SOFTMAX BEYOND THE SIMPLEX: GIBBS CHARTS AND POLAR TRANSPORT ON EUCLIDEAN JORDAN ALGEBRAS

John Zachary Fitch

Independent Researcher

zack@internetuniverse.org

ORCID 0009-0007-7953-1531

October 2026

Abstract

Softmax is the normalized exponential of the simplicial Euclidean Jordan algebra $\mathbb{R}^n$. We study the same construction on an arbitrary finite-dimensional Euclidean Jordan algebra J. Its normalized Jordan exponential descends to a real-analytic diffeomorphism

$$\Gamma_J : J/\mathbb{R}e \longrightarrow \mathscr{B}_J^\circ, \qquad \Gamma_J(x) = \frac{\exp_J(x)}{\mathrm{tr}_J(\exp_J x)},$$

which we call the Gibbs chart. Within continuous separable spectral normalizers, invariance under common score shifts determines the temperature family $\Gamma_J(\beta\,\cdot)$, and first-order calibration at the barycenter selects $\beta = 1$. The associated log-partition function is Legendre dual to spectral negative entropy.

The Hessian of that log-partition function separates into the usual softmax covariance on spectral weights and logarithmic-mean operators on the off-diagonal Peirce blocks. This yields an explicit Jacobian and a real Monge–Ampère equation whose source is the determinant of the frame block. The source is identically one precisely for the simplicial carrier, the unique Euclidean Jordan carrier with no continuous frame directions; there the centered gradient image is the extremal Ehrhart simplex.

We then study canonical normalization of cone-valued query–key data. Adjoining the aggregate positive coordinate makes effect normalization lossless, while the normalized effects retain an $\mathrm{Aut}(J)$-valued polar transporter. It is trivial for spectrally commuting pairs — and identically trivial exactly on the simplicial carrier — reduces to the polar unitary for matrix carriers, and is F_4-valued for the Albert algebra. It is invariant under separate rescalings of query and key, hence a function on the projectivized cone, while the radial factor carries the scale; the two channels couple at second order through the curvature bracket of the cone. At the identity query its linearization acts blockwise on the Peirce frame tangent and is injective when the aggregate key has simple spectrum. Representations of the transporter yield positive Jordan instruments, exact additive moment reductions for central values, relative-frame readouts, and positive semidefinite kernels from transported Peirce projections. Ordinary softmax remains the spectral law; non-simplicial carriers add canonical frame geometry and transport.

---

1 INTRODUCTION: SOFTMAX IS A CHART

Softmax is a chart.

The pair $(\mathbb{R}^n, \text{componentwise product})$ is a Euclidean Jordan algebra. Every Euclidean Jordan algebra carries, canonically and with no further data, a symmetric cone of squares, a trace, and a spectral exponential; and the normalized exponential

$$\Gamma_J(x) \;=\; \frac{\exp_J(x)}{\mathrm{tr}_J(\exp_J x)}$$

is a real-analytic diffeomorphism from the algebra modulo its unit line onto the interior of the trace-one slice of that cone. For $J = \mathbb{R}^n$ the slice is the probability simplex, and the diffeomorphism is

$$\mathrm{softmax}(x)_i \;=\; \frac{e^{x_i}}{\sum_j e^{x_j}}.$$

The map entered machine learning as the "normalised exponential" output law for classification networks [13] and is now the routing law of attention [46]; as a normalized exponential of an energy it is far older, and Corollary 6.6 recovers the thermodynamic reading [23].

Read this way, the three stipulations usually made about softmax stop being stipulations. The simplex is not a target chosen for softmax to land in; it is the trace-one base of the carrier's own cone. Within the separable spectral class considered here, the exponential is not one positive function among many; it is the lift that carries an element with no positivity into the cone, and the theorem below says that under separability it is the only such lift compatible with common score shifts. The identity $\mathrm{softmax}(x + a\mathbf{1}) = \mathrm{softmax}(x)$, ordinarily described as the numerical convenience licensing max-subtraction, is the statement that the map factors through the projectivized cone: the additive gauge $\mathbb{R}e$ upstairs becomes the multiplicative gauge $\mathbb{R}_{>0}$ downstairs, and normalization is the choice of trace-one representative on each ray.

What is left variable is the algebra. We call the chosen Euclidean Jordan algebra J the *carrier*, and the simplicial carrier is chosen when token states are modeled as independent scalar channels with componentwise positivity and multiplication. Under that modeling convention the simplicial carrier is already fixed. The consequences of that choice are sharp. Write r for the rank of J and $\mathfrak{F}_J$ for its ordered-frame space. On the simple-spectrum locus,

$$\dim \mathscr{B}_J^{\circ} \;=\; \underbrace{(r-1)}_{\text{spectral weights}} \;+\; \underbrace{(\dim J - r)}_{\text{frame directions}},$$

and the second summand vanishes exactly when $J \cong \mathbb{R}^r$ (Theorem 7.4). Ordinary softmax is the unique frameless Gibbs chart. Every other Euclidean Jordan carrier produces normalized states carrying, in addition to their weights, a positive-dimensional frame; for the Albert algebra $H_3(\mathbb{O})$ the split is 2 weight dimensions against 24 frame dimensions inside a 26-dimensional state space.

Saying exactly what those frame directions do is the substance of what follows.

The paper has three parts. First, we identify the normalized Jordan exponential as a global Gibbs chart, characterize its temperature family by common-score-shift invariance, and select

unit temperature by barycentric calibration. Convex duality yields spectral entropy, Jordan relative entropy, a direct-sum hierarchy, and a zero-temperature selection rule on exposed faces. Second, we show that changing the carrier adds Peirce frame directions to the familiar spectral weights. One quantity then does three jobs: the logarithmic mean of two Gibbs weights is the metric coefficient on the Peirce block joining them, the factor that block contributes to the Jacobian, and the source term the block contributes to a real Monge–Ampère equation satisfied by the rescaled potential. The source is absent exactly for the simplicial carrier, whose centered gradient image is the simplex realizing the extremal value in Ehrhart's volume conjecture. Third, canonical normalization of carrier-valued query–key scores separates radial data from an $\mathrm{Aut}(J)$-valued polar transporter. The normalized effects act through a direct-sum Jordan instrument. At the identity query, the transporter linearization is injective on the full frame tangent when the aggregate key has simple spectrum, and it admits representation-valued and spectral-code readouts as well as an exact additive-moment reduction for central values. A spectral frame of the aggregate key yields blockwise positive Peirce-projection kernels, while the label-free total frame kernel has second-order degeneracy exactly equal to that of the transporter linearization. For the Albert algebra these directions form three triality-related eight-dimensional channels.

The carrier does not by itself prohibit other maps into its cone: sparsemax [29] and the α-entmax family [33] are simplex-valued and shift-invariant, and arise from a different set of design axioms. Theorem 5.1 is a uniqueness statement inside the separable spectral class, and Remark 5.2 draws that line precisely.

Sections 2 to 5 build the chart and establish its rigidity; Sections 6 to 8 read the resulting geometry off it; and Section 9 builds the polar transporter and everything downstream of it.

Trace-one slices of symmetric cones and entropy-regularized normalized exponential updates have appeared before, in symmetric-cone online optimization and game learning. Symmetric-cone multiplicative weights identifies the normalized Jordan exponential with follow-the-regularized-leader and online mirror descent for symmetric-cone negative entropy [14], and the same closed-form update, together with the corresponding strong-convexity theory, drives optimistic symmetric-cone games [6]. Our use of the same map is geometric rather than algorithmic. We treat it as a global quotient chart, characterize it within the separable spectral class by common-shift invariance, compute its complete spectral and Peirce-block differential and Jacobian, and use the cone-group polar factor to expose frame transport and its representation-valued readouts. That the carrier's state-space geometry is fixed by the algebra is itself a theorem: strongly symmetric spectral convex bodies are exactly the state spaces of simple Euclidean Jordan algebras and the simplices [8]. Another route assumes neither spectrality nor self-duality: a homogeneous cone whose normalized automorphisms act transitively on pure states is self-dual, hence Jordan [9]. Remark 2.2 sets these hypotheses beside the two transports used here, and Remark 2.3 reads the Peirce data of a carrier as Vinberg's description of its cone.

2 EUCLIDEAN JORDAN CARRIERS

We recall only what is used. The algebras below are exactly the formally real Jordan algebras classified by Jordan, von Neumann, and Wigner [24]. Faraut and Korányi [17] is the standard reference for their cone theory, with Koecher's lectures [26] as a second route; [22, 30] give the algebraic background and [42, 5] the exceptional case.

Definition 2.1 (Euclidean Jordan algebra). A *Euclidean Jordan algebra* (EJA) is a finite-dimensional real vector space J with a commutative bilinear product $\circ$, a unit e, satisfying the Jordan identity $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, and carrying an inner product for which every multiplication operator $L_x : y \mapsto x \circ y$ is self-adjoint.

Jordan algebras are power-associative, so $x^{\circ k}$ is unambiguous. The *spectral theorem* [17, Thm. III.1.2] states that every $x \in J$ can be written

$$x \;=\; \sum_{k=1}^{r} \lambda_k c_k, \tag{1}$$

where $\lambda_k \in \mathbb{R}$ and $(c_1, \dots, c_r)$ is a *Jordan frame*: a family of primitive idempotents with $c_i \circ c_j = \delta_{ij} c_i$ and $\sum_k c_k = e$. The integer $r = \operatorname{rank} J$ is independent of x. The λ_k are the *eigenvalues* of x, determined with multiplicity; the frame is determined uniquely when the eigenvalues are distinct. Write $\mathfrak{F}_J$ for the set of ordered Jordan frames.

The *canonical trace* is $\operatorname{tr}_J(x) = \sum_k \lambda_k$, so that $\operatorname{tr}_J(c) = 1$ for every primitive idempotent c and $\operatorname{tr}_J(e) = r$. The trace form $T(x, y) = \operatorname{tr}_J(x \circ y)$ is positive definite and *associative*:

$$T(x \circ y, z) \;=\; T(x, y \circ z), \tag{2}$$

so $(x, y, z) \mapsto T(x \circ y, z)$ is a symmetric trilinear form. We use $\langle x, y \rangle := T(x, y)$ throughout; Remark 8.1 records why this matters in implementation.

The *symmetric cone* is $\Omega_J = \operatorname{int}\{\, x \circ x \,:\, x \in J \,\} = \{x \,:\, \lambda_k(x) > 0 \ \forall k\}$, an open convex self-dual cone. Spectral calculus defines $f(x) = \sum_k f(\lambda_k) c_k$ for any $f : \mathbb{R} \to \mathbb{R}$, a well-defined element independent of the chosen spectral decomposition; in particular

$$\exp_J(x) \;=\; \sum_{k \ge 0} \frac{x^{\circ k}}{k!} \;=\; \sum_{k=1}^{r} e^{\lambda_k} c_k, \tag{3}$$

and $\exp_J : J \to \Omega_J$ is a real-analytic diffeomorphism with inverse $\log_J$ [17, Ch. III]. We write $P(a) = 2L_a^2 - L_{a^2}$ for the *quadratic representation*, which for matrix carriers is $P(a)b = aba$; each $P(a)$ with $a \in \Omega_J$ is a linear automorphism of the cone, positive definite and self-adjoint for the trace form, and $P(a^{-1/2})a = e$ [17, Prop. II.3.1, Thm. III.2.1].

Remark 2.1 (The exponential is intrinsic, and in what sense). The cone Ω_J is not a group under the Jordan product, which is nonassociative, nor (for matrix carriers) under ordinary multiplication, which leaves the carrier. It is a homogeneous space: its linear automorphism group $G(\Omega_J)$ acts transitively, and it carries a canonical invariant Riemannian metric $g_a(u, v) = \langle P(a)^{-1}u, v \rangle$. For a *simple* carrier the Riemannian exponential of this geometry at a is $\operatorname{Exp}_a(v) = P(a^{1/2}) \exp_J(P(a^{-1/2})v)$, a diffeomorphism $J \to \Omega_J$; at $a = e$ it reduces to $\exp_J$ [3]. So the Jordan

exponential is the symmetric-space exponential chart at the identity, extended factorwise over the simple-ideal decomposition in general. If a literal Lie-group statement is wanted, one is available by transporting the additive structure of J through $\exp_J$; but that group is created by the chart rather than found in the cone, and nothing below uses it.

Given a primitive idempotent c, the operator L_c has spectrum contained in $\{0, 1/2, 1\}$ and J decomposes into Peirce spaces $J = V_1(c) \oplus V_{1/2}(c) \oplus V_0(c)$ with $V_1(c) = \mathbb{R}c$. Relative to a full frame,

$$J \;=\; \bigoplus_{k=1}^{r} \mathbb{R}c_k \;\oplus\; \bigoplus_{i<j} J_{ij}, \qquad d_{ij} := \dim J_{ij}, \tag{4}$$

where $J_{ij} = \{v : L_{c_i}v = L_{c_j}v = \tfrac{1}{2}v\}$, and $V_{1/2}(c_i) = \bigoplus_{j\neq i} J_{ij}$. Counting (4),

$$\dim J \;=\; r + \sum_{i<j} d_{ij}. \tag{5}$$

The block dimensions need not be equal: in $J = \mathbb{R} \oplus \mathrm{Sym}_2(\mathbb{R})$ a frame contains the scalar primitive and two matrix primitives, the two cross-factor blocks vanish, and the matrix-factor block has dimension 1. When J is *simple*, all off-diagonal blocks share one dimension d, the *Peirce multiplicity*, and (5) becomes $\dim J = r + \tfrac{1}{2}r(r-1)d$. When $J \cong \mathbb{R}^r$ all blocks vanish and $d = 0$.

Definition 2.2 (Frame-transitive carrier). An EJA J is *frame-transitive* if $\mathrm{Aut}(J)$ acts transitively on its ordered Jordan frames. These are exactly the simple Euclidean Jordan algebras and the simplicial algebras $\mathbb{R}^r$ [17, Ch. IV–V]. On a general reducible carrier, automorphisms preserve the decomposition into isotypic families of simple ideals and may permute isomorphic summands, but they cannot mix primitive idempotents freely across the full frame; consequently the action on ordered frames is not transitive unless the carrier is simple or simplicial.

Every EJA is a direct sum of simple ideals, and choosing a frame in each summand produces a frame of J; every element is conjugate, by an automorphism acting factorwise, into the resulting fixed factor-compatible diagonal. What frame transitivity buys is stronger: transitivity of $\mathrm{Aut}(J)$ on the *full* ordered frame space $\mathfrak{F}_J$, hence an action of the full symmetric group S_r on all spectral slots [17, Ch. IV]. Both fail for $\mathbb{R} \oplus \mathrm{Sym}_2(\mathbb{R})$, where no automorphism moves the scalar primitive into the matrix factor; and it is exactly that failure, not any failure of diagonalization, which prevents the global eigenvalue multiset from being a complete orbit invariant (Example 7.1).

We call $\mathbb{R}^r$ with the componentwise product the *simplicial* carrier, since its cone $\mathbb{R}^r_{>0}$ is a simplicial cone and its frame is a basis defined over the ground field. The alternative name *split* is standard for the same reason but collides with the split Albert algebra $H_3(\mathbb{O}_s)$ over the split octonions, which does not appear here — every carrier in this paper is Euclidean, hence a formally real form — so we avoid it.

Example 2.1. $\mathbb{R}^n$ with $x \circ y = (x_i y_i)$ has $r = n$ and $d = 0$. $\mathrm{Sym}_r(\mathbb{R})$ with $X \circ Y = \tfrac{1}{2}(XY + YX)$ has rank r and multiplicity 1; $\mathrm{Herm}_r(\mathbb{C})$ and $\mathrm{Herm}_r(\mathbb{H})$ have rank r and multiplicities 2 and 4. The spin factor $\mathrm{V}_m = \mathbb{R} \oplus \mathbb{R}^m$, with $(a, u) \circ (b, v) = (ab + \langle u, v\rangle,\ av + bu)$ — a Lorentz product, not the componentwise direct-sum product — has rank 2 and multiplicity $m - 1$; it is simple for $m \geq 2$, while $\mathrm{V}_1 \cong \mathbb{R}^2$ is simplicial. The Albert algebra $H_3(\mathbb{O})$ has $r = 3$, $d = 8$, $\dim = 3 + 24 = 27$, and $\mathrm{Aut}(H_3(\mathbb{O})) = F_4$.

Remark 2.2 (Why the carriers are Jordan). This paper assumes Jordan structure, and of the cone it then uses two transitive actions. The linear automorphism group $G(\Omega_J)$ acts transitively on Ω_J (Remark 2.1), and the polar transporter takes values in $\mathrm{Aut}(J) = \mathrm{Stab}_G(e)$ (Theorem 9.2), which is exactly the group of cone automorphisms preserving the closed state space $\mathscr{B}_J$ of Definition 3.1. Indeed $\mathrm{tr}_J(gx) = \langle g^* e, x\rangle$, so $g \in G(\Omega_J)$ preserves $\mathscr{B}_J$ exactly when its adjoint fixes e; the cone is self-dual, so $g^* \in G(\Omega_J)$, hence $g^* \in \mathrm{Aut}(J)$, and $\mathrm{Aut}(J)$ consists of orthogonal maps, so $g = (g^*)^{-1}$ lies in $\mathrm{Aut}(J)$ as well.

Barnum, Ududec and van de Wetering [9, Thm. 25] prove a converse with self-duality removed from the hypotheses. Let C be a closed, pointed, generating cone in a finite-dimensional real space and $\mathscr{B}$ a base of C. Suppose C is *homogeneous*: its linear automorphisms act transitively on its interior. Suppose $\mathscr{B}$ has *pure transitivity*: the automorphisms of C that preserve $\mathscr{B}$ act transitively on its extreme points. Then C is self-dual, and by the Koecher–Vinberg theorem [25, 47] it is the closed cone of squares of a Euclidean Jordan algebra. The proof runs through Vinberg's theory of homogeneous cones [48, 49]. The cone has a self-dual subcone, its *kernel*, which is invariant under the base-preserving automorphisms and contains a pure state of C, and pure transitivity carries that pure state onto every other, so the kernel is all of C. Refined: pure transitivity holds exactly for the direct sums $W \oplus \cdots \oplus W$ of copies of one simple algebra, and *continuous* pure transitivity, a continuous path of base-preserving images between any two pure states, holds exactly for the simple algebras [9, Cor. 44–45].

Read against this paper, the theorem says that the carrier class need not be postulated. A cone that supports a transitive positive transport, as the sections $P(x^{1/2})$ do here, and whose state-space-preserving automorphisms move any pure state to any other, is already one of the carriers of this section. Three qualifications apply.

(a) The hypothesis concerns the group $\mathrm{Aut}(J)$, not the set of values of the transporter. The paper does not show that the values $\mathscr{U}(q, K)$ act transitively on pure states, and the theorem does not need them to.

(b) The class the theorem delivers is larger than the frame-transitive class of Definition 2.2. The simple carriers are exactly those with continuous pure transitivity, and the simplicial carrier $\mathbb{R}^r = \mathbb{R} \oplus \cdots \oplus \mathbb{R}$ has pure transitivity through permutations. But $\mathrm{Herm}_2(\mathbb{C}) \oplus \mathrm{Herm}_2(\mathbb{C})$ has pure transitivity too, through the automorphisms of each summand and the swap, and it is not frame-transitive: no automorphism carries an ordered frame whose first two primitives lie in one summand to an ordered frame whose first two primitives lie in different summands. The frame-transitive carriers are the pure-transitive ones with a single summand or with repeated summand $\mathbb{R}$.

(c) The route of [8] quoted in the introduction assumes spectrality and strong symmetry, which is transitivity on ordered sets of perfectly distinguishable pure states and so includes transitivity on ordered frames. Consistently with (b), it returns exactly the frame-transitive carriers, while the route here assumes transitivity only on single pure states and returns the larger class.

Remark 2.3 (The Peirce data as a Vinberg T-algebra). Vinberg describes every homogeneous cone by a *T-algebra* [48, Ch. III]. This is a bigraded real algebra $T = \bigoplus_{i,j \le r} T_{ij}$ with $T_{ij} T_{jk} \subseteq T_{ik}$,

diagonal spaces $T_{ii} \cong \mathbb{R}$, an involution $*$ exchanging T_{ij} and T_{ji}, and a trace, subject to axioms that include

$$t(uw) = (tu)w, \qquad t(uu^*) = (tu)u^* \qquad (t, u, w \text{ upper triangular}). \tag{6}$$

The cone is $\{tt^* : t \text{ upper triangular with nonnegative diagonal}\}$ [9, Props. 32–33]. Write $n_{ij} = \dim T_{ij}$. An irreducible homogeneous cone is self-dual exactly when all n_{ij} with $i \neq j$ are equal and nonzero [49], [9, §3.3].

For the matrix carriers $\mathrm{Sym}_r(\mathbb{R})$, $\mathrm{Herm}_r(\mathbb{C})$, $\mathrm{Herm}_r(\mathbb{H})$ and for $H_3(\mathbb{O})$, take T to be the $r \times r$ matrices over the coefficient algebra, with the matrix product, every product landing on the diagonal replaced by its real part (as $T_{ii} \cong \mathbb{R}$ requires), and conjugate transpose as $*$. Then T_{ij} with $i < j$ is the Peirce space J_{ij} of (4) for the standard frame, $n_{ij} = d_{ij}$, and $\{tt^*\}$ is the Cholesky description of $\overline{\Omega}_J$; Section C checks the Albert case numerically. So the Peirce data (r, d_{ij}) used throughout — a graph on the frame idempotents with edge weights d_{ij} — is the dimension data of a T-algebra. Self-duality is the statement that the edge weights are constant and nonzero on each simple summand, as (5) records. Homogeneous cones that are not self-dual are T-algebras whose edge dimensions vary inside an irreducible block, and none of them is a carrier.

The two axioms in (6) are conditions on paths in this graph. An entry of $t(uw)$ is a sum of products $t_{ij}u_{jk}w_{kl}$ along paths $i \le j \le k \le l$ that only move upward. Such a product has three off-diagonal factors only when the four indices are distinct. At rank $r \le 3$ every product has at most two off-diagonal factors and real numbers elsewhere, so the first axiom holds over any coefficient algebra. The second axiom goes back along an edge it has used: its $(1,2)$ entry compares $(t_{12}u_{23})\bar{u}_{23}$ with $t_{12}(u_{23}\bar{u}_{23})$, the right alternative law, and its diagonal entries differ by associators, whose real parts vanish. Hence over $\mathbb{O}$ the second axiom holds at rank 3, because $\mathbb{O}$ is alternative. The first fails at rank 4, in the $(1,4)$ entry, by the associator of three octonions. Over the sedenions, which are not alternative, the second axiom fails already at rank 3, in the $(1,2)$ entry. These bounds agree with Jacobson's coordinatization theorem: a Jordan algebra with $r \ge 3$ connected supplementary idempotents is $H_r(D)$ with D alternative, and associative when $r \ge 4$ [22, 30]. Vinberg's axioms separate the two bounds. The rank-4 bound on $\mathbb{O}$ comes from the open-path axiom, and the rank-3 exclusion of the sedenions from the returning-path axiom.

Deleting the first row and column exhibits the cone as a *Siegel cone* over the cone of the remaining rows [9, §§3.1–3.2]. Take the anchor $T_{11} \cong \mathbb{R}$, the off-diagonal space $\mathbb{R}^p = T_{21} \oplus \cdots \oplus T_{r1}$, and the remaining T-algebra V, with $B(u,u) = uu^*$. Then $(a, u, t) \in V \times \mathbb{R}^p \times \mathbb{R}$ lies in the cone exactly when $t \ge 0$ and $ta - B(u,u)$ lies in the cone of V. In Peirce terms this is the decomposition $J = V_1(c_1) \oplus V_{1/2}(c_1) \oplus V_0(c_1)$ relative to the first primitive of the frame. For $H_3(\mathbb{O})$ it reads $27 = 1 + 16 + 10$, with $V = H_2(\mathbb{O}) \cong \mathsf{V}_9$. Truong and Tunçel's description of the extreme rays of a Siegel cone [43, Thm. 2], [9, Thm. 41] splits the pure states into two kinds. The first kind are the pure states of V, with $t = 0$ and therefore $u = 0$. The second kind are the rays through $(B(u,u), u, 1)$ with $u \in \mathbb{R}^p$. For $H_3(\mathbb{O})$ the second kind are the elements vv^* with $v = (1, u_1, u_2)$ and $(u_1, u_2) \in \mathbb{O}^2 \cong \mathbb{R}^{16}$, which after normalization form the affine chart of the octonionic projective plane. The first kind are the pure states of $H_2(\mathbb{O})$, an 8-sphere, the projective line at infinity. This is the cell decomposition $\mathbb{O}P^2 = \mathbb{R}^{16} \cup S^8$. It parametrizes the pure states of the Albert carrier by one primitive of a frame and its Peirce space $V_{1/2}(c_1)$.

Once J is fixed, its cone, trace, spectral calculus and Gibbs chart require no further choice. A carrier arrives with its cone, its trace, and its exponential already fixed, and the next section shows that this is enough to produce the map.

3 THE GIBBS CHART AS A PROJECTIVIZED EXPONENTIAL

Definition 3.1 (State space and Gibbs chart). The *open state space* of J is

$$\mathscr{B}_J^\circ \;=\; \{\, p \in \Omega_J \,:\, \operatorname{tr}_J(p) = 1 \,\},$$

the trace-one section of the symmetric cone, and its *Gibbs chart* is $\Gamma_J = \pi \circ \exp_J$, where $\pi : \Omega_J \to \mathscr{B}_J^\circ$, $\pi(y) = y/\operatorname{tr}_J(y)$. Write $J_0 := \ker \operatorname{tr}_J$ for the traceless hyperplane and $\mathscr{B}_J = \{p \in \overline{\Omega}_J : \operatorname{tr}_J p = 1\}$ for the *closed* state space, $\overline{\Omega}_J$ denoting the closed cone. Boundary states, where some eigenvalue vanishes, lie in $\mathscr{B}_J \setminus \mathscr{B}_J^\circ$; the chart never reaches them, but Theorem 6.4 maximizes over them.

Theorem 3.2 — Global Gibbs chart

For every Euclidean Jordan algebra J of rank r:

(a) $\Gamma_J = \pi \circ \exp_J$, and both factors are surjective onto their stated targets;

(b) $\Gamma_J(x + ae) = \Gamma_J(x)$ for every $a \in \mathbb{R}$;

(c) the restriction $\Gamma_J|_{J_0} : J_0 \longrightarrow \mathscr{B}_J^\circ$ is a real-analytic diffeomorphism, with inverse

$$\Lambda_J(p) \;=\; \log_J p \;-\; \frac{1}{r}\operatorname{tr}_J(\log_J p)\, e.$$

Equivalently, the induced map $J/\mathbb{R}e \to \mathscr{B}_J^\circ$ is a real-analytic diffeomorphism.

In spectral coordinates, if $x = \sum_k \lambda_k c_k$ then

$$\Gamma_J(x) \;=\; \sum_{k=1}^{r} \frac{e^{\lambda_k}}{\sum_\ell e^{\lambda_\ell}}\, c_k. \tag{7}$$

Proof. **Step 1** (*the two factors*)**.** $\exp_J$ maps J onto Ω_J [17, Ch. III]. For $y \in \Omega_J$ all eigenvalues are positive, so $\operatorname{tr}_J(y) > 0$ and $y/\operatorname{tr}_J(y)$ again lies in Ω_J, now with trace 1; hence π maps Ω_J onto $\mathscr{B}_J^\circ$, and it is onto because π fixes $\mathscr{B}_J^\circ$ pointwise. Composing gives (a), and (3) with $\operatorname{tr}_J(c_k) = 1$ gives (7).

Step 2 (*the scalar gauge*)**.** Let $x = \sum_k \lambda_k c_k$ be a spectral decomposition. Since $e = \sum_k c_k$ for *any* frame, the same frame diagonalizes $x + ae = \sum_k (\lambda_k + a)c_k$. Therefore $\exp_J(x + ae) = e^a \exp_J(x)$, and the positive scalar e^a cancels between numerator and denominator of Γ_J. This is (b).

Step 3 (Λ_J *is a right inverse*)**.** Given $p \in \mathscr{B}_J^\circ$, set $m = 1/r \operatorname{tr}_J(\log_J p)$ and $x = \log_J p - me$. Then $\operatorname{tr}_J x = 0$, so $x \in J_0$, and by Step 2 applied to $\log_J p = x + me$ we get $\exp_J(x) = e^{-m}p$. Since $\operatorname{tr}_J p = 1$, the trace of $\exp_J(x)$ is e^{-m}, and $\Gamma_J(x) = e^{-m}p/e^{-m} = p$.

Step 4 (Λ_J *is a left inverse*)**.** Let $x \in J_0$, put $Z(x) = \operatorname{tr}_J(\exp_J x)$ and $p = \Gamma_J(x)$. In a frame diagonalizing x, $p = \sum_k (e^{\lambda_k}/Z)c_k$, so $\log_J p = x - (\log Z)e$. Taking traces and using $\operatorname{tr}_J x = 0$ gives $\operatorname{tr}_J(\log_J p) = -r \log Z$, hence $\Lambda_J(p) = (x - (\log Z)e) + (\log Z)e = x$.

Step 5 (*regularity and the quotient*)**.** $\exp_J$ and $\log_J$ are real-analytic, tr_J is linear, and Z is nowhere zero, so $\Gamma_J|_{J_0}$ and Λ_J are real-analytic; Steps 3–4 show they are mutually inverse bijections. Since $J = J_0 \oplus \mathbb{R}e$ and, by (b), Γ_J is constant on cosets of $\mathbb{R}e$, the induced map on $J/\mathbb{R}e$ is the same diffeomorphism read through the linear isomorphism $J/\mathbb{R}e \cong J_0$. □

The factorization in Theorem 3.2(a) is the part of the theorem that carries architectural content, so we name its halves. The map $\exp_J$ is a *cone lift*: it takes an element of J, which has no positivity, and returns an element of Ω_J, which does. The map π is *normalization*: it rescales an element that is already positive onto the trace-one slice. A model emitting cone elements directly may omit the lift, having thereby changed its score coordinates; Section 9 takes up what it then owes. For the moment we take the chart as given and ask what it looks like in the carrier everyone already uses.

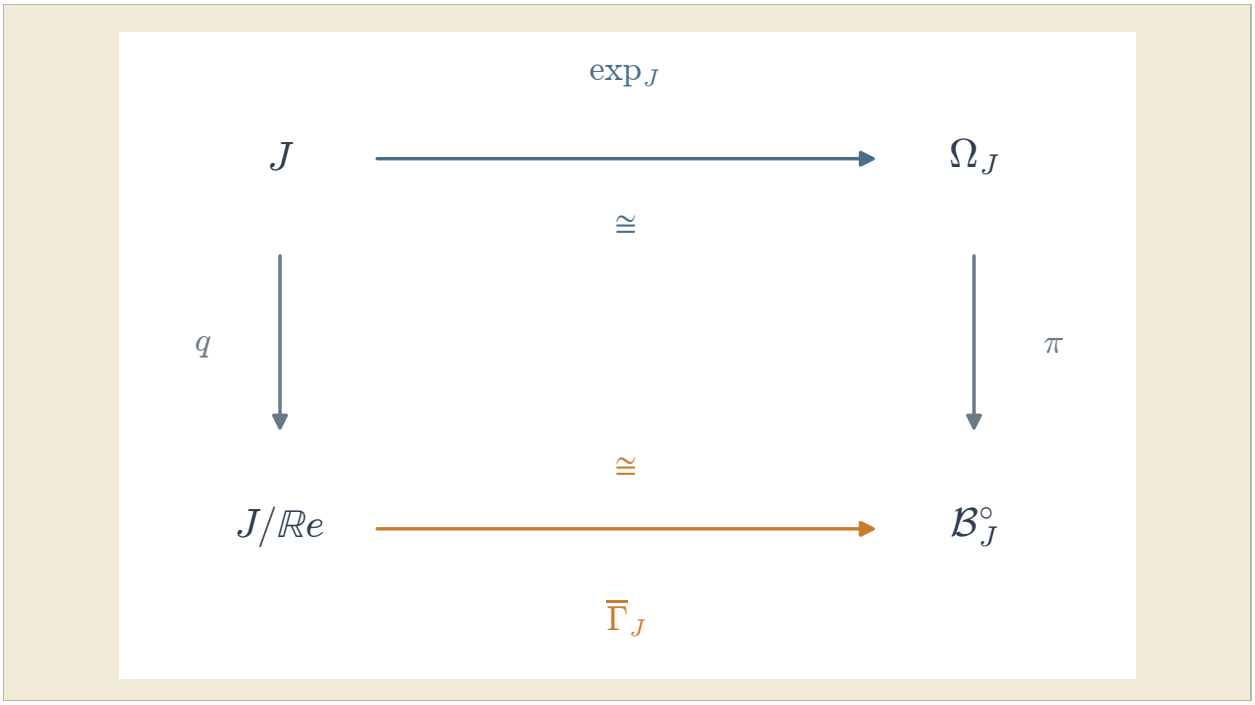


FIG. 1. *The Gibbs chart as a cone lift followed by a normalization. The horizontal maps are diffeomorphisms; the vertical maps quotient by the additive gauge* $\mathbb{R}e$ *on the left and the multiplicative gauge* $\mathbb{R}_{>0}$ *on the right. Softmax is this square for* $J = \mathbb{R}^n$.

The chart also factorizes over a reducible carrier, in a way that puts ordinary softmax on the outside.

