When Does Adaptive Spectral Normalization Define a Fixed Convex Geometry?

Anonymous authors
Paper under double-blind review

Abstract

Finite spectral updates can approximate the polar direction without implementing one fixed convex geometry. We classify this distinction for $q_i(s) = h(s_i/\alpha(s))$, where the scale depends on the spectrum. With one initial Frobenius normalization in dimension at least three, one cubic Newton–Schulz step admits a norm-derived direction, whereas repeated cubic steps and every positive-depth Taylor-quintic composition do not, even after positive scalar rescaling and despite preserving order. More generally, an odd real-analytic filter in dimension at least four, or an odd polynomial in dimension at least three, admits a positive C^1 integrating factor exactly when $h(t) = at + ct^{2k+1}$. We also classify adaptive normalizers whose displayed response belongs to one fixed convex subdifferential. Under stated regularity, boundary, and critical-point assumptions in dimension at least two, the principal modular family is exhaustive and yields unitarily invariant norms. Exact incompatibility has an important limit: a classical perspective construction gives one norm per depth whose gradient approximates the Taylor fields uniformly as depth increases at fixed dimension, including near rank loss. Additional shape assumptions yield a sharp filter-preserving conditioning frontier. Exact Jordan certificates and dependent replay of 4,728 unique stored spectra characterize geometric separation, not a training-performance penalty.

1 Introduction

Let $G = U\,\mathrm{diag}(\sigma)V^\top$. The exact polar direction $UV^\top$ has an unusually clean variational interpretation:

$$UV^\top \in \arg\max_{\|D\|_{\mathrm{op}}\le 1} \langle G, D\rangle_F \subseteq \partial\|G\|_*. \tag{1}$$

This fixed spectral-norm geometry supports steepest-descent, convergence, and implicit-bias analyses of idealized Muon (Fan et al., 2025; Gronich & Vardi, 2026). Practical spectral optimizers, however, rarely compute the polar factor exactly. They normalize a gradient or momentum matrix, apply a small number of matrix polynomials, and may deliberately reshape rather than flatten its spectrum. The resulting map can still point downhill without being the duality map of any norm.

Polar approximation does not answer the structural question:

> *For which scalar filters and adaptive normalizers does the displayed update implement one fixed convex optimization geometry?*

At fixed scale, integrating a coordinatewise monotone transform is separable. An endogenous $\alpha(s)$ couples all coordinates: Jacobian symmetry ties h' to $\nabla\alpha$, making the scale part of the geometry.

The classification tests when fixed-norm arguments can be transferred to a normalized polynomial update. Two tests differ: does the displayed field belong to one convex subdifferential, and can positive scaling make it norm-derived? Such multipliers enter blockwise learning rates and disappear under exact line search (Boyd & Vandenberghe, 2004, Section 9.4). In $d \ge 3$, the simplest separation is one versus two cubic Newton–Schulz steps: both preserve spectral order, but only the first has a norm-derived direction after one initial Frobenius normalization (table 1). More accurate polar approximation therefore need not preserve exact fixed-norm representability.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Table 1: **Two different tests for fixed geometry.** Exact-real responses in $d \geq 3$. Polynomial rows use one initial Frobenius normalization, except the matched Jordan row, which uses α_C. A "yes" in the last column allows a positive scalar multiplier. These are representation tests, not training-performance rankings.

response	convex field	norm-derived ray
exact polar	yes	yes
one Taylor cubic stage	no	yes
two or more Taylor cubic stages	no	no
one or more Taylor quintic stages	no	no
literal five-stage Jordan	no	no
matched principal Jordan	yes	yes

Polynomial entries follow from corollaries 3 and 10; Jordan entries from corollaries 7 and 13.

We then apply the classification to the literal exact-real Jordan response $H = f^{\circ 5}$, where:

$$f(x) = 3.4445x - 4.7750x^3 + 2.0315x^5. \tag{2}$$

Jordan's known order reversal already excludes convex geometry after any positive scalar rescaling; we give the short argument in corollary 13. The integrating-factor classification goes further: even increasing filters can fail to admit any potential. We also derive a sharp conditioning frontier within the classified unscaled family and an exact spectral-gauge distance. Forward modular differentiation, spectral lifting, and order preservation are classical. The first-fold exclusion itself follows from coordinate-line monotonicity and boundary continuity; the substantive classification claims are the resulting exhaustive family and the exact integrating-factor boundary.

2 The Mathematical Object and the Meaning of Geometry

Let $d = \min\{m, n\}$, and let $G = U \operatorname{diag}(\sigma_1, \ldots, \sigma_d) V^\top$ be a thin SVD. Write $H = f^{\circ 5}$ for (2). The main case-study object is the exact-real, epsilon-free spectral map:

$$Q_0(G) = U \operatorname{diag}\left(H\left(\frac{\sigma_i}{\|G\|_F} \right) \right) V^\top, \qquad Q_0(0) = 0. \tag{3}$$

Equation (3) describes five exact-real matrix recurrences, not the full optimizer (Jordan et al., 2024). The pinned implementation also uses momentum, an additive stabilizer, BF16 arithmetic, transposition, and shape scaling. Theorems 16 and 18 handle exact-real stabilizer variants; no theorem below automatically transfers to rounded kernels.

We distinguish two fixed convex geometries. A convex-potential interpretation requires a finite convex Ψ with $Q(G) \in \partial\Psi(G)$. A fixed-body support-oracle interpretation—equivalently an LMO queried at $-G$ under the usual descent convention—requires one compact convex set K such that:

$$Q(G) \in \arg\max_{D \in K} \langle G, D \rangle_F. \tag{4}$$

The latter is a special case of the former because $Q(G)$ is then a subgradient of the support function of K. Both imply monotonicity:

$$\langle Q(A) - Q(B), A - B \rangle_F \geq 0. \tag{5}$$

For zero-homogeneous fields on a cone, monotonicity already yields a support-function representation; bounded range makes the representing body compact (Kushnir & Krishnamoorthy, 2024). This is not our contribution (appendix A.1). It does not determine which filter–normalizer pairs are monotone. A local nonconvex potential or positive gradient correlation is weaker.

3 Related Work

Finite polar transforms. Kim and Oh prove convergence for a monotone Taylor–Newton–Schulz family and note that the Jordan coefficients violate their premise (Kim & Oh, 2026). Polar Express

and CANS instead design high-accuracy matrix-sign polynomials (Amsel et al., 2026; Grishina et al., 2025); Do et al. (2026) cover generalized finite Newton–Schulz updates, including the implementation polynomial, with nonconvergence examples and error bounds. We claim neither a new convergence framework nor a superior approximation polynomial. Similar tested endpoint losses for exact polar, Jordan, and cheaper filters further motivate a structural question (Huang, 2026).

Fixed geometry versus successful non-LMOs. The closest positive result is Li & Tsuchiya (2026): for their monotone finite Newton–Schulz transform, a fixed scale gives a separable convex potential, and their Theorem 5.7 gives an FTRL identity for any positive scale sequence, holding the scale fixed when differentiating each round's potential. Their regret theorem additionally needs an inter-iterate scale condition that practical Frobenius scaling can violate. We instead ask for one potential for the entire endogenous-scale field, and prove no convergence theorem. Fixed spectral constraints and non-Euclidean/LMO-based optimizer families are studied by Chen et al. (2025); Crawshaw et al. (2025); Riabinin et al. (2025).

Shumaylov et al. (2026, Theorem 2.1) show that a unitarily invariant steepest-descent response preserves singular-value order, so order reversal is not our novelty; they also show that non-LMO or randomized spectral maps can converge by alignment and boundedness. Related equivariant cyclically monotone maps obey an order constraint (Bao et al., 2026). Peyré studies fixed-gauge spectral Wasserstein metrics and separately gives a primitive for finite Newton–Schulz at fixed scale, leaving adaptive scale open (Peyré, 2026). Muse varies a fixed Frobenius-isometric representation before polar (Chang et al., 2026); Hu et al. (2026) derive ideal polar from dissipation and change temporal memory. These differ from shared-normalizer integrability. Changing or nonseparable geometries such as GO-Muon and DynMuon lie outside one static shared denominator (Che, 2026; Wu et al., 2026).

Closest integrability antecedents. Forward modular gauges motivate the compatibility equation (Musielak, 1983; Kesselheim et al., 2024). More precisely, the compatible potential itself is a scaled classical Amemiya/Orlicz norm (Hudzik & Maligranda, 2000); the supplement supplies its infimum formula and dual ball. The modular primitive Φ_h need not be convex, so the normalizer itself need not be an Orlicz norm. We claim global inverse exhaustiveness, not novelty of the forward norm construction. Fally's one-aggregator demand classification uses an integrating factor, symmetry, affine coordinate dependence, and exclusions of certain case mixtures (Fally, 2022, Proposition 1 and proof, pp. 44–49). Its budget identity, strict derivative signs, and invertibility assumptions differ from our prescribed spectral-scale problem. The direct budget-normalized identification already fails for $h(t) = t + t^3$ (remark 29); this excludes that mapping, not every possible reduction to demand theory. We credit the proof architecture and compare the explicit hypotheses, rather than claim novelty from application vocabulary or globality alone. Power/log restrictions also arise from classical separability (Blackorby et al., 1978).

Homothetic rationalization and integrating-factor theory are established (Alcantud & Rodríguez-Palmero, 2010; Hosoya, 2021), including locally Lipschitz integrability. Positive-scalar invariance of the Jacobi condition and tangent curvature is explicit in Hosoya (2022, Lemma 3). Our contribution is the explicit odd-analytic filter classification at prescribed Frobenius scale, not general potential existence. The polynomial case also follows from classical cubic containment (Elekes & Szabó, 2024) plus the Euler-derivative reduction in appendix D.

Classical ingredients. Convex-subgradient monotonicity, Rademacher differentiation, rank-one inertia, Schatten duality, and spectral lifting are prior tools (Rockafellar, 1966; Aziznejad & Unser, 2021; Lewis, 1995; Ding et al., 2018). Likewise, the origin-derivative inequality is classical (Klurman, 2012); only the Jordan specialization is used here.

4 Regular Compatibility: The Normalizer Is Not Free

Work first on the positive singular-value cone $\Omega_d = (0,\infty)^d$, $d \geq 2$. Let $\alpha : \Omega_d \to (0,\infty)$ be C^2, positive, and one-homogeneous; let $h \in C^2((0,\infty))$. Define:

$$x = \frac{s}{\alpha(s)}, \qquad q_i(s) = h(x_i), \qquad D(x) = \operatorname{diag}(h'(x_i)), \qquad \Delta(x) = x^\top D(x) x. \tag{6}$$

Lemma 1 (Positive-diagonal restriction)**.** *For a fixed matrix shape, let* $\mathsf{D}s$ *denote the rectangular diagonal embedding. If a matrix field satisfies* $Q(\mathsf{D}s) = \mathsf{D}q(s)$ *for every* $s \in \Omega_d$ *and* $Q(G) \in \partial\Psi(G)$ *for one finite convex full-matrix function* Ψ*, then* $N(s) = \Psi(\mathsf{D}s)$ *is convex on* Ω_d *and* $q(s) \in \partial N(s)$*. Thus every cone obstruction below is necessary for a full-matrix convex realization. The converse is not automatic: it additionally requires an absolutely symmetric convex extension, compatible boundary behavior, and the usual spectral-function hypotheses.*

Proof. Restrict the subgradient inequality for Ψ to rectangular diagonal matrices and use $\langle \mathsf{D}q(s), \mathsf{D}(t-s)\rangle_F = q(s)^\top (t-s)$. □

Let the unit surface $S_\alpha = \{x : \alpha(x) = 1\}$ be connected, and assume $\Delta \neq 0$ there. Choose a primitive Φ satisfying $\Phi'(t) = th'(t)$.

Theorem 2 (Regular Normalizer–Filter Compatibility)**.** *The following are equivalent:*

1. *q is conservative on Ω_d;*
2. *its Jacobian is symmetric;*
3. *the normalizer satisfies:*
$$\nabla\alpha(s) = \frac{D(x)x}{x^\top D(x)x}; \tag{7}$$
4. *one constant C makes the unit surface lie in the regular level below, with α selecting its positive radial branch:*
$$\sum_{i=1}^{d} \Phi(x_i) = C, \qquad \sum_{i=1}^{d} \Phi\left(\frac{s_i}{\alpha(s)}\right) = C. \tag{8}$$

When these conditions hold, the one-homogeneous representative of the potential and its Hessian are forced:

$$\mathcal{N}_h(s) = \sum_i s_i h(x_i), \qquad \nabla^2\mathcal{N}_h(s) = \frac{1}{\alpha(s)}\left(D - \frac{Dxx^\top D}{x^\top Dx}\right). \tag{9}$$

Proof sketch. For $i \neq j$, differentiating $q_i(s) = h(s_i/\alpha(s))$ and imposing $\partial_j q_i = \partial_i q_j$ shows that $D(x)x$ is parallel to $\nabla\alpha(s)$. Euler's identity for the one-homogeneous α fixes the proportionality constant and gives the displayed compatibility equation. Integrating $\Phi'(t) = th'(t)$ makes the modular constant on the connected unit surface. A second differentiation yields the displayed Hessian. The supplement proves both directions under the stated regularity and nondegeneracy hypotheses.

Corollary 3 (Fixed Schatten-Normalizer Classification)**.** *Let $d \geq 3$, $\alpha_p(s) = \|s\|_p$, and $h \in C^2((0,1))$. For $1 \leq p < \infty$, the field $q_i(s) = h(s_i/\|s\|_p)$ is conservative exactly for:*

$$h(t) = \begin{cases} At^{p-1} + B, & 1 < p < \infty, \\ C\log t + B, & p = 1. \end{cases} \tag{10}$$

For $p > 1$, the potential is $A\|s\|_p + B\sum_i s_i$ and is convex on the open positive cone exactly when $A \geq 0$; there the linear B term imposes no sign condition. Its natural spectral formula on positive full-rank matrices is $A\|G\|_{S_p} + B\|G\|_*$. A globally convex full-matrix lift additionally requires $B \geq 0$ (as well as $A \geq 0$). Thus positive-cone convexity is not being promoted silently to matrix-space convexity. For $p = 1$, a potential is

$$C\left[\sum_i s_i \log s_i - \left(\sum_i s_i\right)\log\left(\sum_i s_i\right)\right] + B\sum_i s_i,$$

and it is convex on the open positive cone exactly when $C \geq 0$.

The $B = 0$ power case recovers the fixed Schatten geometry underlying Muon^{p-1} (Dong & Sawin, 2026). Over unrestricted filters and normalizers, no universal price of fixed geometry is possible: t^γ, $0 < \gamma < 1$, retain fixed Schatten geometry and approach polar as $\gamma \downarrow 0$. Our quantitative result below instead fixes the literal Jordan filter $H = f^{\circ 5}$ and varies only the matched shared normalizer forced by the classification.

Stabilizer corollaries. Theorems 16 and 18 classify additive and hard-clamped Frobenius denominators in exact arithmetic, with Jordan and pinned NVIDIA witnesses (NVIDIA, 2026). Their dimensional exceptions, matrix-lift signs, and nonhomogeneous scope are stated there.

5 A Filter-Generic No-Gluing Theorem

Fix $d \geq 2$. Let $h \in C^1([0,\infty))$ have positive derivative near zero, let h' be negative somewhere. Let α be positive, one-homogeneous, and locally Lipschitz on Ω_d, with a strictly positive continuous extension to the nonzero closed positive cone. Define $q_i(s) = h(s_i/\alpha(s))$ and:

$$\zeta_h = \inf\{t > 0 : h'(t) < 0\}, \qquad \Phi_h(t) = \int_0^t u h'(u)\, du. \tag{11}$$

Lemma 4 (Coordinate-line first-fold exclusion)**.** *If q is monotone on Ω_d, every attained coordinate $s_i/\alpha(s)$ is at most ζ_h.*

Proof. Fix the other $d-1$ positive coordinates and put $z(t) = t/\alpha(s_{-i}, t)$ for $0 \leq t \leq T$. Boundary positivity gives $z(0) = 0$, and monotonicity makes $h(z(t))$ nondecreasing in $t > 0$. For $0 < u < v \leq z(T)$, the first hitting time of v has an earlier hitting time of u. Hence $h(u) \leq h(v)$. Thus h is nondecreasing on $[0, z(T)]$, which implies $z(T) \leq \zeta_h$. No monotonicity of z is assumed. □

Theorem 5 (No-Gluing and Principal-Branch Exhaustiveness)**.** *Suppose a finite convex $N : \Omega_d \to \mathbb{R}$ satisfies $q(s) \in \partial N(s)$ everywhere. Then: (i) $N \in C^{1,1}_{\mathrm{loc}}$ and $\nabla N = q$; (ii) on the normalizer's unit surface,*

$$\Delta(x) = \sum_i x_i^2 h'(x_i) \geq 0, \qquad h'(x_i) \geq 0 \quad \textit{for every } i. \tag{12}$$

$\Delta = 0$ holds exactly when every attained coordinate is scalar-critical. If all positive critical points are strict extrema, then $\Delta > 0$. Moreover, (iii) every positive interior spectrum obeys $0 < s_i/\alpha(s) < \zeta_h$; and (iv) one constant C is forced globally:

$$\sum_{i=1}^{d} \Phi_h\left(\frac{s_i}{\alpha(s)}\right) = C, \qquad 0 < C \leq \Phi_h(\zeta_h). \tag{13}$$

Therefore a locally Lipschitz branch selector cannot retain a later branch across the first negative-slope band while realizing a convex potential on the positive singular-value cone. Restricting a full-matrix convex realization to positive diagonal matrices makes this a necessary full-matrix obstruction.

Proof. A continuous convex-subgradient selection is the ordinary gradient. By lemma 4, every $x_i \leq \zeta_h$. At a differentiability point of α, monotonicity gives sym $J_q \succeq 0$, while homogeneity gives $J_q x = 0$. Thus $J_q^\top x = 0$ and the quotient rule forces $Dx = \Delta \nabla \alpha$, without division by Δ. Consequently, the locally Lipschitz function $R(s) = \sum_i \Phi_h(x_i)$ satisfies

$$\nabla R = \frac{Dx - \Delta \nabla \alpha}{\alpha} = 0 \quad \text{almost everywhere.}$$

Connectedness of Ω_d makes $R = C$ everywhere. Taking a rank-one boundary limit gives $C = \Phi_h(1/\alpha(e_j)) \in (0, \Phi_h(\zeta_h)]$. Since $\Phi_h(t) > 0$ for $0 < t \leq \zeta_h$, an interior $x_i = \zeta_h$ would give $R > \Phi_h(\zeta_h)$, a contradiction. The slope and Δ claims follow. Details appear in appendix B.4. Neither finite criticality nor a dimension-specific exceptional-set argument is required for this necessity theorem. □

6 Full-Matrix Classification and Directional Rigidity

The cone obstruction becomes an exact matrix classification under the natural sign condition needed for spectral lifting.

Theorem 6 (Filter-generic full-matrix classification). *Let $d \geq 2$ and $h \in C^1([0,\infty))$ satisfy $h(0) = 0$, $h' > 0$ near zero, and $h' < 0$ somewhere. Assume its positive critical set is finite and consists of strict extrema. Let $\alpha : \Omega_d \to (0,\infty)$ be permutation invariant, locally Lipschitz and one-homogeneous, with a strictly positive continuous extension to the nonzero closed positive cone. Put $C_h = \Phi_h(\zeta_h)$. Using this extension on boundary spectra, the field $Q_{h,\alpha}(G) = U\,\mathrm{diag}(h(\sigma_i/\alpha(\sigma)))V^\top$, with $Q_{h,\alpha}(0) = 0$, admits a finite convex full-matrix realization if and only if $\alpha = \alpha_C$ for a unique $C \in (0, C_h]$, where the principal root satisfies*

$$\sum_i \Phi_h(s_i/\alpha_C(s)) = C, \qquad 0 < s_i/\alpha_C(s) < \zeta_h. \tag{14}$$

For every such C, the root is unique, $\mathcal{N}_C(s) = \sum_i s_i h(s_i/\alpha_C(s))$ extends to a symmetric-gauge norm, and $\Psi_C(G) = \mathcal{N}_C(\sigma(G))$ is a unitarily invariant norm whose subgradient contains the aligned field. This includes the endpoint $C = C_h$ by continuous boundary extension.

The strict-extremum assumption makes $h' > 0$ before the first fold. The implicit modular equation then has one principal radial root, its compatibility Hessian is positive semidefinite, and coordinate monotonicity gives the symmetric-gauge lift. Necessity is theorem 5; the supplement proves existence, boundary continuity, SVD independence, and the support inequality.

For the literal rational-decimal Jordan coefficients, $H = f^{\circ 5}$ means five compositions within one update, not a long-depth limit. Exact certification gives $0.004111351553545 < \zeta_J < 0.004111351553546$.

Corollary 7 (Jordan full-matrix classification). *The exact Jordan response satisfies theorem 6. Thus for $d \geq 2$ its admissible normalizers are exactly α_C, $0 < C \leq C_* = \Phi_J(\zeta_J)$. Each $\Psi_C(G) = \mathcal{N}_C(\sigma(G))$ is a unitarily invariant norm whose aligned field $Q_C(G)$ belongs to $\partial\Psi_C(G)$, including the endpoint. Here admissibility means the normalizer class explicitly stated in theorem 6; no other member has a finite convex realization.*

Remark 8 (Why $d = 1$ is excluded). For $d = 1$, $\alpha(s) = s$ gives $Q(g) = H(1)\,\mathrm{sign}(g) \in \partial(H(1)|g|)$, yet this normalizer is a convex later-branch realization because its normalized coordinate is $1 > \zeta_J$. Thus $d \geq 2$ is essential, not a notational convenience.

6.1 Which directions can a scalar multiplier repair?

Theorem 9 (Analytic projective rigidity). *Let h be real analytic and odd on $(-1,1)$, with $d \geq 4$; alternatively, let h be an odd polynomial with $d \geq 3$. Set $q_i(s) = h(s_i/\|s\|_2)$ on Ω_d. There exist $b \in C^1(\Omega_d)$, $b > 0$, and $V \in C^2(\Omega_d)$ with $bq = \nabla V$ if and only if $h(t) = at + ct^{2k+1}$ on $(-1,1)$ for some $k \geq 1$ and real a, c (including $c = 0$). Neither convexity nor homogeneity of b, V is required. If $h > 0$ on $(0,1)$ and its Taylor expansion contains two nonlinear powers, no positive scalar rescaling can be a subgradient selection of any norm, even without assuming regularity of that scalar.*

Corollary 10 (A depth boundary for monotone Newton–Schulz). *Set $g(t) = (3t - t^3)/2$ and $f(t) = (15t - 10t^3 + 3t^5)/8$. Normalize the input once by its Frobenius radius. In $d \geq 3$, g admits a norm-derived direction, but $g^{\circ L}$ for $L \geq 2$ and $f^{\circ L}$ for $L \geq 1$ admit no positive C^1 multiplier to any C^2 potential, even nonconvex. All these filters preserve order and have zero pointwise isotonic angle gap. Corollary 32 proves the full Taylor-stage classification.*

For one cubic stage the positive side is explicit: $N(s) = [(3\|s\|_2^4 - \sum_i s_i^4)/2]^{1/4}$ is a norm and $\nabla N = (\|s\|_2/N)^3 g(s/\|s\|_2)$. Its convexity follows from a weighted fourth-moment representation (corollary 32). A second composition introduces two nonlinear powers, crossing the exact boundary despite preserving order.

Write $x = s/\|s\|_2$, $w = x \odot h'(x)$, and $h(t) = tP(t^2)$. The condition $\omega \wedge d\omega = 0$ for $\omega = q^\top ds$ forces every triple determinant of (q, x, w) to vanish. A fourth coordinate fills the sphere, so the analytic determinant of $(P(y_i), 1, y_i P'(y_i))$ vanishes on an open set. Hence $yP' = AP + B$; its Taylor coefficients leave at most one nonconstant monomial. For polynomials, the top homogeneous part of the determinant also closes the constrained $d = 3$ case. Conversely, $V = (a\|s\|_2^{2k+2} + c\sum_i s_i^{2k+2})/(2k+2)$ has $\nabla V = \|s\|_2^{2k+1} q$. The supplement gives the coefficient identity, regularity steps, and sharp dimension-boundary example.

Exact obstruction versus approximate geometry. For a Taylor composition H_L, set $\Psi_L(t) = \int_0^t H_L(u)\,du$, $\Phi_L(t) = tH_L(t) - \Psi_L(t)$, and $\delta_L = \int_0^1 (1 - H_L(t))\,dt$. Write $r = \|s\|_2$, $x = s/r$. A nonlinear perspective (Briceño-Arias et al., 2023) gives an explicit symmetric norm:

$$N_L(s) = r\left[\sum_i \Psi_L(x_i) + d\delta_L\right], \qquad \nabla N_L = H_L(x) + \left[d\delta_L - \sum_i \Phi_L(x_i)\right] x. \tag{15}$$

The bracket lies in $[0, d\delta_L]$, so the gradient error is uniformly at most $d\delta_L$, including rank loss (proposition 34). At fixed dimension this tends to zero, with the prior fixed-order bound $\delta_L \le ((f_k'(0))^L + 1)^{-1}$ (Li & Tsuchiya, 2026). The correction is additive, not an asserted integrating factor. Its rank-one absolute error is exactly $(d-1)\delta_L$, although its angle there is zero. Thus exact incompatibility need not imply a persistent approximation penalty; neither does this bound promise a small practical one.

For $s^\top q(s) > 0$, pairing normalization $\beta_q(s) = q(s)/(s^\top q(s))$ removes positive output multipliers. A differentiable norm N satisfies $\beta_{\nabla N} = \nabla \log N$. Hence every nonconstant closed piecewise C^1 loop Γ obeys

$$\sup_{s\in\Gamma} \|\beta_q(s) - \nabla \log N(s)\|_2 \ge |\oint_\Gamma \beta_q^\top ds|/L(\Gamma). \tag{16}$$

This uniform ray distance is at least $375/120472576$ on an explicit three-dimensional rectangle for the monotone Taylor quintic (proposition 37). The conservative bound is not angular or evidence of a large practical effect.

