PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 142, Number 8, August 2014, Pages 2581–2591
S 0002-9939(2014)11984-8
Article electronically published on May 13, 2014

QUASI-CONVEX FREE POLYNOMIALS

S. BALASUBRAMANIAN AND S. MCCULLOUGH

(Communicated by Richard Rochberg)

ABSTRACT. Let $\mathbb{R}\langle x\rangle$ denote the ring of polynomials in g freely noncommuting variables $x = (x_1, \dots, x_g)$. There is a natural involution $*$ on $\mathbb{R}\langle x\rangle$ determined by $x_j^* = x_j$ and $(pq)^* = q^*p^*$, and a free polynomial $p \in \mathbb{R}\langle x\rangle$ is symmetric if it is invariant under this involution. If $X = (X_1, \dots, X_g)$ is a g tuple of symmetric $n \times n$ matrices, then the evaluation $p(X)$ is naturally defined and further $p^*(X) = p(X)^*$. In particular, if p is symmetric, then $p(X)^* = p(X)$. The main result of this article says if p is symmetric, $p(0) = 0$ and for each n and each symmetric positive definite $n\times n$ matrix A the set $\{X : A-p(X) \succ 0\}$ is convex, then p has degree at most two and is itself convex, or $-p$ is a hermitian sum of squares.

1. Introduction

Let $\mathbb{R}\langle x\rangle$ denote the ring of polynomials over $\mathbb{R}$ in the freely noncommuting variables $x = (x_1, \dots, x_g)$. A $p \in \mathbb{R}\langle x\rangle$ is a **free polynomial** and is a finite sum

$$p = \sum p_w w,$$

over words w in x with coefficients $p_w \in \mathbb{R}$. The empty word, which plays the role of the multiplicative identity, will be denoted $\emptyset$.

For a word

$$w = x_{j_1} x_{j_2} \cdots x_{j_k}, \tag{1.1}$$

let

$$w^* = x_{j_k} \cdots x_{j_2} x_{j_1}.$$

The operation * extends naturally to an involution on $\mathbb{R}\langle x\rangle$ by

$$p^* = \sum p_w w^*.$$

Let $\mathbb{S}_n(\mathbb{R}^g)$ denote the set of g-tuples $X = (X_1, \dots, X_g)$ of $n \times n$ symmetric matrices. For a word w as in (1.1), substituting X_j for x_j gives

$$X^w = w(X) = X_{j_1} X_{j_2} \cdots X_{j_k}.$$

This evaluation extends to $\mathbb{R}\langle x\rangle$ in the obvious way,

$$p(X) = \sum p_w w(X).$$

Observe, for $0 \in \mathbb{S}_n(\mathbb{R}^g)$, that $p(0) = p_\emptyset I_n$, where I_n is the $n \times n$ identity.

Received by the editors March 6, 2012 and, in revised form, August 11, 2012.
2010 *Mathematics Subject Classification.* Primary 15A24, 47A63, 08B20.
Key words and phrases. Free polynomials, quasi-convex, free real algebraic geometry.
The research of the second author was supported by NSF grants DMS 0758306 and 1101137.

The transpose operation * on matrices is compatible with the involution * on $\mathbb{R}\langle x\rangle$ in that

$$p(X)^* = p^*(X).$$

A polynomial $p \in \mathbb{R}\langle x\rangle$ is symmetric if $p = p^*$, and in this case $p(X)^* = p^*(X) = p(X)$ so that p takes symmetric values.

Let $\mathbb{S}_n(\mathbb{R})$ denote the collection of symmetric $n \times n$ matrices. Given $S \in \mathbb{S}_n(\mathbb{R})$ the notation $S \succ 0$ and $S \succeq 0$ indicate that S is positive definite and positive semidefinite respectively. A symmetric $p \in \mathbb{R}\langle x\rangle$ is *quasi-convex* if $p(0) = 0$, and for each n and positive definite matrix $n \times n$ matrix A the set

$$\mathcal{D}(A) = \{X \in \mathbb{S}_n(\mathbb{R}^g) : A - p(X) \succ 0\}$$

is convex.

