Article

Reciprocal Convex Costs for Ratio Matching: Axiomatic Characterization

Jonathan Washburn [1] and Amir Rahnamai Barghi [1,*]

[1] Recognition Physics Research Institute, Austin, TX, USA; jon@recognitionphysics.org (J.W.); arahnamab@gmail.com (A.R.B.)
* Correspondence: arahnamab@gmail.com

Abstract

We study ratio-induced mismatch costs functions of the form $c(s,o) = J\big(\iota_S(s)/\iota_O(o)\big)$ built from positive scale maps $\iota_S : S \to \mathbb{R}_{>0}$ and $\iota_O : O \to \mathbb{R}_{>0}$ and a penalty $J : (0,\infty) \to [0,\infty)$. Assuming inversion symmetry, strict convexity, coercivity, normalization at 1, and a multiplicative d'Alembert identity, we show that $f(u) := 1 + J(e^u)$ is continuous and satisfies the additive d'Alembert equation; hence, by a classical classification theorem, there exists $a > 0$ such that $J(x) = \cosh(a \log x) - 1 = \frac{1}{2}\big(x^a + x^{-a}\big) - 1, \quad x > 0$. We then analyze the associated argmin mapping over feasible scale sets: existence under explicit subspace-closedness assumptions, an explicit geometric-mean decision geometry for finite dictionaries with stability away from boundaries, exact compositionality for product models, and an optimal sequential mediation principle described by a geometric mean (or its log-space projection when infeasible). The paper is purely mathematical; any semantic interpretation is optional and external to the theorems proved here.

Keywords: functional equations; d'Alembert equation; reciprocal convex cost; ratio-based optimization; geometric-mean decision boundaries; compositionality; sequential mediation

MSC: Primary 39B52, 49J40; Secondary 26A51, 90C25, 94A17, 90C31

1. Introduction

This section introduces the optimization-based model of reference, fixes terminology and standing assumptions, and outlines the main results and organization.

The paper's goal is to make the ratio-matching paradigm mathematically explicit. We fix ratio-induced costs of the form $c(s,o) = J(\iota_S(s)/\iota_O(o))$ and define meanings by the argmin rule. The first question is structural: under inversion symmetry, convexity/regularity, and a multiplicative compatibility axiom, which mismatch penalties J are admissible, and how canonical is the resulting form? The second question is geometric: once J is fixed, what decision boundaries and stability properties are forced for finite dictionaries, and how do these behave under products and sequential mediation? The intended contribution is a self-contained set of theorems that separate what is proved inside the axioms from any external semantic or empirical interpretation.

We start with two sets:

- a *configuration (token) space* S (words, codes, internal states, messages, ...),
- an *object space* O (candidate referents, concepts, states of affairs, ...).

Academic Editor: Academic Editor:

Received:
Revised:
Accepted:
Published:

Terminology.

Throughout we use *configuration* (or *token*) for an arbitrary element $s \in S$. We reserve the term *symbol for o* for a configuration s satisfying the predicate in Definition 8, i.e. $o \in \mathrm{Mean}(s)$ together with the compression inequality $J_S(s) < J_O(o)$.

Model ingredients and notation

For quick reference, the functionals and maps used throughout are organized as follows.

- *Scale maps:* $\iota_S : S \to \mathbb{R}_{>0}$ and $\iota_O : O \to \mathbb{R}_{>0}$.
- *Mismatch penalty:* $J : (0, \infty) \to [0, \infty)$ (axioms in Definition 1, explicit choice in Definition 2).
- *Reference cost:* $c(s, o) := J(\iota_S(s)/\iota_O(o))$ (1).
- *Meaning set (argmin rule):* $\mathrm{Mean}(s) := \arg\min_{o \in O} c(s, o)$ (Definition 7).
- *Intrinsic costs:* $J_S(s)$ and $J_O(o)$ (Definition 3); in the canonical setting these are induced by scales via $J_S(s) = J(\iota_S(s))$ and $J_O(o) = J(\iota_O(o))$ (see Definition 6).
- *Symbol predicate:* s is a *symbol for o* if $o \in \mathrm{Mean}(s)$ and the compression inequality $J_S(s) < J_O(o)$ holds (Definition 8).

Each space is equipped with a positive *scale* map $\iota_S : S \to \mathbb{R}_{>0}$ and $\iota_O : O \to \mathbb{R}_{>0}$, interpreted as an intrinsic "size/complexity" in a common currency. Fix a cost functional $J : (0, \infty) \to [0, \infty)$ with the properties stated in Section 2 (symmetry under inversion, strict convexity, and a unique minimum at 1). We then define a *ratio-induced reference cost*

$$c(s, o) := J\left(\frac{\iota_S(s)}{\iota_O(o)}\right), \qquad (s, o) \in S \times O. \tag{1}$$

Meaning as minimization.

The *meaning set* of a configuration s is the set of objects achieving minimal cost:

$$\mathrm{Mean}(s) := \arg\min_{o \in O} \; c(s, o).$$

Equivalently, $o \in \mathrm{Mean}(s)$ iff $c(s, o) \le c(s, o')$ for all $o' \in O$ (Definition 7). Ties are allowed: meaning is set-valued unless uniqueness is proved under additional hypotheses.

Interpretive content (and its limits).

Because J is minimized at 1, low reference cost forces *scale matching*: a configuration can only refer cheaply to objects whose scale is close to its own. This yields an explicit, checkable constraint on admissible reference patterns. The framework is deliberately *axiomatic*: the scale maps and the chosen J are inputs.

1.1. A toy example: three-object dictionary

Let $O = \{o_1, o_2, o_3\}$ with scales $y_i := \iota_O(o_i)$ satisfying $0 < y_1 < y_2 < y_3$. For a configuration s with scale $x := \iota_S(s)$, the meaning rule compares the three costs $J(x/y_i)$. For the explicit functional (3), the boundary between preferring o_1 and o_2 occurs at the *geometric mean* $\sqrt{y_1 y_2}$, and similarly between o_2 and o_3 at $\sqrt{y_2 y_3}$ (Theorem 8). Thus the model induces a piecewise-constant semantic partition of the positive line in the configuration ratio x, with stability away from the boundary points.

1.2. Relation to prior work

Classical analyses of reference emphasize logical form and truth conditions (e.g. Frege and Russell) [1,2]. The symbol-grounding literature highlights that purely formal symbol manipulation does not by itself determine what symbols are about [4]. The present paper does not attempt to resolve these debates empirically. Instead, it isolates a mathematically tractable *selection principle*: aboutness is determined by minimizing an explicit mismatch cost. For comparison with contemporary subject-matter/aboutness and truthmaker-semantics accounts (e.g. Yablo [11], Hawke [12], and the *Philosophical Studies* symposium discussion [13–15]), see Section 9. The intended payoff is that, once scales are fixed, aboutness becomes a tractable variational problem with explicit decision boundaries and composition theorems.

This paper adopts an *optimization-first* viewpoint: once a mismatch cost is fixed, semantic *meaning* is defined by an argmin rule (Definition 7). A closely related *measurement-first* stance appears in *Recognition Geometry* [17], which takes recognition events as primitive and derives observable space as a quotient under an operational indistinguishability relation [17, Def. 4]. In the same spirit, the present framework treats mismatch costs as primitive measurements and regards stable meanings as effective equivalence classes of *cost-minimization events*. Both viewpoints emphasize operationally defined structure over a priori metaphysical commitments, and both isolate exactly which axioms must be validated when connecting the formalism to an empirical domain.

1.3. Contributions and what is proved

Within the ratio-induced model (1) (and the explicit choice (3) used throughout), we establish the following structural facts under clearly stated hypotheses:

- **Existence.** If the feasible scale set $\iota_O(O) \subset \mathbb{R}_{>0}$ is nonempty and closed in the usual topology on $(0, \infty)$ and if the minimum is attained (as made precise in Theorem 2), then every configuration admits at least one meaning.
- **Finite-dictionary decision geometry.** For finite ordered dictionaries, decision boundaries are given by geometric means of adjacent object scales, and meanings are locally stable away from these boundaries (Theorem 8 and Corollary 7).
- **Compositionality.** For product symbol/object spaces with separable scales, meaning factorizes componentwise (Theorem 5).
- **Mediation.** For sequential reference through an intermediate representation, the set of optimal mediator ratios is characterized explicitly in log-coordinates; and whenever the balance-point ratio b_{geo} is feasible, mediation weakly decreases the total mismatch cost relative to direct reference (Theorem 6 and Corollary 3).

1.4. Organization

Section 2 states the axioms for J and fixes the explicit mismatch functional (3). Section 3 defines costed spaces, ratio-induced reference, and the meaning relation. Section 4 contains the principal theorems, followed by compositionality (Section 5), extensions, and examples.

2. The mismatch functional J

This section fixes the scalar mismatch functional $J : (0, \infty) \to [0, \infty)$ used throughout to compare configuration and object scales via the ratio-induced cost (1). The role of J here is purely mathematical: it is an explicit penalty for scale mismatch, and no physical, cognitive, or linguistic interpretation is assumed.

2.1. Standard properties and canonicity

The conditions below are recorded as a compact axiom package for the mismatch penalty. They encode inversion symmetry, strict convexity, and a multiplicative compatibility under scale multiplication. After a log change of variables, the compatibility axiom becomes d'Alembert's functional equation, so the resulting class of penalties is classical.

We include a tailored derivation in Appendix A to keep the paper self-contained and to emphasize that the axioms are used only as mathematical assumptions, not as a claim of novelty.

Definition 1 (Cost Functional Axioms)**.** *A* mismatch functional *is a function* $J : (0, \infty) \to [0, \infty)$ *satisfying:*

1. ***Normalization**:* $J(1) = 0$*.*
2. ***Strict convexity**:* J *is strictly convex on* $(0, \infty)$*.*
3. ***Multiplicative d'Alembert identity**: for all* $x, y > 0$*,*

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

The d'Alembert identity (2) is the dominant structural constraint. Inversion symmetry is not *assumed as an axiom; it is derived from (2) and normalization in Lemma 1. We invoke strict convexity only in statements where uniqueness is required; existence and attainment statements are formulated without using strict convexity.*

Lemma 1 (Derived inversion symmetry)**.** *Assume* J *satisfies normalization* $J(1) = 0$ *and the multiplicative d'Alembert identity (2). Then for every* $x > 0$ *one has* $J(x) = J(x^{-1})$*.*

Proof. Set $y = 1$ in (2). Using $J(1) = 0$ we obtain

$$J(x) + J(x^{-1}) = 2J(x) + 2J(1) + 2J(x)J(1) = 2J(x),$$

hence $J(x^{-1}) = J(x)$. □

We first record a basic consequence used repeatedly: under strict convexity, the normalization point $x = 1$ *is the unique zero of* J*.*

Lemma 2 (Uniqueness of the zero-cost point)**.** *If* J *satisfies Definition 1, then* $J(x) = 0$ *implies* $x = 1$*.*

Proof. By (2) and (1), J attains its minimum value 0 at $x = 1$. By strict convexity (3), the minimizer is unique. Hence $J(x) = 0$ forces $x = 1$. □

2.2. The explicit choice used in this paper

Definition 2 (The functional fixed below)**.** *In the remainder of this paper we fix the explicit functional*

$$J(x) = \tfrac{1}{2}(x + x^{-1}) - 1 = \tfrac{(x-1)^2}{2x} \qquad (x > 0). \tag{3}$$

The next proposition verifies that the explicit functional indeed satisfies the axioms, so subsequent sections can treat Definition 1 as established.