Conversely, even a folded cubic can admit a norm-derived direction (appendix D), so no-gluing cannot be extended to arbitrary output multipliers. Analyticity is also essential: a smooth increasing non-binomial filter can admit a norm-derived direction (proposition 35).

For a nonzero nonnegative vector v, let $\kappa_+(v)$ be the ratio of its largest to smallest positive coordinate, with value one on rank-one support. Suppose, in addition to theorem 6, that h is strictly concave on $[0, \zeta_h]$ and its logarithmic elasticity $E_h(t) = th'(t)/h(t)$ decreases strictly. For $K > 1$, define

$$\Phi_h(\tau_{h,K,C}) + \Phi_h(\tau_{h,K,C}/K) = C, \quad \rho_{h,C}(K) = \frac{h(\tau_{h,K,C})}{h(\tau_{h,K,C}/K)}, \tag{17}$$

and let $\rho_h(K) = \rho_{h,C_h}(K)$ and $\eta_h = h(\zeta_h)/(\zeta_h h'(0))$. Set both ρ quantities to one at $K = 1$.

Theorem 11 (Sharp filter-preserving conditioning frontier)**.** *For every classified response q_C and every nonzero nonnegative spectrum s, with $K = \kappa_+(s)$,*

$$\rho_{h,C}(K) \le \kappa_+(q_C(s)) \le K. \tag{18}$$

The lower bound is attained by rank-two spectra. For $K > 1$ it decreases with C, and

$$\eta_h K < \rho_h(K) \le \kappa_+(q_C(s)) \le K, \qquad \lim_{K\to\infty} \rho_h(K)/K = \eta_h. \tag{19}$$

The ratio bounds also hold after positive rescaling of these q_C; they do not classify all multiplier-corrected fields $b(s)h(s/\alpha(s))$.

For $K > 1$, the proof reduces each spectrum to its largest and smallest coordinates. Nonnegative modular terms bound the largest normalized coordinate by $\tau_{h,K,C}$; decreasing elasticity makes $h(t)/h(t/K)$ decrease with t. Rank two exhausts the modular constraint. Strict concavity gives the upper bound and the endpoint ratio η_h. For example, every odd one-fold filter $h_p(t) = at - bt^p$ ($a, b > 0$, odd $p \ge 3$) satisfies the hypotheses and has $\eta_{h_p} = 1 - 1/p$, giving an analytic family of sharp frontiers independent of the Jordan coefficients.

Corollary 12 (Sharp Jordan conditioning frontier)**.** *The Jordan principal branch satisfies the shape hypotheses above. Thus eqs. (18) and (19) hold with $h = H$, $C_h = C_*$, $\rho_{h,C} = \rho_{J,C}$, $\rho_h = \rho_J$, and*

$$\eta_J = \frac{H(\zeta_J)}{\zeta_J H'(0)} = 0.6031454673894552\ldots. \tag{20}$$

The supplement supplies the complete generic proof and coefficient-exact Jordan shape certificate.

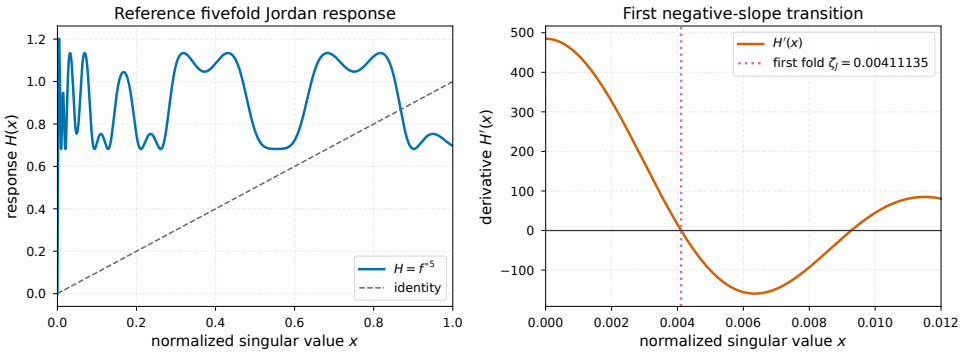


Figure 1: **Literal Jordan response and its first fold.** The right panel resolves the first zero of H' at $\zeta_J \approx 0.00411135$. Every unscaled convex shared-normalizer realization is confined strictly left of this fold on positive full-rank spectra.

For $\vartheta(R) = \arccos((R+1)/(\sqrt{2}\sqrt{R^2+1}))$, a rank-two spectrum $(K, 1)$ therefore remains at least $\vartheta(\rho_J(K))$ from polar. Certified bounds are $36.354979°$ at $K = 10$ and $44.058868°$ at $K = 100$; literal Frobenius-normalized Jordan is only $0.271129°$ and $0.098823°$ from polar, respectively. Thus within the classified unscaled-field family, preserving literal H while imposing fixed convex geometry is not a cosmetic renormalization.

Also $\alpha_C(s) \geq \|s\|_\infty/\zeta_J > 243.229\|s\|_\infty$, with equality in the first bound only at maximum-level rank-one rays. Computational and norm-smooth descent details appear in the supplement.

7 Directional Diagnostics on Stored Spectra

Do these obstructions occur on stored optimizer inputs? We measure two distinct quantities: angle to an admissible norm direction at each input, and a local obstruction to one norm-derived field. Neither is a loss metric. For the first, given positive s, y, let $\mathcal{K}_s = \{z \geq 0 : (s_i - s_j)(z_i - z_j) \geq 0 \ \forall i, j\}$ and let $p = \Pi_{\mathcal{K}_s} y$. Then

$$\inf_{\substack{\gamma \text{ absolutely symmetric norm} \\ 0 \neq z \in \partial\gamma(s)}} \angle(y, z) = \arccos \frac{\|p\|_2}{\|y\|_2}. \tag{21}$$

The infimum is attained by the support function of the signed-permutation orbit of p. Classical invariance and projection arguments (Lewis, 1996) give a sharper quantifier when y is constant on ties of s: over all finite convex potentials and nonzero subgradient selections F, the minimum of $\max_R \angle(F(Rs), Ry)$ is exactly the right side of (21), with R ranging over signed permutations (proposition 52). Mere monotonicity gives the same value. These orbit inputs need not have been encountered.

Source and sampling. The source spectra come from FP32 character-level causal decoders on Tiny Shakespeare: four sizes (0.805–85.072M parameters), five seeds, and 250 steps, with six larger-model trajectories extended to 2,000 steps. The supplement specifies the dependent layer/checkpoint sampling. A hash-gated raw replay analyzes 4,872 records from 26 scientific NPZ payloads (the 27th archived payload is timing-only), with 4,728 unique spectra after 144 deterministic overlaps.

Pointwise angular separation. For $y_i = H(s_i/(\|s\|_2 + 10^{-7}))$, replay of the exact angle formula (21) ranges from $7.3807°$ to $10.8858°$ (median $9.1352°$); every one of the 20 unique scale–seed trajectory summaries exceeds the preregistered $5°$ threshold. This is a dependent, pointwise encounter statement: the attaining gauge may change with s, and the replay estimates neither a population quantity nor a training penalty.

Field-level obstruction. For the epsilon-free field $q = H(s/\|s\|_2)$, we separately replay a multiplier-invariant obstruction to any fixed norm-derived direction:

$$\mathcal{P}(s) = \frac{r\|\omega \wedge d\omega\|}{\|q\|_2^2} = \frac{\sqrt{\det \mathrm{Gram}(q, x, x \odot H'(x))}}{\|q\|_2^2}. \tag{22}$$

Since $(b\omega) \wedge d(b\omega) = b^2\omega \wedge d\omega$, this quantity is invariant under *every* positive C^1 output multiplier. Its minimum, median, and maximum are 0.650488, 0.783272, and 1.111339. The integrable cubic control has nonzero ordinary curl but $\mathcal{P} \le 1.85 \times 10^{-16}$. This dependent float64 replay establishes encounter, not a training effect; zero would not certify global convexity. Exploratory replay of g, $g^{\circ 2}$, f, and $f^{\circ 5}$ from corollary 10 preserves order, while the latter three have nonzero $\mathcal{P}$ (table 2). These counterfactual evaluations use Jordan-generated spectra, not trajectories trained with the alternative filters.

Jordan still satisfies $\langle G, Q_0(G)\rangle_F \ge H(1)\|G\|_F$. This positive pairing with the input, possibly momentum, does not guarantee a finite-step decrease of the training loss.

8 Implications for Analysis and Design

Table 1 distinguishes norm-derived rays from fields requiring alignment, approximation, or roundwise-potential arguments. One cubic step admits scalar repair; deeper Taylor fields do not. Jordan's exclusion is elementary.

Corollary 13 (Order reversal survives positive scaling)**.** *For $d = \min\{m, n\} \ge 2$, no positive scalar function λ makes the exact-real epsilon-free Jordan field λQ_0 a selection of one finite convex subdifferential on all full-rank matrices. In particular it cannot be an exact support oracle of any fixed matrix-norm ball. Neither symmetry nor regularity of λ is assumed.*

Proof. The certified witness in proposition 50 has $s_i > s_j$ but $q_i < q_j$. For the transposition P of these coordinates,

$$\langle b_s q - b_{Ps} Pq, s - Ps\rangle = (b_s + b_{Ps})(s_i - s_j)(q_i - q_j) < 0$$

for every $b_s, b_{Ps} > 0$. This contradicts subgradient monotonicity. Adding sufficiently small positive coordinates preserves the strict reversal by continuity; rectangular diagonal embedding preserves the pairing. □

9 Limitations

The exact-real classifications require the stated dimension, regularity, critical-point, and boundary assumptions. Rounded BF16 kernels, full optimizer dynamics, and stateful or nonspectral rules are outside scope. Conservativity does not imply convexity; the conditioning frontier additionally requires strict pre-fold concavity and decreasing logarithmic elasticity.

Even a large polar-angle gap need not imply a large support-oracle objective gap: on $s = (K, 1)$, an operator-normalized response $(1, 1/R)$ achieves the fraction $(K + 1/R)/(K + 1)$ of polar's linear objective, tending to one as $K \to \infty$ when $R \asymp K$. Our angles are geometric, not loss penalties. Dependent FP32 Tiny-Shakespeare replay establishes no training effect, independent replication, native-BF16 behavior, or cross-corpus universality; the deterministic witnesses are likewise noncausal.

10 Conclusion

Under the stated hypotheses, shared normalization forces a principal convex modular family and its sharp conditioning frontier. Frobenius-scale polynomial rigidity starts at $d = 3$, beyond order preservation. Fixed geometry, polar proximity, descent, and training quality remain distinct.

Reproducibility Statement

The stated assumptions and complete proofs appear below. Two anonymous artifacts separate theorem certificates from empirical evidence: the first contains exact and interval verifiers, while the second contains all 27 stored singular-value payloads and one command that replays all diagnostic metrics. Digest manifests cover both bundles. The limitations separate reproducibility checks from claims that are not made.

AI Usage Disclosure

Generative-AI systems assisted with research ideation, mathematical derivations and proof drafting, literature discovery, methodology and experiment design, code and certificate construction, red-teaming, and manuscript editing. AI-assisted mathematical and numerical claims were checked against written derivations and targeted exact-rational, interval, or replay verifiers; those checks are not external proof certification. Responsibility for the research questions, protocols, final content, and all claims and artifacts rests with the authors. AI systems are not authors.

References

José Carlos R. Alcantud and Carlos Rodríguez-Palmero. Complete solution of the integrability problem for homothetic demand functions. *International Journal of Economic Theory*, 6(2):263–271, 2010. doi: 10.1111/j.1742-7363.2010.00134.x. URL https://doi.org/10.1111/j.1742-7363.2010.00134.x.

Noah Amsel, David Persson, Christopher Musco, and Robert M. Gower. The Polar Express: Optimal matrix sign methods and their application to the Muon algorithm. In *International Conference on Learning Representations*, 2026. URL https://arxiv.org/abs/2505.16932. arXiv:2505.16932.

Shayan Aziznejad and Michael Unser. Duality mapping for schatten matrix norms. *Numerical Functional Analysis and Optimization*, 42(6):679–695, 2021. doi: 10.1080/01630563.2021.1922438. URL https://doi.org/10.1080/01630563.2021.1922438.

Han Bao, Amirreza Eshraghi, and Yutong Wang. Brenier isotonic regression. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, 2026. URL https://arxiv.org/abs/2603.10452. AISTATS 2026 poster; arXiv:2603.10452.

Charles Blackorby, Daniel Primont, and R. Robert Russell. *Duality, Separability, and Functional Structure: Theory and Economic Applications*. North-Holland, 1978.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. URL https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

H. W. Braden and J. G. B. Byatt-Smith. On a functional differential equation of determinantal type. *Bulletin of the London Mathematical Society*, 31(4):463–470, 1999. doi: 10.1112/S0024609398005542. URL https://arxiv.org/abs/math/9804082.

Luis M. Briceño-Arias, Patrick L. Combettes, and Francisco J. Silva. Perspective functions with nonlinear scaling. *arXiv preprint arXiv:2210.16937*, 2023. URL https://arxiv.org/abs/2210.16937v2. Revised version, October 17, 2023.

Luis A. Caffarelli and Michael G. Crandall. Distance functions and almost global solutions of eikonal equations. *Communications in Partial Differential Equations*, 35(3):391–414, 2010. doi: 10.1080/03605300903253927. URL https://web.math.ucsb.edu/~crandall/papers/deik.pdf.

Da Chang, Qiankun Shi, Lvgang Zhang, Di He, Yaoshuai Ma, Ganzhao Yuan, and Yongxiang Liu. Muse: Representation geometry of Muon beyond normalized momentum, 2026. URL https://arxiv.org/abs/2607.14536.

Tong Che. Second-order Muon done right: A principled marriage of spectral geometry and curvature, 2026. URL https://arxiv.org/abs/2608.09763.

Lizhang Chen, Jonathan Li, and Qiang Liu. Muon optimizes under spectral norm constraints, 2025. URL https://arxiv.org/abs/2506.15054.

Michael Crawshaw, Chirag Modi, Mingrui Liu, and Robert M. Gower. An exploration of non-euclidean gradient descent: Muon and its many variants, 2025. URL https://arxiv.org/abs/2510.09827.

Chao Ding, Defeng Sun, Jie Sun, and Kim-Chuan Toh. Spectral operators of matrices. *Mathematical Programming*, 168:509–531, 2018. doi: 10.1007/s10107-017-1162-3. URL https://arxiv.org/abs/1401.2269.

Thang Do, Steffen Dereich, and Arnulf Jentzen. On MUON optimization: From non-convergence to an error analysis with Polar Express and the Newton–Schulz polynomial from implementations, 2026. URL https://arxiv.org/abs/2608.04607.

Yihe Dong and Will Sawin. Muon^p: Muon with fractional spectral powers, 2026. URL https://arxiv.org/abs/2606.13867.

György Elekes and Endre Szabó. On triple lines and cubic curves: The orchard problem revisited. *Discrete & Computational Geometry*, 72:743–763, 2024. doi: 10.1007/s00454-023-00556-3. URL https://doi.org/10.1007/s00454-023-00556-3.

Thibault Fally. Generalized separability and integrability: Consumer demand with a price aggregator. *Journal of Economic Theory*, 203:105471, 2022. doi: 10.1016/j.jet.2022.105471. URL https://doi.org/10.1016/j.jet.2022.105471.

Chen Fan, Mark Schmidt, and Christos Thrampoulidis. Implicit bias of spectral descent and Muon on multiclass separable data. In *Advances in Neural Information Processing Systems*, 2025. URL https://arxiv.org/abs/2502.04664. Spotlight; arXiv:2502.04664.

Ekaterina Grishina, Matvey Smirnov, and Maxim Rakhuba. Accelerating Newton–Schulz iteration for orthogonalization via Chebyshev-type polynomials, 2025. URL https://arxiv.org/abs/2506.10935.

Eitan Gronich and Gal Vardi. The implicit bias of Adam and Muon on smooth homogeneous neural networks, 2026. URL https://arxiv.org/abs/2602.16340.

Yuhki Hosoya. Consumer optimization and a first-order PDE with a non-smooth system. *Operations Research Forum*, 2:66, 2021. doi: 10.1007/s43069-021-00104-w. URL https://doi.org/10.1007/s43069-021-00104-w.

Yuhki Hosoya. On Gale's contribution in revealed preference theory, 2022. URL https://arxiv.org/abs/2208.07970.

Yinze Hu, Hongjun Xiang, Xingao Gong, and Hongyu Yu. A physical response-and-memory model for Muon optimization, 2026. URL https://arxiv.org/abs/2608.22994.

Hua Huang. How much orthogonalization does Muon need?, 2026. URL https://arxiv.org/abs/2606.00371.

Henryk Hudzik and Lech Maligranda. Amemiya norm equals Orlicz norm in general. *Indagationes Mathematicae*, 11(4):573–585, 2000. doi: 10.1016/S0019-3577(00)80026-9.

Keller Jordan, Yuchen Jin, Vlado Boza, Jiacheng You, Franz Cesista, Laker Newhouse, and Jeremy Bernstein. Muon: An optimizer for hidden layers in neural networks, 2024. URL https://github.com/KellerJordan/Muon/blob/f90a42b28e00b8d9d2d05865fe90d9f39abcbcbd/muon.py. Official citation and reference implementation; source pinned at commit f90a42b28e00b8d9d2d05865fe90d9f39abcbcbd.

Thomas Kesselheim, Marco Molinaro, and Sahil Singla. Supermodular approximation of norms and applications. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pp. 1841–1852, 2024. doi: 10.1145/3618260.3649734. URL https://doi.org/10.1145/3618260.3649734.

Gyu Yeol Kim and Min-hwan Oh. Convergence of Muon with Newton–Schulz. In *International Conference on Learning Representations*, 2026. URL https://arxiv.org/abs/2601.19156. arXiv:2601.19156.

Oleksiy Klurman. V. Markov's Problem for Monotone Polynomials, 2012. URL https://arxiv.org/abs/1205.0846. Theorem 2.1 gives the sharp pointwise derivative inequality whose origin evaluation specializes to the monotone odd-polynomial constants used in the auxiliary barrier calculation.

Alexey Kushnir and Vinod Krishnamoorthy. A simple characterization of supply correspondences. *Mathematics of Operations Research*, 49(2):826–831, 2024. doi: 10.1287/moor.2023.1379. URL https://doi.org/10.1287/moor.2023.1379.

Adrian S. Lewis. The convex analysis of unitarily invariant matrix functions. *Journal of Convex Analysis*, 2(1–2):173–183, 1995. URL https://people.orie.cornell.edu/~aslewis/publications/95-convex.pdf.

Adrian S. Lewis. Group invariance and convex matrix analysis. *SIAM Journal on Matrix Analysis and Applications*, 17(4):927–949, 1996. URL https://people.orie.cornell.edu/aslewis/publications/96-group.pdf.

Mingyi Li and Taira Tsuchiya. Muon with finite Newton–Schulz: The smoothing benefit in nonsmooth nonconvex optimization, 2026. URL https://arxiv.org/abs/2608.26288.

Juan Enrique Martínez-Legaz. Fenchel subdifferential operators: A characterization without cyclic monotonicity. *Journal of Applied and Numerical Optimization*, 5(1):181–187, 2023. doi: 10.23952/jano.5.2023.1.12. URL https://jano.biemdas.com/issues/JANO2023-1-12.pdf.

Julian Musielak. *Orlicz Spaces and Modular Spaces*, volume 1034 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 1983. doi: 10.1007/BFb0072210. URL https://doi.org/10.1007/BFb0072210.

NVIDIA. NeMo emerging optimizers: Muon utilities, 2026. URL https://github.com/NVIDIA-NeMo/Emerging-Optimizers/blob/170d1b864049ab58fab134941eb6f268b3e64dee/emerging_optimizers/orthogonalized_optimizers/muon_utils.py. Hard-clamp path with normalize-in-double disabled; pinned source commit 170d1b864049ab58fab134941eb6f268b3e64dee, accessed 9 August 2026.

Gabriel Peyré. Muon dynamics as a spectral Wasserstein flow, 2026. URL https://arxiv.org/abs/2604.04891v2. Version 2, Appendix B.

Artem Riabinin, Egor Shulgin, Kaja Gruntkowska, and Peter Richtárik. Gluon: Bridging theory and practice of LMO-based optimizers for LLMs, 2025. URL https://arxiv.org/abs/2505.13416.

R. Tyrrell Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific Journal of Mathematics*, 17(3):497–510, 1966. doi: 10.2140/pjm.1966.17.497. URL https://msp.org/pjm/1966/17-3/pjm-v17-n3-p08-p.pdf.

Zakhar Shumaylov, Nathaël Da Costa, Peter Zaika, Bálint Mucsányi, Alex Massucco, Yoav Gelberg, Carola-Bibiane Schönlieb, Yarin Gal, and Philipp Hennig. Muon is not that special: Random or inverted spectra work just as well, 2026. URL https://arxiv.org/abs/2605.11181.

Fangzhou Wu, Rikhav Shah, Sandeep Silwal, and Qiuyi Zhang. DynMuon: A dynamic spectral shaping view of Muon, 2026. URL https://arxiv.org/abs/2605.17109.

A Complete proof chain for compatibility and no-gluing

This supplement supplies the full load-bearing proof chain for the regular normalizer–filter compatibility theorem and the locally Lipschitz no-gluing theorem in the main paper. It deliberately repeats hypotheses at the point where they are used. No differentiability on the unit surface is inferred from surface measure. The only Rademacher argument below takes place in the ambient open cone and reaches the unit surface by radial normalization. The direct necessity proof in appendix B.4 uses only coordinate-line monotonicity, positive boundary continuity, and ambient almost-everywhere compatibility. It does not need an inertia/topology route or any finite-critical-set hypothesis. The inertia identity below is retained only to describe local Hessians, not as a prerequisite or an additional novelty claim. Earlier longer derivations remain in `theory/matched_normalizer_duality.md` for the smooth equivalence and `theory/abstract_spectral_geometry_classification.md` for the locally Lipschitz no-gluing chain.

A.1 Notation and hypotheses

Fix an integer $d \geq 2$ and write

$$\Omega_d = (0, \infty)^d.$$

Let $\alpha : \Omega_d \to (0, \infty)$ be one-homogeneous:

$$\alpha(\lambda s) = \lambda \alpha(s) \qquad (s \in \Omega_d,\ \lambda > 0).$$

For a scalar response h, define

$$x(s) = \frac{s}{\alpha(s)}, \qquad q_i(s) = h(x_i(s)), \tag{23}$$

and, at a normalized point x,

$$D(x) = \operatorname{diag}(h'(x_1), \ldots, h'(x_d)), \quad w(x) = D(x)x, \quad \Delta(x) = x^\top D(x) x. \tag{24}$$

The unit surface is

$$S = \{x \in \Omega_d : \alpha(x) = 1\}. \tag{25}$$

Whenever a primitive is needed, Φ denotes a function satisfying

$$\Phi'(t) = t h'(t). \tag{26}$$

Known support representation for homogeneous monotone fields. The passage from monotonicity to convex geometry is already known for zero-homogeneous fields (Kushnir & Krishnamoorthy, 2024, Theorem 1). Here is the short argument on our open cone, without requiring a value at the origin. If q is monotone and $q(\lambda t) = q(t)$, then $(q(s) - q(t))^\top (s - \lambda t) \geq 0$ for every $\lambda > 0$. Letting $\lambda \downarrow 0$ gives

$$s^\top q(s) \geq s^\top q(t), \qquad N(s) := \sup_{t \in \Omega_d} s^\top q(t) = s^\top q(s).$$

Thus N is finite on Ω_d, convex and one-homogeneous, with $q(s) \in \partial N(s)$. If the range is bounded, its closed convex hull is compact and supplies the same support function. Conversely, a convex-subgradient selection is monotone. Related Banach-space support characterizations are given by Martínez-Legaz (2023, Corollary 2.1 and Theorem 2.2); those statements use an origin-range condition, with maximality for equality of operators. None of these representation results determines the allowable shared normalizers. The inverse classification below concerns that additional functional restriction, not a new general implication from monotonicity to cyclic monotonicity.

A.2 The regular compatibility theorem

In this subsection only, assume that α is C^2, positive, and one-homogeneous on Ω_d, that $h \in C^2(I)$ on an interval containing every attained normalized coordinate, that S is connected, and that

$$\Delta(x) \neq 0 \qquad (x \in S). \tag{27}$$

Here "q is conservative" means that there is a C^2 scalar function F on Ω_d with $\nabla F = q$.

Theorem 14 (Regular normalizer–filter compatibility)**.** *Under the hypotheses just stated, the following are equivalent.*

1. *The field q is conservative on Ω_d.*
2. *The Jacobian J_q is symmetric on Ω_d.*
3. *The normalizer satisfies*

$$\nabla\alpha(s) = \frac{D(x)x}{x^\top D(x)x} = \frac{w(x)}{\Delta(x)}, \qquad x = \frac{s}{\alpha(s)}. \tag{28}$$

4. *There is a constant C such that S is contained in the regular level surface*

$$\sum_{i=1}^{d} \Phi(x_i) = C, \tag{29}$$

and $\alpha(s)$ selects a positive regular radial branch of

$$\sum_{i=1}^{d} \Phi\left(\frac{s_i}{\alpha(s)}\right) = C. \tag{30}$$

When these conditions hold, a one-homogeneous potential is

$$\mathcal{N}_h(s) = \sum_{i=1}^{d} s_i h\left(\frac{s_i}{\alpha(s)}\right), \tag{31}$$

and

$$\nabla\mathcal{N}_h(s) = q(s), \qquad \nabla^2\mathcal{N}_h(s) = \frac{1}{\alpha(s)}\left(D - \frac{Dxx^\top D}{x^\top Dx}\right). \tag{32}$$

If h' has one fixed nonzero sign on an interval $I \subset (0,\infty)$ containing every attained normalized coordinate, then, for each fixed s, there is at most one radial root whose candidate coordinates s_i/a all lie in I. Positive roots that leave I are not excluded, and the unit surface need not equal the entire positive modular level surface.