A symmetric polynomial p is a *(hermitian) sum of squares* if there exists an m and $h_1, \dots, h_m \in \mathbb{R}\langle x\rangle$ such that

$$p = \sum h_j^* h_j.$$

Evidently such a p is positive in the sense that for each n and $X \in \mathbb{S}_n(\mathbb{R}^g)$,

$$p(X) \succeq 0.$$

The following theorem is the main result of this article.

Theorem 1.1. *If p is quasi-convex, then either $-p$ is a sum of squares, or there exist a linear polynomial $\ell \in \mathbb{R}\langle x\rangle$ and finitely many linear polynomials $s_j \in \mathbb{R}\langle x\rangle$ such that*

$$p(x) = \ell(x) + \sum s_j^*(x) s_j(x). \tag{1.2}$$

Thus, p is a hermitian sum of squares of linear polynomials plus a linear term.

Further, if there is an N such that for each $n \geq N$ there is a $B \in \mathbb{S}_n(\mathbb{R})$ such that $B \not\succeq 0$ and $\{X \in \mathbb{S}_n(\mathbb{R}^g) : B - p(X) \succ 0\}$ is convex, then $-p$ is a sum of squares if and only if $p = 0$.

Remark 1.2. It is easy to see that if p is a hermitian sum of squares of linear polynomials plus a linear term, then p is quasi-convex.

1.1. **Related results and remarks.**

Theorem 1.1 falls within the emerging fields of free analysis and free semialgebraic geometry. Free semialgebraic geometry is, by analogy to the commutative case, the study of free polynomial inequalities. For instance there are now a number of free Positivstellensatze for which [2], [15], [16] and [11] are just a few references. In this regard, see also [22] and its Proposition 17. There is also a theory of free rational functions. Recent developments in this direction have been related to noncommutative multi-variate systems theory. See for instance [1]. Free rational functions actually appeared much earlier in the context of finite automata. See for instance [20]. Issues of convexity in the context of free polynomials and rational functions naturally arise in systems theory problems [12, 14] and mathematically are related to the theory of operator spaces and systems and matrix convexity [6]. More generally, there is a theory of free analytic functions which arise naturally in several contexts, including free probability. A sampling of references includes [23], [13], [19] and [18].

Some systems theory problems present as a free polynomial inequality (or more realistically as a system of matrix-valued free polynomial inequalities) involving two classes of free (freely noncommuting) variables, say the a variables and the x variables. The a variables are thought of as known (system or state) parameters and the x variables as unknowns. For a given free polynomial q, of interest is the case that, for each fixed A in some distinguished collection of known parameters, the inequality $q(A,x) \succ 0$ is convex in x. Thus this article considers the simplest such case. Namely, there is just one a variable and $q(a,x) = a - p(x)$ for a polynomial p in the variables x alone. For comparison, a main result of [8] says, generally, if $q(A,x)$ is convex in x for each fixed A, then $q(a,x) = L(a,x) + \sum h_j(a,x)^* h_j(a,x)$ where L has degree at most one in x and each h_j is linear in x. The articles [3–5, 9] contain results for polynomials f whose positivity set, namely the set of those X such that $f(X) \succ 0$, is convex.

A symmetric polynomial p is **matrix convex** if for each n, each pair $X, Y \in \mathbb{S}_n(\mathbb{R}^g)$, and each $0 \le t \le 1$,

$$p(tX + (1-t)Y) \preceq tp(X) + (1-t)p(Y).$$

The following theorem, pointed out by the referee, generalizes the main result of [10].

