Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of Functional Analysis 255 (2008) 2861–2871
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

The dth linear polarization constant of $\mathbb{R}^d$

Yuk J. Leung, Wenbo V. Li [1], Rakesh *

Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA

Received 16 January 2008; accepted 13 June 2008

Available online 16 July 2008

Communicated by K. Ball

Abstract

In [C. Benítez, Y. Sarantopoulos, A. Tonge, Lower bounds for norms of products of polynomials, Math. Proc. Cambridge Philos. Soc. 124 (3) (1998) 395–408] it was conjectured that for all unit vectors $u_1, \dots, u_d$ in $\mathbb{R}^d$,

$$\mathcal{X}(u_1, \dots, u_d) := \sup_{x \in \mathbb{R}^d,\, |x|^2 = d} \prod_{i=1}^{d} \langle x, u_i \rangle^2 \geqslant 1$$

with equality occurring iff $u_1, \dots, u_d$ are orthonormal. We relate this to a conjecture about solutions of $Ay = y^{-1}$, where $A = (\langle u_i, u_j \rangle)$, and show that if the conjecture fails then the $u_1, \dots, u_d$ minimizing $\mathcal{X}$ must be linearly dependent. We also show $\mathcal{X}(u_1, \dots, u_d) \geqslant 1$ for certain families of $u_1, \dots, u_d$.

Keywords: Polarization constant; Product of linear functionals

1. Introduction

In the study of polynomials on Banach spaces in [4], Benítez, Sarantopoulos and Tonge introduced the concept of a linear polarization constant. For any positive integer d and any Banach space X, the dth linear polarization constant for X is $c_d(X)$ defined by

* Corresponding author.
E-mail addresses: yleung@math.udel.edu (Y.J. Leung), wli@math.udel.edu (W.V. Li), rakesh@math.udel.edu (Rakesh).
[1] Supported in part by NSF Grant DMS-0505805.

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2008.06.011

$$c_d(X)^{-1} := \inf_{f_i \in X^*,\, \|f_i\|=1} \|f_1 \cdots f_d\| = \inf_{f_i \in X^*,\, \|f_i\|=1} \sup_{x \in X,\, \|x\|=1} \left|f_1(x) \cdots f_d(x)\right|.$$

If $\mathbb{R}^d$ is endowed with the standard inner product then it was conjectured in [4] that $c_d(\mathbb{R}^d) = d^{d/2}$ (independently, [11] also discussed this) which may be restated as the following conjecture.

Conjecture 1. *If $u_1, \ldots, u_d$ are unit vectors in $\mathbb{R}^d$ then*

$$\mathcal{X}(u_1, \ldots, u_d) := \sup_{x \in \mathbb{R}^d,\, |x|^2=d} \left|\langle x, u_1\rangle \cdots \langle x, u_d\rangle\right|^2 \geqslant 1, \tag{1}$$

with equality occurring iff $u_1, \ldots, u_d$ are orthonormal.

One may easily check that if $u_1, \ldots, u_d$ are orthonormal then the maximizing x are $\sum_{i=1}^d \pm u_i$ and hence $\mathcal{X}(u_1, \ldots, u_d) = 1$; the main difficulty in the conjecture is to show that $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$ for all $u_1, \ldots, u_d$—see Section 6.2.

The article [15] has a proof of Conjecture 1 when $d \leqslant 5$ and Theorem 2 in [14] shows that $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$ if $u_1, \ldots, u_d$ are close to orthonormal unit vectors. If the span of $u_1, \ldots, u_d$ is two-dimensional then [1] shows that $\mathcal{X}(u_1, \ldots, u_d) \geqslant 4(d/4)^d$ which is attained iff $\pm u_1, \ldots, \pm u_d$ are $2d$ equally spaced vectors in a plane. Conjecture 1, for $\mathbb{C}^d$ instead of $\mathbb{R}^d$, was shown to be true in [2], where it was also observed that this lead to a lower bound for $\mathcal{X}$ for the $\mathbb{R}^d$ case too. Using these ideas, for the $\mathbb{R}^d$ case, [7] and [11] gave better bounds, culminating in [16], where it was shown that $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1/2^{d-2}$ for all $u_1, \ldots, u_d$. Using an averaging argument, the recent article [6] gives the improved lower bound $\mathcal{X}(u_1, \ldots, u_d) \geqslant (d/2)^d \Gamma(d/2)/\Gamma(3d/2)$ for the $\mathbb{R}^d$ case. Using a geometric approach, for the $\mathbb{R}^d$ case, [13] gives an interesting lower bound $\mathcal{X}(u_1, \ldots, u_d) \geqslant \det(\langle u_i, u_j\rangle)$—also see [12] for some older results. Ball's solution of the complex plank problem in [3], where he proved the existence of an $x \in \mathbb{C}^d$, $|x|^2 = d$ with $|\langle x, u_i\rangle| \geqslant 1$ for all $i = 1, \ldots, d$, is a result stronger than Conjecture 1 for the complex case.

Below all matrices and vectors will have real entries and all vectors will be thought of as column vectors. We prove three results in this article. The first relates to the conjecture for all $u_1, \ldots, u_d$ and the other two show that $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$ when $u_1, \ldots, u_d$ are restricted to certain families of unit vectors.

Theorem 1. *If there exist linearly independent unit vectors $u_1, \ldots, u_d$ in $\mathbb{R}^d$ so that $\mathcal{X}(u_1, \ldots, u_d) < 1$ then there exist linearly dependent unit vectors $v_1, \ldots, v_d$ in $\mathbb{R}^d$ so that $\mathcal{X}(v_1, \ldots, v_d) < \mathcal{X}(u_1, \ldots, u_d)$.*

This shows that if Conjecture 1 is false then the $u_1, \ldots, u_d$ which minimize $\mathcal{X}(u_1, \ldots, u_d)$ must be linearly dependent. However, our intuition suggests that given a linearly dependent set of $u_1, \ldots, u_d$ one should be able to find a linearly independent set of $w_1, \ldots, w_d$ with $\mathcal{X}(w_1, \ldots, w_d) < \mathcal{X}(u_1, \ldots, u_d)$; if this could be shown then Conjecture 1 would follow from Theorem 1. We have made some progress in this direction which is described in Section 6.1.