Proof. Write $a = \alpha(s)$ and $\alpha_j = \partial\alpha/\partial s_j$. Since $x_i = s_i/a$, the quotient rule gives

$$\frac{\partial x_i}{\partial s_j} = \frac{\delta_{ij}a - s_i\alpha_j}{a^2} = \frac{1}{a}(\delta_{ij} - x_i\alpha_j). \tag{33}$$

The chain rule therefore gives

$$\frac{\partial q_i}{\partial s_j} = \frac{h'(x_i)}{a}(\delta_{ij} - x_i\alpha_j). \tag{34}$$

For $i \neq j$, symmetry of J_q is, by (34), equivalent to

$$x_i h'(x_i)\alpha_j = x_j h'(x_j)\alpha_i. \tag{35}$$

Equation (35) says that every two-by-two minor of the two-column matrix $[w, \nabla\alpha]$ vanishes. By (27), $\Delta \neq 0$, hence $w \neq 0$. Consequently $\nabla\alpha = \lambda w$ for some scalar λ.

Euler's identity for a differentiable one-homogeneous function states that $s^\top\nabla\alpha(s) = \alpha(s) = a$. Since $s = ax$ and $a > 0$, division by a gives

$$x^\top\nabla\alpha(s) = 1.$$

Substituting $\nabla\alpha = \lambda w$ yields $1 = \lambda x^\top w = \lambda\Delta$, so $\lambda = 1/\Delta$. This proves that Jacobian symmetry implies (28). Conversely, substituting (28) into (35) makes all mixed partials equal, so it implies Jacobian symmetry.

If $q = \nabla F$ with $F \in C^2$, its Jacobian is the Hessian of F and is symmetric. This proves (1)⇒(2). For the reverse implication, Ω_d is convex and therefore simply connected. The Poincaré lemma for C^1

one-forms on a simply connected open subset of Euclidean space says that a C^1 field with symmetric Jacobian is conservative. Hence (2)⇒(1).

Define $R(x) = \sum_i \Phi(x_i)$. By (26), $\nabla R(x) = w(x)$. Equation (28) on S reads

$$\nabla\alpha(x) = \frac{\nabla R(x)}{x^\top \nabla R(x)}. \tag{36}$$

Both denominators in this statement are nonzero by (27). Therefore $\nabla\alpha$ and ∇R have the same tangent kernel at every point of S. Along every C^1 curve γ in S, the identity $\nabla\alpha(\gamma(t))^\top\gamma'(t) = 0$ and (36) imply $\frac{d}{dt}R(\gamma(t)) = 0$. Thus R is constant on each path component of S.

For completeness, S is path connected under the present hypotheses. Let $P = \{u \in (0,\infty)^d : \sum_i u_i = 1\}$. The maps

$$u \longmapsto \frac{u}{\alpha(u)}, \qquad x \longmapsto \frac{x}{\sum_i x_i}$$

are continuous inverses between the convex set P and S; the inverse identities follow directly from one-homogeneity. Hence S is path connected, so one constant C gives (29). One-homogeneity sends every s to $x = s/\alpha(s) \in S$, which gives (30).

Conversely, suppose $\alpha(s)$ selects a differentiable regular radial branch of (30). Differentiate that equation with respect to s_j. Using (33) and $\Phi'(x_i) = x_i h'(x_i)$ gives

$$0 = \frac{1}{a}\left[x_j h'(x_j) - \alpha_j \sum_i x_i^2 h'(x_i)\right].$$

Multiplication by $a > 0$ and division by $\Delta \neq 0$ give (28). This proves (4)⇒(3), and the previous paragraph proves (3)⇒(4).

Differentiate (31). For every j,

$$\begin{aligned}
\frac{\partial \mathcal{N}_h}{\partial s_j} &= h(x_j) + \sum_i s_i h'(x_i)\frac{\partial x_i}{\partial s_j} \\
&= h(x_j) + \sum_i a x_i h'(x_i)\frac{1}{a}(\delta_{ij} - x_i\alpha_j) \\
&= h(x_j) + x_j h'(x_j) - \alpha_j \sum_i x_i^2 h'(x_i) \\
&= h(x_j),
\end{aligned}$$

where the last equality is (28). Thus $\nabla\mathcal{N}_h = q$. Finally, substitute $\nabla\alpha = w/\Delta$ into the matrix form of (34):

$$J_q = \frac{1}{a}(D - w\nabla\alpha^\top) = \frac{1}{a}\left(D - \frac{ww^\top}{\Delta}\right).$$

Since $w = Dx$ and $D = D^\top$, this is (32).

For fixed s, the derivative with respect to a candidate radial scale a of $\sum_i \Phi(s_i/a)$ is

$$-\frac{1}{a}\sum_i (s_i/a)^2 h'(s_i/a).$$

On the interval of candidate scales for which every s_i/a lies in I, the fixed nonzero sign of h' makes this derivative have one fixed nonzero sign. The radial modular is therefore strictly monotone there and has at most one root there. This argument makes no assertion about candidate scales whose normalized coordinates leave I. □

A.3 Fixed Schatten normalizers

The following supplies the proof of the fixed-normalizer corollary in the main paper, including the nuclear-normalized endpoint. Let $d \geq 3$, let $1 \leq p < \infty$, put $\alpha_p(s) = \|s\|_p$, and write $x = s/\alpha_p(s)$.

Corollary 15 (Fixed Schatten-normalizer classification)**.** *Let $h \in C^2((0,1))$ and define $q_i(s) = h(s_i/\|s\|_p)$ on Ω_d. If $1 < p < \infty$, then q is conservative if and only if*

$$h(t) = At^{p-1} + B \qquad (0 < t < 1) \tag{37}$$

for constants $A, B \in \mathbb{R}$. Its potential is

$$F_p(s) = A\|s\|_p + B\sum_i s_i, \tag{38}$$

and it is convex on Ω_d if and only if $A \geq 0$. For $p = 1$, conservativity is equivalent to

$$h(t) = C\log t + B \qquad (0 < t < 1), \tag{39}$$

with potential

$$F_1(s) = C\left[\sum_i s_i \log s_i - \Big(\sum_i s_i\Big)\log\Big(\sum_i s_i\Big)\right] + B\sum_i s_i. \tag{40}$$

This latter potential is convex on Ω_d if and only if $C \geq 0$. On positive full-rank matrices, (38) has the spectral formula $A\|G\|_{S_p} + B\|G\|_*$. When the matrix space has at least two singular-value coordinates, this formula extends to a globally convex full-matrix function if and only if $A \geq 0$ and $B \geq 0$.

Proof. First suppose $1 < p < \infty$. Differentiation of the p-norm gives

$$\frac{\partial \alpha_p}{\partial s_j} = x_j^{p-1}. \tag{41}$$

Substituting (41) into the quotient-rule calculation (34), symmetry of the (i,j) and (j,i) entries for $i \neq j$ is equivalent to

$$x_i h'(x_i)x_j^{p-1} = x_j h'(x_j)x_i^{p-1}, \quad \text{or equivalently} \quad h'(x_i)x_i^{2-p} = h'(x_j)x_j^{2-p}. \tag{42}$$

Let arbitrary $u, v \in (0,1)$ be given. Choose $\tau > 0$ so small that $u^p + \tau^p < 1$ and $v^p + \tau^p < 1$. Because $d \geq 3$, each of (u,τ) and (v,τ) can be completed, using positive remaining coordinates, to a positive p-unit vector. Applying (42) first to (u,τ) and then to (v,τ) gives

$$h'(u)u^{2-p} = h'(\tau)\tau^{2-p} = h'(v)v^{2-p}.$$

Thus one constant K satisfies $h'(t) = Kt^{p-2}$ on $(0,1)$. Integration gives (37), with $A = K/(p-1)$. Conversely, direct substitution in (42) proves symmetry of the Jacobian for every response of the form (37); the Poincaré lemma on the convex set Ω_d then gives conservativity. More directly,

$$\frac{\partial F_p}{\partial s_i} = A\frac{s_i^{p-1}}{\|s\|_p^{p-1}} + B = Ax_i^{p-1} + B = q_i(s),$$

which proves (38).

For convexity, $\|s\|_p$ is convex, so $A \geq 0$ is sufficient. To see necessity, fix $s \in \Omega_d$ and choose a nonzero vector z satisfying $\sum_i s_i^{p-1} z_i = 0$. Such a vector exists because $d \geq 2$. A direct second differentiation gives

$$z^\top \nabla^2 \|s\|_p z = (p-1)\|s\|_p^{1-p}\sum_i s_i^{p-2} z_i^2 > 0.$$

The linear B-term has zero Hessian, so $A < 0$ produces a negative second directional derivative and cannot be convex.

Now let $p = 1$. Here $\alpha_1(s) = \sum_i s_i$ and $\partial_j \alpha_1 = 1$. Equality of the mixed partials is therefore

$$x_i h'(x_i) = x_j h'(x_j). \tag{43}$$

Given $u, v \in (0,1)$, choose $\tau > 0$ with $u + \tau < 1$ and $v + \tau < 1$, and complete (u,τ) and (v,τ) to positive unit 1-norm vectors. Applying (43) twice shows that $th'(t)$ is one constant C. Hence $h'(t) = C/t$, which integrates to (39). Differentiating (40), with $S = \sum_i s_i$, gives

$$\frac{\partial F_1}{\partial s_i} = C(\log s_i - \log S) + B = C\log x_i + B = q_i(s).$$

Finally,

$$z^\top \nabla^2 F_1(s) z = C \left(\sum_i \frac{z_i^2}{s_i} - \frac{(\sum_i z_i)^2}{S} \right).$$

The bracket is nonnegative by weighted Cauchy–Schwarz. It is strictly positive, for example, for any nonzero z with $\sum_i z_i = 0$. Therefore convexity is equivalent to $C \geq 0$.

For $p > 1$, let $G = U \operatorname{diag}(s) V^\top$ have positive singular values. At a simple singular value, the first-order SVD identity is $ds_i = u_i^\top (dG) v_i$. Therefore differentiating the following spectral function gives

$$\widetilde{F}_p(G) = A\|G\|_{S_p} + B\|G\|_*, \qquad \nabla \widetilde{F}_p(G) = U \operatorname{diag}\Big(A x_i^{p-1} + B \Big) V^\top.$$

When singular values coincide, the displayed diagonal coefficient is equal on the corresponding singular subspace, so the expression is independent of the choice of singular vectors; continuity gives the same formula throughout the positive full-rank stratum. Thus $\widetilde{F}_p$ is precisely the matrix lift of (38). If $A, B \geq 0$, this is a nonnegative linear combination of two norms and is therefore globally convex. Conversely, the positive-diagonal calculation above already proves $A \geq 0$: its B-term is linear there, while $A < 0$ gives the displayed negative second directional derivative. To force the remaining sign, suppose the matrix space has at least two singular-value coordinates and restrict $\widetilde{F}_p$ to the affine line

$$G(t) = \operatorname{diag}(1, t, 0, \ldots, 0).$$

Along this line,

$$\widetilde{F}_p(G(t)) = A(1 + |t|^p)^{1/p} + B(1 + |t|).$$

The first term has derivative zero at $t = 0$. Hence the left and right derivatives of this restriction at zero are respectively $-B$ and B. A convex real function has left derivative no larger than its right derivative, so $-B \leq B$, equivalently $B \geq 0$. This proves the claimed full-matrix condition without extending positive-cone convexity across a rank-loss stratum by assertion. □

B Complete proofs for stabilized Frobenius maps

This section proves the two implementation-level propositions stated in the main paper. They are not consequences of the homogeneous compatibility theorem: the additive denominator is not homogeneous, while the hard clamp is piecewise homogeneous and normally creates a derivative jump at its clamp sphere. Throughout this section

$$\Omega_d = (0, \infty)^d, \qquad d \geq 2, \qquad \varepsilon > 0.$$

B.1 The additive Frobenius denominator

Let $h \in C^1([0, \infty))$, set $r = \|s\|_2$, $a = r + \varepsilon$, and define

$$x_i = \frac{s_i}{a}, \qquad q_i(s) = h(x_i). \tag{44}$$

Theorem 16 (Additive-denominator classification)**.** *The field in (44) is conservative on Ω_d if and only if*

$$h(t) = At + B \qquad (0 < t < 1) \tag{45}$$

for constants $A, B \in \mathbb{R}$. In that case every potential differs by a constant from

$$F_\varepsilon(s) = A \left[r - \varepsilon \log\left(\frac{r + \varepsilon}{\varepsilon} \right) \right] + B \sum_{i=1}^d s_i. \tag{46}$$

It is convex on Ω_d if and only if $A \geq 0$. Equivalently, a finite convex function N can satisfy $q(s) \in \partial N(s)$ everywhere if and only if (45) holds with $A \geq 0$, and then $N - F_\varepsilon$ is constant.

Proof. First identify exactly which normalized vectors are attained. If $x = s/(\|s\|_2 + \varepsilon)$, then $x_i > 0$ and $\|x\|_2 = r/(r + \varepsilon) < 1$. Conversely, if $x_i > 0$ and $\rho = \|x\|_2 < 1$, set

$$s = \frac{\varepsilon x}{1 - \rho}. \tag{47}$$

Then $\|s\|_2 = \varepsilon\rho/(1-\rho)$ and direct substitution in (44) returns x. Thus the attained set is the open positive unit ball.

For $i \neq j$, the quotient and chain rules give

$$\frac{\partial x_i}{\partial s_j} = -\frac{s_i s_j}{r(r+\varepsilon)^2}, \tag{48}$$

$$\frac{\partial q_i}{\partial s_j} = -\frac{s_i s_j}{r(r+\varepsilon)^2} h'(x_i). \tag{49}$$

Every displayed prefactor is nonzero on Ω_d. Hence symmetry of the Jacobian is equivalent to

$$h'(x_i) = h'(x_j) \tag{50}$$

whenever the two coordinates occur in one attained vector.

Let arbitrary $u, v \in (0,1)$ be given. Choose $t > 0$ so small that both $u^2 + t^2 < 1$ and $v^2 + t^2 < 1$. In dimension two, (u,t) and (v,t) already belong to the attained ball. In higher dimension, append sufficiently small positive coordinates so that the resulting vectors still have norm below one. Applying (50) to the first two coordinates of these two vectors gives

$$h'(u) = h'(t) = h'(v).$$

Since u, v were arbitrary, h' is a constant A on $(0,1)$, and integration gives (45). A conservative C^1 field has symmetric Jacobian, so this proves necessity.

Conversely, suppose (45) holds. Differentiating (46) gives

$$\frac{\partial F_\varepsilon}{\partial s_i} = A\left(1 - \frac{\varepsilon}{r+\varepsilon}\right)\frac{s_i}{r} + B = A\frac{s_i}{r+\varepsilon} + B = q_i(s).$$

Thus the field is conservative. The Hessian of the nonlinear radial term is

$$A\left[\frac{1}{r+\varepsilon} I - \frac{ss^\top}{r(r+\varepsilon)^2}\right]. \tag{51}$$

On $s^\perp$ its eigenvalue is $A/(r+\varepsilon)$, and on $\operatorname{span}\{s\}$ its eigenvalue is $A\varepsilon/(r+\varepsilon)^2$. Therefore (46) is convex exactly when $A \geq 0$.

It remains to justify the subgradient formulation rather than assume it. Suppose a finite convex N satisfies $q(s) \in \partial N(s)$ everywhere. Because $q \in C^1$, it has a Lipschitz constant L on a compact neighborhood of each fixed s. For every sufficiently small increment u, the subgradient inequalities at the two endpoints give

$$0 \leq N(s+u) - N(s) - q(s)^\top u \leq (q(s+u) - q(s))^\top u \leq L\|u\|_2^2.$$

Dividing the remainder by $\|u\|_2$ directly yields Fréchet differentiability and $\nabla N = q$. Since $q \in C^1$, the function N is C^2 and the preceding Jacobian and Hessian arguments apply. Finally, two differentiable functions with the same gradient on connected Ω_d differ by a constant. The converse follows from the convex potential already constructed. □

Corollary 17 (Natural full-matrix lift)**.** *On the positive full-rank stratum, the affine case lifts to the spectral potential*

$$\mathcal{F}_\varepsilon(G) = A\left[\|G\|_F - \varepsilon \log\left(\frac{\|G\|_F + \varepsilon}{\varepsilon}\right)\right] + B\|G\|_*. \tag{52}$$

It is finite and convex on the whole matrix space when $A \geq 0$ and $B \geq 0$; these signs are necessary for this natural lift. A choice-independent continuous gradient extension through arbitrary rank loss forces $B = 0$.

Proof. On the positive full-rank stratum, the standard differential formulas $\nabla\|G\|_F = G/\|G\|_F$ and $\nabla\|G\|_* = UV^\top$ give exactly the lifted response. For $A \geq 0$, the scalar radial function in (52) is convex and nondecreasing, so its composition with the Frobenius norm is convex; for $B \geq 0$, the nuclear term is convex. Necessity of $A \geq 0$ follows by restriction to the positive full-rank stratum and (51). On the line $G = tuv^\top$, the radial term has derivative zero at $t = 0$, whereas the nuclear term is $B|t|$. Convexity requires the left derivative $-B$ not to exceed the right derivative B, hence $B \geq 0$. If $B \neq 0$, approaching a rank-deficient matrix while rotating the vanishing singular directions changes the limiting $BUV^\top$ term; hence no choice-independent continuous gradient extension exists. □

B.2 The hard-clamped Frobenius denominator

Let $h \in C^1([0,1])$, set $r = \|s\|_2$, and define

$$q_i(s) = h\left(\frac{s_i}{\max\{r, \varepsilon\}}\right). \tag{53}$$

Write $\Omega_- = \{s : r < \varepsilon\}$, $\Omega_+ = \{s : r > \varepsilon\}$, and $\Gamma = \{s : r = \varepsilon\}$. Let

$$\Psi(t) = \int_0^t h(u)\, du. \tag{54}$$

Theorem 18 (Hard-clamp classification and gluing)**.** *The field (53) has a C^1 potential on Ω_d if and only if, for every positive unit vector x,*

$$h'(x_i) = h'(x_j) \qquad (1 \le i, j \le d). \tag{55}$$

For $d \ge 3$, this is equivalent to $h(t) = At + B$ on $(0,1)$. For $d = 2$, it is equivalent to the reflection identity

$$h'(t) = h'\left(\sqrt{1-t^2}\right) \qquad (0 < t < 1). \tag{56}$$

Under these equivalent conditions the potential is convex if and only if $h'(t) \ge 0$ on $(0,1)$. The same condition is necessary and sufficient for a finite convex N satisfying $q(s) \in \partial N(s)$ everywhere.

Proof. Inside the clamp sphere the denominator is constant and

$$F_-(s) = \varepsilon \sum_{i=1}^{d} \Psi(s_i/\varepsilon) \tag{57}$$

satisfies $\nabla F_- = q$. Thus the inner region is conservative for every h.

On Ω_+, put $x = s/r$. The quotient rule gives

$$\frac{\partial x_i}{\partial s_j} = \frac{1}{r}(\delta_{ij} - x_i x_j), \tag{58}$$

and hence

$$\frac{\partial q_i}{\partial s_j} = \frac{h'(x_i)}{r}(\delta_{ij} - x_i x_j). \tag{59}$$

For $i \ne j$, all of x_i, x_j, r are positive, so symmetry of the outer Jacobian is exactly (55). Necessity now follows from equality of mixed partials for an existing C^1 potential, which is C^2 on Ω_+ because q is C^1 there.

Suppose first that $d \ge 3$. Given arbitrary $u, v \in (0,1)$, choose $\tau > 0$ such that $u^2 + \tau^2 < 1$ and $v^2 + \tau^2 < 1$. Complete (u, τ) and (v, τ) to positive unit vectors by distributing the remaining squared mass among the other $d - 2$ coordinates. Condition (55) gives $h'(u) = h'(\tau) = h'(v)$. Thus h' is constant, which is equivalent to the stated affine form; conversely an affine response plainly satisfies the condition.

If $d = 2$, every positive unit vector is uniquely $(t, \sqrt{1-t^2})$ for $0 < t < 1$. Substitution into (55) gives (56), and every positive unit vector has been covered. This proves both necessity and sufficiency of the reflection identity at the differential level. It also exhibits genuine nonaffine examples: for instance $h'(t) = 1 + t^2(1-t^2)$ is invariant under the reflection.

We next construct the global potential, so no simple-connectivity assertion about the outer region is needed. Assume (55). For a positive unit vector x set

$$K(x) = \sum_i \big[\Psi(x_i) - x_i h(x_i)\big]. \tag{60}$$

If v is tangent to the unit sphere at x, then $x^\top v = 0$. The directional derivative is

$$DK(x)[v] = -\sum_i x_i h'(x_i) v_i.$$

Condition (55) makes every $h'(x_i)$ one common scalar $c(x)$. Consequently $DK(x)[v] = -c(x)x^\top v = 0$. The positive unit sphere is path connected for $d \geq 2$; therefore K is one constant κ_h.

Define

$$F(s) = \begin{cases} \varepsilon \sum_i \Psi(s_i/\varepsilon), & r \leq \varepsilon, \\ \sum_i s_i h(s_i/r) + \varepsilon\kappa_h, & r \geq \varepsilon. \end{cases} \tag{61}$$

On Ω_+, let $F_0(s) = s^\top q(s)$. Zero-homogeneity gives $Dq(s)s = 0$. The sphere condition and (59) make $Dq(s)$ symmetric, hence $Dq(s)^\top s = 0$. Differentiating F_0 gives $\nabla F_0 = q + Dq(s)^\top s = q$.

At $s = \varepsilon x \in \Gamma$, the inner and outer values in (61) coincide because

$$\varepsilon \sum_i x_i h(x_i) + \varepsilon\kappa_h = \varepsilon \sum_i \Psi(x_i).$$

Both one-sided gradients tend to $(h(x_1), \ldots, h(x_d))$. More explicitly, each piece has, on its own side, $F(s+z) = F(s) + q(s)^\top z + o(\|z\|_2)$; splitting an arbitrary sequence into the two sides gives the same expansion without a side restriction. Continuity of q then proves that $F \in C^1(\Omega_d)$ and $\nabla F = q$. This completes sufficiency and the gluing proof.

It remains to classify convexity. In the inner region,

$$\nabla^2 F_-(s) = \frac{1}{\varepsilon} \operatorname{diag}(h'(s_1/\varepsilon), \ldots, h'(s_d/\varepsilon)). \tag{62}$$

For every $t \in (0,1)$ there is an inner point having $s_1/\varepsilon = t$ and all other positive coordinates sufficiently small. Thus convexity forces $h'(t) \geq 0$. Conversely, under this sign condition the inner Hessian is positive semidefinite. On the outer region, the sphere condition turns (59) into

$$\nabla^2 F_+(s) = \frac{c(x)}{r}(I - xx^\top) \succeq 0. \tag{63}$$

For a line segment in Ω_d, the sphere $r = \varepsilon$ is crossed at at most two points. Away from those points the restriction of F has nonnegative second derivative; at a crossing its first derivative is continuous because $F \in C^1$. Hence the first derivative along the whole segment is nondecreasing. The restriction is convex, and convexity on every line segment proves convexity on Ω_d.

Finally, if a finite convex N selects q as a subgradient, the same two-endpoint subgradient squeeze used in theorem 16 gives $\nabla N = q$. The conservative and convexity classifications just proved apply, and $N - F$ is constant. The converse is immediate from the convex C^1 potential. □

Corollary 19 (Affine Huber splice)**.** *For $h(t) = At + B$, the potential (61) is*

$$F(s) = A\rho_\varepsilon(\|s\|_2) + B\sum_i s_i, \qquad \rho_\varepsilon(r) = \begin{cases} r^2/(2\varepsilon), & 0 \leq r \leq \varepsilon, \\ r - \varepsilon/2, & r \geq \varepsilon. \end{cases} \tag{64}$$

In particular, $\nabla F = As/\max\{\|s\|_2, \varepsilon\} + B\mathbf{1}$.

Proof. Choose $\Psi(t) = At^2/2 + Bt$. For a unit vector x,

$$\kappa_h = \sum_i \left(\frac{A}{2}x_i^2 + Bx_i - x_i(Ax_i + B) \right) = -\frac{A}{2}\sum_i x_i^2 = -\frac{A}{2}.$$

Substitution into the two pieces of (61) gives (64); direct differentiation gives the stated gradient. This also shows explicitly that the potential is C^1 at the clamp radius, although its radial second derivative generally jumps there. □

Corollary 20 (Rank-two absolutely symmetric matrix lift)**.** *Let $d = 2$ and suppose the reflection identity (56) and $h' \geq 0$ hold. The natural extension of F in (61) to the closed positive cone induces a finite convex absolutely symmetric spectral potential on the full matrix space if and only if*

$$h(0) \geq 0. \tag{65}$$

When $h(0) > 0$ the lift is generally nonsmooth at rank loss; this does not affect the convex-subgradient statement.