Theorem 1.3. *For a symmetric polynomial p the following are equivalent:*

(i) $p - p(0)$ *has the form in* (1.2);
(ii) p *is matrix convex;*
(iii) $\mathcal{D}(A)$ *is convex for every* $A \in \mathbb{S}_n(\mathbb{R}^g)$;
(iv) $p(x) - p(0)$ *is quasi-convex and* $p(0) - p(x)$ *is not a nonzero sum of squares.*

A remark is in order about the definition of quasi-convex used here. A classical definition says that a function f of several (commuting) variables is quasi-convex if each of its sub-level sets is convex. (The interested reader can work out the relationships between this definition of quasi-convex and the seemingly more popular one $f(tx + (1-t)y) \le \max\{f(x), f(y)\}$ for $0 \le t \le 1$.) In considering a free analog here, because of the role that positivity plays in the arguments, it was convenient to make the harmless normalization that $p(0) = 0$ and then only require convexity of the (level) sets $\mathcal{D}(A)$ for A positive definite in the definition of quasi-convex.

1.2. **Reader's guide.** The remainder of this paper is organized as follows. Issues surrounding sums of squares are dealt with in Section 2. Just as in the commutative case, convexity is related to positivity of a Hessian. The necessary definitions and basic results appear in Section 3. Section 4 examines membership in the boundary of the set $\mathcal{D}(A)$ as well as consequences of the convexity hypothesis. Theorem 1.1 is proved in the final section, Section 5.

2. The sum of squares case

The following proposition dispenses with the alternative that $-p$ is a sum of squares.

Proposition 2.1. *If there is an N such that for each $n \ge N$ there exists a $B \in \mathbb{S}_n(\mathbb{R})$ such that $B \not\preceq 0$ and the set $\mathcal{D}(B)$ is convex, then $-p$ is not a nonzero sum of squares.*

Proof. Arguing by contradiction, suppose $-p$ is a sum of squares. Consider the polynomial

$$-p(tx) = \sum_{j=1}^{2d} p_j(x)t^j,$$

where the p_j are homogeneous of degree j polynomials in the free variables x. Since $-p$ is a sum of squares, $-p(tx)$ has even degree as a polynomial in t, and so, without loss of generality, we may assume that p_{2d} is nonzero. Now p_{2d} is itself a sum of squares and hence it takes positive semi-definite values. If, on the other hand, $p_{2d}(X)$ is never positive definite, then $\det(p_{2d}(X)) = 0$ for all n and $X \in \mathbb{S}_n(\mathbb{R}^g)$. An application of the Guralnick-Small lemma as found in [8] then gives the contradiction that p_{2d} is the zero polynomial. Thus there is an n and an $X \in \mathbb{S}_n(\mathbb{R}^g)$ such that $p_{2d}(X) \succ 0$. By assumption, there is a $B \in \mathbb{S}_n(\mathbb{R})$ such that $B \not\succeq 0$ and $\mathcal{D}(B)$ is convex. Choosing t sufficiently large, it may be assumed that both $B - p(tX)$ and $B - p(-tX)$ are positive definite. In this case tX and $-tX$ are in $\mathcal{D}(B)$ and thus $0 = \frac{1}{2}(tX + (-tX)) \in \mathcal{D}(B)$, contradicting the assumption that B is not positive semidefinite and completing the proof. □

2.1. **When $-p$ is not a sum of squares.** Given n, let $\mathcal{K}(n)$ denote the set of those $X \in \mathbb{S}_n(\mathbb{R}^g)$ such that $p(X)$ has a positive eigenvalue. The $\mathcal{K}(n)$ are open sets, the issue is whether they are empty or not. In the free setting, and unlike in the case of several commuting variables, positive polynomials are sums of squares with [21], [22] and [7] as a very small sampling of the references. (For a reference which explicitly treats the case of the symmetric variables ($x_j^* = x_j$) used here, see [17].) In particular, each $\mathcal{K}(n)$ is empty if and only if $-p$ is a sum of squares.

The conclusion of the following lemma will be used later in the proof of Theorem 1.1 when $-p$ is not a sum of squares.

Lemma 2.2. *Suppose $q \in \mathbb{R}\langle x\rangle$ and $-p$ is not a sum of squares. If* $\det(q(X)) = 0$ *for every n and $X \in \mathcal{K}(n)$, then $q = 0$.*

Before beginning the proof of Lemma 2.2 we record the following simple fact.

