Recognition Science predicate, we will specify by using the "RS" prefix. Standard language uses (e.g. "true" referring to the classical Boolean notion or "exists" referring to existential quantification of some mathematical entity such as a limit) are written without the "RS" prefix.

Define

$$\text{RSExists}(x) \iff (0 < x) \wedge (J(x) = 0), \tag{33}$$

where J is defined in (1) and is proven to be unique under appropriate axioms in T5.

Note that this is more restrictive than the definition in [11], where an "Exists" condition is defined only requiring $J(x) < \infty$, and the condition that $J(x) = 0$ or $x = 1$ is referred to as

"Balance" (the definition used here is from the Lean database[13]). Based on the definition of (1), $\text{RSExists}(x) \implies x = 1$. However, the variable x is typically a ratio (for instance as in [40]), so it is helpful to extend the idea directly to configurations.

Let $\mathcal{C}$ be a configuration space, $R : \mathcal{C} \to \mathcal{E}$ an observable map, $\iota : \mathcal{E} \to \mathbb{R}_{>0}$ a scale map, and $c_{\text{ref}} \in \mathcal{C}$ a reference configuration. Define the cost bridge

$$\chi_{R,\iota,c_{\text{ref}}}(c) := \frac{\iota(R(c))}{\iota(R(c_{\text{ref}}))} \in \mathbb{R}_{>0}. \tag{34}$$

Then define configuration-level RSExistence by

$$\text{RSExists}_{\mathcal{C}}(c) \iff \text{RSExists}\big(\chi_{R,\iota,c_{\text{ref}}}(c)\big). \tag{35}$$

It is an important distinction that at the level of the scalar cost function input, only $x = 1$ RSExists. However, there may be multiple RSExistent configurations mapping to 1 through (34).

The condition "RSTrue" requires a few more steps to define. Suppose there is a map between configurations $B : \mathcal{C} \to \mathcal{C}$ and a map between configurations and classical Booleans $P : \mathcal{C} \to \{\text{true,false}\}$; P may therefore be thought of as a recognizer where the event space is the classical Booleans. We say a configuration sequence $c_i = B^i(c_0)$ **stabilizes** [2] under P if, for $n, N \in \mathbb{N}$,

$$\text{Stab}(B, P, c_0, c_*) \iff \exists N \in \mathbb{N}\ \forall n \geq N,\ P(B^n(c_0)) = P(c_*). \tag{36}$$

Further, P is **RS-decidable** at (c_*, B, c_0) if:

1. $\text{RSExists}_{\mathcal{C}}(c_*)$, and
2. $\text{Stab}(B, P, c_0, c_*)$.

With $B, \chi_{R,\iota,c_{\text{ref}}}$, and c_0 fixed, define

$$\text{RSTrue}_{B,\chi,c_0}(P; c_*) \iff \text{RSExists}_{\mathcal{C}}(c_*) \wedge P(c_*) \wedge \Big(\exists N \in \mathbb{N}\ \forall n \geq N,\ P(B^n(c_0)) = P(c_*)\Big). \tag{37}$$

Therefore, the RSTrue condition requires configuration level RSExistence (35) of c_*, Boolean truth of the P map acting on the state c_*, and stabilization of the sequence $B^n(c_0)$. Alternatively, RSTrue applies when we have RS-decidability and $P(c_*)$. Notice that we will sometimes suppress the subscripts B, χ, c_0 to reduce clutter.

Finally, "RSReal" requires a discrete skeleton $\mathcal{D}$. Formally,

$$\text{RSReal}_{\mathcal{D}}(x) \iff \text{RSExists}(x) \wedge (x \in \mathcal{D}). \tag{38}$$

One example of a discrete skeleton is the "ϕ ladder"

$$\mathcal{D}_\phi = \{\phi^n, n \in \mathbb{Z}\} \tag{39}$$

where $\phi = (1 + \sqrt{5})/2$. Note the discrete skeleton is not restrictive where x is treated as its own scalar entity and the skeleton is the ϕ ladder: $1 = \phi^0 \in \mathcal{D}_\phi$. However, a more general process could apply a discrete skeleton $\mathcal{D}_U \subseteq U$ to a component space U, then define x through a synthesis map $F : U \to \mathbb{R}_{>0}$. In this case,

$$\text{RSReal}_{F,\mathcal{D}_U}(x) \iff \Big(\exists u \in \mathcal{D}_U :\ x = F(u)\Big) \wedge \text{RSExists}(x). \tag{40}$$

One concrete example of this would be a configuration space $U = \mathbb{R}^2$, such that $u = (y, z)$, the discrete skeleton $\mathcal{D}_U : y, z \in \mathbb{Z}$, and $F(u) = |z + y| + 1$. In this example, $\text{RSExists}(x) \Leftrightarrow z = -y$ and membership $u \in \mathcal{D}_U$ are independent, so the $\text{RSReal}_{F,\mathcal{D}_U}(x)$ could fail for either reason.

The definition of RSTrue can be written as

$$\text{RSTrue}_{B,\chi,c_0}(P; c_*) \iff \underbrace{\text{RSExists}_{\mathcal{C}}(c_*)}_{(A)} \wedge \underbrace{P(c_*) = 1}_{(B)} \wedge \underbrace{\text{Stab}(B, P, c_0, c_*)}_{(C)}. \tag{41}$$

[2]The RS notion of stability is similar to the criterion from language identification in the limit[41] that a "learner" eventually settles on a correct "guess".

3.3.1 Negation

Proposition 3.1. *Given state map B, cost bridge χ, and reference c_0, for all P, c_*:*

$$\mathrm{RSTrue}(\neg P;\, c_*) \implies \neg\mathrm{RSTrue}(P;\, c_*).$$

Proof If $\mathrm{RSTrue}(\neg P;\, c_*)$, then $P(c_*) = 0$. But $\mathrm{RSTrue}(P;\, c_*)$ requires $P(c_*) = 1$, a contradiction. □

Theorem 3.2 (Conditional negation law)**.** *If P is RS-decidable at (c_*, B, c_0), then*

$$\mathrm{RSTrue}(\neg P;\, c_*) \iff \neg\mathrm{RSTrue}(P; c_*).$$

Proof ($\Rightarrow$) is Proposition 3.1.