Proposition 3.3 (Direct-sum factorization). Let $J = \bigoplus_{a=1}^{s} J_a$ with each J_a a Euclidean Jordan algebra, write $x = (x_1, \dots, x_s)$ and $Z_a = \operatorname{tr}_{J_a} \exp_{J_a}(x_a)$. Then

$$\Gamma_J(x) \;=\; \bigoplus_{a=1}^{s} \alpha_a \, \Gamma_{J_a}(x_a), \qquad \alpha_a \;=\; \frac{Z_a}{\sum_b Z_b} \;=\; \operatorname{softmax}\big(\log Z_1, \dots, \log Z_s\big)_a. \tag{8}$$

Consequently

$$\mathscr{B}_J^\circ \;\cong\; \Delta_{s-1}^\circ \times \prod_{a=1}^{s} \mathscr{B}_{J_a}^\circ, \qquad (\alpha, p_1, \dots, p_s) \longmapsto \bigoplus_a \alpha_a p_a,$$

a real-analytic diffeomorphism. A reducible carrier is thus an outer simplicial softmax over the factor log-partition functions, followed by an inner Gibbs chart inside each factor.

Proof. The Jordan product, spectral calculus and trace of J are all factorwise, so $\exp_J(x) = \bigoplus_a \exp_{J_a}(x_a)$ and $\operatorname{tr}_J \exp_J(x) = \sum_a Z_a$. Dividing and inserting Z_a/Z_a in the a-th block gives (8); the softmax form is $e^{\log Z_a} = Z_a$. The displayed map is inverse to $p \mapsto (\alpha, \alpha_1^{-1} p_1, \dots, \alpha_s^{-1} p_s)$ with $\alpha_a = \operatorname{tr}_{J_a}(p_a)$, which is positive on $\mathscr{B}_J^\circ$; both directions are real-analytic. □

Remark 3.1 (What relative shifts between ideals do). Adding t to the a-th block alone multiplies Z_a by e^t, so it moves α by an ordinary softmax logit shift and leaves every inner state $\Gamma_{J_a}(x_a)$ untouched. Only the global shift $x \mapsto x + ae$ is invisible, which is Theorem 3.2(b); relative shifts between ideals are the visible remainder.

4 THE SIMPLICIAL SPECIALIZATION: ORDINARY SOFTMAX

Theorem 4.1 — Softmax is the Gibbs chart of $\mathbb{R}^n$

Let $J_n = \mathbb{R}^n$ with $x \circ y = (x_1 y_1, \dots, x_n y_n)$, unit $e = (1, \dots, 1)$ and trace $\operatorname{tr}_J(x) = \sum_i x_i$. Then $\Omega_{J_n} = \mathbb{R}^n_{>0}$, the open state space $\mathscr{B}_{J_n}^\circ$ is the interior of the probability simplex Δ_{n-1}, and

$$\Gamma_{J_n}(x)_i \;=\; \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \;=\; \operatorname{softmax}(x)_i.$$

Proof. The standard basis is the unique unordered Jordan frame of J_n: each c_i is a primitive idempotent, they are pairwise orthogonal, and they sum to e. So $\operatorname{rank} J_n = n$ and $x = \sum_i x_i c_i$ is already the spectral decomposition (1), spectral calculus being coordinatewise. Squares are $y \circ y = \sum_i y_i^2 c_i$, so the closed cone of squares is $\mathbb{R}^n_{\geq 0}$ with interior $\mathbb{R}^n_{>0}$ and trace-one section the open simplex, and (7) reads $\Gamma_{J_n}(x) = \sum_i e^{x_i} c_i / \sum_j e^{x_j}$. □

Corollary 4.2. $\operatorname{softmax}(x + a\mathbf{1}) = \operatorname{softmax}(x)$ for all $a \in \mathbb{R}$, and this identity says exactly that softmax depends on x only through the ray $\mathbb{R}_{>0} \exp_J(x) \subset \Omega_{J_n}$.

Proof. Theorem 3.2(b) with $e = \mathbf{1}$, read along the right-hand vertical map of Figure 1: $\exp_J$ carries the additive gauge $x \sim x + a\mathbf{1}$ to the multiplicative gauge $\exp_J(x) \sim e^a \exp_J(x)$, and π chooses the trace-one representative on each ray. □

The formula is therefore not an ansatz that happens to work: it is what the carrier produces. Whether some other formula would have worked equally well is the next question.

5 SHIFT INVARIANCE CHARACTERIZES THE GIBBS FAMILY

Theorem 4.1 identifies softmax with an object the carrier already possessed. The next theorem says the identification is rigid, and on every carrier at once: within the class of separable spectral normalizers, shift invariance forces the exponential family, with a single real parameter as the only residual freedom. For carriers of rank at least two, common-score-shift invariance characterizes, among continuous separable spectral normalizers, the one-parameter Gibbs family $\Gamma_J(\beta\,\cdot)$; order preservation selects $\beta > 0$, and a temperature convention selects $\beta = 1$. The rigidity is conditional on that class, and Remark 5.2 draws the line precisely; inside it, the only thing to tune is temperature.

For continuous $\varphi : \mathbb{R} \to \mathbb{R}_{>0}$, define its *spectral lift* $\varphi_J(x) = \sum_k \varphi(\lambda_k)c_k$ and the associated *separable spectral normalizer* $\mathcal{N}_{\varphi,J}(x) = \varphi_J(x)/\operatorname{tr}_J \varphi_J(x)$.

Theorem 5.1 — Shift invariance forces the Gibbs temperature family

Let J be a Euclidean Jordan algebra of rank $r \geq 2$ and let $\varphi : \mathbb{R} \to \mathbb{R}_{>0}$ be continuous. If

$$\mathcal{N}_{\varphi,J}(x + ae) \;=\; \mathcal{N}_{\varphi,J}(x) \qquad \text{for all } x \in J,\ a \in \mathbb{R},$$

then there exist $C > 0$ and $\beta \in \mathbb{R}$ with $\varphi(s) = Ce^{\beta s}$, and consequently $\mathcal{N}_{\varphi,J} = \Gamma_J(\beta\,\cdot)$. The constant C is invisible; the residual scalar β is the Gibbs parameter, and every $\beta \in \mathbb{R}$ occurs.

Proof. **Step 1** (*extract a coefficient ratio*)**.** Fix a Jordan frame $(c_1, \dots, c_r)$, whose members satisfy $\langle c_i, c_j\rangle = \delta_{ij}$. For $s, t \in \mathbb{R}$ put $x = sc_1 + tc_2$, with remaining eigenvalues 0. Pairing with the fixed frame members,

$$\big\langle \mathcal{N}_{\varphi,J}(x),\, c_1\big\rangle \big/ \big\langle \mathcal{N}_{\varphi,J}(x),\, c_2\big\rangle \;=\; \frac{\varphi(s)}{\varphi(t)},$$

the common denominator cancelling. Since $x + ae$ is diagonalized by the same frame with eigenvalues shifted by a, the hypothesis gives $\varphi(s + a)/\varphi(t + a) = \varphi(s)/\varphi(t)$ for all s, t, a.

Step 2 (*a multiplicative Cauchy equation*)**.** Set $t = 0$: $\varphi(s + a) = \varphi(s)\varphi(a)/\varphi(0)$. Put $g(s) = \varphi(s)/\varphi(0)$, continuous and positive with $g(0) = 1$ and $g(s + a) = g(s)g(a)$.

Step 3 (*solve it*)**.** $h = \log g$ is continuous and additive, hence $\mathbb{R}$-linear: $h(s) = \beta s$. Therefore $\varphi(s) = \varphi(0)e^{\beta s}$.

Step 4 (*normalize*)**.** Substituting, the constant cancels, and by (7) the result is $\Gamma_J(\beta x)$. □

Corollary 5.2 (Barycentric calibration selects unit temperature). For $h \in J_0$,

$$D\Gamma_J(0)[h] \;=\; \tfrac{1}{r}h, \qquad \text{hence} \qquad D\big(\Gamma_J(\beta\,\cdot)\big)(0)[h] \;=\; \tfrac{\beta}{r}h.$$

Consequently, if a separable spectral normalizer $\mathcal{N}$ satisfies the hypotheses of Theorem 5.1 together with the calibration axiom

$$D\mathcal{N}(0)\big|_{J_0} \;=\; \tfrac{1}{r}\,\mathrm{id}_{J_0}, \tag{9}$$

then $\beta = 1$ and $\mathcal{N} = \Gamma_J$ exactly.

Proof. Let $h \in J_0$. Analyticity of $\exp_J$ gives $\exp_J(th) = e + th + O(t^2)$, and $\operatorname{tr}_J h = 0$ gives $\operatorname{tr}_J \exp_J(th) = r + O(t^2)$, so

$$\Gamma_J(th) \;=\; \frac{e}{r} + \frac{t}{r}\,h + O(t^2)$$

and $D\Gamma_J(0)[h] = h/r$. The chain rule gives the β-scaled version, and Theorem 5.1 says every admissible $\mathscr{N}$ is of that form, so (9) forces $\beta = 1$. □

Equivalently, (9) asks that the Hessian metric of Corollary 7.11 at the barycenter e/r be the canonical trace metric scaled by $1/r$, since $g_{e/r}(h,k) = \langle h, D\Gamma_J(0)k\rangle = 1/r\langle h,k\rangle$; that reading selects the same branch.

The exact statement is therefore in two stages: *common-shift invariance determines the Gibbs temperature family; barycentric first-order calibration determines the unit-temperature Gibbs chart.*

Remark 5.1 (Rank one, and how little the theorem needs). If rank $J = 1$ then $J = \mathbb{R}e$, $\mathscr{B}_J^\circ = \{e\}$, and every positive separable normalizer is the same constant map $x \mapsto e$. Shift invariance therefore has no classificatory force in rank one, and the hypothesis $r \geq 2$ in Theorem 5.1 is exactly the condition under which there is something to classify.

At the other end, the proof consumes very little. Step 1 evaluates only on the flat $\mathbb{R}c_1 \oplus \mathbb{R}c_2$ spanned by two orthogonal primitive idempotents, so shift invariance on one fixed rank-two diagonal slice already forces the exponential family globally. Continuity may be replaced by any regularity hypothesis excluding the pathological additive solutions of Cauchy's equation — measurability, or local boundedness on a set of positive measure — since Step 3 needs only that h be additive and non-pathological.

Corollary 5.3 (The order-preserving branch). In the setting of Theorem 5.1: $\beta = 0$ gives the constant map $x \mapsto e/r$; nonconstancy, equivalently surjectivity onto $\mathscr{B}_J^\circ$, forces $\beta \neq 0$; and strict preservation of score order forces $\beta > 0$. Under the thermodynamic identification $x = -H$ for an energy observable H, the parameter β is inverse temperature: $\beta > 0$ the positive-temperature branch, $\beta = 0$ the uniform infinite-temperature state, $\beta < 0$ the negative-temperature, population-inverted branch.

Proof. For $\beta = 0$ every weight is $1/r$. For $\beta \neq 0$, precomposition with the invertible scaling $x \mapsto \beta x$ leaves Theorem 3.2(c) intact, so the map is surjective and nonconstant. The weight on c_i is $e^{\beta\lambda_i}/\sum_k e^{\beta\lambda_k}$, strictly increasing in λ_i exactly when $\beta > 0$. □

Classification and branch selection are kept apart deliberately: order preservation is a property of ordinary attention, not a hypothesis of the functional equation, and the $\beta \leq 0$ members are states of the same family.

Remark 5.2 (What the hypothesis excludes). Separability is the substantive restriction, and shift-invariant surjections onto the simplex that are not separable exist. The Euclidean projection $\operatorname{sparsemax}(z) = \arg\min_{p\in\Delta} \|p - z\|^2$ [29] is shift-invariant — on Δ the cross term $-2a\langle \mathbf{1}, p\rangle$ is constant in p, so the minimizer does not move — and its coordinates depend on a threshold computed from the whole vector, not on z_i alone. The α-entmax family [33] interpolates between it and softmax; in the dual description of Section 6 these arise by replacing Shannon entropy

with Tsallis entropy [44] in the variational problem, and they map onto the *closed* simplex, which Γ_J by construction does not. A normalized positive kernel $z \mapsto \kappa(z)/\sum_i \kappa(z)_i$ is simplex-valued for any positive κ; what Theorem 5.1 adds is that among the *separable coordinatewise* kernels $\kappa(z)_i = \varphi(z_i)$, only the exponential family is shift-invariant. Nonseparable positive kernels are not covered by the theorem and may be shift-invariant for other reasons, as sparsemax is.

Rigidity of this kind usually means the object solves a variational problem, and this one does.

6 VARIATIONAL DUALITY AND ZERO-TEMPERATURE FACES

The variational characterization below is not new as an identity: it is the symmetric-cone negative entropy underlying symmetric-cone multiplicative weights [14, 6]. What follows is a self-contained proof, chosen for its equality analysis, which is what the geometry of Sections 7 and 9 needs and which spares the reader an excursion into the optimization literature.

The Gibbs chart is the gradient of a convex potential, and the potential is the Legendre transform of the spectral negative entropy. A single overlap-matrix argument gives convexity, the gradient identity, uniqueness of the maximizer, and the conjugate, with no differentiation of Jordan power series.

With $h(t) = t\log t$ for $t > 0$ and $h(0) = 0$, define the *spectral negative entropy*

$$\eta_J(p) \;=\; \operatorname{tr}_J h(p) \;=\; \sum_k \mu_k \log \mu_k, \qquad p = \sum_k \mu_k d_k \in \mathscr{B}_J.$$

Its negative $S_J(p) = -\eta_J(p)$ is the Jordan entropy; for $J = \mathrm{Herm}_r(\mathbb{C})$ it is the von Neumann entropy and for $J = \mathbb{R}^n$ the Shannon entropy.

We use one standard fact about the cone, the complementarity condition familiar from symmetric-cone programming.

Lemma 6.1 (Cone orthogonality). If $a, b \in \overline{\Omega}_J$ and $\langle a, b\rangle = 0$, then $a \circ b = 0$.

Proof. **Step 1** (*a single primitive idempotent*)**.** Let f be a primitive idempotent and write $b = y^{\circ 2}$ with $y = b^{1/2} \in \overline{\Omega}_J$, which exists by spectral calculus. Associativity (2) of the trace form gives

$$0 \;=\; \langle f, b\rangle \;=\; \langle f, y \circ y\rangle \;=\; \langle L_f y,\, y\rangle.$$

The operator L_f is self-adjoint with spectrum in $\{0, 1/2, 1\}$, hence positive semidefinite, so $\langle L_f y, y\rangle = 0$ forces $L_f y = 0$, that is $y \in V_0(f)$. Since $V_0(f)$ is a Jordan subalgebra [17, Prop. IV.1.1], also $b = y^{\circ 2} \in V_0(f)$, and therefore $f \circ b = 0$.

Step 2 (*the general case*)**.** Write $a = \sum_i \alpha_i f_i$ with $\alpha_i > 0$ on the support and (f_i) orthogonal primitive idempotents. Self-duality of the cone gives $\langle f_i, b\rangle \geq 0$ for every i, so $0 = \langle a, b\rangle = \sum_i \alpha_i \langle f_i, b\rangle$ is a sum of nonnegative terms and each vanishes. By Step 1, $f_i \circ b = 0$ on the support of a, whence $a \circ b = \sum_i \alpha_i\, f_i \circ b = 0$. □

Remark 6.1. Positivity of b is used, not decorative. The quadratic representation $P(f)$ annihilates the whole Peirce $1/2$-space $V_{1/2}(f)$, so $P(f)b = 0$ does not by itself give $f \circ b = 0$: any nonzero $b \in V_{1/2}(f)$ has $P(f)b = 0$ and $f \circ b = {}^1\!/_2 b \neq 0$. Such a b is of course not in the closed cone.

The next definition names an object that the proof below produces and that Sections 7 and 9 use again: it is the elementary carrier of *relative-frame* information between two spectrally decomposed elements.

Definition 6.2 (Overlap matrix). Let $x = \sum_i \alpha_i c_i$ and $y = \sum_j \beta_j d_j$ be spectral decompositions in J, with respect to (possibly different) Jordan frames c and d. Their *overlap matrix* is $M_{ij} = \langle c_i, d_j \rangle$.

Lemma 6.3 (The overlap matrix is doubly stochastic). M is doubly stochastic, and for all $f, g : \mathbb{R} \to \mathbb{R}$,

$$\big\langle f(x), g(y) \big\rangle \;=\; \sum_{i,j} f(\alpha_i) g(\beta_j) M_{ij} \;=\; f(\alpha)^{\top} M\, g(\beta). \tag{10}$$

M is a permutation matrix exactly when the two frames coincide up to order; for $J \cong \mathbb{R}^r$ that is the only possibility, and whenever $\dim \mathfrak{F}_J > 0$ the matrix M varies continuously with the relative frame.

Proof. The cone is self-dual and $c_i, d_j \in \overline{\Omega}_J$, so $M_{ij} \geq 0$; and $\sum_j M_{ij} = \langle c_i, e \rangle = \operatorname{tr}_J c_i = 1$, symmetrically for columns. The pairing (10) is bilinearity of $\langle \cdot, \cdot \rangle$ applied to $f(x) = \sum_i f(\alpha_i) c_i$ and $g(y) = \sum_j g(\beta_j) d_j$. A doubly stochastic matrix is a permutation matrix exactly when every entry is 0 or 1; since $\langle c_i, d_j \rangle = 1$ with both of unit norm forces $c_i = d_j$ by the equality case of Cauchy–Schwarz, this happens exactly when the frames agree up to order. In $\mathbb{R}^r$ the frame is unique up to order (Theorem 7.4), so M is always a permutation matrix. □

Theorem 6.4 — Gibbs variational theorem

For every Euclidean Jordan algebra J and every $x \in J$,

$$\Psi_J(x) \;:=\; \log \operatorname{tr}_J\big(\exp_J x\big) \;=\; \max_{p \in \mathscr{B}_J} \big\{ \langle x, p \rangle - \eta_J(p) \big\},$$

and the maximizer is unique and equals $p = \Gamma_J(x)$. Consequently Ψ_J is convex on J, strictly convex on J_0, real-analytic, and $\nabla \Psi_J(x) = \Gamma_J(x)$ with respect to the trace inner product. The Legendre–Fenchel conjugate [37, §12] is

$$\Psi_J^*(p) \;=\; \begin{cases} \eta_J(p), & p \in \mathscr{B}_J, \\ +\infty, & p \notin \mathscr{B}_J. \end{cases}$$

Proof. The only obstruction is that x and p may be diagonal in different Jordan frames; the overlap matrix converts that mismatch into a doubly stochastic matrix, after which the inequality is classical.

Step 1 (*the overlap matrix*)**.** Write $x = \sum_i \lambda_i c_i$ and $p = \sum_j \mu_j d_j$ in possibly different frames and let M be their overlap matrix (Definition 6.2). By Lemma 6.3 it is doubly stochastic; in particular each column $M_{\cdot j}$ is a probability vector on the index set of the frame of x.

Step 2 (*regroup the objective*)**.** Put $Z = \sum_i e^{\lambda_i}$ and $w_i = e^{\lambda_i}/Z$, so that $\lambda_i = \log w_i + \log Z$. Using $\sum_i M_{ij} = 1$ and $\sum_j \mu_j = 1$,

$$\langle x, p\rangle \;=\; \sum_{i,j} \lambda_i \mu_j M_{ij} \;=\; \log Z \;+\; \sum_j \mu_j \sum_i M_{ij} \log w_i,$$

and therefore

$$\langle x, p\rangle - \eta_J(p) \;=\; \log Z + \sum_j \mu_j \Big[\sum_i M_{ij} \log w_i - \log \mu_j \Big]. \tag{11}$$

Step 3 (*two applications of Jensen*)**.** Set $\widetilde{w}_j = \sum_i M_{ij} w_i$. Then $\widetilde{w}$ is again a probability vector, since $\sum_j \widetilde{w}_j = \sum_i w_i \sum_j M_{ij} = \sum_i w_i = 1$. Concavity of log against the probability vector $M_{\cdot j}$ gives $\sum_i M_{ij} \log w_i \le \log \widetilde{w}_j$, so (11) yields

$$\langle x, p\rangle - \eta_J(p) \;\le\; \log Z + \sum_j \mu_j \log \frac{\widetilde{w}_j}{\mu_j} \;=\; \log Z - D_{\mathrm{KL}}(\mu \,\|\, \widetilde{w}) \;\le\; \log Z \;=\; \Psi_J(x)$$

for every $p \in \mathscr{B}_J$.

Step 4 (*equality forces the spectral projections to coincide*)**.** Suppose equality holds. The second inequality gives $\mu = \widetilde{w}$. Every $w_i > 0$ and every column of M sums to 1, so $\widetilde{w}_j = \sum_i M_{ij} w_i > 0$ for every j; hence $\mu_j > 0$ for every j, and no index is exceptional in what follows. In the first inequality log is *strictly* concave, so for every j the probability vector $M_{\cdot j}$ is supported where w_i is constant. Let $v_1 > \cdots > v_s$ be the distinct values of w, let $B_k = \{i : w_i = v_k\}$ and $P_k = \sum_{i \in B_k} c_i$, let $k(j)$ be the block containing the support of $M_{\cdot j}$, and let $Q_k = \sum_{j\,:\,k(j)=k} d_j$. Then $\mu_j = \widetilde{w}_j = v_{k(j)}$, so $p = \sum_k v_k Q_k$.

For $k \neq l$, $\langle P_k, Q_l\rangle = \sum_{i \in B_k} \sum_{j\,:\,k(j)=l} M_{ij} = 0$, and both elements lie in $\overline{\Omega}_J$; by Lemma 6.1 this forces $P_k \circ Q_l = 0$. Since $\sum_l P_l = e$,

$$Q_k \;=\; Q_k \circ e \;=\; \sum_l Q_k \circ P_l \;=\; Q_k \circ P_k,$$

and symmetrically $P_k = P_k \circ Q_k$, so $P_k = Q_k$ for every k. Hence $p = \sum_k v_k P_k = \sum_i w_i c_i = \Gamma_J(x)$ by (7): the maximizer exists, is unique, and is the Gibbs chart value.

Step 5 (*convexity, strict convexity, and the gradient*)**.** As a supremum of affine functions of x, Ψ_J is convex; it is real-analytic because $\exp_J$ and tr_J are and $Z > 0$. A finite convex function whose subdifferential at each point contains the unique maximizing state, and which is differentiable by analyticity, has $\nabla \Psi_J(x) = \Gamma_J(x)$. If Ψ_J were affine on a nondegenerate segment in J_0, its gradient would be constant there, so Γ_J would take one value at two distinct points of J_0, contradicting the injectivity in Theorem 3.2(c). Hence Ψ_J is strictly convex on J_0; it is affine along $\mathbb{R}e$, so not on all of J.

Step 6 (*the conjugate on the state space*)**.** For $p \in \mathscr{B}_J^\circ$, taking $x = \log_J p$ gives $\langle x, p\rangle - \Psi_J(x) = \mathrm{tr}_J(p \circ \log_J p) - \log \mathrm{tr}_J\, p = \eta_J(p)$, so with Step 2 the conjugate equals $\eta_J(p)$ there. For $p = \sum_i \mu_i c_i$ on the boundary of $\mathscr{B}_J$, take $x_\varepsilon = \sum_i \log(\mu_i + \varepsilon)\, c_i$ for $\varepsilon > 0$. Then $\exp_J(x_\varepsilon) = \sum_i (\mu_i + \varepsilon) c_i$, so $\Psi_J(x_\varepsilon) = \log(1 + r\varepsilon)$, and

$$\langle x_\varepsilon, p\rangle - \Psi_J(x_\varepsilon) \;=\; \sum_i \mu_i \log(\mu_i + \varepsilon) - \log(1 + r\varepsilon) \;\xrightarrow[\varepsilon \downarrow 0]{}\; \sum_i \mu_i \log \mu_i \;=\; \eta_J(p),$$

termwise (the sum is finite, and $0 \log 0 = 0$). Together with Step 2 this gives $\Psi_J^*(p) = \eta_J(p)$ on all of $\mathscr{B}_J$.

Step 7 (*the conjugate off the state space*)**.** If $\operatorname{tr}_J p \neq 1$, take $x = ae$: $\langle x, p\rangle - \Psi_J(x) = a(\operatorname{tr}_J p - 1) - \log r$ is unbounded above in a. If $\operatorname{tr}_J p = 1$ but $p \notin \overline{\Omega}_J$, pick a spectral decomposition $p = \sum \mu_k d_k$ with $\mu_{k_0} < 0$ and take $x = -t\, d_{k_0}$: then $\langle x, p\rangle = -t\mu_{k_0} \to +\infty$ while $\Psi_J(-t\, d_{k_0}) = \log(r - 1 + e^{-t})$ stays bounded. □

Remark 6.2 (Why not Birkhoff). A doubly stochastic overlap matrix invites the classical route: Birkhoff's theorem [12] writes M as a convex combination of permutations, and the rearrangement inequality [20, Thm. 368] then bounds $\langle x, p\rangle$ by the similarly-ordered pairing. That yields the same inequality, but its equality case only says that some permutation is optimal, which does not by itself identify p. The two applications of Jensen above give the bound and the equality case together, and the equality case is what Lemma 6.1 converts into $P_k = Q_k$.

Remark 6.3. Ψ_J is the spectral lift of the symmetric function $\mathrm{LSE}(\lambda) = \log \sum_i e^{\lambda_i}$ — a joint function of the whole spectrum, not a separable sum $\sum_i f(\lambda_i)$. The transfer of convexity from symmetric functions of eigenvalues to spectral functions goes back to Davis [15] for Hermitian matrices, with the modern conjugacy formulation due to Lewis [27]; Baes [4] carries the theory to arbitrary Euclidean Jordan algebras. The proof above makes these citations corroborative rather than load-bearing.

The deficit in the variational bound is a divergence.

Corollary 6.5 (Jordan relative entropy). For $p \in \mathscr{B}_J$ and $q \in \mathscr{B}_J^\circ$ define

$$D_J(p \,\|\, q) \;=\; \eta_J(p) - \langle p, \log_J q\rangle.$$

Then $D_J(p\|q) \geq 0$, with equality if and only if $p = q$, and for every $x \in J$

$$\Psi_J(x) - \langle x, p\rangle + \eta_J(p) \;=\; D_J\big(p \,\|\, \Gamma_J(x)\big).$$

For $J = \mathbb{R}^n$ this is the Kullback–Leibler divergence, and for $J = \mathrm{Herm}_r(\mathbb{C})$ the Umegaki quantum relative entropy [45].

Proof. Take $x = \log_J q$ in Theorem 6.4. Since $\operatorname{tr}_J q = 1$ we have $\Psi_J(\log_J q) = \log \operatorname{tr}_J q = 0$, so the theorem reads $0 \geq \langle \log_J q, p\rangle - \eta_J(p)$, which is $D_J(p\|q) \geq 0$; equality holds exactly at the unique maximizer $p = \Gamma_J(\log_J q) = q$. For the identity, write $P = \Gamma_J(x) = \exp_J(x)/Z$, so $\log_J P = x - (\log Z)e$ and

$$D_J(p\|P) = \eta_J(p) - \langle p, x\rangle + (\log Z)\operatorname{tr}_J p = \eta_J(p) - \langle x, p\rangle + \Psi_J(x),$$

using $\operatorname{tr}_J p = 1$ and $\Psi_J(x) = \log Z$. □

Remark 6.4 (Bregman form). On the affine state space, $D_J(p \,\|\, q)$ is the Bregman divergence generated by the spectral negative entropy η_J. The ambient gradient of η_J is $\nabla\eta_J(q) = \log_J q + e$, so for $p, q \in \mathscr{B}_J^\circ$

$$\eta_J(p) - \eta_J(q) - \big\langle p - q,\ \nabla\eta_J(q)\big\rangle \;=\; \eta_J(p) - \eta_J(q) - \langle p - q, \log_J q\rangle \;=\; D_J(p \,\|\, q),$$

the e-term dropping because $\langle p - q, e\rangle = \operatorname{tr}_J p - \operatorname{tr}_J q = 0$ and the $\eta_J(q)$ terms cancelling against $\langle q, \log_J q\rangle = \eta_J(q)$. This places Γ_J and η_J in the standard dually flat picture, and makes the

mirror-descent reading immediate, without importing any further theorem. For Hermitian matrix carriers the same generator $t \log t$ produces the Bregman representation of the relative operator entropy studied by Petz [36]; the displayed calculation establishes the corresponding statement on an arbitrary Euclidean Jordan carrier directly, without passing through an associative model.

Corollary 6.6 (Thermodynamic form). Let $H \in J$, $\beta \in \mathbb{R}$, and $\rho_{\beta,H} = \Gamma_J(-\beta H)$. Then

$$\rho_{\beta,H} \;=\; \operatorname*{arg\,max}_{p \in \mathscr{B}_J} \big\{ S_J(p) - \beta \langle H, p \rangle \big\},$$

uniformly in the sign of β — the maximum-entropy formulation of statistical mechanics [23], transported to the carrier J. Writing $A(\beta) = \Psi_J(-\beta H) = \log \operatorname{tr}_J(e^{-\beta H})$ for the log-partition, or Massieu, potential and $\operatorname{Var}_\rho(H) = \langle H^{\circ 2}, \rho \rangle - \langle H, \rho \rangle^2$, one has $-A'(\beta) = \langle H, \rho_{\beta,H} \rangle =: U$, $A''(\beta) = \operatorname{Var}_{\rho_{\beta,H}}(H) \geq 0$, and $S_J(\rho_{\beta,H}) = A(\beta) + \beta U$; for $\beta \neq 0$ the corresponding Helmholtz free energy is $\mathscr{F}(\beta) = -\beta^{-1} A(\beta)$. For each fixed $\beta \neq 0$ the map $[H] \mapsto \rho_{\beta,H}$ is a diffeomorphism $J/\mathbb{R}e \to \mathscr{B}_J^\circ$; at $\beta = 0$ it collapses to e/r. Since the state depends only on βH modulo $\mathbb{R}e$, a parameterized theory must fix an energy scale or quotient that redundancy.

Proof. Substitute $x = -\beta H$ in Theorem 6.4 and rearrange; the identities follow by differentiating A using $\nabla \Psi_J = \Gamma_J$, the variance appearing as the second derivative of a log-partition function. The diffeomorphism statement is Theorem 3.2(c) precomposed with $H \mapsto -\beta H$. □

Corollary 6.7 (Zero temperature selects exposed-face barycenters). For $x \in J$ let $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$ be the extremal eigenvalues, $E_{\max}$ and $E_{\min}$ their spectral idempotents, and $m_{\max} = \operatorname{tr}_J E_{\max}$, $m_{\min} = \operatorname{tr}_J E_{\min}$ their multiplicities. Define the exposed faces

$$F_x^+ = \big\{ p \in \mathscr{B}_J \,:\, \langle x, p \rangle = \lambda_{\max}(x) \big\}, \qquad F_x^- = \big\{ p \in \mathscr{B}_J \,:\, \langle x, p \rangle = \lambda_{\min}(x) \big\}.$$

Then

$$F_x^+ = \mathscr{B}_J \cap V_1(E_{\max}), \qquad F_x^- = \mathscr{B}_J \cap V_1(E_{\min}),$$

and each limit below exists and is the unique entropy maximizer of its face:

$$\lim_{\beta \to +\infty} \Gamma_J(\beta x) \;=\; \operatorname*{arg\,max}_{p \in F_x^+} S_J(p) \;=\; \frac{E_{\max}}{m_{\max}}, \qquad \lim_{\beta \to -\infty} \Gamma_J(\beta x) \;=\; \operatorname*{arg\,max}_{p \in F_x^-} S_J(p) \;=\; \frac{E_{\min}}{m_{\min}}.$$

If $x \notin \mathbb{R}e$ the faces are proper and both limits lie on $\partial \mathscr{B}_J$, which Γ_J itself never reaches; if $x \in \mathbb{R}e$ both faces are all of $\mathscr{B}_J$ and both limits equal the interior barycenter e/r. For $J = \mathbb{R}^n$ the positive limit is the uniform distribution on the set of maximizing coordinates — the Jordan form of arg max — and for matrix and exceptional carriers it is the normalized projection onto the extremal eigenspace.

Proof. **Step 1** (*the face is a Peirce state space*)**.** Write $\lambda_{\max} = \lambda_{\max}(x)$ and $y = \lambda_{\max} e - x$, whose eigenvalues $\lambda_{\max} - \lambda_k$ are all nonnegative, so $y \in \overline{\Omega}_J$. Self-duality of the cone gives $\langle y, p \rangle \geq 0$ for every $p \in \mathscr{B}_J$, that is $\langle x, p \rangle \leq \lambda_{\max} \operatorname{tr}_J p = \lambda_{\max}$, with equality exactly when $\langle y, p \rangle = 0$; by Lemma 6.1 that is equivalent to $y \circ p = 0$. So the maximum of the linear functional $\langle x, \cdot \rangle$ on $\mathscr{B}_J$ is $\lambda_{\max}$ and F_x^+ is the face it exposes.

The element y has strictly positive coefficient on every spectral idempotent of x except $E_{\max}$, so its support idempotent is $f = e - E_{\max}$: write $y = \sum_i \alpha_i f_i$ with $\alpha_i > 0$ and (f_i) orthogonal primitive idempotents summing to f. For $p \in \overline{\Omega}_J$ self-duality makes every $\langle f_i, p\rangle \geq 0$, so $\langle y, p\rangle = 0$ forces each of them to vanish, and Step 1 of the proof of Lemma 6.1 then gives $f_i \circ p = 0$ for every i, hence $f \circ p = 0$ and $p \in V_0(f) = V_1(E_{\max})$. Conversely $V_1(f) \ni y$ and $V_1(E_{\max})$ are trace-orthogonal Peirce spaces of f, so every $p \in \mathscr{B}_J \cap V_1(E_{\max})$ satisfies $\langle y, p\rangle = 0$. This is $F_x^+ = \mathscr{B}_J \cap V_1(E_{\max})$.