Proof. The reflection identity makes F permutation invariant, and $h' \geq 0$ together with (65) gives $h(t) \geq h(0) \geq 0$ on $[0,1]$. Hence the closed-cone extension of F is convex and coordinatewise nondecreasing. For vectors $z, w \in \mathbb{R}^2$, coordinatewise $|\lambda z + (1-\lambda)w| \leq \lambda|z| + (1-\lambda)|w|$; monotonicity and convexity of F therefore prove convexity of $\bar{F}(z) = F(|z_1|, |z_2|)$. The standard convex spectral-function theorem then makes $G \mapsto \bar{F}(\sigma(G))$ convex. Conversely, restriction of an absolutely symmetric convex extension to a coordinate line requires its right derivative at the origin, $h(0)$, to be nonnegative. □

B.3 From a continuous convex selection to an ordinary gradient

We now pass to the hypotheses used by the no-gluing result. Assume $h \in C^1([0,\infty))$ and $\alpha : \Omega_d \to (0,\infty)$ is positive, one-homogeneous, and locally Lipschitz. Suppose there is a finite convex function $N : \Omega_d \to \mathbb{R}$ such that

$$q(s) \in \partial N(s) \qquad (s \in \Omega_d). \tag{66}$$

Lemma 21 (A continuous convex subgradient selection is the gradient)**.** *Under these hypotheses, N is differentiable, $\nabla N = q$, and $N \in C^{1,1}_{\mathrm{loc}}(\Omega_d)$.*

Proof. Fix s, v and $t > 0$ small enough that $s + tv \in \Omega_d$. Applying the subgradient inequality at s and then at $s + tv$ gives

$$q(s)^\top (tv) \leq N(s+tv) - N(s) \leq q(s+tv)^\top (tv).$$

Division by $t > 0$ gives

$$q(s)^\top v \leq \frac{N(s+tv) - N(s)}{t} \leq q(s+tv)^\top v. \tag{67}$$

The field q is locally Lipschitz: h is C^1, α is locally Lipschitz and strictly positive, and composition and division preserve local Lipschitzness on compact subsets on which the denominator stays positive. Fix a compact neighborhood of s on which q has Lipschitz constant L. For every sufficiently small increment u, the same two subgradient inequalities give the explicit remainder bound

$$0 \leq N(s+u) - N(s) - q(s)^\top u \leq (q(s+u) - q(s))^\top u \leq L\|u\|_2^2. \tag{68}$$

After division by $\|u\|_2$, the remainder tends to zero uniformly in the direction of u. Thus N is Fréchet differentiable with $\nabla N = q$. Since q is locally Lipschitz, $N \in C^{1,1}_{\mathrm{loc}}$. □

B.4 Direct principal-branch necessity without critical-set assumptions

Assume $d \geq 2$, $h \in C^1([0,\infty))$, $h' > 0$ near zero, and $h' < 0$ somewhere. Let α be positive, one-homogeneous, locally Lipschitz on Ω_d, and continuously and strictly positively extended to the nonzero closed positive cone. Suppose $q \in \partial N$ everywhere for a finite convex N. Set $\zeta = \inf\{t > 0 : h'(t) < 0\}$ and $\Phi(t) = \int_0^t u h'(u)\, du$. No assumption about the number or type of critical points is made here.

Theorem 22 (Direct principal-branch necessity)**.** *Under these hypotheses, every interior normalized coordinate belongs to $(0, \zeta)$, and a single constant C satisfies*

$$\sum_i \Phi(s_i/\alpha(s)) = C, \qquad 0 < C \leq \Phi(\zeta).$$

Moreover $N \in C^{1,1}_{\mathrm{loc}}$, $\nabla N = q$, $h'(x_i) \geq 0$, and $\Delta \geq 0$, with $\Delta = 0$ if and only if all $h'(x_i) = 0$. If every positive critical point is a strict extremum, then $h'(x_i) > 0$ and $\Delta > 0$. The one-homogeneous potential $\mathcal{N}(s) = \sum_i s_i h(s_i/\alpha(s))$ differs from N by a constant.

Proof. First-fold exclusion. Fix an index i, fix all other $d-1$ coordinates positively, and let $s(t)$ denote the resulting coordinate line. For any $T > 0$ put $z(t) = t/\alpha(s(t))$ on $[0,T]$. Since $d \geq 2$, $s(0) \neq 0$; the assumed boundary extension gives continuity and $z(0) = 0$. For $0 < t_1 < t_2$, monotonicity of a convex subgradient implies

$$0 \leq \langle q(s(t_2)) - q(s(t_1)), s(t_2) - s(t_1)\rangle = (t_2 - t_1)\big[h(z(t_2)) - h(z(t_1))\big].$$

For any $0 < u < v \le z(T)$, the first hitting time of v exists and is positive. Continuity between zero and that time supplies an earlier positive time at which $z = u$. Thus $h(u) \le h(v)$. By continuity this also holds at $u = 0$, so h is nondecreasing on $[0, z(T)]$. If $z(T) > \zeta$, the definition of the infimum supplies a point strictly between 0 and $z(T)$ where $h' < 0$, a contradiction. Hence every attained coordinate is at most ζ. This argument requires no monotonicity or injectivity of z, no permutation symmetry, and no homogeneity of α.

Global modular constant without division by Δ. At an ambient differentiability point of α, put $a = \alpha(s)$ and $p = \nabla\alpha(s)$. The quotient rule gives $J_q = (D - wp^\top)/a$. Monotonicity implies $B = (J_q + J_q^\top)/2 \succeq 0$, while homogeneity implies $J_q x = 0$. Therefore $x^\top Bx = 0$, hence $Bx = 0$ and $J_q^\top x = 0$. It follows that $w = \Delta p$, including at $\Delta = 0$. The locally Lipschitz function $R(s) = \sum_i \Phi(s_i/\alpha(s))$ has

$$\nabla R(s) = \frac{w - \Delta p}{a} = 0 \quad \text{almost everywhere on } \Omega_d.$$

The zero weak-gradient theorem on a connected open set makes $R = C$ everywhere. No derivative on the unit surface and no deletion of critical spectra is used.

Strict interior confinement. On $[0, \zeta]$, $h' \ge 0$; because $h' > 0$ near zero, $\Phi(t) > 0$ for every $t \in (0, \zeta]$. Taking $s \to e_j$ through positive spectra gives $C = \Phi(1/\alpha(e_j)) \in (0, \Phi(\zeta)]$. An interior $x_i = \zeta$ would instead give $C = \Phi(\zeta) + \sum_{k \ne i} \Phi(x_k) > \Phi(\zeta)$. Thus all interior $x_i < \zeta$. Nonnegative summands now give the stated characterization of $\Delta = 0$. A nondecreasing function cannot have a strict local extremum inside $(0, \zeta)$, so the strict-extremum hypothesis, when present, gives $h' > 0$ there.

Finally lemma 21 gives $\nabla N = q$ and the claimed regularity. Differentiating $\mathcal{N}$ almost everywhere gives $\nabla\mathcal{N} = q + w - \Delta p = q$; consequently $\mathcal{N} - N$ is constant by the same weak-gradient argument. □

Necessity versus existence. The theorem does not assert that every modular level has a unique regular radial root when h' has flat intervals. Strict pre-fold positivity, used in the full-matrix sufficiency theorem below, gives that additional property. Finite criticality is sufficient to make Φ strictly increasing on its principal interval, but is unnecessary for the necessity theorem in any dimension $d \ge 2$. The argument fails in $d = 1$ because the coordinate-line endpoint is the excluded zero vector.

Order preservation alone is weaker. For $h(t) = t - t^3$ and $\alpha(s) = \frac{5}{4}\sum_i s_i$, every normalized pair satisfies $u + v \le 4/5$, and $h(u) - h(v) = (u - v)(1 - u^2 - uv - v^2)$ preserves its order. Nevertheless attained coordinates can exceed $1/\sqrt{3}$. On one coordinate line in dimension two, $q_1(3, 1) = 48/125 > 357/1000 = q_1(7, 1)$, contradicting monotonicity. Thus the first-hitting proof uses comparisons between different spectra, not merely equivariant order at one spectrum.

B.5 Strictness and radial uniqueness

Corollary 23 (Strict-extremum specialization)**.** *Under the hypotheses of theorem 22, if every positive critical point is a strict extremum, then $h' > 0$ on $(0, \zeta)$, $w \ne 0$, and $\Delta > 0$. The attained principal radial root is unique.*

Proof. Only uniqueness remains to prove. For fixed $s > 0$, whenever all $s_i/a < \zeta$,

$$\frac{d}{da}\sum_i \Phi(s_i/a) = -\frac{1}{a}\sum_i (s_i/a)^2 h'(s_i/a) < 0.$$

Thus there is at most one principal root. □

Corollary 24 (Finite-critical specialization in every dimension)**.** *Under the hypotheses of theorem 22, suppose $Z = \{t > 0 : h'(t) = 0\}$ is finite. Then $A = S \cap Z^d$ is finite, $\Delta > 0$ on $S \setminus A$, and the attained principal radial root is unique in every $d \ge 2$. Points of A may have $\Delta = w = 0$.*

Proof. Nonnegative summands give $\Delta = 0$ exactly on A. On $(0, \zeta)$, $h' \ge 0$ with finitely many zeros, so Φ is strictly increasing. Hence $a \mapsto \sum_i \Phi(s_i/a)$ is strictly decreasing on its principal interval even if its derivative vanishes at isolated values. □

B.6 Classical inertia identity and a local post-fold branch

This classical identity explains local Hessians; it is not needed for global principal-branch necessity.

Lemma 25 (Rank-one inertia without invertibility)**.** *Let D be real symmetric, $\Delta = x^\top Dx \neq 0$, $w = Dx$, and $M = D - ww^\top/\Delta$. Then*

$$\text{inertia}(M) = \begin{cases} (n_+(D) - 1, n_-(D), n_0(D) + 1), & \Delta > 0, \\ (n_+(D), n_-(D) - 1, n_0(D) + 1), & \Delta < 0. \end{cases}$$

No invertibility of D is required.

Proof. Let Z span $\{z : w^\top z = 0\}$. Since $w^\top x = \Delta \neq 0$, $P = [Z \ x]$ is invertible. Direct multiplication gives

$$P^\top DP = \text{diag}(Z^\top DZ, \Delta), \qquad P^\top MP = \text{diag}(Z^\top DZ, 0).$$

Sylvester's law removes the sign of Δ and adds one zero, preserving any existing null directions. □

Local convexity does not supply the boundary hypothesis. For $h(t) = 2t - t^3$, take $\Phi(t) = t^2 - 3t^4/4$, $x^0 = (1, 1/10)$, and $C = 10397/40000$. The equation $\sum_i \Phi(s_i/\alpha(s)) = C$ has a smooth local radial branch with $\alpha(x^0) = 1$, since $\Delta(x^0) = -9803/10000 \neq 0$. At this point its compatibility Hessian factor is exactly

$$M = \frac{197}{9803} \begin{pmatrix} 1 & -10 \\ -10 & 100 \end{pmatrix} \succeq 0.$$

In a sufficiently small neighborhood the diagonal slopes have one negative and one positive entry and $\Delta < 0$. The inertia identity keeps the Hessian positive semidefinite there, so on a convex neighborhood $\mathcal{N}(s) = \sum_i s_i h(s_i/\alpha(s))$ is a convex potential. The response at x^0 is $(1, 199/1000)$ and preserves coordinate order, yet $x_1^0 = 1 > \zeta = \sqrt{2/3}$. This branch cannot extend to a whole-cone convex field with the stated positive boundary extension. Its obstruction is the elementary coordinate-line argument, not a new topological principle.

C Filter-generic full-matrix classification and projective frontier

This section proves the two filter-generic theorems stated in the main text. Let $h \in C^1([0, \infty))$ satisfy $h(0) = 0$, with $h' > 0$ near zero and $h' < 0$ somewhere. Assume the positive critical set is finite and every critical point is a strict local extremum. No convex realization is assumed in the construction below. Write

$$\zeta_h = \inf\{t > 0 : h'(t) < 0\}, \qquad \Phi_h(t) = \int_0^t uh'(u)\, du, \qquad C_h = \Phi_h(\zeta_h). \tag{69}$$

The strict-extremum assumption and positivity of h' near zero imply $h'(t) > 0$ on $(0, \zeta_h)$ and $h'(\zeta_h) = 0$. Hence h and Φ_h are strictly increasing there, with $h(t), \Phi_h(t) > 0$ for $t > 0$.

Theorem 26 (Filter-generic full-matrix classification)**.** *Fix $d \geq 2$. For every $C \in (0, C_h]$ and every $s \in \Omega_d$, there is a unique principal root $\alpha_C(s)$ satisfying*

$$\sum_i \Phi_h(s_i/\alpha_C(s)) = C, \qquad 0 < s_i/\alpha_C(s) < \zeta_h. \tag{70}$$

It is C^1, positive, permutation invariant, and one-homogeneous on Ω_d, with a strictly positive continuous extension to the nonzero closed positive cone. The function

$$\mathcal{N}_C(s) = \sum_i s_i h(s_i/\alpha_C(s)) \tag{71}$$

extends to a coordinatewise nondecreasing symmetric-gauge norm on the closed positive cone. Consequently $\Psi_C(G) = \mathcal{N}_C(\sigma(G))$ is a unitarily invariant norm and

$$Q_C(G) = U \,\text{diag}\big(h(\sigma_i/\alpha_C(\sigma))\big)V^\top \in \partial\Psi_C(G) \tag{72}$$

for every SVD of G, with $Q_C(0) = 0$. At $C = C_h$, the only loss of modular regularity on the nonzero closed cone is at rank-one rays.

Conversely, let α be permutation invariant and satisfy the positivity, homogeneity, local-Lipschitz, and positive-boundary hypotheses above. If its aligned field is a subgradient selection of any finite convex full-matrix function, then $\alpha = \alpha_C$ on Ω_d for one unique $C \in (0, C_h]$.

Proof. For s in the nonzero closed positive cone, put $m = \max_i s_i$, $\bar{a} = m/\zeta_h$, and $R_s(a) = \sum_i \Phi_h(s_i/a)$ for $a \geq \bar{a}$. At $\bar{a}$ at least one coordinate equals ζ_h, so $R_s(\bar{a}) \geq C_h$, with strict inequality when at least two coordinates are positive. Also $R_s(a) \to 0$ as $a \to \infty$, and for $a > \bar{a}$,

$$R_s'(a) = -\frac{1}{a} \sum_i (s_i/a)^2 h'(s_i/a) < 0. \tag{73}$$

Thus the principal root exists uniquely. It is strictly larger than $\bar{a}$ except at $C = C_h$ on a rank-one ray. On the positive cone the implicit-function theorem, uniqueness, and permutation symmetry give the claimed C^1 regularity, homogeneity, and symmetry. Higher regularity follows only when the scalar filter has the corresponding higher regularity.

The root also varies continuously on the nonzero closed cone. By homogeneity it suffices to use the compact simplex $\sum_i s_i = 1$. Choose $\delta \in (0, \zeta_h)$ with $d\Phi_h(\delta) < C$. Every root lies between m/ζ_h and m/δ; subsequential limits solve the limiting radial equation, whose root is unique. This proves continuity and strict boundary positivity. At a maximum-level rank-one ray, its sole positive normalized coordinate is ζ_h, so $\Delta = \zeta_h^2 h'(\zeta_h) = 0$; all other nonzero closed-cone spectra are regular.

In the positive interior set $x = s/\alpha_C(s)$, $D = \operatorname{diag}(h'(x_i))$, and $\Delta = x^\top D x > 0$. Implicit differentiation gives

$$\nabla \alpha_C(s) = \frac{Dx}{\Delta}, \qquad \nabla \mathcal{N}_C(s) = \big(h(x_i)\big)_i, \qquad \nabla^2 \mathcal{N}_C(s) = \frac{1}{\alpha_C(s)} \left(D - \frac{Dxx^\top D}{\Delta} \right). \tag{74}$$

Weighted Cauchy–Schwarz makes the final matrix positive semidefinite, so $\mathcal{N}_C$ is convex. Continuity passes convexity to the boundary. Because $h(0) = 0$ and $h' > 0$ on the attained interior band, the gradient is componentwise positive. Hence the boundary extension is coordinatewise nondecreasing, symmetric, one-homogeneous, convex, and positive away from the origin. For $z, w \in \mathbb{R}^d$, $|\lambda z + (1-\lambda)w| \leq \lambda|z| + (1-\lambda)|w|$ coordinatewise; applying monotonicity and convexity shows that $g_C(z) = \mathcal{N}_C(|z|)$ is a symmetric-gauge norm.

At boundary spectra, approximate by $s + \varepsilon \mathbf{1}$ and pass the interior subgradient inequality to the limit. The response has the same coordinate order as s, equal singular values receive equal responses, and zero singular values receive $h(0) = 0$. Therefore (72) is SVD independent. Von Neumann's trace inequality and the vector subgradient inequality give, for every Z,

$$\langle Q_C(G), Z \rangle_F \leq q_C(\sigma(G))^\top \sigma(Z) \leq \mathcal{N}_C(\sigma(Z)) = \Psi_C(Z),$$

with equality at $Z = G$. This proves the matrix subgradient and dual-ball support-oracle claims.

For necessity, restrict any assumed full-matrix subgradient inequality to positive rectangular diagonal matrices. The resulting cone field satisfies corollary 23, which forces one $C \in (0, C_h]$ and the principal modular equation. Root uniqueness gives $\alpha = \alpha_C$. The modular level is unique because two different levels cannot hold at the same normalized spectrum. □

Classical norm representation and what is inverse. The sufficiency construction has a classical Amemiya/Orlicz interpretation (Hudzik & Maligranda, 2000). Extend the principal restriction of h by the constant $h(\zeta_h)$ beyond ζ_h, call this nondecreasing extension $\widetilde{h}$, and put $J(t) = \int_0^t \widetilde{h}(u)\,du$ for $t \geq 0$. Then J is convex. If $J^*(v) = \sup_{t \geq 0}(tv - J(t))$, Fenchel equality gives $J^*(h(x)) = xh(x) - J(x) = \Phi_h(x)$ on the principal interval. Consequently

$$\mathcal{N}_C(s) = \inf_{a > 0} a \left[C + \sum_i J(|s_i|/a) \right] = \sup_{\sum_i J^*(|v_i|) \leq C} \langle s, v \rangle. \tag{75}$$

Indeed, Fenchel's inequality bounds the supremum by every candidate in the infimum. For $s \neq 0$, at $a = \alpha_C(|s|)$ and the signed aligned principal response, all inequalities become equalities and

$\sum_i J^*(|v_i|) = C$. At $s = 0$ both sides are zero directly. This also handles the rank-one endpoint, without differentiating its normalizer. Thus the potential is C times the Amemiya norm generated by J/C, and its forward construction is not new. In contrast, the primitive Φ_h defining the normalizer can be nonconvex. The no-gluing argument is needed to show that the prescribed nonmonotone filter admits no other admissible global convex realization.

For a nonzero nonnegative vector v, let $\kappa_+(v)$ denote its largest positive coordinate divided by its smallest positive coordinate, with value one on rank-one support.

Theorem 27 (Sharp filter-generic projective repair frontier)**.** *In addition to theorem 26, assume h is strictly concave on $[0, \zeta_h]$ and $E_h(t) = th'(t)/h(t)$ decreases strictly on $(0, \zeta_h)$. For $K > 1$ and $C \in (0, C_h]$, let $\tau_{h,K,C} \in (0, \zeta_h)$ be the unique root*

$$\Phi_h(\tau_{h,K,C}) + \Phi_h(\tau_{h,K,C}/K) = C \tag{76}$$

and define

$$\rho_{h,C}(K) = \frac{h(\tau_{h,K,C})}{h(\tau_{h,K,C}/K)}, \qquad \rho_h(K) = \rho_{h,C_h}(K), \qquad \eta_h = \frac{h(\zeta_h)}{\zeta_h h'(0)}. \tag{77}$$

Set both ρ quantities to one at $K = 1$. Then every classified response and nonzero nonnegative spectrum s, with $K = \kappa_+(s)$, obey

$$\boxed{\rho_{h,C}(K) \le \kappa_+(q_C(s)) \le K}. \tag{78}$$

The lower bound is attained by a rank-two spectrum. For fixed $K > 1$ it decreases strictly with C, and

$$\boxed{\eta_h K < \rho_h(K) \le \kappa_+(q_C(s)) \le K}, \qquad \lim_{K\to\infty} \frac{\rho_h(K)}{K} = \eta_h. \tag{79}$$

The ratio assertions remain true after positive rescaling of these classified q_C. They do not cover simultaneous changes to the normalizer and a state-dependent multiplier.

Proof. Let $x = s/\alpha_C(s)$ and $t = \max_i x_i$. The smallest positive coordinate is t/K. If $K = 1$, all positive input and response coordinates are equal, and the ratio bounds hold directly. Assume henceforth $K > 1$, so the largest and smallest coordinates occur at distinct indices. Nonnegativity of the remaining modular terms gives $\Phi_h(t) + \Phi_h(t/K) \le C$, hence $t \le \tau_{h,K,C}$. With $R_K(t) = h(t)/h(t/K)$,

$$\frac{d\log R_K(t)}{d\log t} = E_h(t) - E_h(t/K) < 0. \tag{80}$$

Therefore $\kappa_+(q_C(s)) = R_K(t) \ge \rho_{h,C}(K)$. Rank two with normalized coordinates $(\tau_{h,K,C}, \tau_{h,K,C}/K)$ attains equality. Strict concavity and $h(0) = 0$ make $r(t) = h(t)/t$ strictly decreasing, so $R_K(t) = Kr(t)/r(t/K) \le K$.

Increasing C increases $\tau_{h,K,C}$, and (80) decreases the response ratio. At $C = C_h$ and finite $K > 1$, both arguments are interior, so

$$\frac{\rho_h(K)}{K} = \frac{r(\tau_{h,K,C_h})}{r(\tau_{h,K,C_h}/K)} > \frac{r(\zeta_h)}{r(0)} = \eta_h.$$

As $K \to \infty$, the root equation forces $\tau_{h,K,C_h} \to \zeta_h$ and $\tau_{h,K,C_h}/K \to 0$, proving the limit. Positive radial rescaling changes no coordinate ratio. □

Corollary 28 (Analytic odd one-fold family)**.** *For $h_p(t) = at - bt^p$ with $a, b > 0$ and odd integer $p \ge 3$,*

$$\zeta_{h_p} = \left(\frac{a}{bp}\right)^{1/(p-1)}, \qquad C_{h_p} = \frac{a\zeta_{h_p}^2(p-1)}{2(p+1)}, \qquad \eta_{h_p} = 1 - \frac{1}{p}. \tag{81}$$

Thus theorems 26 and 27 give an exact full-matrix classification and sharp projective frontier for every member of this family.

Proof. The unique positive critical point solves $a - bpt^{p-1} = 0$ and is a strict maximum. On the principal interval, $h_p''(t) = -bp(p-1)t^{p-2} < 0$. With $y = t^{p-1}$,

$$E_{h_p}(t) = \frac{a - bpy}{a - by}, \qquad \frac{dE_{h_p}}{dy} = -\frac{ab(p-1)}{(a-by)^2} < 0.$$

The three constants follow by direct substitution and integration. □

D Analytic directions modulo positive output scaling

We now keep the Frobenius input normalizer fixed but allow an arbitrary positive scalar output multiplier. This is the relevant equivalence when a scalar can be absorbed into a step size. The generic integrating-factor condition is classical; Jacobi integrability and tangent-curvature invariance under positive scalar multiplication also arise in inverse-demand rationalization (Hosoya, 2022, Lemma 3). The result below is the filter-specific algebraic classification used here.

Remark 29 (A limit of the direct demand identification)**.** Fally's single-aggregator direct-demand hypothesis (Fally, 2022, Assumption A1 and Proposition 1) is not automatic after budget normalization of our field. Take $d = 4$, $h(t) = t + t^3$, and $q_i(s) = h(s_i/\|s\|_2)$. Put $R = \sum_i s_i^2$, $T = R^2 + \sum_i s_i^4$, and $n_i = s_i(R + s_i^2)$. Then $B_i = q_i/(s^\top q) = n_i/T$ satisfies $s^\top B = 1$. Any local C^1 representation $B_i = A_i(s_i, L(s))$ has off-diagonal derivatives $J_{ij} = (\partial_2 A_i)\partial_j L$, hence $J_{12}J_{34} - J_{14}J_{32} = 0$. Direct differentiation instead gives

$$J_{ij} = \frac{(R + 3s_i^2)\delta_{ij} + 2s_i s_j}{T} - \frac{4n_i n_j}{T^2}, \qquad (J_{12}J_{34} - J_{14}J_{32})\big|_{s=(1,2,3,4)} = -\frac{256}{27387987} \neq 0.$$

Thus this budget-normalized example has no such smooth scalar aggregator when s is identified with normalized prices. This only excludes the direct identification, not every change of variables or reduction to demand theory, and does not assert that B itself is rational demand.

Let $r = \|s\|_2$, $x = s/r$, $q_i(s) = h(x_i)$, and $\omega = \sum_i q_i\, ds_i$ on Ω_d. A direct calculation gives

$$J_q = \frac{1}{r}\{D - wx^\top\}, \qquad D = \operatorname{diag}(h'(x_i)), \quad w = x \odot h'(x). \tag{82}$$

For any positive C^1 scalar b,

$$(b\omega) \wedge d(b\omega) = b^2 \omega \wedge d\omega. \tag{83}$$

Thus the vanishing of the three-form is invariant under even a nonhomogeneous output multiplier. When q is zero-homogeneous and $s^\top q > 0$, the stronger classical statement is that a positive one-homogeneous potential exists locally exactly when $q^\top ds/(s^\top q)$ is closed; convexity remains an additional tangent-curvature condition. We use only (83) for necessity below.

Theorem 30 (Analytic projective rigidity)**.** *Let $d \geq 4$ and let h be real analytic and odd on $(-1, 1)$. There exist $b \in C^1(\Omega_d)$, $b > 0$, and $V \in C^2(\Omega_d)$ such that*

$$b(s)h(s/\|s\|_2) = \nabla V(s) \tag{84}$$

if and only if, on $(-1, 1)$,

$$h(t) = at + ct^{2k+1} \tag{85}$$

for some integer $k \geq 1$ and $a, c \in \mathbb{R}$; the linear and zero functions are included by $c = 0$. Neither convexity nor homogeneity of b or V is assumed. The same equivalence holds for odd polynomials already when $d \geq 3$.

