hexagon:2611.00548v1

THE MATSAEV (1971) CONJECTURE IS FALSE FOR EVERY $p \neq 2$, AND OTHER MISCELLANEA

DAVE GOEL

ABSTRACT. The Matsaev (1971) conjecture asserts that for $1 < p < \infty$, every contraction T on an L^p space and every polynomial P satisfy $\|P(T)\| \le \|P(S)\|$, where S is the shift on $\ell^p(\mathbb{N})$. We prove that it fails for every $p \in (1,\infty) \setminus \{2\}$, already for a real 2×2 matrix and a real quadratic, without numerical computation, and that it holds for every quadratic whose constant and leading coefficients have equal modulus; these results and their proofs also form the short paper "The Matsaev (1971) conjecture is false for every $p \neq 2$". This expanded version adds an explicit quadratic and an explicit family of scaled rotations that violate the conjecture on a whole punctured neighbourhood of $p = 2$, with a proof by hand; an exact rational re-proof of Drury's counterexample at $p = 4$; upper bounds, certified in exact rational arithmetic, for the shift norms of Drury's polynomial at $p = 3, 4, 6, \dots, 20$; explicit counterexamples, certified by computer, for every $p \in [40/39, 40]$, $p \neq 2$; and conditions that every 2×2 counterexample must satisfy.[1]

This is the expanded version of the paper "The Matsaev (1971) conjecture is false for every $p \neq 2$" [8].

1. INTRODUCTION

Let $1 < p < \infty$, let S be the right shift on $\ell^p(\mathbb{N})$, $\mathbb{N} = \{0, 1, 2, \dots\}$, given by $(Sx)_n = x_{n-1}$ with $x_{-1} = 0$, and let T be a contraction on an L^p space ($\|T\| \le 1$). The conjecture of V. I. Matsaev appears in Nikol'skii's 1971 list of problems [13, Problem V]; some authors date it to 1966 [17]. It states that

$$\|P(T)\|_{L^p \to L^p} \;\le\; \|P(S)\|_{\ell^p \to \ell^p} \qquad \text{for every polynomial } P(z) = \textstyle\sum_{k=0}^{\deg P} c_k z^k,\ c_k \in \mathbb{C}. \tag{1}$$

We show that (1) fails for every $p \in (1,\infty)$ other than 2, with T a real 2×2 matrix and P a real quadratic, and that it holds for every quadratic P with $|c_0| = |c_2|$ (Figure 1). Part II of this paper adds explicit counterexamples, some certified by computer, upper bounds for shift norms certified in exact arithmetic, and necessary conditions on counterexamples (see the table below).

Notation. We write ℓ^p_2 for $\mathbb{C}^2$ (or $\mathbb{R}^2$) with the norm $\|\xi\|_p = (|\xi_1|^p + |\xi_2|^p)^{1/p}$ (and $\|\xi\|_\infty = \max(|\xi_1|, |\xi_2|)$); it is the L^p space of a measure space with two atoms of mass one. For $p = 1$ and $p = \infty$, S is the right shift on $\ell^p(\mathbb{N})$ given by the same formula as above. For an operator A on an L^p space (in particular on ℓ^p or ℓ^p_2), $\|A\|_p$ denotes its operator norm (so (1) reads $\|P(T)\|_p \le \|P(S)\|_p$), and I is the identity; ℓ^p stands for $\ell^p(\mathbb{N})$. We also write $J(\xi_1, \xi_2) = (\xi_2, -\xi_1)$, the clockwise rotation by a right angle, and $R_\theta = \exp(\theta J) = \cos\theta\, I + \sin\theta\, J$ (as $J^2 = -I$), the clockwise rotation by the angle θ.

Part I is the short paper [8]: the background, Theorems A and B and the idea of their proof below, and Sections 2–5.

Date: First version: October 28, 2026. This version: November 16, 2026.

2020 *Mathematics Subject Classification.* Primary 47A30; Secondary 47A60, 47B37, 46E30.

Key words and phrases. Matsaev's conjecture, von Neumann's inequality, L^p spaces, contractions, shift operator, Bellman function.

[1] Additionally, the results of Part I are proved in Lean 4 from Lean's three standard axioms alone, and checked with `lake comparator`: https://github.com/david44hello/papers_supporting_materials/tree/matsaev-1971-conjecture-v2.00/matsaev-1971-conjecture

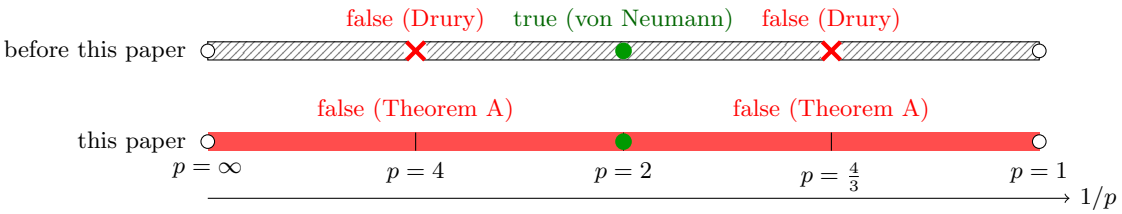


FIGURE 1. Matsaev's conjecture for each $p \in (1,\infty)$, placed by $1/p$, so that p and its dual exponent $p' = p/(p-1)$ lie symmetrically about $p = 2$. Top, before this paper: true for $p = 2$ (von Neumann's inequality), false for $p = 4$ and $p = 4/3$ (Drury [6]), and not known otherwise (hatched). Bottom, this paper: false for every $p \neq 2$ (Theorem A). The endpoints $p = \infty$ and $p = 1$ (open circles) lie outside the conjecture; there (1) holds (see the Background).

Background. For $p = 2$, inequality (1) is von Neumann's inequality. For every $p \in [1,\infty]$ and every contraction T, the triangle inequality and $\|T^k\|_p \le 1$ give the *trivial bound* $\|P(T)\|_p \le \sum_k |c_k|$. For $p = 1$ and $p = \infty$ also $\|P(S)\|_p = \sum_k |c_k|$, so (1) holds trivially: for $p = 1$ apply $P(S)$ to the unit vector $e_0 = (1,0,0,\dots)$; for $p = \infty$, with $d = \deg P$, take $x_n = \overline{c_{d-n}}/|c_{d-n}|$ for $0 \le n \le d$ with $c_{d-n} \neq 0$ and $x_n = 0$ otherwise, so that $(P(S)x)_d = \sum_k |c_k|$. The trivial bound proves (1), for a given p, for every polynomial with $\|P(S)\|_p = \sum_k |c_k|$. Every polynomial of degree at most one is of this kind. To see this, let $P(z) = c_0 + c_1 z$ and choose $\omega, \mu \in \mathbb{C}$ with $|\omega| = |\mu| = 1$, $\omega c_0 = |c_0|$ and $\omega\mu c_1 = |c_1|$. The operator $\omega P(\mu S)$ has the same norm as $P(S)$, because $\mu S = USU^{-1}$ with the isometry $(Ux)_n = \mu^n x_n$. For $x = (1,\dots,1,0,0,\dots)$ with N ones, $\omega P(\mu S)x = (|c_0|,\ |c_0|+|c_1|,\dots,|c_0|+|c_1|,\ |c_1|,0,0,\dots)$, with $N-1$ entries equal to $|c_0|+|c_1|$; so $\|\omega P(\mu S)x\|_p/\|x\|_p \to |c_0|+|c_1|$ as $N \to \infty$, and, with the trivial bound, $\|P(S)\|_p = |c_0|+|c_1|$.

For $p \neq 2$, (1) is known to hold for positive contractions ($Tx \ge 0$ whenever $x \ge 0$), by the dilation theorem of Akcoglu and Sucheston [1] combined with the transference principle of Coifman, Rochberg and Weiss [4], and, more generally, for contractions T with a positive contractive majorant (a positive contraction T_+ with $|Tx| \le T_+|x|$ for every x), because such contractions have isometric dilations (Peller [16]; see [2, Sec. 1]). This class contains the isometries, the Lamperti contractions (which map functions with disjoint supports to functions with disjoint supports) and the absolute contractions (contractions on both L^1 and L^∞); for the last two see also Peller [14, 15].

Drury [6] found a real 2×2 matrix and $P(z) = 1 + 2z - \frac{22}{5}z^2$ (of the form $1 + az - Bz^2$ used below, with $a = 2$, $B = \frac{22}{5}$) that violate (1) for $p = 4$, hence, by duality (Lemma 2.3), for $p = 4/3$; the norms $\|P(T)\|_4$ and $\|P(S)\|_4$ were bounded with computer assistance. Fackler [7, Sec. 2.3] noted that an analytic approach was missing, and, to our knowledge, no counterexample other than Drury's was known (see [2, Sec. 1] and [7, Sec. 2.3]). For polynomials in several commuting contractions, the analogue of (1) is known to fail for every $1 < p < \infty$ once the number of contractions is large enough (Ray [17]); this does not decide the one-variable conjecture studied here.

Theorem A. *Let $p \in (1,\infty)$, $p \neq 2$, $B > 1$, and $P_a(z) = 1 + az - Bz^2$. For every sufficiently small $a > 0$ (how small depends only on p and B) there is a real 2×2 matrix T with $\|T\|_p \le 1$ on ℓ_2^p such that*

$$\|P_a(T)\|_p \;>\; \|P_a(S)\|_p.$$

For $p > 2$ one can take $T = e^{-\varepsilon D_*/2} R_{\pi/2+\varepsilon/2}$, for every sufficiently small $\varepsilon > 0$ (depending on p, B and a), where $D_* = \max_{0\le t\le 1}(t - t^{p-1})/(1+t^p) > 0$ depends only on p (Lemma 3.4). For $1 < p < 2$ one can take the transpose of the matrix above, built for the dual exponent $p' = p/(p-1) > 2$ in place

of p; since the transpose of R_θ is $R_{-\theta}$, this is $T = e^{-\varepsilon D'_*/2} R_{-\pi/2-\varepsilon/2}$, where D'_* is D_* computed for p': again a scaled rotation by slightly more than a right angle, now counterclockwise.

Theorem B. *Let $P(z) = c_0 + c_1 z + c_2 z^2$ with complex coefficients and $|c_0| = |c_2|$. Then $\|P(T)\|_p \le \|P(S)\|_p$ for every $1 \le p \le \infty$ and every contraction T on every L^p space.*

Idea of the proof of Theorem A. Let $(x_n)_{n\in\mathbb{Z}}$ be a two-sided sequence with $x_{n+2} = -x_n$ for all n, and shift it by the same formula, $(Sx)_n = x_{n-1}$. Then $((Sx)_n, (Sx)_{n-1}) = (x_{n-1}, x_{n-2}) = (x_{n-1}, -x_n) = J(x_n, x_{n-1})$: on pairs of consecutive terms of such sequences, S acts as J. Truncating long pieces of such sequences gives $\|P(J)\|_p \le \|P(S)\|_p$ for every P (Lemma 2.2). For $P_0(z) = 1 - Bz^2$ we have $P_0(J) = I - BJ^2 = (1+B)I$, so $1 + B = \|P_0(J)\|_p \le \|P_0(S)\|_p \le 1 + B$ (Lemma 2.2 and the triangle inequality); thus $\|P_0(S)\|_p = 1 + B$, and $P_0(S)$ has many almost-maximising vectors: long truncations x of every sequence with $x_{n+2} = -x_n$ give $\|P_0(S)x\|_p / \|x\|_p$ close to $1 + B$.

For $p \neq 2$ and small $a > 0$ we prove that $P_a(S)$ still has exactly the norm $C_a = \|P_a(J)\|_p$ (Theorem 3.6). For $p > 2$ the upper bound comes from a Bellman function, a function G of two real variables with $|(P_a(S)x)_n|^p \le C_a^p |x_n|^p + G(x_{n-1}, x_{n-2}) - G(x_n, x_{n-1})$ for every real $x \in \ell^p(\mathbb{N})$ (with $x_{-1} = x_{-2} = 0$) and every $n \ge 0$; in the sum over n the G terms cancel, so $\|P_a(S)\|_p \le C_a$ (Lemma 3.1). For $1 < p < 2$ the upper bound follows by duality (Lemma 2.3).

For $p > 2$ we compose J with the rotation $R_{\varepsilon/2}$ and multiply by $e^{-\varepsilon D_*/2}$; this gives the matrix $T_\varepsilon = e^{-\varepsilon D_*/2} R_{\varepsilon/2} J = e^{-\varepsilon D_*/2} R_{\pi/2+\varepsilon/2}$ of Theorem A. The factor $e^{-\varepsilon D_*/2}$ compensates the largest possible rate of growth of the ℓ^p norm under the rotations $R_{\varepsilon/2}$ (at most $D_*/2$ per unit of ε), so that T_ε is still a contraction (Figure 3). To see this, put $\phi(y) = \operatorname{sgn} y\, |y|^{p-1}$ (so that $\frac{d}{dy}|y|^p = p\,\phi(y)$) and, for real $\xi \neq 0$,

$$Q(\xi) = \frac{\phi(\xi_1)\xi_2 - \phi(\xi_2)\xi_1}{\|\xi\|_p^p}, \qquad \text{so that} \qquad Q(1,t) = \frac{t - t^{p-1}}{1 + t^p} \quad (0 \le t \le 1).$$

Since $\frac{d}{d\varepsilon} R_{\varepsilon/2}\xi\big|_{\varepsilon=0} = \frac{1}{2} J\xi = \frac{1}{2}(\xi_2, -\xi_1)$, the chain rule in each coordinate gives

$$\frac{d}{d\varepsilon} \|R_{\varepsilon/2}\xi\|_p^p \Big|_{\varepsilon=0} = \frac{p}{2}\big(\phi(\xi_1)\xi_2 - \phi(\xi_2)\xi_1\big) = \frac{p}{2}\, Q(\xi) \|\xi\|_p^p,$$

hence, dividing by $p\|\xi\|_p^{p-1}$, $\frac{d}{d\varepsilon}\|R_{\varepsilon/2}\xi\|_p\big|_{\varepsilon=0} = \frac{1}{2}\, Q(\xi)\|\xi\|_p$. Lemma 3.4 shows that the maximum of Q over all $\xi \neq 0$ is attained at $\xi = (1, t_*)$ for some $t_* \in (0,1)$; so $\max Q = D_*$, and the derivative above is at most $\frac{D_*}{2}\|\xi\|_p$, with equality at $\xi = (1, t_*)$. As $R_{(\varepsilon+h)/2} = R_{h/2} R_{\varepsilon/2}$, the same bound holds at every ε, with ξ replaced by $R_{\varepsilon/2}\xi$. Thus $\frac{d}{d\varepsilon} \log \|R_{\varepsilon/2}\xi\|_p \le \frac{D_*}{2}$, and integration from 0 to ε gives $\|R_{\varepsilon/2}\xi\|_p \le e^{\varepsilon D_*/2}\|\xi\|_p$ for $\varepsilon \ge 0$. As J is an isometry of ℓ_2^p (it permutes the coordinates up to sign), $\|T_\varepsilon \xi\|_p = e^{-\varepsilon D_*/2}\|R_{\varepsilon/2} J\xi\|_p \le \|J\xi\|_p = \|\xi\|_p$; so T_ε is a contraction on real vectors, and hence, by Lemma 2.1, on ℓ_2^p (see also Lemma 4.1).

Let ξ_a be a real vector at which $\|P_a(J)\xi\|_p / \|\xi\|_p$ is maximal, normalised by $(\xi_a)_1 = 1$ and close to $(1, t_*)$ (such a vector exists by Lemma 3.5). The derivative in ε at $\varepsilon = 0$ of $\|P_a(T_\varepsilon)\xi_a\|_p^p$ is

$$\tfrac{1}{2}\, p\,(p-1)(1+B)^{p-2} E_*\, a\,(B-1) + o(a) \quad (a \to 0), \qquad E_* = \frac{t_*^{p-2}(1+t_*^2)^2}{1 + t_*^p} > 0$$

(Lemma 4.1). At $\varepsilon = 0$ we have $T_0 = J$, so $\|P_a(T_0)\xi_a\|_p = \|P_a(J)\|_p \|\xi_a\|_p$ by the choice of ξ_a. Hence, for $B > 1$ and every sufficiently small $a > 0$, we get $\|P_a(T_\varepsilon)\xi_a\|_p > \|P_a(J)\|_p \|\xi_a\|_p$, and so $\|P_a(T_\varepsilon)\|_p > \|P_a(J)\|_p$, for all sufficiently small $\varepsilon > 0$; how small ε must be depends on a. The gain $\|P_a(T_\varepsilon)\|_p - \|P_a(J)\|_p$ is positive, but it is small: it tends to 0 as $\varepsilon \to 0$, and the argument gives no explicit lower bound for it. Therefore an upper bound $\|P_a(S)\|_p \le \|P_a(J)\|_p + r$ with an error term $r > 0$ would prove nothing unless r were shown to be smaller than this gain. This is why we prove the exact equality $\|P_a(S)\|_p = \|P_a(J)\|_p$. For $B = 1$ the gain vanishes to first order, and indeed for $B = 1$ inequality (1) holds (Theorem B).

Further results. The table lists the results of Part II next to the best results known before.

Best result before this paper	**New result**
No result near $p = 2$.	The quadratic $1+2z-\frac{6}{5}z^2$, with the scaled rotations $R_{\theta_0}/\|R_{\theta_0}\|_p$, $\cos\theta_0 = -\frac{1}{12}$, violates (1) for every $p \neq 2$ close to 2; proof by hand (Section 6)
A computer verification over many patches of the unit sphere [6].	A rational certificate of Drury's counterexample at $p = 4$, verified in exact arithmetic (Section 7)
$\|P(S)\|_4 < 6.5278$, and the guess ≈ 6.4548 [6].	Upper bounds, certified in exact rational arithmetic, for $\|P(S)\|_p$, $P = 1+2z-\frac{22}{5}z^2$, at $p = 3, 4, 6, \ldots, 20$, within $2 \cdot 10^{-6}$ of $\|P(J)\|_p$ for even p (Section 8)
Drury's explicit pair, for $p = 4$ and $p = 4/3$ only [6].	Explicit counterexamples for every $p \in [40/39, 40]$, $p \neq 2$: four explicit real quadratics and explicit real 2×2 contractions, certified by computer (Section 9)
Known classes of contractions that satisfy (1) [16, 2].	Elementary necessary conditions on a 2×2 counterexample (Section 10)

Programs, certificates and logs for Part II are in the ancillary files of this paper: the archive `matsaev_1971_ancillary_files_v1.10.zip` (top folder `ancillary_files/`) in its Zenodo record, https://doi.org/10.5281/zenodo.23691816

2. Basic facts

As in the Introduction, for a real number y put $\phi(y) = \operatorname{sgn} y\,|y|^{p-1}$, so that $\frac{d}{dy}|y|^p = p\,\phi(y)$ for $p > 1$; on real $\xi \in \ell_2^p$, ϕ and $|\cdot|^{p-2}$ act coordinatewise, and a product of two vectors, such as $|\xi|^{p-2}\psi$, is taken coordinatewise. With $\langle \xi, \psi\rangle = \xi_1\psi_1 + \xi_2\psi_2$, $\frac{d}{dh}\|\xi + h\psi\|_p^p\big|_{h=0} = p\langle\phi(\xi), \psi\rangle$ for real ξ, ψ and $p > 1$.

Lemma 2.1 (real and complex norms)**.** *Let $1 \le p < \infty$. A real 2×2 matrix has the same norm on the real and on the complex ℓ_2^p, and for a real polynomial P the operator $P(S)$ has the same norm on the real and on the complex ℓ^p.*

Proof. This is the Marcinkiewicz–Zygmund theorem [11] (see also [9]). For a complex number $z = y_1 + iy_2$ we have $y_1\cos\theta + y_2\sin\theta = |z|\cos(\theta - \arg z)$, so $\int_0^{2\pi}|y_1\cos\theta + y_2\sin\theta|^p\,d\theta = \beta_p|z|^p$ with $\beta_p = \int_0^{2\pi}|\cos\theta|^p\,d\theta$. Let A be a real matrix (a 2×2 matrix, or the infinite matrix of $P(S)$), and let $x = u + iv$ with real vectors u, v. Applying this identity to each coordinate of $Ax = Au + iAv$ gives $\|Ax\|_p^p = \beta_p^{-1}\int_0^{2\pi}\|A(u\cos\theta + v\sin\theta)\|_p^p\,d\theta$, and the same identity for $A = I$ gives $\|x\|_p^p = \beta_p^{-1}\int_0^{2\pi}\|u\cos\theta + v\sin\theta\|_p^p\,d\theta$. Write $\|A\|_{\mathrm{real}}$ for the norm of A on real vectors. Since $\|A(u\cos\theta + v\sin\theta)\|_p \le \|A\|_{\mathrm{real}}\|u\cos\theta + v\sin\theta\|_p$ for every θ, we get $\|Ax\|_p \le \|A\|_{\mathrm{real}}\|x\|_p$: the complex norm of A is at most its real norm. The reverse inequality is clear. □

Lemma 2.2 (a lower bound for $\|P(S)\|_p$)**.** *For every polynomial P and $1 \le p \le \infty$, $\|P(J)\|_p \le \|P(S)\|_p$.*

Proof. Let $\xi \in \mathbb{C}^2$ and let x be the sequence with $(x_n, x_{n-1}) = J^{-n}\xi$ for $n \in \mathbb{Z}$. This is consistent: if $J^{-n-1}\xi = (b_1, b_2)$, then $J^{-n}\xi = J(b_1, b_2) = (b_2, -b_1)$, so both rules give $x_n = b_2$. Then $x_{n+2} = -x_n$ (as $J^{-2} = -I$), and with $X_n = (x_n, x_{n-1})$ we have $X_{n-k} = J^kX_n$, so $y_n := \sum_k c_kx_{n-k} = (P(J)X_n)_1$. Since $X_{n+1} = J^{-1}X_n = -JX_n$ and $P(J)$ commutes with J, $(P(J)X_{n+1})_1 = -(P(J)X_n)_2$. Hence, with $Y_n = P(J)X_n$ (so $Y_{n+1} = -JY_n$ and $Y_{n+2} = -Y_n$), the numbers $|y_n|, |y_{n+1}|, |y_{n+2}|, |y_{n+3}|$ are $|(Y_n)_1|, |(Y_n)_2|, |(Y_n)_1|, |(Y_n)_2|$; since $\|Y_n\|_p = \|J^{-n}P(J)\xi\|_p = \|P(J)\xi\|_p$ (J is an isometry of ℓ_2^p), over any four consecutive indices $\sum|y_n|^p = 2\|P(J)\xi\|_p^p$, and likewise $\sum|x_n|^p = 2\|\xi\|_p^p$. For $j \ge 1$ let $x^{(j)} \in \ell^p(\mathbb{N})$ agree with x for $0 \le n < 4j$ and vanish elsewhere. Then $\|x^{(j)}\|_p^p = 2j\|\xi\|_p^p$, and $(P(S)x^{(j)})_n = y_n$ for $\deg P \le n < 4j$, while $(P(S)x^{(j)})_n = 0$ for $n \ge 4j + \deg P$; so $\|P(S)x^{(j)}\|_p^p = 2j\|P(J)\xi\|_p^p + O(1)$, and the ratio $\|P(S)x^{(j)}\|_p^p/\|x^{(j)}\|_p^p$ tends to $\|P(J)\xi\|_p^p/\|\xi\|_p^p$ as $j \to \infty$ (the sums $\sum_{n\ge0}|(P(S)x^{(j)})_n|^p$ and $\sum_{0\le n<4j}|y_n|^p = 2j\|P(J)\xi\|_p^p$ differ

only in the at most $2\deg P$ terms with $n < \deg P$ or $4j \le n < 4j + \deg P$, and in these terms $|(P(S)x^{(j)})_n|$ and $|y_n|$ are at most $\sum_k |c_k| \max_m |x_m|$). Hence $\|P(S)\|_p \ge \|P(J)\xi\|_p / \|\xi\|_p$ for every $\xi \neq 0$, which is the claim. For $p = \infty$ the inequality is trivial: $\|P(J)\|_\infty \le \sum |c_k| = \|P(S)\|_\infty$. □

For $1 < p < \infty$, Lemma 2.2 also follows from Peller's results [16], as J is an isometry of ℓ_2^p; we give the direct proof because its sequences with $x_{n+2} = -x_n$ reappear in Section 3 (Lemma 3.2).

Lemma 2.3 (duality)**.** *For a polynomial P and $1 < p < \infty$, $\|P(S)\|_p = \|P(S)\|_{p'}$, $p' = p/(p-1)$. If T is a real 2×2 matrix with transpose T^{t}, then $\|P(T^{\mathsf{t}})\|_p = \|P(T)\|_{p'}$ and $\|T^{\mathsf{t}}\|_p = \|T\|_{p'}$ (norms on ℓ_2^p and on $\ell_2^{p'}$).*

Proof. Let $\mathcal{C}$ be convolution by (c_k) on $\ell^p(\mathbb{Z})$, $(\mathcal{C}x)_n = \sum_k c_k x_{n-k}$. Then $\|P(S)\|_p = \|\mathcal{C}\|_p$: $P(S)$ is the restriction of $\mathcal{C}$ to the sequences that vanish at negative indices, so $\|P(S)\|_p \le \|\mathcal{C}\|_p$; conversely, a finitely supported $x \in \ell^p(\mathbb{Z})$ can be translated into $\ell^p(\mathbb{N})$, translations are isometries that commute with $\mathcal{C}$, and finitely supported vectors are dense, so $\|\mathcal{C}\|_p \le \|P(S)\|_p$. The adjoint of $\mathcal{C}$, acting on $\ell^{p'}(\mathbb{Z})$, is convolution by (c_{-k}), which the reflection $n \mapsto -n$ (an isometry) turns back into convolution by (c_k); hence $\|\mathcal{C}\|_p = \|\mathcal{C}\|_{p'}$, and the same holds for $P(S)$. The statements about matrices follow from $\|A^{\mathsf{t}}\|_p = \|A\|_{p'}$ (duality) and $P(T)^{\mathsf{t}} = P(T^{\mathsf{t}})$. □

3. The norm of $P_a(S)$ for small a

Fix $p > 2$, $B > 0$, and $P_a(z) = 1 + az - Bz^2$; put $C_a := \|P_a(J)\|_p$ (by Lemma 2.1, a maximum over real vectors). For a real sequence x, the *window* at n is $(u, v, w) = (x_n, x_{n-1}, x_{n-2})$; then $(P_a(S)x)_n = u + av - Bw$.