($\Leftarrow$) Assume $\neg\mathrm{RSTrue}(P;\, c_*)$ and RS-decidability. By RS-decidability, (A) and (C) from (41) hold. Since $\mathrm{RSTrue}(P;\, c_*)$ is false but (A) and (C) hold, it must be (B) that fails: $P(c_*) \neq 1$, hence $P(c_*) = 0$ (the only other Boolean value). Then all three conjuncts of $\mathrm{RSTrue}(\neg P;\, c_*)$ hold: (A) by assumption, $P(c_*) = 0$ gives $(\neg B)$, and from condition (C) of RS-decidability, there exists N such that $P(B^n(c_0)) = P(c_*)$ for all $n \geq N$. Since we have established $P(c_*) = 0$, this gives $P(B^n(c_0)) = 0$ or equivalently, $(\neg P)(B^n(c_0)) = 1$ for all $n \geq N$, establishing $\mathit{Stab}(B, \neg P, c_0, c_*)$. □

3.3.2 Conjunction

Theorem 3.3 (RSTrue distributes over conjunction)**.**

$$\mathrm{RSTrue}(P \wedge Q;\, c_*) \iff \mathrm{RSTrue}(P;\, c_*) \;\wedge\; \mathrm{RSTrue}(Q;\, c_*).$$

Proof ($\Rightarrow$) RSExistence (A) is based on c_* so it is shared. $P(c_*) \wedge Q(c_*) = 1$ implies both $P(c_*) = 1$ and $Q(c_*) = 1$. Stabilization of $P \wedge Q$ to the value at c_* implies stabilization of each component (since $a \wedge b = a' \wedge b'$ with $a' = b' = 1$ forces $a = 1$ and $b = 1$).

($\Leftarrow$) Take $N = \max(N_P, N_Q)$. For $n \geq N$, both $P(B^n(c_0)) = P(c_*)$ and $Q(B^n(c_0)) = Q(c_*)$, hence $(P \wedge Q)(B^n(c_0)) = (P \wedge Q)(c_*)$. □

3.3.3 Disjunction

Proposition 3.4. *Given state map B, cost bridge χ, reference c_0, state c_*, and predicates P and Q*

$$\mathrm{RSTrue}(P;\, c_*) \vee \mathrm{RSTrue}(Q; c_*) \implies \mathrm{RSTrue}(P \vee Q;\, c_*).$$

Proof Suppose $\mathrm{RSTrue}(P;\, c_*)$ holds. Then $\mathrm{RSExists}_{\mathcal{C}}(c_*)$ holds, $P(c_*) = 1$, and $P(B^n(c_0)) = 1$ for $n > N$. $\mathrm{RSExists}_{\mathcal{C}}(c_*)$ is condition (A) of $\mathrm{RSTrue}(P \vee Q;\, c_*)$. $P(c_*) = 1$ requires $(P \vee Q)(c_*) = 1$, i.e. condition (B) of $\mathrm{RSTrue}(P \vee Q;\, c_*)$, and $P(B^n(c_0)) = 1$ for $n \geq N$ requires $(P \vee Q)(B^n(c_0)) = 1$, i.e. condition (C) of $\mathrm{RSTrue}(P \vee Q;\, c_*)$. Since all three conjuncts apply, then $\mathrm{RSTrue}(P \vee Q;\, c_*)$ must hold.

If $\mathrm{RSTrue}(P;\, c_*)$ does not hold, then $\mathrm{RSTrue}(Q;\, c_*)$ must hold, and the previous reasoning applies with P replaced by Q in the relevant locations. □

Theorem 3.5 (Conditional disjunction law)**.** *If P and Q are both RS-decidable at (c_*, B, c_0), then*

$$\mathrm{RSTrue}(P;\, c_*) \vee \mathrm{RSTrue}(Q; c_*) \;\Leftrightarrow\; \mathrm{RSTrue}(P \vee Q;\, c_*).$$

Proof ($\Rightarrow$) is Proposition 3.4.

($\Leftarrow$) If P and Q are RS-decidable, then conditions (A) and (C) of $\mathrm{RSTrue}(P;\, c_*)$, $\mathrm{RSTrue}(Q;\, c_*)$ are satisfied, and $\mathrm{RSTrue}(P;\, c_*) \vee \mathrm{RSTrue}(Q; c_*)$ becomes

$$(1 \wedge (P(c_*) = 1) \wedge 1) \vee (1 \wedge (Q(c_*) = 1) \wedge 1) \Leftrightarrow (P(c_*) = 1) \vee (Q(c_*) = 1) \tag{42}$$

Condition (B) of $\mathrm{RSTrue}(P \vee Q;\, c_*)$, which we are assuming to be true, is $(P \vee Q)(c_*) = 1$, from which $(P(c_*) = 1) \vee (Q(c_*) = 1)$. However, this is exactly the required condition from (42). □

4 The Boundary(T1)

The formal definition of **T1** is as follows:

(T1: Cost of the zero limit). $\lim_{x\to 0^+} J(x) = \infty$

Proof The proof is constructive: for any bound $M > 0$, choose $\varepsilon = 1/(2(M+2))$. For $0 < x < \varepsilon$, we have $x^{-1} > 2(M+2)$, so

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

This proves $J(x) \to +\infty$ as $x \to 0^+$. □

4.1 Simplistic models of cost minimizing dynamics

If starting near the $x = 0$ boundary and moving towards it has an unbounded increase in cost, then we would expect any sort of dynamics for which J cost is the objective to be repelled from the boundary. In the following, we define some examples that we label "weak" and "strong" versions of cost minimization and examine the implications of the boundary divergence. It is important to keep in mind that these are illustrative classes, meant to demonstrate general trends rather than realistically model any specific system.

A dynamical process can be minimally modeled as a discrete or continuous sequence of distinct states labeled by a monotonic parameter, for instance

$$c = c(t). \tag{43}$$

Suppose there is some equilibrium configuration with zero cost c_{eq}, with respect to some recognizer R_{dyn} and scale map ι_{dyn}, such that

$$x(t) = \frac{\iota_{dyn}(R_{dyn}(c(t)))}{\iota_{dyn}(R_{dyn}(c_{eq}))} \tag{44}$$

A weak formalization of "cost minimizing" is

$$lim_{s\to\infty} J(x(s)) = J_\infty, \qquad J_\infty \leq J(x(t)) \quad \forall t. \tag{45}$$

This allows temporary increases in cost due to, for instance, inertial or stochastic effects, so local maxima and minima in J may occur at intermediate times. However, any finite amplitude oscillations must dampen, because they would prevent a well defined distant future limit J_∞. Further, the limit J_∞ must be less than or equal to any local minimum that occur in intermediate times, because (45) holds for ALL t, including hypothetical intermediate time local minima.

A stronger version of cost minimizing dynamics would be

$$\begin{cases} J(x(t)) = 0, & J(x(t+\Delta t)) = 0 \quad \forall\ \Delta t > 0 \\ J(x(t)) > 0, & J(x(t+\Delta t)) < J(x(t)) \quad \forall\ \Delta t > 0 \end{cases}. \tag{46}$$

In this case, the cost at any future time must be less than the cost at present time, unless the cost at present time is already zero, which fixes the future cost at zero. Temporary increase in cost, or local extrema, are forbidden here. A deterministic continuous time version of (46) could be a one dimensional first order differential equation $\frac{dx}{dt} = -h(\frac{dJ}{dx})$ where h is some injective function.

