hexagon:2611.00547v1

THE MATSAEV (1971) CONJECTURE IS FALSE FOR EVERY $p \neq 2$

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 right shift on $\ell^p(\mathbb{N})$; for $p = 2$ this is von Neumann's inequality. Drury (2011) disproved it for $p = 4$, and hence, by duality, for $p = 4/3$, with a computer-assisted example. We prove that it fails for every $p \in (1,\infty)\setminus\{2\}$, even with T a real 2×2 matrix and P a real quadratic polynomial.[1] Let $P_a(z) = 1+az-Bz^2$ with $B > 0$. For $p \neq 2$ and small $a > 0$ we show that $\|P_a(S)\| = \|P_a(J)\|$, where J is the rotation of ℓ^p_2 ($\mathbb{C}^2$ with the ℓ^p norm) by a right angle. For $p > 2$ the bound $\|P_a(S)\| \le \|P_a(J)\|$ comes from a Bellman function: a function G of two real variables such that, for every real sequence $x \in \ell^p(\mathbb{N})$ (with $x_{-1} = x_{-2} = 0$) and every $n \ge 0$, $|(P_a(S)x)_n|^p \le \|P_a(J)\|^p|x_n|^p + G(x_{n-1},x_{n-2}) - G(x_n,x_{n-1})$; the G terms cancel in the sum over n. For $1 < p < 2$ the bound follows by duality. If $B > 1$, a rotation by an angle slightly larger than $\pi/2$, scaled down to a contraction, gives a real 2×2 matrix T with $\|P_a(T)\| > \|P_a(J)\| = \|P_a(S)\|$. In the other direction, the conjecture holds for every quadratic whose constant and leading coefficients have equal modulus.[2]

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 [9, Problem V]; some authors date it to 1966 [13]. 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 (Theorem A), and that it holds for every quadratic P with $|c_0| = |c_2|$ (Theorem B).

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. The table below compares our two main results with what was known before, and Figure 1 shows the conjecture for each p before and after this paper.

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, Bellman function.

[1] Expanded version, with further results [6]: https://doi.org/10.5281/zenodo.23691816

[2] Additionally, all results of this paper 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

Best result before this paper	**New result**
Known to fail for $p = 4$ and $p = 4/3$, by a computer-assisted proof (Drury 2011 [4]).	Matsaev's conjecture fails for every $p \in (1,\infty) \setminus \{2\}$, with a real 2×2 contraction and a real quadratic (Theorem A, proved in Section 5)
For $p = 2$, known for every P (von Neumann's inequality). For $p \neq 2$, to our knowledge, known to hold for all contractions only in the trivial case $\|P(S)\|_p = \sum |c_k|$ (see the Background below), for example when $\deg P \leq 1$, or when P is real with coefficients all of one sign or of alternating signs [2, Prop. 6.1].	Matsaev's conjecture holds for every quadratic $P = c_0 + c_1 z + c_2 z^2$ with $\|c_0\| = \|c_2\|$, for every p and every contraction (Theorem B, proved in Section 5)