Lemma 3.1 (a Bellman-function bound for $\|P_a(S)\|_p$)**.** *Let $C > 0$, and let $G : \mathbb{R}^2 \to \mathbb{R}$ be continuous with $G(0,0) = 0$. If*

$$F(u,v,w) := C^p|u|^p + G(v,w) - G(u,v) - |u + av - Bw|^p \ \ge\ 0 \qquad \text{for all } (u,v,w) \in \mathbb{R}^3, \tag{2}$$

then $\|P_a(S)\|_p \le C$. We call such a G a Bellman function *for P_a and C.*

Proof. For a finitely supported real $x \in \ell^p(\mathbb{N})$ (with $x_{-1} = x_{-2} = 0$, and $x_n = 0$ for $n > n_0$) sum (2) over the windows at $0 \le n \le n_0 + 2$: the G terms telescope to $G(0,0) - G(0,0) = 0$, so $\|P_a(S)x\|_p^p \le C^p \|x\|_p^p$. By density this holds for all real $x \in \ell^p(\mathbb{N})$, and by Lemma 2.1 also for complex x. □

Lemma 3.2 (the case $a = 0$)**.** *With $G_0(v,w) = B(1+B)^{p-1}(|v|^p + |w|^p)$ and $C = 1 + B$, (2) holds for $a = 0$: $F_0(u,v,w) = (1+B)^{p-1}\big(|u|^p + B|w|^p\big) - |u - Bw|^p \ge 0$, with equality exactly on the plane $\Pi = \{w = -u\}$, the set of windows of the sequences with $x_{n+2} = -x_n$.*

Proof. The identities $G_0(v,w) - G_0(u,v) = B(1+B)^{p-1}(|w|^p - |u|^p)$ and $(1+B)^p - B(1+B)^{p-1} = (1+B)^{p-1}$ give the formula for F_0. Write $u - Bw = (1+B)\big(\frac{1}{1+B}u + \frac{B}{1+B}(-w)\big)$; strict convexity of $|\cdot|^p$ ($p > 1$) gives $|u - Bw|^p \le (1+B)^{p-1}(|u|^p + B|w|^p)$, with equality if and only if $u = -w$. □

Lemma 3.3 (growth of F_0 away from Π)**.** *With $\nu = u + w$, there is $\gamma_1 = \gamma_1(p, B) > 0$ with $F_0(u,v,w) \ge \gamma_1(|u| + |\nu|)^{p-2}\nu^2$ for all real u, v, w.*

Proof. The function F_0 does not depend on v, and both sides are homogeneous of degree p in (u, ν), so it suffices to consider $\mathcal{K} = \{|u| + |\nu| = 1\}$. On $\mathcal{K}$ with $\nu \neq 0$, $F_0 > 0$ by Lemma 3.2. Near $\nu = 0$ on $\mathcal{K}$, $|u|$ is close to 1, and the bases of the three p-th powers in F_0 (the numbers y in its terms $|y|^p$), namely u, $w = \nu - u$ and $u - Bw = (1+B)u - B\nu$, are close to $\pm 1, \mp 1, \pm(1+B)$; so F_0 is real-analytic there, $F_0 = 0$ at $\nu = 0$, $\partial_\nu F_0 = pB(1+B)^{p-1}\phi(\nu - u) + pB\,\phi\big((1+B)u - B\nu\big)$, which at $\nu = 0$ is $-pB(1+B)^{p-1}\phi(u) + pB(1+B)^{p-1}\phi(u) = 0$, and $\partial_\nu^2 F_0 = p(p-1)B(1+B)^{p-2}|u|^{p-2} > 0$ at $\nu = 0$. Hence F_0/ν^2 extends to a positive continuous function on $\mathcal{K}$, and γ_1 is its minimum. □

The function Q. As in the Introduction, for $\xi \in \mathbb{R}^2 \setminus \{0\}$ put $Q(\xi) = \langle \phi(\xi), J\xi\rangle / \|\xi\|_p^p = \big(\phi(\xi_1)\xi_2 - \phi(\xi_2)\xi_1\big)/\|\xi\|_p^p$. Since $J^2 = -I$, $P_a(J) = (1+B)I + aJ$, so $\|P_a(J)\xi\|_p^p = (1+B)^p \|\xi + hJ\xi\|_p^p$ with $h = a/(1+B)$. The derivative formula of Section 2 gives the term of first order in h, namely $hp\langle \phi(\xi), J\xi\rangle = hp\, Q(\xi)\|\xi\|_p^p$; the second derivative in h, namely $p(p-1)\sum_{i=1}^2 |\xi_i + h(J\xi)_i|^{p-2}(J\xi)_i^2$, is bounded for ξ on the Euclidean unit circle and $|h| \le 1$ because $p > 2$. Taylor's formula in h, and homogeneity in ξ, therefore give

$$\|P_a(J)\xi\|_p^p = (1+B)^p \|\xi\|_p^p \Big(1 + \tfrac{pa}{1+B}\, Q(\xi)\Big) + O(a^2)\|\xi\|_p^p, \tag{3}$$

so, to first order in a, the vectors ξ that maximise $\|P_a(J)\xi\|_p / \|\xi\|_p$ are the maximisers of Q. For real σ put $f(\sigma) = Q(1,\sigma) = \sigma(1 - |\sigma|^{p-2})/(1+|\sigma|^p)$.

Lemma 3.4 (the maximum of Q)**.** *Let $p > 2$. The function f has exactly one critical point t_* in $(0,1)$; it is a nondegenerate maximum, and $D_* := \max_{\xi \ne 0} Q(\xi) = f(t_*) > 0$ is attained exactly on the nonzero multiples of $(1, t_*)$ and of $J(1, t_*)$. Moreover*

$$1 + t_*^{2p-2} = (p-1)\, t_*^{p-2}(1 + t_*^2). \tag{4}$$

Proof. For $0 < t < 1$, $f(t) = (t - t^{p-1})/(1+t^p)$, and by the quotient rule $f'(t) = g(t)/(1+t^p)^2$ with $g(t) = 1 + t^{2p-2} - (p-1)t^{p-2}(1+t^2)$. Next $g'(t) = (p-1)t^{p-3}\chi(t)$ with $\chi(t) = 2t^p - pt^2 - (p-2)$; since $\chi(0) = -(p-2) < 0$ and $\chi'(t) = 2pt\,(t^{p-2} - 1) < 0$ on $(0,1)$, $\chi < 0$ there, so g is strictly decreasing on $(0,1)$, from $g(0^+) = 1$ to $g(1) = 4 - 2p < 0$. Hence g has exactly one zero t_*, which is (4), and $f''(t_*) = g'(t_*)/(1+t_*^p)^2 < 0$. As $f > 0$ on $(0,1)$ and $f(0) = f(1) = 0$, t_* is the only maximiser of f on $[0,1]$. Finally, Q is invariant under $\xi \mapsto J\xi$, $\xi \mapsto -\xi$ and $\xi \mapsto c\xi$ $(c > 0)$. Replacing ξ by $J\xi$ if necessary, we may assume $|\xi_2| \le |\xi_1|$; then ξ is a multiple of $(1, \sigma)$ with $|\sigma| \le 1$, and $Q(\xi) = f(\sigma)$. Since $f(\sigma) \le 0$ for $\sigma \in [-1, 0]$, the maximum of Q is $f(t_*)$, attained exactly at the nonzero multiples of $(1, t_*)$ and of $J(1, t_*)$. □

For example, $t_* = 2\sin(\pi/12)$, $D_* = 2^{-3/2}$ for $p = 4$, and $t_* = (\sqrt{5} - 1)/2$, $D_* = \frac{1}{2}$ for $p = 6$ (Figure 2).

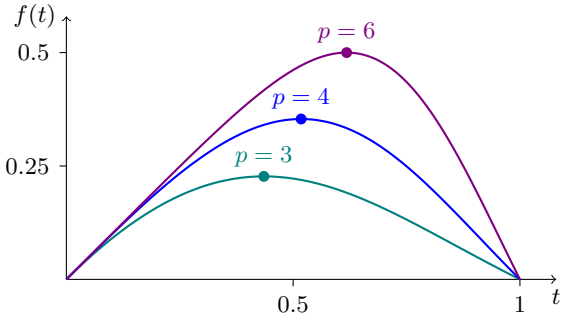


FIGURE 2. The function $f(t) = (t - t^{p-1})/(1+t^p)$ of Lemma 3.4 on $[0,1]$, for $p = 3, 4, 6$. The dot marks its maximum $D_* = f(t_*)$ (for $p = 4$ and $p = 6$, the values given above). Plotted numerically; an illustration, not used in the proofs.

Lemma 3.5 (the maximising vectors of $P_a(J)$)**.** *Let $p > 2$ and $B > 0$. We measure distances on the Euclidean unit circle of $\mathbb{R}^2$ by angle. Let $e_* = (1, t_*)/\|(1, t_*)\|_2$, let $\mathcal{O}_0 = \{\pm e_*, \pm Je_*\}$, and for $s \in \mathbb{R}^2$ put $\Sigma_a(s) = 2\big(C_a^p \|s\|_p^p - \|P_a(J)s\|_p^p\big)$; thus $\Sigma_a \ge 0$ by the definition of C_a. There are $\rho \in (0, \frac{1}{2}\arctan t_*)$, $a_1 > 0$ and $\delta, \eta > 0$ such that for $0 < a < a_1$:*

(i) the function Σ_a vanishes on the unit circle at exactly four points; if s_a denotes the one nearest to e_*, these points are $\mathcal{O}_a = \{\pm s_a, \pm Js_a\}$; they are the maximisers of $\|P_a(J)s\|_p / \|s\|_p$ on the unit circle, and $\mathcal{O}_a$ lies within angular distance $\rho/2$ of $\mathcal{O}_0$;

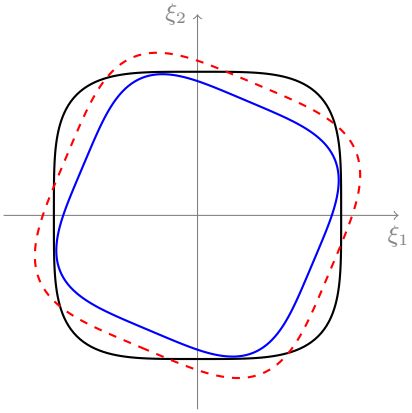


FIGURE 3. The real unit circle of ℓ_2^4 (black); its image under the rotation $R_{\pi/2+\varepsilon/2}$ (red, dashed), which leaves the unit ball; and its image under $T_\varepsilon = e^{-\varepsilon D_*/2} R_{\pi/2+\varepsilon/2}$ (blue), which stays inside it (Lemma 4.1). Here $\varepsilon = 0.8$, larger than in Theorem A, so that the curves are easy to tell apart. Computed numerically; an illustration, not used in the proofs.

(ii) if s is on the unit circle at angular distance at least ρ from $\mathcal{O}_a$, then $\Sigma_a(s) \geq a\delta$;
(iii) if s is on the unit circle at angular distance $r(s) \leq \rho$ from $\mathcal{O}_a$, then $\Sigma_a(s) \geq a\eta\, r(s)^2$;
(iv) with $s_0 := e_*$, the maps $a \mapsto s_a$ and $a \mapsto C_a^p$ are real-analytic on $[0, a_1)$.

Proof. With θ the angle on the unit circle and $\theta_* = \arctan t_*$, we have $Q(\cos\theta, \sin\theta) = f(\tan\theta)$ near θ_*, and $f'(t_*) = 0$ gives $\partial_\theta^2 Q(\theta_*) = f''(t_*)\sec^4\theta_* < 0$; by the invariance of Q under $\xi \mapsto J\xi$ and $\xi \mapsto -\xi$, the same holds at the other three points of $\mathcal{O}_0$. The points $(0, \pm 1)$, $(\pm 1, 0)$ are at angular distance at least $\arctan t_*$ from $\mathcal{O}_0$. Hence we can fix $\rho \in (0, \frac{1}{2}\arctan t_*)$ and $m > 0$ such that $\partial_\theta^2 Q \leq -m$ on the four closed arcs of angular radius 2ρ centred at the points of $\mathcal{O}_0$ (the *2ρ-arcs*); these arcs are disjoint and avoid the coordinate axes. Write

$$\frac{\Sigma_a(s)}{a} = 2\|s\|_p^p\big(c(a) - H(a,s)\big), \qquad \text{where}$$
$$H(a,s) = \frac{1}{a}\Big(\frac{\|P_a(J)s\|_p^p}{\|s\|_p^p} - (1+B)^p\Big), \qquad c(a) = \max_{s\neq 0} H(a,s) = \frac{C_a^p - (1+B)^p}{a}.$$

By (3), $H(a,s) \to H(0,s) := p(1+B)^{p-1}Q(s)$ uniformly on the unit circle, hence $c(a) \to p(1+B)^{p-1}D_*$ and, uniformly on the unit circle,

$$\Sigma_a(s)/a \;\longrightarrow\; \Sigma'(s) := 2p(1+B)^{p-1}\|s\|_p^p\big(D_* - Q(s)\big) \qquad (a \to 0^+).$$

(As $\Sigma_0 = 0$, $\Sigma'(s)$ is the right derivative of $\Sigma_a(s)$ at $a = 0$.) On the 2ρ-arcs both coordinates of s have absolute value at least $\sin(\arctan t_* - 2\rho)$, so for $|a| < \frac{1}{2}(1+B)\sin(\arctan t_* - 2\rho)$ no coordinate of $P_a(J)s = (1+B)s + aJs$ vanishes there. For such a, $\|P_a(J)s\|_p^p/\|s\|_p^p$ is real-analytic in (a,s) and equals $(1+B)^p$ at $a = 0$, so H (with the value $H(0,s)$ above) is real-analytic in (a,s), also at $a = 0$. In particular $H(a,\cdot) \to H(0,\cdot)$ in C^2 on these arcs. By Lemma 3.4, $\Sigma' \geq 0$ vanishes on the unit circle exactly on $\mathcal{O}_0$. Put

$$\delta = \tfrac{1}{2}\min\{\Sigma'(s) :\ \|s\|_2 = 1,\ s \text{ at angular distance} \geq \rho/2 \text{ from } \mathcal{O}_0\} > 0, \qquad m_1 = \tfrac{1}{2}p(1+B)^{p-1}m.$$

By the uniform convergence, and the C^2 convergence with $\partial_\theta^2 H(0,\cdot) = p(1+B)^{p-1}\partial_\theta^2 Q \leq -2m_1$, there is $a_1 > 0$ such that for $0 < a < a_1$: $\Sigma_a/a \geq \delta$ at angular distance $\geq \rho/2$ from $\mathcal{O}_0$, and $\partial_\theta^2 H(a,\cdot) \leq -m_1$ on the closed 2ρ-arcs. Decreasing a_1 if necessary, we may also assume that $a_1 \leq \frac{1}{2}(1+B)\sin(\arctan t_* - 2\rho)$, so that H is real-analytic on $[0, a_1)$ times the 2ρ-arcs.

Proof of (i) and (iv). The maximisers of $H(a,\cdot)$ on the unit circle are the zeros of Σ_a there, so they lie within angular distance $\rho/2$ of $\mathcal{O}_0$; by the invariance under $s \mapsto Js$ and $s \mapsto -s$ there is a maximiser near each point of $\mathcal{O}_0$. As $H(a,\cdot)$ is strictly concave on each 2ρ-arc, the arc around a point $o \in \mathcal{O}_0$ contains exactly one maximiser, namely the only zero of $\partial_\theta H(a,\cdot)$ on that arc. For $a = 0$ this maximiser is o, as $H(0,\cdot) = p(1+B)^{p-1}Q$ (Lemma 3.4). Since $\partial_\theta^2 H(a,\cdot) \le -m_1 < 0$ on the arc for every $a \in [0,a_1)$, the implicit function theorem for $\partial_\theta H = 0$, applied at each $a \in [0,a_1)$, shows that this maximiser is a real-analytic function of $a \in [0,a_1)$ with value o at $a = 0$; for $o = e_*$ it is s_a. Hence $c(a) = H(a,s_a)$ is real-analytic in a, and so is $C_a^p = (1+B)^p + a\,c(a)$.

Proof of (ii). If s is at angular distance $\ge \rho$ from $\mathcal{O}_a$, it is at angular distance $\ge \rho/2$ from $\mathcal{O}_0$, so $\Sigma_a(s) \ge a\delta$.

Proof of (iii). Let o_a be the point of $\mathcal{O}_a$ nearest to s, so that $r(s)$ is the angular distance from s to o_a, and let o be the point of $\mathcal{O}_0$ within $\rho/2$ of o_a. The arc from o_a to s lies within $3\rho/2 < 2\rho$ of o; on this arc $\partial_\theta^2 H(a,\cdot) \le -m_1$, and $\partial_\theta H(a,o_a) = 0$, so $c(a) - H(a,s) \ge \frac12 m_1 r(s)^2$. Hence $\Sigma_a(s) \ge a\eta\, r(s)^2$ with $\eta = m_1 \min_{\|s\|_2=1}\|s\|_p^p$. □

Theorem 3.6 ($\|P_a(S)\|_p = \|P_a(J)\|_p$ for small a)**.** *For every $p > 2$ and $B > 0$ there is $a_0 = a_0(p,B) > 0$ such that $\|P_a(S)\|_p = \|P_a(J)\|_p$ for $0 \le a < a_0$; for $1 < p < 2$ the same holds for $0 \le a < a_0(p',B)$.*

Proof. By Lemma 2.2 only $\|P_a(S)\|_p \le C_a$ needs proof (and only this inequality is used for Theorem A). For $a = 0$ it is Lemma 3.2. For small $a > 0$ we construct $G = G_a$ with $C = C_a$ in Lemma 3.1.

Outline. Write a window as $\hat s + \nu e_3$ with the *state* $s = (u,v)$, $\hat s = (u,v,-u) \in \Pi$, $\nu = u + w$ and $e_3 = (0,0,1)$. At $a = 0$, $F_0 \ge 0$ vanishes exactly on Π and grows like $(|u|+|\nu|)^{p-2}\nu^2$ off it (Lemmas 3.2 and 3.3). First, along a sequence with $x_{n+2} = -x_n$ the states run through the cycle $s, -Js, -s, Js$, and the sum of F (with $C = C_a$) over the four windows does not depend on G: it is $\Sigma_a(s)$. Second, we take $G_a = G_0 + \frac12\Delta_a + k_a$, with Δ_a and k_a defined in Steps 1 and 2, and F_a denotes F with $C = C_a$ and $G = G_a$. Then F_a takes the same value $\frac14\Sigma_a(s)$ at each of the four windows of the cycle; in particular $F_a(\hat s) = \frac14\Sigma_a(s) \ge 0$, which is zero exactly when $s = 0$ or s lies on a ray through a point of $\mathcal{O}_a$, and grows quadratically off these rays (Lemma 3.5). Third, since we need $F_a \ge 0$, the gradient of F_a must vanish wherever F_a vanishes. Along Π this holds automatically. In the direction e_3, across Π, it is achieved by the term k_a, which is of size $O(a)$ and does not change F_a on Π. Fourth, $F_a - F_0$ and its gradient are of size $O(a)$ on bounded sets (Step 5). Away from $\mathcal{O}_a$, $\Sigma_a \ge a\delta$ absorbs this error for small ν, and F_0 absorbs it for large ν; near $\mathcal{O}_a$, F_a is bounded below by a positive quadratic form in the distance to $\mathcal{O}_a$ and ν. Throughout, $J(u,v) = (v,-u)$, and ρ, a_1, δ, η, $\mathcal{O}_a$, s_a are those of Lemma 3.5; the constant a_1 will be decreased finitely many times.

Step 1: the sum over a cycle of windows. For a sequence with $x_{n+2} = -x_n$, if the window at n is $\hat s$, then the window at $n+1$ is $\widehat{-Js}$, and $(v,w) = (v,-u) = Js$ on Π. Thus, with

$$\Delta_a(s) = C_a^p|u|^p - |(1+B)u + av|^p,$$

$F(\hat s) = \Delta_a(s) + G(Js) - G(s)$ on Π, for $C = C_a$. Since Δ_a is even and $-Js, -s, Js$ complete the cycle of s, the G terms telescope around the cycle and

$$\sum_{s' \in \{s,-Js,-s,Js\}} F(\widehat{s'}) = 2\big(\Delta_a(s) + \Delta_a(Js)\big) = \Sigma_a(s), \tag{5}$$

because $\Delta_a(s)+\Delta_a(Js) = C_a^p(|u|^p+|v|^p) - |(1+B)u+av|^p - |(1+B)v-au|^p$ and $P_a(J) = (1+B)I + aJ$.

Step 2: the Bellman function. By Step 1, $F(\hat s) = \frac14\Sigma_a(s)$ for every s if $G(Js) - G(s) = \frac12\big(\Delta_a(Js) - \Delta_a(s)\big)$; the choice $G = \frac12\Delta_a$ satisfies this. Adding to G a function invariant under J does not change F on Π. Accordingly, with G_0 as in Lemma 3.2, we put

$$G_a = G_0 + \tfrac12\Delta_a + k_a, \qquad k_a = \lambda_1(a)k^{(1)} + \lambda_2(a)k^{(2)}, \quad k^{(1)}(s) = (u^2+v^2)^{p/2},\ k^{(2)}(s) = |u|^p + |v|^p,$$

with $\lambda(a) = (\lambda_1(a), \lambda_2(a))$ to be fixed in Step 4. The functions G_0, $k^{(1)}$, $k^{(2)}$ are even, J-invariant and homogeneous of degree p; the functions $k^{(1)}$ and $k^{(2)}$ are C^2 (as $p > 2$); and $G_a(0,0) = 0$. For $a = 0$, F_a is F_0 of Lemma 3.2, as $\Delta_0 = 0$ and $\lambda(0) = 0$ (see Step 4). On Π the J-invariant parts cancel, so

$$F_a(\hat{s}) = \Delta_a(s) + \tfrac{1}{2}\big(\Delta_a(Js) - \Delta_a(s)\big) = \tfrac{1}{4}\Sigma_a(s) \ \geq 0. \tag{6}$$

Step 3: the bases of the p-th powers near the zeros of F_a. The function $|y|^p$ is real-analytic in y for $y \neq 0$, and we use its analyticity only there. We therefore show that the bases of all the p-th powers in F_a stay away from 0 near the windows $\hat{s}$, $s \in \mathcal{O}_a$. For $a > 0$ these bases are u, v, w (from $|u|^p$, G_0, $k^{(2)}$), $(1+B)u + av$, $(1+B)v + aw$ (from $\Delta_a(u,v)$ and $\Delta_a(v,w)$) and $u + av - Bw$. The function $k^{(1)}$ is smooth where $u^2 + v^2$ and $v^2 + w^2$ are positive. For $\nu_0 > 0$ let

$$\mathcal{U}(\nu_0) = \{(s,\nu) : \ \|s\|_2 = 1, \ s \text{ within angular distance } 2\rho \text{ of } \mathcal{O}_0, \ |\nu| \leq \nu_0\},$$

with $s = (u,v)$ and $w = \nu - u$. At $a = 0$ and $\nu = 0$ the six bases are $u, v, -u, (1+B)u, (1+B)v, (1+B)u$, nonzero as s avoids the axes. Hence we can fix $\nu_0 \in (0,1)$ and $\kappa, q_* > 0$ such that on $\mathcal{U} = \mathcal{U}(\nu_0)$ the six numbers u, v, w, $(1+B)u$, $(1+B)v$, $u - Bw$ have absolute value at least κ, and $\partial_\nu^2 F_0 \geq 4q_*$ (at $\nu = 0$, $\partial_\nu^2 F_0 = p(p-1)B(1+B)^{p-2}|u|^{p-2}$, by the proof of Lemma 3.3). For $a > 0$ the bases $(1+B)u + av$, $(1+B)v + aw$ and $u + av - Bw$ differ from $(1+B)u$, $(1+B)v$ and $u - Bw$ by av, aw and av; as $|v| \leq 1$ and $|w| \leq 2$ on $\mathcal{U}$, these differences are at most $2a$ in absolute value. Therefore all six bases have absolute value at least $\kappa/2$ on $\mathcal{U}$ once $a \leq \kappa/4$, and we shrink a_1 to at most $\kappa/4$. The terms $k^{(1)}(u,v)$ and $k^{(1)}(v,w)$ are also real-analytic on $\mathcal{U}$, as $u^2 + v^2 = 1$ and $v^2 + w^2 \geq w^2 \geq \kappa^2$ there.

Step 4: the derivative across Π. At the windows $\hat{s}$, $s \in \mathcal{O}_a$, $F_a = \frac{1}{4}\Sigma_a = 0$ by (6) and Lemma 3.5(i); since we need $F_a \geq 0$, its gradient must vanish there. Along Π it does, by (6); across Π we impose it by the choice of $\lambda(a)$. Let $N_a(s) = \partial_\nu F_a(\hat{s} + \nu e_3)|_{\nu=0}$. The term $k_a(v,w) - k_a(u,v)$ contributes $\partial_2 k_a(Js)$ to $N_a(s)$, where ∂_1, ∂_2 are the partial derivatives in the first and second argument. For a J-invariant C^1 function k (such as k_a), differentiating $k(J\xi) = k(\xi)$ gives $\nabla k(J\xi) = J\nabla k(\xi)$ (as $J^{\mathrm{t}} = J^{-1}$), so this contribution is $-\partial_1 k_a(s)$, and its contribution to $N_a(Js)$ is $-\partial_1 k_a(Js) = -\partial_2 k_a(s)$. Write N_a^0 for N_a with $k_a = 0$. We choose $\lambda(a)$ so that

$$\nabla k_a(s_a) = \big(N_a^0(s_a),\, N_a^0(Js_a)\big), \quad \text{i.e. } N_a(s_a) = N_a(Js_a) = 0.$$

This is a 2×2 linear system. Here $s_0 = e_*$ is a positive multiple of $(1, t_*)$; at $(1, t_*)$ the gradients of $k^{(1)}, k^{(2)}$ are $p(1+t_*^2)^{(p-2)/2}(1, t_*)$ and $p(1, t_*^{p-1})$, and at s_0 they are positive multiples of these, so they are independent because $0 < t_* < 1$ and $p > 2$. Moreover $N_0^0 = \partial_\nu F_0 = 0$ on all of Π ($F_0 \geq 0$ vanishes on Π), and N_a^0 and s_a are real-analytic near $a = 0$ (Lemma 3.5(iv)); so, after strictly decreasing a_1, the system is uniquely solvable for $0 \leq a \leq a_1$, $\lambda(a)$ is real-analytic there, $\lambda(0) = 0$, and $|\lambda(a)| \leq \Lambda a$ there, for some constant Λ. As F_a is even, $N_a(-s) = -N_a(s)$, so $N_a = 0$ on all of $\mathcal{O}_a$.