Suppose at some t_0 we have $x(t_0)$. Then under (45)

$$J(x(t_0)) = \frac{1}{2}(x(t_0) + \frac{1}{x(t_0)}) - 1 \geq \frac{1}{2}(x(t_\infty) + \frac{1}{x(t_\infty)}) - 1 = J(x(t_\infty)), \tag{47}$$

whereas under (46), we can replace the infinite limit with any finite future time $t_1 > t_0$

$$\Big(J(x(t_0)) = \frac{1}{2}(x(t_0) + \frac{1}{x(t_0)}) - 1 > \frac{1}{2}(x(t_1) + \frac{1}{x(t_1)}) - 1 = J(x(t_1))\Big) \vee \Big(J(x(t_0)) = J(x(t_1)) = 0\Big). \tag{48}$$

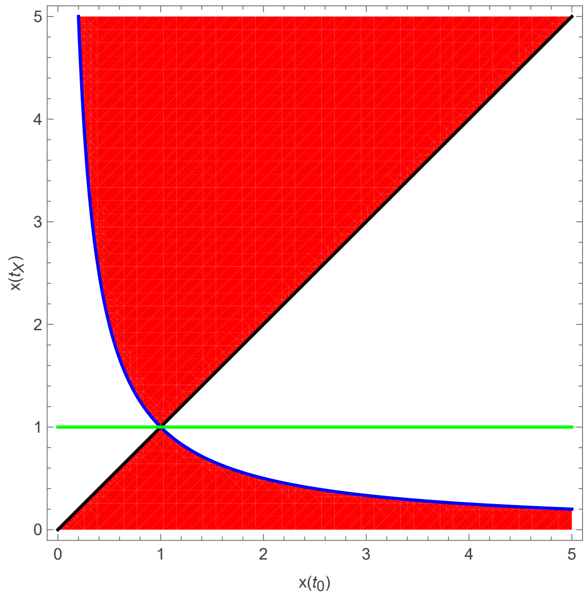

Fig. 1 The allowed (white) and forbidden (red) regions for $x(t_X)$ given $x(t_0)$ where $X = 1$ under (46) and $X = \infty$ under (45). The boundary curves $x(t_0)$ and $1/x(t_0)$ are shown in black and blue respectively. Continuous dynamics satisfying (46) cannot cross the equilibrium (shown in green). Plot made in Wolfram Mathematica 14.3[42].

The allowed region for $x(t_\infty)$ under (45) and $x(t_1)$ under (46), given $x(t_0)$, are shown in white in Figure 1 and the forbidden region is in red. The boundary curves $x(t_0)$ and $1/(x(t_0))$ are shown as the black and blue curves. Under (45), $x(t_\infty)$ may lie on a boundary curve, but under (46) $x(t_1)$ may not, except at the intersection of both boundary curves at $x(t_0) = x(t_X) = 1$. Continuous time dynamics (e.g. differential equations) obeying (46) cannot cross equilibrium, but discrete time dynamics (e.g. iterative maps) can. Under (45), $x(t_\infty) \in [x(t_0)^w, x(t_0)^{-w}]$ where $w = sign(1 - x(t_0))$ so the $x \to 0$ boundary is not approached in the long run if $x(t_0) < 1$. Under (46), $x(t_1) \in (x(t_0)^w, x(t_0)^{-w})$ so we are strictly repelled away from the $x \to 0$ boundary any time $x(t_0) < 1$.

4.2 The boundary divergence vanishing scale limits

Remark 4.1. *The historic Meta-Principle of Recognition Science—"Nothing cannot recognize itself" was originally developed through linguistic reasoning. The item "nothing" cannot have properties. However, "nothing" recognizing itself would give it properties (being a recognizer and being recognized). Therefore, "nothing" cannot recognize itself.*

A previously presented mathematical interpretation of the Meta Principle from [11] involved the empty set. A recognition event is an ordered pair of a "recognizer" a from set A and a "recognized" b from set B. The total set of recognition events is $\text{Recognition}(A, B) = A \times B$*. If either A or B is the empty set* $\emptyset$*, then the* $\text{Recognition}(A, B)$ *set is also the empty set. In particular,* $\text{Recognition}(\emptyset, \emptyset) = \emptyset$*.*

Another heuristic version of the Meta-Principle, is to consider the cost function J and identify $x = 0$ *with "nothing". Therefore, under this heuristic, "nothing" has undefined cost (it is formally outside of the domain of J), and the approach to "nothing" has infinite cost.*

While this is reasonable as a heuristic or philosophical point, it is prudent to examine what $x = 0$ and $x \to 0$ actually mean in terms of states. Looking at the definition (25), and the fact that the ι functions are positive scale maps, we can determine that $x = 0$ exactly implies a domain problem for the ι functions which are meant to be positive. Examining 0 as the limit is more interesting,

$$x_{nab} \to 0 \implies (\iota_n(R_n(a)) \to 0) \vee (\iota_n(R_n(b)) \to \infty). \tag{49}$$

The ratio x approaching 0 (or ∞, as reciprocals have the same cost) corresponds to an infinitely bad match between the a and b states. There are a few ways with which this could occur:

1. Keep a, b, and R_n constant, but re-parameterize the scale map ι_n.
2. Keep a and b, but allow the recognizer to change $R_n \to R_m$. Notice that if the recognizer changes, the scale map must also change to ι_m because ι_n is defined to map from E_n to $\mathbb{R}_{>0}$, not from E_m.

3. Keep R_n and ι_n, but vary the state a such that $\iota_n(R_n(a)) \to 0$
4. Keep R_n and ι_n, but vary the state b such that $\iota_n(R_n(b)) \to \infty$
5. A more complicated process where multiple variations occur simultaneously.

Option 1 has the advantage that it would allow for a well defined continuous limit, but it involves neither the state space nor the event space. Option 2 could conceivably be interpreted as an approach to a configuration in event space, but it would involve the requirement on the ι functions that for event space configurations, one of these occurs:

$$(E_{na} \to E_\emptyset \implies \iota_n(E_{na}) \to 0) \tag{50}$$

$$(E_{nb} \to E_\emptyset \implies \iota_n(E_{nb}) \to \infty). \tag{51}$$

We can think of (50) and (51) as "boundary event compatibility constraints on ι_n". It is important to distinguish that $E_\emptyset$ is a limit of a sequence in event space[3] and need not correspond to an element actually present in the event space. Options 3 and 4 would involve a or b approaching specific configurations in the state space itself[4]. This would require that

$$a \to C_\emptyset \implies R_n(a) = E_{na} \to E_\emptyset \tag{52}$$

and (50) or

$$b \to C_\emptyset \implies R_n(b) = E_{nb} \to E_\emptyset \tag{53}$$

and (51). The conditions (52) and (53) are effectively "boundary state compatibility constraints on R_n". Similarly to $E_\emptyset$, $C_\emptyset$ is the limit of a sequence in state space and need not be actually present in the state space.