Proposition 1 (Verification of the axioms)**.** *The function (3) satisfies Definition 1.*

Proof. Normalization and inversion symmetry are immediate from (3), and (3) shows $J(x) \geq 0$ for all $x > 0$. Differentiating $J(x) = \frac{1}{2}(x + x^{-1}) - 1$ gives

$$J'(x) = \tfrac{1}{2} - \tfrac{1}{2x^2}, \qquad J''(x) = \tfrac{1}{x^3} > 0 \quad (x > 0),$$

so J is strictly convex on $(0, \infty)$. For (4), set $C(x) = 1 + J(x) = \frac{1}{2}(x + x^{-1})$. Then

$$C(xy) + C(x/y) = \tfrac{1}{2}\Big(xy + \tfrac{1}{xy} + \tfrac{x}{y} + \tfrac{y}{x}\Big) = \tfrac{1}{2}\big(x + \tfrac{1}{x}\big)\big(y + \tfrac{1}{y}\big) = 2C(x)C(y),$$

which is equivalent to (2) after substituting $C = 1 + J$ and expanding. □

Proposition 2 (Classical characterization of J). *Assume $J : (0, \infty) \to [0, \infty)$ satisfies Definition 1. Then there exists a constant $a > 0$ such that for all $x > 0$,*

$$J(x) = \cosh(a \log x) - 1 = \tfrac{1}{2}(x^a + x^{-a}) - 1.$$

Moreover, if we replace the scale maps by $\tilde{\iota}_S := \iota_S^a$ and $\tilde{\iota}_O := \iota_O^a$, then the ratio-induced model with parameter a becomes the same model written with parameter 1. Consequently, one may take $a = 1$ without loss of generality at the level of the induced reference costs.

Proof. See Appendix A. □

Example 1 (Small-mismatch regime). *For $|u| \ll 1$ one has*

$$J(1 + u) = \frac{u^2}{2} + O(u^3),$$

so near balance the mismatch cost behaves like a quadratic penalty in the relative deviation.

3. Costed spaces and reference structures

We now formalize the axioms of the model introduced in Section 1. Throughout, the mismatch functional J is fixed as in Section 2. The intent is to make precise which pieces of data are inputs (configuration/object spaces and their scale maps) and which pieces are derived (reference costs and meaning).

3.1. Costed spaces

Definition 3 (Costed space). *Fix a mismatch functional $J : (0, \infty) \to [0, \infty)$ (Section 2). A* costed space *is a triple (C, J_C, ι_C) consisting of:*

- *a set C of configurations,*
- *a map $\iota_C : C \to \mathbb{R}_{>0}$ called the* scale map*,*
- *a cost function $J_C : C \to \mathbb{R}_{\geq 0}$ satisfying $J_C(c) = J(\iota_C(c))$ for all $c \in C$.*

Equivalently, once ι_C is fixed, J_C is determined by J; we retain J_C in the notation since later statements compare configuration costs and object costs directly.

Notation 1. *We write $\mathcal{S} = (S, J_S, \iota_S)$ for a configuration (token) costed space and $\mathcal{O} = (O, J_O, \iota_O)$ for an object costed space.*

Throughout, we identify $\mathbb{R}_{>0}$ with $(0, \infty)$ and equip $\mathbb{R}_{>0}$ (and $(\mathbb{R}_{>0})^d$) with the usual Euclidean topology on $(0, \infty)$ (equivalently, the Euclidean subspace topology inherited from $\mathbb{R}$). Accordingly, when we say that a set $Y \subset \mathbb{R}_{>0}$ is closed, *we mean* closed in the usual topology on $(0, \infty)$ *(equivalently, $Y = (0, \infty) \cap F$ for some closed $F \subset \mathbb{R}$). Likewise, for $Y \subset (\mathbb{R}_{>0})^d$ the term* closed *means* closed in the usual topology on $(0, \infty)^d$.

Example 2 (Ratio space). *The canonical example is $C = \mathbb{R}_{>0}$ with $\iota_C = \mathrm{id}$ and $J_C = J$.*

The next example isolates a small neighborhood of the balanced point; it will serve as a convenient test class for stability statements.

Example 3 (Near-balanced configurations). *For $\epsilon > 0$ let $C_\epsilon := \{x \in \mathbb{R}_{>0} : |x - 1| < \epsilon\}$. Then every $c \in C_\epsilon$ satisfies $J_C(c) = J(c) < J(1 + \epsilon)$.*

3.2. Reference structures

Definition 4 (Reference structure)**.** *A reference structure from $\mathcal{S}$ to $\mathcal{O}$ is a function*

$$c_{\mathcal{R}} : S \times O \to \mathbb{R}_{\geq 0}, \tag{4}$$

called the reference cost. *It assigns to each pair (s, o) the cost of using s to refer to o.*

In the remainder of the paper we focus on the ratio-induced costs generated by J and the scale maps.

Definition 5 (Ratio-induced reference)**.** *Given scale maps ι_S and ι_O, the* ratio-induced *reference structure is defined by*

$$c^J_{\mathcal{R}}(s, o) := J\Big(\frac{\iota_S(s)}{\iota_O(o)}\Big). \tag{5}$$

This is the cost used in the Introduction (Eq. 1).

Link to comparative recognizers.

The ratio-induced reference cost (5) can be viewed as a specific instantiation of a *comparative recognizer* in the sense of Recognition Geometry [17, Axiom 5 (RG4)]. In that framework, a comparative recognizer maps *pairs* of configurations to an *event space* [17, Axiom 2 (RG1)] so as to induce comparative structure (order/distance) from observable events. Here the "event" is the scalar mismatch value $J(\iota_S(s)/\iota_O(o))$, and the induced *indistinguishability* relation [17, Def. 4] corresponds to the zero-cost condition $J(\iota_S(s)/\iota_O(o)) = 0$, which forces exact scale match $\iota_S(s) = \iota_O(o)$ by Lemma 2.
The following admissibility condition specifies when the reference cost is exactly the canonical ratio penalty.

Definition 6 (Admissible reference structure)**.** *A reference structure $\mathcal{R}$ from $\mathcal{S}$ to $\mathcal{O}$ is called* admissible *(with respect to J and the scale maps ι_S, ι_O) if it is ratio-induced, i.e.*

$$c_{\mathcal{R}}(s, o) = J\Big(\frac{\iota_S(s)}{\iota_O(o)}\Big) \qquad \forall (s, o) \in S \times O. \tag{6}$$

Unless stated otherwise, we work with admissible reference structures.

Admissibility transfers the symmetry properties of Jto the reference cost; we record this for later use.

Proposition 3 (Inversion symmetry of the reference cost)**.** *If $\mathcal{R}$ is admissible, then for all $(s, o) \in S \times O$ one has*

$$c_{\mathcal{R}}(s, o) = J\Big(\frac{\iota_S(s)}{\iota_O(o)}\Big) = J\Big(\frac{\iota_O(o)}{\iota_S(s)}\Big).$$

Proof. Immediate from admissibility and inversion symmetry $J(x) = J(x^{-1})$ (Lemma 1). □

3.3. Meaning and the symbol predicate

Definition 7 (Meaning)**.** *Let $\mathcal{R}$ be a reference structure from $\mathcal{S}$ to $\mathcal{O}$. A configuration $s \in S$* means *an object $o \in O$, written* $\mathrm{Mean}_{\mathcal{R}}(s, o)$, *if o minimizes the reference cost among all objects:*

$$\mathrm{Mean}_{\mathcal{R}}(s, o) \iff \forall o' \in O, \quad c_{\mathcal{R}}(s, o) \leq c_{\mathcal{R}}(s, o'). \tag{7}$$

For each $s \in S$ we write

$$\mathrm{Mean}_{\mathcal{R}}(s) := \{o \in O : \mathrm{Mean}_{\mathcal{R}}(s,o)\}$$

for the (possibly multi-valued) meaning set. If $\mathcal{R}$ is admissible, then equivalently

$$\mathrm{Mean}_{\mathcal{R}}(s) = \arg\min_{o \in O} J\Big(\frac{\iota_S(s)}{\iota_O(o)}\Big).$$

Definition 8 (Symbol)**.** *Let $\mathcal{R}$ be a reference structure from $\mathcal{S}$ to $\mathcal{O}$. A configuration $s \in S$ is a* symbol *for an object $o \in O$ (relative to $\mathcal{R}$) if:*

1. ***Reference:*** $\mathrm{Mean}_{\mathcal{R}}(s,o)$.
2. ***Compression:*** $J_S(s) < J_O(o)$.

The compression requirement is a modeling assumption: it enforces that symbols are lower-cost encodings than their referents in the common currency induced by J. No empirical interpretation is asserted; the condition is simply part of the definition used in later results.

4. Main theorems

This section collects the main mathematical consequences of the ratio-induced reference model. Throughout we fix the explicit mismatch functional

$$J(x) = \frac{(x-1)^2}{2x} = \tfrac{1}{2}(x + x^{-1}) - 1 \qquad (x > 0) \tag{8}$$

which satisfies Definition 1, and we assume the reference structure is *admissible*:

$$c_{\mathcal{R}}(s,o) = J\left(\frac{\iota_S(s)}{\iota_O(o)}\right). \tag{9}$$

Thus, for each $s \in S$, the meaning set $\mathrm{Mean}_{\mathcal{R}}(s)$ is the set of minimizers of $o \mapsto J(\iota_S(s)/\iota_O(o))$.