Step 5: a uniform C^1 bound. Put $\mathcal{E}_a = F_a - F_0$, i.e.

$$\mathcal{E}_a = (C_a^p - (1+B)^p)|u|^p + \tfrac{1}{2}\big(\Delta_a(v,w) - \Delta_a(u,v)\big) + k_a(v,w) - k_a(u,v) - \big(|u + av - Bw|^p - |u - Bw|^p\big).$$

With $\|\cdot\|_\infty$ the maximum norm on $\mathbb{R}^3$ there is L, depending on p, B and Λ, with

$$|\mathcal{E}_a(\mathbf{x})| \leq La\|\mathbf{x}\|_\infty^p, \qquad |\nabla \mathcal{E}_a(\mathbf{x})| \leq La\|\mathbf{x}\|_\infty^{p-1} \qquad (\text{windows } \mathbf{x} = (u,v,w) \in \mathbb{R}^3, \ 0 \leq a < a_1). \tag{7}$$

Indeed, $|C_a - (1+B)| \leq a\|J\|_p = a$, so $|C_a^p - (1+B)^p| \leq p(1+B+a)^{p-1}a$ by the mean value theorem. By the mean value theorem, $\big||y+h|^p - |y|^p\big| \leq p(|y| + |h|)^{p-1}|h|$ and (as $p > 2$) $|\phi(y+h) - \phi(y)| \leq (p-1)(|y| + |h|)^{p-2}|h|$; with $y = u - Bw$, $h = av$, $|y| \leq (1+B)\|\mathbf{x}\|_\infty$ and $|h| \leq a\|\mathbf{x}\|_\infty$, these bound $|u + av - Bw|^p - |u - Bw|^p$ and its gradient (whose remaining part, $ap\,\phi(u + av - Bw)$ in the v-direction, is at most $ap(1+B+a)^{p-1}\|\mathbf{x}\|_\infty^{p-1}$); after we write $\Delta_a(s) = (C_a^p - (1+B)^p)|u|^p - \big(|(1+B)u + av|^p - |(1+B)u|^p\big)$, they also bound the terms with Δ_a. The terms with k_a are bounded because $|\lambda(a)| \leq \Lambda a$.

Step 6: $F_a \geq 0$. Since F_a is homogeneous of degree p, it suffices to consider windows $\hat{s}+\nu e_3$ with $\|s\|_2 = 1$ or $s = 0$. If $s = 0$, $F_a(0,0,w) = F_0(0,0,w)+\mathcal{E}_a(0,0,w) \geq \big(B(1+B)^{p-1} - B^p - La\big)|w|^p \geq 0$ once $La \leq B(1+B)^{p-1} - B^p$, a positive number as $1+B > B$. If $\|s\|_2 = 1$ and $\mathbf{x} = \hat{s}+\nu e_3$, then $F_a(\mathbf{x}) = F_0(\mathbf{x}) + \mathcal{E}_a(\hat{s}) + \big(\mathcal{E}_a(\mathbf{x}) - \mathcal{E}_a(\hat{s})\big)$ with $\mathcal{E}_a(\hat{s}) = F_a(\hat{s}) = \frac{1}{4}\Sigma_a(s)$; along the segment from $\hat{s}$ to $\mathbf{x}$ the maximum norm is at most $1+|\nu|$, so (7) and Lemma 3.3 give

$$F_a(\mathbf{x}) \;\geq\; \tfrac{1}{4}\Sigma_a(s) + \gamma_1(|u|+|\nu|)^{p-2}\nu^2 - La(1+|\nu|)^{p-1}|\nu|. \tag{8}$$

Case 1: s at angular distance $\geq \rho$ from $\mathcal{O}_a$. (This case contains $s = (0,\pm 1)$ and $s = (\pm 1, 0)$, where $u = 0$ or $v = 0$ and F_a need not be real-analytic; the argument uses only (8).) Let $\nu_1 = \min\big(1, \delta/(2^{p+2}L)\big)$. For $|\nu| \leq \nu_1$, (8) and Lemma 3.5(ii) give $F_a \geq a\delta/4 - 2^{p-1}La\nu_1 \geq a\delta/8$. For $|\nu| \geq \nu_1$, $(|u|+|\nu|)^{p-2}\nu^2 \geq |\nu|^p$ and $(1+|\nu|)^{p-1} \leq (1+1/\nu_1)^{p-1}|\nu|^{p-1}$, so $F_a \geq |\nu|^p\big(\gamma_1 - La(1+1/\nu_1)^{p-1}\big) \geq 0$ for small a.

Case 2: s within angular distance ρ of $\mathcal{O}_a$. Write $r = r(s)$. As $\mathcal{O}_a$ lies within $\rho/2$ of $\mathcal{O}_0$, s lies within $3\rho/2 < 2\rho$ of $\mathcal{O}_0$, so for $|\nu| \leq \nu_0$ the point (s,ν) is in $\mathcal{U}$, where all bases of all p-th powers in F_a are at least $\kappa/2$ in absolute value (Step 3). Then F_a is real-analytic in (a,s,ν) there, and by Taylor's formula in ν,

$$F_a(\hat{s}+\nu e_3) = \tfrac{1}{4}\Sigma_a(s) + \nu N_a(s) + \nu^2 q_a(s,\nu), \qquad q_a(s,\nu) = \int_0^1 (1-\tau)\,\partial_\nu^2 F_a(\hat{s}+\tau\nu e_3)\,d\tau.$$

By Step 3, $\partial_\nu^2 F_0 \geq 4q_*$ on $\mathcal{U}$; as F_a is real-analytic in (a,s,ν) on the compact set $[0,a_1]\times\mathcal{U}$, there is L_2 with $|\partial_\nu^2(F_a - F_0)| \leq L_2 a$ there, so $q_a \geq q_*$ on $\mathcal{U}$ once $L_2 a \leq 2q_*$. Since $N_0 = N_0^0 = 0$ on all of Π (Step 4, as $\lambda(0) = 0$) and N_a is real-analytic in a, the derivative of N_a along the unit circle is $O(a)$ on the 2ρ-arcs. As N_a vanishes on $\mathcal{O}_a$, there is a constant L_1 with $|N_a(s)| \leq L_1 a r$. With Lemma 3.5(iii),

$$F_a \;\geq\; \tfrac{1}{4}a\eta r^2 - L_1 a r|\nu| + q_*\nu^2 \;\geq\; 0 \qquad \text{once } (L_1 a)^2 \leq a\eta q_*,$$

since then the discriminant $(L_1 a)^2 - 4\cdot\frac{1}{4}a\eta\cdot q_*$ of this quadratic form in $(r,|\nu|)$ is ≤ 0. For $|\nu| \geq \nu_0$, (8) with $\Sigma_a \geq 0$, $(|u|+|\nu|)^{p-2}\nu^2 \geq |\nu|^p$ and $(1+|\nu|)^{p-1} \leq (1+1/\nu_0)^{p-1}|\nu|^{p-1}$ gives $F_a \geq |\nu|^p\big(\gamma_1 - La(1+1/\nu_0)^{p-1}\big) \geq 0$ for small a.

Order of the constants. The constants are chosen in this order: γ_1 (Lemma 3.3); ρ, m, δ, η and a_1 (Lemma 3.5); ν_0, κ, q_* (Step 3); Λ (Step 4, after the last shrinking of a_1); L (Step 5); L_1, L_2, ν_1 (Step 6); finally, a_0 is the smallest of a_1 and the finitely many upper bounds on a imposed in Step 6. For $0 < a < a_0$, (2) holds with $C = C_a$, and Lemma 3.1 gives $\|P_a(S)\|_p \leq C_a = \|P_a(J)\|_p$.

The case $1 < p < 2$. By Lemma 2.3 and the case $p' > 2$, $\|P_a(S)\|_p = \|P_a(S)\|_{p'} = \|P_a(J)\|_{p'} = \|P_a(J^{\mathsf{t}})\|_p$ for $0 \leq a < a_0(p', B)$. Since $J^{\mathsf{t}} = -J = VJV$ with the isometry $V = \mathrm{diag}(1,-1)$ of ℓ_2^p, $P_a(J^{\mathsf{t}}) = VP_a(J)V$ has the norm of $P_a(J)$. □

4. A scaled rotation by more than a right angle

Lemma 4.1 (a contraction T_ε with $\|P_a(T_\varepsilon)\|_p > \|P_a(J)\|_p$)**.** *Let $p > 2$, $B > 1$, $K_* = \frac{1}{2}(J - D_* I)$ and $T_\varepsilon = \exp(\varepsilon K_*)J$. Then $\|T_\varepsilon\|_p \leq 1$ for all $\varepsilon \geq 0$, on both the real and the complex ℓ_2^p, and there is $a_2 > 0$ such that for each $0 < a < a_2$, $\|P_a(T_\varepsilon)\|_p > \|P_a(J)\|_p$ for all sufficiently small $\varepsilon > 0$ (depending on a).*

Proof. The matrix T_ε is a contraction. For real ξ and $\xi(\varepsilon) = \exp(\varepsilon K_*)\xi$, $\frac{d}{d\varepsilon}\|\xi(\varepsilon)\|_p^p = p\langle\phi(\xi(\varepsilon)), K_*\xi(\varepsilon)\rangle = \frac{p}{2}\|\xi(\varepsilon)\|_p^p\,(Q(\xi(\varepsilon)) - D_*) \leq 0$. Hence $\exp(\varepsilon K_*)$, and so $T_\varepsilon = \exp(\varepsilon K_*)J$ (J is an isometry), is a contraction of the real ℓ_2^p, and by Lemma 2.1 also of the complex ℓ_2^p. (As $\exp(\varepsilon K_*) = e^{-\varepsilon D_*/2}R_{\varepsilon/2}$, $T_\varepsilon = e^{-\varepsilon D_*/2}R_{\pi/2+\varepsilon/2}$.)

Choice of test vector. By Lemma 3.5(i), (iv), for $0 < a < a_1$ the real maximisers of $\|P_a(J)\xi\|_p/\|\xi\|_p$ are exactly the nonzero real multiples of s_a and Js_a, with s_a real-analytic in a; we take $\xi_a = s_a/(s_a)_1$, so $\xi_a = \xi_0 + a\dot{\xi} + o(a)$ with $\xi_0 = (1, t_*)$ and $\dot{\xi} = \frac{d}{da}\xi_a|_{a=0}$. Put $y_a = P_a(J)\xi_a$.

Derivative in ε. Recall $P_a(T_\varepsilon) = I + aT_\varepsilon - BT_\varepsilon^2$ and $T_0 = J$. Since $\frac{d}{d\varepsilon}T_\varepsilon|_0 = K_*J$ and $\frac{d}{d\varepsilon}T_\varepsilon^2|_0 = K_*J^2 + JK_*J = -K_* + JK_*J$,

$$\tfrac{d}{d\varepsilon}\|P_a(T_\varepsilon)\xi_a\|_p^p\Big|_{\varepsilon=0} = p\,\Gamma(a), \qquad \Gamma(a) = \big\langle \phi(y_a), M_a\xi_a\big\rangle, \quad M_a = aK_*J + B(K_* - JK_*J).$$

Since $J^2 = -I$, $K_*J = -\frac{1}{2}(I + D_*J)$ and $K_* - JK_*J = J - D_*I$. At $a = 0$, $y_0 = (1+B)\xi_0$ and $\Gamma(0) = (1+B)^{p-1}B\|\xi_0\|_p^p\,(Q(\xi_0) - D_*) = 0$. Differentiating $\Gamma(a) = \langle\phi(y_a), M_a\xi_a\rangle$ at $a = 0$, with $\frac{d}{da}y_a|_0 = (1+B)\dot\xi + J\xi_0$, $d\phi(y) = (p-1)|y|^{p-2}dy$, $\frac{d}{da}M_a = K_*J$ and $M_0 = B(J - D_*I)$, gives

$$\begin{aligned}\Gamma'(0) = {}& (p-1)(1+B)^{p-2}B\big\langle |\xi_0|^{p-2}\big((1+B)\dot\xi + J\xi_0\big), (J - D_*I)\xi_0\big\rangle \\ & + (1+B)^{p-1}\langle\phi(\xi_0), K_*J\xi_0\rangle + (1+B)^{p-1}B\langle\phi(\xi_0), (J - D_*I)\dot\xi\rangle.\end{aligned}$$

The two terms containing $\dot\xi$ add up to $(1+B)^{p-1}B$ times the derivative at ξ_0, in the direction $\dot\xi$, of $\Psi(\xi) = \langle\phi(\xi), (J - D_*I)\xi\rangle = \|\xi\|_p^p(Q(\xi) - D_*)$. As $\Psi \le 0$ and $\Psi(\xi_0) = 0$, ξ_0 is a maximum of Ψ, so this derivative is 0. To evaluate the remaining terms, put $t = t_*$. Then $\|\xi_0\|_p^p = 1 + t^p$, $\langle\phi(\xi_0), J\xi_0\rangle = D_*(1+t^p)$, and

$$E_* := \langle|\xi_0|^{p-2}J\xi_0, (J - D_*I)\xi_0\rangle = t^2 + t^{p-2} - D_*(t - t^{p-1}) = \frac{t^{p-2}(1+t^2)^2}{1+t^p};$$

the last equality follows from $D_* = (t - t^{p-1})/(1+t^p)$ and the identity $(t^2 + t^{p-2})(1+t^p) - (t - t^{p-1})^2 = t^{p-2}(1+t^2)^2$, valid for every $t > 0$. Using $K_*J = -\frac{1}{2}(I + D_*J)$, the remaining terms are $(p-1)(1+B)^{p-2}BE_* - \frac{1}{2}(1+B)^{p-1}\big(\|\xi_0\|_p^p + D_*\langle\phi(\xi_0), J\xi_0\rangle\big)$, that is

$$\Gamma'(0) = (1+B)^{p-2}\Big[-\tfrac{1}{2}(1+B)(1+t^p)(1+D_*^2) + (p-1)BE_*\Big].$$

By (4), $(1+t^p)^2(1+D_*^2) = (1+t^p)^2 + (t - t^{p-1})^2 = (1+t^2)(1+t^{2p-2}) = (p-1)t^{p-2}(1+t^2)^2$, that is $(1+t^p)(1+D_*^2) = (p-1)E_*$. Therefore

$$\Gamma(a) = a\,(1+B)^{p-2}(p-1)E_*\,\frac{B-1}{2} + o(a),$$

which is positive for $0 < a < a_2$, for some $a_2 \in (0, a_1]$, when $B > 1$. For such a, $p\,\Gamma(a)$ is positive, so for all small $\varepsilon > 0$, $\|P_a(T_\varepsilon)\|_p \ge \|P_a(T_\varepsilon)\xi_a\|_p/\|\xi_a\|_p > \|P_a(J)\xi_a\|_p/\|\xi_a\|_p = \|P_a(J)\|_p$. □

5. Proofs of Theorems A and B

Proof of Theorem A. Let $p > 2$ and $B > 1$. For $0 < a < \min(a_0, a_2)$, Lemma 4.1 gives a contraction T_ε of ℓ_2^p with $\|P_a(T_\varepsilon)\|_p > \|P_a(J)\|_p$, and Theorem 3.6 gives $\|P_a(J)\|_p \ge \|P_a(S)\|_p$. Since ℓ_2^p is an L^p space, this contradicts (1). This holds for real and for complex scalars alike: T_ε is a contraction of both spaces (Lemma 4.1), and the real operators $P_a(T_\varepsilon)$, $P_a(J)$ and $P_a(S)$ have the same norms on the real and on the complex spaces (Lemma 2.1). For $1 < p < 2$, apply this at $p' > 2$ and use Lemma 2.3: $T_\varepsilon^{\mathsf t}$ is a contraction of ℓ_2^p and $\|P_a(T_\varepsilon^{\mathsf t})\|_p = \|P_a(T_\varepsilon)\|_{p'} > \|P_a(S)\|_{p'} = \|P_a(S)\|_p$. Finally, if T is any of these contractions, $T \oplus 0$ is a contraction of $\ell_2^p \oplus_p \ell^p = \ell^p$ with $P_a(T\oplus 0) = P_a(T)\oplus I$, hence $\|P_a(T\oplus 0)\|_p = \max(\|P_a(T)\|_p, 1) \ge \|P_a(T)\|_p$, so (1) also fails on ℓ^p itself. □

Proof of Theorem B. If $c_0 = c_2 = 0$, then $P(z) = c_1 z$, and $\|P(T)\|_p \le |c_1| = \|P(S)\|_p$ because S is an isometry. Otherwise choose $\zeta \in \mathbb{C}$ with $|\zeta| = 1$ and $\zeta^2 = -c_2/c_0$; then $P(z) = c_0\widetilde P(\zeta z)$ with $\widetilde P(z) = 1 + \alpha z - z^2$, $\alpha = c_1/(c_0\zeta) \in \mathbb{C}$. Since ζT is a contraction and $\zeta S = USU^{-1}$ with the isometry $(Ux)_n = \zeta^n x_n$ (as in the Background), it suffices to treat $\widetilde P$: if $\|\widetilde P(T')\|_p \le \|\widetilde P(S)\|_p$ for every contraction T', then $\|P(T)\|_p = |c_0|\|\widetilde P(\zeta T)\|_p \le |c_0|\|\widetilde P(S)\|_p = |c_0|\|U\widetilde P(S)U^{-1}\|_p = |c_0|\|\widetilde P(\zeta S)\|_p = \|P(S)\|_p$. For a contraction T on $L^p = L^p(\Omega, \mathfrak m)$ and $x \in L^p$,

$$\widetilde P(T)x = \tfrac{1}{2}\big[(2x + \alpha Tx) + T(\alpha x - 2Tx)\big].$$

By convexity of $|\cdot|^p$ (pointwise) and $\|T\|_p \le 1$,

$$\|\widetilde{P}(T)x\|_p^p \le \tfrac{1}{2}\|2x+\alpha Tx\|_p^p + \tfrac{1}{2}\|\alpha x - 2Tx\|_p^p = \tfrac{1}{2}\int_\Omega \|M(x(\omega),Tx(\omega))\|_p^p \, d\mathfrak{m}(\omega)$$
$$\le \tfrac{1}{2}\|M\|_p^p\big(\|x\|_p^p + \|Tx\|_p^p\big) \le \|M\|_p^p\|x\|_p^p,$$

where $M = \begin{pmatrix} 2 & \alpha \\ \alpha & -2 \end{pmatrix}$ acts pointwise on (x, Tx). As $M = V\widetilde{P}(J)$ with $V = \mathrm{diag}(1,-1)$, $\|M\|_p = \|\widetilde{P}(J)\|_p \le \|\widetilde{P}(S)\|_p$ by Lemma 2.2. (For $p=\infty$, replace each average of p-th powers by a maximum and the integral by an essential supremum.) □

Corollary 5.1 (exact norms for $|c_0| = |c_2|$)**.** *Let $P(z) = c_0 + c_1 z + c_2 z^2$ with $|c_0| = |c_2| > 0$, and let $\zeta \in \mathbb{C}$ satisfy $|\zeta| = 1$ and $\zeta^2 = -c_2/c_0$. For $1 \le p \le \infty$, $\|P(S)\|_p = \|P(\bar\zeta J)\|_p = \max_{|\mu|=1}\|P(\mu J)\|_p$.*

Proof. With $P(z) = c_0\widetilde{P}(\zeta z)$ as in the proof of Theorem B, $\|P(S)\|_p = |c_0|\|\widetilde{P}(S)\|_p$. The proof of Theorem B with $T = S$ gives $\|\widetilde{P}(S)\|_p \le \|\widetilde{P}(J)\|_p$, and Lemma 2.2 gives equality; here $c_0\widetilde{P}(J) = P(\bar\zeta J)$. For every $|\mu| = 1$, Lemma 2.2 applied to $z \mapsto P(\mu z)$ gives $\|P(\mu J)\|_p \le \|P(\mu S)\|_p = \|P(S)\|_p$. (The factor $\bar\zeta$ cannot be omitted: for $P(z) = 1 + z^2$, $P(J) = 0$, while, with $\zeta = i$, $\|P(S)\|_p = \|P(-iJ)\|_p = \|2I\|_p = 2$.) □

Remark 5.2 (Theorem B is not the trivial bound)*.* Theorem B does not reduce to the trivial bound $\|P(T)\|_p \le \sum|c_k|$: for $\widetilde{P}(z) = 1 + \alpha z - z^2$ with real $\alpha \ne 0$ ($\zeta = 1$ in Corollary 5.1) and $1 < p < \infty$, Corollary 5.1 gives $\|\widetilde{P}(S)\|_p = \|2I + \alpha J\|_p < 2 + |\alpha| = \sum|c_k|$ (consistent with [2, Prop. 6.1]). Indeed, by strict convexity of ℓ_2^p, $\|2\xi + \alpha J\xi\|_p = 2\|\xi\|_p + |\alpha|\|J\xi\|_p$ for a real $\xi \ne 0$ would make $J\xi$ a real multiple of ξ, impossible as $J^2 = -I$; so the maximum of $\|(2I + \alpha J)\xi\|_p$ over the compact real unit sphere is $< 2 + |\alpha|$, and Lemma 2.1 gives the same on the complex ℓ_2^p.

Part II: further results

Sections 6–10 prove the further results listed in the table of Section 1. Each section of Part II fixes its own local constants and functions; S, J, R_θ, ℓ_2^p, $\|\cdot\|_p$ and the results of Part I keep their meaning, and ℓ_n^p is $\mathbb{C}^n$ (or $\mathbb{R}^n$) with the ℓ^p norm.

6. A counterexample proved by hand on a punctured neighbourhood of $p = 2$

The next two remarks are in the notation of the proof of Theorem 3.6.

Remark 6.1. The term k_a cannot be omitted: $N_a^0(s_a)$ need not vanish. For $B = 1$, the maximiser identity $P_a(J)^{\mathsf{t}}\phi(P_a(J)s) = C_a^p\phi(s)$ gives $N_a^0 = \frac{p}{2}(C_a^p - 2^p)\phi(u)$ on $\mathcal{O}_a$, which is nonzero since $C_a > 2$ by (3), as $D_* > 0$. As $F_a(\hat s_a) = 0$, a nonzero derivative across Π there makes F_a negative nearby.

Remark 6.2. As $p \to 2^+$, $Q \to 0$ and the curvature of Σ' at $\mathcal{O}_0$ tends to 0. For $B \ne 1$, let $\bar a(p) = \sup\{b \ge 0 :\ \|P_a(S)\|_p = \|P_a(J)\|_p \text{ for all } a \in [0,b]\}$. Then $\bar a(p) \to 0$ as $p \to 2$. Indeed, for $p = 2$, $B \ne 1$ and every $a > 0$, $\|P_a(S)\|_2 = \max_{|z|=1}|P_a(z)| > |P_a(i)| = \|P_a(J)\|_2$, because $\frac{d}{d\theta}|P_a(e^{i\theta})|^2 = 2a(B-1) \ne 0$ at $\theta = \pi/2$; and both norms depend continuously on p (for $\|P_a(S)\|_p$ by the Riesz–Thorin theorem [3, Thm. 1.1.1]). (The assumption $B \ne 1$ is needed: for $B = 1$, Corollary 5.1 gives $\|P_a(S)\|_p = \|P_a(J)\|_p$ for every $a \ge 0$ and every p.)

By Remark 6.2, in the proof of Theorem A the admissible coefficients a shrink to 0 as $p \to 2$, so that proof gives no single polynomial for a whole punctured neighbourhood of $p = 2$. Here an explicit quadratic and an explicit family of scaled rotations violate (1) for every $p \ne 2$ in a neighbourhood of $p = 2$, with a proof by hand. Throughout this section $M_P = \max_{|z|=1}|P(z)|$, and ε denotes $p - 2$ (it is not the rotation parameter ε of Theorem A).

Theorem 6.3. *Let $P(z) = 1 + 2z - \frac{6}{5}z^2$, let $\theta_0 \in (\pi/2, \pi)$ with $\cos\theta_0 = -\frac{1}{12}$, and put $T_p = R_{\theta_0}/\|R_{\theta_0}\|_p$. There is $\delta > 0$ such that for every p with $0 < |p-2| < \delta$, T_p is a contraction on ℓ_2^p and $\|P(T_p)\|_p > \|P(S)\|_p$.*

At $p = 2$ both sides equal M_P. Moving p away from 2, the left side grows linearly in $|p-2|$ (Lemma 6.5), while the right side has slope 0 (Lemma 6.8 and the paragraph after its proof), because $|P|$ attains its maximum on the circle at an angle that is an irrational multiple of π. Lemma 6.5 treats $p < 2$ and $p > 2$ alike; Lemma 6.8 reduces $p < 2$ to $p' > 2$ by Lemma 2.3.

6.1. **The norm of a rotation near $p = 2$.** For a probability vector μ let $H(\mu) = -\sum_i \mu_i \log \mu_i$, and put $\mathcal{H}(\omega) = H(\cos^2\omega, \sin^2\omega)$. For $v \neq 0$ in $\mathbb{R}^n$,

$$\tfrac{d}{dp}\log\|v\|_p\big|_{p=2} = -\tfrac{1}{4}H(\mu), \qquad \mu_i = |v_i|^2/\|v\|_2^2, \tag{9}$$

as one sees by differentiating $\frac{1}{p}\log\sum_i |v_i|^p$. The function $\mathcal{H}$ is even and has period $\pi/2$; for $t = \tan\omega \in (0,1]$, $|\mathcal{H}'(\omega)| = |\sin 2\omega \, \log\cot^2\omega| = 4t|\log t|/(1+t^2) \le 4/e$; by these symmetries, $|\mathcal{H}'| \le 4/e$ everywhere.