To summarize, a meaningful identification of the infinite cost, $x \to 0^+$ limit with the limit of a sequence in event space involves an approach to $E_\emptyset$ and a scale map compatibility criterion ((50) or (51)). Extending this to the limit of a sequence in state space requires an approach to $C_\emptyset$ and a compatibility criterion on the recognizer ((52) or (53)).

5 Deriving Structure (T2 & T3)

Having established the global minimum of the cost ($x = 1$), and the divergence of the $x \to 0$ boundary, we now discuss the fundamental structural properties of Discreteness and Conservation.

5.1 Discreteness (T2)

In this framework, discreteness is an operational consequence of recognition, not a claim about an underlying continuum. Discreteness (T2) is not a consequence of the cost axioms alone. It requires the additional postulate of finite local resolution in the recognition quotient (as formalized in Recognition Geometry).

The Recognition Geometry[10] introduces locality via neighborhoods and postulates *finite local resolution*: any recognizer can distinguish only finitely many outcomes within a local region. The formal statement of this, as given in [10], was

Finite local resolution: *For every configuration $c \in \mathcal{C}$ and recognizer R, there exists a neighborhood $U = \mathcal{N}(c)$ such that the image $R(U)$ is a finite set, i.e. $|R(U)| < \infty$.*

A consequence of finite local resolution as stated above is that recognizers are not injective if the underlying state space $\mathcal{C}$ is continuous: a neighborhood $U = \mathcal{N}(c)$ in state space has infinitely many elements, but the event space $R(U) = \mathcal{E}$ has finitely many elements. The event space for a neighborhood is always therefore countable.

Observable space is therefore partitioned into finitely many *resolution cells* at any fixed local scale, such that every point in the state neighborhood is in the fiber $\{c \in \mathcal{C} | R(c) = e\}$ of one of a finite number of events e.

The cost landscape J then governs stability across these cells. In particular, the calibrated curvature

$$J''(1) = 1 \tag{54}$$

[3] The scale map ensures that the sequence $\iota_n(E_{na})$ for $n = 0, 1, 2, \ldots$ consists of positive numbers, so approaching 0 or ∞ is a standard notion. However, event spaces can have a variety of structures or element types, under which the notion of approach may have to be defined in a context specific way.

[4] Like event spaces, state spaces are quite general so the notion of approach may be context specific

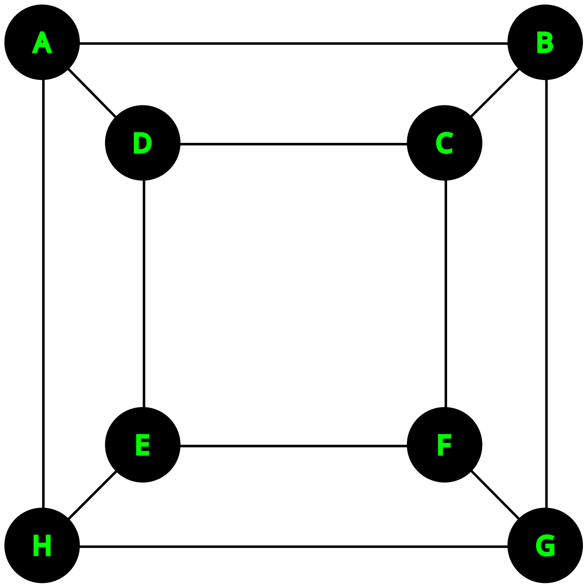

Fig. 2 Graph for examples 1 and 2 made in Inkscape.

sets the local stiffness of the mismatch penalty near perfect match, so that transitions between distinct resolution cells away from equilibrium incur positive cost while variations within a cell are observationally invisible.

$J(x)$ is a continuous function over $\mathbb{R}_{>0}$, but x, when treated as $x_{nA(t)a}$ is not typically continuous. This could be due to fundamental discreteness in the underlying state space $\mathcal{C}$, induced by the finite resolution of the recognizers R_n, or both.

It is easiest to illustrate this behavior with some examples. These examples will also allow us to examine the RS ontological states. It should be kept in mind that the purpose of these examples is pedagogy rather than realistic modeling.

5.1.1 Example 1: Fundamental state space discreteness

Consider a configuration space as the graph isomorphic to the vertices of a cube in Figure 2. Suppose the evolution in state space is

$$A(t_0) = \text{node } b, A(t_1) = \text{node } a, A(t_2) = \text{node } h, A(t_3) = \text{node } g, A(t_4) = \text{node } f. \tag{55}$$

Further, suppose the recognizer R_1 gives the letter of the node, and

$$\iota_1(w) = \begin{cases} 1, w = a \\ 8, w = b \\ 2, w = c \\ 7, w = d \\ 3, w = e \\ 6, w = f \\ 5, w = g \\ 4, w = h \end{cases} \tag{56}$$

allowing our reference configuration to be node a makes it so that

$$x_{1A(t)a} = \frac{\iota_1(R_1(A(t)))}{\iota_1(R_1(a))} = \iota_1(R_1(A(t))), \tag{57}$$

we get the evolution described by Table 1.

This example also allows us to examine the ontological states of RSExists and RSReal. Given the definition in (57), $\text{RSExists}(x_{1Aa}) \implies A = a$, so the only configuration c^* such that $\text{RSExists}_{\mathcal{C}}(c^*)$ is satisfied is $c^* = a$. Because of the underlying discreteness of the graph in Figure 2, a discrete

t	$R_1(A(t))$	$\iota_1(R_1(A(t))) = x_{1A(t)a}$	$J(x_{1A(t)a})$
t_0	b	8	49/16
t_1	a	1	0
t_2	h	4	9/8
t_3	g	5	8/5
t_4	f	6	25/12

Table 1 Evolution for example 1. Notice that the t parameter, configuration space, event space, and cost all move in discrete jumps

t	$R_2(A(t))$	$\iota_2(R_2(A(t))) = x_{2A(t)a}$	$J(x_{2A(t)a})$
t_0	False	2	1/4
t_1	True	1	0
t_2	False	2	1/4
t_3	False	2	1/4
t_4	False	2	1/4

Table 2 Evolution for example 2. In this case, a change in the state space occurs as $t_2 \to t_3$ and $t_3 \to t_4$, but neither have observational consequences under R_2

skeleton $\mathcal{D}_U$ (the graph itself) naturally exists, and the composite synthesis map $F(u) = \iota_1(R_1(u))$ is in (57). Then $\text{RSReal}_{F,\mathcal{D}_U}(x_{1A(t)a})$ again requires $A = a$.