Proof. If (84) holds, then $d(b\omega) = 0$, so (83) gives $\omega \wedge d\omega = 0$. For $i < j < k$, (82) gives the coefficient identity

$$[\omega \wedge d\omega]_{ijk} = -\frac{1}{r} \det \begin{pmatrix} q_i & x_i & w_i \\ q_j & x_j & w_j \\ q_k & x_k & w_k \end{pmatrix}. \tag{86}$$

Thus every three-coordinate minor satisfies

$$\det[\, h(x_I), x_I, x_I \odot h'(x_I)\,] = 0. \tag{87}$$

Odd real analyticity gives $h(t) = tP(t^2)$ for a real-analytic P on $[0, 1)$. Put $y_i = x_i^2$. Since $h'(x_i) = P(y_i) + 2y_i P'(y_i)$, factoring $x_1 x_2 x_3 > 0$ from the rows, subtracting the first column from the third, and dividing by two turns (87) into

$$\det \begin{pmatrix} P(y_1) & 1 & y_1 P'(y_1) \\ P(y_2) & 1 & y_2 P'(y_2) \\ P(y_3) & 1 & y_3 P'(y_3) \end{pmatrix} = 0. \tag{88}$$

Every positive triple with $y_1 + y_2 + y_3 < 1$ can be completed to a positive unit vector using a fourth coordinate (and by splitting the remaining mass when $d > 4$). Hence (88) holds on that open set.

If P is nonconstant, analyticity lets us choose sufficiently small $u, v > 0$ with $P(u) \neq P(v)$. Equation (88) says that every point $(P(y), yP'(y))$ for all sufficiently small $y > 0$ lies on the line through the corresponding points at u, v. Thus $yP'(y) = AP(y) + B$ near zero, and the analytic identity theorem extends this relation throughout $(0, 1)$. Writing $P(y) = \sum_{j\geq 0} c_j y^j$ and comparing coefficients gives $(j - A)c_j = 0$ for every $j \geq 1$. Therefore at most one c_j, $j \geq 1$, is nonzero. If P is constant the same conclusion is immediate. This proves necessity for $d \geq 4$.

The polynomial case in dimension three. Now let $d = 3$ and let P be a polynomial. Denote the determinant in (88) by $D(u, v, w)$ and set $e_1 = u+v+w$. It vanishes on the positive simplex $e_1 = 1$. Thus $D(u, v, 1 - u - v)$ is a bivariate polynomial vanishing on an open triangle, so it is identically zero. Division by the monic polynomial $w + u + v - 1$ gives $D = (e_1 - 1)K$ for a polynomial K.

For integers $1 \leq m < n$, define

$$A_{m,n}(u, v, w) = \det \begin{pmatrix} 1 & u^m & u^n \\ 1 & v^m & v^n \\ 1 & w^m & w^n \end{pmatrix}, \qquad V_0 = (v - u)(w - u)(w - v).$$

We first show that e_1 divides $A_{m,n}$ only for $(m, n) = (1, 3)$. Such divisibility implies $F(t) = A_{m,n}(1, t, -1 - t) \equiv 0$. But $F(0) = (-1)^m - (-1)^n$, so m, n have the same parity, say $\varepsilon = (-1)^m = (-1)^n$. Direct differentiation gives

$$F'(0) = \mathbf{1}_{\{m=1\}}(\varepsilon - 1) + \varepsilon(m - n).$$

This cannot vanish for $m > 1$; for $m = 1$ it equals $n - 3$. Conversely $A_{1,3} = e_1 V_0$, verifying the exceptional pair.

If $P(z) = \sum_j a_j z^j$ has at least two nonzero positive-degree coefficients, let $n > m$ be their two largest indices. Expanding the first and third columns of D, its unique highest homogeneous part is $-(n - m)a_m a_n A_{m,n}$, of degree $m + n$. Since $D = (e_1 - 1)K$, this part is divisible by e_1. The preceding calculation forces $(m, n) = (1, 3)$; hence $P = a_0 + a_1 z + a_3 z^3$ with $a_1 a_3 \neq 0$. Its full determinant is $D = -2a_1 a_3 e_1 V_0$, which at $(u, v, w) = (1/6, 1/3, 1/2)$ equals $-a_1 a_3/54 \neq 0$. This contradiction proves necessity for every polynomial degree in dimension three. The argument uses the highest nonzero degree and does not resolve the general analytic $d = 3$ case.

Conversely, for (85),

$$V(s) = \frac{a\|s\|_2^{2k+2} + c\sum_i s_i^{2k+2}}{2k + 2} \quad \Longrightarrow \quad \nabla V = \|s\|_2^{2k+1} h(s/\|s\|_2),$$

and the multiplier is positive on Ω_d. □

Determinantal antecedents. Braden & Byatt-Smith (1999) classify related ordinary-derivative determinants under a zero-sum constraint. Here recentering $u = x+1/3$ leaves the weighted derivative $(x + 1/3)P'(x + 1/3)$, whereas taking logarithmic coordinates changes the constraint to a sum of exponentials. More generally, Elekes & Szabó (2024, Lemma 3.8) obtain cubic containment for three continuous arcs whose collinearity is described by an Abelian group law, under their standard-system hypotheses. Such containment alone does not give the stated Euler-derivative classification: $(P(t), tP'(t))$ for $P = t + t^3$ lies on a cubic, but its determinant is nonzero on the positive simplex. Nevertheless, the cubic-containment theorem together with the following Euler-derivative argument gives a complete alternative proof of the polynomial classification. We therefore treat this classification as an explicit specialization, not a new general collinearity theorem. The direct highest-degree proof above is shorter and independent of that theorem.

Reduction to cubic containment and elementary algebra. Let P be nonconstant, $Q(t) = tP'(t)$, and $\gamma(t) = (P(t), Q(t))$. Use $D(u, v, w) = \det[1, P(t), Q(t)]_{t=u,v,w}$; this reverses the sign of the earlier column convention but preserves its zeros. Suppose $D = 0$ on the positive simplex $u + v + w = 1$. If the image of γ is a line, then $A + BP + CQ = 0$ for some constants not all zero. Nonconstancy forces $C \neq 0$, and coefficient comparison leaves at most one positive exponent in P. Assume henceforth that the image is not a line.

The polynomial D is nonzero, and write $D = (u+v+w-1)^r K$ with maximal $r \geq 1$. In fact $r = 1$. Otherwise all first partial derivatives of D vanish on the simplex. Fix $u_0 \in (0, 1)$ with $\gamma'(u_0) \neq 0$

and put $S = 1 - u_0$. The points $\gamma(v), \gamma(S - v)$ are distinct except at finitely many v: an identical equality would give both $P'(v) = -P'(S-v)$ and $vP'(v) = (S-v)P'(S-v)$, forcing $SP'(v) = 0$. For every other v, $D = \partial_u D = 0$ puts their secant through $\gamma(u_0)$ parallel to $\gamma'(u_0)$. An interval of the polynomial image would therefore lie on one fixed line, a contradiction.

Choose a positive simplex triple with $K \neq 0$. At this triple all three partials of D equal K, so the image points are distinct, regular, and their tangents are transverse to the common secant. On small disjoint parameter intervals, K remains nonzero and collinearity is equivalent to $u + v + w = 1$. Rotate the secant to the vertical axis. The three arcs become ordered graphs over a common interval $(-\epsilon, \epsilon)$. Shrink it so their height ranges have gap $\eta > 0$ and their slopes have absolute value at most M, with $\eta/(2\epsilon) > M$. Any nonvertical line meeting two arcs then has absolute slope greater than M, so meets each graph at most once; vertical lines do too. This verifies the intersection condition in Elekes & Szabó (2024, Definition 3.4). The group coordinates $\gamma(t) \mapsto t - 1/3$ verify their Definition 3.1 in both directions. Their Lemma 3.8 supplies a nonzero cubic equation for the arcs. Polynomial continuation extends it to the whole image; over $\mathbb{C}$, some irreducible factor $F(X, Y)$ of degree at most three satisfies $F(P, Q) = 0$.

Set $\mathcal{K} = \mathbb{C}(P, Q) \subset \mathbb{C}(t)$ and $\delta = t\, d/dt$. The nonconstant P is transcendental over $\mathbb{C}$, so F depends on Y. Characteristic zero and irreducibility give $F_Y(P, Q) \neq 0$: otherwise the irreducible relation would divide its nonzero lower-degree derivative. Differentiating the relation yields

$$\delta P = Q, \qquad \delta Q = -F_X(P, Q)Q/F_Y(P, Q) \in \mathcal{K}.$$

The quotient rule gives $\delta\mathcal{K} \subseteq \mathcal{K}$. If $P = a_0 + \sum_{i=1}^{r} a_i t^{m_i}$ with distinct positive m_i and nonzero a_i, the r iterates $\delta^j P$, $1 \leq j \leq r$, isolate every t^{m_i} through the invertible Vandermonde matrix (m_i^j). For $k = \gcd(m_1, \ldots, m_r)$, an integer Bezout combination gives $t^k \in \mathcal{K}$. The reverse inclusion is immediate, hence $\mathcal{K} = \mathbb{C}(t^k)$. Write $z = t^k$, $P = A(z)$, $Q = kzA'(z)$, and $m = \deg A$.

Choose a linear coordinate $L = aP + bQ$ with $a + bkm \neq 0$, and a complementary coordinate M. Then $L(z)$ has degree m, while the transformed cubic relation gives $[\mathcal{K} : \mathbb{C}(L)] \leq 3$. But $[\mathbb{C}(z) : \mathbb{C}(L(z))] = m$: the elements $1, z, \ldots, z^{m-1}$ span by polynomial reduction and are independent because the leading degrees of $A_j(L(z))z^j$ have distinct residues modulo m after clearing denominators. Thus $m \leq 3$.

Write $A(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3$, allowing zero coefficients. For $e_1 = \sum_i z_i$, $e_2 = \sum_{i<j} z_i z_j$, and $V = \prod_{i<j}(z_j - z_i)$, direct expansion gives

$$\det[1, A(z_i), kz_i A'(z_i)]_{i=1}^{3} = kV\{a_1 a_2 + 2a_1 a_3 e_1 + a_2 a_3 e_2\}.$$

Cancel V on the open subset of the positive simplex with distinct t_i, and continue the bracket polynomial over the simplex. Only then substitute $(t_1, t_2, t_3) = (t, t, 1 - 2t)$. For the resulting bracket $B(t)$,

$$B(0) = a_1 a_2 + 2a_1 a_3 = 0, \qquad B'(0) = \begin{cases} 2a_2 a_3, & k = 1, \\ -4k a_1 a_3, & k \geq 2. \end{cases}$$

For $k \geq 2$, these identities give $a_1 a_3 = a_1 a_2 = 0$; evaluating inside the simplex, where $e_2 > 0$, gives $a_2 a_3 = 0$. For $k = 1$, $a_2 a_3 = 0$ and $a_1(a_2 + 2a_3) = 0$ give the same conclusion by splitting on $a_3 = 0$. At most one positive-degree coefficient survives. This completes the alternative polynomial proof, not a reduction for general analytic P.

Corollary 31 (Fixed-norm directions and literal Jordan)**.** *Let h satisfy the analytic hypotheses of the preceding theorem. Suppose $h(t) > 0$ on $(0, 1)$ and its Taylor expansion contains at least two nonlinear powers. For $d \geq 4$, or $d \geq 3$ when h is a polynomial, no positive scalar multiple of the Frobenius-normalized field can select the subgradient of any norm. This holds without assuming the scalar is initially differentiable. Literal five-step Jordan satisfies the hypotheses. Its exclusion from even nonconvex potentials already holds globally for $d = 3$ under the integrating-factor regularity above.*

Proof. If a norm N had $bq \in \partial N$, the norm subgradient identity gives $b = N/(s^\top q)$. Since N and q are continuous and $s^\top q > 0$, this first makes bq continuous even when no regularity of b was assumed. A continuous subgradient selection of a finite convex function is its gradient: the two subgradient inequalities and continuity give the first-order remainder. Hence $N \in C^1$, the quotient makes $b \in C^1$, and $bq = \nabla N$ makes $N \in C^2$. The theorem applies.

For the frozen rational coefficients $f(t) = at + bt^3 + ct^5$, the discriminant $b^2 - 4ac = -2594691/500000 < 0$ and $a, c > 0$, so f, and therefore $H = f^{\circ 5}$, is positive on the positive axis. If A_j, B_j, C_j denote the coefficients of t, t^3, t^5 after j compositions, then

$$A_{j+1} = aA_j, \qquad B_{j+1} = aB_j + bA_j^3, \qquad C_{j+1} = aC_j + 3bA_j^2 B_j + cA_j^5.$$

Starting from $(A_0, B_0, C_0) = (1, 0, 0)$ gives $A_j > 0$, $B_j < 0$, and $C_j > 0$ for $1 \le j \le 5$. Thus H has two nonlinear powers. As a separate check for $d = 3$, the exact rational evaluation at $x = (2, 3, 6)/7$ gives

$$\frac{1652223}{10^6} < \det[H(x), x, x \odot H'(x)] < \frac{1652224}{10^6},$$

so (83) directly excludes a global integrating factor. The accompanying check computes the unreduced-composition rational exactly rather than trusting the decimal. □

Corollary 32 (Scalar repair of monotone Taylor–Newton–Schulz compositions)**.** *For integers $k \ge 1$, let*

$$c_j = 4^{-j}\binom{2j}{j}, \qquad f_k(t) = t\sum_{j=0}^{k} c_j(1 - t^2)^j.$$

Fix a finite schedule $(k_1, \dots, k_L)$, and set $H = f_{k_L} \circ \cdots \circ f_{k_1}$, with $H(t) = t$ when $L = 0$. Normalize the input once: $q_i(s) = H(s_i/\|s\|_2)$ on Ω_d, $d \ge 3$. There are $b \in C^1(\Omega_d)$, $b > 0$, and $V \in C^2(\Omega_d)$ with $bq = \nabla V$ exactly when $L = 0$ or $(L, k_1) = (1, 1)$. These two cases admit norm-derived directions. In every other case, no positive scalar rescaling is a norm-subgradient selection, even without initial regularity of the scalar. Every such H is strictly increasing on $(0, 1)$ and preserves $[0, 1]$, so its pointwise isotonic angle gap is zero.

Proof. The relation $2(j + 1)c_{j+1} = (2j + 1)c_j$ telescopes to

$$f_k'(t) = (2k + 1)c_k(1 - t^2)^k.$$

Thus the derivative is positive on $(0, 1)$; the endpoints $f_k(0) = 0$ and $f_k(1) = 1$ prove interval invariance, preserved by composition. Write $f_k(t) = a_k t + b_k t^3 + e_k t^5 + O(t^7)$, where

$$a_k = \sum_{j=0}^{k} c_j > 0, \qquad b_k = -\sum_{j=1}^{k} jc_j < 0, \qquad e_k = \sum_{j=2}^{k}\binom{j}{2}c_j \ge 0.$$

Here $e_k = 0$ exactly when $k = 1$. If a partial composition has coefficients (A, B, C) for (t, t^3, t^5), the next stage gives

$$A_+ = a_k A, \qquad B_+ = a_k B + b_k A^3, \qquad C_+ = a_k C + 3b_k A^2 B + e_k A^5.$$

Starting at $(1, 0, 0)$, the first stage has $A > 0$, $B < 0$, and $C \ge 0$, with $C > 0$ unless $k_1 = 1$. From the second stage onward, $3b_k A^2 B > 0$, so $A > 0$, $B < 0$, $C > 0$ persist. Except for the two stated cases, H therefore contains two nonzero nonlinear powers. Theorem 30 excludes an integrating factor; positivity of H gives the norm-regularity bootstrap in the preceding proof.

For $L = 0$, $V = \|s\|_2$ suffices. For one cubic step, set $r = \|s\|_2$, $x = s/r$, $A(x) = 3/2 - \sum_i x_i^4/2 \ge 1$, and

$$N(s) = \left(\frac{3r^4 - \sum_i s_i^4}{2}\right)^{1/4}, \qquad \nabla N = A(x)^{-3/4}\frac{3x - x^{\odot 3}}{2}.$$

For independent uniform signs ε_i, $\mathbb{E}|\varepsilon^\top s|^4 = 3r^4 - 2\sum_i s_i^4$, and hence

$$N(s)^4 = \tfrac{1}{4}\mathbb{E}|\varepsilon^\top s|^4 + \tfrac{3}{4}r^4.$$

Minkowski's inequality for this weighted ℓ_4 sum proves that N is an absolutely symmetric norm. Its multiplier is positive and smooth off zero. □

The Taylor family and derivative identity are classical; see Kim & Oh (2026, Definition 2) and Li & Tsuchiya (2026, Appendix D.1). The additional conclusion is the integrating-factor depth boundary at the *endogenous Frobenius input scale*. The fixed external scale in Li–Tsuchiya's Algorithm 1 and the roundwise FTRL representation in their Theorem 5.7 concern different maps and quantifiers. Our statement uses exact arithmetic, no additive epsilon, and no renormalization between stages; it neither invalidates those results nor implies a convergence or training penalty.

Proposition 33 (Exact nonexistence without a persistent approximation gap)**.** *For the Taylor schedule in corollary 32, write* $r = \|s\|_2$, $x = s/r$, $q_L = H(x)$, *and* $M_L = \prod_{j=1}^{L}(k_j + 1)$, *with* $M_0 = 1$. *Fix* $0 < \ell \le d^{-1/2}$ *and restrict to spectra with* $\min_i x_i \ge \ell$. *Set* $u_L = (1-\ell^2)^{M_L}$ *and* $a_L = \sqrt{1-u_L}$. *For the single fixed norm* $N_*(s) = \|s\|_1$ *on the positive cone,*

$$\angle(q_L, \mathbf{1}) \le \arcsin \frac{1-a_L}{1+a_L}, \tag{89}$$

$$r \left\| \frac{q_L}{s^\top q_L} - \nabla \log N_*(s) \right\|_2 \le \sqrt{d}\,(1-a_L). \tag{90}$$

Both bounds tend to zero as depth increases, although every finite depth beyond the exceptions in corollary 32 has no exact integrating factor. In contrast, for every fixed finite schedule,

$$\sup_{s \in \Omega_d} \angle(q_L(s), \mathbf{1}) = \arccos(1/\sqrt{d}).$$

This last equality concerns approximation to polar, not a lower bound against all norm-derived directions.

Proof. The scalar Taylor residual estimate is classical (Kim & Oh, 2026, Theorem 2 and residual analysis). A direct proof also fixes the mixed-schedule scope. Put $u = 1-t^2$ and $P_k(u) = \sum_{j=0}^{k} c_j u^j$. If $P_k^2 = \sum_{n=0}^{2k} D_n u^n$, the generating function $(1-u)^{-1/2}$ gives $D_n = 1$ for $n \le k$. For $k \le n < 2k$, the convolution $D_n = \sum_{i=n-k}^{k} c_i c_{n-i}$ decreases with n, since the next sum loses a term and each remaining second factor decreases. Setting $D_{2k+1} = 0$, the polynomial

$$1 - (1-u)P_k(u)^2 = \sum_{n=k+1}^{2k+1} (D_{n-1} - D_n) u^n$$

has nonnegative coefficients summing to one. Hence $0 \le 1 - f_k(t)^2 \le (1-t^2)^{k+1}$ on $[0,1]$. Iteration gives $1 - H(t)^2 \le (1-t^2)^{M_L}$. Also $f_k(t) \ge t$ and $f_k([0,1]) \subseteq [0,1]$, so $a_L \le q_{L,i} \le 1$ and $q_{L,i} \ge x_i$.

For any vector $q \in [a,1]^d$, let $\mu = d^{-1}\sum_i q_i$ and $v = d^{-1}\sum_i q_i^2$. The coordinate inequality $(q_i - a)(q_i - 1) \le 0$ gives $v \le (1+a)\mu - a$ and therefore $\cos^2 \angle(q, \mathbf{1}) = \mu^2/v \ge 4a/(1+a)^2$. This proves (89). For $S = \sum_i x_i$ and $Q = x^\top q_L$, the weighted average $Q/S \in [a_L, 1]$ and $Q \ge \sum_i x_i^2 = 1$. Thus

$$r\left(\frac{q_L}{s^\top q_L} - \frac{\mathbf{1}}{rS}\right) = \frac{q_L - (Q/S)\mathbf{1}}{Q},$$

which proves (90). Since $M_L \ge 2^L$, both bounds vanish uniformly on the stated region. Positive output scaling changes neither comparison; the factor r removes inverse-input-scale units from the pairing metric.

At fixed finite depth, continuity, $H(0) = 0$, and $H(1) = 1$ give $H(\sqrt{1-(d-1)\epsilon^2}, \epsilon, \ldots, \epsilon) \to (1, 0, \ldots, 0)$. The angle tends to $\arccos(1/\sqrt{d})$; nonnegative vectors cannot have a larger angle to $\mathbf{1}$. This proves the supremum and shows why depth limits and unrestricted spectral suprema cannot be interchanged. □

These are consequences of prior polar convergence, not a new convergence theorem. On any fixed loop in Ω_d, compactness supplies $\ell > 0$ and $r \ge r_{\min} > 0$. Its normalized circulation is at most $\sqrt{d}(1-a_L)/r_{\min}$, because $\nabla \log N_*$ has zero circulation. Thus the finite-depth certificate below cannot have a depth-independent positive lower bound on that loop. Exact incompatibility does not preclude accurate inexact norm oracles or imply worse optimization.

Proposition 34 (Uniform approximation by a norm at each depth)**.** *Fix* $d \ge 2$. *Let* $H : [-1,1] \to [-1,1]$ *be odd,* C^1, *and nondecreasing, with* $H(1) = 1$ *and* $H(t) \ge t$ *on* $[0,1]$. *Write*

$$\Psi(t) = \int_0^t H(u)\,du, \qquad \Phi(t) = tH(t) - \Psi(t), \qquad \delta = \Phi(1) = \int_0^1 (1 - H(t))\,dt, \qquad \varepsilon = d\delta.$$

For $s \neq 0$ let $r = \|s\|_2$, $x = s/r$, $q = H(x)$, and define

$$N_H(s) = r\left(\sum_i \Psi(x_i) + \varepsilon\right), \quad N_H(0) = 0, \qquad b(x) = \varepsilon - \sum_i \Phi(x_i). \tag{91}$$

Then N_H is an absolutely symmetric norm, C^2 off zero, and

$$\nabla N_H = q + bx, \qquad r\nabla^2 N_H = P\,\mathrm{diag}(H'(x_i))P + bP, \qquad P = I - xx^\top, \qquad 0 \le b \le \varepsilon. \tag{92}$$

Uniformly over all nonzero signed spectra, including zero coordinates,

$$\|\nabla N_H - q\|_2 \le \varepsilon, \qquad 0 \le N_H(s) - \|s\|_1 \le \varepsilon r, \tag{93}$$

$$\angle(q, \nabla N_H) \le \arcsin\frac{\varepsilon}{2+\varepsilon}, \qquad r\left\|\frac{q}{s^\top q} - \nabla \log N_H\right\|_2 \le \sqrt{d-1}\,\frac{\varepsilon}{1+\varepsilon}. \tag{94}$$

The error, value, angle, and pairing bounds hold for a fixed rectangular matrix shape with d the maximum rank, the unitarily invariant norm $\mathcal{N}_H = N_H \circ \sigma$, and Frobenius inner products and norms. Its gradient at every nonzero matrix, including rank-deficient matrices, is $\nabla\mathcal{N}_H(G) = Q_H(G) + bG/\|G\|_F$.

Proof. Both Ψ and Φ are even, and $\Phi(t) = \int_0^{|t|} uH'(u)\,du \in [0, \delta]$. Thus $0 \le b \le \varepsilon$. Put $A = \sum_i \Psi(x_i) + \varepsilon$ and $D = \mathrm{diag}(H'(x_i))$. The identities $\nabla r = x$ and $D_s x = P/r$ give

$$\nabla N_H = Ax + Pq = q + bx, \qquad \nabla_s b = -PDx/r,$$

so $rD_s(q + bx) = DP - xx^\top DP + bP = PDP + bP$. The Hessian is positive semidefinite. Restrictions to lines avoiding zero are therefore convex; on lines through zero, evenness and homogeneity give a positive multiple of the absolute value. To check positivity and the norm-value bound, set $E(t) = t - \Psi(t)$ for $0 \le t \le 1$. Then $0 \le E(t) \le E(1) = \delta$ and

$$N_H(s) - \|s\|_1 = r\left(\varepsilon - \sum_i E(|x_i|)\right) \in [0, \varepsilon r].$$

This also proves continuity at zero. Convexity, absolute symmetry, positivity, and homogeneity prove the norm assertion. A C^1 extension of H past ± 1 justifies the differentiation there; only its values on $[-1, 1]$ enter the formulas.

For the angle let $a = x^\top q \ge 1$ and $v = q - ax$, so $x^\top v = 0$, $\|q\|_2^2 \le d$, and $\|v\|_2^2 \le d - 1$. The two fields are $ax + v$ and $(a + b)x + v$. Their acute angle θ satisfies

$$\tan\theta = \frac{b\|v\|_2}{a(a+b) + \|v\|_2^2} \le \frac{b}{2\sqrt{a(a+b)}}, \qquad \sin\theta \le \frac{b}{2a+b} \le \frac{\varepsilon}{2+\varepsilon}.$$

Euler's identity also gives the exact pairing-error vector

$$r\left(\frac{q}{s^\top q} - \nabla \log N_H\right) = \frac{bv}{a(a+b)},$$

which proves (94). The factor r is essential; no scale-independent bound on the ordinary pairing distance is claimed.