Theorem 2. *Suppose $u_1, \ldots, u_d$ are unit vectors in $\mathbb{R}^d$ and ε some ± 1 vector in $\mathbb{R}^d$ such that $\varepsilon_i \varepsilon_j \langle u_i, u_j\rangle \geqslant 0$ for all $i, j = 1, \ldots, d$. Then $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$ with equality occurring iff $u_1, \ldots, u_d$ are orthonormal; in fact there exists an $x \in \mathbb{R}^d$, $|x|^2 = d$, so that $|\langle x, u_i\rangle| \geqslant 1$ for all $i = 1, \ldots, d$.*

Theorem 2 asserts that $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$ if, after direction reversal, the u_i make acute angles with each other. When, say $\varepsilon_i = 1$ for all $i = 1, \ldots, d$, one is tempted to use as x a vector parallel to $\sum_{i=1}^d u_i$; $\sum_{i=1}^d u_i$ has the longest length amongst all vectors $\sum_{i=1}^d \sigma_i u_i$ with $\sigma_i = \pm 1$. However, this vector can fail to satisfy even the condition $\prod_{i=1}^d \langle x, u_i \rangle^2 \geqslant 1$ as shown in Theorem 3 in [14].

Our last result is that $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$ if $\pm u_1, \ldots, \pm u_d$ form a root system. Please refer to [8] for additional details about root systems and the connection with finite reflection groups.

Definition 2. Suppose n and d are positive integers with $n < d$. For any unit vector $u \in \mathbb{R}^n$ let S_u be the orthogonal transformation on $\mathbb{R}^n$ consisting of reflection across the hyperplane $\langle x, u \rangle = 0$. A subset $\Phi = \{\pm u_1, \ldots, \pm u_d\}$ of the unit vectors in $\mathbb{R}^n$ will be called a root system if $S_u \Phi = \Phi$ for all $u \in \Phi$.

Let $e_1, \ldots, e_n$ be the standard basis for $\mathbb{R}^n$ and u a unit vector in $\mathbb{R}^d$ and let $S_u e_i = \sum_{j=1}^n \alpha_{ij} e_j$, for $i = 1, \ldots, n$. Then S_u may also be considered a homomorphism on the algebra $\mathbb{R}[x_1, \ldots, x_n]$ of multinomials in $x_1, \ldots, x_n$ with real coefficients, by taking $S_u x_i := \sum_{j=1}^n \alpha_{ji} x_j$ for $i = 1, \ldots, n$. The subalgebra of $\mathbb{R}[x_1, \ldots, x_n]$ consisting of multinomials fixed by all the S_{u_i}, $i = 1, \ldots, d$, is generated by the constant 1 and n algebraically independent homogeneous multinomials $f_1, \ldots, f_n$ with positive degrees $d_1, \ldots, d_n$ (Chevalley's theorem). While there is more than one choice for the f_i, the d_i are uniquely determined by the root system.

Theorem 3. *Suppose n, d are positive integers with $n < d$ and $u_1, \ldots, u_d$ are unit vectors in $\mathbb{R}^n$ so that $\{\pm u_1, \ldots, \pm u_d\}$ forms a root system. Then*

$$\mathcal{X}(u_1, \ldots, u_d) = 2^{-2d} \prod_{i=1}^n d_i^{d_i}$$

and in particular $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$.

The root systems are built by combining orthogonal collections of 11 basic root systems corresponding to the 11 possible, basic finite reflection groups—see [8]. So Theorem 3 proves $\mathcal{X}(u_1, \ldots, u_d) \geqslant 1$ only for those $u_1, \ldots, u_d$ which can be constructed from orthogonal unions of a finite number of irreducible root systems, but it does provide the exact value of $\mathcal{X}(u_1, \ldots, u_d)$ in these cases. Further, $u_1, \ldots, u_d$ lead to non-trivial root systems only if they are linearly dependent and as seen in Theorem 1 the linearly dependent case is important and not well understood. The proof of Theorem 3 uses an integral relation conjectured by Macdonald—see [5].

Using Lagrange multipliers, a maximizing x in the definition of $\mathcal{X}(u_1, \ldots, u_d)$ satisfies $\sum_{j=1}^d \frac{1}{\langle x, u_j \rangle} u_j = \lambda x$ for some λ. Taking the dot product of this relation with x and noting that $|x|^2 = d$, we obtain $\lambda = 1$ and hence

$$\sum_{j=1}^d \frac{1}{\langle x, u_j \rangle} u_j = x. \tag{2}$$

Taking the dot product of this relation with u_i, we obtain

$$\sum_{j=1}^{d} \frac{\langle u_i, u_j \rangle}{\langle x, u_j \rangle} = \langle x, u_i \rangle, \quad i = 1, \dots, d. \tag{3}$$

Hence all maximizers are also solutions of (3). It is certainly possible (if $u_1, \dots, u_d$ are linearly dependent) that solutions of (3) may not satisfy (2). However, if x satisfies (3) then $x = z + \sum_{i=1}^{d} \frac{u_i}{\langle x, u_i \rangle}$ for some vector z which is orthogonal to all the u_i. Now the choice of z does not affect $\langle x, u_i \rangle$, so clearly if $|x|^2 = d$ then the maximizer should have no component perpendicular to the u_i. So the maximizers $x \in \mathbb{R}^d$ are to be found amongst the solutions of (3) which lie in the span of the u_i.