Lemma 2.3. *Suppose $q \in \mathbb{R}\langle x\rangle$ and m is a positive integer. If $\mathcal{K}(m)$ is nonempty and* $\det(q(X)) = 0$ *for all $X \in \mathcal{K}(m)$, then* $\det(q(X)) = 0$ *for all* $X \in \mathbb{S}_m(\mathbb{R}^g)$.

Proof. The function $\mathbb{S}_m(\mathbb{R}^g) \ni X \mapsto \det(q(X))$ is a polynomial in the entries of X. Hence, if it vanishes on an open set it must be identically zero. □

Given integers k, ℓ, let $m = k + \ell$ and consider the subspace

$$S = \mathbb{S}_k(\mathbb{R}^g) \oplus \mathbb{S}_\ell(\mathbb{R}^g)$$

of $\mathbb{S}_m(\mathbb{R}^g)$. Each tuple $X \in S$ is a direct sum $X = (Y_1 \oplus Z_1, \dots, Y_g \oplus Z_g)$, where $Y = (Y_1, \dots, Y_g) \in \mathbb{S}_k(\mathbb{R}^g)$ and $Z = (Z_1, \dots, Z_g) \in \mathbb{S}_\ell(\mathbb{R}^g)$ and where

$$Y_j \oplus Z_j = \begin{pmatrix} Y_j & 0 \\ 0 & Z_j \end{pmatrix}.$$

Proof of Lemma 2.2. Since $-p$ is not a sum of squares, by the remarks at the outset of this section, there is an m and a $Y \in \mathcal{K}(m)$ such that $p(Y)$ has a positive eigenvalue. First observe that for any positive integer k and $X \in \mathbb{S}_k(\mathbb{R}^g)$ that $p(X \oplus Y) = p(X) \oplus p(Y)$ too has a positive eigenvalue. Thus, $\mathcal{K}(n)$ is nonempty for each $n \geq m$. Lemma 2.3 now implies that $\det(q(X)) = 0$ for each $n \geq m$ and $X \in \mathbb{S}_n(\mathbb{R}^g)$.

Now suppose $1 \leq \ell < m$. Since $n = m\ell \geq m$, $\det(q(X)) = 0$ for all $X \in \mathbb{S}_n(\mathbb{R}^g)$. Given $X \in \mathbb{S}_\ell(\mathbb{R}^g)$, the tuple

$$\tilde{X} = \bigoplus_1^m X \in \mathbb{S}_n(\mathbb{R}^g).$$

Thus $(\det(q(X)))^m = \det(q(\tilde{X})) = 0$. Hence $\det(q(X)) = 0$ for every $X \in \mathbb{S}_\ell(\mathbb{R}^g)$.

Since $\det(q(X)) = 0$ for every n and $X \in \mathbb{S}_n(\mathbb{R}^g)$, the Guralnick-Small lemma in [8] implies q is the zero polynomial. □

3. Directional derivatives and the Hessian

Given $p \in \mathbb{R}\langle x\rangle$, another set of freely noncommuting variables $h = (h_1, \ldots, h_g)$ and the real parameter t,

$$p(x + th) = \sum p_j(x)[h]t^j,$$

where $p_j(x)[h]$ are polynomials in the variables $(x, h) = (x_1, \ldots, x_g, h_1, \ldots, h_g)$ (which are of course freely noncommuting). The notation indicates the different role that these variables play. Indeed, observe that $p_j(x)[h]$ is homogeneous of degree j in h.

The polynomial $p_1(x)[h]$ is the *directional derivative* or simply the *derivative* of p (in the direction h) and is denoted $p'(x)[h]$. The polynomial $2p_2(x)[h]$ is the *Hessian* of p and is denoted by $p''(x)[h]$.