Lemma 6.4. *For every θ, $\|R_\theta\|_p = 1 + |p-2|\,\sigma(\theta) + o(p-2)$ as $p \to 2$, where $\sigma(\theta) = \frac{1}{4}\max_\omega\big(\mathcal{H}(\omega) - \mathcal{H}(\omega-\theta)\big) \ge 0$. Moreover σ depends only on the distance from θ to $\frac{\pi}{2}\mathbb{Z}$, and $\sigma(\theta) \le \operatorname{dist}(\theta, \frac{\pi}{2}\mathbb{Z})/e$.*

Proof. By Lemma 2.1 we may work over $\mathbb{R}$. With $x_\omega = (\cos\omega, \sin\omega)$, $R_\theta x_\omega = x_{\omega-\theta}$, so $\|R_\theta\|_p = \max_\omega \psi_\omega(p)$ with $\psi_\omega(p) = \|x_{\omega-\theta}\|_p/\|x_\omega\|_p$. Each $\psi_\omega(2) = 1$, and $(\omega, p) \mapsto \psi_\omega(p)$ is C^1 (as $t \mapsto t^p \log t$ is continuous on $[0,1]$), with $\partial_p\psi_\omega(2) = \frac{1}{4}(\mathcal{H}(\omega) - \mathcal{H}(\omega-\theta))$ by (9). By uniform continuity of $\partial_p\psi_\omega(p)$ on compact sets, $\psi_\omega(2+\varepsilon) = 1 + \varepsilon\,\partial_p\psi_\omega(2) + o(\varepsilon)$ uniformly in ω, so $\|R_\theta\|_{2+\varepsilon} = 1 + \varepsilon\max_\omega \partial_p\psi_\omega(2) + o(\varepsilon)$ for $\varepsilon > 0$ and $1 + \varepsilon\min_\omega \partial_p\psi_\omega(2) + o(\varepsilon)$ for $\varepsilon < 0$ [5]. Here $\max_\omega \partial_p\psi_\omega(2) = \sigma(\theta)$ by definition, and $\min_\omega \partial_p\psi_\omega(2) = -\sigma(\theta)$, since the substitution $\omega \mapsto \theta - \omega$ changes the sign of $\mathcal{H}(\omega) - \mathcal{H}(\omega-\theta)$ ($\mathcal{H}$ is even). The last two claims follow from the symmetries of $\mathcal{H}$ and the bound on $|\mathcal{H}'|$. □

Lemma 6.5. *Let P be a real polynomial and $|P|$ maximal on the circle at $z_0 = e^{i\theta_0}$, with $\sin\theta_0 \neq 0$. Put $\rho = z_0P'(z_0)/P(z_0)$, $\alpha = \arg P(z_0)$, $\tau_p = 1/\|R_{\theta_0}\|_p$, $T_p = \tau_p R_{\theta_0}$. Then $\rho \ge 0$ is real,*

$$\|P(T_p)\|_p = |P(\tau_p z_0)|\;\|R_{\alpha(\tau_p)}\|_p, \qquad \alpha(\tau) = \arg P(\tau z_0), \tag{10}$$

and $\|P(T_p)\|_p \ge M_P\big(1 + g\,|p-2|\big) + o(p-2)$ with $g = \sigma(\alpha) - \rho\,\sigma(\theta_0)$.

Proof. The span of I and J is an algebra isomorphic to $\mathbb{C}$ ($J^2 = -I$), with $R_\theta \leftrightarrow e^{i\theta}$; so $P(\tau R_{\theta_0}) = \operatorname{Re}P(\tau z_0)\,I + \operatorname{Im}P(\tau z_0)\,J = |P(\tau z_0)|\,R_{\alpha(\tau)}$, which gives (10). At a maximum of $\log|P(e^{it})|$ its t-derivative, $-\operatorname{Im}\rho$, vanishes, and $|P(rz_0)| \le M_P$ for $r < 1$ gives $\operatorname{Re}\rho \ge 0$. Hence $|P(\tau z_0)| = M_P\big(1 - \rho(1-\tau)\big) + O((1-\tau)^2)$, and $1 - \tau_p = |p-2|\,\sigma(\theta_0) + o(p-2)$ by Lemma 6.4. For the second factor a lower bound suffices. For $p > 2$ fix ω_* with $\sigma(\alpha) = \frac{1}{4}(\mathcal{H}(\omega_*) - \mathcal{H}(\omega_* - \alpha))$, and for $p < 2$ fix ω_* with $-\sigma(\alpha) = \frac{1}{4}(\mathcal{H}(\omega_*) - \mathcal{H}(\omega_* - \alpha))$ (Lemma 6.4 and its proof); let $\Psi(p,\tau) = \|R_{\alpha(\tau)}x_{\omega_*}\|_p/\|x_{\omega_*}\|_p \le \|R_{\alpha(\tau)}\|_p$. Ψ is C^1 near $(2,1)$, $\Psi(2,\tau) = 1$ for all τ, and $\partial_p\Psi(2,1) = \operatorname{sgn}(p-2)\,\sigma(\alpha)$, so $\Psi(p,\tau_p) = 1 + |p-2|\,\sigma(\alpha) + o(p-2)$. Multiplying the two expansions gives the claim. □

Remark 6.6 (the family P_η). For $P_\eta(z) = 1 + 2z - (1+\eta)z^2$, $\eta > 0$, one has $|P_\eta(e^{it})|^2 = (2 - \eta\cos t)^2 + (2+\eta)^2\sin^2 t$, maximal exactly at $\cos t = -\eta/(2(1+\eta))$; there $P_\eta(z_0)/z_0 = 2(2+\eta)\cos(\theta_0/2)e^{-i\theta_0/2}$ and $P_\eta'(z_0) = 2(1+\eta)(1 + \cos\theta_0 - i\sin\theta_0)$ (recall $\rho = P_\eta'(z_0)/(P_\eta(z_0)/z_0)$), so $\alpha = \theta_0/2$, $M_{P_\eta} = 2(2+\eta)\cos(\theta_0/2)$ and $\rho = 2(1+\eta)/(2+\eta)$ for every η. The value $\eta = \frac{7}{10}$ is the polynomial of item (a) of Theorem 9.1, and $\eta = \frac{17}{5}$ is Drury's polynomial; for the latter, numerically $g \approx -0.038 < 0$, so Lemma 6.5 gives no counterexample with Drury's polynomial near $p = 2$ (Drury's matrix, which is not a scaled rotation, gives a counterexample at $p = 4$).

Lemma 6.7. *For $P(z) = 1 + 2z - \frac{6}{5}z^2$, $|P|$ is maximal on the circle exactly at $e^{\pm i\theta_0}$ with $\cos\theta_0 = -\frac{1}{12}$, and $M_P = \frac{11\sqrt{66}}{30}$, $\rho = \frac{12}{11}$, $\alpha = \theta_0/2$, and $g \geq 0.1366 > 0$.*

Proof. With $u = \cos t$, $|P(e^{it})|^2 = \frac{221}{25} - \frac{4}{5}u - \frac{24}{5}u^2$, maximal at $u = -\frac{1}{12}$ with value $M_P^2 = \frac{1331}{150}$. At $z_0 = e^{i\theta_0}$, $P(z_0)/z_0 = \frac{11}{60}(11 - \sqrt{143}\,i)$ and $P'(z_0) = \frac{1}{5}(11 - \sqrt{143}\,i)$, so $\rho = \frac{12}{11}$; and $\tan(\theta_0/2) = \sin\theta_0/(1+\cos\theta_0) = \sqrt{143}/11$ gives $\arg(P(z_0)/z_0) = -\theta_0/2$, so $\alpha = \theta_0/2$. Let $\delta_0 = \theta_0 - \frac{\pi}{2} = \arcsin\frac{1}{12} < 0.083431$. By Lemma 6.4, $\sigma(\theta_0) \leq \delta_0/e < 0.030693$. Taking $\omega = \pi/4$ in the definition of σ, $\sigma(\alpha) \geq \frac{1}{4}\big(\mathcal{H}(\frac{\pi}{4}) - \mathcal{H}(\frac{\pi}{4} - \alpha)\big) = \frac{1}{4}\big(\log 2 - \mathcal{H}(\frac{\delta_0}{2})\big)$, since $\frac{\pi}{4} - \alpha = -\frac{\delta_0}{2}$ and $\mathcal{H}$ is even. Here $\sin^2\frac{\delta_0}{2} = \frac{1}{2}(1 - \frac{\sqrt{143}}{12}) < 0.0017392$, so $\mathcal{H}(\frac{\delta_0}{2}) < 0.012792$ and $\sigma(\alpha) > 0.170088$. Hence $g > 0.170088 - \frac{12}{11}\cdot 0.030693 > 0.1366$. The decimal bounds are roundings in the safe direction, checked by hand from $\arcsin x \leq x + \frac{x^3}{6(1-x^2)}$ for $0 \leq x < 1$ (the Taylor coefficients of arcsin after the first are at most $\frac{1}{6}$), $e > 2.71828$, $\log 2 > 0.693147$, and $\mathcal{H}(\omega) \leq s\,(1 + \log\frac{1}{s})$ for $s = \sin^2\omega \in (0,1)$ (as $-\log(1-s) \leq \frac{s}{1-s}$; the right side increases in s), with $\log\frac{1}{0.0017392} < 6.3544$. □

6.2. An upper bound for $\|P(S)\|_p$ near $p = 2$.

Lemma 6.8. *Let $P(z) = c_0 + c_1 z + c_2 z^2$ be real with $c_0 c_2 < 0$, and suppose $|P|$ attains its maximum M_P on the circle exactly at $e^{\pm i\theta_0}$, with θ_0/π irrational. Then $\|P(S)\|_p \leq M_P + o(p-2)$ as $p \to 2$.*

Proof. By Lemma 2.3 it suffices to take $p = 2+\varepsilon$, $\varepsilon > 0$. We use windows $(u, v, w) = (x_n, x_{n-1}, x_{n-2})$ and $y = c_0 u + c_1 v + c_2 w$, as in Lemma 3.1 (whose proof applies verbatim to any real quadratic P).

(Z1) $p = 2$, exactly. Put $q^2 = -c_0 c_2 > 0$. As a quadratic in $\cos t$, $|P(e^{it})|^2$ is concave ($4c_0c_2 < 0$) with maximiser $\cos\theta_0 = -c_1(c_0+c_2)/(4c_0c_2)$, so $M_P^2 - |P(e^{it})|^2 = 4q^2(\cos t - \cos\theta_0)^2$. Matching coefficients shows that the quadratic form $\mathcal{G}_0(s,t) = (M_P^2 - c_0^2 + c_0c_2)s^2 + (4q^2\cos\theta_0 - 2c_0c_1)st + c_2(c_2 - c_0)t^2$ satisfies

$$F_0(u,v,w) := M_P^2 u^2 + \mathcal{G}_0(v,w) - \mathcal{G}_0(u,v) - y^2 = q^2\,(u - 2v\cos\theta_0 + w)^2 \;\geq 0. \tag{11}$$

Its zero set is the plane $\Pi_{\theta_0} = \{u - 2v\cos\theta_0 + w = 0\}$: the windows of the sequences $x_n = r\cos(\theta + n\theta_0)$. On such a window, with $\mathbf{e}(\theta) = (\cos\theta, \cos(\theta - \theta_0))$, the two states are $(u, v) = r\,\mathbf{e}(\theta)$ and $(v, w) = r\,\mathbf{e}(\theta - \theta_0)$, and $y = rM_P\cos(\theta + \gamma)$ for a fixed γ, since $|c_0 + c_1 e^{-i\theta_0} + c_2 e^{-2i\theta_0}| = M_P$. The curve $E = \{\mathbf{e}(\theta)\}$ is an ellipse (as $\sin\theta_0 \neq 0$); let N be its gauge ($N = 1$ on E) and $\theta(s)$ the angle of $s/N(s)$ on E.

(Z2) The ansatz. Fix $\kappa > 0$ and put $C_\varepsilon = M_P(1 + \kappa\varepsilon)$, $\mathcal{G}_\varepsilon(s) = N(s)^\varepsilon\big(\mathcal{G}_0(s) + \varepsilon N(s)^2 k(\theta(s))\big)$ with a continuous k on $\mathbb{R}/2\pi\mathbb{Z}$ chosen below, and $F_\varepsilon = C_\varepsilon^p|u|^p + \mathcal{G}_\varepsilon(v,w) - \mathcal{G}_\varepsilon(u,v) - |y|^p$. Then $F_\varepsilon = F_0 + \varepsilon F_1 + R_\varepsilon$ with F_1 continuous and $|R_\varepsilon| \leq K\varepsilon^2$ on the Euclidean unit sphere of $\mathbb{R}^3$ (the terms are $|t|^{2+\varepsilon}$, $N^\varepsilon\mathcal{G}_0$ and $\varepsilon N^{\varepsilon+2}k$, and $|t|^2\log^2|t|$ is bounded near 0). On Π_{θ_0} with $r = 1$ both states lie on E, so $F_1 = \Omega(\theta) + k(\theta - \theta_0) - k(\theta)$ with

$$\Omega(\theta) = M_P^2\big[(\log M_P + 2\kappa)\cos^2\theta + \cos^2\theta\log|\cos\theta| - \cos^2(\theta+\gamma)\big(\log M_P + \log|\cos(\theta+\gamma)|\big)\big],$$

and $\int_0^{2\pi}\Omega = 2\pi\kappa M_P^2$: the other terms cancel by translation invariance.

(Z3) A coboundary. As $\theta_0/2\pi$ is irrational, the rotation $\theta \mapsto \theta - \theta_0$ is uniquely ergodic, so the averages $A_L\Omega = \frac{1}{L}\sum_{j<L}\Omega(\cdot - j\theta_0)$ converge to κM_P^2 uniformly as $L \to \infty$. Fix L with $A_L\Omega \geq \kappa M_P^2/2$ and put $k = \frac{1}{L}\sum_{j=1}^{L-1}\sum_{i<j}\Omega(\cdot - i\theta_0)$. Telescoping gives $\Omega + k(\cdot - \theta_0) - k = A_L\Omega \geq \kappa M_P^2/2$.

(Z4) Positivity. F_ε is $(2+\varepsilon)$-homogeneous and $F_0 = 0$ on Π_{θ_0}, so $F_1(\lambda x) = \lambda^2 F_1(x)$ there; a window of Π_{θ_0} with $r = 1$ has Euclidean norm at most $\sqrt{3}$, so $F_1 \geq \kappa M_P^2/6$ on Π_{θ_0} and the unit sphere, and $F_1 \geq \kappa M_P^2/12$ within some distance r_0 of it. Off that neighbourhood $F_0 \geq m_0 > 0$ on the sphere, by (11). Hence $F_\varepsilon \geq 0$ on the sphere, and so on $\mathbb{R}^3$, for all small $\varepsilon > 0$.

(Z5) Conclusion. Summing over all windows of a finitely supported real sequence, the $\mathcal{G}_\varepsilon$ terms telescope ($\mathcal{G}_\varepsilon(0) = 0$), so $\|P(S)x\|_p \leq C_\varepsilon\|x\|_p$ for every finitely supported real x, hence for every

real $x \in \ell^p(\mathbb{N})$, and Lemma 2.1 extends this to complex x. Hence $\|P(S)\|_{2+\varepsilon} \le M_P(1+\kappa\varepsilon)$ for all small $\varepsilon > 0$; as $\kappa > 0$ was arbitrary, the lemma follows. □

Conversely, $\|P(S)\|_p \ge \|P(S)\|_2 = M_P$ for every $p \in (1,\infty)$: $\log\|P(S)\|_{1/s}$ is convex in s by the Riesz–Thorin theorem [3, Thm. 1.1.1], it takes the same value at $s = 1/p$ and $s = 1/p'$ by Lemma 2.3, and $s = \frac{1}{2}$ is the midpoint. So $\|P(S)\|_p = M_P + o(p-2)$, that is, the slope at $p = 2$ is 0. Only the upper bound is used.

Proof of Theorem 6.3. By Lemma 6.7, $|P|$ is maximal on the circle exactly at $e^{\pm i\theta_0}$ and $g > 0$; $c_0c_2 = -\frac{6}{5} < 0$; and θ_0/π is irrational, since otherwise $2\cos\theta_0 = -\frac{1}{6}$ would be an algebraic integer. By Lemmas 6.5 and 6.8, $\|P(T_p)\|_p \ge M_P(1+g|p-2|) + o(p-2) > M_P(1+\frac{g}{2}|p-2|) \ge \|P(S)\|_p$ for all $p \neq 2$ close to 2. □

Remark 6.9. The constant δ of Theorem 6.3 is not explicit. Item (a) of Theorem 9.1 gives an explicit range, $p \in [21/11, 2) \cup (2, 2.1]$, for the polynomial with $\eta = \frac{7}{10}$, by certified computation. Numerically, for P of Theorem 6.3, $g = 0.1425$ (the bound above is 0.1366).

7. Drury's counterexample at $p = 4$ with exact arithmetic

Drury [6] disproved Matsaev's conjecture at $p = 4$, and so at $p = 4/3$, with the polynomial

$$P(z) = 1 + 2z - \tfrac{22}{5}z^2 \tag{12}$$

and a real 2×2 matrix T_{Dr} with six-decimal entries. His proof needs two upper bounds, $\|T_{\mathrm{Dr}}\|_4 \le 1$ and $\|P(S)\|_4 < 6.5278$. Both are certified by a general algorithm for the norm of a finite real matrix: a Stečkin-type estimate on patches of the unit sphere, searched depth first in exact rational arithmetic in Maple, in under 24 hours on a PC [6, Sec. 3–8]. The only step that is not exact is the final comparison $6.52773\,\|\nu\|_{4/3}\|\mu\|_{4/3}^{-1} < 6.5278$ [6, p. 324 and footnote 3, p. 329], where μ and ν are Drury's two weight vectors (Lemma 7.1 and Remark 7.2(3) below).

In this section we prove Drury's result again, with the same polynomial and with a two-decimal matrix close to Drury's, so that every step is a short finite computation in exact rational arithmetic. On the matrix side, the proof needs a 3×3 integer Gram matrix and an integer test vector (Proposition 7.4); every number can be checked by hand. On the shift side, the proof uses the weighted window bound of Lemma 7.1, which is Drury's own argument, with a 15×15 positive definite Gram matrix (Proposition 7.3). Remark 7.6 treats Drury's own matrix in the same way. No new value of p is obtained here (Theorem A covers every $p \neq 2$, without numerical computation).

7.1. **A weighted window bound for the shift.** Let $1 < p < \infty$, $p' = p/(p-1)$, let $P(z) = \sum_{i=0}^d c_i z^i$ have complex coefficients, and put $c_t = 0$ for $t \notin \{0,\dots,d\}$. Let A be the convolution by (c_i) on $\ell^p(\mathbb{Z})$, $(Ax)_n = \sum_{i=0}^d c_i x_{n-i}$. As in the proof of Lemma 2.3, $\|P(S)\|_p = \|A\|_{\ell^p(\mathbb{Z})}$. Below we use only the inequality $\le$: A maps the sequences supported in $\mathbb{N}$ into themselves and acts on them as $P(S)$. For $K \ge 1$ let $W_0 = W_0(c,K)$ be the $(K+d)\times K$ matrix with entries $W_0[j,k] = c_{j-k}$ $(1 \le j \le K+d,\ 1 \le k \le K)$. For a vector v, D_v is the diagonal matrix with the entries of v on its diagonal.

Lemma 7.1 (weighted window bound)**.** *Let $K \ge 1$, $\mu \in (0,\infty)^K$ and $\nu \in (0,\infty)^{K+d}$. Put $s_1 = \sum_{k=1}^K \mu_k^{p'}$, $s_2 = \sum_{j=1}^{K+d} \nu_j^{p'}$ and $\widetilde{W} = D_\nu^{-1} W_0 D_\mu$, and let $\|\widetilde{W}\|_p$ be the norm of $\widetilde{W}$ from ℓ^p_K to ℓ^p_{K+d}. Then*

$$\|P(S)\|_p \;\le\; \|A\|_{\ell^p(\mathbb{Z})} \;\le\; \Big(\frac{s_2}{s_1}\Big)^{1/p'} \|\widetilde{W}\|_p.$$

This is the argument of Drury [6, Prop. 2], stated in general (any p, d, K and weights); Drury's case is $p = 4$, $d = 2$, $K = 5$.

Proof. The idea is to write A as an average of translated copies of $\widetilde{W}$ and to apply Hölder's inequality at each output site. Let $x \in \ell^p(\mathbb{Z})$ be finitely supported and $y = Ax$; all sums below are finite.

(a) Since $s_1^{-1}\sum_k \mu_k^{p'} = 1$, we have $y_n = s_1^{-1}\sum_{k=1}^{K}\sum_{i=0}^{d} c_i\,\mu_k^{p'}\,x_{n-i}$.

(b) Index the term (k,i) by the window $m = n-i-k$ and the slot $j = i+k = n-m$. Then $x_{n-i} = x_{m+k}$ and $c_i = c_{j-k}$. For fixed n and k, the map $i \mapsto m$ is a bijection from $\{0,\dots,d\}$ onto $\{n-k-d,\dots,n-k\}$, and $c_{n-m-k} = 0$ for every other m. Hence

$$y_n = s_1^{-1}\sum_{m\in\mathbb{Z}}\sum_{k=1}^{K} c_{n-m-k}\,\mu_k^{p'}\,x_{m+k},$$

and only the m with $1 \le n-m \le K+d$ contribute (as $0 \le j-k \le d$ and $1 \le k \le K$).

(c) As $p'-1 = p'/p$, $\mu_k^{p'} = \mu_k \cdot \mu_k^{p'/p}$. Define $w^{(m)} \in \mathbb{C}^K$ by $w_k^{(m)} = \mu_k^{p'/p}x_{m+k}$. Then $\sum_k c_{j-k}\mu_k w_k^{(m)} = (W_0 D_\mu w^{(m)})_j = \nu_j(\widetilde{W}w^{(m)})_j$, so

$$y_n = s_1^{-1}\sum_{m:\,1\le n-m\le K+d} \nu_{n-m}\left(\widetilde{W}w^{(m)}\right)_{n-m}.$$

(d) In this sum $j = n-m$ runs exactly once over $1,\dots,K+d$, so $\sum_m \nu_{n-m}^{p'} = s_2$, and Hölder's inequality with exponents p' and p gives

$$|y_n| \le s_1^{-1}s_2^{1/p'}\Big(\sum_{m:\,1\le n-m\le K+d}\big|(\widetilde{W}w^{(m)})_{n-m}\big|^p\Big)^{1/p}.$$

(e) Take p-th powers, sum over n and exchange the sums; each pair (m,j) occurs once, with $n = m+j$. So

$$\|y\|_p^p \le s_1^{-p}s_2^{p/p'}\sum_m \|\widetilde{W}w^{(m)}\|_p^p \le s_1^{-p}s_2^{p/p'}\|\widetilde{W}\|_p^p\sum_m\|w^{(m)}\|_p^p.$$

Here $\sum_m\|w^{(m)}\|_p^p = \sum_m\sum_k \mu_k^{p'}|x_{m+k}|^p = s_1\|x\|_p^p$, because each site l lies in slot k of the window $m = l-k$, once for each k.

(f) Hence, as $(1-p)/p = -1/p'$,

$$\|y\|_p^p \le s_1^{1-p}s_2^{p/p'}\|\widetilde{W}\|_p^p\|x\|_p^p, \qquad \text{that is,} \qquad \|Ax\|_p \le (s_2/s_1)^{1/p'}\|\widetilde{W}\|_p\,\|x\|_p.$$

The operator A is bounded ($\|A\| \le \sum_i |c_i|$), and the finitely supported sequences are dense in $\ell^p(\mathbb{Z})$ ($p < \infty$), so the bound holds for $\|A\|$. □

Remark 7.2. (1) The bound does not change if μ or ν is multiplied by a positive constant. For $K = 1$ it is at least $\sum_i |c_i|$, by Hölder's inequality, so only $K \ge 2$ can be useful.

(2) The bound is sharp as $K \to \infty$. The flat weights $\mu_k^{p'} = 1/K$, $\nu_j^{p'} = 1/(K+d)$ give $s_1 = s_2 = 1$ and the bound $((K+d)/K)^{1/p'}\|W_0\|$. For x supported in $\{1,\dots,K\}$, Ax is supported in $\{1,\dots,K+d\}$ and equals W_0x there, so $\|W_0(c,K)\| \le \|A\|$; and $\sup_K\|W_0(c,K)\| = \|A\|$, since every finitely supported vector can be translated into some $\{1,\dots,K\}$. Hence the infimum of the right side of Lemma 7.1 over K and the weights equals $\|A\| = \|P(S)\|_p$: every value $\|P(S)\|_p$ can be bounded from above, to any accuracy, by the norm of a single finite matrix. The difficulty moves to certifying a bound for $\|\widetilde{W}\|_p$.

(3) In Drury's Proposition 2 the 5×7 matrix $k_{m,n} = \mu_m c_{n-m}/\nu_n$ is the transpose of $\widetilde{W}$, and he bounds its norm on $\ell^{4/3}$, which is $\|\widetilde{W}\|_4$, by 6.52773. His weights are normalised so that s_1 and s_2 are 1 up to rounding; the factor $(s_2/s_1)^{3/4} = \|\nu\|_{4/3}/\|\mu\|_{4/3}$ is the one in his footnote 3.