If $A = (\langle u_i, u_j \rangle)$ then $\operatorname{diag}(A) = I$ and $A \geqslant 0$; conversely, every $A \geqslant 0$ with $\operatorname{diag}(A) = I$ arises in this fashion. Define $y \in \mathbb{R}^d$ by

$$y := \left[\langle x, u_1 \rangle^{-1}, \dots, \langle x, u_d \rangle^{-1} \right];$$

then (2) and (3) may be written as $x = \sum_{j=1}^{d} y_j u_j$ and $Ay = y^{-1}$ where y^{-1} is to be taken componentwise. Also note that $\prod_{j=1}^{d} y_j^{-2} = \prod_{j=1}^{d} \langle x, u_j \rangle^2$.

The following proposition will be useful in the proof of our results and is of interest on its own. For any ± 1 vector $\varepsilon = [\varepsilon_1, \dots, \varepsilon_d]$ in $\mathbb{R}^d$, the set

$$Q := \left\{ y \in \mathbb{R}^d \;\middle|\; \operatorname{sign}(y_i) = \varepsilon_i \text{ or } 0 \text{ for all } i = 1, \dots, d \right\}$$

will be called a quadrant of $\mathbb{R}^d$. For any $d \times d$ matrix $A \geqslant 0$, we define, respectively, the hollow and the solid ellipsoids

$$\mathcal{A} := \left\{ y \in \mathbb{R}^d \;\middle|\; y^{\mathrm{T}} A y = d \right\}, \qquad \mathcal{E} := \left\{ y \in \mathbb{R}^d \;\middle|\; y^{\mathrm{T}} A y \leqslant d \right\}.$$

Proposition 3. *If $A \geqslant 0$ is a $d \times d$ matrix and Q a closed quadrant of $\mathbb{R}^d$, then*:

(a) *the equation $Ay = y^{-1}$, $y \in \mathbb{R}^d$, has at most one solution in Q*;
(b) *$\mathcal{E} \cap Q$ is bounded iff $Q \cap \ker A = \{0\}$*;
(c) *$Ay = y^{-1}$ has a solution in Q iff $\mathcal{E} \cap Q$ is bounded*;
(d) *$Ay = y^{-1}$ has a solution in Q iff $\sup_{y \in \mathcal{E} \cap Q} \prod_{i=1}^{d} y_i^2$ is finite. Further, the maximizer is unique and is the solution of $Ay = y^{-1}$.*

In view of this proposition and our work just before that, Conjecture 1 is equivalent to the following conjecture.

Conjecture 4. *For every $d \times d$ matrix $A \geqslant 0$ with* $\operatorname{diag}(A) = I$,

$$\mathcal{Y}(A) := \min_{y \in \mathbb{R}^d,\, Ay = y^{-1}} \prod_{i=1}^{d} y_i^2 \leqslant 1 \tag{4}$$

and equality occurs iff $A = I$.

It is clear that $\mathcal{Y}(I) = 1$ and the main difficulty is the proof of $\mathcal{Y}(A) \leqslant 1$. For future reference we explicitly state the connection between the two conjectures. Given unit vectors $u_1, \ldots, u_d$ in $\mathbb{R}^d$, let $A = (\langle u_i, u_j \rangle)$; then x is a maximizer in the definition of $\mathcal{X}(u_1, \ldots, u_d)$ iff $x = \sum_{j=1}^d y_j u_j$ where $y \in \mathbb{R}^d$ is a solution of $Ay = y^{-1}$ with the smallest $\prod_{i=1}^d y_i^2$ amongst all the solutions. Further $\langle x, u_j \rangle = y_j^{-1}$ and $\mathcal{X}(u_1, \ldots, u_d) = \mathcal{Y}(A)^{-1}$.

Ball's proof of the complex plank problem and in particular of Conjecture 1 for the complex case, was based on showing that $A\bar{z} = z^{-1}$ has a solution z with $|z_i| \leqslant 1$ for all $i = 1, \ldots, d$. However, in the complex case, $A\bar{z} = z^{-1}$ has an infinite number of solutions whereas in the real case $Ay = y^{-1}$ has at most 2^d solutions and, for some A, has no solutions with $|y_i| \leqslant 1$ for all $i = 1, \ldots, d$. The equation $Ay = y^{-1}$, in the real case, also shows up in an optimization problem studied by the authors in [9]; [10] has other interesting properties of the solutions of $Ay = y^{-1}$.

We state theorems equivalent to Theorems 1 and 2; these are the theorems we will prove.

Theorem 4. *If there is a matrix $A > 0$, $\operatorname{diag}(A) = I$, such that $\mathcal{Y}(A) > 1$ then there is a singular matrix $A_s \geqslant 0$ of the same size, $\operatorname{diag}(A_s) = I$, so that $\mathcal{Y}(A_s) > \mathcal{Y}(A)$.*

Theorem 5. *Suppose $A = (a_{ij}) \geqslant 0$ is a $d \times d$ matrix with $\operatorname{diag}(A) = I$ and ε some ± 1 vector in $\mathbb{R}^d$ such that $\varepsilon_i \varepsilon_j a_{ij} \geqslant 0$ for all $i, j = 1, \ldots, d$. Then $\mathcal{Y}(A) \leqslant 1$ with equality occurring iff $A = I$; in fact $Ay = y^{-1}$ has a solution with $|y_i| \leqslant 1$ for all $i = 1, \ldots, d$.*

One consequence of Theorem 5 is that $\mathcal{Y}(A) \leqslant 1$ for all tri-diagonal matrices $A \geqslant 0$ with $\operatorname{diag}(A) = I$.