Given $X \in \mathbb{S}_n(\mathbb{R}^g)$ and $v \in \mathbb{R}^n$, let

$$\mathcal{T}(X, v) = \{H \in \mathbb{S}_n(\mathbb{R}^g) : p'(X)[H]v = 0\} \subset \mathbb{S}_n(\mathbb{R}^g).$$

In the case that $(A - p(X))v = 0$, the subspace $\mathcal{T}(X, v)$ is the *clamped tangent plane* to $\mathcal{D}(A)$ at (X, v) [3]. In this case, if one chooses $H \in \mathcal{T}(X, v)$, then

$$\langle (A - p(X + tH))v, v\rangle = -\frac{1}{2}t^2\langle p''(X)[H]v, v\rangle + t^3 e(t),$$

for some polynomial $e(t)$. This identity, much as in the commutative case, provides a link between convexity and positivity of the Hessian of p.

4. The boundaries

Fix p satisfying the hypothesis of Theorem 1.1. In particular, $p(0) = 0$.

Lemma 4.1. *Let n and a positive definite $A \in \mathbb{S}_n(\mathbb{R})$ be given. A given $X \in \mathbb{S}_n(\mathbb{R}^g)$ is in the boundary of $\mathcal{D}(A)$ if and only if $A - p(X)$ is positive semidefinite and has a nontrivial kernel.*

Proof. Suppose that X is in the boundary of $\mathcal{D}(A)$. It follows that $A - p(X) \succeq 0$. It must be the case that $A - p(X)$ has a nontrivial kernel, because otherwise, $X \in \mathcal{D}(A)$, which is an open set.

To prove the converse, suppose $A - p(X)$ is positive semidefinite and has a nontrivial kernel. Clearly, $X \notin \mathcal{D}(A)$. For positive integers n, let $A_n = \left(\frac{n+1}{n}\right) A$. Then $A_n - p(X)$ is positive definite. Hence $X \in \mathcal{D}(A_n)$ and by convexity of $\mathcal{D}(A_n)$, for a fixed $0 < s < 1$, $A_n - p(sX) \succ 0$. Letting n tend to infinity, it follows that $A - p(sX) \succeq 0$.

Consider the function $\psi : \mathbb{R} \to \mathbb{R}$ defined by $\psi(t) = \det(A - p(tX))$. From what has already been proved, $\psi(t) \geq 0$ for $0 \leq t < 1$. Since $\psi(t)$ is a polynomial in the variable t, either it vanishes everywhere on $\mathbb{R}$ or only on a finite subset of $\mathbb{R}$. If $\psi(t)$ vanishes everywhere, then $\psi(0) = \det(A) = 0$, which contradicts the positive definiteness of A. Thus $\psi(t) > 0$ except for finitely many points in $(0, 1)$, and thus there is a sequence (s_n) from $(0, 1)$ such that each $s_n X \in \mathcal{D}(A)$ and $s_n X \to X$. Hence X is in the boundary of $\mathcal{D}(A)$. ☐

Suppose that $X \in \mathbb{S}_n(\mathbb{R}^g)$ is in the boundary of $\mathcal{D}(A)$ and $v \neq 0$ is a vector in $\mathbb{R}^n$ such that $Av = p(X)v$.

Proposition 4.2. *With X and v as above, if the dimension of the kernel of $A - p(X)$ is one, then there exists a subspace $\mathcal{H}$ of $\mathcal{T}(X, v)$ of codimension one (in $\mathcal{T}(X, v)$) such that, for $H \in \mathcal{H}$,*

$$\langle p''(X)[H]v, v\rangle \geq 0. \tag{4.1}$$

Remark 4.3. Since $\mathcal{T}(X, v)$ has codimension at most n in $\mathbb{S}_n(\mathbb{R}^g)$, it turns out that $\mathcal{H}$ will have codimension at most $n + 1$ in $\mathbb{S}_n(\mathbb{R}^g)$. In fact, a slight modification of the proof below shows that there is a subspace $\mathcal{K}$ of $\mathbb{S}_n(\mathbb{R}^g)$ of codimension at most n for which equation (4.1) holds. The key point is, with Λ as in the proof of the proposition, if $\Lambda(H) = 0$, then $\langle p'(X)[H]v, v\rangle = 0$.