7.2. **An upper bound for $\|P(S)\|_4$.**

Proposition 7.3. *For P as in* (12)*,*

$$\|P(S)\|_4 \;\le\; C_S := \frac{6529466499}{10^9} = 6.529466499.$$

The data. Take $p = 4$ (so $p' = 4/3$), $d = 2$, $K = 5$, and

$$a = \Big(\tfrac{6411}{10000}, \tfrac{3639}{5000}, \tfrac{3559}{5000}, \tfrac{6711}{10000}, \tfrac{5493}{10000}\Big), \qquad b = \Big(\tfrac{501}{1250}, \tfrac{5271}{10000}, \tfrac{6701}{10000}, \tfrac{1429}{2000}, \tfrac{6829}{10000}, \tfrac{157}{250}, \tfrac{623}{1250}\Big),$$
$$\mu_k = a_k^3, \qquad \nu_j = b_j^3, \qquad t_S = \tfrac{653}{100}, \qquad \lambda_S = \tfrac{9999183}{10^7}.$$

The numbers a_k and b_j are the cube roots of Drury's weights, rounded to four decimals. Taking cubes makes $\mu_k^{4/3} = a_k^4$ and $\nu_j^{4/3} = b_j^4$ rational. The last datum is a symmetric 15×15 rational matrix Q_S, whose rows and columns are indexed by the 15 monomials $v_i v_j$ $(i \le j)$ of $v \in \mathbb{R}^5$; it is given in the ancillary file `e_p4_exact/certificates/cert_S_p4_v020.txt`.

Proof. Each step is a finite computation in exact rational arithmetic.

(S1) $\widetilde{W} = D_\nu^{-1} W_0 D_\mu$ has the entries $\widetilde{W}[j,k] = c_{j-k}\, a_k^3 / b_j^3$, with $(c_0, c_1, c_2) = (1, 2, -\frac{22}{5})$; it is a rational 7×5 matrix.

(S2) The quartic form $F_S(v) = t_S^4 \sum_{k=1}^5 v_k^4 - \sum_{j=1}^7 (\widetilde{W}v)_j^4$ on $\mathbb{R}^5$ is identically equal to $Z^{\mathrm{t}} Q_S Z$, where $Z = (v_i v_j)_{i \le j}$. This compares the 70 coefficients of a quartic form in five variables.

(S3) Q_S is positive definite: symmetric Gaussian elimination without row exchanges gives 15 pivots, all positive (the smallest is about 5.87). Hence $F_S \ge 0$ on $\mathbb{R}^5$, that is, $\|\widetilde{W}v\|_4 \le t_S \|v\|_4$ for real v. As $\widetilde{W}$ is real, the proof of Lemma 2.1, which applies to every real matrix, gives $\|\widetilde{W}\|_4 \le t_S$ on complex ℓ_5^4.

(S4) Put $s_1 = \sum_k a_k^4$ and $s_2 = \sum_j b_j^4$. Then $\lambda_S^4 s_1^3 \ge s_2^3$, that is, $(s_2/s_1)^{3/4} \le \lambda_S$.

(S5) By Lemma 7.1, $\|P(S)\|_4 \le (s_2/s_1)^{3/4} \|\widetilde{W}\|_4 \le \lambda_S t_S = 6529466499/10^9$. □

Steps (S1)–(S5) were checked by two programs written independently of each other, each in exact rational arithmetic and each written from the statement above: `e_p4_exact/programs/verify_S.py` (log `e_p4_exact/logs_rerun/verify_S_run.log`) and `e_p4_exact/programs/verify_certS.py` (log `e_p4_exact/logs/verify_certS_v020.log`). As controls, a coarser certificate, with a_k and b_j rounded to three decimals (and a smaller λ), fails (S3) (certificate `e_p4_exact/certificates/cert_S_p4_v010_FAILS_coarse.txt`, log `e_p4_exact/logs_rerun/verify_S_control_coarse_run.log`), and lowering λ_S by 10^{-7} fails (S4), so λ_S is tight. The matrix Q_S was proposed by a semidefinite program in floating point; its entries were rounded to 10^{-5} and then corrected exactly so that (S2) holds, which gives some entries long denominators. How Q_S was found plays no role in the proof. Step (S4) is the step that Drury did not do exactly; here it is exact because the weights are rational cubes.

7.3. **A contraction T' with $\|P(T')\|_4 > 6.53254$.** Let

$$T' = \begin{pmatrix} -0.27 & 0.83 \\ -0.96 & -0.11 \end{pmatrix}.$$

This is Drury's matrix T_{Dr} (Remark 7.6) rounded to two decimals, $\left(\begin{smallmatrix} -0.29 & 0.81 \\ -0.96 & -0.11 \end{smallmatrix}\right)$, with the entries t_{11} and t_{12} each moved by $+0.02$; it was found by a search over the 5^4 two-decimal matrices whose entries differ from those of the rounded matrix by at most 0.02.

Proposition 7.4. *$\|T'\|_4 < 1$, on real and on complex ℓ_2^4. For P as in* (12) *and $v = (50, -53)$, $\|P(T')v\|_4 / \|v\|_4 > 6.53254$.*

Proof. The norm of T'. For real (x,y), $100\,T'(x,y)^{\mathsf t} = (-27x+83y,\ -96x-11y)^{\mathsf t}$. So $\|T'(x,y)\|_4 \le \|(x,y)\|_4$ is the inequality $\Phi(x,y)\ge 0$ for the integer binary quartic

$$\begin{aligned}\Phi(x,y) &:= 10^8(x^4+y^4) - (27x-83y)^4 - (96x+11y)^4\\ &= 14533903\,x^4 - 32393628\,x^3y - 36823302\,x^2y^2 + 61241892\,xy^3 + 52527038\,y^4.\end{aligned}$$

(For example, $10^8 - 27^4 - 96^4 = 10^8 - 531441 - 84934656 = 14533903$.) With $Z = (x^2, xy, y^2)$ we have $\Phi = Z^{\mathsf t} H Z$ for the integer matrix

$$H = \begin{pmatrix} 14533903 & -16196814 & -27500000\\ -16196814 & 18176698 & 30620946\\ -27500000 & 30620946 & 52527038\end{pmatrix}.$$

Indeed H_{11}, $2H_{12}$, $2H_{13}+H_{22}$, $2H_{23}$, H_{33} are the five coefficients of Φ; the only free entry is $H_{13} = -27500000$, and then $H_{22} = -36823302 + 2\cdot 27500000$. The leading principal minors of H are

$$m_1 = 14533903, \qquad m_2 = 1841581841698, \qquad m_3 = \det H = 899424407723373376,$$

all positive, so H is positive definite. Explicitly, with $r = H_{11}H_{23} - H_{12}H_{13} = -370526067762$,

$$\Phi = \frac{1}{m_1}\bigl(m_1x^2 - 16196814\,xy - 27500000\,y^2\bigr)^2 + \frac{1}{m_1m_2}\bigl(m_2\,xy + r\,y^2\bigr)^2 + \frac{m_3}{m_2}\,y^4,$$

as one checks by expanding. If $y\neq 0$ the last term is positive, and if $y=0\neq x$ the first is $m_1x^4>0$. So $\Phi>0$ off the origin, and by homogeneity and compactness of the unit sphere $\|T'\|_4<1$ on real vectors. As T' is real, Lemma 2.1 gives the same on complex vectors. (Numerically, $1-\|T'\|_4^4 = 5.0\times 10^{-4}$.)

The test vector. $T'^2 = \left(\begin{smallmatrix} -0.7239 & -0.3154\\ 0.3648 & -0.7847\end{smallmatrix}\right)$, so

$$P(T') = I + 2T' - \tfrac{22}{5}T'^2 = \begin{pmatrix} 3.64516 & 3.04776\\ -3.52512 & 4.23268\end{pmatrix}, \qquad 50000\,P(T') = \begin{pmatrix} 182258 & 152388\\ -176256 & 211634\end{pmatrix}.$$

For $v=(50,-53)$ this gives $50000\,P(T')v = (1036336,\ -20029402)$, and $\|v\|_4^4 = 50^4+53^4 = 14140481$. Hence

$$\frac{\|P(T')v\|_4^4}{\|v\|_4^4} = \frac{1036336^4 + 20029402^4}{50000^4\cdot 14140481} = 1821.0876\ldots, \qquad \frac{\|P(T')v\|_4}{\|v\|_4} = 6.5325495\ldots.$$

As integers, $(1036336^4 + 20029402^4)\cdot 10^{20} > 653254^4\cdot 14140481\cdot 50000^4$, which is the claim. □

Theorem 7.5 (Drury [6]; exact certificate)**.** *For $P(z) = 1+2z-\frac{22}{5}z^2$, the matrix T' is a contraction on ℓ_2^4 and*

$$\|P(T')\|_4 \;>\; 6.53254 \;>\; 6.529466499 \;\ge\; \|P(S)\|_4.$$

So Matsaev's conjecture fails at $p=4$; with the transpose $T'^{\mathsf t}$ it also fails at $p=4/3$. Every step of the proof is a finite computation in exact rational arithmetic.

Proof. Propositions 7.4 and 7.3; ℓ_2^4 is the L^4 space of two atoms. For $p=4/3$, by Lemma 2.3, $\|T'^{\mathsf t}\|_{4/3} = \|T'\|_4 < 1$ and $\|P(T'^{\mathsf t})\|_{4/3} = \|P(T')\|_4 > \|P(S)\|_4 = \|P(S)\|_{4/3}$. □

The relative margin is small, about 0.047%. The test vector also beats Drury's own bound $\|P(S)\|_4 < 6.5278$, so Proposition 7.4 together with [6, Prop. 2] is a second proof.

Remark 7.6 (Drury's own matrix). Drury's matrix is $T_{\mathrm{Dr}} = \left(\begin{smallmatrix} -0.292187 & 0.812218\\ -0.958489 & -0.113428\end{smallmatrix}\right)$ [6, Prop. 1], with exactly these entries (they are also the entries checked in his verification worksheet[2]); he shows $\|P(T_{\mathrm{Dr}})\|_4 > 6.5381$ with the test vector $\xi = (0.821725, -0.858838)$. The same two exact checks work for T_{Dr}. Write the unscaled quartic $x^4+y^4-(t_{11}x+t_{12}y)^4-(t_{21}x+t_{22}y)^4$, with the entries

[2] https://www.math.mcgill.ca/drury/research/matsaev/matsaev.html

t_{jk} of T_{Dr}, as $Z^{\mathsf{t}} H_{\mathrm{Dr}} Z$ with the free entry $(H_{\mathrm{Dr}})_{13} = -7243/25000$ (so $(H_{\mathrm{Dr}})_{22}$ is the coefficient of x^2y^2 minus $2(H_{\mathrm{Dr}})_{13}$). The three leading principal minors of H_{Dr} are exact rationals (with denominators of 24, 47 and 69 digits), and are positive: 1.486996×10^{-1}, 1.034277×10^{-5} and 9.811364×10^{-10}. So $\|T_{\mathrm{Dr}}\|_4 < 1$. Numerically $\|T_{\mathrm{Dr}}\|_4 = 0.999990231$, so all six decimals are needed. Already $v = (1, -1)$ gives $\|P(T_{\mathrm{Dr}})v\|_4 / \|v\|_4 = 6.533390 > C_S$ (an exact comparison); numerically $\|P(T_{\mathrm{Dr}})\|_4 = 6.538155009$, a margin of about 0.13% over C_S. The program `e_p4_exact/programs/tside.py` repeats these exact checks for T_{Dr}, and the Gram certificate of Proposition 7.4 for T' (log `e_p4_exact/logs/tside_v020.log`).

Remark 7.7 (Drury's guess). Drury wrote: "We think that $\|P(S)\|_{\ell^4 \to \ell^4} \approx 6.4548$ but this is a guess" [6, p. 324]. Theorem 8.1 bears this out: $\|P(S)\|_4 \leq 6.454826$, which is within 2×10^{-6} of the numerically computed $\|P(J)\|_4 = 6.454824957$, a lower bound for $\|P(S)\|_4$ by Lemma 2.2. So the margin of Drury's matrix over the shift is about 1.3%, not 0.13%. The bound 6.529466499 of Proposition 7.3 is about 1.2% above $\|P(J)\|_4$.

8. Shift norms of Drury's polynomial

In this section $P(z) = 1 + 2z - \frac{22}{5}z^2$ is Drury's polynomial (12). Since $J^2 = -I$, $P(J) = \frac{27}{5}I + 2J$, and Lemma 2.2 gives the lower bound $\|P(S)\|_p \geq \|P(J)\|_p$. We give upper bounds at $p = 3$ and at the even $p = 4, 6, \ldots, 20$, by window certificates in the sense of Lemma 3.1. Section 9 uses these bounds as upper bounds for $\|P(S)\|_p$ in item (c) of Theorem 9.1.

Theorem 8.1 (computer-assisted)**.** *Let $P(z) = 1 + 2z - \frac{22}{5}z^2$.*

(a) For $p \in \{4, 6, \ldots, 20\}$, $\|P(S)\|_p \leq C_p$, with C_p as in Table 1.

(b) $\|P(S)\|_3 \leq C_3 := 248143/40000 = 6.203575$.

By Lemma 2.3 the same bounds hold on $\ell^{p'}$.

The proof is a computer check of explicit certificates, in exact rational arithmetic (for (b) also in ball arithmetic); the two subsections below describe the certificates and the programs.

p	C_p	C_p (decimal)	size of $\mathcal{Q}$	$\|P(J)\|_p$ (numerical)
4	3227413/500000	6.454826	6	6.454824957
6	6737251/1000000	6.737251	10	6.737249200
8	13783/2000	6.891500	15	6.891498511
10	6988033/1000000	6.988033	21	6.988031138
12	1410781/200000	7.053905	28	7.053903233
14	3550829/500000	7.101658	36	7.101656124
16	3568923/500000	7.137846	45	7.137844301
18	7166211/1000000	7.166211	55	7.166209723
20	7189041/1000000	7.189041	66	7.189039566

Table 1. The bounds $\|P(S)\|_p \leq C_p$ for $P(z) = 1 + 2z - \frac{22}{5}z^2$, proved by the certificates of Definition 8.3; the size of their Gram matrix $\mathcal{Q}$ (the number of monomials of degree $p/2$ in three variables), and the lower bound $\|P(J)\|_p$ of Lemma 2.2, computed in floating point.

In every row, C_p is within 2×10^{-6} of the numerically computed $\|P(J)\|_p$, which is a lower bound for $\|P(S)\|_p$ by Lemma 2.2; so numerically $\|P(S)\|_p = \|P(J)\|_p$ at these p. At $p = 3$ the numerically computed value is $\|P(J)\|_3 = 6.202999138$, and the bound of part (b) is within 0.0093% of it.

Remark 8.2. The equality $\|P(S)\|_p = \|P(J)\|_p$ cannot hold for every p. At $p = 2$, $\|P(S)\|_2 = \max_{|z|=1} |P(z)| = \sqrt{19683/550} = 5.98224646\ldots$, while $\|P(J)\|_2 = |\frac{27}{5} + 2i| = 5.758472\ldots$.

8.1. **Sum-of-squares window certificates at even p.** Let $p = 2m$ be even and let $P(z) = c_0 + c_1 z + c_2 z^2$ be real. As in Lemma 3.1 we use windows $(u, v, w) = (x_n, x_{n-1}, x_{n-2})$ and $y = c_0 u + c_1 v + c_2 w$. Let $Z(u, v, w)$ be the vector of the $\binom{m+2}{2}$ monomials of degree m in u, v, w.

Definition 8.3. A *TISOS certificate* (translation-invariant sum of squares, window length 3) for the bound $\|P(S)\|_p \le C$ is a pair $(G, \mathcal{Q})$, where $G(s, t)$ is a real homogeneous polynomial of degree p in two variables and $\mathcal{Q}$ is a real symmetric positive semidefinite matrix indexed by the monomials in Z, such that, as polynomials in (u, v, w),

$$C^p u^p - y^p + G(v, w) - G(u, v) = Z(u, v, w)^{\mathsf{t}}\, \mathcal{Q}\, Z(u, v, w). \tag{13}$$

Write $\mathcal{Q} = \sum_i \lambda_i q_i q_i^{\mathsf{t}}$ with $\lambda_i \ge 0$ (spectral theorem). Then the right side of (13) is $\sum_i \lambda_i\, (q_i^{\mathsf{t}} Z)^2$, a sum of squares of polynomials. The name refers to the sum over all windows of a finitely supported real sequence x: the G terms telescope, as in the proof of Lemma 3.1, and

$$C^p \|x\|_p^p - \|P(S)x\|_p^p = \sum_n Z(x_n, x_{n-1}, x_{n-2})^{\mathsf{t}}\, \mathcal{Q}\, Z(x_n, x_{n-1}, x_{n-2}),$$

a single sum of squares whose terms are translates of each other.

Proposition 8.4. *If a TISOS certificate for $\|P(S)\|_p \le C$ exists, then $\|P(S)\|_p \le C$ (on complex ℓ^p).*

Proof. As p is even, $u^p = |u|^p$ and $y^p = |y|^p$, so the left side of (13) is the function F of (2) for this P and this G. By (13), $F \ge 0$ on $\mathbb{R}^3$, as $\mathcal{Q}$ is positive semidefinite. G is continuous, and $G(0, 0) = 0$ since G is homogeneous of degree $p > 0$. Lemma 3.1, whose proof applies verbatim to every real quadratic P, gives $\|P(S)\|_p \le C$. □

How the certificates are stored. The certificate files give C and $\mathcal{Q}$, but not G; they record (13) in a folded form, which determines G. Two monomials $u^{\alpha_1} v^{\alpha_2} w^{\alpha_3}$ of the same degree are in the same *translation class* if their exponent vectors $(\alpha_1, \alpha_2, \alpha_3)$ agree after the zero entries at both ends are deleted. For example, $u^3 v$ and $v^3 w$ are in the same class, but $u^3 v$ and $u v^3$ are not. The substitution $(u, v) \mapsto (v, w)$, which moves the window by one index, maps $u^a v^b$ to $v^a w^b$. So the classes are $\{u^p, v^p, w^p\}$; the pairs $\{u^a v^b,\ v^a w^b\}$ with $a, b \ge 1$; and the single monomials $u^a v^b w^c$ with $a, c \ge 1$.

Lemma 8.5. *Let $E(u, v, w)$ be a homogeneous polynomial of degree $p \ge 1$. Suppose that for each translation class, the coefficients of E on the monomials of that class add up to 0. Then $E = G(v, w) - G(u, v)$ for a homogeneous polynomial G of degree p.*

Proof. We treat the classes one at a time and add the results. A single monomial $u^a v^b w^c$ with $a, c \ge 1$ has coefficient 0. On a pair $\{u^a v^b, v^a w^b\}$ with $a, b \ge 1$, E has the form $e\, u^a v^b - e\, v^a w^b$, which is $G(v, w) - G(u, v)$ for $G(s, t) = -e\, s^a t^b$. On $\{u^p, v^p, w^p\}$, E has the form $e_u u^p + e_v v^p + e_w w^p$ with $e_u + e_v + e_w = 0$, which is $G(v, w) - G(u, v)$ for $G(s, t) = -e_u s^p + e_w t^p$. □

Apply Lemma 8.5 to $E = Z^{\mathsf{t}} \mathcal{Q} Z - (C^p u^p - y^p)$. If, for each translation class, the coefficients of $Z^{\mathsf{t}} \mathcal{Q} Z$ and of $C^p u^p - y^p$ have the same sum over the class (the *folded identity*), then (13) holds for some G. (Conversely, (13) implies the folded identity, since every $G(v, w) - G(u, v)$ has class sums 0; this is not needed.)

The computer check. For $P(z) = 1 + 2z - \frac{22}{5} z^2$ and $p = 4, 6, \ldots, 20$, the certificates are the ancillary files `cert_pp_W3_CC`$_p$`.txt` (for example `cert_p4_W3_C6.454826.txt`) in the folder `c_drury_28_20/certificates/`. Each gives p, the window length 3, the coefficients $1, 2, -\frac{22}{5}$, the number C_p, the list of the monomials of degree m and the upper triangle of $\mathcal{Q}$. A monomial is listed by the multiset of its window positions $0, 1, 2$; position i holds x_{n+i} in the window (x_n, x_{n+1}, x_{n+2}), so positions $0, 1, 2$ are w, v, u in our notation, and $y = c_0 u + c_1 v + c_2 w$. (The identity holds in this orientation, not in the mirrored one; the two orientations bound the same

norm, by the reflection $n \mapsto -n$ in the proof of Lemma 2.3.) Each certificate is checked in exact rational arithmetic by two programs written independently of each other, and once more by a third program; they check that

(i) the listed monomials are all the monomials of degree m in three variables, each once;
(ii) the folded identity holds, class by class;
(iii) $\mathcal{Q}$ is positive definite (all leading principal minors, or all pivots of symmetric elimination, are positive).

By Lemma 8.5 and Proposition 8.4 this proves Theorem 8.1(a). The programs and their logs, all in `c_drury_28_20/`, are:

- `tisos_verify.py`: leading principal minors by fraction-free (Bareiss) elimination; the coefficients of $C^p u^p - y^p$ by multiplying out dense polynomials. Log `logs/tisos_bellman_verify_v010.log`.
- `verify_tisos.py`: LDL^{t} elimination; it also checks the summed identity, with no folding, on 40 random integer sequences. Logs `logs/verify_tisos_1939.log`, `logs/verify_tisos_high1_2017.log` and `logs/verify_tisos_high2_2017.log`.
- `tisos_read.py`: its own parser and elimination; it tries both orientations of the window. Logs `logs_rerun/tisos_read_cert_p`p`_W3_C`C_p`_run.log`.

As controls, lowering C by 10^{-6}, or changing a single entry of $\mathcal{Q}$ by 10^{-6}, breaks (ii). The matrices $\mathcal{Q}$ were found by a semidefinite program in floating point and then made exact; how they were found plays no role in the check.

8.2. **A Bellman certificate at $p = 3$.** For odd p, $|t|^p$ is not a polynomial and (13) is not available. Instead we take an explicit G and check $F \geq 0$ in Lemma 3.1 by a certified branch-and-bound computation. We give a separate bound at $p = 3$ because interpolation (Riesz–Thorin) between the bounds at $p = 2$ and $p = 4$ gives only $\|P(S)\|_3 \leq 6.2933$; Section 9 uses it on $[2.8, 3.6]$ (Table 5).

Proposition 8.6. *Let $C_3 = 248143/40000$, let $h_3(\sigma_1, \sigma_2) = \sum_{j=1}^{8} \omega_j \, |\alpha_j \sigma_1 + \beta_j \sigma_2|^3$ with the rational numbers of Table 2, and $y = u + 2v - \frac{22}{5} w$. Then*

$$F_3(u, v, w) := C_3^3 \, |w|^3 - |y|^3 + h_3(w, v) - h_3(v, u) \;\geq\; 0 \qquad \text{for all } (u, v, w) \in \mathbb{R}^3, \tag{14}$$

and hence $\|P(S)\|_3 \leq C_3 = 6.203575$.

j	α_j	β_j	ω_j
1	1	0	$-5479839/125000$
2	231/250	383/1000	$-1674041/125000$
3	707/1000	707/1000	$-10856541/100000$
4	383/1000	231/250	55056591/1000000
5	0	1	19116003/500000
6	$-383/1000$	231/250	$-113888233/1000000$
7	$-707/1000$	707/1000	45200287/500000
8	$-231/250$	383/1000	$-372648/15625$

TABLE 2. The data of $h_3(\sigma_1, \sigma_2) = \sum_j \omega_j |\alpha_j \sigma_1 + \beta_j \sigma_2|^3$ in Proposition 8.6 (the three columns of the file `bell_p3_M8_m4.txt`).

Proof of the implication. Put $G(s, t) = h_3(t, s) + C_3^3\big(|s|^3 + |t|^3\big)$. It is continuous, $G(0, 0) = 0$, and $G(v, w) - G(u, v) = h_3(w, v) - h_3(v, u) + C_3^3\big(|w|^3 - |u|^3\big)$. Hence

$$C_3^3 |u|^3 - |y|^3 + G(v, w) - G(u, v) = F_3(u, v, w) \geq 0,$$

and Lemma 3.1 (for $p = 3$ and the real quadratic P) gives $\|P(S)\|_3 \le C_3$. The weights ω_j have both signs; no step needs $h_3 \ge 0$. □

The computer check of (14). F_3 is even and $F_3(t\xi) = |t|^3 F_3(\xi)$ for real t. Every $\xi \ne 0$ in $\mathbb{R}^3$ is $\pm \max_i |\xi_i|$ times a point of one of the three faces $\{\xi_k = 1,\ |\xi_i| \le 1\ (i \ne k)\}$ of the cube $[-1,1]^3$. So it suffices that $F_3 > 0$ on these faces. Three programs, written independently of each other, check this. Each covers the faces by boxes and accepts a box only with a proved positive lower bound for F_3 on it.

(i) `bellman_verify.py`, in ball arithmetic (python-flint `arb`): adaptive bisection; a box is accepted when a lower bound for F_3 at its centre exceeds $\sum_i r_i \sup_{\text{box}} |\partial F_3/\partial \xi_i|$, with r_i the half-widths (mean value theorem; F_3 is C^1). Log `logs/tisos_bellman_verify_v010.log`.

(ii) `verify_bellman_p3.py`, in exact rational arithmetic: the same kind of branch and bound, with F_3 at the centre computed exactly ($p = 3$ is an integer) and exact interval bounds for the derivatives. Log `logs/verify_bellman_p3_0031.log`. In its control run (`logs/verify_bellman_p3_controlgrid_0031.log`), C_3 is lowered to $6.2029 < \|P(J)\|_3$, and an exact search on a 200×200 grid of each face finds a point with $F_3 < 0$.

(iii) `bellman_bernstein.py`, in exact rational arithmetic, by a different method: on a box where a linear form $\alpha\sigma_1 + \beta\sigma_2$ has constant sign (decided at the corners), the term $\omega|\alpha\sigma_1 + \beta\sigma_2|^3$ is a polynomial; the sum of these terms is written in the tensor Bernstein basis of degree $(3,3)$ on the box, and its least coefficient is a lower bound; a term $\omega|\alpha\sigma_1 + \beta\sigma_2|^3$ whose form changes sign on the box contributes the lower bound $\min(0,\omega) \max_{\text{corners}} |\alpha\sigma_1 + \beta\sigma_2|^3$. The three faces need 28, 70 and 82 boxes. Log `logs/bellman_bernstein_p3_v020.log`. Its control runs (`logs/bellman_bernstein_p3_controls_v020.log`) fail: for example, with $C = 6.2029 < \|P(J)\|_3$ it finds rational points where $F_3 < 0$.