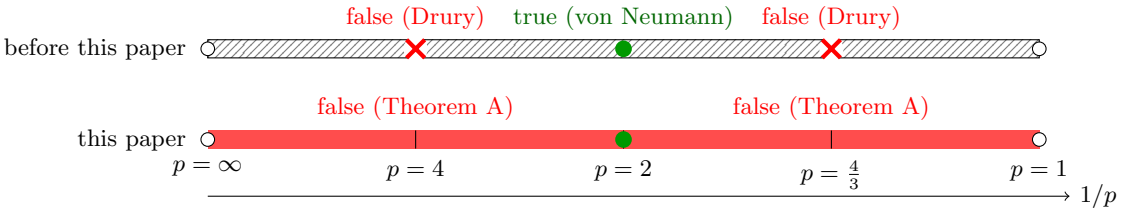


FIGURE 1. Matsaev's conjecture for each $p \in (1,\infty)$, placed by $1/p$, so that p and its dual exponent $p' = p/(p-1)$ lie symmetrically about $p = 2$. Top, before this paper: true for $p = 2$ (von Neumann's inequality), false for $p = 4$ and $p = 4/3$ (Drury [4]), and not known otherwise (hatched). 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 $\|T^k\|_p \leq 1$ give the *trivial bound* $\|P(T)\|_p \leq \sum_k |c_k|$. 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 $0 \leq n \leq d$ 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 ($Tx \geq 0$ whenever $x \geq 0$), by the dilation theorem of Akcoglu and Sucheston [1] combined with the transference principle of Coifman, Rochberg and Weiss [3], and, more generally, for contractions T with a positive contractive majorant (a positive contraction T_+ with $|Tx| \leq T_+|x|$ for every x), because such contractions have isometric dilations (Peller [12]; 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 [10, 11].

Drury [4] 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 [5, 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 [5, 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 [13]); 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 $\|T\|_p \leq 1$ 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 $D_* = \max_{0\leq t\leq 1}(t - t^{p-1})/(1+t^p) > 0$ 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 $\|P(T)\|_p \leq \|P(S)\|_p$ for every $1 \leq p \leq \infty$ 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 $\|P(J)\|_p \leq \|P(S)\|_p$ 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 $1 + B = \|P_0(J)\|_p \leq \|P_0(S)\|_p \leq 1 + B$ (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 $|(P_a(S)x)_n|^p \leq C_a^p |x_n|^p + G(x_{n-1}, x_{n-2}) - G(x_n, x_{n-1})$ for every real $x \in \ell^p(\mathbb{N})$ (with $x_{-1} = x_{-2} = 0$) and every $n \geq 0$; in the sum over n the G terms cancel, so $\|P_a(S)\|_p \leq C_a$ (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 $\xi \neq 0$,

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

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 $\frac{d}{d\varepsilon}\log\|R_{\varepsilon/2}\xi\|_p \leq \frac{D_*}{2}$, and integration from 0 to ε gives $\|R_{\varepsilon/2}\xi\|_p \leq e^{\varepsilon D_*/2}\|\xi\|_p$ for $\varepsilon \geq 0$. As J is an isometry of ℓ_2^p (it permutes the coordinates up to sign), $\|T_\varepsilon\xi\|_p = e^{-\varepsilon D_*/2}\|R_{\varepsilon/2}J\xi\|_p \leq \|J\xi\|_p = \|\xi\|_p$; 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 and relation to earlier work.

- For $1 < p < \infty$, $p \neq 2$, we show in two cases that $\|P(S)\|_p$ equals the norm of an explicit 2×2 matrix: for every quadratic with $|c_0| = |c_2|$ (Corollary 5.1), and for $P_a(z) = 1 + az - Bz^2$, $B > 0$, with $0 < a < a_0$ (Theorem 3.6; $a_0 > 0$ depends on p and B and is not explicit). This norm can be strictly smaller than $\sum |c_k|$ (Remark 5.2). To our knowledge, these are the first exact formulas for $\|P(S)\|_p$, $p \neq 2$, other than the trivial formula $\|P(S)\|_p = \sum |c_k|$; for real P with nonzero coefficients, the polynomials with $\|P(S)\|_p = \sum |c_k|$ are characterised in [2, Prop. 6.1].
- The proof of Theorem B is elementary: it writes $P(T)x$ as an average of two vectors and uses only the convexity of $|\cdot|^p$ and Lemma 2.2; unlike the earlier positive results for classes of contractions recalled in the Background, it uses no dilation.
- An expanded version of this paper, with many other results, is posted on Zenodo [6]. It gives explicit counterexamples for every $p \in [40/39, 40]$, $p \neq 2$ (previously, to our knowledge, the only explicit counterexample was Drury's, for $p = 4$ and $p = 4/3$ [4]). It also gives, with a proof by hand, an explicit counterexample in which the polynomial and the rotation angle do not depend on p: $P(z) = 1 + 2z - \frac{6}{5}z^2$ and $T = R_\theta/\|R_\theta\|_p$, where $\theta \in (\pi/2, \pi)$ and $\cos\theta = -\frac{1}{12}$, violate (1) for every $p \neq 2$ close enough to 2. These results are not used here.