Unlike a related argument in [3], the proof here does not rely on choosing a curve lying in the boundary of a convex set, thus eliminating the need for a corresponding smoothness hypothesis.

Proof. Since X is in the boundary of the convex set $\mathcal{D}(A)$, there is a linear functional $\Lambda : \mathbb{S}_n(\mathbb{R}^g) \to \mathbb{R}$ such that $\Lambda(Z) < 1$ for $Z \in \mathcal{D}(A)$ and $\Lambda(X) = 1$. The subspace

$$\mathcal{H} = \{H \in \mathcal{T}(X, v) : \Lambda(H) = 0\}$$

has codimension one in $\mathcal{T}(X, v)$.

Fix $H \in \mathcal{H}$ and define $F : \mathbb{R} \to \mathbb{S}_n(\mathbb{R})$ by $F(t) = A - p(X + tH)$. Thus, F is a matrix-valued polynomial in the real variable t. Let $[v]$ denote the one dimensional subspace of $\mathbb{R}^n$ spanned by the vector v. Write $F(t)$, with respect to the orthogonal decomposition of $\mathbb{R}^n$ as $[v]^\perp \oplus [v]$, as

$$F(t) = \begin{pmatrix} Q(t) & g(t) \\ g(t)^* & f(t) \end{pmatrix},$$

where Q is a square matrix-valued polynomial, g is a vector, and f is a scalar-valued polynomial. The assumption $(A - p(X))v = 0$ implies that f and g vanish at 0. The further assumption that $H \in \mathcal{T}(X, v)$ implies that f and g actually vanish to second order at 0. In particular, there are polynomials β and γ such that $g(t) = t^2\beta(t)$ and $f(t) = t^2\gamma(t)$.

Observe that

$$\gamma(0) = -\langle p''(X)[H]v, v\rangle.$$

Thus, to complete the proof of the theorem it suffices to use the choice of Λ (and thus the convexity hypothesis on $\mathcal{D}(A)$) and the assumption on the dimension of the kernel of $A - p(X)$ to show that $\gamma(0) \leq 0$. Indeed, since the kernel of $A - p(X)$ has dimension one, it follows that $Q(0) \succ 0$. Therefore, there exists an $\epsilon > 0$ such that if $|t| < \epsilon$, then $Q(t) \succ 0$. On the other hand, $\Lambda(X + tH) = \Lambda(X) = 1$ for all t.

Thus $X + tH \notin \mathcal{D}(A)$, which means $F(t) = A - p(X + tH) \not\succeq 0$. Hence, the Schur complement of F is nonpositive; i.e.,

$$t^2[\gamma(t) - t^2\beta^*(t)Q^{-1}(t)\beta(t)] \leq 0.$$

It follows that, for $|t| < \epsilon$,

$$\gamma(t) \leq t^2\beta^*(t)Q^{-1}(t)\beta(t),$$

and hence $\gamma(0) \leq 0$. $\square$

We end this section with the following simple observation.

Lemma 4.4. *Suppose $X \in \mathbb{S}_n(\mathbb{R}^g)$ and $0 \neq v \in \mathbb{R}^n$. If there is a $\lambda > 0$ such that $p(X)v = \lambda v$, then there exists a positive definite $A \in \mathbb{S}_n(\mathbb{R})$ such that X is in the boundary of $\mathcal{D}(A)$ and v spans the kernel of $A - p(X)$. Hence, for the triple (A, X, v) the conclusion of Proposition* 4.2 *holds.*

Further, if, for a given positive definite $A \in \mathbb{S}_n(\mathbb{R})$, X is in the boundary of $\mathcal{D}(A)$, and v is a nonzero vector such that $(A - p(X))v = 0$, then for each $\epsilon > 0$ there is a $A_\epsilon > 0$ such that $\|A - A_\epsilon\| < \epsilon$, X is in the boundary of $\mathcal{D}(A_\epsilon)$ and the kernel of $(A_\epsilon - p(X))$ is spanned by v.