All files are in `c_drury_28_20/`. In the certificate file `certificates/bell_p3_M8_m4.txt`, the inequality is written for the window $(x_n, x_{n+1}, x_{n+2}) = (w, v, u)$, as $C^3|u_0|^3 - |c_2u_0 + c_1u_1 + c_0u_2|^3 + h(u_0,u_1) - h(u_1,u_2)$, which is (14). In floating point the minimum of F_3 on the faces is about 0.04, that is about 1.7×10^{-4} of $C_3^3 = 238.74$.

9. Explicit counterexamples for every $p \in [40/39, 40]$, $p \ne 2$

Theorem A is an existence statement. For each $p \ne 2$ it needs a small coefficient a, and the admissible range $0 < a < a_0$ must shrink as $p \to 2$ (Remark 6.2); a single polynomial can still violate (1) on a whole punctured neighbourhood of $p = 2$ (Theorem 6.3; item (a) below). This section gives explicit counterexamples: four explicit real quadratics and explicit real 2×2 contractions, which together cover every $p \in [40/39, 40]$ with $p \ne 2$. It gives in full the explicit counterexamples announced in [8], and it contains everything needed to check the claims: the definitions, the criterion that each computer check verifies, a worked example, the anchor values, the margins, and the names of the files that certify and re-check each claim. Nothing in this section is used in the proofs of Theorems A and B.

The proof is computer-assisted. It reduces to finitely many inequalities, each checked in exact rational arithmetic or in interval (ball) arithmetic, and each checked by at least two programs written independently of each other. All certificates, programs and logs are in the ancillary files of this paper (Section 1). Some file names contain run labels (such as `A7`, `R2` or `v070`); they have no mathematical meaning.

9.1. The four polynomials and the main theorem.

Throughout this section $s = 1/p$ and $p' = p/(p-1)$; $\|A\|_p$ and R_θ are as in Section 1. By Lemma 2.1 all norms below are the same over

$\mathbb{R}$ and over $\mathbb{C}$. The four polynomials are

$$\begin{aligned} P_{7/10}(z) &= 1 + 2z - \tfrac{17}{10}z^2, & P_A(z) &= 1 + \tfrac{13043}{6250}z - \tfrac{1789439}{500000}z^2, \\ P_{\mathrm{Dr}}(z) &= 1 + 2z - \tfrac{22}{5}z^2, & P_D(z) &= 1 + z - 8z^2. \end{aligned} \tag{15}$$

(The subscripts 7/10, A, Dr, D are names; they are not values of the coefficient a of P_a in Part I.) P_{Dr} is Drury's polynomial [6], and $P_A = 1 + 2.08688\,z - 3.578878\,z^2$ in decimals. All four have the form $1 + az - Bz^2$ with $a, B > 0$, the form used in Lemma 3.1.

Theorem 9.1. *Let $p \in [40/39, 40]$, $p \neq 2$, and let $q = \max(p, p')$, so that $q \in (2, 40]$. Let (P, T) be given by the first item that applies:*

- *(a) $q \in (2, 2.1]$: $P = P_{7/10}$ and $T = T_q = R_\varphi/U(q)$, where $\varphi \in (\pi/2, \pi)$, $\cos\varphi = -\frac{7}{34}$, and $U(q) \geq \|R_\varphi\|_q$ is the explicit bound of Definition 9.12;*
- *(b) $q \in [2.1, 2.8]$: $P = P_A$, and $T = T_i/\varrho_i$, where $[p_{\mathrm{lo}}, p_{\mathrm{hi}}]$ is a piece of item (b) (Definition 9.9; there are 15) with $p_{\mathrm{lo}} \leq q \leq p_{\mathrm{hi}}$, and T_i, ϱ_i are its matrix and its norm bound;*
- *(c) $q \in [2.8, 20]$: $P = P_{\mathrm{Dr}}$, and $T = T_i/\varrho_i$, where $[p_{\mathrm{lo}}, p_{\mathrm{hi}}]$ is a piece of item (c) (there are 445) with $p_{\mathrm{lo}} \leq q \leq p_{\mathrm{hi}}$, and T_i, ϱ_i are its matrix and its norm bound;*
- *(d) $q \in [20, 40]$: $P = P_D$, and $T = T_i/\varrho_i$, where $[p_{\mathrm{lo}}, p_{\mathrm{hi}}]$ is a piece of item (d) (there are 100) with $p_{\mathrm{lo}} \leq q \leq p_{\mathrm{hi}}$, and T_i, ϱ_i are its matrix and its norm bound.*

In (b)–(d), T_i is a real matrix with six-decimal entries and ϱ_i is a number with 8 decimals; both are printed on the line of the piece in the log named in Section 9.6, 9.7 or 9.8. Put $T_* = T$ if $p > 2$ and $T_* = T^{\mathsf{t}}$ if $p < 2$. Then

$$\|T_*\|_p \leq 1 \qquad \text{and} \qquad \|P(T_*)\|_p \;>\; \|P(S)\|_p,$$

so Matsaev's inequality (1) *fails at p, for a real quadratic and a real 2×2 contraction.*

For every piece of (b)–(d), the programs `verify_R2.py` and `R2_read.py` read p_{lo}, p_{hi}, T_i and ϱ_i from the log and prove that $\varrho_i \geq \max(\|T_i\|_{p_{\mathrm{lo}}}, \|T_i\|_{p_{\mathrm{hi}}})$ and that $L > \widehat{C}(p_{\mathrm{hi}})$ for T_i/ϱ_i (Definition 9.9, Lemma 9.10, Section 9.2).

For $p < 2$ the four ranges are $[21/11, 2)$, $[14/9, 21/11]$, $[20/19, 14/9]$ and $[40/39, 20/19]$, the duals of $(2, 2.1]$, $[2.1, 2.8]$, $[2.8, 20]$ and $[20, 40]$. The relative margins (Definition 9.9) are between 0.2900% and 1.6958% in (b), between 0.0139% and 1.9101% in (c), and between 0.0228% and 0.0994% in (d); in (a) the relative margin exceeds 0.0329% for $p \in [400/199, 2.1]$ and tends to 0 as $p \to 2$.

The rest of the section is organised as follows. Section 9.2 defines anchors, window certificates, the interpolation between anchors and the pieces, and proves the criterion that every computer check verifies. Section 9.3 works out one piece of item (c) in full. Section 9.4 lists all anchors. Sections 9.5–9.8 treat items (a)–(d). Section 9.9 lists what was checked twice and by which programs, and Section 9.10 assembles the proof of Theorem 9.1.

9.2. **Anchors, window certificates and pieces.**

Anchors.

Definition 9.2 (anchor)**.** Let P be a real polynomial. An *anchor* for P is a value $p_j \in [2, \infty]$ together with a number $C(p_j)$ and a proof that $\|P(S)\|_{p_j} \leq C(p_j)$. At $p_j = 2$ we take $C(p_j) = M_P := \max_{|z|=1} |P(z)|$, at $p_j = \infty$ we take $C(p_j) = \sum |c_k|$ (Lemma 9.3), and at $2 < p_j < \infty$ the proof is a window certificate (Definitions 9.4 and 8.3).

Only *upper* bounds for the shift are ever used in this section.

Lemma 9.3 (the anchors at $p = 2$ and $p = \infty$)**.** *For every polynomial $P = \sum c_k z^k$, $\|P(S)\|_2 = M_P$ and $\|P(S)\|_\infty \leq \sum |c_k|$. If $P = c_0 + c_1 z + c_2 z^2$ is real with $c_0 c_2 < 0$, then, with $x = \cos t$,*

$$|P(e^{it})|^2 = (c_0 - c_2)^2 + c_1^2 + 2c_1(c_0 + c_2)\,x + 4c_0c_2\,x^2, \tag{16}$$

a concave function of x; if its vertex $x_* = -c_1(c_0+c_2)/(4c_0c_2)$ lies in $[-1,1]$, then $M_P^2 = (c_0 - c_2)^2 + c_1^2 - c_1^2(c_0+c_2)^2/(4c_0c_2)$ is attained exactly at $\cos t = x_*$.

Proof. On $\ell^2(\mathbb{Z})$, $P(S)$ is convolution by (c_k); by Plancherel's theorem it is multiplication by $P(e^{it})$ on L^2 of the circle, with norm M_P. The norm on $\ell^2(\mathbb{N})$ is the same (translate finitely supported vectors, as in the proof of Lemma 2.3). For $p=\infty$, $|(P(S)x)_n| \le \sum_k |c_k||x_{n-k}| \le \sum |c_k|\,\|x\|_\infty$. Formula (16) follows from $|P(e^{it})|^2 = c_0^2+c_1^2+c_2^2+2c_1(c_0+c_2)\cos t + 2c_0c_2\cos 2t$ and $\cos 2t = 2x^2-1$; it is concave in x because $c_0c_2<0$, and its maximum on $[-1,1]$ is at the vertex when the vertex lies in $[-1,1]$. □

P	$x_* = \cos\varphi_P$	M_P^2	$M_P \approx$	$\sum \lvert c_k\rvert$
$P_{7/10}$	$-7/34$	$19683/1700$	3.4026806	$47/10$
P_A	$-16818152877/44735975000$	$\frac{7645501823722262953633}{2795998437500000000000}$	5.2291914	$3332879/500000$
P_{Dr}	$-17/44$	$19683/550$	5.9822465	$37/5$
P_D	$-7/32$	$2673/32$	9.1395432	10

TABLE 3. The exact anchors at $p=2$ and $p=\infty$ (Lemma 9.3); $\varphi_P \in (\pi/2,\pi)$ is the angle where $|P(e^{it})|$ is maximal.

Bellman certificates.

Definition 9.4 (Bellman certificate)**.** Let $P = c_0 + c_1 z + c_2 z^2$ be real. A *Bellman certificate* for $\|P(S)\|_p \le C$ consists of rational numbers $p>1$, $C>0$ and finitely many rational triples $(\alpha_j,\beta_j,\omega_j)$. Put

$$h(\sigma_1,\sigma_2) = \sum_j \omega_j\,|\alpha_j\sigma_1 + \beta_j\sigma_2|^p,$$
$$F(u_0,u_1,u_2) = C^p|u_0|^p - |c_2u_0 + c_1u_1 + c_0u_2|^p + h(u_0,u_1) - h(u_1,u_2). \tag{17}$$

The certificate is *valid* if $F \ge 0$ on $\mathbb{R}^3$.

The certificate files list p, the coefficients, C, the number of terms and a line $\alpha_j\ \beta_j\ \omega_j$ for each term; the anchors of this section use 8 or 12 terms (Table 4). Here $(u_0,u_1,u_2) = (x_{n-2},x_{n-1},x_n)$, so $c_2u_0 + c_1u_1 + c_0u_2 = (P(S)x)_n$.

Lemma 9.5. *If a Bellman certificate for $P = 1 + az - Bz^2$ is valid, then $\|P(S)\|_p \le C$.*

Proof. This is the proof of the implication in Proposition 8.6, with 3, C_3 and h_3 replaced by p, C and h. In the window $(u,v,w) = (x_n,x_{n-1},x_{n-2}) = (u_2,u_1,u_0)$ of Lemma 3.1, the continuous function $G(\sigma_1,\sigma_2) = h(\sigma_2,\sigma_1) + C^p(|\sigma_1|^p + |\sigma_2|^p)$, which has $G(0,0)=0$, gives $C^p|u|^p + G(v,w) - G(u,v) - |u + av - Bw|^p = F(w,v,u) \ge 0$. This is (2) with $a=c_1$ and $B=-c_2$, and Lemma 3.1 gives $\|P(S)\|_p \le C$ (its proof uses only $G(0,0)=0$; all anchors of this section have $p>2$, as in that lemma). □

Lemma 9.6 (reduction to three faces)**.** *The function F of (17) satisfies $F(-u) = F(u)$ and $F(tu) = t^pF(u)$ for $t>0$. Hence $F \ge 0$ on $\mathbb{R}^3$ if and only if $F \ge 0$ on the three faces $\mathcal{C}_k = \{u : u_k = 1,\ |u_i| \le 1\ (i \ne k)\}$, $k=0,1,2$, of the cube $[-1,1]^3$.*

Proof. Every term of F is a multiple of $|\ell(u)|^p$ with ℓ linear, which gives both identities. If $u \ne 0$, let k be an index with $|u_k| = \max_i |u_i|$; then $u = \pm|u_k|\,u'$ with $u' \in \mathcal{C}_k$, and $F(u) = |u_k|^pF(u')$. □

How a Bellman certificate is checked. Each face is cut into boxes, and on each box a rigorous lower bound for F is computed; a box whose bound is not positive is cut into smaller boxes, and the check fails if the subdivision does not terminate. All data are rational, the box corners are exact rationals, and the evaluation uses ball arithmetic (the Arb library [10] through python-flint) or exact rational arithmetic. Three kinds of lower bound are used by the different programs:

(i) first order (mean value theorem; F is C^1 for $p > 1$): $F \geq F(u^\circ) - \sum_i \sup_{\text{box}} |\partial_i F|\, r_i$, with u° the centre and r_i the half-widths; the derivative of $|\ell|^p$ in ℓ, $p\,\mathrm{sgn}(\ell)|\ell|^{p-1}$, is increasing, so it is bounded by its values at the extreme values of ℓ, which are taken at box corners (`bellman_check.py`, `bellman_verify.py`; for integer p in exact rationals, `verify_bellman_p3.py`, `verify_bellman_evenp.py`);

(ii) second order (for $p \geq 2$): each term $\omega|\ell|^p$ is replaced by a quadratic lower bound on the box. If ℓ keeps a strict sign on the box and $\min|\ell| \geq 2\max|\delta|$, where $\delta = |\ell| - |\ell(u^\circ)|$, u° is the centre and min, max are taken over the box, the bound is the Taylor polynomial of degree 2 of $|\ell|^p$ at u° with the remainder bound $p(p-1)(p-2)\max\bigl(\min|\ell|^{p-3}, \max|\ell|^{p-3}\bigr)\max|\delta|^3/6$. Otherwise it uses convexity: $|\ell|^p$ is at least its tangent at $\ell(u^\circ)$ (used if $\omega > 0$) and at most $M^{p-2}\ell^2$ with $M = \max|\ell|$ (used if $\omega < 0$). The resulting quadratic is bounded below through its coefficients in the Bernstein basis (`bellman_taylor.py`);

(iii) for integer p: on a box where ℓ has constant sign, $|\ell|^p = \pm\ell^p$ is a polynomial, expanded exactly in the Bernstein basis of the box (`bellman_bernstein.py`).

Translation-invariant sum-of-squares certificates. The anchors at $p = 4, 6, \ldots, 20$ of item (c) and at $p = 20, 22, \ldots, 32$ of item (d) are TISOS certificates in the sense of Definition 8.3: for $p = 2m$, a positive semidefinite matrix $\mathcal{Q}$, indexed by the $(m+1)(m+2)/2$ monomials of degree m in three variables, for which (13) holds. By Proposition 8.4 such a certificate proves $\|P(S)\|_p \leq C$. The files store (13) in the folded form of Lemma 8.5, and in their names `W3` records the window length 3. The files of item (d) also have a line `tau`, which gives a rational number γ with $0 < \gamma \leq C^p$; in these files γ takes the place of C^p in (13), so each is a TISOS certificate for the bound $\gamma^{1/p}$, and $\gamma^{1/p} \leq C$.

How a TISOS certificate is checked (all steps in exact rational arithmetic; for item (c) the programs are listed in Section 8). (V1) The listed monomials are exactly all monomials of degree m in three variables, each once. (V2) The folded identity holds, class by class (Lemma 8.5). (V3) $\mathcal{Q}$ is positive definite: all leading principal minors are positive (Sylvester's criterion), computed by fraction-free elimination, or an exact LDL^{t} factorisation has positive pivots. (V4) For the files with a line `tau`, $0 < \gamma \leq C^p$. The program `verify_tisos.py` also verifies the summed identity directly on random integer sequences, with no folding. Two orientations of the convolution are possible in (13) ($P(S)$ or its reflection $n \mapsto -n$); they have the same norm (proof of Lemma 2.3), and the program `tisos_read.py` tries both.

Interpolation between anchors.

Definition 9.7 (interpolated shift bound)**.** Let $2 = p_0 < p_1 < \cdots < p_r \leq \infty$ be anchors for P with values $C(p_0) \leq C(p_1) \leq \cdots \leq C(p_r)$, and $s_j = 1/p_j$ ($s_r = 0$ if $p_r = \infty$). For $p \in [p_{j-1}, p_j]$ put $\lambda = (s_{j-1} - 1/p)/(s_{j-1} - s_j) \in [0,1]$ and

$$\widehat{C}(p) = C(p_{j-1})^{1-\lambda}\, C(p_j)^{\lambda}, \qquad \text{that is,} \qquad \log\widehat{C}(1/s) = \tfrac{s-s_j}{s_{j-1}-s_j}\log C(p_{j-1}) + \tfrac{s_{j-1}-s}{s_{j-1}-s_j}\log C(p_j).$$

Lemma 9.8 (log-convexity)**.** *(i) $\|P(S)\|_p \leq \widehat{C}(p)$ for every $p \in [2, p_r]$. (ii) $\widehat{C}$ is non-decreasing in p. (iii) For a real 2×2 matrix T, $s \mapsto \log\|T\|_{1/s}$ is convex on $[0,1]$; so for $p_{\mathrm{lo}} \leq p \leq p_{\mathrm{hi}}$, $\|T\|_p \leq \max(\|T\|_{p_{\mathrm{lo}}}, \|T\|_{p_{\mathrm{hi}}})$.*

Proof. (i) and (iii): by the Riesz–Thorin theorem [3, Thm. 1.1.1], for a linear operator A on complex ℓ^p spaces, $\log\|A\|_{1/s}$ is a convex function of $s \in [0,1]$. Apply this to $A = P(S)$ on $\ell^p(\mathbb{N})$ and to

$A = T$ on ℓ_2^p (complex norms equal real norms by Lemma 2.1; at $p = \infty$ the bound $\sum |c_k|$ of Lemma 9.3 holds for complex sequences too). Between s_j and s_{j-1} a convex function lies below its chord, which gives $\log\|P(S)\|_p \le \log \widehat{C}(p)$; and a convex function on an interval is at most the larger of its values at the ends. (ii) $\log \widehat{C}(1/s)$ is continuous and piecewise linear in s, and on $[s_j, s_{j-1}]$ its slope $(\log C(p_{j-1}) - \log C(p_j))/(s_{j-1} - s_j)$ is ≤ 0, because $C(p_{j-1}) \le C(p_j)$. So $\widehat{C}(1/s)$ does not increase with s, that is, $\widehat{C}(p)$ does not decrease with p. □

The anchor values of all four polynomials increase with p (Table 4), and the certifier program asserts this before it uses them.

Pieces.

Definition 9.9 (piece, certified norm bound, test vector, relative margin)**.** A *piece* is an interval $[p_{\mathrm{lo}}, p_{\mathrm{hi}}] \subset [2, \infty)$ with rational ends, together with a real 2×2 matrix T with decimal entries. A *certified norm bound* for T on the piece is a number ϱ with a proof (in interval arithmetic) that $\varrho \ge \|T\|_{p_{\mathrm{lo}}}$ and $\varrho \ge \|T\|_{p_{\mathrm{hi}}}$. A *test vector* is any $v \in \mathbb{R}^2 \setminus \{0\}$. The *matrix-side bound* of the piece is

$$L = \frac{\|P(T/\varrho)\, v\|_{p_{\mathrm{hi}}}}{\|v\|_{p_{\mathrm{lo}}}}$$

(the programs compute an interval enclosure of L and use its lower end), the *shift-side bound* is $\widehat{C}(p_{\mathrm{hi}})$, and the *relative margin* of the piece is $L/\widehat{C}(p_{\mathrm{hi}}) - 1$, printed in percent.

Lemma 9.10 (piece criterion)**.** *If $L > \widehat{C}(p_{\mathrm{hi}})$, then for every $p \in [p_{\mathrm{lo}}, p_{\mathrm{hi}}]$*

$$\|T/\varrho\|_p \le 1 \qquad \text{and} \qquad \|P(T/\varrho)\|_p \;\ge\; L \;>\; \widehat{C}(p_{\mathrm{hi}}) \;\ge\; \|P(S)\|_p.$$

In particular $\|P(T/\varrho)\|_p/\|P(S)\|_p - 1$ is at least the relative margin, for every p in the piece.

Proof. (1) By Lemma 9.8(iii), $\|T\|_p \le \varrho$, so T/ϱ is a contraction. (2) For a fixed vector y, $p \mapsto \|y\|_p$ is non-increasing; hence $\|P(T/\varrho)v\|_p \ge \|P(T/\varrho)v\|_{p_{\mathrm{hi}}}$ and $\|v\|_p \le \|v\|_{p_{\mathrm{lo}}}$, so $\|P(T/\varrho)\|_p \ge \|P(T/\varrho)v\|_p/\|v\|_p \ge L$. (3) By Lemma 9.8(i)–(ii), $\|P(S)\|_p \le \widehat{C}(p) \le \widehat{C}(p_{\mathrm{hi}})$. □

Lemma 9.11 ($p < 2$)**.** *Let T be a real 2×2 matrix and P a real polynomial. If $\|T\|_p \le 1$ and $\|P(T)\|_p > \|P(S)\|_p$, then $\|T^{\mathsf{t}}\|_{p'} \le 1$ and $\|P(T^{\mathsf{t}})\|_{p'} > \|P(S)\|_{p'}$.*

Proof. Lemma 2.3: $\|T^{\mathsf{t}}\|_{p'} = \|T\|_p$, $\|P(T^{\mathsf{t}})\|_{p'} = \|P(T)\|_p$ and $\|P(S)\|_{p'} = \|P(S)\|_p$. □

The certifier. The program `interval_certify.py` (the same file in parts b, c and d) takes a range of p, a step, a minimal step, the coefficients of P, an anchor table and a table of matrices. It walks through the range in pieces of the given step; for each piece it takes the row of the matrix table whose p is nearest to the midpoint of the piece, proves a norm bound at both ends, chooses a test vector, and checks $L > \widehat{C}(p_{\mathrm{hi}})$. If a piece fails, it halves the step (not below the minimal step). It works in interval arithmetic (mpmath [12], 30 digits); the piece ends are exact rationals. The norm bound: every real direction is $\pm(\cos\vartheta, \sin\vartheta)$ with $\vartheta \in [0, \pi]$; a floating point search gives a target value (the largest ratio found, times $1 + 10^{-6}$), and the interval $[0, \pi]$ is cut into cells until, on every cell, an interval enclosure of $\|T(\cos\vartheta, \sin\vartheta)\|_p/\|(\cos\vartheta, \sin\vartheta)\|_p$ is below the target; then ϱ is the larger of the two targets. The test vector is found by a floating point search (it maximises $\|P(T/\varrho)x\|_p/\|x\|_p$ at the midpoint); only the final inequality is certified, so how v was found does not matter. For each piece the log prints a line with the piece, T, ϱ (as "R="), L (as "T-side>="), $\widehat{C}(p_{\mathrm{hi}})$ (as "shift<=") and the relative margin; at the end it prints that the pieces tile the range with no gap. The printed values of ϱ, L and $\widehat{C}$ are rounded to 8, 6 and 6 decimals; the certifier compares before rounding, and two of the re-checking programs below prove the criterion again with the printed ϱ.

The re-checking programs. Every piece line of every certifier log is re-proved by two further programs: `read_R2_pieces.py` and `verify_R2.py` in item (b), `verify_R2.py` and `R2_read.py` in item (c), `R2_read.py` and `read_pieces.py` in item (d). Each reads p_{lo}, p_{hi} and T from the log and proves the piece criterion again, with its own certified norm bound, its own test vector and its own interpolation, in ball arithmetic [10] with exact rational inputs. Their norm bounds use other methods than the certifier: `verify_R2.py`, `R2_read.py` and `read_R2_pieces.py` write every direction as a multiple of $(1,t)$ or $(t,1)$ with $|t| \le 1$, and `read_pieces.py` uses angle cells with a mean-value bound and bounds the matrix side on 8 sub-pieces of each piece. The anchor values are read from the certificate files, except in `read_R2_pieces.py`, which reads the anchor table and takes the maximum of $\widehat{C}$ over the whole piece. So in this mode the contraction proved by a re-checker is T/ϱ' with its own ϱ'. In a second mode (`--printed-R` and `--verify-printed-rho`), `verify_R2.py` and `R2_read.py` also read the printed ϱ and prove $\varrho \ge \max(\|T\|_{p_{\mathrm{lo}}}, \|T\|_{p_{\mathrm{hi}}})$ and the piece criterion for T/ϱ itself, on every piece of (b), (c) and (d); this is the claim of Theorem 9.1. In their control runs every ϱ is multiplied by 99999/100000. Every re-check proves the piece criterion (Lemma 9.10) on every piece.

9.3. **A worked example.** We go through one piece of item (c) step by step, using only the definitions above and a few lines of arithmetic. The polynomial is Drury's $P_{\mathrm{Dr}}(z) = 1 + 2z - \frac{22}{5}z^2$, and the piece is $[3.59, 3.6] = [359/100, 18/5]$ of the run on $[3.55, 4]$ (line 15 of `c_drury_28_20/logs/cert_left_tisos_v070.log`, broken here into three lines; ϱ is printed as `R=`):

```
[359/100, 18/5] = [3.5900000, 3.6000000]
    T=-0.301279 0.818028 -0.955951 -0.131758  R=1.00000101
    T-side>=6.461403  shift<=6.400525  margin=+0.9511%
```

Step 1: the matrix and its norm bound. The matrix is the row $p = 3.60$ of `c_drury_28_20/tables/T_table_1836.txt`:

$$T = \begin{pmatrix} -0.301279 & 0.818028 \\ -0.955951 & -0.131758 \end{pmatrix}.$$

In floating point, $\|T\|_{3.59} \approx 0.9998831$ and $\|T\|_{3.6} \approx 1.0000000057$, so $\|T\|_{3.6}$ slightly exceeds 1. The certifier proves $\|T\|_{3.59} \le 0.99988412$ and $\|T\|_{3.6} \le \varrho = 1.00000101$. By Lemma 9.8(iii), T/ϱ is a contraction on ℓ_2^p for every $p \in [3.59, 3.6]$.