5.1.2 Example 2: Fundamental and Induced Discreteness

In this example, we still consider the same state space (the nodes from Figure 2) and the same state evolution (55), but the recognizer is coarser: R_2 returns the classical Booleans "true" if the node label is a vowel and "false" if the node label is a consonant. The scale map is now

$$\iota_2(w) = \begin{cases} 1, w = \text{true} \\ 2, w = \text{false}, \end{cases} \tag{58}$$

which is bijective, and hence injective. Keeping a as our reference configuration and the same state sequence (55), the evolution in this case is shown in Table 2, which has an important difference that there is no observed change between t_2 and t_3 for R_2.

Because R_2 returns classical Booleans, we can use this example to illustrate some of the properties of the RSTrue condition. Define these three state to state maps

$$\begin{aligned} B_- &: \text{If at node a, stay at node a,} \\ &\quad \text{Else, go to node with the letter that precedes the current node's letter} && (59) \\ B_+ &: \text{If at node h, stay at node h,} \\ &\quad \text{Else, go to node with the letter that succeeds the current node's letter} && (60) \\ B_c &: \text{If at node a, go to node h,} \\ &\quad \text{Else, go to node with the letter that precedes the current node's letter.} && (61) \end{aligned}$$

Further define

$$\chi_{2p}(q) = \frac{\iota_2(R_2(q))}{\iota_2(R_2(p))}. \tag{62}$$

Consider $\text{RSTrue}_{B_-,\chi_2,c_0}(R_2, q)$. The conditions for this to be satisfied are

$$\begin{aligned} \text{RSExists}(\chi_{2p}(q)) \Leftrightarrow \chi_{2p}(q) = 1 \Leftrightarrow \iota_2(R_2(q)) = \iota_2(R_2(p)) \Leftrightarrow R_2(q) = R_2(p) && (63) \\ R_2(q) = 1 && (64) \\ \exists N \in \mathbb{N}\ \forall n \geq N,\ R_2(B_-^n(c_0)) = R_2(q) && (65) \end{aligned}$$

where we have used the injectivity of ι_2 in the final implication of (63). The condition (63) is satisfied if p and q are the same type of node, i.e. both labeled by vowels or both labeled by consonants. From

t	$R_3(c(t))$	$\iota_3(R_3(c(t))) = x_{3c0}$	$J(x_{3c0})$
$0 \le t < 0.121$	0	1	0
$0.121 \le t < 0.791$	1	3/2	1/12
$0.791 \le t \le 1$	0	1	0
$1 < t < 1.209$	-1	1/2	1/4
$1.209 \le t < 1.879$	-2	0.295	0.842
$1.879 \le t < 2$	-1	1/2	1/4

Table 3 Evolution for example 3. While the state space undergoes continuous evolution, the finite resolution of the recognizer induces discrete "jumps" in the event space, which extend to the recognition ratio and cost. Numbers without a simple rational form are rounded to 3 decimal places

the definition of R_2, we have that $R_2(q)$ being true (64) implies q is a vowel node; node a or node e. The condition (65) also requires q to be a vowel node because $B^n_-(c_0) =$ node a for all c_0 on the graph and $n > 7$. Taken together, $\text{RSTrue}_{B_-,\chi_2,c_0}(R_2, q)$ is satisfied when q and p are both vowel nodes and is not satisfied otherwise.

The conditions for the modified version $\text{RSTrue}_{B_+,\chi_2,c_0}(R_2, q)$ to be satisfied are still (63) and (64), with the modification to the third condition

$$\exists N \in \mathbb{N}\ \forall n \ge N,\ R_2(B^n_+(c_0)) = R_2(q) \tag{66}$$

Because $B^n_+(c_0) =$ node h for all c_0 on the graph and $n > 7$, (66) is now only satisfied if q is a consonant node. This is potentially compatible with the RSExistence (63) if p is also a consonant node, but it is not compatible with (64), so $\text{RSTrue}_{B_+,\chi_2,c_0}(R_2, q)$ fails. However, an altered condition $\text{RSTrue}_{B_+,\chi_2,c_0}(\neg R_2, q)$ would be satisfied if p and q are consonant nodes.

Finally, $\text{RSTrue}_{B_c,\chi_2,c_0}(R_2, q)$ cannot be satisfied because $R_2(B^n_c(c_0))$ never stabilizes. We likewise cannot satisfy $\text{RSTrue}_{B_c,\chi_2,c_0}(\neg R_2, q)$ for the same reason.

5.1.3 Example 3: Purely induced Discreteness with piecewise evolution

In this case, we allow for a continuous state space $c \in \mathbb{R}$ and continuous trajectory parameter $0 \le t < 2$. This particular example will show the emergence of an observable piecewise continuous evolution.

The state evolution obeys

$$c(t) = 5t(t-1)(t-2) \tag{67}$$

Further define a recognizer R_3 as the "floor function"

$$R_3(c) = \lfloor c \rfloor, \tag{68}$$

which is compatible with the finite resolution. Further, let the scale map be

$$\iota_3(w) = \frac{2}{\pi}\arctan(w) + 1 \tag{69}$$

which is positive (specifically, $\iota_3 \in (0, 2)$ for $w \in \mathbb{R}$), and let the reference configuration be 0 such that

$$x_{3c0} = \frac{\iota_3(R_3(c))}{\iota_3(R_3(0))} = \iota_3(R_3(c)). \tag{70}$$

For comparison, we can also define an unphysical infinite resolution recognizer and cost bridge

$$R_\infty(c) = c, \quad \iota_\infty(w) = \iota_3(w), \quad x_{\infty c0} = \frac{\iota_\infty(R_\infty(c))}{\iota_\infty(R_\infty(0))} = \iota_\infty(R_\infty(c)) \tag{71}$$

Plots of the configuration in state space c, event space R_3, recognition ratio x_{3c0}, and cost J for this example are shown in Figure 3, with a table giving specific values and the relevant intervals of t in Table 3.