For matrices, absolute symmetry gives the standard spectral norm lift. The displayed gradient is a subgradient by von Neumann's trace inequality: its singular coefficients are nonnegative, ordered like $\sigma(G)$, and the vector subgradient inequality bounds its pairing with any Z by $N_H(\sigma(Z))$, with equality at G. Repeated singular values receive equal coefficients and null singular spaces receive zero coefficients. Consequently the expression is SVD independent, without a full-rank or simple-spectrum assumption. Uniform odd-polynomial approximation to H makes Q_H a locally uniform limit of continuous matrix-polynomial maps off zero, so the subgradient selection is continuous. The two subgradient inequalities then prove differentiability with this gradient. Aligned SVDs preserve the relevant inner products and Frobenius errors. At $G = 0$ one may select the zero subgradient; no gradient, angle, or pairing is asserted. □

Prior construction and the quantifiers. The convex construction is a specialization of nonlinear perspectives, not a new convexification principle. Extend H by $\operatorname{sign}(t)$ outside $[-1,1]$ and let F be its primitive with $F(0)=0$. Then $\varphi(u)=\sum_i F(u_i)+d\delta \geq \|u\|_1 = \operatorname{rec}\varphi(u)$, and $\varphi \neq \operatorname{rec}\varphi$ since $\delta > 0$. Corollary 5.3(i) of Briceño-Arias et al. (2023), with scaling $s(y)=\|y\|_2$, followed by restriction to $u=y$, gives exactly (91). The direct proof above supplies the deployed-field and angular bounds, including rank loss.

For any Taylor schedule in proposition 33, the residual estimate there implies

$$\delta_L \leq \int_0^1 (1-t^2)^{M_L}\,dt = \frac{4^{M_L}(M_L!)^2}{(2M_L+1)!} \leq \frac{\sqrt{\pi}}{2\sqrt{M_L}}, \qquad M_L = \prod_{j=1}^{L}(k_j+1) \geq 2^L. \tag{95}$$

Here $1-H_L \leq 1-H_L^2$, and $1-t^2 \leq e^{-t^2}$ proves the last bound. For repeated fixed order k, the stronger prior bound is $\delta_L \leq ((f_k'(0))^L+1)^{-1}$ (Li & Tsuchiya, 2026, Lemma D.1); the primitive and its scalar deficit already occur in that work. These are not new scalar convergence rates. At fixed d, (94) therefore gives

$$\inf_{N\in\mathcal{A}_d}\ \sup_{s\neq 0} \angle\big(H_L(s/\|s\|_2), \nabla N(s)\big) \longrightarrow 0,$$

where $\mathcal{A}_d$ denotes absolutely symmetric norms C^2 off zero. One norm works for every input at each depth, but the norm changes with depth. It converges in value to $\|s\|_1$ on the unit sphere; its gradients need not converge uniformly to the full-rank polar gradient near rank loss. Indeed, the unrestricted positive-spectrum polar-angle supremum remains $\arccos(1/\sqrt{d})$ at every finite depth by the preceding proposition.

This additive input-radial correction does not assert an exact integrating factor for the original field; it does not contradict integrating-factor nonexistence. Nor is the bound dimension uniform: at a rank-one input its absolute error is exactly $(d-1)\delta_L$. Here d must remain the fixed matrix shape's maximum rank, not the varying active rank of each input. Small error at practical depth and large matrix dimension, efficient evaluation, and an optimization benefit are not established.

A folded convex exception. The theorem does not say that every first fold is intrinsically nonintegrable. With $h(t)=2t-t^3$ and $x=s/\|s\|_2$, define

$$N(s)=(2\|s\|_2^4-\|s\|_4^4)^{1/4}, \qquad b(s)=(2-\textstyle\sum_i x_i^4)^{-3/4}. \tag{96}$$

Then $b(s)h(x)=\nabla N(s)$. For independent uniform signs ε_i,

$$N(s)^4=\tfrac{1}{2}\mathbb{E}|\varepsilon^\top s|^4+\tfrac{1}{2}\|s\|_2^4.$$

Minkowski's inequality for this ℓ_4 direct sum proves that N is a symmetric-gauge norm, so its spectral lift is a unitarily invariant matrix norm. Yet at $s=(3,1)$ the first normalized coordinate has $h'=-7/10<0$. Moreover, at input condition number K the response condition number is $K(K^2+2)/(2K^2+1)\sim K/2$, while the classified unscaled cubic family has endpoint constant $2/3$. This exact counterexample prevents extending either no-gluing or the earlier frontier to all multiplier-corrected fields.

The analytic dimension restriction in theorem 30 is a proof boundary, not a claim that four is minimal for analytic filters. The polynomial threshold is three, and the following example proves it sharp. In dimension two every three-form vanishes. In dimension three the simplex constraint prevents the open three-variable identity step, but the highest-degree argument resolves polynomials. Vanishing of the three-form alone does not certify global convexity.

An explicit rank-two analytic exception. For $h(t)=t+t^3+t^5$ and $s\in\mathbb{R}^2\setminus\{0\}$, put $r=\|s\|_2$, $x=s/r$, $v=x_1^2x_2^2$, $A=3-5v$, and

$$N(s)=rA^{1/5}, \qquad N(0)=0.$$

Since $0\leq v\leq 1/4$, $A\geq 7/4$. Direct differentiation gives

$$\nabla N=A^{-4/5}h(x), \qquad \nabla^2 N=\frac{3(1+v)(1+3v)}{rA^{9/5}}(I-xx^\top).$$

For completeness, on the unit circle $x = (\cos\theta, \sin\theta)$, write $q = h(x)$, $a = x^\top q = A$, and $b = x_\perp^\top q$. Then $A' = 5b$, so $\psi = A^{1/5}$ obeys $\psi'/\psi = b/a$ and $\nabla N = \psi x + \psi' x_\perp = A^{-4/5} q$. Moreover, $\det(q, q') = 3(1+v)(1+3v)$ and $\psi + \psi'' = \psi \det(q, q')/a^2$, proving the Hessian formula. The definition is absolutely symmetric, real analytic off zero, and one-homogeneous. Its positive tangential Hessian proves convexity away from zero; lines through zero satisfy $N(ts) = |t|N(s)$. Thus N is a norm with strictly convex unit ball. In contrast, for the same filter in dimension three, $x = (2, 3, 6)/7$ gives $\det[h(x), x, x \odot h'(x)] = -311040/40353607 \neq 0$. This explicit dimension-two exception does not settle the general dimension-three analytic classification.

Proposition 35 (Smooth escape: analyticity is essential)**.** *For every $d \geq 2$ there is an odd C^∞ filter h, positive and strictly increasing on $(0, 1)$ but not of the binomial form in theorem 30, whose Frobenius-normalized field admits a positive smooth multiplier making it the gradient of an absolutely symmetric norm. The norm is smooth off zero and its unit ball is strictly convex.*

Proof. Fix $\varepsilon > 0$ and define

$$b(u) = \begin{cases} 0, & u \leq 1/2, \\ e^{-1/(2u-1)}, & u > 1/2, \end{cases} \qquad \phi(u) = 1 + \varepsilon b(u), \qquad h(t) = t\phi(t^2).$$

Then $\phi \geq 1$, $\phi' \geq 0$, and $h' > 0$. The filter equals t for $|t| \leq 1/\sqrt{2}$ but differs beyond, so it is not an analytic binomial. Put $A(u) = 1 - u + u\phi(u)$ and

$$K(u) = \int_0^u \frac{\phi(v) - 1}{2A(v)}\, dv, \qquad N(s) = r \exp\left(\sum_i K(u_i)\right), \quad r = \|s\|_2, \quad u_i = s_i^2/r^2,$$

with $N(0) = 0$. At most one $u_i > 1/2$. Since K is smooth and flat at $1/2$, the sum defines a smooth, even, permutation-invariant function off zero, including ties and zero coordinates. Writing $\psi = \exp(\sum_i K(u_i))$ and $D = 1 + \sum_i u_i(\phi(u_i) - 1)$ gives $2K'(u_i) = (\phi(u_i) - 1)/D$ and hence

$$\partial_j N = \psi \frac{s_j}{r} \left[1 + 2K'(u_j) - 2\sum_i u_i K'(u_i)\right] = \frac{\psi}{D} h(s_j/r).$$

Thus the multiplier ψ/D is smooth and positive.

For convexity, the all-inactive region has $N = r$. In an active cap $u = u_j > 1/2$, set $x = s/r$, $v = e_j - x_j x$, and $k = K'(u)$. Homogeneity gives a radial Hessian zero. On $x^\perp$,

$$\frac{r}{\psi} \nabla^2 N = (1 - 2uk) I_{x^\perp} + [2k + 4u(k' + k^2)] v v^\top.$$

Its transverse eigenvalue is $1/A$, and its meridional eigenvalue is

$$1 + 2(1 - 2u)k + 4u(1 - u)(k' + k^2) = \frac{\phi + 2u(1 - u)\phi'}{A^2} > 0.$$

At an axis ($u = 1$), both expressions agree by continuity. This proves positive tangential curvature everywhere off zero. Also $r \leq N(s) \leq e^{K(1)} r$, so N is continuous at zero. Lines avoiding zero are convex by the Hessian; on a line through zero, even homogeneity gives $N(tv) = |t|N(v)$. Thus N is a norm with strictly convex unit ball. □

This example explains the analytic identity step: at most one coordinate can enter the non-linear region, so every three-coordinate Pfaff determinant has two proportional rows. Smoothness alone permits this escape; analytic continuation forbids it. This is an assumption-boundary example, not a polynomial implementation or a polar-approximation improvement.

Proposition 36 (Quantitative projective loop separation)**.** *Let $q : \Omega \to \mathbb{R}^d$ be continuous on an open positive cone and assume $s^\top q(s) > 0$. Define the pairing-normalized direction*

$$\beta_q(s) = \frac{q(s)}{s^\top q(s)}. \tag{97}$$

For fixed s, $d_s([u],[v]) = \|u/(s^\top u) - v/(s^\top v)\|_2$ is a metric on rays represented by vectors with positive pairing against s. If N is a positive one-homogeneous function in $C^1(\Omega)$, then $\beta_{\nabla N} = \nabla \log N$. Consequently, for every closed piecewise C^1 loop $\Gamma \subset \Omega$ of Euclidean length $L(\Gamma) > 0$,

$$\sup_{t\in[0,1]} d_{\Gamma(t)}([q(\Gamma(t))],[\nabla N(\Gamma(t))]) \geq \frac{\left|\oint_\Gamma \beta_q^\top ds\right|}{L(\Gamma)}. \tag{98}$$

Both sides are invariant under arbitrary positive pointwise rescaling of q. Thus nonzero circulation of β_q quantitatively separates the displayed direction field from every differentiable norm-gradient field.

Proof. Each such ray intersects the affine section $s^\top z = 1$ exactly once, so Euclidean distance on that section proves the metric claim. Euler's identity gives $s^\top \nabla N = N$, hence $\beta_{\nabla N} = \nabla N/N = \nabla \log N$. Its integral around a closed loop is zero. Subtract this zero from the circulation of β_q, apply Cauchy–Schwarz pointwise, and bound the integral by $L(\Gamma)$ times the displayed supremum. Positive rescaling cancels in (97). □

Proposition 37 (A nonzero loop certificate for the monotone Taylor quintic)**.** *Let $d = 3$, $q_i(s) = f(s_i/\|s\|_2)$, where $f(t) = (15t - 10t^3 + 3t^5)/8$, and let Γ be the boundary of the rectangle*

$$\mathcal{R} = \{(s_1, s_2, 1) : 1/4 \leq s_1 \leq 1/3, \quad 1/2 \leq s_2 \leq 2/3\}.$$

For every positive one-homogeneous $N \in C^1(\Omega_3)$, in particular every differentiable norm,

$$\sup_{s\in\Gamma} \|\beta_q(s) - \nabla \log N(s)\|_2 \geq \frac{375}{120472576} > 3.1127 \times 10^{-6}.$$

The bound is unchanged by arbitrary positive pointwise output rescaling.

Proof. For a general quintic $at + bt^3 + ct^5$, put $R = \sum_i s_i^2$, $T = \sum_i s_i^4$, $U = \sum_i s_i^6$, and $F = aR^3 + bRT + cU$. Where $F > 0$, pairing normalization cancels powers of the radius, giving

$$\beta_i = \frac{s_i(aR^2 + bRs_i^2 + cs_i^4)}{F}, \qquad \partial_1\beta_2 - \partial_2\beta_1 = \frac{2bc\, s_1 s_2 s_3^2 (s_2^2 - s_1^2)(s_3^2 - s_1^2)(s_3^2 - s_2^2)}{F^2}.$$

The last identity follows by differentiating the displayed rational functions and collecting terms. For $(a, b, c) = (15/8, -5/4, 3/8)$, positivity of f on $(0, 1)$ gives $F > 0$. On $\mathcal{R}$, $R \leq 14/9$, $U \leq R^3$, and

$$F \leq (a + c)R^3 \leq 686/81, \qquad 2s_1 s_2 (s_2^2 - s_1^2)(1 - s_1^2)(1 - s_2^2) \geq 2 \cdot \tfrac{1}{4} \cdot \tfrac{1}{2} \cdot \tfrac{5}{36} \cdot \tfrac{8}{9} \cdot \tfrac{5}{9} = \tfrac{25}{1458}.$$

The curl has constant negative sign since $bc = -15/32$. Green's theorem, the area $1/72$, and the perimeter $1/2$ therefore give

$$\frac{|\oint_\Gamma \beta_q^\top ds|}{L(\Gamma)} \geq \frac{1}{36} \frac{(15/32)(25/1458)}{(686/81)^2} = \frac{375}{120472576}.$$

Apply proposition 36. □

This is a conservative certificate on a specified compact slice, not an optimal constant. It separates the monotone field from norm-derived rays without relying on order reversal. Its units depend on input scale; it is not an angular-error bound, a practical effect-size estimate, or a loss gap.

E Output normalization and fixed geometry

The conditioning and angular frontiers are invariant under positive radial output scaling. Conservativity and convex-subgradient realizability are not. We give elementary proofs to prevent those distinct statements from being conflated. All assertions in this section concern exact-real fields.

Lemma 38 (Radial rescaling of a homogeneous gradient)**.** *Let Ω be a connected open cone, let $N \in C^1(\Omega)$ be positive and one-homogeneous, and let $a : \Omega \to (0, \infty)$ be continuous and zero-homogeneous. Then $a\nabla N$ is conservative on Ω if and only if a is constant.*

Proof. The constant case has potential aN. Conversely suppose $\nabla F = a\nabla N$ for some $F \in C^1(\Omega)$ and denote this zero-homogeneous gradient by T. Since $\nabla_s(F(2s)-2F(s)) = 2T(2s)-2T(s) = 0$, connectedness makes $F(2s) - 2F(s) = k$ constant. Integration on the ray from s to $2s$ gives $F(2s) - F(s) = s^\top T(s)$; hence $F(s) + k = s^\top T(s) = a(s)N(s)$ by Euler's identity. In particular $a = (F + k)/N$ is C^1. Differentiating this equality and using $\nabla F = a\nabla N$ gives $N\nabla a = 0$. Positivity of N and connectedness force a to be constant. □

Theorem 39 (One negative mode after shell normalization)**.** *Assume theorem 26 and strict concavity of h on the principal interval. Fix $C \in (0, C_h]$ and a radius $R > 0$ independent of s. Write $N = \mathcal{N}_C$, $q = \nabla N$, $A = \nabla^2 N$, and $u = q/\|q\|_2$. The shell-normalized response $T(s) = Rq(s)/\|q(s)\|_2$ satisfies*

$$J_T = \frac{R}{\|q\|_2}(I - uu^\top)A, \qquad \|J_T - J_T^\top\|_F = \frac{\sqrt{2}R}{\|q\|_2^2}\|(I - uu^\top)Aq\|_2. \tag{99}$$

At every nonflat $s \in \Omega_d$, its symmetric part has inertia $(d-1, 1, 0)$ and the curl norm in (99) is strictly positive. At flat spectra the inertia is $(d-1, 0, 1)$ and the curl is zero. Thus this postnormalized field has no convex potential or any C^1 potential on the entire positive cone.

Proof. The C^1 principal normalizer and C^1 filter give $q \in C^1$ and $N \in C^2$ in the positive interior. There $D > 0$ and $\Delta > 0$ in (74). Weighted Cauchy–Schwarz, including its equality condition, shows $A \succeq 0$ with $\ker A = \operatorname{span}\{s\}$. Also $q^\top s = N(s) > 0$. Differentiating the Euclidean normalization gives the first formula in (99). Antisymmetrizing and decomposing Aq into its components parallel and perpendicular to q gives the second.

In the orthogonal decomposition $\mathbb{R}^d = \operatorname{span}\{u\} \oplus u^\perp$, write

$$A = \begin{pmatrix} a_0 & b^\top \\ b & B \end{pmatrix}.$$

The restriction B is positive definite: its null vector would have to be parallel to s, which is not orthogonal to u. Up to the positive scalar $R/\|q\|_2$, the symmetric part of the shell Jacobian is

$$\begin{pmatrix} 0 & b^\top/2 \\ b/2 & B \end{pmatrix}.$$

Its Schur complement is $-b^\top B^{-1} b/4$, strictly negative if $b \neq 0$. If $b = 0$, then Au is parallel to u. Taking its inner product with s and using $As = 0$ and $s^\top u > 0$ forces $Au = 0$, so u is parallel to s. The converse is immediate.

Strict concavity and $h(0) = 0$ make $h(t)/t$ strictly decreasing on the principal interval. Consequently q is parallel to s precisely when all coordinates of s are equal. This proves both inertia statements and the curl assertion. Nonzero curl excludes a potential; a negative quadratic direction of sym J_T gives nearby points violating monotonicity. A continuous selection of a convex subdifferential would be its ordinary gradient, so nonsmooth convex potentials cannot evade this test. □

Exact small witness and scope. For $h(t) = 3t - t^3$, choose $s = (1/2, 1/4)$ and $C = 429/1024$. Here $\alpha_C(s) = 1$, $q = (11/8, 47/64)$, and

$$A = \begin{pmatrix} 15/28 & -15/14 \\ -15/14 & 15/7 \end{pmatrix}.$$

The artifact checks the nonzero antisymmetric part and negative determinant of sym$[(I - qq^\top/(q^\top q))A]$ in exact arithmetic, plus independent finite-difference and higher-dimensional regression checks. These checks illustrate the written proof, not replace it. Strict concavity matters: a linear filter gives a Euclidean radial gradient, which remains conservative after shell normalization. Ordinary constant rescaling likewise preserves the potential. No floating-point optimizer trajectory is inferred.

Proposition 40 (Constant-radius full-space rigidity)**.** *Let E be a finite-dimensional real inner-product space of dimension at least two. Let $T : E \setminus \{0\} \to E$ be continuous and zero-homogeneous, with $\|T(x)\| = R > 0$ and $\langle x, T(x)\rangle > 0$. If T is a subgradient selection of one finite convex function on E, then $T(x) = Rx/\|x\|$ everywhere outside zero.*

Proof. Continuous convex-subgradient selection makes T an ordinary gradient outside zero. Integrate on each ray and use continuity of the finite convex potential at zero to obtain the positive one-homogeneous representative $N(x) = \langle x, T(x)\rangle$. Its minimum and maximum on the compact Euclidean unit sphere are attained. At either extremum, tangential derivatives vanish, so $\nabla N(x) = N(x)x$. Positivity and the constant gradient norm force both extreme values to equal R. Thus $N(x) = R\|x\|$ and differentiation proves the claim. □

Full-space continuity, including rank loss in a matrix application, is essential. On the positive cone the linear potential $N(s) = \sum_i s_i$ is a counterexample to full-space radial rigidity. Exact polar on full-rank matrices has fixed Frobenius norm but lacks a continuous extension through rank loss; an active-rank-dependent shell also fails the constant-radius hypothesis. The sphere argument is an elementary homogeneous instance of eikonal rigidity; much stronger almost-global classifications are classical (Caffarelli & Crandall, 2010). We claim no new general eikonal classification.

For matrix dimension $d = \min\{m, n\} \geq 2$, the classified field Q_C has a continuous, positive-pairing extension through rank loss, since equal singular values receive equal responses and $h(0) = 0$. Its Frobenius-shell normalization retains these properties but is nonradial on nonflat spectra, and therefore cannot define a fixed convex full-matrix potential. On active support of rank r, the only possible continuous constant-radius convex gradient, the normalized source itself, has polar angle

$$\arccos \frac{\|\sigma(G)\|_1}{\sqrt{r}\,\|\sigma(G)\|_2}.$$

This is a diagnostic distinction between displayed update fields, not a reason to prefer their training behavior. A positive scalar multiple of a fixed norm's oracle remains a norm-derived direction and may be analyzed using a state-dependent step size; it need not itself be that oracle.

F Pinned NVIDIA clamp-response certificate

This record makes the two implementation-specific predicates in the main-paper hard-clamp proposition independently replayable without importing the project's response implementation. All displayed decimals in this section are interpreted as exact base-ten rationals. Given a schedule $((a_k, b_k, c_k))_{k=0}^{4}$ and a rational input x, define

$$y_0 = x, \quad d_0 = 1, \qquad \begin{cases} y_{k+1} = a_k y_k + b_k y_k^3 + c_k y_k^5, \\ d_{k+1} = d_k(a_k + 3b_k y_k^2 + 5c_k y_k^4), \end{cases} \tag{100}$$

for $0 \leq k < 5$. Thus $y_5 = H(x)$ and $d_5 = H'(x)$ by the chain rule. At the pinned NVIDIA source revision used here,

`170d1b864049ab58fab134941eb6f268b3e64dee`

the two source-order schedules are

$$\texttt{simple}: \quad (3.4445, -4.7750, 2.0315) \quad \text{repeated five times}, \tag{101}$$

$$\begin{aligned} \texttt{quintic}: \quad & \big((4.0848, -6.8946, 2.9270), (3.9505, -6.3029, 2.6377), \\ & (3.7418, -5.5913, 2.3037), (2.8769, -3.1427, 1.2046), \\ & (2.8366, -3.0525, 1.2012)\big). \end{aligned} \tag{102}$$

Every c_k is nonzero, so both composed responses have exact degree $5^5 = 3125$. Exact evaluation of (100) certifies the four predicates

$$\begin{aligned} H'_{\text{simple}}(3/5) &\neq H'_{\text{simple}}(4/5), & H'_{\text{simple}}(1/5) &< 0, \\ H'_{\text{quintic}}(3/5) &\neq H'_{\text{quintic}}(4/5), & H'_{\text{quintic}}(1/20) &< 0. \end{aligned} \tag{103}$$

The unequal-derivative predicates are the dimension-two reflection failures; the negative derivatives are attained inside the fixed-denominator branch of the hard clamp.

Independently checkable verifier record. The standalone program

`qa/verify_nvidia_clamp_witnesses_independent.py`

contains (100)–(102) literally and uses only Python's exact `fractions.Fraction` arithmetic. For each large output fraction it records the sign, numerator and denominator bit lengths, and the SHA-256 digest of the canonical `numerator/denominator` string. The deterministic archive is

`experiments/geometry/results/33_nvidia_clamp_exact_witnesses.json`

and its SHA-256 is

`bcfd4a18283bdbf786e56f84a8217737934472d2b6e83c214f42db71f0d77147`.

The verifier reconstructs the archive byte for byte and aborts unless every inequality in (103) holds. This is a certificate for the printed exact-rational schedules and the pinned `normalize_in_double=False` source path, not for later commits, binary floating-point coefficients, or the complete optimizer.

G Exact certificate specifications for the Jordan case

This section separates the short mathematical reductions from the finite exact-arithmetic certificates. Decimal coefficients are interpreted as the exact rationals

$$a = \frac{6889}{2000}, \qquad b = -\frac{191}{40}, \qquad c = \frac{4063}{2000}, \tag{104}$$
$$f(x) = ax + bx^3 + cx^5, \qquad H = f^{\circ 5}.$$

No floating-point grid is used in either certificate below.

G.1 Sharp Frobenius pairing

Proposition 41 (Finite exact certificate for the sharp scalar bound)**.** *Let* $P_0(z) = 1$ *and recursively define, in* $\mathbb{Q}[z]$*,*

$$P_{j+1}(z) = P_j(z)\left[a + bzP_j(z)^2 + cz^2P_j(z)^4\right] \qquad (0 \le j < 5). \tag{105}$$

Set $\kappa_J = P_5(1)$ *and*

$$S(z) = \frac{P_5(z) - P_5(1)}{1 - z}. \tag{106}$$

Exact polynomial division has zero remainder, $\deg P_5 = 1562$*, and* $\deg S = 1561$*. The six rational intervals*

$$[0, 1/32],\ [1/32, 1/16],\ [1/16, 1/8],\ [1/8, 1/4],\ [1/4, 1/2],\ [1/2, 1] \tag{107}$$

form a partition of $[0, 1]$*. For every interval* $[\ell, r]$ *in* (107)*, all* 1562 *coefficients of*

$$T_{\ell,r}(t) = (1 + t)^{1561} S\left(\ell + \frac{(r - \ell)t}{1 + t}\right) \tag{108}$$

are strictly positive rational numbers. These finite predicates imply

$$H(x) \ge \kappa_J x \quad (0 \le x \le 1), \tag{109}$$

with equality only at $x = 0$ *or* $x = 1$*.*

Proof. Inductively, if $f^{\circ j}(x) = xP_j(x^2)$, substitution into $f(y) = y(a + by^2 + cy^4)$ gives exactly (105). Since the assertion is true for $j = 0$, it follows that

$$H(x) = xP_5(x^2). \tag{110}$$

For $t \in [0, \infty)$, the rational map $t \mapsto \ell + (r - \ell)t/(1 + t)$ maps onto $[\ell, r)$ and tends to r as $t \to \infty$. Strict positivity of every coefficient of $T_{\ell,r}$ implies $T_{\ell,r}(t) > 0$ for every finite $t \ge 0$; positivity of

its leading coefficient gives the same conclusion at the limiting right endpoint. Hence $S(z) > 0$ on every interval in the exact cover, and therefore on all of $[0, 1]$.

For $0 \le x \le 1$, put $z = x^2$. Equations (106) and (110) give

$$H(x) - \kappa_J x = x\,[P_5(z) - P_5(1)] = x(1-z)S(z) \ge 0.$$

Because $S(z) > 0$, equality is possible only when $x = 0$ or $z = 1$, which on the stated interval means $x = 1$. □

Corollary 42 (Sharp matrix pairing and equality case)**.** *For every nonzero matrix G, the exact-real, epsilon-free, memoryless Jordan map satisfies*

$$\langle G, Q_0(G)\rangle_F \ge \kappa_J \|G\|_F. \tag{111}$$

Equality holds if and only if G has rank one. Thus the constant is sharp simultaneously in every matrix dimension.