Proof. With respect to the decomposition of $\mathbb{R}^n$ as $[v] \oplus [v]^\perp$,

$$p(X) = \begin{pmatrix} \lambda & 0 \\ 0 & T \end{pmatrix}$$

for some symmetric matrix T. Choose $\mu > 0$ so that $\mu - T \succ 0$ and let

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

In particular,

$$A - p(X) = \begin{pmatrix} 0 & 0 \\ 0 & \mu - T \end{pmatrix}$$

is positive semidefinite with a one dimensional kernel spanned by v. From Lemma 4.1, X is in the boundary of $\mathcal{D}(A)$.

As for the further statement, diagonalize with respect to same orthogonal decomposition of $\mathbb{R}^n$ as above,

$$A - p(X) = \begin{pmatrix} 0 & 0 \\ 0 & T \end{pmatrix},$$

for some positive semidefinite T. Let P denote the projection onto $[v]^\perp$ and let $A_\epsilon = A + \epsilon P$. Then

$$A_\epsilon - p(X) = \begin{pmatrix} 0 & 0 \\ 0 & \epsilon + T \end{pmatrix}$$

and the result follows. $\square$

5. Direct sums and linear independence

As in Section 2, let $\mathcal{K}(n)$ denote the set of those $X \in \mathbb{S}_n(\mathbb{R}^g)$ such that $p(X)$ has a positive eigenvalue. From here on, it is assumed that $-p$ is not a sum of squares. Equivalently, $\mathcal{K}(m)$ is not empty for some m.

Let $\hat{\mathcal{K}}(n)$ denote the set of pairs (X, v) such that $X \in \mathcal{K}(n)$ and v is an eigenvector of $p(X)$ corresponding to a positive eigenvalue. By Lemma 4.4, if $(X, v) \in \hat{\mathcal{K}}(n)$, then there exists a positive definite $A \in \mathbb{S}_n(\mathbb{R})$ such that X is in the boundary of

$\mathcal{D}(A)$ and the kernel of $A - p(X)$ is spanned by v. Let $\langle x \rangle_k$ denote the set of words of length at most k.

Lemma 5.1. *Fix a positive integer k. Given $X \in \mathbb{S}_n(\mathbb{R}^g)$ and $v \in \mathbb{R}^n$, there is a nonzero $q \in \mathbb{R}\langle x \rangle$ of degree at most k such that $q(X)v = 0$ if and only if the set $\{w(X)v : w \in \langle x \rangle_k\}$ is linearly dependent.*

If $q \in \mathbb{R}\langle x \rangle$ and $q(X)v = 0$ for all n and $(X, v) \in \hat{\mathcal{K}}(n)$, then $q = 0$.

Proof. The first statement is evident. As for the second, the hypotheses imply that $\det(q(X)) = 0$ for each n and $X \in \mathcal{K}(n)$. Hence by Lemma 2.2, $q = 0$. $\square$

Lemma 5.2. *Let d denote the degree of p. Given a positive integer N, there exists an $n \geq N$ and a pair (X, v) with $X \in \mathbb{S}_n(\mathbb{R}^g)$ and $v \neq 0$ in $\mathbb{R}^n$ such that:*

(i) there is a subspace $\mathcal{H}$ of $\mathcal{T}(X, v)$ of codimension at most one such that, for all $H \in \mathcal{H}$,

$$\langle p''(X)[H]v, v \rangle \geq 0;$$

(ii) if q is of degree at most $d - 1$ such that $q(X)v = 0$, then $q = 0$.