Step 2 (*entropy inside the face*)**.** $V_1(E_{\max})$ is a Euclidean Jordan subalgebra with unit $E_{\max}$ and rank $m_{\max}$ [17, Prop. IV.1.1, Thm. IV.2.1]; its cone is $\overline{\Omega}_J \cap V_1(E_{\max})$ and its trace is the restriction of tr_J, so F_x^+ is exactly the closed state space of that rank-$m_{\max}$ algebra. An element of F_x^+ has the same nonzero eigenvalues there as in J, and $h(0) = 0$, so η_J restricts to its spectral negative entropy. Applying Theorem 6.4 inside $V_1(E_{\max})$ at score 0 gives a unique maximizer of S_J on F_x^+, namely $\Gamma_{V_1(E_{\max})}(0) = E_{\max}/m_{\max}$.

Step 3 (*the limits*)**.** By shift invariance, Theorem 3.2(b), replace x by $x - \lambda_{\max} e$, so that the distinct eigenvalues are $0 = \theta_1 > \theta_2 > \cdots > \theta_s$. By (7),

$$\Gamma_J(\beta x) \;=\; \frac{E_{\max} + \sum_{a\geq 2} e^{\beta\theta_a} E_a}{m_{\max} + \sum_{a\geq 2} m_a e^{\beta\theta_a}},$$

and every $e^{\beta\theta_a} \to 0$ as $\beta \to +\infty$ because $\theta_a < 0$. The second limit is the first applied to $-x$, whose maximal face is F_x^-.

Step 4 (*proper faces and the scalar case*)**.** If $x \notin \mathbb{R}e$ then x has at least two distinct eigenvalues, so $E_{\max} \neq e$, the limit has 0 in its spectrum and lies on $\partial\mathscr{B}_J$, while Γ_J is valued in $\mathscr{B}_J^\circ$ by Theorem 3.2. If $x = ae$ then $\langle x, p\rangle = a$ for every $p \in \mathscr{B}_J$, so both faces are all of $\mathscr{B}_J$; here $E_{\max} = E_{\min} = e$ and $m_{\max} = m_{\min} = r$, and both limits are e/r. □

One map is thus simultaneously the normalized Jordan exponential, the trace-one representative of a ray in the cone, the symmetric-space exponential chart at the identity, the gradient of a log-partition function, and the entropy-regularized argmax. These are five descriptions of one object, fixed by the carrier together with the Gibbs-chart axioms. Everything so far has held the carrier fixed. We now change it, and watch the state space come apart into two pieces.

7 SPECTRAL WEIGHTS AND PEIRCE-FRAME GEOMETRY

We now separate what Γ_J emits from where it emits it. By (7) the chart produces exactly r scalar weights, one per frame element, while the target $\mathscr{B}_J^\circ$ has dimension $\dim J - 1$. The difference is carried by the frame. This section counts it, characterizes when it vanishes, determines what invariant content it holds, equips it with a metric, and computes the determinant of its differential.

Equation (7) has a short reading in spectral–frame coordinates. On the simple-spectrum locus write $x = \sum_k \lambda_k c_k$; then

$$\Gamma_J : (\lambda, c) \;\longmapsto\; \big(\mathrm{softmax}(\lambda),\, c\big), \tag{12}$$

so the chart applies ordinary softmax to the eigenvalues and carries the Jordan frame unchanged. What varies with the carrier is not the scalar law but the geometry of the frame to which its weights are attached. In $\mathbb{R}^3$ the weights are coordinates; in $H_3(\mathbb{O})$ they are only eigenvalues. The two-row comparison in Figure 2 fixes the rank at 3 so that the only variable is the carrier.

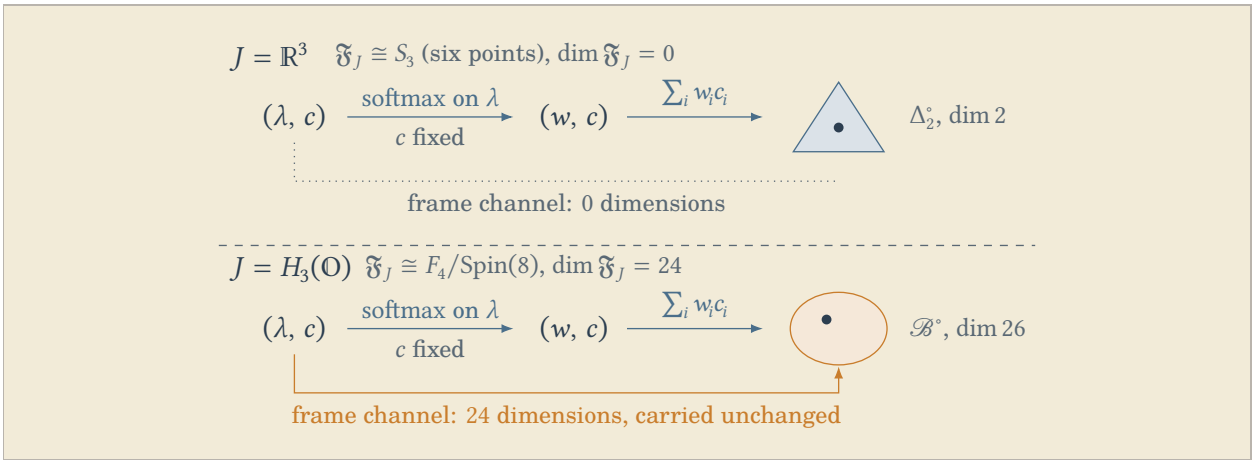


FIG. 2. *The Gibbs chart in spectral–frame coordinates at fixed rank* 3: $(\lambda, c) \mapsto (\mathrm{softmax}\,\lambda,\ c) \mapsto \sum_i w_i c_i$. *Both rows apply the same three-way softmax to the eigenvalue channels and carry the Jordan frame unchanged; only the frame space differs.*

For $J = \mathbb{R}^3$ the frame space is the finite set S_3, so the two independent weights already exhaust Δ_2°. For $J = H_3(\mathbb{O})$ it is $F_4/\mathrm{Spin}(8)$ of dimension 24, and the three idempotents move as one constrained triple rather than independently.

Proposition 7.1 (Equivariance). For every $g \in \mathrm{Aut}(J)$ and $x \in J$, $\Gamma_J(gx) = g\,\Gamma_J(x)$.

Proof. An automorphism preserves $\circ$, hence all powers, hence $\exp_J$; and it carries frames to frames, hence preserves tr_J. Both numerator and denominator of Γ_J are therefore equivariant, the denominator being a scalar. □

Corollary 7.2. $\mathrm{Aut}(\mathbb{R}^n) = S_n$, so Proposition 7.1 specializes to the permutation equivariance of softmax; and $\mathrm{Aut}(H_3(\mathbb{O})) = F_4$, a compact simple Lie group of dimension 52 acting transitively on frames.

Proof. An automorphism of $\mathbb{R}^n$ permutes the primitive idempotents, which are exactly the n standard basis vectors; every permutation is an automorphism. The identification $\mathrm{Aut}(H_3(\mathbb{O})) = F_4$ is classical [42, 5]. □

Theorem 7.3 — Frame tangents are the off-diagonal Peirce blocks

Let J be any Euclidean Jordan algebra and $c = (c_1, \dots, c_r)$ an ordered Jordan frame. Then

$$T_c\mathfrak{F}_J \;\cong\; \bigoplus_{i<j} J_{ij}, \qquad \text{hence} \qquad \dim \mathfrak{F}_J \;=\; \sum_{i<j} d_{ij} \;=\; \dim J - r,$$

the same at every frame; for frame-transitive J the frame space is a single $\mathrm{Aut}(J)$-orbit and $\dim \mathfrak{F}_J = \binom{r}{2} d$. On the *simple-spectrum locus* $\{\lambda_1 > \cdots > \lambda_r\}$ the frame in (1) is unique, the map $(\lambda, c) \mapsto \sum_k \lambda_k c_k$ is a diffeomorphism onto that locus, and restricting to trace one,

$$\dim \mathscr{B}_J^\circ \;=\; (r-1) + (\dim J - r) \;=\; \dim J - 1,$$

with $r-1$ spectral directions and $\dim J - r$ frame directions.

Proof. **Step 1** (*tangent computation*)**.** Let u_i be an infinitesimal variation of c_i through frames. Linearizing $c_i \circ c_i = c_i$ gives $2c_i \circ u_i = u_i$, so $u_i \in V_{1/2}(c_i) = \bigoplus_{j \neq i} J_{ij}$; write $u_i = \sum_{j \neq i} u_{ij}$ with $u_{ij} \in J_{ij}$. Linearizing $c_i \circ c_j = 0$ for $i \neq j$ gives $u_i \circ c_j + c_i \circ u_j = 0$, which on the block J_{ij} reads $u_{ij} + u_{ji} = 0$. So the solution space of the linearized equations is parameterized by free choices $z_{ij} \in J_{ij}$ for $i < j$, via $u_i = \sum_{j>i} z_{ij} - \sum_{j<i} z_{ji}$. The constraint $\sum_k u_k = 0$ is automatic, since each z_{ij} enters u_i and u_j with opposite signs.

Step 2 (*realization*)**.** Every such solution is realized by an actual frame curve. For $z \in J_{ij}$ the operator $D = 4[L_z, L_{c_i}]$ is an inner derivation of J [17, Ch. II], and the Peirce relations $c_i \circ z = c_j \circ z = {}^1\!/_2 z$, $c_k \circ z = 0$ for $k \neq i, j$ give

$$D(c_i) = z, \qquad D(c_j) = -z, \qquad D(c_k) = 0 \quad (k \neq i, j). \tag{13}$$

Derivations integrate to automorphisms, which carry frames to frames, so $t \mapsto \exp(tD) \cdot c$ is a frame curve with the prescribed initial velocity; composing over pairs realizes the general solution. Hence $T_c\mathfrak{F}_J \cong \bigoplus_{i<j} J_{ij}$, and the dimension formula follows from (5).

Step 3 (*chart and slice*)**.** Uniqueness of the frame at simple spectrum is part of the spectral theorem [17, Ch. III]; the evaluation map is then a smooth bijection onto that open set with everywhere-invertible differential, the image of $(\dot\lambda, z)$ being $\sum_k \dot\lambda_k c_k + \sum_{i<j} (\lambda_i - \lambda_j) z_{ij}$ with nonzero coefficients. Intersecting with $\{\mathrm{tr}_J = 1\}$ removes $\sum \lambda_k = 1$ from the spectral factor and leaves the frame factor untouched. □

Remark 7.1 (Simple spectrum, not invertibility). The product decomposition concerns the *simple-spectrum* locus $\{\lambda_1 > \cdots > \lambda_r\}$. We avoid the word *regular* throughout, because in Jordan theory it standardly means invertible or of maximal rank, which is a different and weaker condition. At a repeated eigenvalue only a partial flag of spectral projections is determined, and at e/r every frame diagonalizes the state. Globally $\mathscr{B}_J^\circ$ is a smooth manifold of dimension $\dim J - 1$; it is not globally a product of a simplex with one frame manifold, and no statement below assumes it is.

Theorem 7.4 — The simplicial carrier is the unique frameless carrier

For a Euclidean Jordan algebra J of rank r, the following are equivalent:

(a) $\dim \mathfrak{F}_J = 0$;

(b) $\dim J = r$;

(c) $J_{ij} = 0$ for all $i < j$ (any, hence every, frame);

(d) $V_{1/2}(c) = 0$ for every primitive idempotent c;

(e) $J \cong \mathbb{R}^r$ as a Euclidean Jordan algebra.

In that case, and only in that case, Γ_J is ordinary softmax and every dimension of $\mathscr{B}_J^\circ$ is a spectral weight. The ordered-frame set of $\mathbb{R}^r$ is the finite set of $r!$ orderings of the standard frame; the unordered frame space is a single point.

Proof. (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) is Theorem 7.3 with (5). (c) $\Leftrightarrow$ (d): $V_{1/2}(c_i) = \bigoplus_{j \neq i} J_{ij}$ for frame members, and every primitive idempotent belongs to some frame. (c) $\Rightarrow$ (e): if all off-diagonal blocks vanish, a frame spans J and $(\sum_i x_i c_i) \circ (\sum_i y_i c_i) = \sum_i x_i y_i c_i$, the componentwise product on $\mathbb{R}^r$. (e) $\Rightarrow$ (d): in $\mathbb{R}^r$ the operator L_{c_i} is the i-th coordinate projection, with spectrum $\{0, 1\}$ and no eigenvalue $1/2$. The final statements are Theorem 4.1 and the observation that automorphisms of $\mathbb{R}^r$ permute the unique unordered frame. □

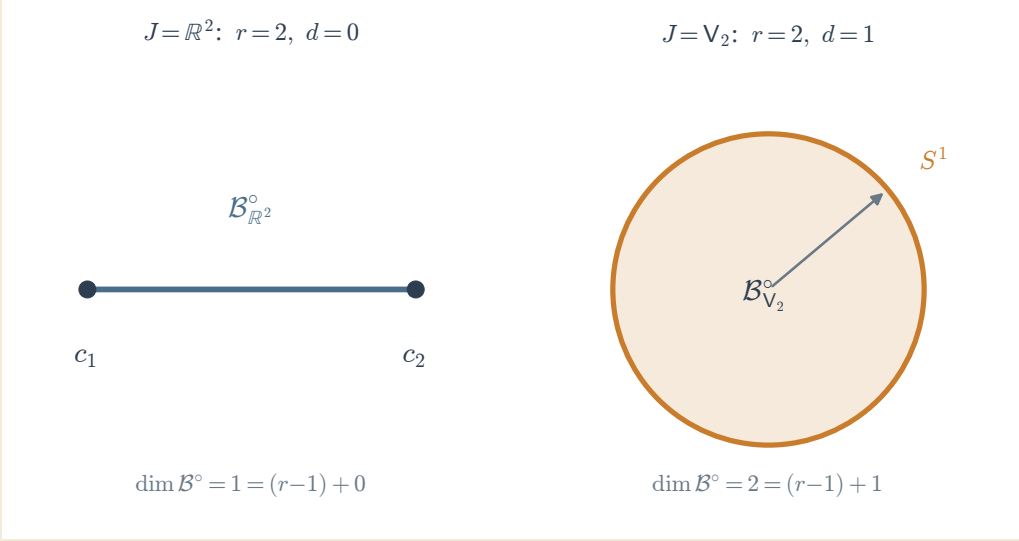


FIG. 3. *Two rank-2 carriers, $\mathbb{R}^2$ and the spin factor $\mathsf{V}_2 = \mathbb{R} \oplus \mathbb{R}^2$, so two weights in each case. On the left the extreme points are two isolated vertices and the interval is entirely weight; on the right they sweep a circle, and the frame contributes one dimension the weights cannot see.*

Proposition 7.5 (Extreme points). Let $\mathscr{B}_J$ be the closed state space of any EJA with $r \geq 2$. Its extreme points are exactly the primitive idempotents of J, each lying in $\partial\Omega_J$, and at each primitive c the tangent space to the primitive-idempotent set is the Peirce $1/2$-space $V_{1/2}(c)$. For simple J the primitive idempotents form a single compact $\mathrm{Aut}(J)$-homogeneous manifold of dimension $(r-1)d$; for $\mathbb{R}^r$ they form one finite transitive S_r-orbit; in general they form a union of $\mathrm{Aut}(J)$-orbits indexed by the isotypic classes of simple ideals, since automorphisms may permute isomorphic ideals.

Proof. Every $p \in \mathscr{B}_J$ is a convex combination $\sum_k \lambda_k c_k$ of primitive idempotents with $\lambda_k \geq 0$, $\sum \lambda_k = 1$; if two λ_k are nonzero, p is a proper convex combination and not extreme. Conversely $\|p\|^2 = \operatorname{tr}_J(p^{\circ 2}) = \sum \lambda_k^2 \leq 1$ on $\mathscr{B}_J$, with equality exactly at primitive idempotents. If a primitive $c = tp + (1-t)q$ with $0 < t < 1$, then by the parallelogram identity

$$1 \;=\; \|c\|^2 \;=\; t\|p\|^2 + (1-t)\|q\|^2 - t(1-t)\|p-q\|^2 \;\leq\; 1 - t(1-t)\|p-q\|^2,$$

forcing $p = q = c$. A primitive idempotent has rank $1 < r$, hence is singular and lies in $\partial\Omega_J$. The defining equations $y \circ y = y$, $\operatorname{tr}_J y = 1$ linearize at c to $2L_c v = v$, identifying the tangent space with $V_{1/2}(c)$, of dimension $(r-1)d$ in the simple case. The orbit statements follow from frame transitivity for frame-transitive J, and in general from the fact that $\mathrm{Aut}(J)$ permutes the simple ideals within each isomorphism class and acts on each class by the wreath-type action; $J = \mathbb{R} \oplus \mathrm{Sym}_2(\mathbb{R})$, whose primitive set is an isolated point together with a copy of $\mathbb{R}P^1$, realizes the disconnected case. □

7.1 ORBIT INVARIANTS AND RELATIVE FRAMES

Proposition 7.6 (Invariant polynomials on a frame-transitive carrier). If J is frame-transitive, then restriction to the diagonal of a fixed frame induces an isomorphism

$$\mathbb{R}[J]^{\mathrm{Aut}(J)} \;\cong\; \mathbb{R}[\lambda_1, \dots, \lambda_r]^{S_r} \;:$$

every $\mathrm{Aut}(J)$-invariant polynomial is a symmetric function of the spectrum.

Proof. Fix a frame $(c_1, \dots, c_r)$ and let $C = \bigoplus_k \mathbb{R}c_k$. By the spectral theorem and frame transitivity, every element of J is $\mathrm{Aut}(J)$-conjugate to an element of C, so restriction to C is injective on invariants. Frame permutations are induced by automorphisms, so the restriction of an invariant is S_r-symmetric. Conversely, the elementary symmetric functions of the eigenvalues are the coefficients of the Jordan characteristic polynomial — polynomial functions on J, invariant under $\mathrm{Aut}(J)$ — and they generate the symmetric polynomials. □

Example 7.1 (The restriction is necessary). Let $J = \mathbb{R} \oplus \mathrm{Sym}_2(\mathbb{R})$, rank 3, dimension 4, and take

$$x = \left(0,\, \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\right), \qquad y = \left(1,\, \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}\right).$$

Both have eigenvalue multiset $\{0, 1, 2\}$. But $P(a, A) = a$ is $\mathrm{Aut}(J)$-invariant, since automorphisms preserve the nonisomorphic simple ideals, and $P(x) = 0 \neq 1 = P(y)$. So $\mathbb{R}[J]^{\mathrm{Aut}(J)} \neq \mathbb{R}[\lambda]^{S_3}$ here, and no automorphism carries x to y despite equal spectra.

The failure can be computed exactly. On an arbitrary carrier the complete one-state invariant is the multiset of *component* spectra within each isotypic family of simple ideals, not the global spectrum.

Proposition 7.7 (Invariants of an arbitrary carrier). Write $J \cong \bigoplus_{a=1}^{s} K_a^{\oplus m_a}$ with the K_a pairwise nonisomorphic simple Euclidean Jordan algebras. Then

$$\mathrm{Aut}(J) \;\cong\; \prod_{a=1}^{s} \big(\mathrm{Aut}(K_a) \wr S_{m_a}\big),$$

and two elements $x, y \in J$ lie in the same $\mathrm{Aut}(J)$-orbit if and only if for every a the unordered multisets $\{\mathrm{spec}(x_{a,1}), \dots, \mathrm{spec}(x_{a,m_a})\}$ and $\{\mathrm{spec}(y_{a,1}), \dots, \mathrm{spec}(y_{a,m_a})\}$ of component spectra agree. Correspondingly

$$\mathbb{R}[J]^{\mathrm{Aut}(J)} \;\cong\; \bigotimes_{a=1}^{s} \left(\mathbb{R}[\lambda^{(1)}]^{S_{r_a}} \otimes \cdots \otimes \mathbb{R}[\lambda^{(m_a)}]^{S_{r_a}} \right)^{S_{m_a}},$$

the multisymmetric functions in the characteristic coefficients of the components, $r_a = \mathrm{rank}\, K_a$.

Proposition 7.6 is the case $s = m_1 = 1$, together with $J \cong \mathbb{R}^r$, where $K_1 = \mathbb{R}$, $\mathrm{Aut}(\mathbb{R}) = 1$ and $\mathrm{Aut}(J) = 1 \wr S_r = S_r$.

Proof. The decomposition of J into simple ideals is unique, and an automorphism carries each simple ideal isomorphically onto a simple ideal; nonisomorphic ideals cannot be exchanged, so $\mathrm{Aut}(J)$ permutes the m_a copies inside each isotypic family and acts on each copy by an automorphism. That is the stated wreath product, and conversely every such element is an automorphism.

For orbits: necessity is immediate, since an automorphism permutes components and preserves each component's spectrum. For sufficiency, permute components inside each isotypic family so that matching spectra are aligned, and then use frame transitivity in each simple K_a [17, Ch. IV] to carry each component of x to the corresponding component of y.

For the invariant ring, $\mathrm{Aut}(J)$ is compact — it is a closed subgroup of the orthogonal group of the trace form — so invariants of a product group acting on a direct sum factorwise are the tensor product of the factors' invariants. Applying Proposition 7.6 to each simple K_a and then taking S_{m_a}-invariants gives the displayed multisymmetric description. In Example 7.1 the two isotypic families are $\{\mathbb{R}\}$ and $\{\mathrm{Sym}_2(\mathbb{R})\}$, and the projection $P(a, A) = a$ is the degree-one invariant of the first. □

Proposition 7.8 (What the weights determine). Let $x \in J$ with weights $w_i = e^{\lambda_i} / \sum_j e^{\lambda_j}$. Then

$$\lambda_i \;=\; \log w_i \;+\; \frac{\mathrm{tr}_J\, x - \sum_j \log w_j}{r},$$

so the weights are exactly the eigenvalues of $p = \Gamma_J(x)$; they determine the spectrum of the gauge class $[x] \in J/\mathbb{R}e$; and together with $\mathrm{tr}_J\, x$ they determine the full spectrum of x. For frame-transitive J, by Proposition 7.6, these spectra exhaust the one-state invariant content of p, of $[x]$, and of x respectively.

Proof. From $\log w_i = \lambda_i - \log Z$, summing over i gives $\sum_j \log w_j = \mathrm{tr}_J\, x - r \log Z$; substitute back. The eigenvalues of p are the w_i by (7). □

Proposition 7.9 (Relative frames contribute tuple invariants). Let J have rank $r \geq 2$. The trace pairing $\langle p, q \rangle$ of two states is invariant under the simultaneous action $(p, q) \mapsto (gp, gq)$ of $\mathrm{Aut}(J)$, and is not in general determined by the two individual spectra. Concretely, fix a frame $(c_1, \dots, c_r)$ and weights

$$\alpha > \beta > 0, \qquad \gamma_k > 0, \qquad \alpha + \beta + \textstyle\sum_{k \geq 3} \gamma_k = 1,$$

so that the three elements

$$p = \alpha c_1 + \beta c_2 + \textstyle\sum_{k\geq 3} \gamma_k c_k, \qquad q = p, \qquad q' = \beta c_1 + \alpha c_2 + \textstyle\sum_{k\geq 3} \gamma_k c_k$$

are interior states of $\mathscr{B}_J^\circ$. Then q and q' have identical spectrum while

$$\langle p, q\rangle - \langle p, q'\rangle \;=\; (\alpha - \beta)^2 \;>\; 0.$$

If J is frame-transitive, then individual spectra classify the one-state orbits, while simultaneous orbits of pairs retain additional relative-frame data.

Proof. Invariance is Proposition 7.1 together with preservation of the trace form. By (10) the pairing is $\langle p, q\rangle = \mu^\top M \nu$ with μ, ν the spectra of p, q and M the overlap matrix of the two frames, so it depends on the two spectra *and* on M; here $M = \mathrm{id}$ for q and M the transposition $(1\,2)$ for q', giving $\langle p, q\rangle - \langle p, q'\rangle = (\alpha^2 + \beta^2) - 2\alpha\beta$. The final sentence is Proposition 7.6 for the one-state claim, and the displayed example for the pair claim. □

On an arbitrary carrier, the complete one-state invariant is the isotypically labeled collection of component spectra,

$$\mathrm{Spec}_{\mathrm{iso}}(x) \;=\; \Big\{ \{\mathrm{spec}(x_{a,1}), \dots, \mathrm{spec}(x_{a,m_a})\} \Big\}_a,$$

which is Proposition 7.7. On a frame-transitive carrier it reduces to the ordinary spectrum, by Proposition 7.6. Covariantly, a simple-spectrum state also determines an equivariant spectral flag: writing $\Delta_{\mathrm{ss}}^\circ$ for the open simplex locus with pairwise distinct weights,

$$p \;\longmapsto\; \big((w_1, \dots, w_r), (c_1, \dots, c_r)\big) \;\in\; \big(\Delta_{\mathrm{ss}}^\circ \times \mathfrak{F}_J\big)/S_r,$$

and (12) says the Gibbs chart acts on it by $(\lambda, c) \mapsto (\mathrm{softmax}\,\lambda, c)$, changing the weights and carrying the frame unchanged. A single state therefore carries no invariant frame coordinate but does carry a canonical equivariant spectral flag; at repeated eigenvalues, replace the primitive frame by the coarse spectral idempotents of Theorem 7.10.

For a single isolated state on a frame-transitive carrier, the frame is pure gauge. For tuples it is not. Lemma 6.3 supplies a canonical first relative-frame invariant, the overlap matrix M, and M is sufficient for every paired spectral functional of the form (10); it need not classify the full simultaneous orbit of two frames, and further relative-frame invariants may remain. Its character differs by carrier. In $\mathbb{R}^r$ every M is a permutation matrix, so the relative datum is the finite set of relative orderings and the displayed example is a transposition. Whenever $\dim \mathfrak{F}_J > 0$ — for a simple carrier, equivalently $d > 0$ — the doubly stochastic M varies continuously, and the frame can be moved with the spectrum held fixed. Proposition 9.21 exploits exactly that difference, and Theorem 9.2 identifies the mechanism that produces it inside a routing construction.

Counting frame dimensions in the state space would mean little if normalization collapsed them. The differential shows it does not, and identifies the resulting bilinear form.

Recall the *logarithmic mean* of $u, v > 0$:

$$\mathscr{L}(u,v) \;=\; \begin{cases} \dfrac{u-v}{\log u - \log v}, & u \neq v, \\ u, & u = v, \end{cases}$$

a continuous, symmetric, strictly positive function with $\min(u,v) \le \mathscr{L}(u,v) \le \max(u,v)$. The second branch is the limit of the first, so $\mathscr{L}$ is continuous across the diagonal; this is what makes the theorem below global rather than generic.

Theorem 7.10 — Spectral and Peirce blocks of the Gibbs differential

Let $x \in J$ be arbitrary, let $\theta_1, \dots, \theta_s$ be its distinct eigenvalues with spectral idempotents $E_1, \dots, E_s$, and put

$$m_a = \operatorname{tr}_J E_a, \qquad Z = \sum_a m_a e^{\theta_a}, \qquad w_a = \frac{e^{\theta_a}}{Z}, \qquad \text{so} \qquad \Gamma_J(x) = \sum_a w_a E_a.$$

Decompose J by the coarse Peirce relations of (E_a),

$$J \;=\; \bigoplus_a J_{aa} \;\oplus\; \bigoplus_{a<b} J_{ab}, \qquad J_{aa} = V_1(E_a), \quad J_{ab} = V_{1/2}(E_a) \cap V_{1/2}(E_b),$$

and write $h = \sum_a h_{aa} + \sum_{a<b} h_{ab}$ accordingly. Then, with $p = \Gamma_J(x)$,

$$D\Gamma_J(x)[h] \;=\; \sum_a w_a h_{aa} \;+\; \sum_{a<b} \mathscr{L}(w_a, w_b)\, h_{ab} \;-\; p \sum_a w_a \operatorname{tr}_J(h_{aa}). \tag{14}$$

In particular, when the spectrum is simple, $E_a = c_a$ and (14) reads, relative to a frame diagonalizing x and with $h = \sum_i h_i c_i + \sum_{i<j} h_{ij}$,

$$D\Gamma_J(x)[h] \;=\; \sum_i w_i \Big(h_i - \sum_k w_k h_k \Big) c_i \;+\; \sum_{i<j} \frac{w_i - w_j}{\lambda_i - \lambda_j}\, h_{ij}, \tag{15}$$

where in the second sum

$$\frac{w_i - w_j}{\lambda_i - \lambda_j} \;=\; \frac{w_i - w_j}{\log w_i - \log w_j} \;=\; \mathscr{L}(w_i, w_j).$$

The frame coefficient is a divided difference of the eigenvalues, and it is exactly the logarithmic mean of the corresponding Gibbs weights, because $\lambda_i - \lambda_j = \log w_i - \log w_j$.

Proof. **Step 1** (*diagonal blocks*)**.** Fix a. The subspace $J_{aa} = V_1(E_a)$ is a Jordan subalgebra with unit E_a, and J contains the orthogonal direct sum $\bigoplus_b V_1(E_b)$ of subalgebras. Since $x = \sum_b \theta_b E_b$ restricts to the central element $\theta_a E_a$ on $V_1(E_a)$ and to $\theta_b E_b$ elsewhere, spectral calculus splits along that sum:

$$\exp_J(x + t\,h_{aa}) \;=\; e^{\theta_a} \exp_{V_1(E_a)}(t\,h_{aa}) \;+\; \sum_{b \neq a} e^{\theta_b} E_b.$$

Differentiating at $t = 0$ gives $e^{\theta_a} h_{aa}$ for the numerator and $e^{\theta_a} \operatorname{tr}_J(h_{aa})$ for $\operatorname{tr}_J \exp_J$; the quotient rule gives $w_a h_{aa} - p\, w_a \operatorname{tr}_J(h_{aa})$.

Step 2 (*off-diagonal blocks*)**.** Fix $a < b$ and $z \in J_{ab}$. Because $J_{ab} = V_{1/2}(E_a) \cap V_{1/2}(E_b)$, the operator $D = 4[L_z, L_{E_a}]$ is an inner derivation with $D(E_a) = z$, $D(E_b) = -z$ and $D(E_c) = 0$ for $c \neq a, b$, exactly as in (13) — the computation only used the Peirce relations, which hold verbatim for the coarse idempotents. Along $x(t) = \sum_c \theta_c E_c(t)$ with $E_c(t) = \exp(tD)E_c$, the spectrum is constant, so $\dot{x}(0) = (\theta_a - \theta_b)z$ and $\dot{p}(0) = (w_a - w_b)z$. When $\theta_a \neq \theta_b$ — which holds by construction, the θ being the *distinct* eigenvalues — the chain rule gives $D\Gamma_J(x)[z] = \frac{w_a - w_b}{\theta_a - \theta_b} z = \mathscr{L}(w_a, w_b)z$, using $\theta_a - \theta_b = \log w_a - \log w_b$. Also $\operatorname{tr}_J z = 0$, so no normalization term appears.

Step 3 (*assembly*)**.** The coarse Peirce decomposition spans J [17, Thm. IV.2.1], so Steps 1–2 determine $D\Gamma_J(x)$, giving (14). Every object in (14) is built from the spectral idempotents alone, so no choice of primitive frame inside a repeated eigenspace enters, and the formula is valid at every $x \in J$ without a genericity hypothesis.

Step 4 (*the simple-spectrum form*)**.** If the spectrum is simple then $E_a = c_a$, $m_a = 1$, $J_{aa} = \mathbb{R}c_a$, $h_{aa} = h_a c_a$ and $\operatorname{tr}_J(h_{aa}) = h_a$, so the correction term is $-\sum_i c_i w_i \sum_k w_k h_k$ and (14) becomes (15). Since $\mathscr{L}$ is continuous across the diagonal with $\mathscr{L}(u, u) = u$, (15) is also the limit of itself as eigenvalues merge, consistent with (14). □

Remark 7.2 (Why the coarse projections). Stating Theorem 7.10 with a primitive frame and extending to repeated eigenvalues by continuity is not quite enough, and the gap is worth naming. Changing the primitive frame inside a repeated eigenspace *mixes* the directions that the primitive-frame formula calls diagonal with those it calls off-diagonal: if $\lambda_i = \lambda_j$ then a rotation inside that block moves J_{ij} into the span of c_i, c_j and back. Constancy of the coefficient on the block makes the resulting operator agree, but only after that mixing has been accounted for. The coarse decomposition removes the issue at the source, because E_a is determined by x.

Corollary 7.11 (The Hessian metric). The symmetric bilinear form

$$g_x(h, k) \;:=\; \big\langle h,\, D\Gamma_J(x)[k] \big\rangle \;=\; \nabla^2 \Psi_J(x)[h, k]$$

satisfies, at simple spectrum and in the notation of (15),

$$g_x(h, h) \;=\; \sum_i w_i h_i^2 - \Big(\sum_i w_i h_i \Big)^2 \;+\; \sum_{i<j} \mathscr{L}(w_i, w_j)\, \|h_{ij}\|^2 \;\geq\; 0,$$

and in general

$$g_x(h, h) \;=\; \sum_a w_a \|h_{aa}\|^2 - \Big(\sum_a w_a \operatorname{tr}_J(h_{aa}) \Big)^2 \;+\; \sum_{a<b} \mathscr{L}(w_a, w_b)\, \|h_{ab}\|^2,$$

with kernel exactly $\mathbb{R}e$, at every $x \in J$. Hence g is a real-analytic Riemannian metric on all of $J_0 \cong J/\mathbb{R}e$, the Hessian metric of the potential Ψ_J; its diagonal block is the softmax covariance and its frame block is diagonal in the Peirce decomposition with logarithmic-mean weights. For $J = H_3(\mathbb{O})$ all 24 frame directions occur in it with strictly positive weight.