4.1. Sublevel geometry of the explicit mismatch cost

Lemma 3 (Sublevel intervals)**.** *Assume J is given by (3) (equivalently (8)). For each $\epsilon > 0$, the sublevel set*

$$L_\epsilon := \{x \in \mathbb{R}_{>0} : \; J(x) \le \epsilon\}$$

coincides with the closed interval $[a_\epsilon, b_\epsilon]$, where

$$b_\epsilon := (1+\epsilon) + \sqrt{\epsilon(2+\epsilon)}, \qquad a_\epsilon := (1+\epsilon) - \sqrt{\epsilon(2+\epsilon)} = \frac{1}{b_\epsilon}.$$

Proof. Using $J(x) = \frac{(x-1)^2}{2x}$, the inequality $J(x) \le \epsilon$ is equivalent (after multiplying by $2x > 0$) to

$$(x-1)^2 \le 2\epsilon x \iff x^2 - 2(1+\epsilon)x + 1 \le 0.$$

The quadratic has discriminant $\Delta = 4\epsilon(2+\epsilon)$ and roots $x_\pm = (1+\epsilon) \pm \sqrt{\epsilon(2+\epsilon)}$. Since it opens upward, the inequality holds exactly for $x \in [x_-, x_+]$. Set $a_\epsilon := x_-$ and $b_\epsilon := x_+$. Then $a_\epsilon b_\epsilon = (1+\epsilon)^2 - \epsilon(2+\epsilon) = 1$, so $a_\epsilon = 1/b_\epsilon$. □

4.2. Meaning constraints from a balanced baseline

Theorem 1 (Scale window for meanings of low-cost configurations)**.** *Assume* $1 \in Y := \iota_O(O)$ *and choose* $o_0 \in O$ *with* $\iota_O(o_0) = 1$. *Let* $s \in S$ *and let* $o \in \mathrm{Mean}_{\mathcal{R}}(s)$. *Then*

$$c_{\mathcal{R}}(s, o) \le c_{\mathcal{R}}(s, o_0) = J(\iota_S(s)) = J_S(s). \tag{10}$$

In particular, for every $\epsilon > 0$, *if* $J_S(s) \le \epsilon$ *then*

$$\frac{\iota_S(s)}{\iota_O(o)} \in [a_\epsilon, b_\epsilon] \tag{11}$$

and hence

$$\frac{\iota_S(s)}{b_\epsilon} \le \iota_O(o) \le \frac{\iota_S(s)}{a_\epsilon}, \tag{12}$$

where $[a_\epsilon, b_\epsilon]$ *is as in Lemma 3.*

Proof. Since $o \in \mathrm{Mean}_{\mathcal{R}}(s)$, by definition $c_{\mathcal{R}}(s, o) \le c_{\mathcal{R}}(s, o_0)$. By admissibility (9) and $\iota_O(o_0) = 1$, $c_{\mathcal{R}}(s, o_0) = J(\iota_S(s)) = J_S(s)$, which gives (10). If $J_S(s) \le \epsilon$, then (10) implies $J(\iota_S(s)/\iota_O(o)) \le \epsilon$, hence (11) by Lemma 3. Rearranging yields (12). □

Corollary 1 (Near-balanced configurations force near-balanced meanings)**.** *Under the hypotheses of Theorem 1, if* $J_S(s) \le \epsilon$ *and* $o \in \mathrm{Mean}_{\mathcal{R}}(s)$, *then*

$$\iota_O(o) \in \left[\frac{1}{b_\epsilon^2}, b_\epsilon^2\right].$$

In particular, as $\epsilon \downarrow 0$, *any meaning of an* ϵ*-cheap symbol must satisfy* $\iota_O(o) \to 1$.

Proof. From $J_S(s) = J(\iota_S(s)) \le \epsilon$ and Lemma 3 we have $\iota_S(s) \in [a_\epsilon, b_\epsilon]$. Combining this with (12) and $a_\epsilon = 1/b_\epsilon$ gives the stated bounds. □

4.3. Existence of meanings under attainment hypotheses

Lemma 4 (Coercivity of J)**.** *Assume* J *is given by (3). Then* $J(x) \to \infty$ *as* $x \to 0^+$ *and as* $x \to \infty$. *In particular, for each* $M \ge 0$ *the sublevel set* $\{x \in \mathbb{R}_{>0} : J(x) \le M\}$ *is compact in* $\mathbb{R}$.

Proof. From (8), $J(x) = \frac{1}{2}(x + x^{-1}) - 1$. As $x \to \infty$ the term $\frac{1}{2}x$ dominates, and as $x \to 0^+$ the term $\frac{1}{2}x^{-1}$ dominates, so in both limits $J(x) \to \infty$. If $J(x) \le M$ then $x + x^{-1} \le 2(M+1)$, hence both x and x^{-1} are bounded; the sublevel set is therefore closed and bounded away from 0 and ∞, hence compact. □

Theorem 2 (Existence of meanings for ratio-induced reference)**.** *Assume* $\mathcal{R}$ *is admissible and that* J *is given by (3). Let* $Y := \iota_O(O) \subset \mathbb{R}_{>0}$ *be nonempty and closed in the usual topology on* $(0, \infty)$. *Then for every* $s \in S$ *there exists* $o \in O$ *such that* $\mathrm{Mean}_{\mathcal{R}}(s, o)$ *(equivalently,* $\mathrm{Mean}_{\mathcal{R}}(s) \neq \emptyset$*). Moreover, if* $x := \iota_S(s) \in Y$, *then any* $o \in O$ *with* $\iota_O(o) = x$ *is a meaning and satisfies* $c_{\mathcal{R}}(s, o) = 0$.

Proof. Fix s and set $x := \iota_S(s)$. Consider $f : Y \to \mathbb{R}_{\ge 0}$ defined by $f(y) := J(x/y)$. The map f is continuous. By Lemma 4, $f(y) \to \infty$ as $y \to 0^+$ or $y \to \infty$, so the infimum of f over Y is achieved on a compact sublevel set. Concretely, choose a minimizing sequence $y_n \in Y$ with $f(y_n) \downarrow \inf_Y f$. Coercivity implies (y_n) is bounded away from 0 and ∞, hence has a convergent subsequence; since Y is closed in $(0, \infty)$, the limit $y_* \in Y$, and continuity gives $f(y_*) = \inf_Y f$. Choose $o \in O$ with $\iota_O(o) = y_*$. Then $c_{\mathcal{R}}(s, o) = f(y_*) \le f(\iota_O(o')) =$

$c_{\mathcal{R}}(s, o')$ for all $o' \in O$, i.e. $\mathrm{Mean}_{\mathcal{R}}(s, o)$. If $x \in Y$, take $y_* = x$; then $J(x/x) = J(1) = 0$, so any o with $\iota_O(o) = x$ is a meaning with zero reference cost. □

Remark 1. *If $Y = \iota_O(O)$ is not closed in $(0, \infty)$, the minimum need not be attained; in that case $\mathrm{Mean}_{\mathcal{R}}(s)$ may be empty even though the infimum exists.*

4.4. A simple total-cost benchmark

Theorem 3 (Balanced reference minimizes the intrinsic+reference sum)**.** *Assume admissible reference (9) and intrinsic costs $J_S(s) = J(\iota_S(s))$, $J_O(o) = J(\iota_O(o))$. Define*

$$C(s, o) := J_S(s) + J_O(o) + c_{\mathcal{R}}(s, o).$$

Then $C(s, o) \geq 0$ for all $(s, o) \in S \times O$, and

$$C(s, o) = 0 \iff \iota_S(s) = 1 \text{ and } \iota_O(o) = 1.$$

In particular, if there exist $s_0 \in S$ and $o_0 \in O$ with $\iota_S(s_0) = \iota_O(o_0) = 1$, then (s_0, o_0) is a global minimizer of C over $S \times O$.

Proof. Each term in C is nonnegative, hence $C \geq 0$. If $C(s, o) = 0$, then all three terms vanish; by Lemma 2 this forces $\iota_S(s) = \iota_O(o) = 1$. The converse is immediate from $J(1) = 0$. □

4.5. A backbone window for near-balanced configuration classes

Definition 9 (Referential capacity)**.** *Given a reference structure $\mathcal{R}$ from $\mathcal{S}$ to $\mathcal{O}$, define the* referential capacity *to be*

$$\mathrm{Cap}(\mathcal{S}, \mathcal{O}; \mathcal{R}) := \left|\{o \in O : \exists s \in S \text{ with } o \in \mathrm{Mean}_{\mathcal{R}}(s)\}\right|.$$

(If O is infinite, this cardinality may be infinite.)

We now show that restricting to near-balanced configurations forces all attainable meanings to lie in an explicit scale window.

Theorem 4 (Backbone window for near-balanced configurations)**.** *Let $\mathcal{S}_\delta = (S_\delta, J_\delta, \iota_\delta)$ be the near-balanced ratio space*

$$S_\delta := \{x \in \mathbb{R}_{>0} : |x - 1| < \delta\}, \qquad \iota_\delta = \mathrm{id}, \qquad J_\delta = J|_{S_\delta}.$$

Let $\mathcal{O} = (O, J_O, \iota_O)$ be a costed space such that $Y := \iota_O(O) \subset \mathbb{R}_{>0}$ is nonempty, closed in the usual topology on $(0, \infty)$, and contains 1. Assume $\mathcal{R}$ is admissible and J is given by (3).

Set $\epsilon_\delta := J(1 + \delta)$ and let $[a_{\epsilon_\delta}, b_{\epsilon_\delta}]$ be as in Lemma 3. Define the window

$$I_\delta := \left[\frac{1-\delta}{b_{\epsilon_\delta}}, \frac{1+\delta}{a_{\epsilon_\delta}}\right].$$

Then:

1. *For every $s \in S_\delta$ the meaning set $\mathrm{Mean}_{\mathcal{R}}(s)$ is nonempty.*
2. *If $s \in S_\delta$ and $o \in \mathrm{Mean}_{\mathcal{R}}(s)$, then $\iota_O(o) \in I_\delta$. Equivalently, if $\iota_O(o) \notin I_\delta$, then no $s \in S_\delta$ can mean o under admissible reference.*

In particular,

$$\mathrm{Cap}(\mathcal{S}_\delta, \mathcal{O}; \mathcal{R}) \leq \left|\{o \in O : \iota_O(o) \in I_\delta\}\right|.$$

Proof. (1) is a direct application of Theorem 2 to the closed (in $(0,\infty)$) nonempty set Y.

For (2), fix $s \in S_\delta$ and write $x := \iota_\delta(s) \in (1-\delta, 1+\delta)$. Let $o \in \mathrm{Mean}_{\mathcal{R}}(s)$ and choose $o_0 \in O$ with $\iota_O(o_0) = 1$ (possible since $1 \in Y$). By Theorem 1,

$$J\left(\frac{x}{\iota_O(o)}\right) = c_{\mathcal{R}}(s,o) \leq c_{\mathcal{R}}(s,o_0) = J(x) \leq J(1+\delta) = \epsilon_\delta.$$

Applying Lemma 3 gives $x/\iota_O(o) \in [a_{\epsilon_\delta}, b_{\epsilon_\delta}]$, hence

$$\frac{x}{b_{\epsilon_\delta}} \leq \iota_O(o) \leq \frac{x}{a_{\epsilon_\delta}}.$$

Using $x \in [1-\delta, 1+\delta]$ yields $\iota_O(o) \in I_\delta$.