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 [8] (see also [7]). 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 $X_{n-k} = J^k X_n$, so $y_n := \sum_k c_k x_{n-k} = (P(J)X_n)_1$. 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 $j \geq 1$ let $x^{(j)} \in \ell^p(\mathbb{N})$ agree with x for $0 \leq n < 4j$ and vanish elsewhere. Then $\|x^{(j)}\|_p^p = 2j\|\xi\|_p^p$, and $(P(S)x^{(j)})_n = y_n$ for $\deg P \leq n < 4j$, while $(P(S)x^{(j)})_n = 0$ for $n \geq 4j + \deg P$; 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 $\sum_{n \geq 0} |(P(S)x^{(j)})_n|^p$ and $\sum_{0 \leq n < 4j} |y_n|^p = 2j\|P(J)\xi\|_p^p$ differ only in the at most $2 \deg P$ terms with $n < \deg P$ or $4j \leq n < 4j + \deg P$, and in these terms $|(P(S)x^{(j)})_n|$ and $|y_n|$ are at most $\sum_k |c_k| \max_m |x_m|$). Hence $\|P(S)\|_p \geq \|P(J)\xi\|_p / \|\xi\|_p$ for every $\xi \neq 0$, which is the claim. For $p = \infty$ the inequality is trivial: $\|P(J)\|_\infty \leq \sum |c_k| = \|P(S)\|_\infty$. □

For $1 < p < \infty$, Lemma 2.2 also follows from Peller's results [12], 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 $\|P(S)\|_p \leq \|\mathcal{C}\|_p$; 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 $\|\mathcal{C}\|_p \leq \|P(S)\|_p$. 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*

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

then $\|P_a(S)\|_p \leq C$. 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 $0 \leq n \leq n_0 + 2$: the G terms telescope to $G(0,0) - G(0,0) = 0$, so $\|P_a(S)x\|_p^p \leq C^p\|x\|_p^p$. 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 $a = 0$: $F_0(u,v,w) = (1+B)^{p-1}\big(|u|^p + B|w|^p\big) - |u - Bw|^p \geq 0$, 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 $|u - Bw|^p \leq (1+B)^{p-1}(|u|^p + B|w|^p)$, 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 $F_0(u, v, w) \geq \gamma_1(|u| + |\nu|)^{p-2}\nu^2$ 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 $|h| \leq 1$ 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 $D_* := \max_{\xi\neq 0} Q(\xi) = f(t_*) > 0$ 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 $|\xi_2| \leq |\xi_1|$; then ξ is a multiple of $(1, \sigma)$ with $|\sigma| \leq 1$, and $Q(\xi) = f(\sigma)$. Since $f(\sigma) \leq 0$ 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).

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 $\Sigma_a \geq 0$ 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 $\mathcal{O}_0$;*
- *(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 $[0, a_1)$.*

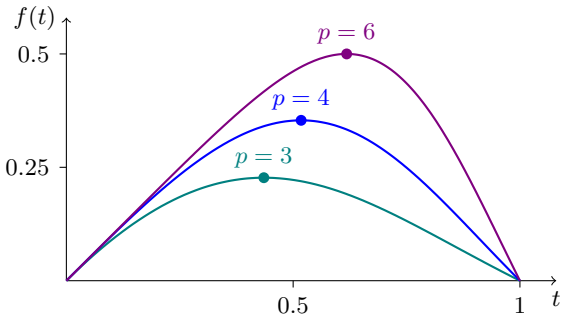


FIGURE 2. The function $f(t) = (t - t^{p-1})/(1 + t^p)$ of Lemma 3.4 on $[0, 1]$, for $p = 3, 4, 6$. The dot marks its maximum $D_* = f(t_*)$ (for $p = 4$ and $p = 6$, the values given above). Plotted numerically; an illustration, not used in the proofs.