Proof. $D\Gamma_J = D(\nabla\Psi_J) = \nabla^2\Psi_J$ by Theorem 6.4, so g_x is symmetric. The Peirce blocks are mutually orthogonal for the trace form and orthogonal to $\bigoplus_a J_{aa}$, so the quadratic form splits as displayed; the diagonal part is the softmax covariance, positive semidefinite with kernel $\{h_{aa} = \text{const} \cdot E_a\}$, and each $\mathscr{L}(w_a, w_b) > 0$. The joint kernel is $\{h_{ab} = 0,\ h_{aa} = \gamma E_a$ with γ independent of $a\} = \mathbb{R}e$, so g is positive definite on J_0, matching the strict convexity of Ψ_J there. □

Corollary 7.12 (Jacobian of the Gibbs chart). Let $x = \sum_i \lambda_i c_i$ have simple spectrum, $w_i = e^{\lambda_i}/\sum_k e^{\lambda_k}$, and $d_{ij} = \dim J_{ij}$. Relative to the trace inner products on J_0 and on $T_{\Gamma_J(x)}\mathscr{B}_J^\circ$,

$$\det\big(D\Gamma_J(x)\big|_{J_0}\big) \;=\; r\Big(\prod_{i=1}^{r} w_i\Big)\prod_{i<j}\mathscr{L}(w_i,w_j)^{d_{ij}}. \tag{16}$$

For a simple carrier of Peirce multiplicity d this reads

$$\det\big(D\Gamma_J(x)\big|_{J_0}\big) \;=\; r\Big(\prod_{i} w_i\Big)\prod_{i<j}\mathscr{L}(w_i,w_j)^{d} \;=\; r\,N\big(\Gamma_J(x)\big)\prod_{i<j}\mathscr{L}(w_i,w_j)^{d},$$

and equivalently, with $n = \dim J$ and $Z = \sum_i e^{\lambda_i}$,

$$\det\nabla^2\Psi_J(x)\big|_{J_0} \;=\; r\,Z^{-n}e^{\operatorname{tr}_J x}\prod_{i<j}\left(\frac{e^{\lambda_i}-e^{\lambda_j}}{\lambda_i-\lambda_j}\right)^{d_{ij}}. \tag{17}$$

Legendre duality (Theorem 6.4) pairs J_0 with the affine trace-one state space. The dual statement is therefore about the Hessian of the restriction $\eta_J|_{\mathscr{B}_J^\circ}$, read on the tangent space $T_p\mathscr{B}_J^\circ = J_0$, rather than about an ambient Hessian on J: at $p = \Gamma_J(x)$ with eigenvalues $\mu_i = w_i$,

$$\det{}_{T_p\mathscr{B}_J^\circ}\nabla^2\eta_J(p) \;=\; \Big[\,r\prod_i \mu_i\prod_{i<j}\mathscr{L}(\mu_i,\mu_j)^{d_{ij}}\Big]^{-1},$$

all determinants again taken relative to the trace inner product. For the Albert algebra,

$$\det\nabla^2\Psi_{H_3(\mathbb{O})}(x)\big|_{J_0} \;=\; 3\,w_1w_2w_3\big[\mathscr{L}(w_1,w_2)\mathscr{L}(w_1,w_3)\mathscr{L}(w_2,w_3)\big]^8.$$

Proof. Let $C = \bigoplus_k \mathbb{R}c_k$ and $C_0 = C \cap J_0$. Since tr_J vanishes on every J_{ij}, we have the trace-orthogonal splitting $J_0 = C_0 \oplus \bigoplus_{i<j} J_{ij}$, and by (15) each summand is invariant under $D\Gamma_J(x)$. The determinant is therefore the product of the block determinants. On J_{ij} the operator is $\mathscr{L}(w_i,w_j)$ times the identity, in dimension d_{ij}.

On $C \cong \mathbb{R}^r$ — an isometry, because $\langle c_i, c_j\rangle = \delta_{ij}$ — the operator is the Laplacian-like matrix $C_w = \operatorname{diag}(w) - ww^\top$, whose kernel is $\mathbb{R}\mathbf{1}$ (as $C_w\mathbf{1} = w - w = 0$) and whose range is $\mathbf{1}^\perp = C_0$. As C_w is symmetric, $\det(C_w|_{C_0})$ is the pseudo-determinant $\operatorname{pdet}(C_w)$.

A symmetric matrix with one-dimensional kernel $\mathbb{R}\mathbf{1}$ has $\operatorname{adj}(C_w) = r^{-1}\operatorname{pdet}(C_w)\,\mathbf{1}\mathbf{1}^\top$, so all its principal cofactors are equal and $\operatorname{pdet}(C_w) = r\cdot C_{rr}$. Deleting the last row and column and writing $w' = (w_1,\dots,w_{r-1})$, the matrix determinant lemma gives

$$C_{rr} \;=\; \det\big(\operatorname{diag}(w') - w'w'^\top\big) \;=\; \Big(\prod_{i<r} w_i\Big)\Big(1-\sum_{i<r} w_i\Big) \;=\; \Big(\prod_{i<r} w_i\Big)w_r \;=\; \prod_{i=1}^{r} w_i,$$

using $\sum_i w_i = 1$. Hence $\operatorname{pdet}(C_w) = r\prod_i w_i$, and multiplying the blocks gives (16). For (17), substitute $\prod_i w_i = e^{\operatorname{tr}_J x}Z^{-r}$ and $\mathscr{L}(w_i,w_j) = Z^{-1}(e^{\lambda_i}-e^{\lambda_j})/(\lambda_i-\lambda_j)$, and use $\sum_{i<j} d_{ij} = n - r$

from (5). The identification with $N(\Gamma_J(x)) = \prod_i w_i$ is Proposition 7.8 with the definition of N in Section 8, and the dual statement is that mutually inverse operators have reciprocal determinants. □

7.3 PEIRCE SOURCES AND THE GIBBS MONGE–AMPÈRE EQUATION

Formula (16) is more than a Monge–Ampère quantity. After one rescaling it becomes a Monge–Ampère *equation*, and the Peirce multiplicities enter it as a source term.

Corollary 7.13 (Peirce multiplicity sources the Gibbs Monge–Ampère equation). Let $n = \dim J$ and define the rescaled potential

$$\varphi_J \;=\; r\,\Psi_J\big|_{J_0} \;-\; n\log r$$

on the traceless slice J_0, of dimension $n-1$. Then at *every* $x \in J_0$, with Gibbs weights $w_1, \dots, w_r$ listed with multiplicity and Peirce multiplicities d_{ij} relative to any Jordan frame diagonalizing x,

$$\det{}_{J_0} \nabla^2\varphi_J(x) \;=\; e^{-\varphi_J(x)}\,\mathfrak{s}_J(w), \qquad \mathfrak{s}_J(w) \;=\; \prod_{i<j} \mathscr{L}(w_i, w_j)^{d_{ij}}, \tag{18}$$

where $\mathscr{L}(w,w) = w$ on the diagonal. Relative to the Gibbs Hessian $\nabla^2\Psi_J$, the spectral covariance block contributes $r\prod_i w_i$ and the Peirce-frame block contributes $\mathfrak{s}_J(w)$. Rescaling the whole Hessian by r on the $(n-1)$-dimensional traceless slice converts the first factor, together with the accumulated powers of r, into $e^{-\varphi_J}$, and leaves $\mathfrak{s}_J$ as the source. It satisfies $0 < \mathfrak{s}_J(w) \le 2^{-(n-r)}$, it equals $r^{-(n-r)}$ at the barycenter $w_i \equiv 1/r$, and for a simple carrier the barycenter is its unique maximizer. Finally $\mathfrak{s}_J \equiv 1$ if and only if $J \cong \mathbb{R}^r$, that is, exactly when the carrier has no continuous frame directions (Theorem 7.4); for those carriers (18) is the source-free real Monge–Ampère equation $\det\nabla^2\varphi = e^{-\varphi}$.

Proof. **Step 1** (*the determinant*)**.** Since $\Gamma_J = \nabla\Psi_J$ and $\varphi_J = r\Psi_J - n\log r$ on J_0, the restricted Hessian is $\nabla^2\varphi_J|_{J_0} = r\,D\Gamma_J(x)|_{J_0}$. As $\dim J_0 = n-1$, (16) gives

$$\det{}_{J_0}\nabla^2\varphi_J \;=\; r^{n-1}\det\big(D\Gamma_J(x)|_{J_0}\big) \;=\; r^n\Big(\prod_i w_i\Big)\mathfrak{s}_J(w),$$

whose two factors before rescaling are the spectral block determinant $r\prod_i w_i$ and the frame block determinant $\mathfrak{s}_J(w)$. On J_0 we have $\operatorname{tr}_J x = 0$, so

$$\prod_i w_i \;=\; e^{\operatorname{tr}_J x}e^{-r\Psi_J(x)} \;=\; e^{-r\Psi_J(x)}, \qquad \text{hence} \qquad r^n\prod_i w_i \;=\; e^{n\log r - r\Psi_J(x)} \;=\; e^{-\varphi_J(x)},$$

which is (18). That $\mathfrak{s}_J$ is the frame-block determinant is Theorem 7.10: the differential acts on J_{ij} by the scalar $\mathscr{L}(w_i,w_j)$ in dimension d_{ij}.

Step 2 (*the bounds*)**.** The logarithmic mean lies between the geometric and arithmetic means, so $0 < \mathscr{L}(w_i,w_j) \le \tfrac{1}{2}(w_i+w_j) \le \tfrac{1}{2}$, and the inequality $\mathscr{L}(w_i,w_j) \le \tfrac{1}{2}(w_i+w_j)$ is strict unless $w_i = w_j$; since $\sum_{i<j} d_{ij} = n-r$ by (5), this gives $0 < \mathfrak{s}_J \le 2^{-(n-r)}$. At $w_i \equiv 1/r$ every $\mathscr{L}(w_i,w_j) = 1/r$, whence $\mathfrak{s}_J = r^{-(n-r)}$. For a simple carrier every $d_{ij} = d$, and $w \mapsto d\sum_{i<j}\log\tfrac{1}{2}(w_i+w_j)$ is concave

on the simplex and invariant under the coordinate permutations, so it is maximized at the barycenter; combined with $\mathscr{L} \le$ arithmetic mean, with equality only for equal arguments, the barycenter is the unique maximizer of $\mathfrak{s}_J$.

Step 3 (*from the simple-spectrum locus to all of* J_0)**.** Steps 1 and 2 use (16), which is stated at simple spectrum. Both sides of (18) are continuous on J_0: the left because Ψ_J is real-analytic, the right because the unordered eigenvalue list of $\Gamma_J(x)$ depends continuously on x and $\mathscr{L}$ is continuous on $(0,\infty)^2$. The discriminant of the characteristic polynomial is a polynomial on J_0 that does not vanish identically when $r \ge 2$, since $\sum_i \lambda_i c_i$ with distinct λ_i summing to zero is traceless with simple spectrum; so the simple-spectrum locus is dense and open in J_0, and equality there extends to all of J_0.

Step 4 (*the source-free case*)**.** $\mathfrak{s}_J \equiv 1$ forces $\mathscr{L}(w_i, w_j)^{d_{ij}} = 1$ for weights that can be moved independently, hence $d_{ij} = 0$ for all $i < j$, which by Theorem 7.4 is $J \cong \mathbb{R}^r$. Conversely $\mathbb{R}^r$ has all $d_{ij} = 0$ and the product is empty. □

On a reducible carrier the source factors over the simple ideals, and the factor masses of Proposition 3.3 appear as exponents.

Proposition 7.14 (Direct-sum factorization of the source). Let $J = \bigoplus_{a=1}^{s} K_a$ be the decomposition into simple ideals, with $r_a = \operatorname{rank} K_a$ and $f_a = \dim K_a - r_a$, and write the Gibbs weights of the a-th factor as $w_{a,i} = \alpha_a v_{a,i}$, where $\alpha_a = \sum_i w_{a,i}$ is the factor mass of Proposition 3.3 and $\sum_i v_{a,i} = 1$. Then

$$\mathfrak{s}_J(w) \;=\; \prod_{a=1}^{s} \alpha_a^{f_a}\, \mathfrak{s}_{K_a}(v_a). \tag{19}$$

Writing $F = \sum_a f_a = \dim J - r$, the source obeys

$$\sup \mathfrak{s}_J \;=\; \prod_{a\,:\,f_a>0} \Big(\frac{f_a}{F\, r_a}\Big)^{f_a}, \tag{20}$$

and if every $f_a > 0$ the supremum is attained, uniquely, at

$$\alpha_a = \frac{f_a}{F}, \qquad v_{a,i} = \frac{1}{r_a}, \qquad \text{that is} \qquad w_{a,i} = \frac{f_a}{F\, r_a}.$$

If some $f_a = 0$ while others are positive, the supremum is approached only on the boundary, with the mass of every frameless factor tending to zero. If $F = 0$ then every $K_a \cong \mathbb{R}$, $J \cong \mathbb{R}^r$, and $\mathfrak{s}_J \equiv 1$.

Proof. Peirce blocks joining different ideals vanish, so the product defining $\mathfrak{s}_J$ splits over the factors. The logarithmic mean is homogeneous of degree one, $\mathscr{L}(\alpha u, \alpha v) = \alpha \mathscr{L}(u, v)$, and the exponents inside the a-th factor sum to $\sum_{i<j} d_{ij}^{(a)} = \dim K_a - r_a = f_a$ by (5) applied to K_a; hence that factor contributes $\alpha_a^{f_a} \mathfrak{s}_{K_a}(v_a)$, which is (19).

For (20), Corollary 7.13 applied to the simple carrier K_a gives $\mathfrak{s}_{K_a}(v_a) \le r_a^{-f_a}$ with equality exactly at $v_{a,i} \equiv 1/r_a$; when $f_a = 0$ both sides are 1. So $\mathfrak{s}_J \le \prod_a (\alpha_a / r_a)^{f_a}$, and it remains to maximize $\sum_a f_a \log \alpha_a$ over the closed simplex. That function is strictly concave in the

coordinates with $f_a > 0$ and constant in the others, so its maximum is at $\alpha_a = f_a/F$ for $f_a > 0$ and $\alpha_a = 0$ otherwise, of value $\prod_{a:\, f_a>0}(f_a/F)^{f_a}$. If every $f_a > 0$ that point is interior, and uniqueness follows from strict concavity together with the uniqueness of the factor barycenters. If some $f_a = 0$ the maximizing α lies on the boundary, where the corresponding factor state leaves $\mathscr{B}_J^\circ$, so the value is a supremum and not attained. Finally $F = 0$ forces every $f_a = 0$, hence every K_a of rank one. □

Taking logarithms in (19),

$$\log \mathfrak{s}_J \;=\; \sum_a f_a \log \alpha_a \;+\; \sum_a \log \mathfrak{s}_{K_a}(v_a),$$

an outer term in the factor masses and an inner term in the factor states, matching the outer–inner split of Proposition 3.3. One consequence is worth stating: on a reducible carrier the global barycenter is generally *not* the source maximizer. The optimal mass on a factor is governed by its frame dimension f_a, not by its rank.

For the simplicial carrier the equation has a second life, as the transport equation of a specific lattice polytope.

Corollary 7.15 (The centered simplicial image is the Ehrhart simplex). Let $J = \mathbb{R}^r$, so that $n = r$ and (18) is source-free. The gradient of $\varphi_{\mathbb{R}^r}$ along J_0 is the trace-orthogonal projection of the ambient gradient,

$$\nabla_{J_0}\varphi_{\mathbb{R}^r}(x) \;=\; r\,\Gamma_{\mathbb{R}^r}(x) - e \;=\; r\,\operatorname{softmax}(x) - e, \tag{21}$$

and it is a real-analytic diffeomorphism from J_0 onto the interior of

$$K_r \;=\; r\Delta^{r-1} - e \;=\; \operatorname{conv}\{rc_1 - e, \dots, rc_r - e\} \;\subset\; J_0.$$

Relative to the root lattice $A_{r-1} = \mathbb{Z}^r \cap J_0$:

(a) the barycenter of K_r is 0, and 0 is its only interior lattice point;

(b) $\operatorname{covol}(A_{r-1}) = \sqrt{r}$ and, in the trace metric, $\operatorname{vol}(K_r) = r^{\,r-1}\sqrt{r}/(r-1)!$, so with $m = r - 1 = \dim J_0$ the lattice-normalized volume is

$$\frac{\operatorname{vol}(K_r)}{\operatorname{covol}(A_{r-1})} \;=\; \frac{r^{\,r-1}}{(r-1)!} \;=\; \frac{(m+1)^m}{m!}.$$

The value $(m+1)^m/m!$ is the one conjectured by Ehrhart [16] to bound the lattice-normalized volume of any m-dimensional convex body whose barycenter is its only interior lattice point, and K_r is the body realizing it; see Remark 7.3 for the status of that bound.

Proof. Since $\operatorname{tr}_J \Gamma_J = 1$, the trace-orthogonal projection of $\nabla\Psi_J = \Gamma_J$ onto J_0 is $\Gamma_J - e/r$, and multiplying by r gives (21); that $\Gamma_{\mathbb{R}^r}|_{J_0}$ is a real-analytic diffeomorphism onto Δ° is Theorem 3.2 with Theorem 4.1, and $y \mapsto ry - e$ is affine and invertible.

(a) The barycenter of Δ^{r-1} is e/r, so that of $K_r = r\Delta^{r-1} - e$ is 0. A point of K_r is $y = rp - e$ with $p \in \Delta^{r-1}$, that is $y \in J_0$ with every $y_i \geq -1$, and y is interior exactly when every $y_i > -1$. If such a y also lies in $\mathbb{Z}^r$ then $y_i \geq 0$ for every i, and $\sum_i y_i = 0$ forces $y = 0$; and 0 is interior.

(b) The lattice A_{r-1} has basis $c_i - c_{i+1}$ with Gram matrix the Cartan matrix of type A_{r-1}, of determinant r, so its covolume is $\sqrt{r}$. The standard simplex Δ^{r-1} has $(r-1)$-dimensional volume $\sqrt{r}/(r-1)!$; dilation by r multiplies it by r^{r-1} and the translation by $-e$ leaves it unchanged. Dividing gives $r^{r-1}/(r-1)!$. □

Remark 7.3 (Peirce multiplicity becomes a Monge–Ampère source). Corollary 7.13 separates the Gibbs volume density into the two pieces the Hessian keeps apart. Before rescaling, the spectral covariance block contributes $r\prod_i w_i$ and each continuous frame channel contributes one logarithmic-mean factor for each of its d_{ij} dimensions; the rescaling by r turns the first of these into $e^{-\varphi_J}$ and leaves the second untouched as the source.

The source degenerates exactly when the frame block disappears, and what remains is the classical real Monge–Ampère equation underlying the Berman–Berndtsson transport potential: $\det \nabla^2\varphi = e^{-\varphi}$ with $\nabla\varphi(\mathbb{R}^m) = \operatorname{int} K$ is the equation they attach to a centered convex body K [10], and Corollary 7.15 exhibits $\varphi_{\mathbb{R}^r}$ as an explicit solution for $K = K_r$. Every carrier with a continuous frame deforms that equation by the factor $\mathfrak{s}_J$.

Two qualifications belong with the Ehrhart reading. The bound $(m+1)^m/m!$ was conjectured by Ehrhart [16] and is known for particular classes of rational polytopes [11], with the equality case in those classes settled in [31]; a proof in every dimension is reported in [32], where the classification of the equality case is left open. Corollary 7.15 does not depend on any of this: it computes the lattice data and the volume of K_r directly. And unlike the Bergman route to the inequality, the convexity used here comes from the Gibbs variational theorem and needs no identification of a distinguished Gram entry with a full Bergman kernel.

The independence also runs the other way, and that direction carries more content. With the inequality reported in every dimension while its equality case is left open, an equality body produced by an unrelated route is a datum for that classification rather than a consequence of it. Corollary 7.15 supplies one in every dimension $m = r - 1$: the body $K_r = r\Delta^{r-1} - e$, its lattice data, and a potential $\varphi_{\mathbb{R}^r}$ solving $\det \nabla^2\varphi = e^{-\varphi}$ with $\nabla\varphi(J_0) = \operatorname{int} K_r$, all read off the Gibbs chart of the simplicial carrier. The claim is about the route and not about the body: K_r is the expected extremal simplex, and Corollary 7.15 classifies nothing.

The multiplicity data in the source are the same data that control symmetric-cone integration. For a *simple* carrier they are the pair (r, d): the rank appears as the prefactor and as the number of weight factors, the single multiplicity d as the common exponent. For an arbitrary carrier the controlling datum is the full array (d_{ij}), constant on each simple ideal and zero between ideals. The same pair governs the Gindikin gamma function and the Weyl-type density of the cone [19, 17]. The difference is that the Gibbs chart contributes logarithmic means of the weights where the cone's own volume element contributes a bare Vandermonde factor.

Remark 7.4 (The Bogoliubov–Kubo–Mori metric). For $J = \operatorname{Herm}_r(\mathbb{C})$, $\Psi_J(X) = \log \operatorname{tr} e^X$ and (15) is the standard eigenbasis expression for $\nabla^2 \log \operatorname{tr} e^X$, whose off-diagonal entries are the logarithmic means of the eigenvalues of $\rho = \Gamma_J(X)$. So g is the *pullback of the Bogoliubov–Kubo–Mori metric under the Gibbs chart*, equivalently its expression in exponential (natural) coordinates [34, 35, 2]. This formulation distinguishes the metric in natural coordinates from its Legendre-dual expression in state coordinates; see Remark 7.5. For $J = H_3(\mathbb{O})$ the same formula gives an F_4-invariant analogue on a 26-dimensional score space, with all 24 frame directions metrically visible.

Remark 7.5 (Which coordinates). The coefficients depend on the chart. In score (natural) coordinates $\nabla^2\Psi_J$ has frame coefficients $\mathscr{L}(w_i, w_j)$, as above. In state coordinates the relevant object is the dual Hessian $\nabla^2\eta_J$, and since Ψ_J and η_J are Legendre conjugates by Theorem 6.4, the two Hessians are mutually inverse at corresponding points: the frame coefficients of $\nabla^2\eta_J$ are the reciprocals $1/\mathscr{L}(w_i, w_j)$. These are one dual pair, not two competing conventions.

The frame directions therefore do more than persist in the state space. They carry an automorphism-invariant metric, weighted by logarithmic means of the same Gibbs weights the chart emits, with total gain (16). The metric is canonical in a precise sense: it is the Hessian of Ψ_J, and the chart is the gradient of Ψ_J. On $H_3(\mathbb{O})$ this is a 26-dimensional F_4-invariant Riemannian metric on score space whose 24 frame directions are metrically visible and whose 2 spectral directions carry the familiar softmax covariance. Softmax is not absent from the Albert chart; it is the spectral half of it. The next section sets the classical carriers side by side before the frame is put to work.

8 CARRIER ATLAS: THE WEIGHT–FRAME TRADEOFF

Table 1 collects the standard carrier families; all entries follow from Theorem 7.3 and Proposition 7.5 with the multiplicities of Example 2.1.

CARRIER J	$\dim J$	r	d	$\dim \mathfrak{F}_J$	$\dim \mathscr{B}^\circ$	EXTREME PTS	$\mathrm{Aut}(J)$
$\mathbb{R}^n$, $n \geq 2$	n	n	0	0	$n-1$	n points	S_n
$\mathrm{Sym}_r(\mathbb{R})$, $r \geq 2$	$\frac{r(r+1)}{2}$	r	1	$\frac{r(r-1)}{2}$	$\frac{r(r+1)}{2}-1$	$\mathbb{R}P^{r-1}$	$PO(r)$
$\mathrm{Herm}_r(\mathbb{C})$, $r \geq 2$	r^2	r	2	$r(r-1)$	r^2-1	$\mathbb{C}P^{r-1}$	$PU(r) \rtimes \mathbb{Z}_2$
$\mathrm{Herm}_r(\mathbb{H})$, $r \geq 3$	$r(2r-1)$	r	4	$2r(r-1)$	$r(2r-1)-1$	$\mathbb{H}P^{r-1}$	$PSp(r)$
V_m, $m \geq 2$	$m+1$	2	$m-1$	$m-1$	m	S^{m-1}	$O(m)$
$H_3(\mathbb{O})$	27	3	8	24	26	$\mathbb{O}P^2$, dim 16	F_4

Table 1. *Standard carrier families. The Gibbs chart emits r weights into a state space of dimension* $(r-1) + \dim \mathfrak{F}_J = \dim J - 1$*; the last two columns are the extreme points of the closed state space and the chart symmetry group* $\mathrm{Aut}(J)$*. Of these columns only d enters the Monge–Ampère source of Corollary 7.13; the group in the last column is what Corollary 9.22 reads instead.*

Two conventions in the table. In the complex row the $\mathbb{Z}_2$ acts by entrywise conjugation and $PU(r)$ is the identity component, while the other listed groups are full automorphism groups. Under the stated restrictions every row but the first is simple; the first is the only frameless one (Theorem 7.4), and its ordered-frame set is finite, with $n!$ orderings, rather than empty. Two boundary cases sit just outside the ranges shown. At $m = 1$ the spin factor is $\mathrm{V}_1 \cong \mathbb{R}^2$, simplicial rather than simple, with zero frame dimension. The quaternionic row starts at $r = 3$ because its rank-two member is a spin factor, $\mathrm{Herm}_2(\mathbb{H}) \cong \mathrm{V}_5$ with $\mathrm{Aut} = O(5)$ and $\mathbb{H}P^1 \cong S^4$; there $PSp(2) \cong SO(5)$ is only the identity component, since $X \mapsto \mathrm{tr}_J(X)e - X$ is an automorphism of determinant -1. The real and complex rows have no such defect, as $PO(2) \cong O(2)$ and $PU(2) \rtimes \mathbb{Z}_2 \cong O(3)$.

Reading the table across a fixed rank separates two things this paper keeps apart throughout and has not yet named. Fix $r = 3$ and read the rows $\mathbb{R}^3$, $\mathrm{Sym}_3(\mathbb{R})$, $\mathrm{Herm}_3(\mathbb{C})$, $\mathrm{Herm}_3(\mathbb{H})$, $H_3(\mathbb{O})$: the Peirce multiplicity runs $d = 0, 1, 2, 4, 8$, and by Corollary 7.13 the Monge–Ampère source of every one of them is

$$\mathfrak{s}_J(w) \;=\; \big[\mathscr{L}(w_1, w_2)\,\mathscr{L}(w_1, w_3)\,\mathscr{L}(w_2, w_3)\big]^d,$$

one function of the same three weights, raised to a power. Every coefficient in Theorem 7.10 and Corollaries 7.11 to 7.13 depends on the Peirce blocks only through their dimensions d_{ij}, so along this ladder the Albert algebra enters as the exponent 8 and in no other way. This does not fail to separate carriers, since for simple J the pair (r, d) already determines J up to isomorphism; it says what the source is a function of, which is a *multiplicity count*.

What a non-simplicial carrier supplies beyond that multiplicity count — the second of the two things — is not a count. The blocks are modules, and the module structure first does work in Corollary 9.22: the subgroup of F_4 fixing an ordered Albert frame pointwise is Spin(8), and the three edges carry $\mathbf{8}_v \oplus \mathbf{8}_s \oplus \mathbf{8}_c$, permuted by triality. No determinant, Jacobian or source can see this, because each of them reads $\dim J_{ij}$ and nothing else. The two readings of what the frame adds — how many dimensions, and what group acts on them how — are used separately below and are not interchangeable.

Three instances are worth writing out.

Example 8.1 (Symmetric matrices: the density matrix). For $J = \mathrm{Sym}_r(\mathbb{R})$ with $X \circ Y = {}^1\!/_2(XY + YX)$, the cone is the positive definite matrices, tr_J is the matrix trace, $\exp_J$ is the matrix exponential, and $\Gamma_J(X) = e^X / \mathrm{tr}(e^X)$. This is a density matrix, not a probability vector: it applies ordinary softmax to the eigenvalues and transports the eigenspaces, and the ${}^1\!/_2 r(r-1)$ frame dimensions record where those eigenspaces point.

Example 8.2 (Spin factor: the radial tanh). For $J = \mathrm{V}_m$ write $x = (a, u)$ and $\rho = \|u\|$. The eigenvalues are $a \pm \rho$, and for $\rho > 0$ the primitive idempotents are $c_\pm = {}^1\!/_2(1, \pm u/\rho)$, so $\exp_J(a, u) = e^a(\cosh\rho,\ \frac{\sinh\rho}{\rho}u)$ and $\mathrm{tr}_J(a, u) = 2a$, whence

$$\Gamma_J(a, u) \;=\; \Big(\tfrac{1}{2},\ \tfrac{\tanh\rho}{2\rho}\,u\Big).$$

At $\rho = 0$ the frame is not determined and u/ρ is undefined, but the final formula extends continuously, since $\tanh\rho/\rho \to 1$ and $\frac{\tanh\rho}{2\rho}u \to 0$: $\Gamma_J(a, 0) = ({}^1\!/_2, 0) = e/2$, the center of the ball. The state space is $\{{}^1\!/_2\} \times B(0, {}^1\!/_2)$: the simplicial normalizer has become a radial tanh into a ball, the first coordinate pinned by the trace condition.

In the rank-3 example below we use the cubic norm. For any EJA of rank r the *determinant* N is the highest-degree characteristic coefficient of the Jordan characteristic polynomial, so $N(x) = \prod_k \lambda_k$ on a spectral decomposition; it is a homogeneous polynomial of degree r, and $N(x) > 0$ for $x \in \Omega_J$. The converse fails: Ω_J is the connected component of e in the invertible locus $\{N \neq 0\}$, and positivity of N alone does not characterize it. In $\mathrm{Sym}_2(\mathbb{R})$, for instance, $N(-I) = \det(-I) = 1 > 0$ while $-I \notin \Omega_J$. Its full polarization is normalized by

$$N_r(x_1, \dots, x_r) \;=\; \frac{1}{r!}\,\frac{\partial^r}{\partial t_1 \cdots \partial t_r}\,N\big(t_1 x_1 + \cdots + t_r x_r\big)\Big|_{t=0}, \tag{22}$$

the unique symmetric r-linear form with $N_r(x,\dots,x) = N(x)$; homogeneity gives $\partial_t^r N(tx)|_0 = r!\,N(x)$, which is what fixes the factor. For $r = 3$ we write N_3.

Example 8.3 (Albert algebra). For $J = H_3(\mathbb{O})$ and $x = \lambda_1 c_1 + \lambda_2 c_2 + \lambda_3 c_3$,

$$\Gamma_J(x) \;=\; \sum_{i=1}^{3} \frac{e^{\lambda_i}}{e^{\lambda_1} + e^{\lambda_2} + e^{\lambda_3}}\, c_i,$$

a three-way softmax, entirely familiar. What has changed is where the weights are attached: the frame (c_1, c_2, c_3) ranges over the 24-dimensional variety $F_4/\mathrm{Spin}(8)$, and the state occupies the 26-dimensional manifold $\mathscr{B}^\circ$. The determinant-one slice $\{y \in \Omega_J \,:\, N(y) = 1\}$ is the Riemannian symmetric space $E_{6(-26)}/F_4$, and $\mathscr{B}^\circ$ is radially diffeomorphic to it via $p \mapsto N(p)^{-1/3} p$, with inverse $y \mapsto y/\operatorname{tr}_J y$. By Propositions 7.6 and 7.8 the three weights exhaust the one-state F_4-invariants; by Proposition 7.9 the 24 frame dimensions become invariantly observable the moment a second state is present; and by Corollary 7.11 they carry an F_4-invariant metric with logarithmic-mean coefficients. The ratio of spectral to total state dimensions is $2/26 = 1/13$; for $\mathbb{R}^n$ it is 1.

Remark 8.1 (Trace form versus dot product). Every inner product above is $\langle x, y\rangle = \operatorname{tr}_J(x \circ y)$. For $H_3(\mathbb{O})$ in natural matrix coordinates — diagonal entries a_i, off-diagonal octonions u_{ij} —

$$\operatorname{tr}_J(X \circ Y) \;=\; \sum_i a_i b_i \;+\; 2\sum_{i<j} \langle u_{ij}, v_{ij}\rangle_{\mathbb{O}},$$

so the Gram matrix of T has eigenvalue 1 on the three diagonal directions and 2 on the twenty-four off-diagonal ones. Pairing states with a plain dot product in these coordinates is off by a factor of 2 on 24 of 27 coordinates, which Corollary 7.11 would then report incorrectly. This is a coordinate convention, not an indictment of dot products: rescaling every off-diagonal coordinate by $\sqrt{2}$ gives an orthonormal basis in which the ordinary dot product *is* the trace form.

The table settles what the frame *is*. What remains is whether it can do anything.

9 POLAR TRANSPORT AND FRAME-SENSITIVE READOUTS

The preceding sections identify the additional degrees of freedom supplied by a non-simplicial carrier. We now determine which of those degrees survive canonical normalization, and how they may act on carrier-valued data. The aim is not to prescribe a particular architecture but to characterize the normalization, transport and readout operations that the symmetric-cone structure makes available.