For the capacity bound, any object counted in $\mathrm{Cap}(\mathcal{S}_\delta, \mathcal{O}; \mathcal{R})$ lies in $\mathrm{Mean}_{\mathcal{R}}(s)$ for some $s \in S_\delta$, hence satisfies $\iota_O(o) \in I_\delta$ by (2). □

5. Compositionality

This section records two elementary composition mechanisms for reference costs: *(i)* product composition (independent coordinates) and *(ii)* sequential mediation through an intermediate space. Both are purely variational constructions: they introduce no semantic primitive beyond the cost function(s).

5.1. Product reference and coordinatewise meaning

Definition 10 (Product reference)**.** *Let $\mathcal{R}_1$ be a reference structure from a configuration (token) set S_1 to an object set O_1, and let $\mathcal{R}_2$ be a reference structure from a configuration (token) set S_2 to an object set O_2. Write their costs as $c_{\mathcal{R}_i}$. The* product reference structure $\mathcal{R}_1 \otimes \mathcal{R}_2 : S_1 \times S_2 \to O_1 \times O_2$ *is defined by*

$$c_{\mathcal{R}_1 \otimes \mathcal{R}_2}\big((s_1,s_2),(o_1,o_2)\big) := c_{\mathcal{R}_1}(s_1,o_1) + c_{\mathcal{R}_2}(s_2,o_2). \tag{13}$$

With the product cost in hand, meaning decomposes coordinatewise; the next theorem makes this precise.

Theorem 5 (Compositionality of product meaning)**.** *For any reference structures $\mathcal{R}_1, \mathcal{R}_2$ and their product $\mathcal{R}_1 \otimes \mathcal{R}_2$, and for every $(s_1,s_2) \in S_1 \times S_2$, the meaning set in the product structure factorizes as the Cartesian product*

$$\mathrm{Mean}_{\mathcal{R}_1 \otimes \mathcal{R}_2}(s_1,s_2) = \mathrm{Mean}_{\mathcal{R}_1}(s_1) \times \mathrm{Mean}_{\mathcal{R}_2}(s_2).$$

Equivalently, viewing meaning as a relation $\mathrm{Mean}_{\mathcal{R}_i} \subseteq S_i \times O_i$, one has equality of relations inside $(S_1 \times S_2) \times (O_1 \times O_2)$:

$$\mathrm{Mean}_{\mathcal{R}_1 \otimes \mathcal{R}_2} = \mathrm{Mean}_{\mathcal{R}_1} \times \mathrm{Mean}_{\mathcal{R}_2},$$

where the right-hand side denotes the Cartesian product relation.

Proof. Fix $(s_1,s_2) \in S_1 \times S_2$ and write

$$A := \mathrm{Mean}_{\mathcal{R}_1 \otimes \mathcal{R}_2}(s_1,s_2) \subseteq O_1 \times O_2, \qquad A_i := \mathrm{Mean}_{\mathcal{R}_i}(s_i) \subseteq O_i (i = 1,2).$$

By definition of the product reference structure, for every $(o_1', o_2') \in O_1 \times O_2$,

$$c_{\mathcal{R}_1 \otimes \mathcal{R}_2}\big((s_1,s_2),(o_1',o_2')\big) = c_{\mathcal{R}_1}(s_1,o_1') + c_{\mathcal{R}_2}(s_2,o_2').$$

Inclusion $A \subseteq A_1 \times A_2$. Let $(o_1, o_2) \in A$. Then for all $(o_1', o_2') \in O_1 \times O_2$,

$$c_{\mathcal{R}_1}(s_1, o_1) + c_{\mathcal{R}_2}(s_2, o_2) \le c_{\mathcal{R}_1}(s_1, o_1') + c_{\mathcal{R}_2}(s_2, o_2').$$

Specializing to $o_2' = o_2$ gives, for all $o_1' \in O_1$,

$$c_{\mathcal{R}_1}(s_1, o_1) \le c_{\mathcal{R}_1}(s_1, o_1'),$$

so $o_1 \in A_1$. Similarly, specializing to $o_1' = o_1$ gives $o_2 \in A_2$. Hence $(o_1, o_2) \in A_1 \times A_2$.
Inclusion $A_1 \times A_2 \subseteq A$. Let $o_1 \in A_1$ and $o_2 \in A_2$. Then for all $o_1' \in O_1$ and all $o_2' \in O_2$,

$$c_{\mathcal{R}_1}(s_1, o_1) \le c_{\mathcal{R}_1}(s_1, o_1'), \qquad c_{\mathcal{R}_2}(s_2, o_2) \le c_{\mathcal{R}_2}(s_2, o_2').$$

Adding yields, for all $(o_1', o_2') \in O_1 \times O_2$,

$$c_{\mathcal{R}_1}(s_1, o_1) + c_{\mathcal{R}_2}(s_2, o_2) \le c_{\mathcal{R}_1}(s_1, o_1') + c_{\mathcal{R}_2}(s_2, o_2'),$$

which is exactly the defining inequality for $(o_1, o_2) \in A$ in the product structure. Thus $A_1 \times A_2 \subseteq A$. Combining the two inclusions gives $A = A_1 \times A_2$, i.e., $\mathrm{Mean}_{\mathcal{R}_1 \otimes \mathcal{R}_2}(s_1, s_2) = \mathrm{Mean}_{\mathcal{R}_1}(s_1) \times \mathrm{Mean}_{\mathcal{R}_2}(s_2)$. □

Corollary 2 (Existence of product meanings under the explicit mismatch cost)**.** *Assume the explicit mismatch cost* (8) *and admissible reference on each component. If, for $i = 1, 2$, the object ratio set $Y_{O_i} := \iota_{O_i}(O_i) \subset \mathbb{R}_{>0}$ is nonempty and closed in the usual topology on $(0, \infty)$, then for every $(s_1, s_2) \in S_1 \times S_2$ the product meaning set* $\mathrm{Mean}_{\mathcal{R}_1 \otimes \mathcal{R}_2}(s_1, s_2)$ *is nonempty.*

Proof. Under the stated hypotheses, Theorem 2 implies $\mathrm{Mean}_{\mathcal{R}_i}(s_i) \neq \emptyset$ for each i. Pick $o_i \in \mathrm{Mean}_{\mathcal{R}_i}(s_i)$. Then Theorem 5 yields $(o_1, o_2) \in \mathrm{Mean}_{\mathcal{R}_1 \otimes \mathcal{R}_2}(s_1, s_2)$. □

5.2. Sequential mediation

Definition 11 (Sequential reference)**.** *Let $\mathcal{R}_1 : \mathcal{S} \to \mathcal{M}$ and $\mathcal{R}_2 : \mathcal{M} \to \mathcal{O}$ be reference structures. Their* sequential composition *$\mathcal{R}_2 \circ \mathcal{R}_1 : \mathcal{S} \to \mathcal{O}$ is defined by the infimal convolution*

$$c_{\mathcal{R}_2 \circ \mathcal{R}_1}(s, o) = \inf_{m \in \mathcal{M}} \left[c_{\mathcal{R}_1}(s, m) + c_{\mathcal{R}_2}(m, o) \right]. \tag{14}$$

A mediator *m is* optimal *for (s, o) if it attains the infimum in* (14)*.*

We next compute the optimal mediator explicitly under the canonical mismatch cost.

Theorem 6 (Geometric-mean mediator for the explicit mismatch cost)**.** *Assume the explicit mismatch functional* (8) *and admissible reference for $\mathcal{R}_1 : \mathcal{S} \to \mathcal{M}$ and $\mathcal{R}_2 : \mathcal{M} \to \mathcal{O}$ with scale maps $\iota_S, \iota_M, \iota_O$. Fix $s \in S$ and $o \in O$ and set $a := \iota_S(s)$ and $c := \iota_O(o)$. Let $Y_M := \iota_M(\mathcal{M}) \subset \mathbb{R}_{>0}$. Assume that Y_M is nonempty and closed in the usual topology on $(0, \infty)$. Set $b_{\mathrm{geo}} := \sqrt{ac}$ and $U := \{\log b : b \in Y_M\} \subset \mathbb{R}$. Then the infimum in* (14) *is attained by at least one mediator $m_* \in \mathcal{M}$. Moreover, a mediator $m \in \mathcal{M}$ with $b := \iota_M(m)$ is optimal if and only if $\log b$ minimizes $|\log b - \log b_{\mathrm{geo}}|$ over U (equivalently, b minimizes $|\log(b/b_{\mathrm{geo}})|$ over Y_M). If $b_{\mathrm{geo}} \notin Y_M$, write $u_0 := \log b_{\mathrm{geo}}$ and let $\delta := \mathrm{dist}(u_0, U)$, where $U = \{\log b : b \in Y_M\} \subset \mathbb{R}$. Let $u_* \in U$ be a* *closest point to u_0 (so $|u_* - u_0| = \delta$) and set $b_* := e^{u_*}$. Then, for the explicit cost, the constrained optimum value admits the closed form*

$$c_{\mathcal{R}_2\circ\mathcal{R}_1}(s,o) = \Big(\cosh\big(\tfrac{1}{2}\log(a/c)+\delta\big)-1\Big) + \Big(\cosh\big(\tfrac{1}{2}\log(a/c)-\delta\big)-1\Big) \\ = 2\cosh\Big(\tfrac{1}{2}\log\frac{a}{c}\Big)\cosh(\delta)-2.$$

In particular, the suboptimality gap relative to the unconstrained geometric mean (i.e., relative to $\delta = 0$) is

$$c_{\mathcal{R}_2\circ\mathcal{R}_1}(s,o) - 2\,J\Big(\sqrt{\tfrac{a}{c}}\Big) = 2\cosh\Big(\tfrac{1}{2}\log\frac{a}{c}\Big)\,(\cosh(\delta)-1) \geq 0.$$

In particular, if $b_{\mathrm{geo}} \in Y_M$, then the optimal mediator ratio is unique and equals b_{geo}; in that case, choosing $m_* \in \mathcal{M}$ with $\iota_M(m_*) = b_{\mathrm{geo}}$ gives

$$c_{\mathcal{R}_2\circ\mathcal{R}_1}(s,o) = J\Big(\frac{a}{b_{\mathrm{geo}}}\Big) + J\Big(\frac{b_{\mathrm{geo}}}{c}\Big) = 2\,J\Big(\sqrt{\frac{a}{c}}\Big).$$

Proof. Under admissibility, the objective in (14) depends on m only through $b := \iota_M(m) \in Y_M$, namely

$$F(b) := J\Big(\frac{a}{b}\Big) + J\Big(\frac{b}{c}\Big).$$

For the explicit penalty (3), one has $J(x) = \cosh(\log x) - 1$. Writing $t := \log a$, $s := \log c$, and $u := \log b$, we obtain

$$F(b) = \big(\cosh(t-u)-1\big) + \big(\cosh(u-s)-1\big) = \cosh(t-u) + \cosh(u-s) - 2.$$