In Section 2 we prove Proposition 3 and the continuity of $\mathcal{X}$ and $\mathcal{Y}$ and in Sections 3–5 we prove Theorems 4, 5 and 3. Given a singular $A \geqslant 0$, with the use of Theorem 4 to prove Conjecture 4 in mind, in Section 6.1 we suggest a procedure to possibly construct a $B \geqslant 0$, $\operatorname{diag}(B) = I$ with $\mathcal{Y}(B) > \mathcal{Y}(A)$; in Section 6.2 we study the equality part of Conjecture 4.

2. Solutions of $Ay = y^{-1}$ and continuity of $\mathcal{X}$ and $\mathcal{Y}$

Proposition 5. *$\mathcal{X}$ and $\mathcal{Y}$ are continuous functions on their domains.*

Proof. The domain of $\mathcal{X}$ consists of d tuples of unit vectors in $\mathbb{R}^d$ and the domain of $\mathcal{Y}$ consists of $d \times d$ matrices $A \geqslant 0$ with $\operatorname{diag}(A) = I$.

To every $A \geqslant 0$, $\operatorname{diag}(A) = I$ we can associate a unique positive semidefinite matrix $U = \sqrt{A}$, and if $u_1, \ldots, u_d$ are the columns of U then $\mathcal{X}(u_1, \ldots, u_d) = \mathcal{Y}(A)^{-1}$. Since $\sqrt{A}$ is a continuous function of A, the continuity of $\mathcal{Y}$ will follow from the continuity of $\mathcal{X}$.

The continuity of $\mathcal{X}$ follows from the following observation. Suppose X and U are compact metric spaces and $f : X \times U \to \mathbb{R}$ is a continuous function. Define $F : U \to \mathbb{R}$ with $F(u) = \sup_{x \in X} f(x, u)$; then F is continuous as shown next.

Suppose u_k is a sequence in U converging to u in U. From the compactness of X, there are x_k and x in X so that $F(u_k) = f(x_k, u_k)$ and $F(u) = f(x, u)$. Hence $f(x, u_k) \leqslant f(x_k, u_k)$; using this and the definitions we have

$$F(u) = f(x, u) = \lim_{k \to \infty} f(x, u_k) = \liminf_{k \to \infty} f(x, u_k) \leqslant \liminf_{k \to \infty} f(x_k, u_k) = \liminf_{k \to \infty} F(u_k).$$

Next, suppose u_{k_n} is a subsequence such that $\limsup_{k\to\infty} F(u_k) = \lim_{n\to\infty} F(u_{k_n})$. Further, let $x_{k_n} \in X$ so that $F(u_{k_n}) = f(x_{k_n}, u_{k_n})$. From the compactness of X, there is a subsequence $x_{k_{n_l}}$ which is convergent to some point x^*. Then

$$\limsup_{k\to\infty} F(u_k) = \lim_{l\to\infty} F(u_{k_{n_l}}) = \lim_{l\to\infty} f(x_{k_{n_l}}, u_{k_{n_l}}) = f\big(x^*, u\big) \leqslant F(u). \qquad \square$$

Proposition 6. *Suppose $B(t)$ is a continuously differentiable matrix for all t in an open interval J around 0, $B(t) \geqslant 0$ and $\operatorname{diag}(B(t)) = I$. If $B(0)y = y^{-1}$ has a solution in a quadrant Q then $B(t)y = y^{-1}$ has a (unique) solution in Q for all t in some interval around 0; further $y(t)$ will be continuously differentiable on this interval.*

Proof. Let $B(t) = (b_{ij}(t))$ and define the map $F : Q \times J \to \mathbb{R}^d$ by

$$F(z,t) = \big(F_1(z,t), \dots, F_d(z,t)\big),$$

where $F_i(z,t) = z_i \sum_{k=1}^d b_{ik}(t) z_k$. Then F is a continuously differentiable function of z and t and a solution of $B(t)z = z^{-1}$ in Q is exactly the solution of $F(z,t) = e$. Now $F(y,0) = e$ and

$$\frac{\partial F_i}{\partial z_j}(y,0) = b_{ij}(0) y_i + \delta_{ij} \sum_{k=1}^d b_{ik}(0) y_k = b_{ij}(0) y_i + \delta_{ij} y_i^{-1},$$

hence

$$F_z(y,0) = (b_{ij} y_i) + D\big(y^{-1}\big) = D(y)\big(B + D\big(y^{-2}\big)\big),$$

where $D(v)$ is the diagonal matrix whose diagonal entries are those of the vector v. Now $D(y)$ is invertible, $B \geqslant 0$ and $D(y^{-2}) > 0$, hence $F_z(y,0)$ is invertible, so an application of the implicit function theorem completes the proof. $\square$

Proof of Proposition 3. (a) Fix a quadrant Q of $\mathbb{R}^d$ and suppose p and q are two solutions of $Ay = y^{-1}$ in Q, so $Ap = p^{-1}$, $Aq = q^{-1}$, and $p_i/q_i > 0$ for $i = 1, \dots, d$. Hence using the AM–GM inequality, we have

$$\begin{aligned}(p-q)^{\mathrm T} A(p-q) &= p^{\mathrm T} A p + q^{\mathrm T} A q - p^{\mathrm T} A q - q^{\mathrm T} A p \\ &= p^{\mathrm T} p^{-1} + q^{\mathrm T} q^{-1} - p^{\mathrm T} q^{-1} - q^{\mathrm T} p^{-1} \\ &= d + d - \sum_{i=1}^d \big(p_i q_i^{-1} + q_i p_i^{-1}\big) \leqslant 2d - 2\sum_{i=1}^d \sqrt{p_i q_i^{-1} q_i p_i^{-1}} = 0.\end{aligned}$$

Further, since we used the AM–GM inequality, the inequality is a strict inequality unless $p_i/q_i = q_i/p_i$ for all $i = 1, \dots, d$; that is unless $p = q$ because p_i and q_i have the same sign. But $A \geqslant 0$ so we cannot have a strict inequality, hence $p = q$.