Step 2: the test vector and the matrix side. Take $v = (1, -1.03683)$ (the certifier's vector, rescaled so that $v_1 = 1$ and rounded to 6 significant digits). Then

$$P_{\mathrm{Dr}}(T/\varrho) = I + 2\tfrac{T}{\varrho} - \tfrac{22}{5}\big(\tfrac{T}{\varrho}\big)^2 \approx \begin{pmatrix} 3.438829 & 3.194691 \\ -3.733330 & 4.100869 \end{pmatrix},$$

$y = P_{\mathrm{Dr}}(T/\varrho)v \approx (0.126478,\ -7.985234)$, $\|y\|_{3.6} \approx 7.985235$ and $\|v\|_{3.59} \approx 1.235836$, so

$$L = \|y\|_{3.6}/\|v\|_{3.59} = 6.4614024\ldots$$

(with the certifier's own unrounded ϱ and v the value is $6.4614025\ldots$; the log prints 6.461403). By step (2) of the proof of Lemma 9.10, $\|P_{\mathrm{Dr}}(T/\varrho)\|_p \ge 6.461402$ for every $p \in [3.59, 3.6]$.

Step 3: the shift side. The anchor table of this run (`c_drury_28_20/tables/anchors_tisos_4_10.txt`) has anchors at $p = 2, 4, 6, 8, 10, \infty$. The two adjacent to $p \in [3.59, 3.6]$ are $p = 2$, with the exact value $M_{P_{\mathrm{Dr}}} = \sqrt{19683/550} \approx 5.9822465$, and $p = 4$, with $C_4 = 3227413/500000 = 6.454826$ from the TISOS certificate `c_drury_28_20/certificates/cert_p4_W3_C6.454826.txt` (Theorem 8.1(a)). For $2 \le p \le 4$, $\lambda = (\frac{1}{2} - \frac{1}{p})/(\frac{1}{2} - \frac{1}{4}) = 2 - \frac{4}{p}$; at $p = 3.6$, $\lambda = \frac{8}{9}$ and

$$\widehat{C}(3.6) = M_{P_{\mathrm{Dr}}}^{1/9}\, C_4^{8/9} = 6.4005254\ldots$$

(the log prints 6.400525). By Lemma 9.8, $\|P_{\mathrm{Dr}}(S)\|_p \le \widehat{C}(p) \le 6.4005254$ for every $p \in [3.59, 3.6]$.

Step 4: conclusion. $6.461402 > 6.4005254$; the relative margin is $L/\widehat{C}(3.6) - 1 = 0.9511\%$. So for every $p \in [3.59, 3.6]$ the real matrix T/ϱ is a contraction on ℓ_2^p with $\|P_{\mathrm{Dr}}(T/\varrho)\|_p > \|P_{\mathrm{Dr}}(S)\|_p$, and by Lemma 9.11, T^{t}/ϱ does the same for every $p \in [18/13, 359/259]$.

The same piece is also a piece of the run on $[2.8, 3.6]$, which has the anchor at $p = 3$ as well ($C_3 = 248143/40000$); there the shift-side bound is 6.369965 and the margin 1.4354%. More anchors give a smaller shift-side bound.

9.4. **Table of all anchors.** Table 4 lists every anchor used in Theorem 9.1. A certificate file `X` of item (a), (b), (c) or (d) is `a_window_2_21/certificates/X`, `b_PA_21_28/certificates/X`, `c_drury_28_20/certificates/X` or `d_PD_20_40/certificates/X`. "Bellman (k)" is a Bellman certificate with k terms (Definition 9.4); "TISOS (k)" is a TISOS certificate with k monomials (Definition 8.3).

Table 4: All anchors: proved upper bounds C for $\|P(S)\|_p$.

item, P	p	$C \geq \|P(S)\|_p$	kind	certificate file
(a) $P_{7/10}$	2	$\sqrt{19683/1700} \approx 3.4026806$	exact	—
	2.1	$17074961/5000000 = 3.4149922$	Bellman (12)	`bell_p21_eta07_M12.txt`
(b) P_A	2	M_{P_A} (Table 3) ≈ 5.2291914	exact	—
	2.1	$52358261/10000000 = 5.2358261$	Bellman (12)	`bell_PA_p21_M12.txt`
	2.2	$26238519/5000000 = 5.2477038$	Bellman (12)	`bell_PA_p22_M12.txt`
	2.35	$52745691/10000000 = 5.2745691$	Bellman (12)	`bell_PA_p235_M12.txt`
	2.5	$26540459/5000000 = 5.3080918$	Bellman (12)	`bell_PA_p25_M12.txt`
	2.65	$26800967/5000000 = 5.3601934$	Bellman (12)	`bell_PA_p265_M12.txt`
	2.8	$13555947/2500000 = 5.4223788$	Bellman (12)	`bell_PA_p28_m3_M12.txt`
	∞	$3332879/500000 = 6.665758$	$\sum \lvert c_k \rvert$	—
(c) P_{Dr}	2	$\sqrt{19683/550} \approx 5.9822465$	exact	—
	3	$248143/40000 = 6.203575$	Bellman (8)	`bell_p3_M8_m4.txt`
	4	$3227413/500000 = 6.454826$	TISOS (6)	`cert_p4_W3_C6.454826.txt`
	6	$6737251/1000000 = 6.737251$	TISOS (10)	`cert_p6_W3_C6.737251.txt`
	8	$13783/2000 = 6.891500$	TISOS (15)	`cert_p8_W3_C6.891500.txt`
	10	$6988033/1000000 = 6.988033$	TISOS (21)	`cert_p10_W3_C6.988033.txt`
	12	$1410781/200000 = 7.053905$	TISOS (28)	`cert_p12_W3_C7.053905.txt`
	14	$3550829/500000 = 7.101658$	TISOS (36)	`cert_p14_W3_C7.101658.txt`
	16	$3568923/500000 = 7.137846$	TISOS (45)	`cert_p16_W3_C7.137846.txt`
	18	$7166211/1000000 = 7.166211$	TISOS (55)	`cert_p18_W3_C7.166211.txt`
	20	$7189041/1000000 = 7.189041$	TISOS (66)	`cert_p20_W3_C7.189041.txt`
	∞	$37/5 = 7.4$	$\sum \lvert c_k \rvert$	—
(d) P_D	2	$\sqrt{2673/32} \approx 9.1395432$	exact	—
	20	$9844747/1000000 = 9.844747$	TISOS (66)	`cert_PD_p20_W3_C9.844747.txt`
	22	$1971627/200000 = 9.858135$	TISOS (78)	`cert_PD_p22_W3_C9.858135.txt`
	24	$986943/100000 = 9.869430$	TISOS (91)	`cert_PD_p24_W3_C9.869430.txt`
	26	$9879081/1000000 = 9.879081$	TISOS (105)	`cert_PD_p26_W3_C9.879081.txt`
	28	$9887417/1000000 = 9.887417$	TISOS (120)	`cert_PD_p28_W3_C9.887417.txt`
	30	$4947343/500000 = 9.894686$	TISOS (136)	`cert_PD_p30_W3_C9.894686.txt`
	32	$4950539/500000 = 9.901078$	TISOS (153)	`cert_PD_p32_W3_C9.901078.txt`
	34	$49534407/5000000 = 9.9068814$	Bellman (8)	`bell_PD_M8_m2e5v2_p34.txt`
	36	$99119789/10000000 = 9.9119789$	Bellman (8)	`bell_PD_M8_m2e5v2_p36.txt`
	38	$49582631/5000000 = 9.9165262$	Bellman (8)	`bell_PD_M8_m2e5v2_p38.txt`
	40	$99205169/10000000 = 9.9205169$	Bellman (8)	`bell_PD_M8_m5_p40.txt`
	∞	10	$\sum \lvert c_k \rvert$	—

The TISOS certificates of item (d) have a line `tau`: they prove $\|P_D(S)\|_p \le \gamma^{1/p}$, and $\gamma \le C^p$ is checked exactly (Section 9.2). The certificates were found by programs that are not included; the checks do not depend on how a certificate was found.

9.5. **The polynomial $P_{7/10}$: $p \in (2, 2.1]$ and $[21/11, 2)$.** Here $P = P_{7/10} = 1 + 2z - \frac{17}{10}z^2$; it is the member $\eta = \frac{7}{10}$ of the family $1 + 2z - (1+\eta)z^2$ of Section 6. Its maximum on the circle is attained at $e^{\pm i\varphi}$ with $\cos\varphi = -\frac{7}{34}$ (Table 3); we take $\varphi \in (\pi/2, \pi)$, so $\sin\varphi = \sqrt{1107}/34$. Theorem 6.3 shows, with a proof by hand, that the member $\eta = \frac{1}{5}$ of this family violates (1) on a punctured neighbourhood of $p = 2$ that is not explicit; item (a) gives an explicit one, $[21/11, 2) \cup (2, 2.1]$, for the member $\eta = \frac{7}{10}$. On this range the margin tends to 0 as $p \to 2$, since at $p = 2$ both sides equal M_P. So here we do not use pieces with a fixed decimal matrix: the matrix is a scaled rotation that depends continuously on p, and the interval next to $p = 2$ is handled with a bound on a derivative.

For $\vartheta \in \mathbb{R}$ put $N_p(\vartheta) = (|\cos\vartheta|^p + |\sin\vartheta|^p)^{1/p} = \|(\cos\vartheta, \sin\vartheta)\|_p$; it is even and has period $\pi/2$. Since $R_\theta(\cos\vartheta, \sin\vartheta) = (\cos(\vartheta - \theta), \sin(\vartheta - \theta))$,

$$\|R_\varphi\|_p = \max_\vartheta \frac{N_p(\vartheta - \varphi)}{N_p(\vartheta)} = \max_\vartheta \frac{N_p(\vartheta + \varphi)}{N_p(\vartheta)} \tag{18}$$

(substitute $\vartheta \mapsto -\vartheta$ and use that N_p is even).

Definition 9.12 (the rotation bound $U(p)$)**.** Let $\bar u = 10064142767/10^{10}$. For $p \in [2, 2.1]$ put

$$\lambda(p) = \frac{\frac12 - \frac1p}{\frac12 - \frac{10}{21}} = 21 - \frac{42}{p} \in [0, 1], \qquad U(p) = \bar u^{\,\lambda(p)}, \qquad T_p = R_\varphi / U(p).$$

The programs `window_cert.py` and `window_check.py` both prove $\|R_\varphi\|_{2.1} \le \bar u$ in interval arithmetic from (18): the interval $\vartheta \in [0, \pi/2]$ is cut into cells until an enclosure of the quotient is below $\bar u$ on every cell. (In fact $\|R_\varphi\|_{2.1} = 1.0064142762007\ldots$, about $5 \cdot 10^{-10}$ below $\bar u$.)

Lemma 9.13. *(i) $\|T_p\|_p \le 1$ for $p \in [2, 2.1]$. (ii) For $\tau > 0$, put $\alpha(\tau) = \arg P(\tau e^{i\varphi})$. For every $\vartheta_* \in \mathbb{R}$ and $p \ge 1$,*

$$\|P(\tau R_\varphi)\|_p \;\ge\; |P(\tau e^{i\varphi})|\, \frac{N_p(\vartheta_* + \alpha(\tau))}{N_p(\vartheta_*)}.$$

(iii) For $p \in [2, 2.1]$, $\|P(S)\|_p \le M_P^{1-\lambda(p)} C_{2.1}^{\lambda(p)}$ with $C_{2.1} = 17074961/5000000$.

Proof. (i) $\|R_\varphi\|_2 = 1$ (R_φ is orthogonal) and $\|R_\varphi\|_{2.1} \le \bar u$ (proved after Definition 9.12). By the convexity in Lemma 9.8(iii), $\log\|R_\varphi\|_{1/s}$ lies below its chord between $s = \frac{10}{21}$ and $s = \frac12$, so $\|R_\varphi\|_p \le 1^{1-\lambda(p)} \bar u^{\,\lambda(p)} = U(p)$. (ii) The span of I and J is an algebra isomorphic to $\mathbb{C}$ ($J^2 = -I$), with $R_\theta \leftrightarrow e^{i\theta}$; so $P(\tau R_\varphi) = |P(\tau e^{i\varphi})|\, R_{\alpha(\tau)}$, as in (10). For the test vector $x = (\cos\vartheta_*, -\sin\vartheta_*)$, $R_\alpha x = (\cos(\vartheta_* + \alpha), -\sin(\vartheta_* + \alpha))$, and $\|x\|_p = N_p(\vartheta_*)$, $\|R_\alpha x\|_p = N_p(\vartheta_* + \alpha)$. (iii) Lemma 9.8(i) with the anchors $p = 2$ (M_P, Lemma 9.3) and $p = 2.1$ (Table 4); between these two anchors the interpolation parameter is the same $\lambda(p)$. □

The function Θ. With $\vartheta_* = 157/200$, $s = 1/p$, $\lambda = 21 - 42s$, $\tau = \bar u^{\,-\lambda} = 1/U(p)$ and $\alpha = \alpha(\tau)$, put

$$\Theta(s) = \log|P(\tau e^{i\varphi})| + \log N_p(\vartheta_* + \alpha) - \log N_p(\vartheta_*) - \big[(1-\lambda)\log M_P + \lambda \log C_{2.1}\big].$$

By Lemma 9.13, $\Theta(s) > 0$ implies that T_p is a contraction with $\|P(T_p)\|_p > \|P(S)\|_p$; the relative margin at p is at least $e^{\Theta(s)} - 1 > \Theta(s)$. At $s = \frac12$: $\lambda = 0$, $\tau = 1$, $|P(e^{i\varphi})| = M_P$ and $N_2 \equiv 1$, so $\Theta(\frac12) = 0$ exactly.

The checks. (W1) The interval $s \in [10/21, 199/400]$, that is $p \in [400/199, 2.1]$, is cut into 400 equal pieces, and on each piece an interval enclosure of Θ (in Arb ball arithmetic, with the whole piece as a single ball) has a positive lower end. The smallest lower end is 0.000330007

(`window_cert.py`) and 0.000329987 (`window_check.py`). (W2) On the last piece $s \in [199/400, 1/2]$, that is $p \in [2, 400/199]$, interval values of Θ cannot be positive, because $\Theta(\frac{1}{2}) = 0$. Instead the programs enclose the derivative $\Theta'(s)$ on the whole piece, by forward-mode differentiation (every quantity is carried together with its derivative in s, both as balls). The enclosure is negative: $\Theta' \in -0.145090 \pm 0.0151$ (`window_cert.py`, a single ball for the whole piece), and the upper ends on 8 sub-pieces are between -0.1434 and -0.1424 (`window_check.py`). On this piece the bases $\cos\vartheta_*$, $\sin\vartheta_*$, $\cos(\vartheta_* + \alpha)$, $\sin(\vartheta_* + \alpha)$ of the p-th powers stay away from 0 (the programs assert it), so Θ is smooth there and

$$\Theta(s) = \Theta(\tfrac{1}{2}) - \int_s^{1/2} \Theta'(\sigma)\, d\sigma = -\int_s^{1/2} \Theta'(\sigma)\, d\sigma > 0 \qquad (199/400 \le s < 1/2).$$

Together, $\Theta > 0$ on $[10/21, 1/2)$, that is, for every $p \in (2, 2.1]$.

For $p \in [21/11, 2)$ the matrix is $T^{\mathrm{t}}_{p'} = R_{-\varphi}/U(p')$ (Lemma 9.11).

Files (in `a_window_2_21/`; commands are run from `programs/`).

- Anchor `certificates/bell_p21_eta07_M12.txt`: checked by `bellman_check.py` (`logs/bellman_check_p21.log`) and by `bellman_taylor.py` (`logs/bellman_taylor_p21_v020.log`); the controls of the latter are in `logs/bellman_taylor_p21_controls_v020.log`.
- The range $(2, 2.1]$:

  ```
  window_cert.py 7/10 21/10 17074961/5000000 785/1000 1/400 400 10064142767/10000000000
  ```

  (arguments: η, $p_1 = 2.1$, $C_{2.1}$, $\vartheta_* = 785/1000 = 157/200$, the width $1/400$ of the last piece in s, the number of pieces, $\bar{u}$), log `logs/window_cert_p21_ubar_v040.log`.
- Independent re-implementation from the written argument: `window_check.py --run`, log `logs/window_check_v030.log`; its control run, with $C_{2.1} = 3.43$ in place of 17074961/5000000, fails as it must (`logs_rerun/window_check_test_run.log`).

9.6. **The polynomial P_A: $p \in [2.1, 2.8]$ and $[14/9, 21/11]$.**

- *Polynomial:* $P_A = 1 + \frac{13043}{6250} z - \frac{1789439}{500000} z^2$.
- *Anchors:* $p = 2$ (exact), 2.1, 2.2, 2.35, 2.5, 2.65, 2.8 (Bellman certificates) and ∞ (Table 4); anchor table `b_PA_21_28/tables/anchors_PA.txt`.
- *Matrices:* the table `b_PA_21_28/tables/T_table_PA_5col_0030.txt` has 22 matrices with six-decimal entries, for $p = 2.05, 2.10, \ldots, 3.10$; each piece uses the row nearest to its midpoint.
- *Run* (one command, from `b_PA_21_28/programs/`):

  ```
  interval_certify.py 2.1 2.8 0.05 0.0001 1,13043/6250,-1789439/500000
      ../tables/anchors_PA.txt ../tables/T_table_PA_5col_0030.txt
  ```

  log `b_PA_21_28/logs/cert_PA_21_28_v080.log`: 15 pieces tiling $[2.1, 2.8]$, namely $[2.1, 2.125]$, $[2.125, 2.15]$ and 13 pieces of width 0.05.
- *Relative margins:* from 0.2900% (on $[2.15, 2.2]$) to 1.6958% (on $[2.55, 2.6]$).
- *Re-checks of the pieces* (logs in `b_PA_21_28/logs/`): `read_R2_pieces.py` (`read_R2_pieces_PA_21_28_v010.log`, with two controls: all anchors raised by 2%, which certifies no piece, and a norm bound below the true norm at $p = 2.1$, which cannot be proved) and `verify_R2.py` (`verify_R2_PA_21_28_v040.log`; its control, with the anchors scaled by 51/50, is `verify_R2_PA_21_28_control_v040.log`). With the printed ϱ: `verify_R2_rho_PA_21_28_v050.log` and `R2_read_rho_PA_21_28.log`, each with a control log whose name adds `_control`.
- *Checks of the anchors:* `bellman_check.py` (`bellman_check_PA_p21.log`, …, `bellman_check_PA_p28.log`) and `bellman_taylor.py` (`bellman_taylor_PA_4_v020.log` for $p = 2.1$ to 2.5, `bellman_taylor_PA_p265_v020.log`, `bellman_taylor_PA_p28_p2725_v020.log`), with controls with lowered C.

9.7. **Drury's polynomial: $p \in [2.8, 20]$ and $[20/19, 14/9]$.**

- *Polynomial:* Drury's $P_{\mathrm{Dr}} = 1 + 2z - \frac{22}{5}z^2$ [6].
- *Anchors:* $p = 2$ (exact, $\max_{|z|=1} |P_{\mathrm{Dr}}(z)| = \sqrt{19683/550}$), $p = 3$ (Bellman, $\leq 248143/40000$), $p = 4, 6, \ldots, 20$ (TISOS, ≤ 6.454826, 6.737251, 6.891500, 6.988033, 7.053905, 7.101658, 7.137846, 7.166211, 7.189041) and $p = \infty$ ($\sum |c_k| = 7.4$); Table 4. The anchors at $p = 3, 4, \ldots, 20$ are those of Theorem 8.1.
- *Runs:* four runs of `interval_certify.py` with the coefficients `1,2,-22/5`, in Table 5; together $80 + 45 + 160 + 160 = 445$ pieces. The first two overlap on $[3.55, 3.6]$; together the runs cover $[2.8, 20]$.

run	step / minimal step	anchors used	pieces	relative margins
$[2.8, 3.6]$	0.01 / 0.0003125	$2, 3, 4$	80	1.4354%–1.9101%
$[3.55, 4]$	0.01 / 0.0003125	$2, 4$	45	0.9148%–1.2489%
$[4, 12]$	0.05 / 0.0003125	$4, 6, 8, 10, \infty$	160	0.0790%–1.0464%
$[12, 20]$	0.05 / 0.0001	$12, 14, 16, 18, 20$	160	0.0139%–0.0942%

TABLE 5. The four runs of item (c); "anchors used" are the anchors next to the range of the run.

The files of the four runs (anchor table; matrix table; log), all in `c_drury_28_20/`:

- $[2.8, 3.6]$: `tables/anchors_tisos_p3_preview.txt`; `tables/T_table_280_440.txt`; `logs/cert_down_p3preview_v070.log`;
- $[3.55, 4]$: `tables/anchors_tisos_4_10.txt`; `tables/T_table_1836.txt`; `logs/cert_left_tisos_v070.log`;
- $[4, 12]$: `tables/anchors_tisos_4_10.txt`; `tables/T_table_280_1200.txt`; `logs/cert_up_A7_4_12_v080.log`;
- $[12, 20]$: `tables/anchors_tisos.txt`; `tables/T_table_10_20_5col.txt`; `logs/cert_up_A7_12_20_v080.log`.

All piece steps were the nominal ones (no piece needed halving). The matrix tables give a matrix with six-decimal entries for each value of p on a grid of step 0.05 (step 0.1 in `T_table_10_20_5col.txt`); a matrix may serve several adjacent pieces. *Relative margins:* from 0.0139% (on $[19.9, 19.95]$ and $[19.95, 20]$, near $p = 20$) to 1.9101% (on $[2.99, 3]$, near $p = 3$).

The anchors used on $[2.8, 3]$ *and on* $[10, 12]$. On $[2.8, 3]$ the shift-side bound interpolates between the anchors at $p = 2$ and $p = 3$. On $[10, 12]$ the run on $[4, 12]$ interpolates between the anchors at $p = 10$ and $p = \infty$, since its anchor table `anchors_tisos_4_10.txt` has no anchor at $p = 12$; at $p = 12$ this gives $6.988033^{5/6}\, 7.4^{1/6} \approx 7.055066$, which is enough on every piece of $[10, 12]$. The re-check `verify_R2.py` also uses the anchor at $p = 12$ ($\|P_{\mathrm{Dr}}(S)\|_{12} \leq 7.053905$); `R2_read.py` uses the same anchors as the run.

- *Re-checks of the pieces* (logs in `c_drury_28_20/logs/`): `verify_R2.py` (`verify_R2_280_360_v030.log`, `verify_R2_355_12_v030.log` for the runs on $[3.55, 4]$ and $[4, 12]$ together, `verify_R2_12_20_v030.log`) and `R2_read.py` (`R2_read_280_360.log`, `R2_read_355_4.log`, `R2_read_4_12.log`, `R2_read_12_20.log`). Each reports all pieces certified, with no failure and an exact tiling. With the printed ϱ: `verify_R2_rho_280_360_v050.log`, `verify_R2_rho_355_12_v050.log`, `verify_R2_rho_12_20_v050.log`, and `R2_read_rho_280_360.log`, `R2_read_rho_355_4.log`, `R2_read_rho_4_12.log`, `R2_read_rho_12_20.log`, each `R2_read` log with a control log whose name adds `_control`.
- *Checks of the anchors:* each TISOS certificate by `tisos_verify.py` (`tisos_bellman_verify_v010.log`) and `verify_tisos.py` (`verify_tisos_1939.log`,

`verify_tisos_high1_2017.log`, `verify_tisos_high2_2017.log`), and again by `tisos_read.py` (`logs_rerun/`); the Bellman certificate at $p = 3$ by `bellman_verify.py` (ball arithmetic; `tisos_bellman_verify_v010.log`), `verify_bellman_p3.py` (exact rationals; `verify_bellman_p3_0031.log`) and `bellman_bernstein.py` (exact Bernstein bounds; `bellman_bernstein_p3_v020.log`).

Exact rational matrices at $p = 14, 16, 18, 20$. At these four values of p there are also matrices with rational entries, for which every step is exact rational arithmetic. For each such $p = 2m$ the certificate `c_drury_28_20/certificates/TSIDE_cert_p<p>.txt` gives a rational T, a rational test vector v, and an exact Gram matrix H, positive definite, such that for $w = (x, y)$

$$\Phi(x, y) := x^p + y^p - (Tw)_1^p - (Tw)_2^p = Z^{\mathsf{t}} H Z, \qquad Z = (x^m, x^{m-1}y, \dots, y^m);$$

as p is even, $\Phi(x, y) = \|w\|_p^p - \|Tw\|_p^p$, so $\|T\|_p \le 1$. A second proof of $\|T\|_p \le 1$: Φ is homogeneous of even degree, $\Phi(1, 0) > 0$, and $\Phi(t, 1)$ has no real root (exact count of real roots), so $\Phi > 0$ off the origin. Finally, $\|P_{\mathrm{Dr}}(T)v\|_p^p / \|v\|_p^p > C^p$ exactly, with C the TISOS anchor at p:

p	T	v	$\|P_{\mathrm{Dr}}(T)\|_p \ge$	C	margin
14	$\begin{pmatrix} -13/1000 & 991/1000 \\ -443/500 & -129/1000 \end{pmatrix}$	$(1,\ 953413/1000000)$	7.103397338	7.101658	0.0245%
16	$\begin{pmatrix} -113/1000 & 899/1000 \\ -993/1000 & -1/100 \end{pmatrix}$	$(-239149/250000,\ 1)$	7.140643266	7.137846	0.0392%
18	$\begin{pmatrix} -39/5000 & 1243/1250 \\ -4567/5000 & -12/125 \end{pmatrix}$	$(1,\ 191789/200000)$	7.168119861	7.166211	0.0266%
20	$\begin{pmatrix} -823/10000 & 37/40 \\ -9957/10000 & -59/10000 \end{pmatrix}$	$(-24031/25000,\ 1)$	7.190460162	7.189041	0.0197%

The certificates are written by `evenp_final.py` (`logs/evenp_final_v030_12to18.log`, `logs/evenp_final_v030_20.log`) and checked independently by `read_evenp_Tside.py` (exact, sympy; it takes T, v, H from the certificate and C from the TISOS file; `logs/read_evenp_Tside_v060.log`, with two controls: $1.01\,T$ and $1.01\,C$) and by `verify_evenp.py` (`logs/verify_evenp_direct_2016.log`). With the TISOS anchors, which are also exact, these four values of p (and their duals) are proved by exact rational arithmetic alone.