Using $\cosh(\alpha) + \cosh(\beta) = 2\cosh\big(\frac{\alpha+\beta}{2}\big)\cosh\big(\frac{\alpha-\beta}{2}\big)$ with $\alpha = t-u$ and $\beta = u-s$ gives

$$F(b) = 2\,\cosh\Big(\frac{t-s}{2}\Big)\,\cosh\Big(u - \frac{t+s}{2}\Big) - 2 = 2\,\cosh\Big(\frac{\log(a/c)}{2}\Big)\,\cosh\big(u - \log b_{\mathrm{geo}}\big) - 2.$$

Since $\cosh\big(\frac{\log(a/c)}{2}\big) > 0$ is constant and cosh is even and strictly increasing on $[0,\infty)$, minimizing F over $b \in Y_M$ is equivalent to minimizing $|u - \log b_{\mathrm{geo}}|$ over $u \in U = \log Y_M$. Because $\log : (0,\infty) \to \mathbb{R}$ is a homeomorphism and Y_M is closed and nonempty, the set U is closed and nonempty in $\mathbb{R}$, hence the distance function $u \mapsto |u - \log b_{\mathrm{geo}}|$ attains its minimum on U. This proves existence of at least one minimizer $u_* \in U$, and the stated characterization of optimal ratios. If $b_{\mathrm{geo}} \in Y_M$ (equivalently $\log b_{\mathrm{geo}} \in U$), then the unique minimizer of $u \mapsto |u - \log b_{\mathrm{geo}}|$ on U is $u = \log b_{\mathrm{geo}}$, hence the optimal mediator ratio is unique and equals b_{geo}. Substituting $b_{\mathrm{geo}} = \sqrt{ac}$ yields $J(a/b_{\mathrm{geo}}) = J(\sqrt{a/c}) = J(b_{\mathrm{geo}}/c)$ and the displayed formula. □

Corollary 3 (Mediation can strictly reduce mismatch)**.** *For every $x > 0$ one has*

$$2\,J(\sqrt{x}) \leq J(x),$$

with equality if and only if $x = 1$. Consequently, in the setting of Theorem 6, if $b_{\mathrm{geo}} \in Y_M$ and a direct admissible reference $\mathcal{R} : \mathcal{S} \to \mathcal{O}$ is available (built from the same J and scale maps), then

$$c_{\mathcal{R}_2\circ\mathcal{R}_1}(s,o) \leq c_{\mathcal{R}}(s,o),$$

with equality if and only if $\iota_S(s) = \iota_O(o)$.

Proof. Let $t := \sqrt{x} > 0$. Using (3), a direct calculation gives

$$J(t^2) - 2\,J(t) = \frac{1}{2}\Big((t-1)^2 + \big(t^{-1}-1\big)^2\Big) \geq 0,$$

with equality if and only if $t = 1$, i.e. $x = 1$. If $b_{\text{geo}} \in Y_M$, Theorem 6 gives $c_{\mathcal{R}_2 \circ \mathcal{R}_1}(s,o) = 2\,J(\sqrt{x})$ with $x = \iota_S(s)/\iota_O(o)$; comparing with $c_{\mathcal{R}}(s,o) = J(x)$ yields the stated inequality. □

6. Extensions: multi-dimensional scales and robustness

The core framework above uses a single positive scale coordinate $\iota(\cdot) \in \mathbb{R}_{>0}$. In some applications one may want a finite list of independent scale coordinates (for instance, a configuration might carry multiple features, each measured in the same "cost currency" through J). This section records a minimal extension of the model to d coordinates and a simple robustness lemma for finite dictionaries.

6.1. Multi-dimensional costed spaces

Definition 12 (Multi-dimensional costed space). *Let $d \in \mathbb{N}$. A d-dimensional costed space is a triple (C, J_C, ι_C) where*

- *C is a set,*
- *$\iota_C : C \to (\mathbb{R}_{>0})^d$ is a scale map, and*
- *$J_C : C \to \mathbb{R}_{\geq 0}$ is the induced (separable) cost*

$$J_C(c) := \sum_{i=1}^{d} J\big(\iota_C(c)_i\big), \qquad c \in C.$$

We extend admissible reference by taking a separable, coordinatewise ratio penalty.

Definition 13 (Multi-dimensional admissible reference). *Let (S, J_S, ι_S) and (O, J_O, ι_O) be d-dimensional costed spaces. A reference structure $\mathcal{R}$ from S to O is* multi-dimensionally admissible *if its reference cost is the coordinatewise ratio cost*

$$c_{\mathcal{R}}(s,o) = \sum_{i=1}^{d} J\bigg(\frac{\iota_S(s)_i}{\iota_O(o)_i}\bigg), \qquad (s,o) \in S \times O. \tag{15}$$

The separable form immediately implies that meanings factor coordinatewise.

Corollary 4 (Coordinatewise meaning for product models). *Assume $S = \prod_{i=1}^{d} S_i$ and $O = \prod_{i=1}^{d} O_i$ and that the scale maps factor coordinatewise: $\iota_S(s)_i = \iota_{S_i}(s_i)$ and $\iota_O(o)_i = \iota_{O_i}(o_i)$. If $\mathcal{R}$ is multi-dimensionally admissible, then*

$$(o_1, \ldots, o_d) \in \mathrm{Mean}_{\mathcal{R}}(s_1, \ldots, s_d) \quad \Longleftrightarrow \quad \forall i,\ o_i \in \mathrm{Mean}_{\mathcal{R}_i}(s_i),$$

where $\mathcal{R}_i$ denotes the induced one-dimensional admissible reference on (S_i, O_i).

Proof. By (15) the cost is a separable sum of d nonnegative terms, each depending only on (s_i, o_i). Thus minimizing over $O = \prod_i O_i$ is equivalent to minimizing each summand over its coordinate; this is the same argument as in Theorem 5. □

6.2. Log-space geometry for the explicit mismatch cost

In this subsection we specialize to the explicit mismatch functional

$$J(x) = \tfrac{1}{2}(x + x^{-1}) - 1 \qquad (x > 0), \tag{16}$$

already used in Sections 2–5.

Lemma 5 (Log-coordinate form)**.** *For all $t \in \mathbb{R}$ one has $J(e^t) = \cosh(t) - 1$.*

Proof. Immediate from (16): $J(e^t) = \frac{1}{2}(e^t + e^{-t}) - 1 = \cosh(t) - 1$. □

Proposition 4 (Quadratic regime with explicit remainder)**.** *For all $t \in \mathbb{R}$,*

$$0 \le J(e^t) - \frac{t^2}{2} \le \frac{t^4}{24}\cosh(|t|).$$

In particular, for $|t| \le 1$,

$$\frac{t^2}{2} \le J(e^t) \le \frac{t^2}{2} + \frac{\cosh(1)}{24}\, t^4.$$

Proof. By Lemma 5 it suffices to estimate $\cosh(t) - 1 - \frac{1}{2}t^2$. Taylor's theorem at 0 with remainder gives

$$\cosh(t) = 1 + \frac{t^2}{2} + \frac{t^4}{24}\cosh(\xi)$$

for some ξ between 0 and t. Since cosh is even and increasing on $\mathbb{R}_{\ge 0}$, one has $\cosh(\xi) \le \cosh(|t|)$, yielding the upper bound. Nonnegativity follows since $\cosh(\xi) > 0$. □

Corollary 5 (Local Euclidean geometry in log-ratio)**.** *For the explicit mismatch cost (16), set $x := \iota_S(s)$ and $y := \iota_O(o)$. If $|\log(x/y)| \le 1$, then*

$$\frac{1}{2}\big(\log(x/y)\big)^2 \le c_{\mathcal{R}}(s,o) \le \frac{1}{2}\big(\log(x/y)\big)^2 + \frac{\cosh(1)}{24}\big(\log(x/y)\big)^4.$$

Thus, in the small-mismatch regime, meanings behave like nearest neighbors in the log-ratio metric.

Proof. For admissible reference, $c_{\mathcal{R}}(s,o) = J(x/y)$ with $x := \iota_S(s)$ and $y := \iota_O(o)$. Write $t := \log(x/y)$. Then $x/y = e^t$ and $|t| \le 1$ by hypothesis. Apply Proposition 4 to obtain $\frac{1}{2}t^2 \le J(e^t) \le \frac{1}{2}t^2 + \frac{\cosh(1)}{24}t^4$, and substitute $t = \log(x/y)$. □

6.3. Margin stability for finite dictionaries

Definition 14 (Decision margin)**.** *Fix a configuration $s \in S$ and a finite object dictionary $O = \{o_1, \dots, o_N\}$. Write $C_k := c_{\mathcal{R}}(s, o_k)$ and let $M := \min_{1 \le k \le N} C_k$. The* decision margin *at s is*

$$\Delta(s) := \min\{C_k - M :\ 1 \le k \le N,\ C_k > M\} \in [0, \infty],$$

with the convention $\Delta(s) = \infty$ if all C_k are equal.

The margin parameter controls how stable the argmin is under perturbations of the cost values.

Proposition 5 (Robustness under bounded perturbations)**.** *In the setting of Definition 14, suppose the costs C_k are perturbed to numbers $\widetilde{C}_k$ satisfying*

$$\max_{1 \le k \le N} |\widetilde{C}_k - C_k| \le \eta.$$

If $\Delta(s) > 2\eta$, then the set of minimizers is unchanged:

$$\{k : C_k = \min_j C_j\} = \{k : \widetilde{C}_k = \min_j \widetilde{C}_j\}.$$

Proof. Let $I := \{k : C_k = M\}$ be the (nonempty) set of original minimizers. For $k \in I$ one has $\widetilde{C}_k \leq M + \eta$. If $k \notin I$, then $C_k \geq M + \Delta(s)$ by definition of $\Delta(s)$, hence $\widetilde{C}_k \geq M + \Delta(s) - \eta$. If $\Delta(s) > 2\eta$ then $M + \Delta(s) - \eta > M + \eta$, so every perturbed minimizer must lie in I and conversely every $k \in I$ remains minimal. □

6.4. Existence (and optional uniqueness) in d dimensions

Here we discuss the multi-dimensional analogue of Theorem 2; it follows by the same attainment argument under the multi-dimensional admissibility and closedness hypotheses.