Proof. Let σ_i be the positive singular values, let $R = \|G\|_F$, and put $x_i = \sigma_i/R$. Then $0 < x_i \le 1$ and $\sum_i x_i^2 = 1$. The spectral calculus and (109) give

$$\langle G, Q_0(G)\rangle_F = R\sum_i x_i H(x_i) \ge R\kappa_J \sum_i x_i^2 = \kappa_J R.$$

For equality, every positive x_i must be an equality point of the strict scalar inequality, hence must equal one. Since their squares sum to one, there is exactly one positive singular value. Conversely, a rank-one matrix has the sole normalized singular value $x_1 = 1$ and attains equality. □

Independently checkable certificate record. The generator `experiments/geometry/26_jordan_sharp_descent.py` reconstructs (105) over $\mathbb{Q}$, performs the exact division in (106), and archives the six-leaf predicates in `experiments/geometry/results/26_jordan_sharp_descent.json`, including SHA-256 digests of P_5, S, every leaf transform, and the leaf cover. The independent program `qa/verify_jordan_sharp_descent_independent.py` imports neither the generator nor its polynomial objects: it rebuilds P_5 with a second exact FLINT implementation, verifies that the archived intervals tile $[0, 1]$, regenerates every transform in (108), and checks all 6×1562 coefficient inequalities. It also brackets

$$0.6964364094697528 < \kappa_J < 0.6964364094697529$$

by exact rational comparison. These checks make the finite computation reproducible; they do not replace review of the reduction proved above.

G.2 Finiteness and strictness of every scalar critical point

Define

$$C(y) = f'(y) = \frac{1}{2000}P(y^2), \qquad P(z) = 20315z^2 - 28650z + 6889, \tag{112}$$

and the squared scalar dynamics

$$R(z) = z(a + bz + cz^2)^2. \tag{113}$$

Proposition 43 (Finite-field strict-critical certificate)**.** *The polynomial H' is squarefree over $\mathbb{Q}$. Consequently it has finitely many real zeros, and every real critical point of H is a strict local maximum or a strict local minimum. Moreover $H'(0) = a^5 > 0$, while an exact rational evaluation gives $H'(1) < 0$; hence H' is positive near zero and negative somewhere.*

Proof. The chain rule gives the unexpanded factorization

$$H'(x) = \prod_{j=0}^{4} C(f^{\circ j}(x)). \tag{114}$$

The discriminant of P is 261022360, which is positive and not a square. Because P is primitive, Gauss's lemma and the quadratic discriminant criterion imply that P is irreducible, and in particular squarefree, over $\mathbb{Q}$. Also $P(0) = 6889 \neq 0$. Thus if $C(y) = 0$, then $y \neq 0$ and

$$C'(y) = \frac{y}{1000} P'(y^2) \neq 0.$$

Let $g_j(x) = C(f^{\circ j}(x))$. If $g_j(x_0) = g_j'(x_0) = 0$, then the last display and the chain rule force $(f^{\circ j})'(x_0) = 0$. By (114) applied to the shorter iterate, some $g_i(x_0) = 0$ with $i < j$. Therefore a repeated root of H' can occur only if two distinct factors g_i, g_j share a root.

Suppose $g_i(x_0) = g_j(x_0) = 0$ with $i < j$, and put $y = f^{\circ i}(x_0)$ and $z = y^2$. Then

$$P(z) = 0, \qquad P(R^{\circ(j-i)}(z)) = 0, \qquad 1 \leq j - i \leq 4. \tag{115}$$

Since P is irreducible, (115) would imply that P divides $P \circ R^{\circ k}$ in $\mathbb{Q}[z]$ for some $1 \leq k \leq 4$.

It remains to exclude these four finite divisibility predicates. Reduce modulo the prime $p = 1000003$. The leading coefficient 20315 and the denominators 40 and 2000 are nonzero modulo p, so rational divisibility would survive this reduction. In the quotient $\mathbb{F}_p[z]/(P)$, the relation is

$$z^2 = 377702 + 620973z. \tag{116}$$

Direct repeated squaring-map evaluation gives the following remainders of $P(R^{\circ k}(z))$ modulo $P(z)$:

$$\begin{aligned} k = 1: &\quad 411265 + 612970z, \\ k = 2: &\quad 880011 + 391268z, \\ k = 3: &\quad 249816 + 837277z, \\ k = 4: &\quad 393970 + 588222z. \end{aligned} \tag{117}$$

Each is nonzero in the two-dimensional quotient algebra. Hence none of the four rational divisibilities is possible, no two factors in (114) share a root, and the preceding single-factor argument proves that H' is squarefree.

The nonzero polynomial H' has finite degree, so it has finitely many real zeros. At every real zero, squarefreeness means the zero is simple. A real continuous function with a simple zero changes sign there; applied to H', this makes the corresponding critical point of H a strict local extremum. Finally, every iterate fixes zero and $f'(0) = a$, so $H'(0) = a^5 > 0$. Exact rational evaluation of the same product at $x = 1$ gives $H'(1) < 0$ (the sign record is independently archived), proving the remaining two assertions. □

Certificate replay and first-fold scope. The finite-field computation in (116)– (117) is generated by `experiments/geometry/27_jordan_critical_squarefree.py`; its compact record is `experiments/geometry/results/27_jordan_critical_squarefree.json`. The independent generic-polynomial replay is `qa/verify_jordan_critical_squarefree_independent.py`. It also iterates both rational bracket endpoints exactly, verifies that the first four orbit positions remain on the increasing branch, and compares the fourth squared iterates with the smaller root of P by exact signs. The exact sign audit `experiments/geometry/results/22_jordan_exact_symbolic_audit.json` records $H'(1) < 0$ and opposite exact signs at

$$\frac{4111351553545}{10^{15}} \quad \text{and} \quad \frac{4111351553546}{10^{15}}.$$

Those signs isolate a fold to the displayed decimal precision once combined with the principal inverse-branch argument below; the squarefreeness certificate alone proves strictness and finiteness, not the ordering of all critical points.

For completeness, let u be the smaller positive critical point of f. The quadratic P has two positive roots in z and its smaller root lies in $(0, 1/3)$ because $P(0) > 0$ and $P(1/3) < 0$. On $[0, 1/3]$ the polynomial $g(z) = a + bz + cz^2$ is decreasing, and $g(1/3) = 18707/9000 > 2$. Hence $f(x) = xg(x^2) > 2x$ for $0 < x \leq u$, while f is strictly increasing on $[0, u]$. Its increasing inverse branch L is therefore defined at u and satisfies

$$0 < L^4(u) < L^3(u) < L^2(u) < L(u) < u.$$

Putting $\zeta_J = L^4(u)$ in (114) shows that every factor is positive for $0 < x < \zeta_J$ and the final factor vanishes at ζ_J. Thus this is the first positive fold. The archived exact signs give

$$0.004111351553545 < \zeta_J < 0.004111351553546.$$

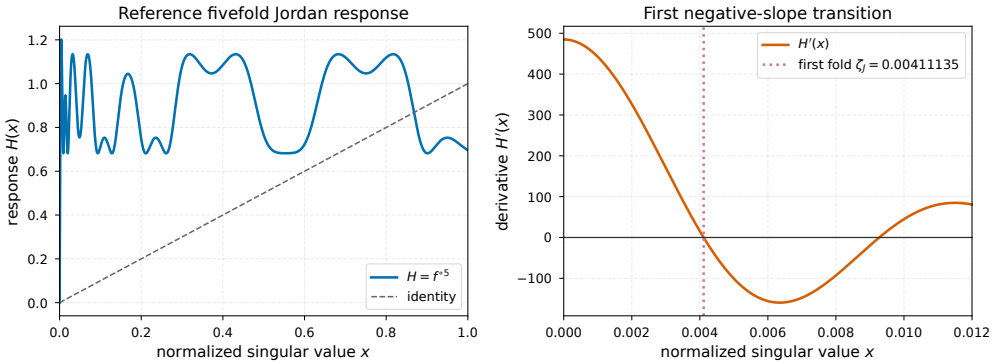


Figure 2: **Jordan NS5 scalar response and first fold.** The right panel resolves the first zero of H' at $\zeta_J \approx 0.00411135$. Under the hypotheses of the no-gluing theorem, every positive normalized singular value in a convex realization lies strictly before this fold.

G.3 Construction and endpoint of the principal matched branch

Take the primitive normalized by $\Phi_J(0) = 0$,

$$\Phi_J(t) = \int_0^t uH'(u)\,du, \qquad C_* := \Phi_J(\zeta_J). \tag{118}$$

The first-fold argument above gives $H'(t) > 0$ for $0 < t < \zeta_J$, so Φ_J is strictly increasing and positive on that interval.

Theorem 44 (Principal Jordan branch, including the boundary level)**.** *Let $d \geq 2$ and $0 < C \leq C_*$. For every $s \in \Omega_d$, there is a unique $\alpha_C(s) > 0$ such that*

$$\sum_{i=1}^{d} \Phi_J\left(\frac{s_i}{\alpha_C(s)}\right) = C, \qquad 0 < \frac{s_i}{\alpha_C(s)} < \zeta_J. \tag{119}$$

The normalizer α_C is smooth, positive, permutation invariant, and one-homogeneous on Ω_d. The function

$$\mathcal{N}_C(s) = \sum_i s_i H\left(\frac{s_i}{\alpha_C(s)}\right) \tag{120}$$

is a convex one-homogeneous potential and satisfies

$$\nabla \mathcal{N}_C(s) = \left(H\left(\frac{s_i}{\alpha_C(s)}\right)\right)_{i=1}^{d}. \tag{121}$$

For $0 < C < C_*$, the construction extends continuously and regularly to $[0,\infty)^d \setminus \{0\}$. For $C = C_*$, it still has a unique smooth root in the positive interior and a continuous convex boundary extension. Its sole loss of modular regularity on the nonzero closed cone is at the rank-one boundary rays: there the sole nonzero normalized coordinate equals ζ_J and $\Delta = 0$.

Conversely, every boundary-positive locally Lipschitz convex realization in the no-gluing theorem for the Jordan response agrees on Ω_d with exactly one of these branches for some $C \in (0, C_*]$.

Proof. Fix first $s \in \Omega_d$, put $m = \max_i s_i$, and define

$$\bar{\rho} = \frac{m}{\zeta_J}, \qquad R_s(\rho) = \sum_i \Phi_J(s_i/\rho) \quad (\rho \geq \bar{\rho}). \tag{122}$$

At $\rho = \bar{\rho}$, at least one normalized coordinate equals ζ_J. Since every other coordinate is positive and $\Phi_J(t) > 0$ for $t \in (0, \zeta_J]$,

$$R_s(\bar{\rho}) > \Phi_J(\zeta_J) = C_* \geq C. \tag{123}$$

Also $R_s(\rho) \to 0$ as $\rho \to \infty$. For $\rho > \bar{\rho}$, every argument lies in $(0, \zeta_J)$, and direct differentiation gives

$$R_s'(\rho) = -\frac{1}{\rho} \sum_i (s_i/\rho)^2 H'(s_i/\rho) < 0. \tag{124}$$

The intermediate value theorem and strict monotonicity therefore give one and only one root of (119). In particular the root satisfies $\alpha_C(s) > \bar{\rho}$, which gives the strict coordinate bounds in that equation even when $C = C_*$.

At the root, define $x = s/\alpha_C(s)$, $D = \operatorname{diag}(H'(x_i))$, and $\Delta = x^\top D x$. Every diagonal entry of D is positive, so $\Delta > 0$. Equation (124) and the implicit-function theorem make α_C smooth. Uniqueness gives, for every $\lambda > 0$ and every coordinate permutation π,

$$\alpha_C(\lambda s) = \lambda \alpha_C(s), \qquad \alpha_C(\pi s) = \alpha_C(s).$$

Thus the normalizer has all the asserted interior properties.

Differentiate (119) with respect to s_j. Since

$$\partial_j x_i = \frac{1}{\alpha_C(s)} \left(\delta_{ij} - x_i \partial_j \alpha_C(s)\right)$$

and $\Phi_J'(t) = tH'(t)$, the chain rule gives

$$0 = \frac{1}{\alpha_C(s)} \left[x_j H'(x_j) - \Delta\, \partial_j \alpha_C(s)\right].$$

Hence

$$\nabla \alpha_C(s) = \frac{Dx}{x^\top D x}. \tag{125}$$

Write $a = \alpha_C(s)$. Differentiating $\mathcal{N}_C(s) = \sum_i s_i H(x_i)$ now gives

$$\begin{aligned} \partial_j \mathcal{N}_C(s) &= H(x_j) + \sum_i s_i H'(x_i) \partial_j x_i \\ &= H(x_j) + x_j H'(x_j) - \partial_j \alpha_C(s) \sum_i x_i^2 H'(x_i) \\ &= H(x_j), \end{aligned}$$

where the last equality uses (125). This proves (121). A second differentiation, again using the displayed formula for $\partial_j x_i$, yields entrywise

$$\partial_j H(x_i) = \frac{1}{a} H'(x_i) \left(\delta_{ij} - x_i \frac{x_j H'(x_j)}{\Delta}\right).$$

Consequently

$$\nabla^2 \mathcal{N}_C(s) = \frac{1}{\alpha_C(s)} \left(D - \frac{Dxx^\top D}{x^\top D x}\right). \tag{126}$$

For every vector z, weighted Cauchy–Schwarz yields

$$(z^\top D x)^2 \leq (z^\top D z)(x^\top D x),$$

because $D \succ 0$. Therefore the Hessian in (126) is positive semidefinite. Convexity follows, and one-homogeneity follows from that of α_C.

It remains to prove the boundary and endpoint statements rather than infer them from the interior implicit-function theorem. Let $s \in [0, \infty)^d \setminus \{0\}$, use the same $m, \bar{\rho}, R_s$, and set $\Phi_J(0) = 0$. Then $R_s(\bar{\rho}) \geq C_*$, while R_s is strictly decreasing on $(\bar{\rho}, \infty)$. If $C < C_*$, there is a unique root strictly larger than $\bar{\rho}$. If $C = C_*$ and at least two coordinates of s are positive, the additional positive contribution makes $R_s(\bar{\rho}) > C_*$, so again the root is strictly larger. If s lies on a rank-one ray, then

$R_s(\bar{\rho}) = C_*$, and $\bar{\rho}$ is the unique root on $[\bar{\rho}, \infty)$. Its only nonzero normalized coordinate is ζ_J, so $\Delta = \zeta_J^2 H'(\zeta_J) = 0$.

For completeness, the resulting root is continuous on the nonzero closed cone, including at those degenerate rank-one points. By one-homogeneity it suffices to work on the compact closed simplex $\sum_i s_i = 1$. Choose $\delta \in (0, \zeta_J)$ so small that $d\Phi_J(\delta) < C$. Every root lies between m/ζ_J and m/δ: at the first scale the radial modular is at least C, while at the second it is at most $d\Phi_J(\delta) < C$. Since $1/d \le m \le 1$, the roots lie in one compact positive interval. If $s_n \to s$, every convergent subsequence of the roots has, by continuity of R_s, the unique root associated with s as its limit. Thus the entire root sequence converges. The potential in (120) consequently extends continuously. Passing the interior convexity inequality to boundary limits proves convexity of the extension. When $C < C_*$, every nonzero attained coordinate is strictly below ζ_J, so $\Delta > 0$ even on the boundary strata; this is the asserted regular case.

Finally, apply the no-gluing theorem. It supplies one $C \in (0, C_*]$, the equation (119), and strict principal-band coordinates at every positive interior point. The root uniqueness just proved forces the given normalizer to equal α_C throughout Ω_d, proving the converse. □

Corollary 45 (Full-matrix lift and within-class Jordan classification)**.** *Under the hypotheses and notation of theorem 44, define on* $\mathbb{R}^d$

$$g_C(z) = \mathcal{N}_C(|z|), \qquad \Psi_C(G) = g_C(\sigma(G)). \tag{127}$$

Then g_C is a symmetric-gauge norm and Ψ_C is a unitarily invariant norm, including when $C = C_*$. For $G \neq 0$, let $s = \sigma(G)$ and set

$$Q_C(G) = U \operatorname{diag}(q_C(s)) V^\top, \qquad (q_C(s))_i = H(s_i/\alpha_C(s)), \tag{128}$$

for any thin SVD $G = U \operatorname{diag}(s) V^\top$, and set $Q_C(0) = 0$. This definition is SVD independent and $Q_C(G) \in \partial\Psi_C(G)$ for every G. Equivalently, it is a linear-minimization-oracle selection for the dual unit ball of Ψ_C.

Conversely, let a permutation-invariant normalizer satisfy the positivity, one-homogeneity, local-Lipschitz, and positive-boundary hypotheses of the Jordan no-gluing theorem. If its aligned Jordan spectral field is a subgradient selection of any finite convex full-matrix function, then its normalizer is α_C for one unique $C \in (0, C_*]$. Thus the family $\{\alpha_C : 0 < C \le C_*\}$ is exactly the family admitting a full-matrix convex-subgradient realization within this shared-normalizer class. This does not assert uniqueness of arbitrary full-matrix potential extensions.

Proof. On the positive interior, $q_C = \nabla\mathcal{N}_C$ and each component is nonnegative because $H(0) = 0$ and $H' \ge 0$ on $[0, \zeta_J]$. If $0 \le u \le v$ coordinatewise, apply the interior gradient inequality along the segment from $u + \varepsilon\mathbf{1}$ to $v + \varepsilon\mathbf{1}$ and let $\varepsilon \downarrow 0$. Hence the closed-cone extension of $\mathcal{N}_C$ is coordinatewise nondecreasing. It is symmetric by uniqueness of the modular root, and it is positive away from the origin because a positive coordinate contributes $s_i H(s_i/\alpha_C(s)) > 0$. Define $\mathcal{N}_C(0) = 0$; convexity and one-homogeneity persist at the origin.

For $z, w \in \mathbb{R}^d$ and $0 \le \lambda \le 1$,

$$|\lambda z + (1-\lambda) w| \le \lambda |z| + (1-\lambda)|w|$$

coordinatewise. Monotonicity and convexity of $\mathcal{N}_C$ therefore show that g_C is convex. It is absolutely symmetric, positive definite, and one-homogeneous, hence a symmetric-gauge norm. Standard spectral lifting then makes $\Psi_C = g_C \circ \sigma$ a unitarily invariant norm (Lewis, 1995).

It remains to verify the displayed selection at boundary spectra, including the endpoint fold. For $s \neq 0$ in the closed positive cone, take $s^\varepsilon = s + \varepsilon\mathbf{1}$. Continuity of the modular root gives $q_C(s^\varepsilon) \to q_C(s)$. Passing the interior subgradient inequality to the limit proves

$$\mathcal{N}_C(t) \ge \mathcal{N}_C(s) + q_C(s)^\top (t - s) \quad \text{for all } t \in [0, \infty)^d. \tag{129}$$

This argument does not differentiate α_C at a rank-one endpoint, where $\Delta = 0$. Since H is nondecreasing on $[0, \zeta_J]$, $q_C(s)$ has the same coordinate order as s. Tied positive singular values receive equal responses, and zero singular values receive $H(0) = 0$, so (128) is independent of all SVD choices.

For any matrix Z, von Neumann's trace inequality and (129) give

$$\langle Q_C(G), Z\rangle_F \le q_C(s)^\top \sigma(Z) \le \mathcal{N}_C(\sigma(Z)) = \Psi_C(Z).$$

At $Z = G$, both inequalities are equalities because $q_C(s)^\top s = \mathcal{N}_C(s) = \Psi_C(G)$. This is precisely $Q_C(G) \in \partial\Psi_C(G)$ for a one-homogeneous convex function. At the origin, $0 \in \partial\Psi_C(0)$. The norm-subgradient identity also gives the dual-ball LMO statement.

For the converse, restrict the assumed full-matrix subgradient inequality to positive rectangular diagonal matrices. The Jordan no-gluing theorem and principal root uniqueness force the normalizer to be exactly one α_C. The construction above supplies a full-matrix realization for every such C, which proves the asserted if-and-only-if classification. □

G.4 Sharp Jordan fixed-geometry repair frontier

The preceding classification is qualitative: it determines every admissible matched normalizer. We now optimize a deployment-invariant quantity over that entire family. For a nonzero nonnegative vector v, define its positive-support condition number by

$$v_{\max} = \max\{v_i : v_i > 0\}, \qquad v_{\min} = \min\{v_i : v_i > 0\}, \qquad \kappa_+(v) = v_{\max}/v_{\min}, \tag{130}$$

with value one when there is only one positive coordinate. This convention discards the common null space and is invariant under every positive scalar rescaling.

Lemma 46 (Strict pre-fold concavity)**.** *The Jordan response $H = f^{\circ 5}$ is strictly increasing and strictly concave on $[0, \zeta_J]$. Consequently*

$$r(0) := H'(0) = a^5, \qquad r(t) := \frac{H(t)}{t} \quad (t > 0) \tag{131}$$

is strictly decreasing on $(0, \zeta_J]$.

Proof. The smaller positive critical point of f is

$$u = \sqrt{\frac{14325 - \sqrt{65255590}}{20315}}. \tag{132}$$

It was shown above that $f' > 0$ on $[0, u)$, $f'(u) = 0$, and $\zeta_J = L^4(u)$ for the increasing inverse branch L. Moreover

$$f''(x) = x\left(-\frac{573}{20} + \frac{4063}{100}x^2\right). \tag{133}$$

Its positive zero has square $2865/4063 = 14325/20315$, which is strictly larger than u^2 by (132). Thus $f''(x) < 0$ for $0 < x \le u$.

For completeness, all intermediate arguments stay in this concavity interval. Write $f(x)/x = g(x^2)$, where $g(z) = a + bz + cz^2$. On $[0, 1/3]$, g is decreasing because

$$g'(1/3) = b + 2c/3 = -\frac{5131}{1500} < 0,$$

and $g(1/3) = 18707/9000 > 2$. Since $u^2 < 1/3$, this gives $f(x) > 2x$ on $(0, u]$. Therefore, for $0 \le x \le \zeta_J$,

$$0 \le x \le f(x) \le f^2(x) \le f^3(x) \le f^4(x) \le u. \tag{134}$$

An increasing concave function composed with another increasing concave function is concave, so H is increasing and concave on the stated interval. Strictness follows directly as well. If $y_{j+1} = f(y_j)$, $A_j = (f^j)'(x)$, and $B_j = (f^j)''(x)$, then

$$A_{j+1} = f'(y_j)A_j, \qquad B_{j+1} = f''(y_j)A_j^2 + f'(y_j)B_j.$$

For $0 < x < \zeta_J$, induction gives $A_j > 0$ and $B_j < 0$ for $j \ge 1$. At the endpoint,

$$H''(\zeta_J) = f''(u)((f^4)'(\zeta_J))^2 < 0,$$

because the term multiplied by $f'(u)$ vanishes. Finally, strict concavity and $H(0) = 0$ imply $H(s) > (s/t)H(t)$ whenever $0 < s < t \le \zeta_J$, proving the last claim. □

Lemma 47 (Strictly decreasing logarithmic elasticity)**.** *With endpoint limits, the elasticity*

$$E_H(t) = \frac{tH'(t)}{H(t)} \tag{135}$$

decreases strictly from one to zero on $[0, \zeta_J]$.

Proof. For one Jordan step, putting $y = x^2$ gives

$$E_f(x) = \frac{xf'(x)}{f(x)} = \frac{20315y^2 - 28650y + 6889}{4063y^2 - 9550y + 6889}. \tag{136}$$

Exact differentiation yields

$$\frac{dE_f}{dy} = -\frac{4\Pi(y)}{(4063y^2 - 9550y + 6889)^2}, \quad \Pi(y) = 19400825y^2 - 55980014y + 32894975. \tag{137}$$

For $0 \le y \le u^2 < 1/3$,

$$\Pi'(y) \le \Pi'(1/3) = -\frac{129138392}{3} < 0, \qquad \Pi(y) \ge \Pi(1/3) = \frac{147515558}{9} > 0.$$

The denominator in (137) is positive, so E_f decreases strictly from one to zero on $[0, u]$. The composition identity

$$E_{f\circ g}(t) = E_f(g(t))E_g(t)$$

and (134) give

$$E_H(t) = \prod_{j=0}^{4} E_f(f^j(t)). \tag{138}$$

Every factor is positive before ζ_J, every orbit coordinate increases strictly with t, and E_f decreases strictly. Hence the product decreases strictly, tends to one at zero, and vanishes at ζ_J. □

For $K > 1$ and $0 < C \le C_*$, let $\tau_{K,C} \in (0, \zeta_J)$ be the unique root

$$\Phi_J(\tau_{K,C}) + \Phi_J(\tau_{K,C}/K) = C, \tag{139}$$

and define

$$\rho_{J,C}(K) = \frac{H(\tau_{K,C})}{H(\tau_{K,C}/K)}, \qquad \rho_J(K) = \frac{H(\tau_K^\star)}{H(\tau_K^\star/K)}, \tag{140}$$

where $\tau_K^\star$ is the root in (139) at $C = C_*$. Set both quantities equal to one at $K = 1$. The root exists uniquely because the left side of (139) is strictly increasing, starts at zero, and at ζ_J exceeds $C_* \ge C$.

Theorem 48 (Sharp Jordan fixed-geometry repair frontier)**.** *Let* $d \ge 2$, $0 < C \le C_*$, *and let* q_C *be the principal matched response in* (128). *For every nonzero nonnegative spectrum* s, *with* $K = \kappa_+(s)$,

$$\boxed{\rho_{J,C}(K) \le \kappa_+(q_C(s)) \le K.} \tag{141}$$

The lower bound is attained by a rank-two spectrum for every fixed (K, C). *For fixed* $K > 1$, $\rho_{J,C}(K)$ *decreases strictly with* C, *so the maximum level gives the strongest-compression envelope* $\rho_J(K) \le \rho_{J,C}(K)$. *Moreover*

$$\boxed{\eta_J K < \rho_J(K) \le \kappa_+(q_C(s)) \le K} \quad (K > 1), \tag{142}$$

where

$$\eta_J = \frac{H(\zeta_J)}{\zeta_J H'(0)} = 0.603145467389455239821680027 4\ldots, \tag{143}$$

and

$$\lim_{K\to\infty} \frac{\rho_J(K)}{K} = \eta_J. \tag{144}$$

Every assertion is unchanged by any positive radial output rescaling, including Frobenius-shell deployment.