The development has three stages. First the normalization: Section 9.1 records what a scalar score keeps and what it drops, Section 9.2 normalizes cone-valued scores into effects, and Section 9.3 identifies the surviving query dependence as an automorphism of the carrier. Then the construction: Section 9.4 gives the radial–effect coordinates in which the normalization is invertible and locates the compact factor on the projectivized cone, Section 9.5 lets the effects act on carrier-valued values, and Section 9.6 exhibits the transporter differential as a filter on Peirce edges. Then the readouts: Section 9.7 reduces the central-value branch to two

additive carrier moments, Section 9.8 treats relative-frame separation and orbit curvature, and Section 9.9 pairs transported Peirce projections into positive kernels. Section 9.10 records two things the construction does not contain: a composite of carriers, and a test of nonclassicality.

9.1 SCALAR NORMALIZATION FORGETS THE COVARIANT FRAME

Suppose a scoring rule, possibly followed by a separately specified positivity map, produces nonnegative scalars $s_{ij} \in \mathbb{R}_{\geq 0}$ for query i and keys $j = 1, \dots, n$. Candidate rules include dot products, determinants, and polarized cubic invariants of Albert-valued data; which of them actually satisfy the nonnegativity hypothesis is a separate question, taken up below. The row $s_i = (s_{i1}, \dots, s_{in})$ lies in the closed cone $\mathbb{R}^n_{\geq 0}$ of the simplicial carrier, and if $s_i \neq 0$ then sum normalization lands in Δ_{n-1}. Scalar normalization therefore returns to the simplicial carrier: whatever algebra manufactured the numbers, the normalization step takes place in $\mathbb{R}^n$, and by Theorem 7.4 that carrier has no frame directions at all. What is dropped is the covariant carrier frame, not every trace of the geometry. Relational scalar invariants computed upstream survive as numbers, and one of them — the trace pairing $\langle p, q \rangle$ — already separates configurations with identical individual spectra (Proposition 9.19). What no longer travels downstream is the equivariant frame itself, and with it the carrier-valued action of Section 9.5.

Two provisos carry real content. Nonnegativity of the kernel on its domain of use requires proof: the polarized cubic $N_3(x, q, k)$ is a real number of no fixed sign, so it can serve as a kernel but is not an element of Ω_J and does not occupy the cone corner of Figure 1. And $s_i = 0$ must be excluded or handled: sum normalization is undefined on the zero row, and unlike the exponential lift — valued in the open cone, which it cannot leave — a construction that omits the lift reaches the cone boundary by design. A boundary rule (uniform fallback, a distinguished null key, a carried-forward state) is part of the specification.

Carrier-valued covariance is retained by keeping the scores in $\overline{\Omega}_J$; cone-valued kernels built from symmetric-cone data have been proposed for linear attention on exactly these grounds [18]. For $q, k \in \Omega_J$ the quadratic representation gives

$$A \;=\; P\big(q^{1/2}\big)k \;\in\; \Omega_J, \tag{23}$$

the Jordan form of the congruence $q^{1/2}kq^{1/2}$, positive because $P(a)$ preserves the cone for invertible a [17, Ch. II–III]. A row of such scores lies in $J^n = J \oplus \cdots \oplus J$, itself an EJA with componentwise operations and trace $\operatorname{tr}_{J^n} = \sum_j \operatorname{tr}_J$, so trace normalization gives carrier-valued weights $W_j = Y_j / \sum_k \operatorname{tr}_J(Y_k)$ with $\sum_j \operatorname{tr}_J(W_j) = 1$ — a normalized tuple of positive elements, not a probability vector. What remains is a normalization adapted to the cone and an action on values. Both are supplied below.

The following is the Jordan analogue of normalizing positive operators into a resolution of the identity. Effects and their pairings with states on Jordan-algebraic state spaces are developed systematically by Alfsen and Shultz [1], and Section 9.4 shows that this normalization discards nothing that cannot be restored.

Proposition 9.1 (Effect normalization). Let $A_1, \dots, A_n \in \overline{\Omega}_J$ with $S = \sum_j A_j \in \Omega_J$, and set $E_j = P(S^{-1/2})A_j$. Then $E_j \in \overline{\Omega}_J$ and $\sum_j E_j = e$, so for every state $\rho \in \mathscr{B}_J$

$$p_j \;=\; \langle \rho, E_j \rangle$$

satisfies $p_j \geq 0$ and $\sum_j p_j = 1$. The construction is equivariant,

$$E_j(gA_1, \dots, gA_n) \;=\; g\, E_j(A_1, \dots, A_n) \qquad \text{for } g \in \mathrm{Aut}(J).$$

Moreover $P(S^{-1/2})$ is the *unique* cone automorphism carrying S to e that is positive definite and self-adjoint for the trace form.

Proof. The operator $P(S^{-1/2})$ is a cone automorphism, so each E_j lies in the closed cone, and $\sum_j E_j = P(S^{-1/2})S = e$ by [17, Prop. II.3.1]. In the matrix case this reads $E_j = S^{-1/2} A_j S^{-1/2}$ with $\sum_j E_j = I$. Nonnegativity of p_j follows from self-duality of the cone, and $\sum_j p_j = \langle \rho, e \rangle = \mathrm{tr}_J\, \rho = 1$. Equivariance follows from $g\, P(a)\, g^{-1} = P(ga)$ and the definition of S.

For uniqueness, suppose $g \in G(\Omega_J)$ satisfies $gS = e$ and is positive definite self-adjoint. Then $h := g\, P(S^{1/2})$ fixes e, since $P(S^{1/2})e = S$; and the stabilizer of e in $G(\Omega_J)$ is $\mathrm{Aut}(J)$ [17, Thm. III.5.1], whose elements are orthogonal for the trace form because they preserve both the product and the trace. Thus $g = h\, P(S^{-1/2})$ exhibits a factorization of g into an orthogonal operator times a positive definite self-adjoint one — a polar decomposition. The polar decomposition of a positive definite self-adjoint operator has orthogonal part the identity, so $h = \mathrm{id}$ and $g = P(S^{-1/2})$. □

The construction specializes as expected: for $J = \mathbb{R}$ it is ordinary scalar sum normalization; for matrix carriers it is the standard square-root normalization $E_j = S^{-1/2} A_j S^{-1/2}$ of positive effects; for $H_3(\mathbb{O})$ it is the exceptional quadratic-representation analogue. It also runs row by row, so the effects may depend on the query: setting

$$A_{ij} = A(q_i, k_j), \qquad S_i = \sum_j A_{ij}, \qquad E_{ij} = P\big(S_i^{-1/2}\big)A_{ij}, \qquad p_{ij} = \langle \rho_i, E_{ij} \rangle,$$

the proof of Proposition 9.1 applies verbatim to each i, and the resulting layer is query–key dependent rather than a state measured against one fixed effect family. Aggregating, $O_i = \sum_j p_{ij} V_j$, and simultaneous transport of all carrier objects by $g \in \mathrm{Aut}(J)$ leaves the probabilities invariant and the output equivariant.

Normalizing the congruenced scores (23) looks at first like discarding the query: the effects sum to e whatever q was. They do not. The query survives exactly, as an automorphism of the carrier, and that automorphism is the mechanism by which relative-frame data enters a routing layer.

Theorem 9.2 — Polar factorization of effect normalization

Let $k_1, \dots, k_m \in \overline{\Omega}_J$ with $K = \sum_j k_j \in \Omega_J$, let $q \in \Omega_J$, and set

$$A_j = P\big(q^{1/2}\big)k_j, \qquad S = \sum_j A_j = P\big(q^{1/2}\big)K, \qquad E_j = P\big(S^{-1/2}\big)A_j, \qquad \widehat{E}_j = P\big(K^{-1/2}\big)k_j.$$

Then $\sum_j \widehat{E}_j = \sum_j E_j = e$, and with

$$\mathcal{U}(q,K) \;:=\; P\big(S^{-1/2}\big)\,P\big(q^{1/2}\big)\,P\big(K^{1/2}\big) \tag{24}$$

one has $\mathcal{U}(q,K) \in \mathrm{Aut}(J)$ and

$$E_j \;=\; \mathcal{U}(q,K)\,\widehat{E}_j \qquad \text{for every } j. \tag{25}$$

Equivalently, (24) is the left polar decomposition

$$P\big(q^{1/2}\big)P\big(K^{1/2}\big) \;=\; P\big(S^{1/2}\big)\,\mathcal{U}(q,K) \tag{26}$$

of the composite cone automorphism, in which $P(S^{1/2})$ is the positive radial factor and $\mathcal{U}(q,K)$ the orthogonal automorphism factor; this is what the name *polar transport* refers to. The construction is equivariant: $\mathcal{U}(gq, gK) = g\,\mathcal{U}(q,K)\,g^{-1}$ for $g \in \mathrm{Aut}(J)$. Writing $G = G(\Omega_J)$ for the linear automorphism group of the cone, $K_J = \mathrm{Aut}(J) = \mathrm{Stab}_G(e)$ and $s_x = P(x^{1/2})$ for the canonical positive section of $G \to G/K_J \cong \Omega_J$, the map

$$\kappa(g,x) \;=\; s_{gx}^{-1}\,g\,s_x \;\in\; K_J$$

satisfies the cocycle identity $\kappa(g_1g_2,x) = \kappa(g_1, g_2x)\,\kappa(g_2,x)$, and $\mathcal{U}(q,K) = \kappa\big(P(q^{1/2}),K\big)$.

Proof. **Step 1** (*the sums*)**.** $P(a^{-1}) = P(a)^{-1}$ and $P(a^{1/2})e = a$ [17, Prop. II.3.1], so $\sum_j \widehat{E}_j = P(K^{-1/2})K = e$ and $\sum_j E_j = P(S^{-1/2})S = e$; each $P(a)$ with $a \in \Omega_J$ preserves $\overline{\Omega}_J$, so all the elements listed lie in the closed cone, and $S \in \Omega_J$ because $P(q^{1/2})$ is a cone automorphism.

Step 2 ($\mathcal{U}$ *fixes* e)**.** Each factor of (24) lies in G, and

$$\mathcal{U}(q,K)\,e \;=\; P(S^{-1/2})P(q^{1/2})P(K^{1/2})e \;=\; P(S^{-1/2})P(q^{1/2})K \;=\; P(S^{-1/2})S \;=\; e.$$

The stabilizer of e in G is $\mathrm{Aut}(J)$ [17, Thm. III.5.1], so $\mathcal{U}(q,K) \in \mathrm{Aut}(J)$.

Step 3 (*the transport identity*)**.** From $\widehat{E}_j = P(K^{-1/2})k_j$ we get $k_j = P(K^{1/2})\widehat{E}_j$, hence

$$E_j \;=\; P(S^{-1/2})P(q^{1/2})k_j \;=\; P(S^{-1/2})P(q^{1/2})P(K^{1/2})\widehat{E}_j \;=\; \mathcal{U}(q,K)\widehat{E}_j.$$

Step 4 (*equivariance and the cocycle*)**.** An automorphism commutes with spectral calculus, so $g(a^{1/2}) = (ga)^{1/2}$, and $g\,P(a)\,g^{-1} = P(ga)$. Hence $S(gq, gK) = P((gq)^{1/2})gK = gS$, and substituting into (24) gives $\mathcal{U}(gq, gK) = g\mathcal{U}(q, K)g^{-1}$. For the cocycle, $s_y^{-1}y = P(y^{-1/2})y = e$ gives $\kappa(g, x)e = s_{gx}^{-1}gx = e$, so $\kappa(g, x) \in K_J$; and

$$\kappa(g_1, g_2x)\kappa(g_2, x) = s_{g_1g_2x}^{-1}g_1s_{g_2x}\,s_{g_2x}^{-1}g_2s_x = s_{g_1g_2x}^{-1}g_1g_2s_x = \kappa(g_1g_2, x).$$

Taking $g = P(q^{1/2})$ and $x = K$, so that $gx = S$, gives $\kappa(g, K) = P(S^{-1/2})P(q^{1/2})P(K^{1/2}) = \mathcal{U}(q, K)$. □

Since $\mathcal{U}(q, K)$ is an automorphism, it preserves every spectrum: the individual spectra of the normalized effects are those of the query-free $\widehat{E}_j$, and all of the query dependence sits in a common transport of their frames. The residue left by normalization is $\mathrm{Aut}(J)$-valued, and the next proposition says exactly when it is nontrivial.

Proposition 9.3 (The transporter across carriers). In the setting of Theorem 9.2:

(a) If q and K are diagonal in a common Jordan frame then $\mathcal{U}(q, K) = \mathrm{id}$. In particular $\mathcal{U} \equiv \mathrm{id}$ on $J \cong \mathbb{R}^r$, where every pair of elements shares the standard frame.

(b) For $J = \mathrm{Herm}_r(\mathbb{C})$ with $S = q^{1/2}Kq^{1/2}$, the element $u = S^{-1/2}q^{1/2}K^{1/2}$ is unitary and $\mathcal{U}(q, K)X = uXu^*$: the transporter is the unitary polar factor of $q^{1/2}K^{1/2}$. The carriers $\mathrm{Sym}_r(\mathbb{R})$ and $\mathrm{Herm}_r(\mathbb{H})$ give the orthogonal and compact-symplectic analogues.

(c) For $J = H_3(\mathbb{O})$, $\mathcal{U}(q, K) \in F_4$.

Proof. (a) Let $q = \sum_k q_kc_k$ and $K = \sum_k K_kc_k$ with $q_k, K_k > 0$. For $a = \sum_k a_kc_k \in \Omega_J$ the operator $P(a)$ acts on $\mathbb{R}c_k$ by a_k^2 and on J_{ij} by a_ia_j [17, Lem. IV.2.2]. Hence $S = P(q^{1/2})K = \sum_k q_kK_kc_k$, and on $\mathbb{R}c_k$ the composite (24) acts by $(q_kK_k)^{-1} \cdot q_k \cdot K_k = 1$, while on J_{ij} it acts by $(q_iK_iq_jK_j)^{-1/2}(q_iq_j)^{1/2}(K_iK_j)^{1/2} = 1$. So $\mathcal{U}(q, K)$ is the identity on every Peirce block, hence on J by (4). For $J \cong \mathbb{R}^r$ the frame is unique (Theorem 7.4), so every pair qualifies.

(b) $uu^* = S^{-1/2}q^{1/2}Kq^{1/2}S^{-1/2} = S^{-1/2}SS^{-1/2} = I$. Since $P(a)X = aXa$ for Hermitian matrix carriers and $u^* = K^{1/2}q^{1/2}S^{-1/2}$,

$$\mathcal{U}(q, K)X = S^{-1/2}q^{1/2}K^{1/2}\,X\,K^{1/2}q^{1/2}S^{-1/2} = uXu^*.$$

Writing $q^{1/2}K^{1/2} = u\,(K^{1/2}qK^{1/2})^{1/2}$ exhibits u as its polar unitary factor.

(c) $\mathrm{Aut}(H_3(\mathbb{O})) = F_4$ (Corollary 7.2) and $\mathcal{U}(q, K) \in \mathrm{Aut}(J)$ by Theorem 9.2. □

Part (a) sharpens Theorem 7.4 on the architectural side. Under the congruence score (23), canonical effect normalization cancels every spectrally commuting component of the query; what survives is precisely the relative-frame transporter, and on the simplicial carrier that survivor is trivial. The exceptional carrier at the other end produces an endogenous F_4-valued transporter from nothing but the query and the keys.

The sharpening is an equivalence, and it closes the circle back to Section 7.

Corollary 9.4 (Framelessness is exactly triviality of the transporter). For a Euclidean Jordan algebra J the following are equivalent:

(a) $\mathcal{U}(q,K) = \mathrm{id}$ for every $q, K \in \Omega_J$;
(b) $\mathcal{U}(\,\cdot\,,K)$ is constant on Ω_J for every $K \in \Omega_J$;
(c) $J \cong \mathbb{R}^r$, equivalently any of the conditions of Theorem 7.4.

Consequently, on the simplicial carrier the query acts on the keys through the radial channel alone: $E_j = \widehat{E}_j$ for every j, so effect normalization removes the query entirely, and by Proposition 9.5 what it removes is exactly $S = q \circ K$. There the compact channel has nothing to act on: by Theorem 7.4 there is no frame tangent, so normalization has nothing to discard beyond the scale.

Proof. (a)⇒(b) is immediate, and (c)⇒(a) is Proposition 9.3(a), since in $\mathbb{R}^r$ every pair of elements is diagonal in the unique frame. For (b)⇒(c) argue contrapositively. Suppose $J \not\cong \mathbb{R}^r$. By Theorem 7.4(c) some Jordan frame (c_k) has a nonzero Peirce block J_{ij}; choose $0 \neq h \in J_{ij}$ and $K = \sum_k \kappa_k c_k \in \Omega_J$ with $\kappa_i \neq \kappa_j$. By Lemma 9.10 — proved in Section 9.6 from K alone, and independent of this corollary —

$$\frac{d}{dt}\Big|_0 \mathcal{U}\big(\exp_J(th), K\big) \;=\; 4\big[L_z, L_{c_i}\big], \qquad z \;=\; \frac{\sqrt{\kappa_i} - \sqrt{\kappa_j}}{2\big(\sqrt{\kappa_i} + \sqrt{\kappa_j}\big)}\, h \;\neq\; 0,$$

and (13) gives $4[L_z, L_{c_i}]c_i = z \neq 0$. So the derivative is nonzero and $t \mapsto \mathcal{U}(\exp_J(th), K)$ is nonconstant. □

Theorem 7.4 and Corollary 9.4 are one degeneracy read in two places. Framelessness is a statement about the state space of the carrier; triviality of the transporter is a statement about what survives a normalization step; and the second holds because the first does.

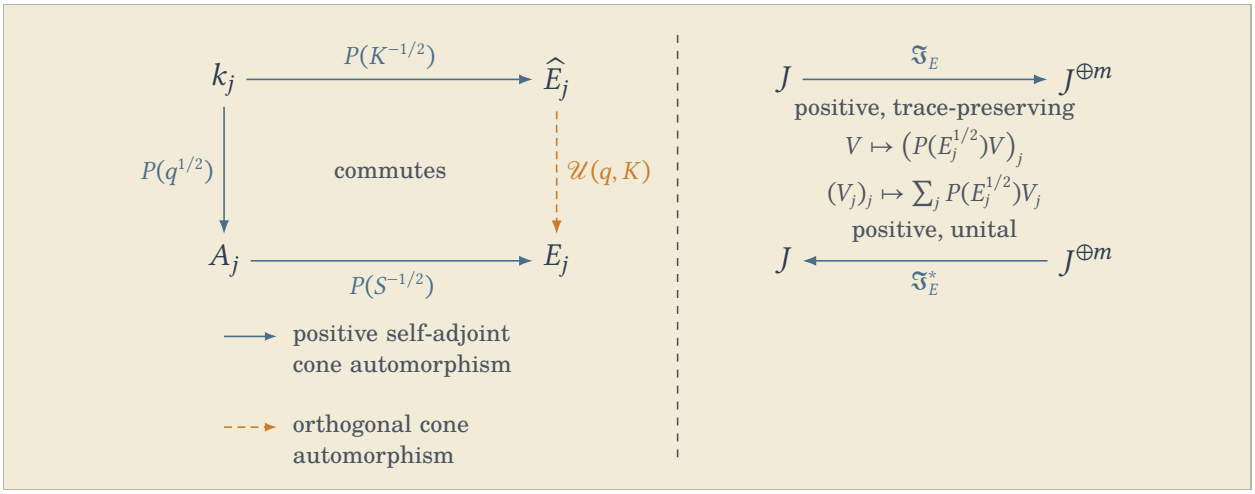


FIG. 4. *Left: query-congruenced and query-free normalization differ by the orthogonal transporter $\mathcal{U}(q,K)$; the other three arrows are positive self-adjoint cone automorphisms (Theorem 9.2). Right: $\mathfrak{T}_E$ is positive and trace-preserving, while $\mathfrak{T}_E^*$ is positive and unital (Proposition 9.9).*

The augmented radial–effect coordinates are lossless: the correspondence $(A_j) \leftrightarrow (S; (E_j))$ is invertible, while projection to the normalized effects alone forgets S. The change of coordinates separates a radial channel from a bounded one. The radial coordinate S ranges over the noncompact cone Ω_J, while the normalized-effect coordinate is bounded: for $m \geq 2$ the strict effect manifold $\mathscr{E}_m^\circ(J)$ is open and hence noncompact, though its closure is compact, and for $m = 1$ it degenerates to the single point $\{e\}$. For query-induced effect families the surviving frame dependence lies in the compact group $\mathrm{Aut}(J)$.

Proposition 9.5 (Radial and effect coordinates). Write

$$\mathscr{E}_m^\circ(J) \;=\; \Big\{(E_1, \dots, E_m) \in \Omega_J^m \;:\; \textstyle\sum_j E_j = e\Big\}$$

for the *strict effect manifold*. The maps

$$S \;=\; \sum_j A_j, \qquad E_j \;=\; P\big(S^{-1/2}\big)A_j \qquad \text{and} \qquad A_j \;=\; P\big(S^{1/2}\big)E_j$$

are mutually inverse and give a real-analytic diffeomorphism

$$\Omega_J^m \;\cong\; \Omega_J \times \mathscr{E}_m^\circ(J).$$

Moreover, writing $g_{q,K} = P(q^{1/2})P(K^{1/2})$ for the composite cone automorphism of Theorem 9.2,

$$g_{q,K}\, e \;=\; S, \qquad g_{q,K} \;=\; P\big(S^{1/2}\big)\,\mathscr{U}(q,K), \qquad A_j \;=\; P\big(S^{1/2}\big)\,\mathscr{U}(q,K)\,\widehat{E}_j. \tag{27}$$

Proof. For $A_j \in \Omega_J$ the sum S lies in Ω_J and $P(S^{-1/2})$ is a cone automorphism, so every $E_j \in \Omega_J$ and $\sum_j E_j \;=\; P(S^{-1/2})S \;=\; e$ [17, Prop. II.3.1]. Conversely, given $S \in \Omega_J$ and $(E_j) \in \mathscr{E}_m^\circ(J)$, the elements $A_j = P(S^{1/2})E_j$ lie in Ω_J and sum to $P(S^{1/2})e = S$. The two composites are the identity because $P(S^{1/2})P(S^{-1/2}) = \mathrm{id}$, and $a \mapsto a^{\pm 1/2}$ and $a \mapsto P(a)$ are real-analytic on Ω_J. The set $\mathscr{E}_m^\circ(J)$ is the intersection of the open set Ω_J^m with the affine subspace $\{\sum_j E_j \;=\; e\}$, hence a real-analytic submanifold of dimension $(m-1)\dim J$; the dimensions match, since $m \dim J = \dim J + (m-1)\dim J$. For (27), $g_{q,K}e = P(q^{1/2})K = S$, and multiplying (24) on the left by $P(S^{1/2})$ gives the factorization, which is (26); the last identity is (25). □

On the larger locus $A_j \in \overline{\Omega}_J$ with $S \in \Omega_J$, the same formulas give a bijection onto $\Omega_J \times \{(E_j) \in \overline{\Omega}_J^m \;:\; \sum_j E_j = e\}$; both directions are restrictions of ambient real-analytic maps in the variables for which $S \in \Omega_J$. That larger target is stratified by the ranks of the E_j rather than being an open analytic manifold, which is why the diffeomorphism is stated on the strict locus. When S itself is singular, normalization still makes sense inside the face that S generates.

Proposition 9.6 (Normalization inside the support face). Let $A_1, \dots, A_m \in \overline{\Omega}_J$ with $S = \sum_j A_j \neq 0$, possibly singular, and let s be the support idempotent of S. Then every A_j lies in the face subalgebra $J_s = V_1(s)$, in which S is invertible, and

$$E_j \;=\; P_{J_s}\big(S^{-1/2}\big)A_j, \qquad \sum_j E_j \;=\; s,$$

the square root and quadratic representation being taken in J_s. Adjoining the null-complement effect $E_0 = e - s$ gives an effect family in J.

Proof. Each A_j lies in $\overline{\Omega}_J$ and $\sum_j A_j = S$, so $0 \le A_j \le S$. Self-duality gives $\langle e - s, A_j\rangle \ge 0$, while $\langle e - s, A_j\rangle \le \langle e - s, S\rangle = 0$ because $S \in V_1(s)$ and the Peirce spaces $V_1(s)$, $V_1(e - s)$ are trace-orthogonal. Hence $\langle e-s, A_j\rangle = 0$, so $(e-s)\circ A_j = 0$ by Lemma 6.1 and $A_j \in V_0(e-s) = V_1(s)$. Inside J_s, a Euclidean Jordan algebra with unit s [17, Prop. IV.1.1], the element S is invertible and Proposition 9.1 applies verbatim. Finally $s + (e - s) = e$. □

This normalizes the boundary stratum by stratum. It does not extend the transporter theorem for free: the support face now moves with the data and has to be carried along, so Theorem 9.2 applies within a fixed face rather than across the stratification.

Normalization becomes lossless after adjoining the radial coordinate S. The projection $(A_j) \mapsto (E_j)$ by itself forgets S; it is the augmented correspondence $(A_j) \leftrightarrow (S; E_j)$ that is invertible. The query accordingly acts on the keys through two channels,

$$S \;=\; P\big(q^{1/2}\big)K \qquad \text{and} \qquad \mathscr{U}(q,K) \in \mathrm{Aut}(J),$$

of which the first is the positive radial factor and the second the compact frame factor; effect normalization keeps the second and discards the first. When q and K spectrally commute the compact factor is trivial (Proposition 9.3(a)), while the radial factor $S = q \circ K$ retains the query dependence. An architecture that wants to keep it can carry

$$\mathscr{T}(q,K) \;=\; \Big(\tfrac{S}{\mathrm{tr}_J S},\ \log \mathrm{tr}_J S,\ \mathscr{U}(q,K)\big|_{J_0}\Big),$$

a normalized cone state, a total scalar gain, and the frame transporter. These three reconstruct S and hence $g_{q,K} = P(S^{1/2})\mathscr{U}$ exactly, so the normalized branch $\mathscr{U}v$ and the full branch $g_{q,K}v = P(S^{1/2})\mathscr{U}v$ — positive for $v \in \Omega_J$, since $g_{q,K}$ is a cone automorphism — can coexist.

The two channels differ in what they are able to represent, and the difference is exact: the radial factor is homogeneous, and the compact factor cannot see scale at all.

Proposition 9.7 (The transporter is a function on the projectivized cone). In the setting of Theorem 9.2, for all $\mu, \lambda > 0$,

$$\mathscr{U}(\mu q,\ \lambda K) \;=\; \mathscr{U}(q,K), \qquad \text{while} \qquad S(\mu q,\ \lambda K) \;=\; \mu\lambda\, S(q,K).$$

Hence $\mathscr{U}$ descends along the normalization π of Definition 3.1 to a map

$$\mathscr{U} \;:\; \big(\Omega_J/\mathbb{R}_{>0}\big) \times \big(\Omega_J/\mathbb{R}_{>0}\big) \;=\; \mathscr{B}_J^\circ \times \mathscr{B}_J^\circ \;\longrightarrow\; \mathrm{Aut}(J).$$

More is true of the cocycle itself. The dilations $\mathbb{R}_{>0}\cdot \mathrm{id}$ lie in $G = G(\Omega_J)$ and are central there, the positive section satisfies $s_{\lambda x} = \lambda s_x$, and consequently

$$\kappa(\lambda\,\mathrm{id},\ x) \;=\; \mathrm{id}, \qquad \kappa(\lambda g,\ \mu x) \;=\; \kappa(g,x) \qquad (\lambda, \mu > 0),$$

so κ is a function on $\big(G/\mathbb{R}_{>0}\big) \times \big(\Omega_J/\mathbb{R}_{>0}\big)$.

Proof. The quadratic representation is homogeneous of degree two, $P(ta) = t^2 P(a)$, because $L_{ta} = tL_a$ and $L_{(ta)^{\circ 2}} = t^2 L_{a^{\circ 2}}$; and the spectral square root is homogeneous of degree one

half. Hence $s_{\lambda x} = P\big((\lambda x)^{1/2}\big) = \lambda P(x^{1/2}) = \lambda s_x$. Scalars commute with every linear map, so $\kappa(\lambda\,\mathrm{id}, x) = s_{\lambda x}^{-1}(\lambda\,\mathrm{id})s_x = (\lambda s_x)^{-1}\lambda s_x = \mathrm{id}$, and, using $g(\mu x) = \mu\, gx$,

$$\kappa(\lambda g, \mu x) \;=\; s_{\lambda\mu\, gx}^{-1}\,\lambda g\, s_{\mu x} \;=\; (\lambda\mu)^{-1}s_{gx}^{-1}\,\lambda g\,\mu s_x \;=\; \kappa(g, x).$$

Take $g = P(q^{1/2})$ and $x = K$. Since $P\big((\mu q)^{1/2}\big) = \mu P(q^{1/2})$, the displayed identity gives

$$\mathscr{U}(\mu q, \lambda K) \;=\; \kappa\big(\mu P(q^{1/2}),\ \lambda K\big) \;=\; \kappa\big(P(q^{1/2}),\ K\big) \;=\; \mathscr{U}(q, K)$$

by Theorem 9.2. The radial identity is $P\big((\mu q)^{1/2}\big)(\lambda K) = \mu\lambda P(q^{1/2})K$, and the descent statement is Theorem 3.2 with $\pi(\Omega_J) = \mathscr{B}_J^\circ$. □

This quotient is the normalization of Definition 3.1 itself, not a second one that resembles it. The multiplicative gauge $\mathbb{R}_{>0}$ of Theorem 3.2(b) — the right-hand vertical of Figure 1, and the image of the additive gauge $\mathbb{R}e$ under $\exp_J$ — is central in the cone group, and the polar section is equivariant for it, so it contributes nothing to κ. The projectivization that turns the cone into the state space and the projectivization that the transporter is blind to are one quotient appearing on both sides of Figure 1. The polar factorization accordingly splits cone-valued data into a channel that carries scale and a channel structurally incapable of carrying it, and effect normalization is exactly the projection onto the second. The coordinate $\log \mathrm{tr}_J S$ in $\mathscr{T}(q, K)$ is what puts the first one back.

That is the whole of the independence between the two channels. It does not survive composition.

Proposition 9.8 (The channels couple, and the coupling is the curvature). Write $\mathfrak{g}$ for the Lie algebra of $G = G(\Omega_J)$. Since Ω_J is self-dual, $\mathfrak{g}$ is stable under $X \mapsto -X^*$ for the trace form, and splits as

$$\mathfrak{g} \;=\; \mathfrak{k} \oplus \mathfrak{p}, \qquad \mathfrak{k} \;=\; \mathfrak{aut}(J) = \mathrm{Der}(J), \qquad \mathfrak{p} \;=\; \{X \in \mathfrak{g} \,:\, X^* = X\} \;\supseteq\; \{L_a \,:\, a \in J\},$$

with $[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}$; in particular $[L_a, L_b] \in \mathrm{Der}(J)$, and every derivation is such a sum [17, Ch. II, Prop. VI.1.2]. Then:

(a) the polar section is the exponential of the radial directions, $s_{\exp_J a} = \exp(L_a)$, so $\kappa(g, e)$ is the orthogonal factor in the polar decomposition of g; moreover $\kappa(k, x) = k$ for $k \in \mathrm{Aut}(J)$ and $\kappa\big(\exp(L_a), e\big) = \mathrm{id}$, so a single radial move away from the base point produces no compact factor at all;

(b) two of them do. Writing $g_i = P\big(q_i^{1/2}\big)$,

$$\kappa(g_1 g_2, K) \;=\; \mathscr{U}\big(q_1,\ S(q_2, K)\big)\,\mathscr{U}(q_2, K),$$

in which the left factor is evaluated at the aggregate score $S(q_2, K) = P\big(q_2^{1/2}\big)K$ that the second query has already moved the base point to, and not at K;

(c) the leading coupling is a bracket: for $a, b \in J$,

$$\frac{\partial^2}{\partial s\,\partial t}\bigg|_{s=t=0} \kappa\big(\exp(sL_a)\exp(tL_b),\ e\big) \;=\; \tfrac{1}{2}\big[L_a, L_b\big] \;\in\; \mathrm{Der}(J).$$

Proof. **Step 1** (*the section is the Cartan section*)**.** Let $a = \sum_k \alpha_k c_k$. By [17, Lem. IV.2.2] the operator $P\big(\sum_k \beta_k c_k\big)$ acts on $\mathbb{R}c_k$ by β_k^2 and on J_{ij} by $\beta_i\beta_j$, while by [17, Thm. IV.2.1] L_a acts on $\mathbb{R}c_k$ by α_k and on J_{ij} by $\tfrac{1}{2}(\alpha_i+\alpha_j)$. Taking $\beta_k = e^{\alpha_k/2}$, both $P\big(\exp_J(a/2)\big)$ and $\exp(L_a)$ act by e^{α_k} on $\mathbb{R}c_k$ and by $e^{(\alpha_i+\alpha_j)/2}$ on J_{ij}, so they agree on every Peirce block and hence on J; and $s_{\exp_J a} = P\big((\exp_J a)^{1/2}\big) = P\big(\exp_J(a/2)\big)$. As $s_e = \mathrm{id}$ we get $\kappa(g,e) = s_{ge}^{-1}g$, in which s_{ge} is positive definite self-adjoint and $\kappa(g,e) \in \mathrm{Aut}(J)$ is orthogonal by Theorem 9.2: that is the polar decomposition of g, which is unique. For $k \in \mathrm{Aut}(J)$, $P(ka) = kP(a)k^{-1}$ and $k(x^{1/2}) = (kx)^{1/2}$ give $s_{kx} = ks_xk^{-1}$, whence $\kappa(k,x) = (ks_xk^{-1})^{-1}ks_x = k$. Finally $\exp(L_a)e = P\big(\exp_J(a/2)\big)e = \exp_J a$, so $\kappa\big(\exp(L_a),e\big) = s_{\exp_J a}^{-1}\exp(L_a) = \mathrm{id}$.