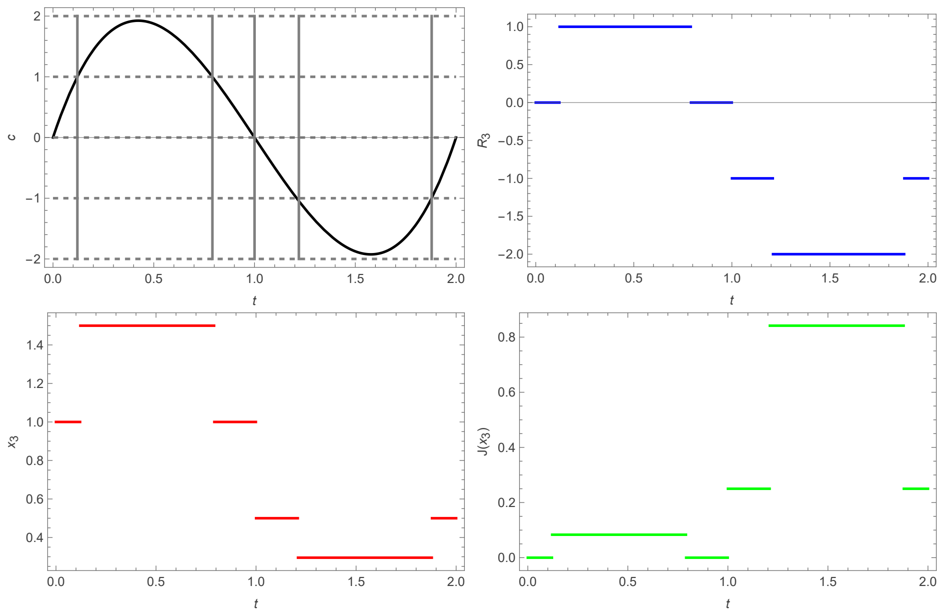

Fig. 3 Graphs of $c(t)$, $R_3(c(t))$, $x_{3c0} = \iota_3(R_3(c(t)))$, and $J(x_{3c0})$. Notice that, even though the state space $C = \mathbb{R}$ is continuous, the event space, x, and J cost all evolve through discrete jumps. The vertical lines on the plot of $c(t)$ show when the resolution cell changes. Plots made in Wolfram Mathematica 14.3[42].

Dynamics in this case proceeds as a sequence of cell-to-cell updates (discrete observable evolution) even if the underlying configuration space admits continuous representatives. The associated cost difference for a cell to cell jump is finite (rather than infinitesimal). This is in contrast with the behavior of the unphysical infinite resolution recognition process (71), where the event space is continuous and the cost barrier to migrate to a different event is infinitesimal. The existence of finite rather than infinitesimal cost barriers separating observable states has been considered as motivation for considering discrete states in previously published work [14] and in the Lean database [13].

Notice that $\iota_3(R_3(c))$ naturally lies on a discrete skeleton isomorphic to the integers $\mathbb{Z}$, so the RSExistent configurations c such that $x_{3c0} = 1$, i.e. $c \in [0, 1)$, are also RSReal with respect to the induced skeleton. However, the unphysical recognizer $R_\infty(c)$ can result in RSExistent configurations that aren't RSreal due to the lack of natural discrete skeleton.

5.1.4 Example 4: Purely induced discreteness with irregular evolution

The finite resolution axiom alone does not guarantee piecewise continuous event space evolution. Consider a state space $[0, 1]$ and state evolution $c = t$ where t is a continuous parameter, with the recognizer

$$R_{\mathbb{Q}}(c) = \begin{cases} \text{true} & c \in \mathbb{Q} \\ \text{false} & c \notin \mathbb{Q} \end{cases} \tag{72}$$

that test whether the number is rational. There are only two possible elements in the event space, so the finite resolution axiom holds, but the resolution cells and hence the evolution are nowhere continuous.

5.2 The Ledger (T3)

The primary exposition of the ledger, including proof-bearing discrete-dynamics, development of atomic ticks, balance-preserving ledger updates on graphs, and the 2^d-tick hypercube period (including the 8-tick $d = 3$ case) is presented in [11]. Here, we summarize some important results.

The reciprocal symmetry

$$J(x) = J(1/x) \tag{73}$$

implies that a deviation by a ratio x carries the same mismatch penalty as its inverse. To connect this reciprocity to conservation, we make explicit the balance invariant of a closed recognition network.

Let a ledger state be a vector of ratios $X = (x_1, \ldots, x_N) \in (\mathbb{R}_{>0})^N$ and define the *balance functional*

$$\mathcal{B}(X) := \prod_{i=1}^{N} x_i. \tag{74}$$

It is also helpful to examine ledger balance in terms of logarithmic coordinates $u = \log(x)$

$$\widetilde{\mathcal{B}}(U) = \sum_{i=1}^{N} u_i. \tag{75}$$

A closed system is balanced when $\mathcal{B}(X) = 1$ or $\widetilde{\mathcal{B}}(U) = 0$. It is a "structural assumption" that the balance $\mathcal{B}(X)$ and $\widetilde{\mathcal{B}}(U)$ are invariant under updates, and it is derived from the Ledger Axioms **L1** (deterministic updates) and **L2** (no metadata) that updates occur in atomic ticks[11]. A local interaction of strength $r > 0$ between two entries $i \neq j$ updates

$$x_i \mapsto x_i r, \qquad x_j \mapsto x_j / r, \tag{76}$$

$$u_i \mapsto u_i + \log(r), \qquad u_j \mapsto u_j - \log(r) \tag{77}$$

leaving all other coordinates fixed. In this sense every "debit" by $\delta x_i = r,\ \delta u_i = \log(r)$ is accompanied by a "credit" by $\delta x_j = 1/r,\ \delta u_j = -\log(r)$, yielding a double-entry ledger form; conservation-like constraints can be represented as invariants of balance-preserving updates. The RCL encodes how the associated recognition costs compose under the paired multiplicative operations $(x, y) \mapsto (xy, x/y)$.

5.2.1 A balanced ledger treatment of electric charge

Consider a 1D chain of N atoms, a neutral reference state c_0, a recognizer R_q that gives the net charge q on the atom, and scale map $\iota_q(q) = e^q$. Then

$$x_n = \frac{\iota_q(R_q(c_n))}{\iota_q(R_q(c_0))} = \frac{e^{q_n}}{1} \tag{78}$$

and

$$u_n = \log(x_n) = q_n, \tag{79}$$

so then U is a vector of the ionization state of each atom in the chain, $U \in \mathbb{Z}^N$. Updates of the form

$$x_i \mapsto x_i e^{\Delta q}, \qquad x_j \mapsto x_j e^{-\Delta q}, \tag{80}$$

$$u_i \mapsto u_i + \Delta q, \qquad u_j \mapsto u_j - \Delta q \tag{81}$$

would preserve the value of balance $\mathcal{B}(X)$ and of logarithmic balance $\widetilde{\mathcal{B}}(u)$, which in this example is literally the total charge in atomic units.

5.2.2 Ledger Laplacian limit

Suppose we have a ledger as an N dimensional hypercube lattice where individual elements are specified by $X \in \mathbb{Z}^N$. Let there be a nonnegative function defined on the elements $\psi(X)$, and a corresponding real function $\xi(X) = \log(\psi(X))$. The total cost of nearest neighbor comparison on the lattice will be

$$J_{NN} = \sum_X \sum_{\hat{i}} J(\psi(X)/\psi(X + \hat{i})) + J(\psi(X)/\psi(X - \hat{i})), \tag{82}$$

where $\hat{i}$ ranges over the unit vectors. Notice this can be rewritten as

$$J_{NN} = \sum_X \sum_{\hat{i}} \widetilde{J}(\xi(X) - \xi(X + \hat{i})) + \widetilde{J}(\xi(X) - \xi(X - \hat{i})). \tag{83}$$

Recall that

$$\widetilde{J}(u) = \cosh(u) - 1 \approx \frac{u^2}{2} + \frac{u^4}{24} + O(u^6) \tag{84}$$