(b) Suppose v is a non-zero vector in $\ker A \cap Q$. Then $\sigma v \in Q$ for all $\sigma > 0$ and also $(\sigma v)^{\mathrm T} A(\sigma v) = 0$ and hence $\mathcal{E} \cap Q$ is unbounded. Conversely, suppose $\mathcal{E} \cap Q$ is unbounded.

Then we have a sequence y_n in Q with $y_n^{\mathrm{T}} A y_n \leqslant d$ and $\lim_{n\to\infty} |y_n| = \infty$. The bounded sequence of unit vectors $u_n = y_n/|y_n|$ in Q has a convergent subsequence in Q; without loss of generality we assume that u_n is convergent to a unit vector u in Q. Then

$$u^{\mathrm{T}} A u = \lim_{n\to\infty} u_n^{\mathrm{T}} A u_n = \lim_{n\to\infty} \frac{1}{|y_n|^2} y_n^{\mathrm{T}} A y_n \leqslant \lim_{n\to\infty} \frac{d}{|y_n|^2} = 0.$$

Since $A \geqslant 0$ so $Au = 0$ and hence Q contains an element of the kernel of A.

(c) and (d) Suppose $Ay = y^{-1}$ has a solution y in Q. If $Q \cap \mathcal{E}$ is unbounded then there is a non-zero vector v in $\ker A \cap Q$. Hence $v_i/y_i \geqslant 0$ for all $i = 1, \ldots, d$ with at least one of these quantities strictly positive. Then using $Av = 0$ and $Ay = y^{-1}$ we have $0 = y^{\mathrm{T}} A v = v^{\mathrm{T}} A y = v^{\mathrm{T}} y^{-1} = \sum_{i=1}^{d} v_i y_i^{-1} > 0$, which is impossible. Hence $Q \cap \mathcal{E}$ is bounded.

Conversely, suppose $\sup_{y\in Q\cap\mathcal{E}} \prod_{i=1}^{d} y_i^2$ is finite. If $Q \cap \ker A \neq (0)$ then there is a non-zero $v \in Q$ so that $Av = 0$. Pick any $z \in Q \cap \mathcal{E}$ all of whose components are non-zero and let $w = z + \sigma v$ for some $\sigma > 0$. Then $w^{\mathrm{T}} A w = z^{\mathrm{T}} A z \leqslant d$ while $\prod_{i=1}^{d} w_i^2$ can be made as large as we wish by increasing σ. Hence we must have $Q \cap \ker A = (0)$ so $\mathcal{E} \cap Q$ must be bounded. Then from compactness, the supremum is attained. Clearly the maximum is positive and is attained on the boundary of $Q \cap \mathcal{E}$ but not on the boundary of Q because $\prod_{i=1}^{d} y_i^2 = 0$ on the boundary of Q. Hence the maximum occurs on the relative interior of $Q \cap \mathcal{A}$. From Lagrange multipliers, the maximizing $y \in Q \cap \mathcal{A}$ must satisfy $Ay = \lambda y^{-1}$; but $y^{\mathrm{T}} A y = d$ so $\lambda = 1$ and hence the optimal y is the unique solution of $Ay = y^{-1}$ in Q. □

3. Proof of Theorem 4

Suppose $\mathcal{Y}(A) > 1$ for some $A > 0$ with $\mathrm{diag}(A) = I$; clearly $A \neq I$. Then $Ay = y^{-1}$ has exactly one solution in each quadrant Q of $\mathbb{R}^d$.

For any real number t define $B(t) = A + t(A - I)$; note that $\mathrm{diag}(B(t)) = I$ and at least for t near 0 we have $B(t) > 0$. Let $\tau > 0$ be the largest number so that $B(t) > 0$ on $[0, \tau)$. For future use we also note that $B(t) = A + t(A - I)$ so $B'(t) = A - I$ and $A - I = (B(t) - I)/(1 + t)$; here $'$ represents the t derivative.

Fix any quadrant Q, and for t in $[0, \tau)$, let $y = y(t)$ be the continuously differentiable solution of $B(t)y = y^{-1}$ in Q. Define $f : [0, \tau) \to \mathbb{R}$ by $f(t) := \prod_{i=1}^{d} y_i^2(t)$. Then

$$\frac{f'(t)}{f(t)} = 2\sum_{i=1}^{d} \frac{y_i'(t)}{y_i(t)} = 2y'(t)^{\mathrm{T}} y(t)^{-1} = 2y'(t)^{\mathrm{T}} B(t) y(t).$$

Now $B(t)y(t) = y(t)^{-1}$ implies $y(t)^{\mathrm{T}} B(t) y(t) = d$. Differentiating this with respect to t we obtain

$$0 = 2y'(t)^{\mathrm{T}} B(t) y(t) + y(t)^{\mathrm{T}} B'(t) y(t)$$

which implies that

$$2y'(t)^{\mathrm{T}} B(t) y(t) = -y(t)^{\mathrm{T}} (A - I) y(t).$$

Hence, using the AM–GM inequality

$$\frac{f'}{f} = -y^{\mathrm{T}}(A-I)y = -\frac{1}{1+t}y^{\mathrm{T}}(B-I)y = \frac{1}{1+t}\big(|y|^2 - y^{\mathrm{T}}By\big)$$
$$= \frac{1}{1+t}\big(|y|^2 - d\big) \geqslant \frac{d}{1+t}\big(f^{1/d} - 1\big). \tag{5}$$

Since $\mathcal{Y}(A) > 1$ so $f(0) > 1$; then using (5) one may show that $f(t)$ is a strictly increasing function on $[0, \tau)$. This is true in every quadrant so $\mathcal{Y}(B(t))$ is a strictly increasing function on $[0, \tau)$. Hence, using the continuity of $\mathcal{Y}$, we have $\mathcal{Y}(B(\tau)) > \mathcal{Y}(A)$. Note that $B(\tau)$ is singular by our definition of τ; of course $B(\tau) \geqslant 0$ and $\operatorname{diag}(B(\tau)) = I$.