Step 2 (*chaining*)**.** The cocycle identity of Theorem 9.2 at $x = K$ reads

$$\kappa(g_1g_2,K) \;=\; \kappa(g_1,\, g_2K)\,\kappa(g_2,K).$$

Now $g_2K = S(q_2,K)$, and $\kappa\big(P(q^{1/2}),y\big) = \mathscr{U}(q,y)$ for every $y \in \Omega_J$, again by Theorem 9.2.

Step 3 (*the bracket*)**.** By Step 1, $\kappa(\,\cdot\,,e)$ is the $\mathrm{Aut}(J)$-component in the decomposition $G = \exp(\mathfrak{p})\,\mathrm{Aut}(J)$, which is the polar decomposition. Each L_a is self-adjoint for the trace form and each derivation is skew, so $L_a, L_b \in \mathfrak{p}$ and $[L_a,L_b] \in \mathfrak{k}$. Baker–Campbell–Hausdorff gives $\exp(sL_a)\exp(tL_b) = \exp(Z)$ with $Z = sL_a + tL_b + \tfrac{1}{2}st[L_a,L_b] + O(3)$, where $O(3)$ collects terms of total degree at least three in (s,t). Splitting $Z = Z_{\mathfrak{p}} + Z_{\mathfrak{k}}$ along $\mathfrak{g} = \mathfrak{k}\oplus\mathfrak{p}$ puts $Z_{\mathfrak{p}} = sL_a + tL_b + O(3)$ and $Z_{\mathfrak{k}} = \tfrac{1}{2}st[L_a,L_b] + O(3)$. Since $\exp(Z_{\mathfrak{p}} + Z_{\mathfrak{k}}) = \exp(Z_{\mathfrak{p}})\exp(Z_{\mathfrak{k}})\exp\big(-\tfrac{1}{2}[Z_{\mathfrak{p}},Z_{\mathfrak{k}}] + \cdots\big)$ and $[Z_{\mathfrak{p}},Z_{\mathfrak{k}}] = O(3)$, the $\mathrm{Aut}(J)$-component is $\exp(Z_{\mathfrak{k}})$ modulo terms of total degree at least three. Applying $\partial_s\partial_t$ at the origin annihilates all of those and leaves $\tfrac{1}{2}[L_a,L_b]$. □

The bracket in (c) is the curvature datum of Ω_J as a Riemannian symmetric space for the invariant metric of Remark 2.1: in the standard normalization the curvature operator of G/K at the base point is $R(X,Y) = -\operatorname{ad}[X,Y]$ on $\mathfrak{p}$ [21, Ch. IV, §4]. So the radial and compact registers of Proposition 9.5 are independent only to first order. At second order the radial channel feeds the compact one, and the rate at which it does so is the curvature of the cone. Lemma 9.10 is the same bracket seen from the other side: its answer $4[L_z,L_{c_i}]$ is a commutator of two multiplication operators, which is what $[\mathfrak{p},\mathfrak{p}] \subseteq \mathfrak{k}$ makes available and the only thing it makes available.

9.5 THE JORDAN INSTRUMENT AND ITS ADJOINT

Effects pair with states to give numbers. To act on carrier-valued data they need the Jordan form of a Lüders operation [28], assembled over the whole effect family rather than one outcome at a time.

Proposition 9.9 (The Jordan instrument and its adjoint). Let $E_1,\dots,E_m \in \overline{\Omega}_J$ with $\sum_j E_j = e$. Give $J^{\oplus m}$ the componentwise Jordan product, the componentwise cone, the unit $(e,\dots,e)$, and the total trace

$$\mathrm{Tr}_{\oplus}(V_1,\dots,V_m) \;=\; \sum_j \mathrm{tr}_J\, V_j,$$

and define the *direct-sum Jordan instrument*

$$\mathfrak{I}_E : J \longrightarrow J^{\oplus m}, \qquad \mathfrak{I}_E(V) \;=\; \big(P(E_1^{1/2})V,\; \dots,\; P(E_m^{1/2})V\big). \tag{28}$$

Then:

(a) $\mathfrak{I}_E$ is positive and trace-preserving, $\mathrm{Tr}_{\oplus}\, \mathfrak{I}_E(V) = \mathrm{tr}_J\, V$;

(b) its adjoint for the two trace forms is $\mathfrak{I}_E^*(V_1, \dots, V_m) = \sum_j P(E_j^{1/2})V_j$;

(c) $\mathfrak{I}_E^*$ is positive and unital, $\mathfrak{I}_E^*(e, \dots, e) = e$;

(d) the common-value channel $T_E(V) = \mathfrak{I}_E^*(V, \dots, V)$ is positive, unital and trace-preserving;

(e) the heterogeneous output $\mathfrak{I}_E^*(V_1, \dots, V_m)$ lies in $\overline{\Omega}_J$ whenever every $V_j \in \overline{\Omega}_J$.

The whole construction is $\mathrm{Aut}(J)$-equivariant, through $gP(a)g^{-1} = P(ga)$ and $g(a^{1/2}) = (ga)^{1/2}$. For $J = \mathbb{R}$ the adjoint is the ordinary scalar weighted sum $\sum_j E_j V_j$, and for $J \cong \mathbb{R}^r$ the componentwise $\sum_j E_j \circ V_j$.

Proof. $P(a)$ preserves $\overline{\Omega}_J$ for $a \in \overline{\Omega}_J$ [17, Prop. III.2.2], which gives the positivity in (a) and, with convexity of the cone, (e); part (c) then follows from $P(E_j^{1/2})e = E_j$ and $\sum_j E_j = e$. Each $P(a)$ is self-adjoint for the trace form, so

$$\mathrm{Tr}_{\oplus}\, \mathfrak{I}_E(V) = \sum_j \big\langle e, P(E_j^{1/2})V \big\rangle = \sum_j \big\langle P(E_j^{1/2})e,\, V \big\rangle = \sum_j \langle E_j, V\rangle = \langle e, V\rangle = \mathrm{tr}_J\, V,$$

which is (a), and the same self-adjointness gives

$$\big\langle \mathfrak{I}_E(V),\, (W_1, \dots, W_m) \big\rangle_{\oplus} \;=\; \sum_j \big\langle P(E_j^{1/2})V,\, W_j \big\rangle \;=\; \Big\langle V,\; \sum_j P(E_j^{1/2})W_j \Big\rangle,$$

which is (b). Part (d) combines (c), (e) and the displayed trace computation read on the constant tuple. Equivariance is $gP(a)g^{-1} = P(ga)$ with $g(E^{1/2}) = (gE)^{1/2}$. In $\mathbb{R}$, $P(t)x = t^2 x$, so $P(E_j^{1/2})V_j = E_j V_j$; in $\mathbb{R}^r$ the same holds coordinatewise. □

Each property belongs to one map in particular. A single operation $V \mapsto P(E_j^{1/2})V$ is positive but in general neither unital nor trace-preserving; the instrument $\mathfrak{I}_E$ is positive and trace-preserving; its adjoint $\mathfrak{I}_E^*$ is positive and unital in the direct-sum sense; and the common-value channel T_E is the one map that is all three at once. Values outside the closed cone are handled by the canonical positive encoding of Section A, which routes the Jordan positive and negative parts through the doubled carrier.

An effect family fixes the outcome probabilities $\langle E_j, p\rangle$ but not the state-update rule, and many positive maps reproduce the same probabilities. Equation (28) is the canonical Jordan Lüders choice: it is built from the E_j alone through the quadratic representation, and is therefore $\mathrm{Aut}(J)$-equivariant and self-adjoint in each component. It is not the only instrument compatible with a given effect family.

The remaining subsections read the transporter through progressively more faithful readouts. Table 2 previews the outcome; Table 4 at the end of the section states it precisely.

READOUT	WHAT IT SEES OF THE FRAME
individual spectra	nothing
scalar overlaps	zero first variation; full second-order sensitivity at simple probe spectrum
effect frame, spectral code	the frame tangent, filtered by key contrast
transported Peirce projections	blockwise positive kernels; the total frame kernel degenerates exactly on $\ker \mathscr{A}_K$

Table 2. *Preview of the readout hierarchy developed in Sections 9.6, 9.8 and 9.9.*

The four steps now compose into a single construction,

$$A_{ij} = P\big(q_i^{1/2}\big)k_j, \qquad S_i = \sum_j A_{ij}, \qquad E_{ij} = P\big(S_i^{-1/2}\big)A_{ij}, \qquad O_i = \sum_j P\big(E_{ij}^{1/2}\big)V_j, \tag{29}$$

normalized by Proposition 9.1, equivariant throughout, and defined without reference to any externally supplied frame. What varies with the query is the transporter, so we treat the transporter first and choose a representation and a readout only afterwards.

Two hypotheses are in play in this subsection, and they are not the same hypothesis. The first constrains the key family only through its aggregate: fix $k_1, \dots, k_m \in \overline{\Omega}_J$ with m arbitrary, whose total

$$K \;=\; \sum_{j=1}^{m} k_j \;=\; \sum_{i=1}^{r} \kappa_i c_i \;\in\; \Omega_J, \qquad \kappa_i > 0, \tag{30}$$

is written in a Jordan frame (c_i) diagonalizing it; and write $\mathscr{A}_K(h) = D_q\mathscr{U}(e,K)[h]$ for the differential of the transporter at $q = e$. The second hypothesis aligns the individual keys with that frame,

$$m = r, \qquad k_i = \kappa_i c_i, \qquad \text{so that} \qquad \widehat{E}_i \;=\; P(K^{-1/2})k_i \;=\; c_i, \tag{31}$$

and it is a genuine restriction rather than a normalization: an attention row has $m \gg r$ and no reason for its keys to be rank one. Everything up to and including Proposition 9.12 uses only (30), because by Theorem 9.2 the transporter depends on the key family solely through K. The alignment (31) is first needed in Proposition 9.13(a), where the normalized effects are identified with the transported frame itself, and it is dropped again in Section 9.9.

The one calculation on which all of the above depends is the derivative of the transporter along a single Peirce block. It uses (30) only. We isolate it before using it.

Lemma 9.10 (Rank-two polar derivative). Let $K = \sum_k \kappa_k c_k \in \Omega_J$ with $\kappa_k > 0$, put $u_k = \sqrt{\kappa_k}$, fix $i < j$ and $h \in J_{ij}$, and set $q(t) = \exp_J(th)$ and $\mathscr{U}(t) := \mathscr{U}\big(q(t), K\big)$. Then

$$\dot{\mathscr{U}}(0) \;=\; 4\big[L_z, L_{c_i}\big], \qquad z \;=\; \frac{u_i - u_j}{2(u_i + u_j)}\,h.$$

Proof. **Step 1** (*the congruence factor*)**.** $q(t) = e + th + O(t^2)$, so $q(t)^{1/2} = e + t/2h + O(t^2)$. For $a = e + su$ one has $L_a = \mathrm{id} + sL_u$ and $a^{\circ 2} = e + 2su + O(s^2)$, so $P(a) = 2L_a^2 - L_{a^{\circ 2}} = \mathrm{id} + 2sL_u + O(s^2)$. With $u = h/2$, $s = t$,

$$P\big(q(t)^{1/2}\big) \;=\; \mathrm{id} + t\,L_h + O(t^2).$$

Step 2 (*the normalizer*)**.** By the Peirce relations $c_k \circ h = 0$ for $k \neq i, j$ and $c_i \circ h = c_j \circ h = {}^1\!/_2 h$, so $h \circ K = \sigma h$ with $\sigma = {}^1\!/_2(\kappa_i + \kappa_j)$. Hence $S(t) = P(q(t)^{1/2})K = K + t\sigma h + O(t^2)$. The derivative of a spectral function at K in a direction lying in J_{ij} is that direction times the divided difference of the function at κ_i, κ_j — this is Step 2 of Theorem 7.10 applied to $s \mapsto s^{-1/2}$ — so

$$\frac{d}{dt}\Big|_0 S(t)^{-1/2} \;=\; \sigma\,\frac{\kappa_i^{-1/2} - \kappa_j^{-1/2}}{\kappa_i - \kappa_j}\,h \;=\; -\rho\,h, \qquad \rho \;=\; \frac{\kappa_i + \kappa_j}{2\,u_i u_j\,(u_i + u_j)},$$

using $\kappa_i^{-1/2} - \kappa_j^{-1/2} = (u_j - u_i)/(u_i u_j)$ and $\kappa_i - \kappa_j = (u_i - u_j)(u_i + u_j)$.

Step 3 (*the product rule*)**.** Since $P(a) = 2L_a^2 - L_{a^{\circ 2}}$, the polarization $P(a,b) = L_a L_b + L_b L_a - L_{a \circ b}$ satisfies $\frac{d}{dt}P(a(t)) = 2P(a, \dot a)$. At $t = 0$ we have $q(0) = e$ and $S(0) = K$, so $\mathscr{U}(0) = P(K^{-1/2})P(K^{1/2}) = \mathrm{id}$, and

$$\dot{\mathscr{U}}(0) \;=\; -2\rho\,P\big(K^{-1/2}, h\big)P\big(K^{1/2}\big) \;+\; P\big(K^{-1/2}\big)L_h\,P\big(K^{1/2}\big). \tag{32}$$

Step 4 (*evaluate on the frame*)**.** $P(K^{\pm 1/2})$ acts on $\mathbb{R}c_k$ by $u_k^{\pm 2}$ and on J_{kl} by $(u_k u_l)^{\pm 1}$. For $n \neq i, j$: $L_h c_n = h \circ c_n = 0$ and $L_{K^{-1/2}} c_n = u_n^{-1} c_n$, so both terms of (32) annihilate c_n. For c_i, the second term gives $u_i^2\,P(K^{-1/2})({}^1\!/_2 h) = \frac{u_i}{2u_j}h$, while

$$P\big(K^{-1/2}, h\big)c_i = \tfrac{u_i^{-1}+u_j^{-1}}{4}h + \tfrac{1}{2u_i}h - \tfrac{u_i^{-1}+u_j^{-1}}{4}h = \tfrac{1}{2u_i}h,$$

so the first term gives $-2\rho u_i^2 \cdot \frac{1}{2u_i}h = -\rho u_i h$. Adding,

$$\dot{\mathscr{U}}(0)c_i \;=\; \frac{u_i}{2u_j}\Big(1 - \frac{u_i^2 + u_j^2}{u_i(u_i + u_j)}\Big)h \;=\; \frac{u_i - u_j}{2(u_i + u_j)}h \;=\; z,$$

and by the symmetry $i \leftrightarrow j$ (which fixes ρ and h), $\dot{\mathscr{U}}(0)c_j = -z$. These are exactly the values of $D_z := 4[L_z, L_{c_i}]$ on the frame, by (13).

Step 5 (*agreeing on the frame is enough here*)**.** $\dot{\mathscr{U}}(0)$ is the derivative at the identity of a curve in $\mathrm{Aut}(J)$, hence a derivation, and so is D_z; put $\Delta = \dot{\mathscr{U}}(0) - D_z$. A derivation killing every c_k preserves every Peirce block: for $x \in J_{kl}$, $c_k \circ \Delta x = \Delta(c_k \circ x) = {}^1\!/_2 \Delta x$, likewise for c_l, and $c_n \circ \Delta x = 0$ otherwise. On the other hand L_h and $L_{K^{-1/2} \circ h}$ are multiplications by elements of J_{ij}, so by the Peirce multiplication rules they carry each block into a *different* block; conjugating by the block-diagonal $P(K^{\pm 1/2})$ preserves that property, so both terms of (32) — and likewise D_z — have zero block-diagonal part. Hence Δ both preserves and shifts blocks, so $\Delta = 0$. □

Theorem 9.11 — The polar transporter is a spectral contrast filter on the frame tangent

For any key family with aggregate (30), and for $h = \sum_{i<j} h_{ij}$ with $h_{ij} \in J_{ij}$,

$$\mathscr{A}_K(h) \;=\; \sum_{i<j} 4\big[L_{z_{ij}}, L_{c_i}\big] \;\in\; \mathfrak{aut}(J), \qquad z_{ij} \;=\; \frac{\sqrt{\kappa_i}-\sqrt{\kappa_j}}{2\big(\sqrt{\kappa_i}+\sqrt{\kappa_j}\big)}\, h_{ij} \;=\; \tfrac{1}{2}\tanh\Big(\tfrac{\log\kappa_i-\log\kappa_j}{4}\Big) h_{ij} \tag{33}$$

and

$$\ker \mathscr{A}_K \;=\; \bigoplus_{\kappa_i=\kappa_j} J_{ij}, \qquad \operatorname{rank} \mathscr{A}_K \;=\; \sum_{\substack{i<j \\ \kappa_i\neq\kappa_j}} d_{ij}. \tag{34}$$

In particular, if the κ_i are pairwise distinct then $h \mapsto \mathscr{A}_K(h)$ is *injective* on the whole frame tangent $\bigoplus_{i<j} J_{ij} \cong T_c\mathfrak{F}_J$. For $J = H_3(\mathbb{O})$ this is an injection of the full 24-dimensional Albert frame tangent into $\mathfrak{f}_4$.
Equivalently, $\mathscr{A}_K$ factors as

$$\mathscr{A}_K \;=\; \iota_c \circ M_K, \tag{35}$$

where

$$\iota_c \;:\; \bigoplus_{i<j} J_{ij} \longrightarrow \operatorname{Der}(J), \qquad \iota_c(z) \;=\; \sum_{i<j} 4\big[L_{z_{ij}}, L_{c_i}\big],$$

is the canonical *frame-determined, aggregate-key-independent* injection of frame velocities into derivations — natural under frame transport, in the sense that $\iota_{gc}(gz) = g\,\iota_c(z)\,g^{-1}$ for every $g \in \operatorname{Aut}(J)$ — and M_K is the diagonal *spectral contrast filter*

$$M_K\big|_{J_{ij}} \;=\; \alpha_{ij}(K)\,\mathrm{id}_{J_{ij}}, \qquad \alpha_{ij}(K) \;=\; \frac{\sqrt{\kappa_i}-\sqrt{\kappa_j}}{2\big(\sqrt{\kappa_i}+\sqrt{\kappa_j}\big)} \;=\; \tfrac{1}{2}\tanh\Big(\tfrac{\log\kappa_i-\log\kappa_j}{4}\Big).$$

Thus $|\alpha_{ij}(K)| < 1/2$ for every key, and near a spectral collision

$$\alpha_{ij}(K) \;=\; \tfrac{1}{8}\big(\log\kappa_i - \log\kappa_j\big) \;+\; O\big((\log\kappa_i - \log\kappa_j)^3\big).$$

Proof. Formula (33) is Lemma 9.10 summed over blocks: the map $h \mapsto D_q\mathscr{U}(e,K)[h]$ is a differential, hence linear in h, so the one-block formula adds. The hyperbolic form is the identity

$$\frac{u_i-u_j}{2(u_i+u_j)} \;=\; \frac{e^{(\log\kappa_i)/2}-e^{(\log\kappa_j)/2}}{2\big(e^{(\log\kappa_i)/2}+e^{(\log\kappa_j)/2}\big)} \;=\; \tfrac{1}{2}\tanh\Big(\tfrac{\log\kappa_i-\log\kappa_j}{4}\Big), \qquad u_k = \sqrt{\kappa_k},$$

obtained by factoring $e^{(\log\kappa_i+\log\kappa_j)/4}$ out of numerator and denominator.

For the kernel, evaluate on the frame. By (13) the summand $4[L_{z_{ij}}, L_{c_i}]$ sends $c_i \mapsto z_{ij}$, $c_j \mapsto -z_{ij}$ and kills the other frame members, so

$$\mathscr{A}_K(h)\,c_k \;=\; \sum_{j>k} z_{kj} \;-\; \sum_{i<k} z_{ik}. \tag{36}$$

The summands on the right lie in the pairwise distinct Peirce blocks $J_{k\ell}$, $\ell \neq k$, which are mutually orthogonal. Hence $\mathscr{A}_K(h) = 0$ forces $\mathscr{A}_K(h)c_k = 0$ for every k, and projecting (36) onto

$J_{k\ell}$ forces every $z_{k\ell} = 0$. Since $\kappa_i \neq \kappa_j$ makes $u_i \neq u_j$ and hence the coefficient in (33) nonzero, $z_{ij} = 0$ gives $h_{ij} = 0$ on exactly those blocks; on blocks with $\kappa_i = \kappa_j$ the coefficient vanishes identically and h_{ij} is unconstrained. That is (34), and the rank count follows because the blocks are orthogonal. Finally $\mathscr{A}_K(h)$ is a sum of inner derivations, hence lies in $\mathfrak{aut}(J) = \mathrm{Der}(J)$, which is $\mathfrak{f}_4$ for $H_3(\mathbb{O})$ by Corollary 7.2. □

The factorization makes the theorem almost self-proving: ι_c is a fixed injection determined by the frame alone, M_K is a scalar gain on each Peirce edge, the kernel is the sum of the edges on which the aggregate key exhibits no contrast, and the rank is the sum of the surviving edge dimensions. The gain is an odd sigmoidal function of the log-eigenvalue gap, saturating in magnitude at $1/2$ and vanishing linearly at coincidence; in terms of the standard logistic $\sigma(t) = (1 + e^{-t})^{-1}$,

$$\alpha_{ij}(K) \;=\; \sigma\Big(\tfrac{\log\kappa_i - \log\kappa_j}{2}\Big) - \tfrac{1}{2}.$$

That same edge indexing runs through the whole paper. Table 3 collects the scalar each construction places on J_{ij}.

ROLE	COEFFICIENT OR BLOCK CONTRIBUTION	DEGREE
Gibbs Hessian (Theorem 7.10)	block multiplier $\mathscr{L}(w_i, w_j)$	1
Jacobian and Monge–Ampère source (Corollary 7.12, Corollary 7.13)	determinant contribution $\mathscr{L}(w_i, w_j)^{d_{ij}}$	d_{ij}
transporter differential (Theorem 9.11)	block multiplier $\alpha_{ij}(K)$	0
spectral-code readout (Corollary 9.14)	block multiplier $(a_i - a_j)\,\alpha_{ij}(K)$	0
relative-overlap curvature (Proposition 9.21)	curvature coefficient $-(\mu_i - \mu_j)^2$, contribution $-(\mu_i - \mu_j)^2\|z_{ij}\|^2$	2

Table 3. *The Peirce edge calculus. The principal differential, determinant, readout and curvature formulas that resolve frame directions share the same indexing by the edges $i < j$ of a Jordan frame; the block dimension d_{ij} enters only as a multiplicity. The last column is the degree of homogeneity of the coefficient in the pair of positive numbers it reads off the edge — the Gibbs weights (w_i, w_j), the aggregate-key eigenvalues (κ_i, κ_j), the probe weights (μ_i, μ_j) — and Remark 9.1 says why the transporter rows are the ones with degree zero. For $H_3(\mathbb{O})$ there are three edges, each of dimension eight, permuted by triality.*

Remark 9.1 (One edge calculus, graded by scale). The rows of Table 3 are not five unrelated coefficients. Each is a function of the two positive numbers sitting at the ends of the edge, and they differ by how that function responds to a common rescaling of the pair.

Write $\delta = \log u - \log v$ for the logarithmic gap of a positive pair. The chart's coefficient factors as a geometric mean times a shape function of the gap,

$$\mathscr{L}(u, v) \;=\; \sqrt{uv}\,\frac{\sinh(\delta/2)}{\delta/2}, \tag{37}$$

obtained by factoring $e^{(\log u+\log v)/2}$ out of numerator and denominator of $(u-v)/\delta$. The second factor is even in δ, equals 1 at coincidence, and increases in $|\delta|$, which is the inequality $\mathscr{L} \geq \sqrt{uv}$ read off the formula. The transporter's coefficient is the shape alone,

$$\alpha_{ij}(K) \;=\; \tfrac{1}{2}\tanh(\delta_\kappa/4), \qquad \delta_\kappa \;=\; \log\kappa_i - \log\kappa_j,$$

with no mean in front of it. So $\mathscr{L}$ is homogeneous of degree one and α_{ij} of degree zero, which is the last column of the table.

The degrees are not a coincidence between two formulas. By Proposition 9.7 the transporter cannot see the size of anything, $\mathscr{U}(\mu q, \lambda K) = \mathscr{U}(q, K)$, so every quantity built from $\mathscr{U}$ alone is a function of log-gaps and of nothing else; α_{ij} had no choice. The chart, by contrast, is the gradient of a potential, and Corollary 7.11 is a metric on a state space that has a size. The last row reads a probe state rather than a key: its coefficient $-(\mu_i - \mu_j)^2$ is a squared difference rather than a gap, homogeneous of degree two, because Proposition 9.21 measures a chord in J and not a spectral ratio.

Remark 9.2 (The logit row is itself a chart). Everything above concerns the carrier J that the values live in. The row of logits has a carrier too, and it is one this paper has already classified, so a single attention head contains two instances of the construction at once. It is worth saying which is which, because they are independent axes and neither asks anything of the other.

Fix a query and let the row range over L keys. By Theorem 4.1 the row softmax is the Gibbs chart of the simplicial carrier $\mathbb{R}^L$: the one-hot key selectors are its unique unordered Jordan frame, so the rank equals the number of keys, and the trace-one slice is the attention simplex Δ°_{L-1}. A head therefore normalizes twice on different carriers — once in $\mathbb{R}^L$ on the logit side, once in J on the value side by Proposition 9.1 — and the generalization pursued in this paper is the second. Nothing in Section 9 asks the logit row to become non-simplicial.

That identification is what makes the shift-invariance quotient available on the logit side. Consider an additive bias of coboundary type, $b_{ij} = \varphi_i - \psi_j$, a query term minus a key term. Within the row indexed by i the query term is constant in j, hence is $\varphi_i \mathbf{1}$ in $\mathbb{R}^L$, hence is invisible by Corollary 4.2; the key term survives as the class of $-\psi$ in $\mathbb{R}^L/\mathbb{R}\mathbf{1}$, nonzero exactly when ψ is nonconstant. So such a bias reduces to an absolute key prior rather than to nothing, and the reduction is not a property of the exponential: by Theorem 5.1, shift invariance already forces the Gibbs form among separable spectral normalizers of rank at least two, so sorting additive bias schemes by what survives is a statement about maps into $J/\mathbb{R}e$ and not about the choice of squashing function. The rank hypothesis is the condition $L \geq 2$, and Remark 5.1 is why it cannot be dropped.

The edge function of Remark 9.1 has a matching reading on that side. For every k,

$$\frac{d}{du}\Big[-k\log\big(2\cosh(u/2)\big)\Big] \;=\; -\tfrac{k}{2}\tanh(u/2), \tag{38}$$

so the one-parameter family of profiles $\big(2\cosh(u/2)\big)^{-k}$ has logarithmic derivatives that are tanh contrasts of gain $k/2$, with a minus sign because the profiles decay. The transporter's filter is the member of gain $1/2$, read at $u = \delta_\kappa/2$, with that sign reversed:

$$\alpha_{ij}(K) \;=\; -\frac{d}{du}\Big[-\log\big(2\cosh(u/2)\big)\Big]_{u=\delta_\kappa/2} \;=\; \tfrac{1}{2}\tanh(\delta_\kappa/4).$$

Equivalently, α_{ij} is the logarithmic derivative of the growing profile $2\cosh(u/2)$, the case $k=-1$ of the family, read at the same point. Differentiating once more in δ_κ gives $d\alpha_{ij}/d\delta_\kappa = {}^1\!/_8\operatorname{sech}^2(\delta_\kappa/4)$, which recovers the linearization $\alpha_{ij}(K) = {}^1\!/_8(\log\kappa_i - \log\kappa_j) + O(\delta_\kappa^3)$ already recorded in Theorem 9.11, the cubic coefficient being $-1/384$. What is shared with a positional profile is the family and the fact that both arguments are gaps on a logarithmic scale; the weight k is a free parameter in it, and α_{ij} fixes it at $k=1$, up to the sign recorded above. No numerical coincidence beyond that is claimed here.

Theorem 9.11 is intrinsic to the transporter and depends only on the aggregate key K: no hypothesis on the individual k_j beyond (30) enters its statement or its proof, and in particular m may exceed r by any amount. Representation-dependent visibility is treated next.

Proposition 9.12 (Representation-valued routing). Let $\rho : \mathrm{Aut}(J) \to GL(W)$ be a smooth finite-dimensional representation, let $v \in W$, and put $O_{\rho,v}(q) = \rho\big(\mathscr{U}(q,K)\big)v$. Then

$$DO_{\rho,v}(e)[h] \;=\; d\rho\big(\mathscr{A}_K(h)\big)v,$$

so

$$\ker DO_{\rho,v}(e) \;=\; \mathscr{A}_K^{-1}\big(\mathfrak{stab}_\rho(v)\big), \qquad \mathfrak{stab}_\rho(v) = \{X \in \mathfrak{aut}(J) \,:\, d\rho(X)v = 0\}. \tag{39}$$

Proof. $\mathscr{U}(e,K) = \mathrm{id}$ and ρ is smooth, so the chain rule along $q(t) = \exp_J(th)$ gives $DO_{\rho,v}(e)[h] = d\rho(\mathscr{A}_K(h))v$; the kernel description is immediate. □

Equation (39) separates three questions that a single-vector statement conflates: which frame transformations the *keys* produce (Theorem 9.11), which the *representation* detects, and which the *probe* v detects. Three choices of ρ are worth recording.

Proposition 9.13 (The effect frame and the traceless module). With the aligned keys (31) in part (a), and any key family with aggregate (30) in part (b):

(a) Each $E_i(q) = \mathscr{U}(q,K)c_i$ is a primitive idempotent and $E_\bullet(q) = (E_1(q),\dots,E_r(q))$ is an ordered Jordan frame, so the normalized effect tuple *is* a transported frame, $E_\bullet : \Omega_J \to \mathfrak{F}_J$. Its differential is $DE_\bullet(e)[h]_i = \mathscr{A}_K(h)c_i$ as in (36), and

$$\operatorname{rank} DE_\bullet(e) \;=\; \sum_{\substack{i<j \\ \kappa_i \neq \kappa_j}} d_{ij},$$

which for pairwise distinct κ_i is the entire frame tangent. Every $E_i(q)$ lies in $\overline{\Omega}_J$ and $\sum_i E_i(q) = e$.

(b) The representation $\rho_0(\mathscr{U}) = \mathscr{U}|_{J_0}$ on $J_0 = \ker \mathrm{tr}_J$ is faithful and orthogonal, so $R(q) = \mathscr{U}(q,K)|_{J_0} \in O(J_0)$ satisfies $\ker DR(e) = \ker \mathscr{A}_K$. For $H_3(\mathbb{O})$ this represents the transporter faithfully on the 26-dimensional traceless Albert module, and detects all 24 frame directions whenever the κ_i are distinct.

Proof. (a) $P(K^{-1/2})$ acts on $\mathbb{R}c_i$ by κ_i^{-1}, so $\widehat{E}_i = c_i$ and $E_i(q) = \mathscr{U}(q,K)c_i$ by (25); an automorphism carries a Jordan frame to a Jordan frame, and preserves the closed cone. Differentiating and using (36), $DE_\bullet(e)[h] = 0$ iff $\mathscr{A}_K(h)c_i = 0$ for all i, which by the argument in Theorem 9.11 is exactly $\ker \mathscr{A}_K$.

(b) Every automorphism fixes e and preserves the trace form, hence preserves J_0 and acts orthogonally on it; if it acts trivially on J_0 then, fixing e, it is trivial on $J = \mathbb{R}e \oplus J_0$, so ρ_0 is faithful. A derivation annihilates e, so $\mathscr{A}_K(h) = 0$ if and only if $\mathscr{A}_K(h)|_{J_0} = 0$. □

Part (a) is a carrier-valued readout that is positive, normalized and endogenous, and that reaches the whole frame tangent using the effect family alone. Introducing values changes the rank, and the change is exactly computable.

Corollary 9.14 (Rank of a spectral-code readout). Take the keys (31) and central values $V_i = a_i e$ with $a_i \in \mathbb{R}$. Then the construction (29) returns

$$O(q) \;=\; \sum_i a_i E_i(q) \;=\; \mathscr{U}(q,K)\,v, \qquad v = \sum_i a_i c_i, \tag{40}$$

whose differential at the identity query is $DO(e)[h] = \sum_{i<j}(a_i - a_j)z_{ij}$, and therefore

$$\operatorname{rank} DO(e) \;=\; \sum_{\substack{i<j \\ a_i \neq a_j,\ \kappa_i \neq \kappa_j}} d_{ij}. \tag{41}$$

Put an edge ij in the *key-contrast graph* when $\kappa_i \neq \kappa_j$ and in the *readout-contrast graph* when $a_i \neq a_j$. Then (41) says that the visible tangent is $\bigoplus J_{ij}$ over the edges present in both, block by block; for $H_3(\mathbb{O})$ every surviving edge contributes 8 dimensions. When the a_i are pairwise distinct, v is a *simple-spectrum code* for the transported frame: the spectral idempotents of $\mathscr{U}v$ are then the $E_i(q)$, so the output determines the transported frame and loses nothing relative to Proposition 9.13(a). The condition $a_i > 0$ makes that code positive, $v \in \Omega_J$ and $O(q) \in \Omega_J$. Thus the a_i-conditions belong to the probe, whereas the intrinsic transporter rank is governed solely by the aggregate key (Theorem 9.11). Since the corollary allows $a_i \in \mathbb{R}$, the genericity statement is that full-frame visibility is open and dense in $(\mathbb{R}_{>0})^r \times \mathbb{R}^r$; if a positive carrier-valued output is additionally required, the positive simple-spectrum locus is open and dense in $(\mathbb{R}_{>0})^{2r}$. In both, the hypothesis on the keys is simple spectrum of the aggregate key $K = \sum_i \kappa_i c_i$, whose spectral coefficients the κ_i are; it is not a condition on the eigenvalues of the individual rank-one keys separately.