Therefore, if the difference in the field value ξ between adjacent elements is much less than 1,

$$J_{NN} \approx \sum_X \sum_{\hat{i}} \frac{((\xi(X) - \xi(X + \hat{i}))^2}{2} + \frac{(\xi(X) - \xi(X - \hat{i}))^2}{2}. \tag{85}$$

Taking the functional derivative with respect to $\xi(X_0)$ gives

$$\frac{\delta J_{NN}}{\delta \xi(X_0)} \approx 2 \sum_{\hat{i}} [2\xi(X_0) - \xi(X_0 + \hat{i}) - \xi(X_0 - \hat{i})]. \tag{86}$$

Notice however that $\sum_{\hat{i}}[2\xi(X_0) - \xi(X_0 + \hat{i}) - \xi(X_0 - \hat{i})]$ is the negative of the Lattice Laplacian operator. If the lattice size is a, then scaling $\frac{1}{a^2}$ times the Lattice Laplacian gives the continuum Laplacian in the $a \to 0$ limit. This means that under more general variational problems, the stationarity of the nearest neighbor comparison cost $\frac{\delta J_{NN}}{\delta \xi(X)} = 0$ gives a Laplacian equation $\nabla^2 \xi = 0$, under the restrictions that the nearest neighbor variation and lattice spacing are small. This result has been formalized in the Lean database [13].

A more elaborate objective, involving both nearest neighbor comparison and other sorts of terms, would result in additional terms in the equation of motion. For instance, consider now a time dependent $\psi(T, X)$, $\xi(T, X)$, where $T \in \mathbb{Z}$, with an objective

$$J_{tot} = \sum_X \left(\frac{A}{\Delta t} J(\psi(T, X)/\psi(T-1, X)) + \frac{B}{a^2} \sum_{\hat{i}} J(\psi(T, X)/\psi(T, X + \hat{i})) + J(\psi(T, X)/\psi(T, X - \hat{i})) \right) \tag{87}$$

The first term $J_{lag} = \sum_X \frac{A}{\Delta t} J(\psi(T, X)/\psi(T-1, X))$ involves comparison of the value of the field at some location to its previous value at that location. Recasting this in terms of ξ and applying the same small difference condition gives $J_{lag} \approx \frac{A}{\Delta t} \sum_X \frac{(\xi(T,X) - \xi(T-1,X))^2}{2}$. Taking the functional derivative of J_{lag} with respect to $\xi(T, X_0)$ now gives

$$\frac{\delta J_{lag}}{\delta \xi(T, X_0)} \approx A \frac{\xi(T, X_0) - \xi(T-1, X_0)}{\Delta t}. \tag{88}$$

Notice that in the continuum $\Delta t \to 0$ limit, this becomes the standard time derivative, so the net equation of motion

$$\frac{\delta J_{tot}}{\delta \xi(T, X_0)} = 0 \Rightarrow -A \frac{\partial \xi}{\partial t} = 2B \nabla^2 \xi \tag{89}$$

is structurally a heat equation if $-2B/A$ is positive or an antidiffusion equation if $-2B/A$ is negative. Modification of the objective with different terms would lead to different partial differential equations in the restriction to small difference and continuum approximations. Deviations of the underlying J cost/ ledger system and the emergent PDEs could be observable at small physical scales or under rapid site to site variation.

6 Conclusion

The primary purpose of this paper was to present and formalize some results from the Recognition Science reasoning chain.

1. **The Primitive:** The Recognition Composition Law (RCL) is the functional equation governing interaction.
2. **The Terrain:** Imposing normalization and calibration on the RCL uniquely forces the cost potential $J(x) = \frac{1}{2}(x + x^{-1}) - 1$.
3. **Identity (T0):** The ground state $J(1) = 0$ defines identity as the lowest cost state.

4. **Existence (T1):** The singularity $J \to \infty$ creates an infinite cost barrier to event or configuration sequences that result in the argument x going to 0
5. **Structure (T2–T3):** Finite local resolution in recognition quotients yields discreteness at the observable level; reciprocity together with a balance-preserving update rule yields a double-entry ledger form. The curvature of J sets the stability scale near perfect match.

6.1 Outlook

The results presented in this paper do not represent a complete physical theory that has been machine checked, or necessarily reduced to the simplest possible form. However, the publicly available Lean database [13] currently uses 12 basic axioms and compiles with zero "sorry"s. It contains modules with preliminary theorems that will be formally analyzed in future publications. In addition to mathematics and physical sciences, there is ongoing research examining applications of the cost function J and the ledger formalism to a variety of other fields, such as artificial intelligence.

Declarations

- **Acknowledgements:** We would like to thank Amir Rahnama for useful discussions about the scale maps and Anil Thapa, Megan Simons, Sebastian Pardo Guerra and Milan Zlatanovic for a careful reading of earlier drafts.

- **Author contributions:** Conceptualization, J.W.; formal analysis, J.W. and P.B.; writing—original draft preparation, J.W.; writing—review and editing, J.W. and P.B.; figures, P.B. All authors have read and agreed to the published version of the manuscript.

- **Funding:** This research received no external funding.

- **Conflict of interest/Competing interests:** The authors declare no conflicts of interest.
- **Data availability:** No datasets were generated or analyzed during the current study.

References

[1] Feynman, R.P., Leighton, R.B., Sands, M.: The Feynman Lectures on Physics, (1963)

[2] Arnold, V.I.: Mathematical Methods of Classical Mechanics. Graduate Texts in Mathematics. Springer, New York (1989). https://doi.org/10.1007/978-1-4757-2063-1

[3] Tong, D.: Lectures on Classical Dynamics. Lecture notes, available at https://www.damtp.cam.ac.uk/user/tong/dynamics.html (2019)

[4] Peskin, M.E., Schroeder, D.V.: An Introduction to Quantum Field Theory. Addison-Wesley, Reading, USA (1995). https://doi.org/10.1201/9780429503559

[5] Lancaster, T., Blundell, S.J.: Quantum Field Theory for the Gifted Amateur. Oxford University Press, Oxford (2014). https://doi.org/10.1093/acprof:oso/9780199699322.001.0001 . https://doi.org/10.1093/acprof:oso/9780199699322.001.0001

[6] Wald, R.M.: General Relativity, (1984)

[7] Carroll, S.M.: Lecture notes on general relativity (1997) arXiv:gr-qc/9712019

[8] Tong, D.: Lectures on Statistical Physics. Lecture notes, available at https://www.damtp.cam.ac.uk/user/tong/statphys.html (2012)

[9] Sakurai, J.J., Napolitano, J.: Modern Quantum Mechanics, 3rd edn. Cambridge University Press, Cambridge (2020)

[10] Washburn, J., Zlatanović, M., Allahyarov, E.: Recognition geometry. Axioms **15**(2) (2026) https://doi.org/10.3390/axioms15020090