4. Proof of Theorem 5

If y is a solution of $Ay = y^{-1}$ then $z = [\varepsilon_1 y_1, \ldots, \varepsilon_d y_d]$ is a solution of $Bz = z^{-1}$ where $B = (\varepsilon_i \varepsilon_j a_{ij})$. Hence there is no loss of generality in assuming that $\varepsilon = [1, \ldots, 1]$. So we are given that $a_{ij} \geqslant 0$ for all i, j and we have to show that $Ay = y^{-1}$ has a solution with $|y_i| \leqslant 1$ for all $i = 1, \ldots, d$.

Let Q be the quadrant in which all components are non-negative. Since all entries of A are non-negative and $\operatorname{diag}(A) = I$, the kernel of A intersects Q only at 0. Hence, from Proposition 3, there is a solution of $Ay = y^{-1}$ in Q, that is with $y_i > 0$ for all $i = 1, \ldots, d$—we continue to call this solution y. Then, since $a_{ii} = 1$ and $a_{ij} \geqslant 0$, we have

$$y_i^2 \leqslant \sum_{j=1}^{d} a_{ij} y_i y_j = 1, \quad i = 1, \ldots, d,$$

so $0 \leqslant y_i \leqslant 1$. Further, if $\mathcal{Y}(A) = 1$ then $y_i = 1$ for all i so $\sum_{j=1}^{d} a_{ij} = 1$ for all i, hence $d = \sum_{i,j=1}^{d} a_{ij}$. But $a_{ij} \geqslant 0$ and $\operatorname{diag}(A) = I$ so all off-diagonal entries of A must be zero.

5. Proof of Theorem 3

We will use the notation $a_k \sim b_k$ to mean that $\lim_{k\to\infty} a_k/b_k = 1$. Define the function $P(x)$ on $\mathbb{R}^n$ by $P(x) := \prod_{i=1}^{d} \langle x, u_i \rangle$. Then, as conjectured by Macdonald and proved by a combination of efforts of several authors (see [5, Eq. (1.20)]), for every positive integer γ, one has

$$\frac{1}{(2\pi)^{n/2}} \int\limits_{\mathbb{R}^n} e^{-|x|^2/2} \big|P(x)\big|^{2\gamma}\, dx = 2^{-d\gamma} \prod_{i=1}^{n} \frac{\Gamma(1 + d_i\gamma)}{\Gamma(1+\gamma)}. \tag{6}$$

Now, using polar coordinates,

$$\int\limits_{\mathbb{R}^n} e^{-|x|^2/2} \big|P(x)\big|^{2\gamma}\, dx = \int\limits_{\theta\in S^{n-1}} \big|P(\theta)\big|^{2\gamma}\, d\theta \int\limits_0^\infty e^{-t^2/2} t^{2d\gamma+n-1}\, dt$$
$$= 2^{d\gamma-1+n/2}\Gamma(d\gamma + n/2) \int\limits_{\theta\in S^{n-1}} \big|P(\theta)\big|^{2\gamma}\, d\theta.$$

Hence (6) implies

$$\int_{\theta\in S^{n-1}} \left|P(\theta)\right|^{2\gamma} d\theta = \frac{(\pi)^{n/2}2^{-2d\gamma+1}}{\Gamma(d\gamma+n/2)}\prod_{i=1}^{n}\frac{\Gamma(1+d_i\gamma)}{\Gamma(1+\gamma)},$$

so

$$\sup_{\theta\in S^{n-1}}\left|P(\theta)\right|^2 = \lim_{\gamma\to\infty}\left(\int_{\theta\in S^{n-1}}\left|P(\theta)\right|^{2\gamma} d\theta\right)^{1/\gamma}$$
$$= \frac{1}{2^{2d}}\lim_{\gamma\to\infty}\left(\frac{1}{(d\gamma-1+n/2)!}\prod_{i=1}^{n}\frac{(d_i\gamma)!}{\gamma!}\right)^{1/\gamma}. \tag{7}$$

From Stirling's formula, for large m one has $m! \sim \sqrt{2\pi}\, m^{m+1/2} e^{-m}$. Hence, for large γ,

$$\left(\frac{(d_i\gamma)!}{\gamma!}\right)^{1/\gamma} \sim \frac{(d_i\gamma)^{d_i+1/(2\gamma)}}{\gamma^{1+1/(2\gamma)}}\frac{e}{e^{d_i}} \sim e^{1-d_i} d_i^{d_i}\gamma^{d_i-1},$$
$$\left(\frac{1}{(d\gamma-1+n/2)!}\right)^{1/\gamma} \sim \frac{e^{d+(n-2)/(2\gamma)}}{(d\gamma+(n-2)/2)^{(d+(n-1)/(2\gamma))}} \sim \frac{e^d}{(d\gamma+(n-2)/2)^d} \sim \frac{e^d}{d^d\gamma^d}$$

because $\lim_{\gamma\to\infty}(d\gamma+a)^{b/\gamma}=1$ by L'Hopital's rule. Hence from (7) and that $\sum_{i=1}^n d_i = d+n$ (see [8, Section 3.9]) we have

$$\sup_{\theta\in S^{n-1}}\left|P(\theta)\right|^2 = \lim_{\gamma\to\infty}\frac{2^{-2d}e^d}{d^d\gamma^d}\prod_{i=1}^{n}e^{1-d_i}d_i^{d_i}\gamma^{d_i-1} = 2^{-2d}d^{-d}\prod_{i=1}^{n}d_i^{d_i}.$$

Noting that $d+n=\sum_{i=1}^n d_i$, we have $2^{-2d}\prod_{i=1}^n d_i^{d_i} = \prod_{i=1}^n(4(d_i/4)^{d_i})$ which exceeds 1 because $4(x/4)^x \geqslant 1$ if $x=1$ or $x\geqslant 2$; so $\mathcal{X}(u_1,\ldots,u_d)\geqslant 1$ for root systems.