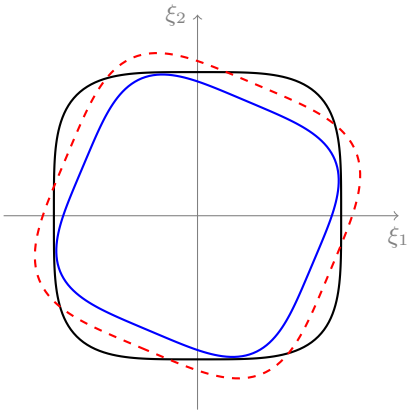


FIGURE 3. The real unit circle of ℓ_2^4 (black); its image under the rotation $R_{\pi/2+\varepsilon/2}$ (red, dashed), which leaves the unit ball; and its image under $T_\varepsilon = e^{-\varepsilon D_*/2} R_{\pi/2+\varepsilon/2}$ (blue), which stays inside it (Lemma 4.1). Here $\varepsilon = 0.8$,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

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 $\partial_\theta^2 H(a,\cdot) \leq -m_1 < 0$ 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 $\geq \rho$ from $\mathcal{O}_a$, it is at angular distance $\geq \rho/2$ from $\mathcal{O}_0$, so $\Sigma_a(s) \geq a\delta$.

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 $\partial_\theta^2 H(a,\cdot) \leq -m_1$, and $\partial_\theta H(a,o_a) = 0$, so $c(a) - H(a,s) \geq \frac{1}{2}m_1 r(s)^2$. Hence $\Sigma_a(s) \geq a\eta\, r(s)^2$ 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 $0 \leq a < a_0$; for $1 < p < 2$ the same holds for $0 \leq a < a_0(p',B)$.*

Proof. By Lemma 2.2 only $\|P_a(S)\|_p \leq C_a$ 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 $a = 0$, $F_0 \geq 0$ 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 $G_a = G_0 + \frac{1}{2}\Delta_a + k_a$, 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 $\frac{1}{4}\Sigma_a(s)$ at each of the four windows of the cycle; in particular $F_a(\hat{s}) = \frac{1}{4}\Sigma_a(s) \geq 0$, 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 $F_a \geq 0$, 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 $\mathcal{O}_a$, $\Sigma_a \geq a\delta$ 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, $F(\hat{s}) = \frac{1}{4}\Sigma_a(s)$ for every s if $G(Js)-G(s) = \frac{1}{2}\big(\Delta_a(Js)-\Delta_a(s)\big)$; the choice $G = \frac{1}{2}\Delta_a$ 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

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

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 $J^{\mathsf{t}} = J^{-1}$), 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 $|\lambda(a)| \le \Lambda a$ 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

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

Indeed, $|C_a - (1+B)| \le a\|J\|_p = a$, so $|C_a^p - (1+B)^p| \le p(1+B+a)^{p-1}a$ by the mean value theorem. By the mean value theorem, $\big||y+h|^p - |y|^p\big| \le p(|y|+|h|)^{p-1}|h|$ and (as $p > 2$) $|\phi(y+h) - \phi(y)| \le (p-1)(|y|+|h|)^{p-2}|h|$; with $y = u - Bw$, $h = av$, $|y| \le (1+B)\|\mathbf{x}\|_\infty$ and $|h| \le a\|\mathbf{x}\|_\infty$, 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 $|\lambda(a)| \le \Lambda a$.

Step 6: $F_a \ge 0$. 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 $s = 0$, $F_a(0,0,w) = F_0(0,0,w) + \mathcal{E}_a(0,0,w) \ge \big(B(1+B)^{p-1} - B^p - La\big)|w|^p \ge 0$ once $La \le B(1+B)^{p-1} - B^p$, 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