[11] Pardo-Guerra, S., Simons, M., Thapa, A., Washburn, J.: Coherent Comparison as Information Cost: A Cost-First Ledger Framework for Discrete Dynamics. arXiv e-prints, 2601–12194 (2026) https://doi.org/10.48550/arXiv.2601.12194 arXiv:2601.12194 [math.IT]

[12] Moura, L.d., Ullrich, S.: The Lean 4 Theorem Prover and Programming Language, pp. 625–635. Springer, Berlin, Heidelberg (2021). https://doi.org/10.1007/978-3-030-79876-5_37 . https://doi.org/10.1007/978-3-030-79876-5_37

[13] Washburn, J.: recognition-science. https://github.com/jonwashburn/recognition-science. Accessed: 2026-03-04 (2026)

[14] Simons, M., Allahyarov, E., Washburn, J.: A discrete informational framework for classical gravity: Ledger foundations and galaxy rotation curve constraints. Entropy **28**(4) (2026) https://doi.org/10.3390/e28040477

[15] Washburn, J., Allahyarov, E.: A Complete Derivation of the Fermion Spectrum from the Recognition Composition Law. arXiv e-prints, 2506–12859 (2025) https://doi.org/10.48550/arXiv.2506.12859 arXiv:2506.12859 [physics.gen-ph]

[16] Washburn, J., Simons, M., Allahyarov, E.: A noble-gas-centered coordinate for within-period ionization energy trends and subshell anomalies. ChemRxiv **2026**(0424) (2026) https://doi.org/10.26434/chemrxiv.15002361/v1 https://chemrxiv.org/doi/pdf/10.26434/chemrxiv.15002361/v1

[17] Rovelli, C.: Relational quantum mechanics. International Journal of Theoretical Physics **35**(8), 1637–1678 (1996) https://doi.org/10.1007/BF02302261 arXiv:quant-ph/9609002 [quant-ph]

[18] Zeilinger, A.: A Foundational Principle for Quantum Mechanics. Found. Phys. **29**(4), 631–643 (1999) https://doi.org/10.1023/A:1018820410908

[19] Jaynes, E.T.: Information theory and statistical mechanics. Phys. Rev. **106**, 620–630 (1957) https://doi.org/10.1103/PhysRev.106.620

[20] Wheeler, J.A.: Information, physics, quantum: The search for links, pp. 354–368. Physical Society of Japan, Tokyo (1990). Physical Society of Japan

[21] Caticha, A.: Entropic Dynamics. Entropy **17**(9), 6110–6128 (2015) https://doi.org/10.3390/e17096110 arXiv:1509.03222 [quant-ph]

[22] Verlinde, E.: On the origin of gravity and the laws of Newton. Journal of High Energy Physics **2011**, 29 (2011) https://doi.org/10.1007/JHEP04(2011)029 arXiv:1001.0785 [hep-th]

[23] Kottos, T., Smilansky, U.: Quantum chaos on graphs. Phys. Rev. Lett. **79**, 4794–4797 (1997) https://doi.org/10.1103/PhysRevLett.79.4794

[24] Barra, F., Gaspard, P.: Classical dynamics on graphs. Physical Review E **63**(6), 066215 (2001) https://doi.org/10.1103/PhysRevE.63.066215 arXiv:nlin/0011045 [nlin.CD]

[25] Kobayashi, T.J., Loutchko, D., Kamimura, A., Horiguchi, S.A., Sughiyama, Y.: Information geometry of dynamics on graphs and hypergraphs. Information Geometry **7**(1), 97–166 (2024)

[26] Wolfram, S.: Statistical mechanics of cellular automata. Rev. Mod. Phys. **55**, 601–644 (1983) https://doi.org/10.1103/RevModPhys.55.601

[27] 't Hooft, G.: The Cellular Automaton Interpretation of Quantum Mechanics. arXiv e-prints, 1405–1548 (2014) https://doi.org/10.48550/arXiv.1405.1548 arXiv:1405.1548 [quant-ph]

[28] Regge, T.: GENERAL RELATIVITY WITHOUT COORDINATES. Nuovo Cim. **19**, 558–571 (1961) https://doi.org/10.1007/BF02733251

[29] Bombelli, L., Lee, J., Meyer, D., Sorkin, R.D.: Space-time as a causal set. Phys. Rev. Lett. **59**, 521–524 (1987) https://doi.org/10.1103/PhysRevLett.59.521

[30] Washburn, J., Zlatanović, M., Allahyarov, E.: The d'alembert inevitability theorem. Mathematics **14**(8) (2026) https://doi.org/10.3390/math14081386

[31] Aczél, J.: Lectures on Functional Equations and Their Applications. Mathematics in science and engineering. Academic Press, New York (1966). https://books.google.com/books?id=0vZQAAAAMAAJ

[32] Washburn, J., Zlatanović, M.: Uniqueness of the canonical reciprocal cost. Mathematics **14**(6) (2026) https://doi.org/10.3390/math14060935

[33] Kannappan, P.: The functional equation $f(xy) + f(xy^{-1}) = 2f(x)f(y)$ for groups. Proceedings of the American Mathematical Society **19** (1968)

[34] Stetkær, H.: D'alembert's functional equation on groups. Banach Center Publications **99** (2013) https://doi.org/10.4064/bc99-0-11

[35] Aczel, J., Dhombres, J.: D'Alembert's functional equation. An application to noneuclidean mechanics. Encyclopedia of Mathematics and its Applications, pp. 103–113. Cambridge University Press, Cambridge (1989)

[36] Itakura, F., Saito, S.: An analysis-synthesis telephony system based on the maximum likelihood method, vol. C, pp. 17–20 (1968)

[37] Févotte, C., Bertin, N., Durrieu, J.-L.: Nonnegative matrix factorization with the itakura-saito divergence: With application to music analysis. Neural Computation **21**(3), 793–830 (2009) https://doi.org/10.1162/neco.2008.04-08-771

[38] Amari, S.-i., Nagaoka, H.: Methods of Information Geometry. Translations of Mathematical Monographs, vol. 191. American Mathematical Society and Oxford University Press, Providence, Rhode Island (2000)

[39] Banerjee, A., Merugu, S., Dhillon, I., Ghosh, J.: Clustering with Bregman Divergences, pp. 234–245. https://doi.org/10.1137/1.9781611972740.22 . https://epubs.siam.org/doi/abs/10.1137/1.9781611972740.22

[40] Washburn, J., Rahnamai Barghi, A.: Reciprocal convex costs for ratio matching: Axiomatic characterization. Axioms **15**(2) (2026) https://doi.org/10.3390/axioms15020151

[41] Gold, E.M.: Language identification in the limit. Information and Control **10**(5), 447–474 (1967) https://doi.org/10.1016/S0019-9958(67)91165-5

[42] Inc., W.R.: Mathematica, Version 14.3. Champaign, IL, 2025. https://www.wolfram.com/mathematica