6. Remarks

6.1. *Variation for linearly dependent $u_1,\ldots,u_d$*

With the use of Theorem 4 towards proving Conjecture 1 in mind, given linearly dependent unit vectors $u_1,\ldots,u_d$ in $\mathbb{R}^d$, in certain situations, we construct unit vectors $v_1,\ldots,v_d$ so that $\mathcal{X}(v_1,\ldots,v_d) < \mathcal{X}(u_1,\ldots,u_d)$. Equivalently, if $A=(\langle u_i,u_j\rangle)$ (which is singular), in certain situations, we construct a $B\geqslant 0$, $\operatorname{diag}(B)=I$ so that $\mathcal{Y}(B) > \mathcal{Y}(A)$. If we were able to do this for all linearly dependent $u_1,\ldots,u_d$, Conjectures 1 and 4 would follow from Theorem 4.

Choose v to be a unit vector orthogonal to $u_1,\ldots,u_d$, choose smooth functions $\theta_i(t)$ with $\theta_i(0)=0$, and define unit vectors

$$v_i(t) := \cos\theta_i(t)u_i + \sin\theta_i(t)v, \quad i=1,\ldots,d.$$

Let $A := (\langle u_i, u_j\rangle)$, $B(t) := (\langle v_i(t), v_j(t)\rangle)$; note that $B(0) = A$, $B(t) \geqslant 0$ and $\operatorname{diag}(B(t)) = I$. Note that $B = D(p)AD(p) + qq^{\mathrm{T}}$ where $D(p)$ denotes the diagonal matrix whose diagonal is the vector p and

$$p(t) := \big[\cos\theta_1(t), \ldots, \cos\theta_d(t)\big], \qquad q(t) := \big[\sin\theta_1(t), \ldots, \sin\theta_d(t)\big]. \tag{8}$$

Differentiating (8) we have

$$p' = -q \circ \theta', \qquad q' = p \circ \theta', \qquad p'' = -q' \circ \theta' - q \circ \theta'', \qquad q'' = p' \circ \theta' + p \circ \theta''.$$

Since $p(0) = e$ and $q(0) = 0$, we obtain

$$p'(0) = 0, \qquad q'(0) = \theta'(0), \qquad p''(0) = -\theta'(0)^2, \qquad q''(0) = \theta''(0). \tag{9}$$

Differentiating the relation $B = D(p)AD(p) + qq^{\mathrm{T}}$ we obtain

$$B' = D(p')AD(p) + D(p)AD(p') + q'q^{\mathrm{T}} + qq'^{\mathrm{T}},$$
$$B'' = D(p'')AD(p) + D(p)AD(p'') + 2D(p')AD(p') + q''q^{\mathrm{T}} + qq''^{\mathrm{T}} + 2q'q'^{\mathrm{T}}.$$

Hence $B'(0) = 0$ and $B''(0) = 2\theta'(0)\theta'(0)^{\mathrm{T}} - D(\theta'(0)^2)A - AD(\theta'(0)^2)$.

If Q is a quadrant with no solutions of $Ay = y^{-1}$ (but $B(t)y = y^{-1}$ may have a solution in Q), then from (d) in Proposition 3 we can find a $z \in Q$ with $\prod_{i=1}^d z_i^2 > \mathcal{Y}(A) + 1$ and $z^{\mathrm{T}}Az < d$. Hence, by continuity, $z^{\mathrm{T}}B(t)z \leqslant d$ for t near zero while $\prod_{i=1}^d z_i^2 > \mathcal{Y}(A) + 1$.

If Q is a quadrant containing a solution of $Ay = y^{-1}$ so that $\prod_{i=1}^d y_i^2 > \mathcal{Y}(A)$, then again continuity implies that $\prod_{i=1}^d y_i(t)^2 > \mathcal{Y}(A)$ for the solution $y(t)$ in Q of $B(t)y = y^{-1}$, for t near 0.

Finally, suppose Q is a quadrant containing a solution of $Ay = y^{-1}$ with $\prod_{i=1}^d y_i^2 = \mathcal{Y}(A)$—we call Q an optimal quadrant and y an optimal solution for A. From Proposition 6, for t near 0, $B(t)y = y^{-1}$ has a smoothly varying solution $y(t)$ in Q. Define $f(t) = \prod_{i=1}^d y_i(t)^2$; then as in the proof of Theorem 4,

$$\frac{f'(t)}{f(t)} = -y^{\mathrm{T}}(t)B'(t)y(t). \tag{10}$$

But $B'(0) = 0$ so $f'(0) = 0$. From (10) and that $f'(0) = 0$, $B'(0) = 0$ we have

$$\begin{aligned}\frac{f''(0)}{f(0)} &= -y^{\mathrm{T}}B''(0)y = y^{\mathrm{T}}\big(D\big(\theta'^2\big)A + AD\big(\theta'^2\big) - 2\theta'\theta'^{\mathrm{T}}\big)y \\ &= y^{\mathrm{T}}D\big(\theta'^2\big)Ay + y^{\mathrm{T}}AD\big(\theta'^2\big)y - 2\langle\theta', y\rangle^2 = 2y^{\mathrm{T}}D\big(\theta'^2\big)y^{-1} - 2\langle\theta', y\rangle^2 \\ &= 2\big(|\theta'|^2 - \langle\theta', y\rangle^2\big) = 2\big(1 - \langle\theta', y\rangle^2\big),\end{aligned} \tag{11}$$

where we have chosen that $|\theta'(0)| = 1$.