Corollary 6. *Let $d \in \mathbb{N}$ and let (S, J_S, ι_S) and (O, J_O, ι_O) be d-dimensional costed spaces. Assume $\mathcal{R}$ is multi-dimensionally admissible in the sense of Definition 13. Let $Y := \iota_O(O) \subset (\mathbb{R}_{>0})^d$ be nonempty and closed in the usual topology on $(0, \infty)^d$. Then for every $s \in S$ the meaning set $\mathrm{Mean}_{\mathcal{R}}(s)$ is nonempty. Moreover, if $x := \iota_S(s)$ lies in Y, then any $o \in O$ with $\iota_O(o) = x$ is a meaning and satisfies $c_{\mathcal{R}}(s, o) = 0$.*

Proof. Fix $s \in S$ and write $x := \iota_S(s) \in (\mathbb{R}_{>0})^d$. Consider the continuous objective on Y,

$$F_x(y) := \sum_{i=1}^{d} J\left(\frac{x_i}{y_i}\right), \qquad y = (y_1, \dots, y_d) \in Y.$$

By Lemma 4, for each $M \geq 0$ the one-dimensional sublevel set $K_M := \{z > 0 : J(z) \leq M\}$ is compact. Hence there exist $0 < a_M \leq 1 \leq b_M < \infty$ such that $K_M \subset [a_M, b_M]$. If $F_x(y) \leq M$ then each term satisfies $J(x_i/y_i) \leq M$, so $x_i/y_i \in K_M \subset [a_M, b_M]$, i.e.

$$\frac{x_i}{b_M} \leq y_i \leq \frac{x_i}{a_M} \qquad (i = 1, \dots, d).$$

Therefore the sublevel set $\{y \in Y : F_x(y) \leq M\}$ is closed and contained in the bounded box $\prod_i [x_i/b_M,\ x_i/a_M]$, so it is compact (Heine–Borel). Thus F_x attains its minimum on Y at some $y_* \in Y$. Choose $o \in O$ with $\iota_O(o) = y_*$; then $o \in \mathrm{Mean}_{\mathcal{R}}(s)$ by (15).

If $x \in Y$, then $F_x(x) = \sum_i J(1) = 0$. Since each term is nonnegative, 0 is the global minimum, so any o with $\iota_O(o) = x$ is a meaning. □

Definition 15 (Log-image and log-convexity)**.** *For $Y \subset (\mathbb{R}_{>0})^d$ define*

$$\log Y := \{(\log y_1, \dots, \log y_d) : \ y \in Y\} \subset \mathbb{R}^d.$$

We call Y log-convex *if $\log Y$ is convex.*

When the log-image of the dictionary is convex, strict convexity yields uniqueness and continuity of the optimizer.

Theorem 7 (Uniqueness and continuity for log-convex dictionaries)**.** *Assume the explicit mismatch cost (16) and the hypotheses of Theorem 6. If $U := \log Y \subset \mathbb{R}^d$ is closed and convex, then the minimizer $y_*(x) \in Y$ of F_x is unique. Equivalently, the meaning set $\mathrm{Mean}_{\mathcal{R}}(s)$ equals the fiber $\{o \in O : \ \iota_O(o) = y_*(\iota_S(s))\}$. Moreover, the optimizer is continuous in log-coordinates: the map $t \mapsto u_*(t)$ is continuous, where $t := \log x$ and $u_*(t) := \log y_*(e^t) \in U$.*

Proof. Let $t := \log x \in \mathbb{R}^d$ and write $u := \log y \in U$. By Lemma 5,

$$F_x(y) = \sum_{i=1}^{d} (\cosh(t_i - u_i) - 1) =: G_t(u).$$

For each i, the map $u_i \mapsto \cosh(t_i - u_i) - 1$ is strictly convex, hence G_t is strictly convex on $\mathbb{R}^d$. Restricting to the convex set U preserves strict convexity, so G_t has at most one minimizer on U; existence follows from Theorem 6. Thus the optimizer $u_*(t)$ is unique, and so is $y_*(x) = e^{u_*(\log x)}$.

For continuity, let $t_n \to t$ and set $u_n := u_*(t_n) \in U$. Fix $u_0 \in U$. Since u_n minimizes G_{t_n} on U, one has $G_{t_n}(u_n) \le G_{t_n}(u_0)$. The right-hand side is bounded because $(t, u) \mapsto G_t(u)$ is continuous and $t_n \to t$. As in the proof of Theorem 6, boundedness of $G_{t_n}(u_n)$ implies boundedness of $\{u_n\}$ in $\mathbb{R}^d$. Passing to a convergent subsequence (still denoted u_n) with limit $\bar{u} \in U$ (closedness), continuity gives $G_t(\bar{u}) = \lim_n G_{t_n}(u_n) \le \lim_n G_{t_n}(u) = G_t(u)$ for all $u \in U$. Hence $\bar{u}$ minimizes G_t on U, and by uniqueness $\bar{u} = u_*(t)$. Therefore every subsequence has the same limit, so $u_n \to u_*(t)$ and continuity holds. □

7. Worked examples

This section gives explicit computations in simple settings. The purpose is not to add new axioms, but to make the definition of meaning $\mathrm{Mean}_{\mathcal{R}}(s) = \arg\min_{o \in O} J(\iota_S(s)/\iota_O(o))$ concrete and to illustrate the decision-geometry proved earlier.

7.1. Continuous ratio model

Proposition 6 (Meaning in the continuous ratio model)**.** *Let $S = O = \mathbb{R}_{>0}$ with $\iota_S = \iota_O = \mathrm{id}$ and intrinsic costs $J_S = J_O = J$. Let $\mathcal{R}$ be admissible (Definition 6), so that*

$$c_{\mathcal{R}}(s, o) = J\Big(\frac{s}{o}\Big).$$

Then for every $s \in \mathbb{R}_{>0}$ there exists a unique *meaning, namely* $\mathrm{Mean}_{\mathcal{R}}(s) = \{s\}$*, and the minimum reference cost equals* 0*.*

Proof. By Lemma 2, one has $J(x) \ge 0$ for all $x > 0$ with equality if and only if $x = 1$. Hence $c_{\mathcal{R}}(s, o) = J(s/o) \ge 0$ with equality if and only if $s/o = 1$, i.e. $o = s$. Therefore $o = s$ is the unique minimizer and the minimum cost is 0. □

7.2. Finite dictionaries and boundary points

Example 4 (Finite object dictionary)**.** *Let $O = \{o_1, \dots, o_n\}$ be finite, set $y_i := \iota_O(o_i)$, and keep $S = \mathbb{R}_{>0}$ with $\iota_S = \mathrm{id}$. Under admissible reference, for a given configuration s with ratio $x := \iota_S(s)$ the meaning set is*

$$\mathrm{Mean}_{\mathcal{R}}(s) = \Big\{ o_i : \ J\Big(\frac{x}{y_i}\Big) = \min_{1 \le j \le n} J\Big(\frac{x}{y_j}\Big) \Big\}.$$

In general, boundary points (where the meaning set is not a singleton) occur when two or more of the values $J(x/y_i)$ tie.

7.3. Geometric-mean boundaries for the explicit mismatch cost

Theorem 8 (Geometric-mean decision boundaries for the explicit mismatch cost)**.** *Assume the explicit mismatch functional (3) and admissible (ratio-induced) reference $c_{\mathcal{R}}(s, o) = J(\iota_S(s)/\iota_O(o))$. Let $O = \{o_1, \dots, o_N\}$ be a finite object set such that the ratios $y_i := \iota_O(o_i)$ are pairwise distinct and ordered $0 < y_1 < \cdots < y_N$. For $x := \iota_S(s) \in \mathbb{R}_{>0}$ define the boundary points*

$$m_i := \sqrt{y_i y_{i+1}} \qquad (i = 1, \dots, N-1),$$

and set $m_0 := 0$, $m_N := +\infty$. Then:

- *If $m_{k-1} < x < m_k$ for some $k \in \{1, \dots, N\}$, then o_k is the* unique *meaning of s.*

- *If $x = m_k$ for some $k \in \{1, \dots, N-1\}$, then s has* exactly two *meanings, namely o_k and o_{k+1}.*

Equivalently, the map $x \mapsto \arg\min_i J(x/y_i)$ is piecewise constant on the open intervals (m_{k-1}, m_k).

Proof. Using (3) one computes, for each i,

$$c_{\mathcal{R}}(s, o_i) = J\left(\frac{x}{y_i}\right) = \frac{\left(\frac{x}{y_i} - 1\right)^2}{2(x/y_i)} = \frac{(x - y_i)^2}{2xy_i}.$$

Fix $i \in \{1, \dots, N-1\}$ and define the adjacent difference

$$\Delta_i(x) := c_{\mathcal{R}}(s, o_{i+1}) - c_{\mathcal{R}}(s, o_i).$$

Multiplying by $2x > 0$ and simplifying gives

$$2x\,\Delta_i(x) = (y_{i+1} - y_i)\left(1 - \frac{x^2}{y_i y_{i+1}}\right).$$

Hence $\Delta_i(x) = 0$ if and only if $x^2 = y_i y_{i+1}$, i.e. $x = m_i$. Moreover, $\Delta_i(x) > 0$ when $x < m_i$ and $\Delta_i(x) < 0$ when $x > m_i$. Therefore:

- if $x < m_i$ then $c_{\mathcal{R}}(s, o_i) < c_{\mathcal{R}}(s, o_{i+1})$ (so the adjacent comparison favors o_i),
- if $x > m_i$ then $c_{\mathcal{R}}(s, o_{i+1}) < c_{\mathcal{R}}(s, o_i)$ (so it favors o_{i+1}).

Fix $k \in \{1, \dots, N\}$ such that $m_{k-1} < x < m_k$. For every $i \le k-1$ we have $x > m_i$, hence $\Delta_i(x) < 0$, so $c_{\mathcal{R}}(s, o_{i+1}) < c_{\mathcal{R}}(s, o_i)$. Iterating these strict inequalities yields $c_{\mathcal{R}}(s, o_k) < c_{\mathcal{R}}(s, o_i)$ for all $i < k$. For every $i \ge k$ we have $x < m_i$, hence $\Delta_i(x) > 0$, so $c_{\mathcal{R}}(s, o_{i+1}) > c_{\mathcal{R}}(s, o_i)$. Iterating yields $c_{\mathcal{R}}(s, o_k) < c_{\mathcal{R}}(s, o_j)$ for all $j > k$. Therefore o_k is the unique minimizer.