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

Case 1: s at angular distance $\ge \rho$ 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 $|\nu| \le \nu_1$, (8) and Lemma 3.5(ii) give $F_a \ge a\delta/4 - 2^{p-1}La\nu_1 \ge a\delta/8$. For $|\nu| \ge \nu_1$, $(|u|+|\nu|)^{p-2}\nu^2 \ge |\nu|^p$ and $(1+|\nu|)^{p-1} \le (1+1/\nu_1)^{p-1}|\nu|^{p-1}$, so $F_a \ge |\nu|^p\big(\gamma_1 - La(1+1/\nu_1)^{p-1}\big) \ge 0$ 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 $|\nu| \le \nu_0$ 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, $\partial_\nu^2 F_0 \ge 4q_*$ 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 $|\partial_\nu^2(F_a - F_0)| \le L_2 a$ there, so $q_a \ge q_*$ on $\mathcal{U}$ once $L_2 a \le 2q_*$. 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 $|N_a(s)| \le L_1 a r$. With Lemma 3.5(iii),

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

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 $|\nu| \ge \nu_0$, (8) with $\Sigma_a \ge 0$, $(|u|+|\nu|)^{p-2}\nu^2 \ge |\nu|^p$ and $(1+|\nu|)^{p-1} \le (1+1/\nu_0)^{p-1}|\nu|^{p-1}$ gives $F_a \ge |\nu|^p\big(\gamma_1 - La(1+1/\nu_0)^{p-1}\big) \ge 0$ 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 $\|P_a(S)\|_p \le C_a = \|P_a(J)\|_p$.

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 $0 \le a < a_0(p', B)$. Since $J^{\mathsf{t}} = -J = VJV$ with the isometry $V = \operatorname{diag}(1, -1)$ 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 $\|T_\varepsilon\|_p \le 1$ for all $\varepsilon \ge 0$, 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 $\xi(\varepsilon) = \exp(\varepsilon K_*)\xi$, $\frac{d}{d\varepsilon}\|\xi(\varepsilon)\|_p^p = p\langle\phi(\xi(\varepsilon)), K_*\xi(\varepsilon)\rangle = \frac{p}{2}\|\xi(\varepsilon)\|_p^p\,(Q(\xi(\varepsilon)) - D_*) \le 0$. 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 $\|P_a(J)\|_p \geq \|P_a(S)\|_p$. 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 $\|P_a(T \oplus 0)\|_p = \max(\|P_a(T)\|_p, 1) \geq \|P_a(T)\|_p$, 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 $\|P(T)\|_p \leq |c_1| = \|P(S)\|_p$ 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 $\|\widetilde{P}(T')\|_p \leq \|\widetilde{P}(S)\|_p$ for every contraction T', then $\|P(T)\|_p = |c_0|\|\widetilde{P}(\zeta T)\|_p \leq |c_0|\|\widetilde{P}(S)\|_p = |c_0|\|U\widetilde{P}(S)U^{-1}\|_p = |c_0|\|\widetilde{P}(\zeta S)\|_p = \|P(S)\|_p$. 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 $\|T\|_p \leq 1$,

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

where $M = \begin{pmatrix} 2 & \alpha \\ \alpha & -2 \end{pmatrix}$ acts pointwise on (x, Tx). As $M = V\widetilde{P}(J)$ with $V = \operatorname{diag}(1, -1)$, $\|M\|_p = \|\widetilde{P}(J)\|_p \leq \|\widetilde{P}(S)\|_p$ 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 $1 \leq p \leq \infty$, $\|P(S)\|_p = \|P(\bar{\zeta}J)\|_p = \max_{|\mu|=1}\|P(\mu J)\|_p$.*

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 $\|\widetilde{P}(S)\|_p \leq \|\widetilde{P}(J)\|_p$, 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 $\|P(\mu J)\|_p \leq \|P(\mu S)\|_p = \|P(S)\|_p$. (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 $\|P(T)\|_p \leq \sum |c_k|$: for $\widetilde{P}(z) = 1 + \alpha z - z^2$ with real $\alpha \neq 0$ ($\zeta = 1$ 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 $\xi \neq 0$ 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.

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.

[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.

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] 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.

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

[5] 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).

[6] D. Goel, The Matsaev (1971) conjecture is false for every $p \neq 2$, and other miscellanea, Zenodo (2026), doi:10.5281/zenodo.23691816

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

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

[9] 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.

[10] 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).

[11] 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.

[12] 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.

[13] 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