Proof. $E_i(q)$ is idempotent, so $E_i^{1/2} = E_i$ and $P(E_i)e = E_i$, giving $P(E_i(q)^{1/2})(a_i e) = a_i E_i(q)$ and hence (40) by linearity of $\mathscr{U}(q,K)$. This is Proposition 9.12 with ρ the natural action on $W = J$ and probe v; applying $\mathscr{A}_K(h)$ to $v = \sum_i a_i c_i$ and using (36) gives $DO(e)[h] = \sum_{i<j}(a_i - a_j)z_{ij}$, and the blocks being orthogonal, the rank is the sum of d_{ij} over the blocks whose coefficient $(a_i - a_j)(u_i - u_j)/2(u_i + u_j)$ is nonzero, which is (41). If the a_i are distinct then v has simple spectrum with spectral idempotents c_i, so the spectral idempotents of $\mathscr{U}v$ are $\mathscr{U}c_i = E_i(q)$. If all $a_i > 0$ then $v \in \Omega_J$ and $\mathscr{U}(q,K) \in \mathrm{Aut}(J)$ preserves Ω_J. Finally, the conditions $\kappa_i \neq \kappa_j$ and $a_i \neq a_j$ for $i < j$ each remove a finite union of proper linear subspaces, so their joint complement is open and dense in $(\mathbb{R}_{>0})^r \times \mathbb{R}^r$, and intersecting with $a \in (\mathbb{R}_{>0})^r$ gives the second statement. □

The representation-theoretic consequence of (40) is immediate, and it constrains the *value interface* rather than the kernel.

Remark 9.3 (Trivial-type confinement under scalar aggregation). Scalar aggregation cannot leave the trivial isotypic component carried by a central value; the carrier-valued effect action can. Concretely, if W is an $\mathrm{Aut}(J)$-module and $0 \neq v_0 \in W^{\mathrm{Aut}(J)}$, then any value interface forming only scalar-weighted combinations $\sum_j \pi_j(q,k)V_j$ of values $V_j = a_j v_0$ has output in the line $\mathbb{R}v_0$, whatever the π_j. For $W = J$ and $v_0 = e$ this says that scalar aggregation of central values stays in $\mathbb{R}e$, whereas the spectral-code output $O(q) = \mathscr{U}(q,K)\sum_i a_i c_i$ of (40) has a nonzero J_0-component whenever the a_i are not all equal: an automorphism fixes e and preserves $J = \mathbb{R}e \oplus J_0$, so $\mathscr{U}(q,K)v \in \mathbb{R}e$ only if $v \in \mathbb{R}e$.

The synthesis map $\Phi_c : \mathbb{R}^r \to J$, $\Phi_c(a) = \sum_i a_i c_i$, makes the split explicit. It is linear and injective with $\Phi_c(\mathbf{1}) = e$ and $\mathrm{tr}_J\, \Phi_c(a) = \sum_i a_i$, so under $\mathbb{R}^r = \mathbb{R}\mathbf{1} \oplus \mathbf{1}^\perp$,

$$\Phi_c(\mathbb{R}\mathbf{1}) \;=\; \mathbb{R}e, \qquad \Phi_c(\mathbf{1}^\perp) \;=\; \mathrm{span}(c_1,\dots,c_r) \cap J_0,$$

identifying the noncentral branch as the transported standard component of the coefficient vector.

Remark 9.4 (What is endogenous). The frame (c_i) is not supplied to the mechanism. Under (30) it is a spectral frame of the aggregate key, intrinsic up to ordering when that spectrum is simple; under (31) it is in addition the frame the individual keys point along. In both cases it appears only in the proofs: the layer (29) reads k_j, q_i, V_j and nothing else.

9.7 EXACT ADDITIVE MOMENTS FOR CENTRAL-VALUE ROUTING

The normalizing operator depends on the key row only through the additive total

$$K \;=\; \sum_j k_j \;\in\; J, \qquad K_{B \sqcup C} \;=\; K_B + K_C,$$

a single element of J with an exactly associative merge law: by Theorem 9.2 the factor $P(S^{-1/2})$ is fixed by K together with the query through $S = P(q^{1/2})K$, so the keys can be reduced once and transformed per query. The individual effects E_j still require the individual k_j. For central values, adjoining $M = \sum_j a_j k_j$ gives an exact sufficient pair for the full output.

Theorem 9.15 — Exact additive moments for central-value polar routing

Let $k_1,\dots,k_m \in \overline{\Omega}_J$ have aggregate $K = \sum_j k_j \in \Omega_J$, let $a_j \in \mathbb{R}$, put $V_j = a_j e$, and set

$$M \;=\; \sum_j a_j k_j \;\in\; J.$$

For $q \in \Omega_J$ write $S(q) = P(q^{1/2})K$ and $E_j(q) = P\big(S(q)^{-1/2}\big)P\big(q^{1/2}\big)k_j$. Then the routing output of (29), namely the adjoint instrument $\mathfrak{T}^*_{E(q)}$ evaluated at $(V_1,\dots,V_m)$, is

$$O(q) \;=\; \sum_j P\big(E_j(q)^{1/2}\big)(a_j e) \;=\; \sum_j a_j E_j(q) \;=\; P\big(S(q)^{-1/2}\big)P\big(q^{1/2}\big)M. \tag{42}$$

Theorem 9.15 — Exact additive moments for central-value polar routing (continued)

The central-value branch therefore depends on the sequence only through the two additive carrier moments

$$(K, M) \;=\; \Big(\sum_j k_j,\; \sum_j a_j k_j\Big) \;\in\; J \oplus J,$$

with the exact block merge law

$$(K, M)_{B \sqcup C} \;=\; \big(K_B + K_C,\; M_B + M_C\big).$$

If every $a_j \geq 0$ then $M \in \overline{\Omega}_J$ and the output is cone-valued. For d_v scalar value channels, $a_j \in \mathbb{R}^{d_v}$, the same statement holds with M replaced by

$$\mathsf{M} \;=\; \sum_j k_j \otimes a_j \;\in\; J \otimes \mathbb{R}^{d_v},$$

the output in channel α being $P\big(S^{-1/2}\big)P\big(q^{1/2}\big)\mathsf{M}_\alpha$.

Proof. For any $E \in \overline{\Omega}_J$ one has $P(E^{1/2})e = (E^{1/2})^{\circ 2} = E$, so $P(E_j(q)^{1/2})(a_j e) = a_j E_j(q)$; no idempotency of E_j is needed. Summing and using linearity of $P(S^{-1/2})P(q^{1/2})$,

$$\sum_j a_j E_j(q) \;=\; P\big(S^{-1/2}\big)P\big(q^{1/2}\big) \sum_j a_j k_j \;=\; P\big(S^{-1/2}\big)P\big(q^{1/2}\big)M,$$

which is (42). The normalizer $S = P(q^{1/2})K$ depends on the sequence only through K, so the right-hand side depends on it only through (K, M); both entries are sums over the block, which is the merge law. If every $a_j \geq 0$ then M is a nonnegative combination of elements of $\overline{\Omega}_J$, hence lies in $\overline{\Omega}_J$, and $P(S^{-1/2})P(q^{1/2})$ is a cone automorphism. The multichannel statement is (42) read in each channel, the operator acting on the J factor of $J \otimes \mathbb{R}^{d_v}$. □

This exact two-moment reduction applies to central-value or scalar-channel routing. It does not collapse the general heterogeneous J-valued instrument, whose dependence on the individual square roots $E_j^{1/2}$ remains nonlinear.

For $H_3(\mathbb{O})$ the pair (K, M) is 54 real accumulators in a coordinate implementation: exactly associative in exact arithmetic, independent of sequence length after reduction, and computed once for the keys and once for the key–value products. The corresponding statement for a classical readout is a single tensor contraction, recorded in Section B.

The curvature computations in this subsection and the next are all the same computation. We isolate it once.

Lemma 9.16 (Curvature of an orthogonal orbit). Let $\rho : G \to O(H)$ be a smooth orthogonal representation on a Euclidean space H, let $g(t)$ be a C^2 curve in G with $g(0) = 1$, and set $X = \frac{d}{dt}\big|_0 \rho(g(t))$. For $v \in H$ put $f_v(t) = \langle v, \rho(g(t))v\rangle$. Then

$$f_v'(0) \;=\; 0, \qquad f_v''(0) \;=\; -\|Xv\|^2,$$

so f_v is stationary at coincidence with negative semidefinite second variation, degenerate exactly when $Xv = 0$. If $g(t) = \exp(tX_0)$ is a one-parameter subgroup, f_v is even.

Proof. $\rho(g(t))$ is orthogonal, so $\|\rho(g(t))v\| = \|v\|$ and $\|\rho(g(t))v - v\|^2 = 2\|v\|^2 - 2f_v(t)$. Writing $u(t) = \rho(g(t))v$ we have $u(0) = v$ and $\dot u(0) = Xv$, so the left-hand side is $t^2\|Xv\|^2 + O(t^3)$ and therefore $f_v(t) = \|v\|^2 - \frac{t^2}{2}\|Xv\|^2 + O(t^3)$. For a one-parameter subgroup $\rho(g(t))^{-1} = \rho(g(t))^\top$, so $f_v(-t) = \langle \rho(g(t))v, v\rangle = f_v(t)$. □

Proposition 9.17 (Orbit kernels). Let $\rho : \mathrm{Aut}(J) \to O(H)$ be a smooth orthogonal representation and $v \in H$ a probe. Then

$$k_v(q, q') \;=\; \big\langle \rho(\mathcal{U}(q, K))v,\ \rho(\mathcal{U}(q', K))v \big\rangle$$

is a positive semidefinite kernel on Ω_J with $k_v(q, q) = \|v\|^2$, and along $q(t) = \exp_J(th)$,

$$k_v\big(e, q(t)\big) \;=\; \|v\|^2 - \tfrac{t^2}{2}\big\|d\rho(\mathscr{A}_K(h))v\big\|^2 + O(t^3),$$

so the second variation degenerates exactly on $h \in \mathscr{A}_K^{-1}\big(\mathfrak{stab}_\rho(v)\big)$.

Proof. Positive semidefiniteness is the Gram identity $\sum_{k,\ell} a_k a_\ell k_v(q_k, q_\ell) = \|\sum_k a_k \rho(\mathcal{U}(q_k, K))v\|^2$, and $k_v(q, q) = \|v\|^2$ by orthogonality. The expansion is Lemma 9.16 applied to $g(t) = \mathcal{U}(q(t), K)$, whose derivative at $t = 0$ is $\mathscr{A}_K(h)$, so that $X = d\rho(\mathscr{A}_K(h))$; the degeneracy condition $Xv = 0$ says exactly that $\mathscr{A}_K(h)$ lies in the infinitesimal stabilizer of v. □

The rest of this subsection and all of Section 9.9 are instances. Proposition 9.21 is the natural representation on $H = J$ with a state as probe; Theorem 9.24 is the conjugation representation on $\mathrm{End}(J)$ with a Peirce projection as probe; Proposition 9.25 packages several probes into one linear feature map. What distinguishes the Peirce projections is that the carrier supplies them: they are the canonical probe-free choice.

Pairing the effects against a state instead of acting on values gives the classical half of the same picture, and it is scalar-valued. Along continuous frame motions it has vanishing first variation at the aligned configuration and negative semidefinite second variation; the latter is negative definite on the full frame tangent when the probe state has simple spectrum. Independently of continuity, the discrete exchange below already separates configurations on every carrier of rank at least two.

Definition 9.18 (Spectrum-only mechanism). A map F defined on m-tuples of elements of J is *spectrum-only* if it factors through the individual spectra,

$$F(x_1, \dots, x_m) \;=\; \widetilde{F}\big(\operatorname{spec}(x_1), \dots, \operatorname{spec}(x_m)\big),$$

each spectrum taken as an unordered multiset of eigenvalues with multiplicity.

Proposition 9.19 (Separation beyond individual spectra). Let J be *any* Euclidean Jordan algebra of rank $r \geq 2$, with Jordan frame $(c_1, \dots, c_r)$; let $\mu_1 > \mu_2 > 0$ and $\mu_k > 0$ for $k \geq 3$ with $\mu_1 + \mu_2 + \sum_{k\geq 3} \mu_k = 1$, so that p, q, q' below are genuine states of $\mathscr{B}_J^\circ$; and let W be a real vector space carrying a linear $\mathrm{Aut}(J)$-action, with values $V_1, V_2 \in W$, $V_1 \neq V_2$. Put

$$p = \sum_k \mu_k c_k, \qquad q = p, \qquad q' = \mu_2 c_1 + \mu_1 c_2 + \textstyle\sum_{k\geq 3} \mu_k c_k,$$

and take the two-key effect families $A_1 = q$, $A_2 = e - q$ and $A_1' = q'$, $A_2' = e - q'$. Then $S = S' = e$, so Proposition 9.1 leaves them unchanged, and with query state $\rho = p$ the outputs $O = \sum_j \langle p, E_j\rangle V_j$ and O' satisfy

$$O - O' \;=\; (\mu_1 - \mu_2)^2\,(V_1 - V_2) \;\neq\; 0.$$

The effects are positive and normalized, and the resulting probabilities $\langle p, E_j\rangle$ are nonnegative and sum to one. They are invariant under simultaneous $\mathrm{Aut}(J)$-transport of the query state and the effects, and transporting the values makes the output equivariant, $O \mapsto gO$; taking $W = \mathbb{R}$ with the trivial action gives an invariant scalar output. Since q and q' have identical spectra, every spectrum-only mechanism in the sense of Definition 9.18 returns $O = O'$. Specializations:

(a) for frame-transitive J the two states q, q' lie in the *same* one-state $\mathrm{Aut}(J)$-orbit (Proposition 7.6), so the distinction is genuinely relational and not a disguised one-state invariant;

(b) for $J = H_3(\mathbb{O})$ the construction is F_4-equivariant;

(c) for $J \cong \mathbb{R}^r$ only the discrete exchange is available, the two configurations differing by a transposition of coordinates;

(d) for every carrier with $J_{12} \neq 0$ the exchange is not merely spectrum-preserving but *connected* to the identity configuration: by Lemma 9.20 there is an isospectral path $q(t) = e^{tD}q$ inside a single rank-two orbit with $q(0) = q$ and $q(\pi/2) = q'$, along which the readout varies nonconstantly by Proposition 9.21.

Proof. Since $0 \leq \mu_k < 1$ for every k, both q and $e - q = \sum_k (1 - \mu_k)c_k$ lie in $\overline{\Omega}_J$, and $A_1 + A_2 = e \in \Omega_J$; likewise for q'. As $P(e^{-1/2}) = \mathrm{id}$, the normalized effects are $E_1 = q$, $E_2 = e - q$, and correspondingly for q'. Using $\langle c_i, c_j\rangle = \delta_{ij}$,

$$\langle p, q\rangle = \sum_k \mu_k^2, \qquad \langle p, q'\rangle = 2\mu_1\mu_2 + \sum_{k\geq 3} \mu_k^2,$$

differing by $(\mu_1 - \mu_2)^2$ as in Proposition 7.9. With two outcomes, $\langle p, E_2\rangle = 1 - \langle p, E_1\rangle$, so $O = \langle p, E_1\rangle(V_1 - V_2) + V_2$ in W, and the difference is as claimed. Invariance of the probabilities is Proposition 7.1 together with Proposition 9.1; equivariance of O follows from linearity of the $\mathrm{Aut}(J)$-action on W. Finally q and q' have the same eigenvalue multiset, and p is common

to both configurations. Nothing in the argument uses the octonions, or any property of J beyond orthonormality of a Jordan frame — which is why the statement is carrier-independent. For (a), q and q' have equal spectra, so on a frame-transitive carrier the frame transitivity carries one to the other; the configurations (p, q) and (p, q') are nevertheless inequivalent, because a simultaneous automorphism would have to fix p and preserve the pairing. Item (b) is Corollary 7.2, and (c) is Theorem 7.4(c). For (d), apply Lemma 9.20 with $(i, j) = (1, 2)$ to any $z \in J_{12}$ with $\|z\|^2 = 2$: the endpoint $q(\pi/2)$ is exactly the q' displayed above, and Proposition 9.21 gives $\frac{d^2}{dt^2}\big|_0 \langle p, q(t)\rangle = -(\mu_1 - \mu_2)^2\|z\|^2 < 0$, so the readout is nonconstant along the path. The discrete exchange is therefore a Weyl transposition realized by an inner-derivation flow, not a separate phenomenon. □

Remark 9.5 (Arbitrary aggregate scores). In Proposition 9.19 the aggregate score is $S = e$ and $P(S^{-1/2}) = \mathrm{id}$, so what the statement exhibits is frame-sensitive effect pairing. The same separation holds with an arbitrary aggregate score. Given any $S \in \Omega_J$, put

$$A_1 = P\big(S^{1/2}\big)q, \qquad A_2 = P\big(S^{1/2}\big)(e - q),$$

both in $\overline{\Omega}_J$ because $P(S^{1/2})$ preserves the cone, with $A_1 + A_2 = P(S^{1/2})e = S$. Proposition 9.1 then returns $E_1 = q$ and $E_2 = e - q$, so the separation is unchanged while the normalizer acts nontrivially.

The separation is not an artifact of the discrete exchange: it survives, at second order, along a continuous spectrum-preserving motion of the frame.

The continuous motion of Proposition 9.21 below is local at $t = 0$. The following lemma supplies the global endpoint, and with it the exact transposition of Proposition 9.19 on a single rank-two orbit.

Lemma 9.20 (Rank-two exchange). Fix $i < j$ with $J_{ij} \neq 0$ and let $z \in J_{ij}$ with $\|z\|^2 = 2$. Put $D = 4[L_z, L_{c_i}]$, $s = c_i + c_j$ and $a = c_i - c_j$. Then $z^{\circ 2} = s$ and

$$D(c_i) = z, \qquad D(c_j) = -z, \qquad D(z) = 2(c_j - c_i), \qquad D(c_k) = 0 \ (k \neq i, j),$$

so D annihilates s and rotates the plane $\mathrm{span}\{a, z\}$ at angular speed 2. Consequently

$$e^{tD}c_i = \frac{s}{2} + \frac{\cos 2t}{2}\, a + \frac{\sin 2t}{2}\, z, \qquad e^{tD}c_j = \frac{s}{2} - \frac{\cos 2t}{2}\, a - \frac{\sin 2t}{2}\, z,$$

and at $t = \pi/2$ the automorphism $e^{(\pi/2)D} \in \mathrm{Aut}(J)$ exchanges c_i and c_j and fixes every other frame member. Hence for $p = \sum_k \mu_k c_k$ the isospectral path $q(t) = e^{tD}p$ satisfies

$$q(0) = p, \qquad q(\pi/2) = \mu_j c_i + \mu_i c_j + \sum_{k \neq i,j} \mu_k c_k.$$

Proof. For $z \in J_{ij}$ one has $z^{\circ 2} = {}^1\!/_2\|z\|^2(c_i + c_j)$ [17, Ch. IV], which is s under the normalization $\|z\|^2 = 2$. The values on the frame are (13). For $D(z)$, using $c_i \circ z = {}^1\!/_2 z$,

$$D(z) = 4\big(z \circ (c_i \circ z) - c_i \circ z^{\circ 2}\big) = 4\big(\tfrac{1}{2}z^{\circ 2} - c_i \circ s\big) = 4\big(\tfrac{1}{2}s - c_i\big) = 2(c_j - c_i).$$

Thus $D(s) = z - z = 0$, $D(a) = 2z$ and $D(z) = -2a$, so D acts on $\mathrm{span}\{a, z\}$ as the generator of a rotation of angular speed 2 and $e^{tD}a = \cos(2t)a + \sin(2t)z$. Writing $c_i = {}^1\!/_2(s + a)$ and $c_j = {}^1\!/_2(s - a)$

gives the displayed flows; at $t = \pi/2$, $\cos 2t = -1$ and $\sin 2t = 0$, so $e^{(\pi/2)D} c_i = {}^1\!/_2(s-a) = c_j$ and symmetrically. Since D is an inner derivation, $e^{tD} \in \mathrm{Aut}(J)$ for every t, so $q(t)$ is isospectral; evaluating at $t = \pi/2$ termwise gives the endpoint. □

Proposition 9.21 (Relative-frame overlap has negative-definite second variation). Let $p = \sum_k \mu_k c_k \in \mathscr{B}_J^\circ$ have simple spectrum $\mu_1 > \cdots > \mu_r$. Let $z = (z_{ij})_{i<j} \in \bigoplus_{i<j} J_{ij}$ be any nonzero element of the frame tangent, set

$$D_z \;=\; \sum_{i<j} 4\big[L_{z_{ij}}, L_{c_i}\big], \qquad q(t) = \exp(tD_z)\, p.$$

Then $q(t)$ has constant spectrum $\{\mu_1, \dots, \mu_r\}$ for *all* t, both $q(t)$ and $e - q(t)$ remain in $\overline{\Omega}_J$ for all t, $D_z p = \sum_{i<j}(\mu_i - \mu_j) z_{ij}$, and

$$\frac{d^2}{dt^2}\Big|_{t=0} \langle p, q(t)\rangle \;=\; -\,\|D_z p\|^2 \;=\; -\sum_{i<j} (\mu_i - \mu_j)^2\, \|z_{ij}\|^2 \;<\; 0.$$

The second variation is therefore negative definite on the entire frame tangent, not merely on one Peirce block. Moreover $t \mapsto \langle p, q(t)\rangle$ is an even function, so its first derivative vanishes at $t = 0$ and

$$\langle p, q(t)\rangle \;=\; \langle p, p\rangle \;-\; \tfrac{1}{2} \sum_{i<j} (\mu_i - \mu_j)^2 \|z_{ij}\|^2\, t^2 \;+\; O(t^4),$$

the displayed second derivative being the leading variation. In the setting of Proposition 9.19, the first-outcome probability of the layer built from $A_1 = q(t)$, $A_2 = e - q(t)$ is a nonconstant function of t for every nonzero frame tangent z, while each individual spectrum in the configuration is held fixed.

Proof. D_z is a sum of inner derivations, hence a derivation, so $\exp(tD_z) \in \mathrm{Aut}(J)$ and $q(t)$ is $\mathrm{Aut}(J)$-conjugate to p for every t; automorphisms preserve the spectrum and fix e, so $q(t)$ and $e - q(t) = \exp(tD_z)(e - p)$ lie in $\overline{\Omega}_J$ for *all* t, not merely small t, and $A_1 + A_2 = e$ throughout. By (13) the summand $4[L_{z_{ij}}, L_{c_i}]$ sends $c_i \mapsto z_{ij}$, $c_j \mapsto -z_{ij}$ and kills every other frame member, so $D_z p = \sum_{i<j}(\mu_i - \mu_j) z_{ij}$; the Peirce blocks are mutually orthogonal, so $\|D_z p\|^2 = \sum_{i<j}(\mu_i - \mu_j)^2 \|z_{ij}\|^2$, which is strictly positive because the spectrum is simple and $z \neq 0$.

The remaining assertions are Lemma 9.16 applied to the natural representation of $\mathrm{Aut}(J)$ on $H = J$, with $g(t) = \exp(tD_z)$ and probe $v = p$: the first variation vanishes, the second is $-\|D_z p\|^2$, and f is even because g is a one-parameter subgroup, so the expansion has no odd terms. □

Corollary 9.22 (Albert triality channels). For $J = H_3(\mathbb{O})$ the three Peirce blocks J_{12}, J_{13}, J_{23} are each 8-dimensional and together span the 24-dimensional frame tangent of Theorem 7.3, so Proposition 9.21 reaches every frame direction at once, with second-derivative magnitude $\sum_{i<j}(\mu_i - \mu_j)^2 \|z_{ij}\|^2$ — the coefficient of t^2 in the Taylor expansion being half that — and the mechanism of Proposition 9.19 is F_4-equivariant. The three blocks are moreover inequivalent as modules for the frame stabilizer: the subgroup of F_4 fixing an ordered Jordan frame pointwise is Spin(8), acting on (J_{23}, J_{13}, J_{12}) through a triality triple of eight-dimensional representations, so that

$$T_c \mathfrak{F}_{H_3(\mathbb{O})} \;=\; J_{12} \oplus J_{13} \oplus J_{23} \;\cong\; \mathbf{8}_v \oplus \mathbf{8}_s \oplus \mathbf{8}_c$$

up to a permutation of the labels, the outer S_3 of Spin(8) permuting the three [50, Thms. 1.16.2, 1.16.4, 2.7.1]. The transporter differential (33) scales these three channels by

$$\alpha_{ij} \;=\; \frac{\sqrt{\kappa_i}-\sqrt{\kappa_j}}{2\big(\sqrt{\kappa_i}+\sqrt{\kappa_j}\big)},$$

and (41) specializes to

$$\operatorname{rank} DO(e) \;=\; 8\,\#\big\{ij \;:\; \kappa_i \neq \kappa_j,\ a_i \neq a_j\big\}.$$

For $J \cong \mathbb{R}^r$, by contrast, Proposition 9.19 still applies — the discrete exchange is a transposition of coordinates — but every J_{ij} vanishes, so Proposition 9.21 is vacuous and no continuous spectrum-preserving frame motion exists at all.

Proof. The block dimensions and their sum are Theorem 7.3 with $r = 3$, $d = 8$; F_4-equivariance is Corollary 7.2; and $J_{ij} = 0$ for $\mathbb{R}^r$ is Theorem 7.4(c). For the module structure, an automorphism fixing every c_i preserves each Peirce block, and in the standard matrix model of $H_3(\mathbb{O})$ those blocks are the three off-diagonal octonion slots. The stabilizer of the diagonal frame in F_4 is Spin(8), realized as the group of triples $(\alpha_1, \alpha_2, \alpha_3) \in SO(8)^3$ satisfying $(\alpha_1 x)(\alpha_2 y) = \alpha_3(xy)$ and acting on the three slots by α_1, α_2, α_3 respectively [50, Thms. 1.16.2 and 2.7.1]; the first component is the vector representation and the other two the half-spin representations, and the outer automorphism group S_3 permutes them [50, Thm. 1.16.4]. The infinitesimal form is the principle of triality in $\mathfrak{so}(8)$ [50, Thm. 1.3.6, Prop. 2.3.7]. The rank formula is (41) with every $d_{ij} = 8$. □

Each surviving edge of the three-slot contrast graph therefore carries one of the three triality modules. What the exceptional carrier contributes is neither twenty-four extra coordinates nor the size of F_4: it is a rank-three system of three coupled eight-dimensional Peirce channels, permuted by the triality symmetry of their common stabilizer. The classical separation of Proposition 9.19, by contrast, is not exceptional at all: it holds on every carrier of rank at least two, the simplicial one included. What the exceptional carrier supplies is continuity — a 24-dimensional space of spectrum-preserving directions along which to take that separation, each weighted by the metric of Corollary 7.11.

What a scalar feature family sees of the transporter is settled by the same differential. For a representation $\rho : \operatorname{Aut}(J) \to GL(W)$, a probe $v \in W$ and a functional $\phi \in W^*$, the scalar matrix coefficient

$$f_{\phi,v}(q,K) \;=\; \phi\big(\rho(\mathcal{U}(q,K))v\big)$$

has differential

$$Df_{\phi,v}(e)[h] \;=\; \phi\big(d\rho(\mathscr{A}_K(h))v\big)$$

by Proposition 9.12. A chosen scalar feature family is locally complete on the frame tangent exactly when its matrix coefficients separate the subspace $d\rho\big(\mathscr{A}_K(T_c\mathfrak{F}_J)\big)$. Scalarization is therefore not a binary expressivity class: frame visibility is controlled by the representation types and the coefficient functionals a readout retains. Table 4, at the end of the section, collects the resulting hierarchy.

Every readout so far scalarizes the transporter by pairing it against a chosen probe. A different scalarization keeps an entire Peirce block and pairs two transported copies of it. It needs no probe, and its positivity is a Gram identity rather than a computation.

Only the aggregate key and one of its spectral frames enter, so this subsection drops the rank-one hypothesis on the individual keys. Fix $K \in \Omega_J$ *with simple spectrum*, and let (c_i) be its spectral frame, $K = \sum_i \kappa_i c_i$ with the $\kappa_i > 0$ pairwise distinct; that frame is then intrinsic up to ordering. Remark 9.7 treats repeated spectrum. Set

$$\Pi_{ij} \;=\; 4L_{c_i}L_{c_j} \quad (i \neq j), \qquad \Pi_{\mathfrak{F}} \;=\; \sum_{i<j} \Pi_{ij}. \tag{43}$$

Proposition 9.23 (Peirce projectors). For $i \neq j$ the operator Π_{ij} is the orthogonal projection of J onto J_{ij} for the trace inner product, and $\Pi_{\mathfrak{F}}$ is the orthogonal projection onto the frame tangent $T_c\mathfrak{F}_J = \bigoplus_{i<j} J_{ij}$ of Theorem 7.3. In particular $\operatorname{Tr}\Pi_{ij} = d_{ij}$, $\operatorname{Tr}\Pi_{\mathfrak{F}} = \dim J - r$, and $\Pi_{ij}\Pi_{k\ell} = 0$ for distinct index pairs.

Proof. The operators $L_{c_1}, \dots, L_{c_r}$ of a Jordan frame commute and are simultaneously diagonalized by the Peirce decomposition $J = \bigoplus_{k\le\ell} J_{k\ell}$, with L_{c_i} acting on $J_{k\ell}$ by $\tfrac{1}{2}(\delta_{ik} + \delta_{i\ell})$ [17, Thm. IV.2.1]. Hence $4L_{c_i}L_{c_j}$ acts on $J_{k\ell}$ by $(\delta_{ik} + \delta_{i\ell})(\delta_{jk} + \delta_{j\ell})$, which is 1 when $\{k,\ell\} = \{i,j\}$ and 0 otherwise. So Π_{ij} is the identity on J_{ij} and zero on every other Peirce block; it is self-adjoint for the trace form because each L_a is and the two factors commute. Distinct index pairs therefore have orthogonal ranges, so their projectors multiply to zero. Summing over $i < j$ gives $\Pi_{\mathfrak{F}}$, and the traces are $\dim J_{ij} = d_{ij}$ and $\sum_{i<j} d_{ij} = \dim J - r$ by (5). □

Theorem 9.24 — Positive kernels from transported Peirce projections

For $q \in \Omega_J$ put $\Pi_{ij}(q) = \mathcal{U}(q,K)\,\Pi_{ij}\,\mathcal{U}(q,K)^{-1}$ and $\Pi_{\mathfrak{F}}(q) = \mathcal{U}(q,K)\,\Pi_{\mathfrak{F}}\,\mathcal{U}(q,K)^{-1}$, and define

$$\mathscr{K}_{ij}(q,q') \;=\; \operatorname{Tr}\big(\Pi_{ij}(q)\,\Pi_{ij}(q')\big), \qquad \mathscr{K}_{\mathfrak{F}}(q,q') \;=\; \operatorname{Tr}\big(\Pi_{\mathfrak{F}}(q)\,\Pi_{\mathfrak{F}}(q')\big). \tag{44}$$

Then, writing Π for either Π_{ij} or $\Pi_{\mathfrak{F}}$, $\mathscr{K}$ for the corresponding kernel and $\delta = \operatorname{Tr}\Pi$:

(a) $\Pi(q)$ is again an orthogonal projection, onto $\mathcal{U}(q,K)\Pi(J)$, and $\mathscr{K}$ is a positive semidefinite kernel on Ω_J: for all $q_1, \dots, q_N \in \Omega_J$ and $a \in \mathbb{R}^N$,

$$\sum_{k,\ell} a_k a_\ell\, \mathscr{K}(q_k,q_\ell) \;=\; \Big\| \sum_k a_k \Pi(q_k) \Big\|_{\mathrm{HS}}^2 \;\ge\; 0;$$

(b) $\mathscr{K}(q,q) = \delta$ and $0 \le \mathscr{K}(q,q') \le \delta$, with $\mathscr{K}(q,q') = \delta$ if and only if $\Pi(q) = \Pi(q')$;

(c) for $h \in T_c\mathfrak{F}_J$ and $q(t) = \exp_J(th)$,

$$\mathscr{K}\big(e, q(t)\big) \;=\; \delta \;-\; \frac{t^2}{2}\,\big\|[\mathscr{A}_K(h), \Pi]\big\|_{\mathrm{HS}}^2 \;+\; O(t^3),$$

so the kernel is stationary at coincidence with negative semidefinite second variation, degenerate exactly on the directions h with $[\mathscr{A}_K(h), \Pi] = 0$;

Theorem 9.24 — Positive kernels from transported Peirce projections (continued)

(d) for the total projector the degeneracy is exact: $[\mathscr{A}_K(h), \Pi_{\mathfrak{F}}] = 0$ if and only if $\mathscr{A}_K(h) = 0$, so

$$\ker\Big(-D^2\mathscr{K}_{\mathfrak{F}}(e,e)\big|_{T_c\mathfrak{F}_J}\Big) \;=\; \ker \mathscr{A}_K \;=\; \bigoplus_{\kappa_i=\kappa_j} J_{ij},$$

and the second variation of $\mathscr{K}_{\mathfrak{F}}$ is negative definite on the entire frame tangent;

(e) the construction is equivariant: transporting the aggregate key, the frame and both queries by a common $g \in \mathrm{Aut}(J)$ leaves $\mathscr{K}$ unchanged.