If $x = m_k$ for some $k \in \{1, \dots, N-1\}$, then for every $i < k$ we still have $x > m_i$ and the costs strictly decrease up to index k, while for every $i \ge k+1$ we have $x < m_i$ and the costs strictly increase from index $k+1$ onward. At $i = k$ one has $\Delta_k(m_k) = 0$, i.e. $c_{\mathcal{R}}(s, o_k) = c_{\mathcal{R}}(s, o_{k+1})$. Hence the argmin consists of exactly two meanings, $\{o_k, o_{k+1}\}$. □

Corollary 7 (Stability away from boundaries)**.** *Under the hypotheses of Theorem 8, if $m_{k-1} < x < m_k$ then there exists $\delta > 0$ such that every x' with $|x' - x| < \delta$ satisfies $m_{k-1} < x' < m_k$ and hence has the same unique meaning o_k.*

Proof. Since (m_{k-1}, m_k) is open and contains x, choose $\delta := \min\{x - m_{k-1},\, m_k - x\}/2 > 0$. Then $|x' - x| < \delta$ implies $x' \in (m_{k-1}, m_k)$, and the conclusion follows from Theorem 8. □

Finite local resolution and discrete meaning cells.

The emergence of stable decision regions around geometric means (Theorem 8) provides a concrete realization of the *Finite Local Resolution* axiom of Recognition Geometry [17, Axiom 4 (RG3)]. While classical geometry typically assumes the idealization of infinite measurement precision, Recognition Geometry posits that local distinguishing power is always finite [17, Axiom 4 (RG3)]. Our results show that, under a cost-minimization dynamic with a finite dictionary $\iota_O(O) = \{y_1, \dots, y_N\}$, this discreteness emerges naturally: the continuous ratio axis $\mathbb{R}_{>0}$ is partitioned into open intervals on which the argmin is constant, separated by the discrete boundary set of geometric means $\{\sqrt{y_i y_{i+1}}\}$. In particular, *meanings* form discrete stable cells with a positive stability margin away from boundaries (Corollary 7).

7.4. Numerical micro-example (three-object dictionary)

Take $O = \{o_1, o_2, o_3\}$ with ratios $y_1 = \frac{1}{4} < y_2 = 1 < y_3 = 4$, and keep $S = \mathbb{R}_{>0}$ with $\iota_S = \mathrm{id}$. The boundary points are $m_1 = \sqrt{y_1 y_2} = \frac{1}{2}$ and $m_2 = \sqrt{y_2 y_3} = 2$. Thus a configuration with ratio x means o_1 for $0 < x < \frac{1}{2}$, means o_2 for $\frac{1}{2} < x < 2$, and means o_3 for $x > 2$ (with ties at the boundary points).

x	$c_{\mathcal{R}}(s, o_1)$	$c_{\mathcal{R}}(s, o_2)$	$c_{\mathcal{R}}(s, o_3)$	meaning(s)
$\frac{3}{10}$	$\frac{1}{60}$	$\frac{49}{60}$	$\frac{1369}{240}$	o_1
$\frac{3}{2}$	$\frac{25}{12}$	$\frac{1}{12}$	$\frac{25}{48}$	o_2
3	$\frac{121}{24}$	$\frac{2}{3}$	$\frac{1}{24}$	o_3

Example 5 (Mediation can sharply reduce cost in a toy case)**.** *Let* $a := \iota_S(s) = 4$ *and* $c := \iota_O(o) = \frac{1}{4}$*, so the direct admissible reference cost is* $J(a/c) = J(16) = \frac{225}{32}$*. If the mediator space contains a configuration* m *with ratio* $b_{\mathrm{geo}} := \sqrt{ac} = 1$*, then Theorem 6 gives an optimal sequential cost*

$$c_{\mathcal{R}_2 \circ \mathcal{R}_1}(s, o) = 2\, J\left(\sqrt{\frac{a}{c}}\right) = 2\, J(4) = \frac{9}{4},$$

which is strictly smaller, in accordance with Corollary 3.

8. Applications

This section records short corollaries and interpretive remarks that follow directly from the formal definitions and theorems; it makes no empirical or metaphysical claims beyond the stated axioms.

This section collects immediate, checkable consequences of the formal development. Each statement below follows from earlier definitions and theorems, and no external or empirical claim is being made. The meaning rule is an optimization rule (Definition 7) driven by the canonical mismatch penalty J(Definition 2); the axiomatic characterization of Jis classical and recorded for completeness in Appendix A.

8.1. Symbol grounding as a criterion

We treat "grounding" as an internal consistency condition in this model: a token s is grounded for an object o when (i) o is a meaning of s (Definition 7) and (ii) the symbol condition $J_S(s) < J_O(o)$ holds (Definition 8).

Corollary 8 (Grounding criterion under admissible reference)**.** *Fix an admissible reference structure* $\mathcal{R}$ *(Definition 6). Then, for* $s \in S$ *and* $o \in O$*,*

$$(s, o) \text{ is a symbol (Definition 8)} \iff o \in \mathrm{Mean}_{\mathcal{R}}(s) \text{ and } J_S(s) < J_O(o).$$

Proof. This is immediate from Definition 8 and Definition 7. □

Corollary 9 (Grounding rule for finite object dictionaries)**.** *Assume the finite-dictionary hypotheses of Theorem 8. As the configuration ratio* $x = \iota_S(s)$ *varies, the meaning set* $\mathrm{Mean}_{\mathcal{R}}(s)$ *is piecewise constant: it is a singleton on each interval* (m_{i-1}, m_i) *and can change only at the geometric-mean boundaries* $m_i = \sqrt{y_i y_{i+1}}$*. In particular, away from the boundaries the meaning is stable under small perturbations (Corollary 7).*

Proof. Immediate from Theorem 8 and Corollary 7. □

8.2. Mathematical effectiveness via low-cost primitives

The next corollary records a purely internal "near-balance" restriction: if a configuration has small intrinsic cost, then any of its meanings must lie in the corresponding low-mismatch window determined by the sublevel sets of J.

Corollary 10 (Near-balance restricts possible referents). *Assume $\mathcal{R}$ is admissible and that the hypotheses of Theorem 4 hold. If $s \in S$ satisfies $J_S(s) \le \epsilon$, and if o is a meaning of s, then*

$$J\left(\frac{\iota_S(s)}{\iota_O(o)}\right) \le \epsilon,$$

so $\iota_O(o)$ must lie in the corresponding bounded sublevel window determined by ϵ (as in Theorem 4).

Proof. This is a direct restatement of Theorem 4. □

Remark 2 (Compositional "range expansion" (model-dependent)). *In a continuous ratio model where ratios can be realized densely (e.g. $S = O = \mathbb{R}_{>0}$ with $\iota = \mathrm{id}$ as in Proposition 6), large mismatches can be decomposed into many small mismatches: write a target ratio $r = e^t$ as a product $r = (e^{t/k})^k$. Since $J(e^u) = \cosh(u) - 1 \to 0$ as $u \to 0$, choosing k large makes each primitive step low-cost. Coupled with the compositionality results (Theorem 5) and optimal mediation (Corollary 3), this shows that, in the continuous ratio model, large ratios can be factored into many small-ratio steps, each incurring small mismatch cost. This is an interpretive program; empirical relevance depends on what ratios are actually realizable in the intended application domain.*

8.3. Information-theoretic interpretation

Although our framework is stated in intrinsic-cost terms, the canonical mismatch penalty admits a simple log-ratio form. We record the identity as a proposition; any further links to coding/learning are interpretive and not used in the proofs.

Proposition 7 (Log-ratio form of the canonical mismatch cost). *For $x > 0$ write $x = e^t$. Then the canonical cost satisfies*

$$J(x) = J(e^t) = \cosh(t) - 1.$$

In particular, J is a convex even function of the log-ratio $t = \log x$ and vanishes exactly at $t = 0$.

Proof. Substitute $x = e^t$ into $J(x) = \frac{1}{2}(x + x^{-1}) - 1$ (Definition 2). □

9. Related work and positioning

This section places the framework in context, highlighting connections to aboutness in formal semantics, truthmaker-style ideas, and compression-based modeling, and clarifying what is new in the present optimization-based formulation.

Relation to *Recognition Geometry*.

Recognition Geometry [17] develops an axiomatic recognition-first framework in which observable space is derived from recognition events via an operational quotient construction. While the present paper does not attempt to construct an ambient geometry, it shares the same operational posture: the fundamental primitive is a measurable comparison (here the mismatch cost), and the induced semantic categories are those determined by minimizing or equating that comparison. The comparative-recognizer formalism of [17] provides a natural abstract home for the reference costs used here; we make this link explicit in Section 3.

This section positions the paper relative to standard themes in semantics and information theory. We do *not* present the mismatch penalty as novel: the axiom package in Definition 1 is a convenient specification whose solutions are classical (Appendix A). The contribution of the paper is instead the explicit *optimization semantics* (Definition 7) and the structural theorems derived from it (existence, stability geometry, compositionality, and mediation).

Symbol grounding and operational meaning rules.

The symbol grounding problem concerns how tokens acquire meaning without a homunculus [4]. The present work is compatible with grounding motivations, but it is formulated as a *mathematical model*: the meaning of s is *defined* as an argmin under an explicit cost. Any interpretation as a cognitive mechanism requires extra hypotheses beyond those stated.

Compression principles.

The general idea that effective representations trade off succinctness and fidelity is classical in information theory (Shannon [5]) and in algorithmic notions of complexity [6]; MDL makes this tradeoff concrete in model selection [7]. Our setup uses a different primitive: a ratio map ι into $\mathbb{R}_{>0}$ and a fixed mismatch penalty J, with compression enforced by the symbol condition $J_S(s) < J_O(o)$. Within this model, reference and compositional behavior become theorem-level consequences.

Remark 3 (Coding/learning viewpoint)**.** *In coding theory and learning, one often selects representations by minimizing a tradeoff between description length and distortion (e.g. Shannon [5] and MDL [7]). Our framework instantiates a specific distortion—$J(\iota_S(s)/\iota_O(o))$—that is symmetric in under-/over-shooting and naturally expressed in log-scale (Proposition 7). This suggests interpreting meanings as "best matches" under a fixed mismatch penalty, with compression enforced by the symbol condition $J_S(s) < J_O(o)$.*

Subject-matter/aboutness and truthmaker-semantics literature.

There is a substantial contemporary literature on "aboutness"/"subject matter" in semantics and logic, including Yablo's monograph [11] and subsequent discussion and refinements (e.g. Rothschild [13], Fine [14], and Yablo's reply [15]); see also Hawke's survey [12]. Related frameworks connect hyperintensional content with truthmakers/truthmaker semantics (e.g. Fine [16]). The present paper does not attempt to adjudicate between these accounts. Rather, it provides an explicit optimization layer which, once a modeling choice of scale maps is made, selects a subject matter/referent by minimizing a mismatch cost.

Novelty signal and conceptual payoff.

Many of the analytic lemmas are consequences of the specific penalty J and convexity. The intended novelty is the resulting *checkable decision geometry* and *compositional calculus* for meanings: finite dictionaries induce geometric-mean boundaries and stability margins, product models factorize exactly, and sequential mediation admits an explicit optimizer. These consequences are the main mathematical payoffs of the framework, and they make clear which modeling assumptions (the scale maps and admissibility hypotheses) must be checked in any intended application.