If we can find a unit vector $\theta'(0)$ so that $|\langle\theta'(0), y\rangle| < 1$ for all optimal y for A then $f''(0) > 0$; hence $f(t) > f(0)$ for all t near zero, in all optimal quadrants for A. Hence for t near 0 we have $\mathcal{Y}(B(t)) > \mathcal{Y}(A)$.

If the span of the set of optimal y is of dimension $d-1$ or less then choosing $\theta'(0)$ to be orthogonal to this span will work. For example, if two or more of the u_i are the same, say $u_1 = u_2$, then $y_1 = y_2$ for all solutions of $Ay = y^{-1}$ and hence $\theta'(0) = [1, -1, 0, \dots, 0]/\sqrt{2}$ will work.

Since the ellipsoid $y^{\mathrm{T}}Ay = d$ is unbounded in the quadrants containing the kernel of A, one conjectures that the optimal y would lean away from $\ker A$; hence one may consider taking $\theta'(0)$ to be an element of the kernel of A.

6.2. Equality in Conjecture 4

We show here that if $\mathcal{Y}(A) \leqslant 1$ for all $A \geqslant 0$ with $\mathrm{diag}(A) = I$ then $A > 0$ and $\mathcal{Y}(A) = 1$ implies $A = I$—we are not able to extend this to singular A.

Suppose $A > 0$, $\mathrm{diag}(A) = I$ and $\mathcal{Y}(A) = 1$, but $A \neq I$. We will use the construction and the calculations in the proof of Theorem 4. Suppose Q is an optimal quadrant for A and $y(t)$ the solution of $B(t)z = z^{-1}$ in Q, for $t \in [0, \tau)$. Now $f(0) = 1$; if $y(0)$ is not a ± 1 vector then from (5) we have $f'(0) > 0$ because we used the AM–GM inequality in (5). So if $y(0)$ is not a ± 1 vector then $f(t)$ is strictly increasing. Since $\mathcal{Y}(B(t)) \leqslant 1$, we conclude that there must be a ± 1 vector ε so that $A\varepsilon = \varepsilon^{-1} = \varepsilon$ and hence $B(t)\varepsilon = \varepsilon$. This argument also shows that $\mathcal{Y}(B(\tau)) = 1$ and the optimal solutions of $B(\tau)y = y^{-1}$ must be ± 1 vectors. Since $B(\tau)$ is singular, there is a unit vector z so that $B(\tau)z = 0$; then for any ± 1 vector ε with $B(\tau)\varepsilon = \varepsilon$, one has $z^{\mathrm{T}}\varepsilon = z^{\mathrm{T}}A\varepsilon = \varepsilon^{\mathrm{T}}Az = 0$. Hence z is orthogonal to the span of the optimal vectors of $B(\tau)$ and so from Section 6.1 we must have $\mathcal{Y}(B(\tau)) < 1$. This is a contradiction, hence $A = I$.

References

[1] V.A. Anagnostopoulos, S.G. Révész, Polarization constants for products of linear functionals over $\mathbb{R}^2$ and $\mathbb{C}^2$ and Chebyshev constants of the unit sphere, Publ. Math. Debrecen 68 (1–2) (2006) 63–75.

[2] J. Arias-de-Reyna, Gaussian variables, polynomials and permanents, Linear Algebra Appl. 285 (1–3) (1998) 107–114.

[3] K.M. Ball, The complex plank problem, Bull. London Math. Soc. 33 (4) (2001) 433–442.

[4] C. Benítez, Y. Sarantopoulos, A. Tonge, Lower bounds for norms of products of polynomials, Math. Proc. Cambridge Philos. Soc. 124 (3) (1998) 395–408.

[5] P.J. Forrester, S.O. Warnaar, The importance of the Selberg integral, arXiv: 0710.3981v1 [math.CA], 2007.

[6] P.E. Frenkel, Pfaffians, hafnians and products of real linear functionals, Math. Res. Lett. 15 (2) (2008) 351–358.

[7] J.C. García-Vázquez, R. Villa, Lower bounds for multilinear forms defined on Hilbert spaces, Mathematika 46 (1999) 315–322.

[8] J.E. Humphreys, Reflection Groups and Coxeter Groups, Cambridge Univ. Press, 1992.

[9] Y.J. Leung, W.V. Li, Rakesh, Optimal ellipsoids and decomposition of positive definite matrices, J. Math. Anal. Appl. 331 (2) (2007) 1452–1466.

[10] Y.J. Leung, W.V. Li, Rakesh, On solutions of $Ay = y^{-1}$ for positive semidefinite matrices A, preprint, 2007.

[11] A.E. Litvak, V.D. Milman, G. Schechtman, Averages of norms and quasi-norms, Math. Ann. 312 (1998) 95–124.

[12] M. Matolcsi, Linear polarization constant of $\mathbb{R}^n$, Acta Math. Hungar. (1–2) (2005) 129–136.

[13] M. Matolcsi, A geometric estimate on the norm of product of functionals, Linear Algebra Appl. 405 (2005) 304–310.

[14] M. Matolcsi, G. Muñoz, On the real polarization problem, Math. Inequal. Appl. 9 (3) (2006) 485–494.

[15] A. Pappas, S.G. Révész, Linear polarization constants of Hilbert spaces, J. Math. Anal. Appl. 300 (1) (2004) 129–146.

[16] S.G. Révész, Y. Sarantopoulos, Plank problems, polarization, and Chebyshev constants, J. Korean Math. Soc. 41 (1) (2004) 157–174.