Proof. Write $x = s/\alpha_C(s)$ and let $t = \max_i x_i$. The common positive normalizer and $H(0) = 0$ preserve support, so $\kappa_+(x) = \kappa_+(s) = K$ and the smallest positive coordinate is t/K. If $K > 1$, these extrema occur at distinct coordinates. Since all modular terms are nonnegative,

$$\Phi_J(t) + \Phi_J(t/K) \le \sum_i \Phi_J(x_i) = C. \tag{145}$$

Strict monotonicity in (139) gives $t \le \tau_{K,C}$. As H is increasing pre-fold,

$$\kappa_+(q_C(s)) = R_K(t), \qquad R_K(t) = \frac{H(t)}{H(t/K)}. \tag{146}$$

Lemma 47 gives

$$\frac{d\log R_K(t)}{d\log t} = E_H(t) - E_H(t/K) < 0. \tag{147}$$

Thus $R_K(t) \ge R_K(\tau_{K,C}) = \rho_{J,C}(K)$. Lemma 46 also gives

$$R_K(t) = K\frac{r(t)}{r(t/K)} \le K,$$

which proves (141). Equality in its lower bound occurs for the rank-two normalized spectrum $(\tau_{K,C}, \tau_{K,C}/K)$, which exhausts the modular constraint. The $K = 1$ case is immediate.

Increasing C increases $\tau_{K,C}$ and (147) then decreases $\rho_{J,C}(K)$. For finite $K > 1$, both arguments in (140) are strictly inside their endpoint limits, so

$$\frac{\rho_J(K)}{K} = \frac{r(\tau_K^\star)}{r(\tau_K^\star/K)} > \frac{r(\zeta_J)}{r(0)} = \eta_J.$$

As $K \to \infty$, (139) with $C = C_*$ forces $\tau_K^\star \to \zeta_J$ and $\tau_K^\star/K \to 0$, which proves (144). Positive radial rescaling cannot change any coordinate ratio.

If zeros are counted in the ordinary condition number, both input and response are singular; (130) states the nontrivial common-support result. At $C = C_*$, support at least two forces $t < \zeta_J$, because one coordinate at ζ_J would exhaust the modular level. A rank-one boundary ray has $K = 1$. Thus the proof never differentiates the degenerate endpoint normalizer. □

For $R \ge 1$, define

$$\vartheta(R) = \arccos\left(\frac{R+1}{\sqrt{2}\sqrt{R^2+1}}\right). \tag{148}$$

This is the angle between $(R, 1)$ and the polar direction $(1, 1)$, and it increases with R.

Corollary 49 (Sharp projective and weak-mode price)**.** *For a rank-two spectrum $s = (K, 1)$, every deployed principal matched response satisfies*

$$\angle(q_C(s), (1,1)) \ge \vartheta(\rho_{J,C}(K)) \ge \vartheta(\rho_J(K)), \tag{149}$$

and its weak coordinate after radius-$\sqrt{2}$ shell normalization obeys

$$[\sqrt{2}q_C(s)/\|q_C(s)\|_2]_{\text{weak}} \le \frac{\sqrt{2}}{\sqrt{\rho_{J,C}(K)^2+1}} \le \frac{\sqrt{2}}{\sqrt{\rho_J(K)^2+1}}. \tag{150}$$

The first inequalities in (149) and (150) are sharp for each fixed C. The envelope inequalities involving ρ_J are attained at $C = C_*$; for $C < C_*$ they are strict.

Proof. Theorem 48 gives the response ratio lower bounds. The angle in (148) increases, and $\sqrt{2}/\sqrt{R^2+1}$ decreases, with $R \ge 1$. Radial output normalization changes neither direction nor ratio. The rank-two equality case in the theorem proves the fixed-C sharpness, and setting $C = C_*$ proves the envelope sharpness. □

At the maximum-compression level C_*, exact rational root brackets and 384-bit Arb evaluation certify

K	$\rho_J(K)$	angle from polar	weak shell coordinate	literal Jordan angle
10	> 6.57723348933760	$> 36.3549793216°$	< 0.2125735823	$< 0.2711290178°$
100	> 60.8742229952083	$> 44.0588688668°$	< 0.0232285969	$< 0.0988228548°$

Here the literal comparison is the exact-real, epsilon-free response

$$q_0(K) = \left(H\left(\frac{K}{\sqrt{K^2+1}} \right), H\left(\frac{1}{\sqrt{K^2+1}} \right) \right). \tag{151}$$

The spherical triangle inequality consequently separates every classified repair from q_0 by more than $36.0838503038°$ at $K = 10$ and $43.9600460121°$ at $K = 100$.

Exact certificate and scope. The independent verifier `qa/verify_jordan_price_of_fixed_geometry.py` expands the degree-3125 response and degree-3126 primitive over $\mathbb{Q}$, checks the concavity and elasticity signs exactly, isolates ζ_J in a rational interval of width 10^{-32}, and proves exact rational brackets for both finite-K roots. It then evaluates (143) and the displayed angular bounds with 384-bit Arb balls. Its deterministic record is `experiments/geometry/results/34_jordan_price_of_fixed_geometry.json`. This establishes a tax only for preserving the literal smooth Jordan filter inside the classified homogeneous shared-normalizer family. It is not a universal price of convex geometry: fractional-power responses $H(t) = t^\gamma$, $0 < \gamma < 1$, have fixed Schatten geometry and approach polar as $\gamma \downarrow 0$.

G.5 Quantitative two-point separation

For the literal Jordan witness, take the exact decimal $v = 0.0041072578$, let $u = \sqrt{1-v^2}$, and set $A = \mathrm{Diag}(u, v)$, $B = \mathrm{Diag}(v, u)$. Here $u > v > 0$ and both matrices have Frobenius norm one. The exact verifier encloses v in $[0.0041072577, 0.0041072579]$ and certifies $H(v) - H(u) > 0.5059196414013767$.

Proposition 50 (Sharp two-point reduction)**.** *Let Q be any field, let $A \neq B$, and suppose $\gamma = -\langle Q(A) - Q(B), A - B \rangle_F > 0$. Every field T that is monotone on this pair obeys*

$$\max\{\|Q(A) - T(A)\|_F, \|Q(B) - T(B)\|_F\} \geq \frac{\gamma}{2\|A - B\|_F},$$

and the constant is sharp when only this pairwise constraint is imposed. For the swapped Jordan witness specified above, the bound equals $(H(v) - H(u))/\sqrt{2} > 0.3577392$.

Proof. Write $E_X = Q(X) - T(X)$ and $D = A - B$. Monotonicity gives

$$\gamma \leq -\langle E_A - E_B, D \rangle_F \leq (\|E_A\|_F + \|E_B\|_F)\|D\|_F.$$

Equality is attained on the two-point domain by taking $E_A = -\gamma D/(2\|D\|_F^2)$ and $E_B = \gamma D/(2\|D\|_F^2)$, for which the pairing of $T(A) - T(B)$ with D is zero. For the swapped-diagonal witness, $\gamma = 2(u - v)(H(v) - H(u))$ and $\|D\|_F = \sqrt{2}(u - v)$. Exact rational interval propagation gives $H(v) - H(u) > 0.5059196414013767$, and exact squaring proves the displayed 0.3577392 bound. □

The anonymous artifact's standalone exact-rational verifier uses only `fractions.Fraction`: it checks the outward enclosure of $u = \sqrt{1 - v^2}$ by exact squaring, propagates the five rational Jordan recurrences, verifies the negative pairing, and proves $(H(v) - H(u))^2 > 2(0.3577392)^2$; the claim manifest identifies and hashes both the verifier and its compact record.

G.6 Exact pointwise projective comparator

We state the tied-coordinate version to make the scope exact. For $s \in (0, \infty)^d$, define the closed convex preorder cone

$$\mathcal{K}_s = \{z \in [0, \infty)^d : (s_i - s_j)(z_i - z_j) \geq 0 \text{ for all } i, j\}. \tag{152}$$

If the coordinates of s are distinct and sorted increasingly, this is the nonnegative part of the usual isotonic cone. Tied coordinates impose no artificial order within their tie block.

Proposition 51 (Exact convex-spectral projective distance)**.** *Let $s \in (0,\infty)^d$, let $y \in (0,\infty)^d$, and set $p = \Pi_{\mathcal{K}_s} y$. Across all convex absolutely symmetric gauges γ and all nonzero $z \in \partial\gamma(s)$,*

$$\inf_{\gamma,\, z \in \partial\gamma(s)\setminus\{0\}} \angle(y,z) = \arccos \frac{\|p\|_2}{\|y\|_2} = \arcsin \frac{\|y-p\|_2}{\|y\|_2}. \tag{153}$$

The infimum is attained by an explicit symmetric gauge. For every chosen SVD of a matrix with singular values s, its unitarily invariant spectral lift attains the same Frobenius angle to the response obtained by aligning y with those singular vectors. If s has ties and y is not constant within each tied block, that aligned target depends on the chosen basis; the matrix assertion is then basis-relative rather than intrinsic.

Proof. Let $z \in \partial\gamma(s)$. If P swaps coordinates i, j, absolute permutation invariance and the subgradient inequality give

$$0 = \gamma(Ps) - \gamma(s) \geq \langle z, Ps - s\rangle = -(s_i - s_j)(z_i - z_j).$$

If R_i flips the sign of coordinate i, the same argument gives $0 \geq -2s_i z_i$, hence $z_i \geq 0$ because $s_i > 0$. Thus every admissible subgradient lies in $\mathcal{K}_s$.

The Moreau decomposition for this closed convex cone writes $y = p + r$ with $p \in \mathcal{K}_s$, $r \in \mathcal{K}_s^\circ$, and $\langle p, r\rangle = 0$. Hence, for every nonzero $z \in \mathcal{K}_s$,

$$\langle y, z\rangle \leq \langle p, z\rangle \leq \|p\|_2 \|z\|_2.$$

The projection is nonzero because $y > 0$ and the positive constant ray lies in $\mathcal{K}_s$. Taking $z = p$ attains both inequalities, and orthogonality gives the sine expression in (153).

It remains to show that this cone optimum is an actual gauge subgradient. Let $\mathcal{H}_d$ be the finite group of signed coordinate permutations and define

$$\gamma_p(x) = \max_{R \in \mathcal{H}_d} \langle Rp, x\rangle. \tag{154}$$

This is the support function of the convex hull of the signed-permutation orbit of p; since $p \neq 0$, it is a convex absolutely symmetric norm. The pairwise inequalities in (152), positivity, and the rearrangement inequality show that p itself maximizes (154) at s (permutations within a tied block give the same pairing). Therefore $p \in \partial\gamma_p(s)$. Standard subgradient calculus for unitarily invariant functions (Lewis, 1995) then gives $U\operatorname{diag}(p)V^\top \in \partial[\gamma_p \circ \sigma](U\operatorname{diag}(s)V^\top)$. For completeness, this aligned lift is optimal even among unaligned matrix subgradients. Fix the displayed thin SVD $G = U\operatorname{diag}(s)V^\top$ and the target $Y = U\operatorname{diag}(y)V^\top$. For any unitarily invariant norm $F = \gamma \circ \sigma$ and $0 \neq Z \in \partial F(G)$, put $z_i = u_i^\top Z v_i$. Restricting its subgradient inequality to $U\operatorname{diag}(t)V^\top$ for all $t \in \mathbb{R}^d$ gives $z \in \partial\gamma(s) \subset \mathcal{K}_s$. The norm identity $s^\top z = F(G) > 0$ gives $z \neq 0$. Since the matrices $u_i v_i^\top$ are Frobenius orthonormal, $\|z\|_2 \leq \|Z\|_F$, while $\langle Y, Z\rangle_F = y^\top z > 0$. Consequently

$$\cos\angle_F(Y,Z) \leq \frac{y^\top z}{\|y\|_2\|z\|_2} \leq \frac{\|p\|_2}{\|y\|_2}.$$

The aligned lift of p attains equality. This proves the matrix statement, including rectangular matrices and repeated singular values. When s has a tied block, simultaneous rotations of the corresponding columns of U and V leave the matrix unchanged but can change the aligned target unless y is constant on that block. Hence the preceding conclusion is relative to the displayed SVD in that case; it is basis-independent whenever s is untied or y is block-constant on every tie. □

If both aligned responses are rescaled to Frobenius norm $\sqrt{d}$, their distance divided by $\sqrt{d}$ is $\sqrt{2 - 2\cos\theta} = 2\sin(\theta/2)$. This makes the metric invariant to every positive radial output rescaling. The result remains pointwise: the orbit-support gauge may depend on (s, y), and the projected responses at several matrices need not be cyclically monotone or arise from one fixed gauge. It also does not cover arbitrary nonspectral convex bodies.

The following consequence removes symmetry and homogeneity of the candidate potential by changing the question from one input to a finite orbit. Its proof uses classical finite-group averaging and orbit geometry (Lewis, 1996), not new integrability machinery.

Proposition 52 (Sharp worst-orbit approximation)**.** *Let $s, y \in (0,\infty)^d$, assume y is constant on every tie block of s, and set $p = \Pi_{\mathcal{K}_s} y$ and $\theta = \arccos(\|p\|_2/\|y\|_2)$. Let $\mathcal{H}_d$ be the signed permutation group. Across all finite convex potentials $\Psi : \mathbb{R}^d \to \mathbb{R}$ and functions $F : \mathcal{H}_d s \to \mathbb{R}^d \setminus \{0\}$ satisfying $F(x) \in \partial\Psi(x)$,*

$$\inf_{\Psi, F} \max_{R \in \mathcal{H}_d} \angle(F(Rs), Ry) = \theta.$$

The infimum is attained. Neither symmetry nor homogeneity of Ψ, nor control of the positive response magnitudes, is required. The same value holds for arbitrary finite convex matrix potentials and their nonzero subgradient selections on the rectangular-diagonal orbit $\{\mathrm{Diag}_{m,n}(Rs) : R \in \mathcal{H}_d\}$, with target $\mathrm{Diag}_{m,n}(Ry)$ and $d = \min\{m, n\}$.

Proof. Write $\delta = \max_R \angle(F(Rs), Ry)$. Since $y > 0$, the preceding projection argument gives $p \neq 0$ and $\theta < \pi/2$. If $\delta \geq \pi/2$, the lower bound is immediate. Otherwise define

$$\bar{\Psi}(x) = \frac{1}{|\mathcal{H}_d|} \sum_R \Psi(Rx), \qquad v_R = R^\top F(Rs), \qquad z = \frac{1}{|\mathcal{H}_d|} \sum_R v_R.$$

Averaging the subgradient inequalities gives $z \in \partial\bar{\Psi}(s)$. Absolute symmetry of $\bar{\Psi}$ and the reflection and transposition arguments above imply $z \in \mathcal{K}_s$; these arguments do not require homogeneity. With $\widehat{y} = y/\|y\|_2$,

$$\langle z, \widehat{y} \rangle \geq \frac{\cos\delta}{|\mathcal{H}_d|} \sum_R \|v_R\|_2 \geq \cos\delta \, \|z\|_2.$$

The middle expression is strictly positive, so $z \neq 0$ and $\theta \leq \angle(y, z) \leq \delta$.

For attainment use the orbit-support norm γ_p in (154). Rearrangement gives $p \in \partial\gamma_p(s)$ and hence $Rp \in \partial\gamma_p(Rs)$. Every element stabilizing s preserves both y and $\mathcal{K}_s$; uniqueness of projection therefore makes it preserve p. Thus $F(Rs) = Rp$ is a well-defined function on the orbit and attains θ at every input.

For matrices, restrict a candidate potential to the rectangular diagonal subspace and project each selected subgradient onto that subspace. If every matrix error is acute, the projection is nonzero and cannot increase its angle to the diagonal target: it preserves the positive pairing and contracts the Frobenius norm. The vector lower bound applies; a nonacute matrix error already exceeds θ. The unitarily invariant spectral lift of γ_p attains the bound by the preceding subgradient construction. □

In fact the optimum is unchanged if competitors are merely monotone maps on the finite orbit. Define $\bar{F}(x) = |\mathcal{H}_d|^{-1} \sum_R R^\top F(Rx)$ there. It remains monotone and is equivariant. Writing $z = \bar{F}(s)$, pairwise monotonicity against a sign flip gives $4 s_i z_i \geq 0$, and against a transposition gives $2(s_i - s_j)(z_i - z_j) \geq 0$. Thus $z \in \mathcal{K}_s$. The same acute-angle average is nonzero and proves the lower bound; the norm above still attains it. Diagonal projection preserves the monotonicity pairings in the matrix version. This particular approximation criterion therefore detects an order obstruction, not an additional integrability constraint.

This is a worst-orbit obstruction, not a pointwise bound against arbitrary potentials: $\Psi(x) = y^\top x$ matches the single response at s exactly. The signed-permuted inputs need not occur in a training trajectory. No training penalty follows. Positive rescaling of target or selected responses does not change the angles, although rescaling a selection need not preserve its subgradient property for the same potential.

Experiment 31 uses FP32 pre-norm causal decoders with tied character embedding/head on Tiny Shakespeare (65-character vocabulary, context 64, batch 16). The four widths/layer counts are $128/4$, $256/6$, $512/8$, and $768/12$ (0.805M, 4.758M, 25.249M, and 85.072M parameters). The 250-step screen uses seeds $401, 409, 419, 421, 431$ and every hidden matrix at steps $1, 10, 50, 100, 200, 225, 250$. The two larger models additionally run to 2,000 steps with seeds $401, 419, 431$: first, middle, and last blocks at steps $1, 250, 500, 1000, 1600, 1800$, and every hidden matrix at step 2,000. The stored sources are post-momentum, pre-filter matrices, not model weights. These are short character-level trajectories, not training-scale evidence.

For Experiment 31 the preregistered full-spectrum response is $y_i = H(s_i/(\|s\|_2 + 10^{-7}))$. Raw replay verifies 4,872 strictly positive spectrum records with no exact ties. Six 2,000-step trajectories

Table 2: **Exploratory monotone-filter replay.** Minimum, median, and maximum $\mathcal{P}$ over the 4,728 unique stored spectra, rounded to six significant figures. The order-gap zeros are analytic; the other columns are float64 QR evaluations, not certified interval bounds.

filter	isotonic angle	min $\mathcal{P}$	median $\mathcal{P}$	max $\mathcal{P}$
g	0°	9.18525×10^{-19}	1.06113×10^{-17}	6.07826×10^{-17}
$g^{\circ 2}$	0°	3.70745×10^{-6}	4.83793×10^{-4}	1.46513×10^{-2}
f	0°	9.53296×10^{-7}	9.79145×10^{-5}	3.89262×10^{-3}
$f^{\circ 5}$	0°	4.12453×10^{-2}	1.75361×10^{-1}	2.88637×10^{-1}

repeat the first two checkpoints of their corresponding 250-step records, producing 144 bit-identical overlaps and 4,728 unique spectrum vectors. On those unique vectors, $\mathcal{K}_s$ reduces to the ordinary isotonic cone after sorting, and the deterministic full-spectrum angle has minimum 7.3806547631°, median 9.1352486279°, 95th percentile 10.5632809966°, and maximum 10.8858049475°. The minimum equal-norm chordal gap is 0.1287276804. All 20 unique scale–seed trajectory summaries exceed the preregistered 5° threshold; six are longitudinal extensions, not independent confirmation replicates. These dependent measurements quantify an encountered spectral-direction gap. They do not estimate a population effect or a training penalty.

G.7 A multiplier-invariant offline diagnostic

For the Frobenius-normalized field in (82), the Euclidean exterior norm gives

$$\mathcal{P}(s) := \frac{r\|\omega \wedge d\omega\|}{\|q\|_2^2} = \frac{\sqrt{\det \mathrm{Gram}(q, x, w)}}{\|q\|_2^2}. \tag{155}$$

Equation (83) proves invariance under every positive C^1 output multiplier; homogeneity proves invariance under input rescaling. The determinant is evaluated as the volume of the three columns $q/\|q\|_2, x, w/\|q\|_2$ by QR, not by subtracting nearly equal terms in a Gram determinant.

A separate hash-gated replay of the same 26 payloads and 4,872 records gives minimum, median, and maximum Jordan values 0.6504879075, 0.7832722345, and 1.1113389618. It also evaluates the two-term folded control $h(t) = 2t - t^3$. That control's ordinary input-scale-invariant curl has minimum 0.0334446, while its maximum $\mathcal{P}$ is 1.85×10^{-16}; the linear-control maximum is 9.59×10^{-32}. Thus the replay distinguishes ordinary curl that an integrating factor removes from the Jordan directional obstruction.

The same evidentiary restrictions as above apply. These are dependent float64 measurements of a formula, not interval enclosures or causal training effects. A zero value is only necessary for a local integrating factor in this use; it does not certify global convexity. Every three-form vanishes in dimension two, and spectra with at most two distinct positive coordinate values can also be blind spots. The replay found no such Jordan failure case, but theorem 30 does not assert pointwise positivity of $\mathcal{P}$ on every possible spectrum.

Monotone filters on the same stored inputs. An exploratory counterfactual replay evaluates $g(t) = (3t - t^3)/2$, $g^{\circ 2}$, $f(t) = (15t - 10t^3 + 3t^5)/8$, and $f^{\circ 5}$ after one epsilon-free Frobenius normalization. These four filters were fixed before this replay; all are reported, without a fitted threshold. The inputs are the same 4,728 unique, dependent spectra generated by Jordan trajectories, not trajectories trained with these monotone filters. Corollary 32 gives an exactly zero isotonic angle for all four.

For g, $w = 3q - 3x$ makes $\mathcal{P}$ exactly zero; its recorded values are numerical residuals, not obstructions. For the other three filters, the global nonexistence result follows from corollary 32, not from a numerical cutoff. This non-preregistered replay illustrates what order preservation misses on a common input set. It establishes no training effect, practical penalty, or independent replication. The empirical artifact contains the replay script and its complete per-record output.

G.8 Circulation quantitatively separates a field from potentials

The local curl obstruction also has a direct metric consequence that does not require convexity.

Proposition 53 (Loop-circulation separation)**.** *Let Ω be open in a finite-dimensional real inner-product space E, let $Q : \Omega \to E$ be continuous, and let $\Gamma : [0,1] \to \Omega$ be an absolutely continuous closed loop of positive length*

$$L(\Gamma) = \int_0^1 \|\Gamma'(t)\|\, dt.$$

Then every $\varphi \in C^1(\Omega)$ satisfies

$$\sup_{t\in[0,1]} \|Q(\Gamma(t)) - \nabla\varphi(\Gamma(t))\| \geq \frac{\left|\int_0^1 \langle Q(\Gamma(t)), \Gamma'(t)\rangle\, dt\right|}{L(\Gamma)}. \tag{156}$$

The same statement holds for any continuous conservative comparison field whose line integral around Γ is zero.

Proof. The chain rule for an absolutely continuous curve and a C^1 function gives

$$\int_0^1 \langle \nabla\varphi(\Gamma(t)), \Gamma'(t)\rangle\, dt = \varphi(\Gamma(1)) - \varphi(\Gamma(0)) = 0,$$

where the last equality uses $\Gamma(1) = \Gamma(0)$. Subtracting this zero integral from the circulation of Q, applying the triangle inequality and then the definition of the supremum yields

$$\begin{aligned}\left|\int_0^1 \langle Q(\Gamma(t)), \Gamma'(t)\rangle\, dt\right| &\leq \int_0^1 \|Q(\Gamma(t)) - \nabla\varphi(\Gamma(t))\|\, \|\Gamma'(t)\|\, dt \\ &\leq L(\Gamma) \sup_t \|Q(\Gamma(t)) - \nabla\varphi(\Gamma(t))\|.\end{aligned}$$

Division by the positive length proves (156). □

Remark 54 (Evidentiary scope)**.** Proposition 53 is deterministic and applies to any fully specified loop. The supplemental diagnostic `experiments/geometry/results/31_integrability_diagnostic.json` verifies 26 scientific raw-spectrum payloads and evaluates 4,872 observations after an independent singular-value replay. Its square loops give nonzero circulation for the Jordan field and zero circulation, to numerical tolerance, for exact polar, an affine Frobenius-gradient field, and a coordinatewise smooth-polar field. The replay also records the corresponding normalized lower bound $|\oint Q \cdot ds|/L(\Gamma)$ from Proposition 53. This is an analytic-formula float64 diagnostic of the exact-real model abstraction, not a theorem check, low-precision backend claim, or training-performance result.

G.9 Dependency and gap audit

The direct necessity proof in appendix B.4 uses convex-subgradient monotonicity, the intermediate value theorem, homogeneity, the quotient rule, Rademacher differentiation on the ambient cone, and the zero weak-gradient theorem. It does not use inertia, punctured connectivity, or exclusion of simultaneous-critical spectra. Regular compatibility additionally uses the Poincaré lemma and implicit differentiation; the local inertia identity uses Sylvester's law with explicit congruences, never D^{-1}. The full-matrix construction uses weighted Cauchy–Schwarz, convexity under absolute symmetrization, and the standard convex spectral-function theorem. These are classical tools.

Within the stated scalar, normalizer, boundary, and convex-selection hypotheses, no unstated lemma is needed for the proof chain above. Principal-branch necessity now holds without critical-set assumptions, including in rank two. A convex absolutely symmetric matrix lift still requires the additional existence, sign, boundary-extension, and spectral-function hypotheses of the full-matrix theorem. Necessity alone does not supply a unique regular root when the scalar derivative has flat intervals.