9.8. **The polynomial P_D: $p \in [20, 40]$ and $[40/39, 20/19]$.**

- *Polynomial:* $P_D = 1 + z - 8z^2$.
- *Anchors:* $p = 2$ (exact, $\sqrt{2673/32}$), $p = 20, 22, \dots, 32$ (TISOS, with a line `tau`), $p = 34, 36, 38, 40$ (Bellman) and $p = \infty$ (10); Table 4; anchor table `d_PD_20_40/tables/anchors_PD_20_40.txt`.
- *Matrices:* `d_PD_20_40/tables/T_table_PD_20_40_5col.txt`, 101 matrices with six-decimal entries for $p = 20.0, 20.2, \dots, 40.0$.
- *Run* (one command, from `d_PD_20_40/programs/`):

```
interval_certify.py 20.0 40.0 0.2 0.0005 1,1,-8
    ../tables/anchors_PD_20_40.txt ../tables/T_table_PD_20_40_5col.txt
```

  log `d_PD_20_40/logs/cert_PD_20_40_v080.log`: 100 pieces of width 0.2 tiling $[20, 40]$.
- *Relative margins:* from 0.0228% (on $[39.8, 40]$) to 0.0994% (on $[20, 20.2]$).
- *Re-checks of the pieces* (logs in `d_PD_20_40/logs/`): `R2_read.py` (`R2_read_PD_20_40.log`; the anchors are read from the eleven certificate files) and `read_pieces.py` (`read_pieces_PD_20_40_0546.log`; its own smallest margin is 0.0229%; its control, with every anchor multiplied by 1.001, has worst margin −0.0770%). With the printed ϱ: `verify_R2_rho_PD_20_40_v050.log` and `R2_read_rho_PD_20_40.log`, each with a control log whose name adds `_control`.

- *Checks of the anchors:* TISOS $p = 20, \dots, 32$ by `tisos_read.py` (`tisos_read_PD_0146.log`) and `verify_tisos.py` (`verify_tisos_PD_0154.log`); Bellman $p = 34, 36, 38$ by `verify_bellman_evenp.py` (exact rationals; `verify_bellman_evenp_PD_0147.log`) and `bellman_check.py` (`bellman_check_PD_p34.log`, `bellman_check_PD_p36.log`, `bellman_check_PD_p38.log`); Bellman $p = 40$ by `bellman_check.py` (`bellman_check_PD_p40.log`) and `bellman_taylor.py` (`bellman_taylor_PD_p40_v020.log`). Each has a control with lowered C.

9.9. **Independent checks.** Table 6 lists, for each kind of claim, the programs that check it. Programs in the same row were written independently of each other; they use different methods or different arithmetic.

Table 6: Independent checks. Logs are in the folder `logs/` of the item's part unless said otherwise.

claim	first check	second (and third) check
(a) anchor $p = 2.1$	`bellman_check.py`: first-order box bounds, balls	`bellman_taylor.py`: second-order Taylor and Bernstein bounds, balls
(a) the range $(2, 2.1]$	`window_cert.py`: 400 pieces; an enclosure of Θ' on the last piece	`window_check.py`: re-implementation from the written argument; enclosures of Θ' on 8 sub-pieces
(b) six Bellman anchors	`bellman_check.py`	`bellman_taylor.py`
(b) 15 pieces	`interval_certify.py`: mpmath intervals	`read_R2_pieces.py` and `verify_R2.py`: own norm bound, own test vector, balls and exact rationals
(c) TISOS anchors $p = 4, \dots, 20$	`tisos_verify.py`: exact, folded identity, Sylvester	`verify_tisos.py`: exact, folded and summed identity, LDL^{t}; `tisos_read.py`: exact
(c) Bellman anchor $p = 3$	`bellman_verify.py`: balls	`verify_bellman_p3.py`: exact rationals; `bellman_bernstein.py`: exact Bernstein bounds
(c) 445 pieces	`interval_certify.py` (four runs)	`verify_R2.py` and `R2_read.py`
(c) exact matrices $p = 14, \dots, 20$	`evenp_final.py`: exact Gram matrix and Sturm count	`read_evenp_Tside.py`: exact (sympy), own Gram check and root count; `verify_evenp.py`
(d) TISOS anchors $p = 20, \dots, 32$	`tisos_read.py`	`verify_tisos.py`
(d) Bellman anchors $p = 34, 36, 38$	`verify_bellman_evenp.py`: exact rationals	`bellman_check.py`: balls
(d) Bellman anchor $p = 40$	`bellman_check.py`	`bellman_taylor.py`
(d) 100 pieces	`interval_certify.py`	`R2_read.py` and `read_pieces.py`
(b)–(d) the printed ϱ_i (Theorem 9.1)	`verify_R2.py --printed-R`: balls	`R2_read.py --verify-printed-rho`: balls

Controls. Every control run named in this paper fails, as it must. They use a changed constant C, a scaled weight of a Bellman certificate, anchors raised by a small factor, a lowered norm bound, or a perturbed matrix. A log with "control" in its name is such a run.

Re-runs. The folder `logs/` of each part holds the logs of the runs this section relies on. The folder `logs_rerun/` holds the same checks run again with exactly the shipped files, one log per command; its first line names the program, its md5 and the full command, and its last line the exit code. The certifier re-runs reproduce every piece line of the logs. Every log in `logs/` prints the md5 of the program, anchor table and matrix table as they were run; `MANIFEST.tsv` relates these

to the shipped files, which differ from the originals only in comments, printed texts and, for a few programs, the paths they read (`README.txt` lists every change). The programs need Python 3.12, python-flint 0.9.0, mpmath 1.2.1, numpy 2.5.4 and sympy 1.12, and run on one thread; each is run from the `programs/` folder of its part.

9.10. **Proof of Theorem 9.1.** Let $q = \max(p, p') \in (2, 40]$. By Lemma 9.11 it suffices to prove the claim at q with $T_* = T$.

If $q \in (2, 2.1]$: the checks (W1)–(W2) of Section 9.5 give $\Theta(1/q) > 0$, and Lemma 9.13 gives $\|T_q\|_q \le 1$ and $\|P_{7/10}(T_q)\|_q > \|P_{7/10}(S)\|_q$.

If $q \in [2.1, 40]$: the pieces of the runs of items (b)–(d) tile $[2.1, 2.8]$, $[2.8, 3.6] \cup [3.55, 4] \cup [4, 12] \cup [12, 20]$ and $[20, 40]$ with no gap (each certifier log ends with this statement, and each re-checker checks the tiling again). So q lies in some piece, and on that piece the anchors of Table 4 (valid by Lemmas 9.3 and 9.5, Proposition 8.4 and the checks of Section 9.9) and the two facts $\varrho_i \ge \max(\|T_i\|_{p_{\mathrm{lo}}}, \|T_i\|_{p_{\mathrm{hi}}})$ and $L > \widehat{C}(p_{\mathrm{hi}})$ for T_i/ϱ_i, proved for the printed ϱ_i by `verify_R2.py` and `R2_read.py` (Section 9.2), give the claim by Lemma 9.10. □

10. Necessary conditions for matrix counterexamples

Theorem A and Drury's theorem [6] give real 2×2 counterexamples to (1). This section collects simple necessary conditions for matrix counterexamples, mostly of size 2×2, and it checks where the known counterexamples lie. All the conditions come from a single known class of contractions that satisfy (1): those whose entrywise modulus is a contraction (Lemma 10.1). These conditions are elementary; they show where a search need not look.

Throughout this section $1 < p < \infty$ and $p' = p/(p-1)$; $\|\cdot\|_p$ is the norm of ℓ^p_n or of an $n \times n$ matrix acting on it. By the proof of Lemma 2.1, which applies to every real matrix, a real matrix has the same norm on $\mathbb{R}^n$ and on $\mathbb{C}^n$, so everything below applies to real matrices on the real ℓ^p_n as well. For an $n \times n$ matrix $T = (t_{jk})$ let $|T| = (|t_{jk}|)$ be its entrywise modulus, and for $x \in \mathbb{C}^n$ let $|x| = (|x_1|, \dots, |x_n|)$. We say that T *satisfies* (1) if $\|P(T)\|_p \le \|P(S)\|_p$ for every polynomial P. A *counterexample at p* is a matrix T with $\|T\|_p \le 1$ that does not satisfy (1).

10.1. The entrywise modulus.

Lemma 10.1 (a contractive modulus)**.** *Let T be an $n \times n$ matrix with $\||T|\|_p \le 1$. Then T is a contraction on ℓ^p_n, and it satisfies* (1)*.*

Proof. This is known; we give only the reduction to the literature. For $x \in \mathbb{C}^n$ and each j, $|(Tx)_j| \le \sum_k |t_{jk}|\,|x_k| = (|T|\,|x|)_j$. So $\|Tx\|_p \le \||T|\,|x|\|_p \le \|x\|_p$, and $|T|$ is a positive contractive majorant of T (that is, a positive contraction with $|Tx| \le |T|\,|x|$ coordinatewise for every x). For contractions with a positive contractive majorant, (1) is known (Peller [16]; see [2, Sec. 1]); this is the class recalled in the Background paragraph of Section 1. If $T = |T|$, it is the case of positive contractions (Akcoglu–Sucheston dilation [1] with Coifman–Rochberg–Weiss transference [4]). □

Corollary 10.2 (rank one)**.** *Every $n \times n$ contraction of rank at most one satisfies* (1)*. In particular, every nilpotent 2×2 contraction does.*

Proof. If $T = 0$, then $|T| = 0$. Otherwise $T = uv^{\mathsf{t}}$ with nonzero $u, v \in \mathbb{C}^n$, so $Tx = (v^{\mathsf{t}}x)\,u$. Hölder's inequality and its equality case give $\|T\|_p = \|u\|_p\|v\|_{p'}$. The modulus $|T| = |u|\,|v|^{\mathsf{t}}$ has the same form, so $\||T|\|_p = \||u|\|_p\||v|\|_{p'} = \|u\|_p\|v\|_{p'} = \|T\|_p \le 1$. Now apply Lemma 10.1. A nilpotent 2×2 matrix is not invertible, so its rank is at most one. □

Corollary 10.2 is the rank-one case of Lemma 10.1. We record it because nilpotent matrices are a natural place to start looking for a counterexample (if $T^2 = 0$, then $P(T) = c_0 I + c_1 T$), and the corollary rules out that family.

10.2. **The phase test for 2×2 matrices.** Lemma 10.1 applies to every contraction T that differs from $|T|$ only by a change of the phases of the coordinates, on both sides. For 2×2 matrices this is a condition on a single complex number. Let $\mathcal{U}_n$ be the set of diagonal $n \times n$ matrices whose diagonal entries have modulus one. They are isometries of ℓ^p_n, since they do not change the moduli of the coordinates.

Lemma 10.3 (the phase test)**.** *(a) If $T = U_1 A U_2$ with $U_1, U_2 \in \mathcal{U}_n$ and A entrywise nonnegative, then $A = |T|$ and $\||T|\|_p = \|T\|_p$. Hence, if such a T is a contraction, it satisfies* (1)*.*

(b) A 2×2 matrix T is of the form in (a) if and only if some entry of T is 0 or $t_{11}t_{22}/(t_{12}t_{21})$ is a positive real number; equivalently, if and only if $t_{11}t_{22}\,\overline{t_{12}t_{21}} \in [0,\infty)$.

Proof. (a) Write $U_1 = \operatorname{diag}(e^{i\alpha_1}, \dots, e^{i\alpha_n})$ and $U_2 = \operatorname{diag}(e^{i\beta_1}, \dots, e^{i\beta_n})$. Then $t_{jk} = e^{i(\alpha_j+\beta_k)}a_{jk}$, so $|t_{jk}| = a_{jk}$ and $A = |T|$. As $A = U_1^{-1}TU_2^{-1}$ and $U_1^{\pm 1}$, $U_2^{\pm 1}$ are isometries, $\|A\|_p = \|T\|_p$. If $\|T\|_p \le 1$, Lemma 10.1 applies.

(b) Let $n = 2$. For real numbers $\alpha_1, \alpha_2, \beta_1, \beta_2$ form U_1, U_2 as in (a). The entries of $U_1^{-1}TU_2^{-1}$ are $e^{-i(\alpha_j+\beta_k)}t_{jk}$, and they are all nonnegative if and only if

$$\alpha_j + \beta_k \equiv \arg t_{jk} \pmod{2\pi} \qquad \text{for every } (j,k) \text{ with } t_{jk} \neq 0. \tag{19}$$

So T is of the form in (a) if and only if the system (19) has a real solution.

All four entries nonzero. If (19) holds, add its equations for $(1,1)$ and $(2,2)$ and subtract those for $(1,2)$ and $(2,1)$. This gives $\arg t_{11} + \arg t_{22} - \arg t_{12} - \arg t_{21} \equiv 0$, that is, $t_{11}t_{22}/(t_{12}t_{21}) > 0$. Conversely, if this ratio is positive, put $\alpha_1 = 0$, $\beta_1 = \arg t_{11}$, $\beta_2 = \arg t_{12}$ and $\alpha_2 = \arg t_{21} - \arg t_{11}$. Then (19) holds for $(1,1)$, $(1,2)$ and $(2,1)$ by construction, and for $(2,2)$ it reads $\arg t_{21} - \arg t_{11} + \arg t_{12} \equiv \arg t_{22}$, which is the assumption.

Some entry zero. Consider the graph whose four vertices are the two rows and the two columns, with an edge from row j to column k whenever $t_{jk} \neq 0$. The unknown α_j sits at row j, and β_k at column k. The four possible edges form a cycle of length four; with at most three of them, the graph has no cycle. In each connected component, set the unknown at one vertex equal to 0 and walk outward: each edge leads to a vertex not yet reached, and (19) for that edge fixes the unknown there. Unknowns at isolated vertices are arbitrary. (For example, if $t_{11} = 0$: $\alpha_1 = 0$, $\beta_2 = \arg t_{12}$, $\alpha_2 = \arg t_{22} - \beta_2$, $\beta_1 = \arg t_{21} - \alpha_2$.)

Finally, $t_{11}t_{22}\,\overline{t_{12}t_{21}} = 0$ if some entry is 0, and $t_{11}t_{22}\,\overline{t_{12}t_{21}} = |t_{12}t_{21}|^2\, t_{11}t_{22}/(t_{12}t_{21})$ otherwise; this gives the equivalent form. □

Corollary 10.4 (necessary conditions for a 2×2 counterexample)**.** *Let T be a 2×2 counterexample at some $p \in (1,\infty)$. Then:*

(i) $\||T|\|_p > 1$, and T is invertible;
(ii) all four entries of T are nonzero, and $t_{11}t_{22}/(t_{12}t_{21}) \notin (0,\infty)$;
(iii) if T is real, then $t_{11}t_{12}t_{21}t_{22} < 0$, so T has one or three negative entries.

Proof. (i) follows from Lemma 10.1 and Corollary 10.2, since a 2×2 matrix of rank two is invertible. (ii) is Lemma 10.3. (iii) For real nonzero entries, the ratio in (ii) is positive if and only if $t_{11}t_{22}$ and $t_{12}t_{21}$ have the same sign, that is, if and only if $t_{11}t_{12}t_{21}t_{22} > 0$. This product is negative exactly when an odd number of the four entries is negative. □

Remark 10.5. (1) A 2×2 matrix of rank one with four nonzero entries has $t_{11}t_{22} = t_{12}t_{21}$, so the ratio in Lemma 10.3(b) is 1. Thus, for $n = 2$, Lemma 10.3 contains Corollary 10.2. (2) Among real matrices with four nonzero entries, Corollary 10.4(iii) excludes half of the 16 sign patterns. For complex matrices the excluded set is small: the ratio is a positive real number only on a set of measure zero. (3) For $n\times n$ matrices, T has the form in Lemma 10.3(a) if and only if the system (19) is solvable on the bipartite graph of the nonzero entries of T (rows and columns as vertices). By the standard criterion for such systems on a graph (solve along a spanning tree of each component,

as above; the remaining equations are conditions on cycles), this holds if and only if, around every cycle of that graph, the alternating sum of the arguments of the entries is 0 modulo 2π.

10.3. **Real sign patterns, and the known counterexamples.**

Remark 10.6 (two sign classes). Three operations take a real 2×2 counterexample at p to another real counterexample at the same p.

(m1) $T\mapsto VTV$ with $V=\operatorname{diag}(1,-1)$. It changes the signs of t_{12} and t_{21}. Here $V=V^{-1}$ is an isometry of ℓ_2^p, and $P(VTV)=VP(T)V$.

(m2) $T\mapsto XTX$ with the coordinate exchange $X=\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$. It exchanges t_{11} with t_{22}, and t_{12} with t_{21}. Again $X=X^{-1}$ is an isometry, and $P(XTX)=XP(T)X$.

(m3) $T\mapsto -T$, with P replaced by $\widetilde{P}(z)=P(-z)$. Then $\widetilde{P}(-T)=P(T)$, and $\|\widetilde{P}(S)\|_p=\|P(-S)\|_p=\|P(S)\|_p$, because $-S=USU^{-1}$ with the isometry $U=\operatorname{diag}((-1)^n)$ of $\ell^p(\mathbb{N})$.

In addition, transposition takes a counterexample at p to one at p', by Lemma 2.3. All four operations keep the signs of $t_{11}t_{22}$ and of $t_{12}t_{21}$, and by Corollary 10.4(iii) exactly one of these two products is negative. Hence (m1) and (m3) carry every real 2×2 counterexample at p to a counterexample in one of two classes:

$$\textit{Class I: } t_{11},t_{22},t_{12}>0>t_{21}; \qquad \textit{Class II: } t_{11}>0>t_{22} \text{ and } t_{12},t_{21}>0.$$

Indeed, if $t_{12}t_{21}<0$, then $t_{11}t_{22}>0$, and (m3), if needed, makes $t_{11}>0$, hence $t_{22}>0$; then (m1), if needed, also makes $t_{12}>0>t_{21}$. If $t_{11}t_{22}<0$, then $t_{12}t_{21}>0$; (m3), if needed, makes $t_{11}>0>t_{22}$, and then (m1), if needed, makes $t_{12},t_{21}>0$. The sign of $t_{11}t_{22}$ separates the two classes; equivalently, $\det T>0$ in class I and $\det T<0$ in class II. We say that a real 2×2 matrix *belongs to* class I (or II) if (m1) and (m3) carry it into that class. So a search among real 2×2 matrices may fix all signs and search two families of matrices with positive entries.

Remark 10.7 (the known counterexamples are in class I). (1) *The matrices of Theorem A.* The matrix of Lemma 4.1 is $T_\varepsilon=\exp(\varepsilon K_*)J=e^{-\varepsilon D_*/2}R_{\varepsilon/2}J=e^{-\varepsilon D_*/2}R_{\pi/2+\varepsilon/2}$, that is,

$$T_\varepsilon=e^{-\varepsilon D_*/2}\begin{pmatrix}-\sin(\varepsilon/2)&\cos(\varepsilon/2)\\-\cos(\varepsilon/2)&-\sin(\varepsilon/2)\end{pmatrix}.$$

For $0<\varepsilon<\pi$ all four entries are nonzero, three of them are negative, and $t_{11}t_{22}/(t_{12}t_{21})=-\tan^2(\varepsilon/2)<0$. So T_ε passes the tests (ii) and (iii) of Corollary 10.4, and it belongs to class I ($-VT_\varepsilon V$, obtained by (m3) and (m1), satisfies the inequalities of class I). The same holds for its transpose, which Theorem A uses for $p<2$. At $\varepsilon=0$ the matrix is J, whose diagonal is zero. So J has the form in Lemma 10.3(a) (indeed $|J|=X$), and the lemma gives $\|P(J)\|_p\le\|P(S)\|_p$, in agreement with Lemma 2.2. Thus the matrices of Lemma 4.1 leave the set of Lemma 10.3(a), where no counterexample can be, as soon as $\varepsilon>0$. More generally, a scaled rotation cR_θ with $c>0$ and $\sin\theta\cos\theta\neq 0$ has $t_{11}t_{22}=c^2\cos^2\theta>0>-c^2\sin^2\theta=t_{12}t_{21}$, so it belongs to class I. The scaled rotations of Theorem 6.3 and of Theorem 9.1(a) are of this kind, with $\cos\theta<0$.

(2) *Drury's matrix.* Drury's contraction at $p=4$ [6] is the matrix of Remark 7.6,

$$T_{\mathrm{Dr}}=\begin{pmatrix}-0.292187&0.812218\\-0.958489&-0.113428\end{pmatrix}.$$

Its products are $t_{11}t_{22}=0.033142187036>0$ and $t_{12}t_{21}=-0.778502018602<0$ (exact products of the decimals), so the ratio is about -0.0426, three entries are negative, and T_{Dr} belongs to class I: $-VT_{\mathrm{Dr}}V$ satisfies the inequalities of class I. By Lemma 10.3 the ratio must be negative, since T_{Dr} is a counterexample. In the same way, Corollary 10.4(i) forces $\||T_{\mathrm{Dr}}|\|_4>1$.

(3) *The matrices of Theorem 9.1.* The matrices T_i of all 560 pieces of Theorem 9.1(b)–(d), as printed in the logs named in Sections 9.6–9.8, and the four rational matrices at $p=14,16,18,20$ in Section 9.7 have the sign pattern $\left(\begin{smallmatrix}-&+\\-&-\end{smallmatrix}\right)$; so does the scaled rotation T_q of item (a), as $\cos\varphi<0<\sin\varphi$.

So T_ε, T_{Dr} and the matrices of Theorem 9.1 have the same sign pattern, $\left(\begin{smallmatrix} - & + \\ - & - \end{smallmatrix}\right)$, and their transposes, used for $p < 2$, have the transposed pattern. All known counterexamples belong to class I; the results of this section do not exclude class II.

Acknowledgements. The author thanks S. K. Ray for correspondence about this work.[3]

This work was done overwhelmingly by AI models and supervised by the author. The models were Anthropic's Claude and OpenAI's ChatGPT. The author takes full responsibility for the paper.

References

[1] M. A. Akcoglu and L. Sucheston, Dilations of positive contractions on L_p spaces, Canad. Math. Bull. 20 (1977), 285–292.

[2] C. Arhancet, On Matsaev's conjecture for contractions on noncommutative L^p-spaces, J. Operator Theory 69 (2013), 387–421 (arXiv:1009.1292).

[3] J. Bergh and J. Löfström, Interpolation Spaces. An Introduction, Grundlehren der Mathematischen Wissenschaften 223, Springer, Berlin, 1976.

[4] R. R. Coifman, R. Rochberg and G. Weiss, Applications of transference: the L^p version of von Neumann's inequality and the Littlewood–Paley–Stein theory, in: Linear Spaces and Approximation (Oberwolfach 1977), P. L. Butzer and B. Sz.-Nagy (eds.), Internat. Ser. Numer. Math. 40, Birkhäuser, Basel, 1978, 53–67.

[5] J. M. Danskin, The Theory of Max-Min and its Application to Weapons Allocation Problems, Springer, 1967.

[6] S. W. Drury, A counterexample to a conjecture of Matsaev, Linear Algebra Appl. 435 (2011), 323–329.

[7] S. Fackler, Regularity properties of sectorial operators: counterexamples and open problems, in: Operator Semigroups Meet Complex Analysis, Harmonic Analysis and Mathematical Physics, W. Arendt, R. Chill and Y. Tomilov (eds.), Operator Theory: Advances and Applications 250, Birkhäuser, Cham, 2015, 171–197 (arXiv:1407.1142).

[8] D. Goel, The Matsaev (1971) conjecture is false for every $p \neq 2$, Zenodo (2026), https://doi.org/10.5281/zenodo.23691821

[9] O. Holtz and M. Karow, Real and complex operator norms, preprint (2005), arXiv:math/0512608.

[10] F. Johansson, Arb: efficient arbitrary-precision midpoint-radius interval arithmetic, IEEE Trans. Comput. 66 (2017), 1281–1292.

[11] J. Marcinkiewicz and A. Zygmund, Quelques inégalités pour les opérations linéaires, Fund. Math. 32 (1939), 115–121.

[12] The mpmath development team, mpmath: a Python library for arbitrary-precision floating-point arithmetic, version 1.2.1 (2021), https://mpmath.org

[13] N. K. Nikol'skii, Five problems about invariant subspaces, Zap. Nauchn. Sem. LOMI 23 (1971), 115–127 (Russian); English: Five problems on invariant subspaces, J. Soviet Math. 2 (1974), 441–450.

[14] V. V. Peller, An analogue of J. von Neumann's inequality for the space L_p, Dokl. Akad. Nauk SSSR 231 (1976), 539–542 (Russian).

[15] V. V. Peller, Approximations by isometries and V. I. Matsaev's hypothesis for absolute contractions of the space L^p, Funktsional. Anal. i Prilozhen. 12 (1978), 38–50 (Russian); English: Funct. Anal. Appl. 12 (1978), 29–38.

[16] V. V. Peller, An analogue of an inequality of J. von Neumann, isometric dilation of contractions, and approximation by isometries in spaces of measurable functions, Trudy Mat. Inst. Steklov. 155 (1981), 103–150 (Russian); English: Proc. Steklov Inst. Math. (1983), no. 1, 101–145.

[17] S. K. Ray, On multivariate Matsaev's conjecture, Complex Anal. Oper. Theory 14 (2020), Paper No. 42 (arXiv:1703.00733).

Email address: david44hello@gmail.com

[3]When the author sent S. K. Ray a preprint of this paper on 28 October 2026, he replied that he had arrived at the result independently, by a somewhat different approach.