For $J = H_3(\mathbb{O})$ the construction gives an F_4-equivariant family of three rank-eight block kernels, permuted with the triality channels of Corollary 9.22, together with the label-free total frame kernel of diagonal value 24. All scalar values are invariant under simultaneous transport of the aggregate key, the frame and the query pair.

Proof. **Step 1** (*projections and positivity*)**.** $\mathscr{U}(q,K) \in \mathrm{Aut}(J)$ is orthogonal for the trace form, so conjugation carries orthogonal projections to orthogonal projections and preserves ranks; $\Pi(q)$ projects onto the image of $\Pi(J)$. Writing $\langle A, B\rangle_{\mathrm{HS}} = \mathrm{Tr}(AB^{\top})$ and using $\Pi(q)^{\top} = \Pi(q)$, the kernel is the Hilbert–Schmidt inner product $\langle \Pi(q), \Pi(q')\rangle_{\mathrm{HS}}$, so the quadratic form in (a) is the squared norm of $\sum_k a_k \Pi(q_k)$.

Step 2 (*range and equality*)**.** $\mathscr{K}(q,q) = \mathrm{Tr}\,\Pi(q)^2 = \mathrm{Tr}\,\Pi(q) = \delta$. For two orthogonal projections P, Q one has $\mathrm{Tr}(PQ) = \mathrm{Tr}(QPPQ) = \|PQ\|_{\mathrm{HS}}^2 \geq 0$ and

$$\mathrm{Tr}(PQ) \;=\; \mathrm{Tr}(P) - \mathrm{Tr}\big(P(I-Q)\big) \;=\; \delta - \big\|(I-Q)P\big\|_{\mathrm{HS}}^2 \;\leq\; \delta,$$

with equality exactly when $(I-Q)P = 0$, that is when the range of P lies in that of Q; as the ranks agree this forces $P = Q$.

Step 3 (*second variation*)**.** Apply Lemma 9.16 to the conjugation representation of $\mathrm{Aut}(J)$ on $H = \mathrm{End}(J)$ with the Hilbert–Schmidt inner product, which is orthogonal because each $g \in \mathrm{Aut}(J)$ is; take $g(t) = \mathscr{U}(q(t), K)$, whose derivative at $t = 0$ is $A := \mathscr{A}_K(h)$, and probe $v = \Pi$. Then $d\rho(A)\Pi = [A, \Pi]$ and $f_{\Pi}(t) = \mathscr{K}(e, q(t))$, so the first variation vanishes and the second is $-\|[A,\Pi]\|_{\mathrm{HS}}^2$. The degeneracy condition is therefore $[A,\Pi] = 0$.

Step 4 (*exact degeneracy for the total projector*)**.** If $A = \mathscr{A}_K(h) = 0$ the commutator vanishes. Conversely suppose $[A, \Pi_{\mathfrak{F}}] = 0$. Each frame idempotent c_k lies in J_{kk}, hence in $\ker \Pi_{\mathfrak{F}}$, so $\Pi_{\mathfrak{F}} A c_k = A \Pi_{\mathfrak{F}} c_k = 0$. But (36) places $Ac_k = \sum_{j>k} z_{kj} - \sum_{i<k} z_{ik}$ in the frame tangent, which is the range of $\Pi_{\mathfrak{F}}$; therefore $Ac_k = \Pi_{\mathfrak{F}} Ac_k = 0$. Projecting (36) onto the mutually orthogonal blocks $J_{k\ell}$ forces every $z_{k\ell} = 0$, and then $A = 0$ exactly as in the proof of Theorem 9.11. Combining with (c) and Theorem 9.11 gives the stated kernel, and simple spectrum of K makes $\ker \mathscr{A}_K = 0$.

Step 5 (*equivariance*)**.** $\mathscr{U}(gq, gK) = g\,\mathscr{U}(q,K)\,g^{-1}$ and $\Pi_{ij}^{gc} = g\,\Pi_{ij}^{c}\,g^{-1}$, since $L_{gc_i} = gL_{c_i}g^{-1}$. Hence $\Pi(gq) = g\,\Pi(q)\,g^{-1}$ and the operator trace is conjugation-invariant. If g permutes the frame, $c_i \mapsto c_{\sigma(i)}$, then $g\Pi_{ij}g^{-1} = \Pi_{\sigma(i)\sigma(j)}$, so the block labels are permuted while $\Pi_{\mathfrak{F}}$ is fixed.

For $H_3(\mathbb{O})$ the three block dimensions are $d_{ij} = 8$ and their sum is 24, by Corollary 9.22, and $\mathrm{Aut}(J) = F_4$ by Corollary 7.2. □

The three block kernels and the total kernel are entries and aggregates of one operator-valued object, which makes the phrase "coupled triality channels" literal.

Proposition 9.25 (The cross-channel Gram kernel). Index the frame blocks by $\alpha = (i, j)$ with $i < j$, write $N = \binom{r}{2}$ for their number, and define the matrix-valued kernel

$$\mathbf{K}(q, q')_{\alpha\beta} \;=\; \mathrm{Tr}\big(\Pi_\alpha(q)\,\Pi_\beta(q')\big), \qquad \mathbf{K}(q, q') \in \mathbb{R}^{N\times N}. \tag{45}$$

Let $\Phi(q) : \mathbb{R}^N \to \mathrm{End}(J)$ be the linear map with $\Phi(q)e_\alpha = \Pi_\alpha(q)$, the target carrying the Hilbert–Schmidt inner product. Then:

(a) $\mathbf{K}(q, q') = \Phi(q)^*\Phi(q')$, so $\mathbf{K}$ is a positive semidefinite operator-valued kernel: for all $q_1, \dots, q_M \in \Omega_J$ and $v_1, \dots, v_M \in \mathbb{R}^N$,

$$\sum_{k,\ell} \langle v_k, \mathbf{K}(q_k, q_\ell) v_\ell \rangle \;=\; \Big\| \sum_k \Phi(q_k) v_k \Big\|_{\mathrm{HS}}^2 \;\geq\; 0;$$

(b) $\mathbf{K}(q, q) = \mathrm{diag}(d_\alpha)$, which for $J = H_3(\mathbb{O})$ is $8I_3$;

(c) the scalar kernels of Theorem 9.24 are recovered as $\mathscr{K}_\alpha = \mathbf{K}_{\alpha\alpha}$ and $\mathscr{K}_{\mathfrak{F}} = \sum_{\alpha,\beta} \mathbf{K}_{\alpha\beta}$, the sum of all entries rather than the trace; the trace $\sum_\alpha \mathbf{K}_{\alpha\alpha}$ is the sum of the block kernels, and the off-diagonal entries are exactly the cross-channel content that the scalar kernels discard;

(d) an automorphism that normalizes the chosen frame acts on the block labels through the induced permutation σ, and therefore conjugates $\mathbf{K}$ by the corresponding permutation matrix. In the Albert case the setwise stabilizer of the frame has quotient S_3 over its pointwise Spin(8) stabilizer, and this quotient permutes the three triality channels. Under a general F_4-transformation the frame, its channel spaces and the kernel are transported together; the total frame kernel is label-free and invariant under simultaneous transport.

Proof. (a) is the definition of Φ together with $\langle \Pi_\alpha(q), \Pi_\beta(q') \rangle_{\mathrm{HS}} = \mathrm{Tr}(\Pi_\alpha(q)\Pi_\beta(q'))$, using self-adjointness of the projectors; the displayed identity is the squared norm of $\sum_k \Phi(q_k)v_k$. For (b), conjugation by $\mathscr{U}(q, K)$ preserves the orthogonality of Proposition 9.23, so $\Pi_\alpha(q)\Pi_\beta(q) = 0$ for $\alpha \neq \beta$ and $\mathrm{Tr}\,\Pi_\alpha(q)^2 = d_\alpha$. For (c), $\Pi_{\mathfrak{F}} = \sum_\alpha \Pi_\alpha$ gives $\mathrm{Tr}(\Pi_{\mathfrak{F}}(q)\Pi_{\mathfrak{F}}(q')) = \sum_{\alpha,\beta} \mathbf{K}_{\alpha\beta}$. Part (d) is frame-normalizer equivariance: the permutation computation is Step 5 of Theorem 9.24, and full F_4-invariance is a statement about simultaneous transport of the frame together with the queries, or about the label-free total kernel, not about a fixed 3×3 matrix of labeled channels. □

Remark 9.6 (What the Gram argument does and does not supply). Positivity here holds block by block and for a single reason: the kernel is an inner product of two transported projections in Hilbert–Schmidt space. A finer hierarchy of kernels indexed by higher $\mathrm{Aut}(J)$-types would have to be built from actual equivariant intertwiners between those types; formal analogues of the block channels carry no positivity guarantee of their own.

Remark 9.7 (Repeated aggregate-key spectrum). At repeated spectrum the primitive frame of K is no longer intrinsic, and there are two consistent options. One may retain a key-encoded primitive refinement, as in (31); the formulas above then remain valid verbatim but depend on that choice. Alternatively one may pass to the intrinsic coarse spectral idempotents $E_1, \dots, E_s$ of K and use the coarse Peirce projections $\Pi_{ab} = 4L_{E_a}L_{E_b}$ for $a \neq b$, which project onto the

coarse blocks J_{ab}. The Gram-kernel, positivity, curvature and equivariance arguments of Theorem 9.24 and Proposition 9.25 admit direct coarse-block analogues after replacing the primitive blocks J_{ij}, the projectors Π_{ij} and the primitive frame tangent by the coarse blocks J_{ab}, the projectors Π_{ab} and the tangent of the coarse spectral flag. The resulting ranks and channel counts are the coarse-block ranks, and therefore change at spectral collisions: for $H_3(\mathbb{O})$ a simple 1+1+1 aggregate-key spectrum gives three channels of rank eight, a 2+1 collision gives a single intrinsic coarse channel of rank sixteen, and a scalar spectrum gives no nontrivial coarse off-diagonal channel at all. The two options differ in what they retain: the coarse projections are blind to frame rotations *inside* a repeated eigenspace, which is exactly the part of the frame tangent that $\mathscr{A}_K$ annihilates.

Table 4 now records the full hierarchy, from the readouts that see nothing of the frame to the projection kernels whose second-order degeneracy is exactly that of the transporter.

REPRESENTATION / READOUT	FRAME RESPONSE
individual spectra	constant on isospectral frame orbits
aligned overlap scalar	first derivative zero; second variation $-\sum_{i<j}(\mu_i-\mu_j)^2\|z_{ij}\|^2$, negative definite at simple probe spectrum
spectral-code vector	rank $\sum_{a_i\neq a_j,\ \kappa_i\neq\kappa_j} d_{ij}$
effect frame	rank $\sum_{\kappa_i\neq\kappa_j} d_{ij}$
faithful operator on J_0	the transporter image, at its intrinsic rank
selected matrix coefficients	rank set by their annihilator on that image
Peirce-block kernel	first variation zero; degenerate where $[\mathscr{A}_K(h),\Pi_{ij}]=0$
total Peirce-projection kernel	first variation zero; second-variation kernel $\ker\mathscr{A}_K$, hence rank $\sum_{\kappa_i\neq\kappa_j} d_{ij}$

Table 4. *Frame response by readout, for an aggregate key $K=\sum_i \kappa_i c_i$ and, in rows three and four, the diagonal keys* (31)*. The rows are Proposition 7.8, Proposition 9.21, Corollary 9.14, Proposition 9.13(a), Proposition 9.13(b),* (39)*, Theorem 9.24(c) and Theorem 9.24(d). For $H_3(\mathbb{O})$ with simple aggregate-key spectrum, simple probe spectrum and pairwise distinct code coefficients the first four read* 0*;* 0 *at first order but negative definite at second order;* 24*; and* 24*, and the last row is negative definite on the whole* 24*-dimensional frame tangent.*

9.10 COMPOSITION ACROSS TOKENS, AND AN OPEN NONCLASSICALITY TEST

Two limits of the construction are recorded here because they are structural rather than incidental. Neither is used elsewhere in the paper.

Remark 9.8 (Across tokens the construction is a classical mixture). Every construction above that involves more than one carrier element is a direct sum, never a composite of state spaces. The score row of Section 9.1 lives in $J^n = J\oplus\cdots\oplus J$; the signed-value lift of Section A lives in the doubled carrier $J\oplus J$; non-simple carriers split as direct sums by Proposition 3.3; and the tensor products that do appear, $J\otimes\mathbb{R}^{d_v}$ in Section 9.7 and $J\otimes W$ in Section B, pair the carrier with a value space that carries no cone of its own, the operator acting on the J factor alone, so that $J\otimes\mathbb{R}^{d_v}$ is d_v parallel copies of J. A direct sum of Euclidean Jordan algebras is a classical

mixture of its summands: its state space is the convex hull of the summands' state spaces, with no joint states and in particular nothing that plays the role of entanglement. Equivalently, J^n is the composite of J with the simplicial carrier $\mathbb{R}^n$, and the token index is a classical register.

For the exceptional carrier this is forced and not merely a choice. Barnum, Graydon and Wilce [7] show that no composite of Euclidean Jordan algebras, in their sense, has the exceptional algebra $H_3(\mathbb{O})$ as a direct summand, and that a composite of two Euclidean Jordan algebras one of which has an exceptional summand exists only when the other is a direct sum of one-dimensional algebras, that is, simplicial. So $H_3(\mathbb{O})$ has no composite with itself, nor with $\mathrm{Herm}_2(\mathbb{C})$, and the only composites available to an Albert-valued sequence model are those with classical carriers, of which J^n is the one used here. For the special carriers $\mathrm{Herm}_r(\mathbb{C})$ composites do exist, and a tensor-product layer across tokens is possible there in principle; it is not studied in this paper. The precise definition of a composite, and the hypotheses under which the exclusion holds, are those of [7].

Remark 9.9 (Simplex embeddability of the accessible fragment: an open test). The paper contains no test of nonclassicality, and the material above poses one. Fix keys $k_1, \dots, k_m$ and a set $Q \subseteq \Omega_J$ of queries. Each query contributes the effects $E_j(q) = \mathcal{U}(q, K)\widehat{E}_j$ of Theorem 9.2, and a set $R \subseteq \mathscr{B}_J$ of probe states pairs against them as in Proposition 9.19. The pair $(R, \{E_j(q)\}_{q \in Q,\, j \le m})$ is a prepare-measure fragment in the sense of Selby et al. [39]. As q ranges over Ω_J the transporter moves the effect family through a continuous $\mathrm{Aut}(J)$-orbit, so the layer accesses a continuous family of measurements on a non-simplicial state space.

A fragment admits a generalized-noncontextual ontological model exactly when it is simplex-embeddable [38], and for finitely many states and effects simplex embeddability is decided by a linear program [40]. For the full theory of a carrier, with every state and every effect, the answer is already known: a theory satisfying the no-restriction hypothesis is noncontextual exactly when it is simplicial [41], so the full theory of every non-simplicial J is contextual. The open question is about the fragment a layer actually reaches. The identity query with the aligned keys (31) reaches only the commuting family $E_j(e) = c_j$, whose statistics $\langle \rho, c_j \rangle$ factor through the diagonal of ρ in one frame; that fragment is simplex-embeddable. Sweeping the query rotates the family by $\mathcal{U}(q, K)$, and Proposition 9.19(d) and Proposition 9.21 show that the statistics then depend on the relative frame continuously and nontrivially. A layer whose fragment fails the test — and failure of any finite subfragment suffices, since a restriction of a simplex embedding is a simplex embedding — would realize routing statistics that no noncontextual hidden-variable model reproduces. We expect this to happen for a non-simplicial carrier with distinct aggregate-key spectra, but it is not proved here.

10 CONCLUSION

Ordinary softmax is not replaced by the Gibbs-chart construction. It is the spectral law on every Euclidean Jordan carrier. What changes with the carrier is the geometry to which those weights are attached. The simplicial carrier has no continuous frame directions; every non-simplicial carrier does, and the same Peirce blocks govern the Hessian metric, the Jacobian, and the source in the Gibbs Monge–Ampère equation (Theorem 7.10 and Corollaries 7.12 and 7.13).

Canonical effect normalization exposes a parallel decomposition of cone-valued data. The aggregate score supplies a positive radial coordinate, while the normalized effects retain a compact $\mathrm{Aut}(J)$-valued transporter. The two channels are graded by scale: the radial factor is homogeneous and the transporter is a function on the projectivized cone, which is the same quotient that produces the state space (Proposition 9.7). They are independent registers only to first order, coupling at second order through the curvature bracket of the cone (Proposition 9.8); and the transporter is trivial precisely on the simplicial carrier, where Theorem 7.4 leaves it nothing to act on (Corollary 9.4). At the identity query, simple spectrum of the aggregate key makes its linearization injective on the full frame tangent, where it acts as a contrast filter on Peirce edges (Theorem 9.11). Representations of this transporter produce effect frames, positive instruments, additive-moment reductions, relative-frame readouts, and positive orbit kernels (Proposition 9.17 and Theorem 9.24).

The conclusion is structural rather than prescriptive. Choosing a carrier fixes, before any learned parameter is introduced, which notions of positivity, normalization, symmetry, frame and transport are available. The carrier $\mathbb{R}^r$ exposes only weights. The Albert carrier exposes the same three-way softmax together with three triality-related eight-dimensional frame channels. Softmax is therefore the frameless member of a carrier-dependent geometric mechanism.

Two boundaries of the construction are recorded in Section 9.10. Across tokens it is a classical mixture of carriers, and for the Albert carrier no other composite exists (Remark 9.8). Whether the measurement fragment a trained layer reaches is simplex-embeddable remains open (Remark 9.9). The carrier class itself need not be assumed: a homogeneous cone with pure transitivity is already a direct sum of copies of one simple carrier (Remark 2.2).

A POSITIVE ENCODING OF SIGNED VALUES

Proposition 9.9 returns a cone element when the values lie in $\overline{\Omega}_J$. Signed values admit a canonical positive encoding, obtained from the Jordan positive and negative parts rather than from an arbitrary shift. For $V \in J$ put

$$|V| \;=\; \big(V^{\circ 2}\big)^{1/2}, \qquad V^+ \;=\; \frac{|V|+V}{2}, \qquad V^- \;=\; \frac{|V|-V}{2}.$$

Proposition A.1 (Canonical positive–negative lift). For every $V \in J$,

$$V^{\pm} \in \overline{\Omega}_J, \qquad V \;=\; V^+ - V^-, \qquad V^+ \circ V^- \;=\; 0,$$

and the lift $\iota(V) = (V^+, V^-) \in \overline{\Omega}_{J\oplus J}$ is $\mathrm{Aut}(J)$-equivariant, $\iota(gV) = (gV^+, gV^-)$. Given effects $E_1, \dots, E_m \in \overline{\Omega}_J$ with $\sum_j E_j = e$, the doubled family $\widetilde{E}_j = (E_j, E_j)$ satisfies $\sum_j \widetilde{E}_j = (e,e)$, the unit of $J \oplus J$, so it is an effect family there, and

$$\mathfrak{T}^*_{\widetilde{E}}\big(\iota(V_1), \dots, \iota(V_m)\big) \;=\; \Big(\sum_j P\big(E_j^{1/2}\big)V_j^+,\; \sum_j P\big(E_j^{1/2}\big)V_j^-\Big) \;\in\; \overline{\Omega}_{J\oplus J},$$

whose two components differ by $\sum_j P(E_j^{1/2})V_j$.

Proof. Write $V = \sum_k \lambda_k c_k$. Spectral calculus gives $V^{\circ 2} = \sum_k \lambda_k^2 c_k$ and $|V| = \sum_k |\lambda_k| c_k$, so $V^{\pm} = \sum_k \max(\pm\lambda_k, 0) c_k$ lies in $\overline{\Omega}_J$, the difference is V, and the two have disjoint supports in a common frame, whence $V^+ \circ V^- = 0$. Automorphisms commute with spectral calculus, which gives the equivariance. The doubled effects sum to (e,e) componentwise, so Proposition 9.9 applies in $J \oplus J$ and acts summand by summand; subtracting the two components gives the last claim. □

The decoder is the difference map, which is not itself positive. The doubled instrument remains positive and trace-preserving, its adjoint aggregation remains positive and unital, and the encoding is canonical rather than shift-dependent.

B TENSOR MOMENTS FOR CLASSICAL READOUTS

Theorem 9.15 reduces the central-value branch to two additive carrier moments. The classical readout — pairing the effects against a state rather than acting on values — admits the corresponding contraction for arbitrary values in a real vector space W. For a state $\rho \in \mathscr{B}_J$ set

$$\mathsf{M} \;=\; \sum_j k_j \otimes V_j \;\in\; J \otimes W, \qquad \xi_{q,\rho} \;=\; \big[P(S^{-1/2})P(q^{1/2})\big]^*\rho \;=\; P\big(q^{1/2}\big)P\big(S^{-1/2}\big)\rho \;\in\; J,$$

the second equality because each $P(a)$ is self-adjoint for the trace form.

Proposition B.1 (Readout by contraction). With $E_j = P(S^{-1/2})P(q^{1/2})k_j$ as in Theorem 9.2,

$$\sum_j \langle \rho, E_j \rangle V_j \;=\; \big(\langle \xi_{q,\rho}, \cdot \rangle \otimes \mathrm{id}_W\big)\mathsf{M}.$$

The sequence therefore enters the classical readout only through $K = \sum_j k_j$, which fixes $\xi_{q,\rho}$, and through the single tensor M; both are additive over blocks.

Proof. Self-adjointness of $P(a)$ gives $\langle \rho, E_j \rangle = \langle \rho, P(S^{-1/2})P(q^{1/2})k_j \rangle = \langle \xi_{q,\rho}, k_j \rangle$, and summing against V_j is the displayed contraction. The normalizer $S = P(q^{1/2})K$ depends on the sequence only through K. □

C VINBERG'S AXIOMS OVER THE OCTONIONS AND THE SEDENIONS

This appendix records the numerical check behind Remark 2.3. For a Cayley–Dickson algebra $A \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}, \mathbb{S}\}$, with $\mathbb{S}$ the sedenions, let T be the $r \times r$ matrices with off-diagonal entries in A and real diagonal entries. The product is the matrix product in A, with every product that lands on the diagonal replaced by its real part. The involution is conjugate transpose and the trace is the sum of the diagonal. Table 5 reports the largest entrywise discrepancy of the two axioms (6), and of trace associativity $\mathrm{tr}(a(bc)) = \mathrm{tr}((ab)c)$, over 60 random draws per row with standard normal entries.

ALGEBRA	RANK	$t(uw) = (tu)w$	$t(uu^*) = (tu)u^*$	TRACE ASSOCIATIVITY
$\mathbb{R}$	3, 4	2×10^{-15}	4×10^{-15}	4×10^{-15}
$\mathbb{C}$	3, 4	4×10^{-15}	7×10^{-15}	7×10^{-15}
$\mathbb{H}$	3, 4	7×10^{-15}	7×10^{-15}	1×10^{-14}
$\mathbb{O}$	3	7×10^{-15}	9×10^{-15}	2×10^{-14}
$\mathbb{O}$	4	36.5, entry $(1,4)$	48	4×10^{-14}
$\mathbb{S}$	3	7×10^{-15}	41, entry $(1,2)$	3×10^{-14}
$\mathbb{S}$	4	87	104	6×10^{-14}
$\mathbb{O}$, no projection	3	5×10^{-15}	54, entry $(1,1)$	1×10^{-14}

Table 5. *Largest entrywise discrepancy of Vinberg's axioms on the Hermitian-style matrix algebra over each coefficient algebra. Values near 10^{-14} are rounding; values of order 10 are failures, located at the entry named. The last row omits the projection of diagonal products to their real parts and fails the second axiom on the diagonal by an associator with zero real part.*

For $\mathbb{O}$ at rank 3 the script also checks the following.

(i) For 200 random upper-triangular t with positive diagonal, tt^* is Hermitian and lies in the Albert cone, with trace, sum of principal 2×2 minors and determinant all nonnegative. Moreover $\det(tt^*) = (t_{11}t_{22}t_{33})^2$ to within 3.2×10^{-13}. The parameter count is $3 + 3 \cdot 8 = 27 = \dim H_3(\mathbb{O})$.

(ii) For 300 interior elements X, each a positive combination of 3 to 6 random pure states drawn from the whole projective plane, the Cholesky factorization computed from the bottom row up gives $tt^* = X$ to relative accuracy 3.3×10^{-16}. Every sampled interior element of the Albert cone therefore factors in this T-algebra.

(iii) For 200 random $(u_1, u_2) \in \mathbb{O}^2$, the element $X = vv^*$ with $v = (1, u_1, u_2)$ satisfies $X \circ X = \mathrm{tr}(X)X$, and its quadratic invariant and determinant vanish, to within 2.3×10^{-13}. So the affine chart of Remark 2.3 consists of primitive elements.

These are tests at random points, not proofs; the algebraic reasons are those given in Remark 2.3. They test the construction against the axioms as stated in [9, Defn. 31]. We have not compared it with the T-algebra that Vinberg writes down for the exceptional cone.

Ancillary code. Two Python supplements, `verify_core_identities.py` and `verify_polar_extensions.py`, numerically cross-check the identities used in Sections 6–9 and Appendices A–B. A third, `verify_talgebra.py`, runs the checks of Section C. They accompany this manuscript as ancillary files, together with the transcripts of their runs, and are also at `github.com/johnzfitch/gibbs-chart`. The scripts are not used in the proofs.

REFERENCES

[1] E. M. Alfsen and F. W. Shultz, *Geometry of State Spaces of Operator Algebras*, Mathematics: Theory & Applications, Birkhäuser, 2003.

[2] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of Mathematical Monographs **191**, American Mathematical Society, 2000.

[3] H. Bae and Y. Lim, On some Finsler structures of symmetric cones, *Forum Math.* **13** (2001), 629–639.

[4] M. Baes, Convexity and differentiability properties of spectral functions and spectral mappings on Euclidean Jordan algebras, *Linear Algebra Appl.* **422** (2007), 664–700.

[5] J. C. Baez, The octonions, *Bull. Amer. Math. Soc.* **39** (2002), 145–205.

[6] A. Barakat, W. Lin, J. Lazarsfeld, and A. Varvitsiotis, Optimistic online learning in symmetric cone games, arXiv:2504.03592 (2025).

[7] H. Barnum, M. A. Graydon, and A. Wilce, Composites and categories of Euclidean Jordan algebras, *Quantum* **4** (2020), 359.

[8] H. Barnum and J. Hilgert, Spectral properties of convex bodies, *J. Lie Theory* **30** (2020), 315–344.

[9] H. Barnum, C. Ududec, and J. van de Wetering, Self-duality and Jordan structure of quantum theory follow from homogeneity and pure transitivity, arXiv:2306.00362 [quant-ph] (2023; v2, 2026).

[10] R. J. Berman and B. Berndtsson, Real Monge–Ampère equations and Kähler–Ricci solitons on toric log Fano varieties, *Ann. Fac. Sci. Toulouse Math.* (6) **22** (2013), no. 4, 649–711.

[11] R. J. Berman and B. Berndtsson, The volume of Kähler–Einstein Fano varieties and convex bodies, *J. Reine Angew. Math.* **723** (2017), 127–152.

[12] G. Birkhoff, Tres observaciones sobre el álgebra lineal, *Univ. Nac. Tucumán Rev. Ser. A* **5** (1946), 147–151.

[13] J. S. Bridle, Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition, in F. Fogelman-Soulié and J. Hérault (eds.), *Neurocomputing: Algorithms, Architectures and Applications*, NATO ASI Series F **68**, Springer, 1990, 227–236.

[14] I. Canyakmaz, W. Lin, G. Piliouras, and A. Varvitsiotis, Multiplicative updates for online convex optimization over symmetric cones, arXiv:2307.03136 (2023).

[15] C. Davis, All convex invariant functions of hermitian matrices, *Arch. Math.* **8** (1957), 276–278.

[16] E. Ehrhart, Une généralisation probable du théorème fondamental de Minkowski, *C. R. Acad. Sci. Paris* **258** (1964), 4885–4887.

[17] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Clarendon Press, 1994.

[18] A. Ghriss and S. Chakraborty, Kernelized linear attention: breaking the capacity wall with symmetric cones, arXiv:2607.17419v1 [cs.LG] (19 July 2026).

[19] S. G. Gindikin, Analysis in homogeneous domains, *Uspekhi Mat. Nauk* **19** (1964), no. 4, 3–92; English transl., *Russian Math. Surveys* **19** (1964), no. 4, 1–89.

[20] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd ed., Cambridge University Press, 1952.

[21] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Graduate Studies in Mathematics **34**, American Mathematical Society, 2001; corrected reprint of the 1978 Academic Press edition.

[22] N. Jacobson, *Structure and Representations of Jordan Algebras*, AMS Colloquium Publications **39**, 1968.

[23] E. T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* **106** (1957), 620–630.

[24] P. Jordan, J. von Neumann, and E. Wigner, On an algebraic generalization of the quantum mechanical formalism, *Ann. of Math. (2)* **35** (1934), 29–64.

[25] M. Koecher, Die Geodätischen von Positivitätsbereichen, *Math. Ann.* **135** (1958), 192–202.

[26] M. Koecher, *The Minnesota Notes on Jordan Algebras and Their Applications* (A. Krieg and S. Walcher, eds.), Lecture Notes in Mathematics **1710**, Springer, 1999.

[27] A. S. Lewis, Convex analysis on the Hermitian matrices, *SIAM J. Optim.* **6** (1996), 164–177.

[28] G. Lüders, Über die Zustandsänderung durch den Meßprozeß, *Ann. Physik (6)* **8** (1951), 322–328; English transl. by K. A. Kirkpatrick, Concerning the state-change due to the measurement process, *Ann. Phys. (Leipzig)* **15** (2006), 663–670.

[29] A. F. T. Martins and R. F. Astudillo, From softmax to sparsemax: a sparse model of attention and multi-label classification, *Proc. 33rd Int. Conf. Machine Learning (ICML)*, PMLR **48** (2016), 1614–1623.

[30] K. McCrimmon, *A Taste of Jordan Algebras*, Universitext, Springer, 2004.

[31] B. Nill and A. Paffenholz, On the equality case in Ehrhart's volume conjecture, *Adv. Geom.* **14** (2014), 579–586.

[32] OpenAI, The sharp inequality in Ehrhart's volume conjecture, Chapter 8 of *Ten advances in mathematics and theoretical computer science*, technical report, 1 August 2026; Lean formalizations at `https://github.com/openai/ten-proofs`.

[33] B. Peters, V. Niculae, and A. F. T. Martins, Sparse sequence-to-sequence models, *Proc. 57th Ann. Meeting Assoc. Comput. Linguistics (ACL)* (2019), 1504–1519.

[34] D. Petz, Geometry of canonical correlation on the state space of a quantum system, *J. Math. Phys.* **35** (1994), 780–795.

[35] D. Petz, Monotone metrics on matrix spaces, *Linear Algebra Appl.* **244** (1996), 81–96.

[36] D. Petz, Bregman divergence as relative operator entropy, *Acta Math. Hungar.* **116** (2007), 127–131.

[37] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[38] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal, and R. W. Spekkens, Characterization of noncontextuality in the framework of generalized probabilistic theories, *PRX Quantum* **2** (2021), 010331.

[39] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Accessible fragments of generalized probabilistic theories, cone equivalence, and applications to witnessing nonclassicality, *Phys. Rev. A* **107** (2023), 062203.

[40] J. H. Selby, E. Wolfe, D. Schmid, A. B. Sainz, and V. P. Rossi, A linear program for testing nonclassicality and an open-source implementation, *Phys. Rev. Lett.* **132** (2024), 050202.

[41] F. Shahandeh, Contextuality of general probabilistic theories, *PRX Quantum* **2** (2021), 010330.

[42] T. A. Springer and F. D. Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, Springer Monographs in Mathematics, 2000.

[43] V. A. Truong and L. Tunçel, Geometry of homogeneous convex cones, duality mapping, and optimal self-concordant barriers, *Math. Program. Ser. A* **100** (2004), 295–316.

[44] C. Tsallis, Possible generalization of Boltzmann–Gibbs statistics, *J. Stat. Phys.* **52** (1988), 479–487.

[45] H. Umegaki, Conditional expectation in an operator algebra IV: entropy and information, *Kōdai Math. Sem. Rep.* **14** (1962), 59–85.

[46] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, Attention is all you need, *Advances in Neural Information Processing Systems* **30** (2017), 5998–6008.

[47] E. B. Vinberg, Homogeneous cones, *Dokl. Akad. Nauk SSSR* **133** (1960), 9–12; English transl. *Soviet Math. Dokl.* **1** (1960), 787–790.

[48] E. B. Vinberg, The theory of convex homogeneous cones, *Trudy Moskov. Mat. Obshch.* **12** (1963), 303–358; English transl. *Trans. Moscow Math. Soc.* **12** (1963), 340–403.

[49] E. B. Vinberg, The structure of the group of automorphisms of a homogeneous convex cone, *Trudy Moskov. Mat. Obshch.* **13** (1965), 56–83; English transl. *Trans. Moscow Math. Soc.* **13** (1965), 63–93.

[50] I. Yokota, Exceptional Lie groups, arXiv:0902.0431 [math.DG] (2009).