What is mathematically concrete here.

Two examples of explicit structure are: (i) for finite object dictionaries under the canonical mismatch penalty, decision boundaries occur at geometric means (Theorem 8) and meanings are locally stable away from them (Corollary 7); (ii) for sequential mediation, the optimal intermediate ratio is explicit (Theorem 6) and strictly improves over direct reference when the mediator set contains the balance point (Corollary 3).

Interpretation layer.

Sections 8 and 10 illustrate how the proved statements can be read once a modeling choice for ι is fixed. These illustrations are optional: removing them does not affect the correctness of the theorems.

10. Discussion

This section clarifies scope and interpretation: which parts are mathematical consequences of the axioms, which parts are modeling choices, and what additional assumptions would be needed to connect the formalism to empirical systems.

This section clarifies scope: which statements are proved inside the model and which statements are interpretation. It also records limitations and concrete mathematical extensions.

10.1. What is proved vs. what is modeled

The core mathematical content consists of the definitions and theorems in Sections 2–7. In particular, meaning is defined by optimization (Definition 7); existence is conditional on an attainment hypothesis (Theorem 2); and explicit geometry, stability, compositionality, and mediation statements follow for admissible reference structures and the canonical mismatch penalty (Theorems 8, 5, 6).

By contrast, any claim that a given real-world domain *does* admit a scale map ι with the required properties, or that agents *compute* meaning by solving the optimization problem, is an interpretation and is outside the theorem-level scope of this paper.

10.2. Limitations

1. **Ratio embedding**: Our framework requires configurations to embed into $\mathbb{R}_{>0}$ via a ratio map. Not all semantic domains naturally admit such embeddings.
2. **Single penalty**: We work with the canonical mismatch penalty J. Alternative penalties may be appropriate in domains where under- and over-shooting are not symmetric.
3. **Static analysis**: The theory is synchronic. Incorporating learning or time-evolution requires additional structure (e.g., dynamics for ι or for admissible reference classes).

10.3. Open problems

To make the forward-looking agenda explicit, we record a few concrete open problems aligned with the motivation above.

1. **Penalty universality beyond d'Alembert.** Identify alternative axiom packages (weaker than Definition 1(4)) that still force a small, classifiable family of penalties, and determine which decision-geometry and compositionality results remain valid.
2. **Structure of argmin ties.** Characterize, in terms of $\iota_O(O)$ and J, when the meaning set $\mathrm{Mean}(s)$ is multi-valued and how tie sets propagate under products and sequential mediation.
3. **Stability under perturbations of ι.** Quantify how errors in the scale maps affect decision boundaries and compositionality: derive uniform Lipschitz/margin bounds in log-space over admissible reference classes.

10.4. Future Directions

1. **Broader admissible reference.** Classify reference structures beyond the ratio-induced form (Definition 6) for which analogues of the stability and compositionality theorems remain true.
2. **Multi-dimensional ratios.** Extend the decision-geometry and boundary descriptions to $\iota : C \to (\mathbb{R}_{>0})^d$ with non-separable penalties, and quantify how coupling between coordinates affects stability margins.
3. **Learning the scale map.** Given data of successful/unsuccessful references, formulate and analyze estimation procedures for ι (and admissible reference parameters) that preserve the proved invariances.

11. Conclusion

This section summarizes the contributions and limitations of the model and records a few directions for refinement and application within the axioms fixed above.

We have developed a mathematical *model* of reference grounded in cost minimization. The theorem-level contributions are internal to the stated axioms and hypotheses.

We summarize the main points:

1. **Reference as compression**: Symbols are low-cost encodings of high-cost objects.
2. **Canonical mismatch geometry**: The canonical penalty $J(x) = \frac{1}{2}(x + x^{-1}) - 1$ yields explicit decision boundaries and stability regions for finite dictionaries (Theorem 8).
3. **Universal backbone**: Near-balanced configurations provide a provable backbone window around balance under admissible reference (Theorem 4). Global descriptive reach is obtained by composing many such low-cost primitives (Section 5).
4. **Compositionality**: Reference structures compose via products and sequences.

The framework connects a simple optimization semantics with explicit geometric and compositional structure. Any application to a specific empirical domain requires specifying an appropriate scale map ι and verifying that the admissibility assumptions reasonably match that domain.

Acknowledgments: We briefly acknowledge contributions and feedback that improved the exposition. We thank colleagues and readers for helpful discussions and feedback on earlier drafts.

Author Contributions: Conceptualization, J.W.; formal analysis, J.W. and A.R.B.; writing—original draft preparation, J.W.; writing—review and editing, J.W. and A.R.B. All authors have read and agreed to the published version of the paper.

Funding: TODO: Provide funding information (or state that there was no external funding).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: TODO: Declare conflicts of interest (or state none).

Appendix A. Classical characterization of the mismatch penalty

This appendix records a classical functional-equation characterization showing that the explicit mismatch penalty used in the paper is essentially forced (up to scale) by the stated axioms.

We prove Proposition 2. The underlying functional-equation step is classical; see, for example, Aczél [8] or Kuczma [9]. We include the argument here to keep the paper self-contained and to clarify that the mismatch penalty is not introduced as a new object.

Lemma A1 (Convexity implies continuity)**.** *Let $I \subset \mathbb{R}$ be an open interval and let $g : I \to \mathbb{R}$ be finite-valued and convex. Then g is continuous on I. (See, e.g., Rockafellar [10, Thm. 10.1].)*

We apply this standard convexity fact to the mismatch penalty to obtain the regularity needed for the functional-equation classification.

Lemma A2 (Regularity for the log-transformed d'Alembert equation)**.** *Assume J satisfies Definition 1. Define $C : (0, \infty) \to \mathbb{R}$ by $C(x) := 1 + J(x)$ and $f : \mathbb{R} \to \mathbb{R}$ by $f(u) := C(e^u)$. Then f is continuous, satisfies*

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

and obeys $f(0) = 1$. In particular, the hypotheses of Lemma A3 (and of the classical theorems of Acz'el and Kuczma) apply to f.

Proof. By strict convexity (Definition 1(3)), J is convex and finite-valued on $(0, \infty)$, hence continuous by Lemma A1; therefore $C = 1 + J$ and $f(u) = C(e^u)$ are continuous. The multiplicative identity in Definition 1(4) is equivalent to (A1) for C, and substituting $x = e^u$, $y = e^v$ yields the displayed d'Alembert equation for f. Finally $f(0) = C(1) = 1$ by normalization. □

Lemma A3 (Continuous solutions of d'Alembert's equation)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and satisfy*

$$f(t + s) + f(t - s) = 2f(t)f(s) \qquad (t, s \in \mathbb{R}),$$

with $f(0) = 1$. Then either $f \equiv 1$, or there exists $a > 0$ such that $f(t) = \cos(at)$ for all $t \in \mathbb{R}$, or there exists $a > 0$ such that $f(t) = \cosh(at)$ for all $t \in \mathbb{R}$.

Proof. This classification is classical; see Aczél [8, Ch. 2] or Kuczma [9, Ch. 13]. □

Proof of Proposition 2. Let J satisfy Definition 1. Define

$$C(x) := 1 + J(x) \qquad (x > 0).$$

Then (2) is equivalent to the multiplicative identity

$$C(xy) + C(x/y) = 2C(x)C(y) \qquad (x, y > 0). \tag{A1}$$

Define $f : \mathbb{R} \to \mathbb{R}$ by $f(t) := C(e^t)$. By Lemma A2, f is continuous. Substituting $x = e^t$ and $y = e^s$ into (A1) yields d'Alembert's functional equation

$$f(t + s) + f(t - s) = 2f(t)f(s) \qquad (t, s \in \mathbb{R}). \tag{A2}$$

Moreover, $f(0) = C(1) = 1$ and $f(t) \geq 1$ for all t since $J \geq 0$.

By Lemma A3, the continuous solutions of (A2) with $f(0) = 1$ are $f \equiv 1$, $f(t) = \cos(at)$, or $f(t) = \cosh(at)$ (for some $a > 0$, with the constant solution corresponding to $a = 0$). The constraint $f(t) \geq 1$ rules out the cosine family unless $a = 0$, and strict convexity rules out the constant solution. Hence there exists $a > 0$ such that $f(t) = \cosh(at)$ for all t.

Undoing the change of variables gives

$$C(x) = f(\log x) = \cosh(a \log x), \qquad x > 0,$$

and therefore

$$J(x) = C(x) - 1 = \cosh(a \log x) - 1 = \tfrac{1}{2}\left(x^a + x^{-a}\right) - 1.$$

Finally, note that

$$\cosh(a \log(\iota_S/\iota_O)) - 1 = \cosh\left(\log((\iota_S)^a/(\iota_O)^a)\right) - 1,$$

so replacing ι_S, ι_O by $\tilde{\iota}_S := \iota_S^a$ and $\tilde{\iota}_O := \iota_O^a$ absorbs the parameter a into the scale maps and produces the normalized choice $a = 1$ at the level of ratio-induced reference costs. □

References

1. G. Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50, 1892.
2. B. Russell. On denoting. *Mind*, 14(56):479–493, 1905.
3. E. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. *Communications on Pure and Applied Mathematics*, 13(1):1–14, 1960.
4. S. Harnad. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346, 1990.
5. C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423; 27(4):623–656, 1948.
6. A.N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1(1):1–7, 1965.
7. J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.
8. J. Aczél. Lectures on Functional Equations and Their Applications. Academic Press, 1966.
9. M. Kuczma. An Introduction to the Theory of Functional Equations and Inequalities: Cauchy's Equation and Jensen's Inequality. 2nd edition, Birkhäuser, 2009.
10. R.T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
11. S. Yablo. *Aboutness*. Princeton University Press, 2014.
12. P. Hawke. Theories of aboutness. *Australasian Journal of Philosophy*, 96(4):697–723, 2018. doi:10.1080/00048402.2017.1388826.
13. D. Rothschild. Yablo's semantic machinery. *Philosophical Studies*, 174(3):787–796, 2017. doi:10.1007/s11098-016-0759-3.
14. K. Fine. Yablo on subject-matter. *Philosophical Studies*, 177(1):129–171, 2020. doi:10.1007/s11098-018-1183-7.
15. S. Yablo. Reply to Fine on aboutness. *Philosophical Studies*, 175(6):1495–1512, 2018. doi:10.1007/s11098-017-0922-5.
16. K. Fine. Truth-maker semantics for intuitionistic logic. *Journal of Philosophical Logic*, 43(2–3):549–577, 2014. doi:10.1007/s10992-013-9281-7.
17. Washburn, J.; Zlatanović, M.; Allahyarov, E. *Recognition Geometry*. *Axioms* 2026, 15(2), 90. doi: 10.3390/axioms15020090