Proof. Let $\mathcal{P}$ denote the vector space of polynomials in g variables of degree at most $d - 1$. Given $(Y, w) \in \hat{\mathcal{K}}(n)$, let

$$Q(Y, w) = \{q \in \mathcal{P} : q(Y)w = 0\}.$$

Thus, $Q(Y, w)$ is a subspace of the finite dimensional vector space $\mathcal{P}$. Further, by Lemma 5.1,

$$\bigcap \{Q(Y, w) : (Y, w) \in \hat{\mathcal{K}}(n), \quad n \in \mathbb{N}\} = \{0\}.$$

Because of finite dimensionality, there are positive integers t and $n_1, \ldots, n_t$ and $(Y^j, w^j) \in \hat{\mathcal{K}}(n_j)$ such that

$$\bigcap_{j=1}^{t} Q(Y^j, w^j) = \{0\}.$$

In particular, if $q \in \mathbb{R}\langle x \rangle$ has degree at most $d - 1$ and $q(Y^j)w^j = 0$ for $j = 1, \ldots, t$, then $q = 0$.

Let $Z = \bigoplus Y^j$ and $z = \bigoplus w^j$. Thus Z acts on a space of dimension $n' = \sum n_j$. Choose a positive integer k such that $n = kn' \geq N$ and let $X = \bigoplus_1^k Z$ and $v = \bigoplus_1^k z$. From the definition of $\hat{\mathcal{K}}(n)$ and by Lemma 4.4, for each j there is a positive definite $A_j \in \mathbb{S}_{n_j}(\mathbb{R})$ such that Y^j is in the boundary of $\mathcal{D}(A_j)$ and $(A_j - p(Y^j))w^j = 0$. Let $B = \bigoplus A_j$ and $A' = \bigoplus_1^k B$. Then $(A' - p(X))v = 0$ and $A' - p(X) \succeq 0$. Moreover, if q has degree at most $d - 1$ and $q(X)v = 0$, then $q = 0$.

Finally, choose a positive definite $A \in \mathbb{S}_n(\mathbb{R})$ by the second part of Lemma 4.4 such that X is in the boundary of $\mathcal{D}(A)$ and the kernel of $(A - p(X))$ is spanned by v. In particular, X is in the boundary of $\mathcal{D}(A)$. The triple (A, X, v) satisfies the hypotheses of Proposition 4.2. Hence there is a subspace $\mathcal{H}$ of $\mathcal{T}(X, v)$ of codimension at most one such that

$$\langle p''(X)[H]v, v \rangle \geq 0$$

for all $H \in \mathcal{H}$. $\square$

The symmetric polynomial

$$r(x)[h] = p''(x)[h]$$

in the $2g$ variables $(x_1, \dots, x_g, h_1, \dots, h_g)$ is homogeneous of degree two in h. It admits a representation of the form

$$r(x)[h] = \begin{bmatrix} V_0(x)[h]^T \cdots V_{d-2}(x)[h]^T \end{bmatrix} Z(x) \begin{bmatrix} V_0(x)[h] \\ \vdots \\ V_{d-2}(x)[h] \end{bmatrix},$$

where $Z(x)$ is a (uniquely determined square symmetric) matrix of free polynomials and $V_j(x)[h]$ is the vector with entries $h_\ell w$ over free words w of the variables $x_1, \dots, x_g$ of length j and $1 \le \ell \le g$. (For details see [4].) The matrix $\mathcal{Z} = Z(0)$ is the **middle matrix**.

Lemma 5.3. *If $\mathcal{Z}$ is positive semidefinite, then p has degree at most two and, moreover, p has the form in* (1.2)*.*

A proof can be found in [10]. The idea is that the middle matrix $\mathcal{Z}$ has an *antidiagonal* structure which implies, if it is positive semidefinite, that only its nonzero entries correspond to $V_0[h]$, which is linear in h and independent of x. Thus,

$$r(x)[h] = r[h] = V_0[h]^T \mathcal{Z} V_0[h]$$

and it can be shown that $\mathcal{Z}$ must be positive semidefinite. Writing $\mathcal{Z}$ as a sum of squares and using

$$p(x) = \ell(x) + \frac{1}{2} r(x)[x]$$

expresses p in the form of (1.2).

The following lemma is a consequence of Lemma 7.2 from [4].

Lemma 5.4. *There is an integer ν depending only upon the degree d of the polynomial p and the number g of variables such that the following holds. If*

(i) n *satisfies* $\frac{\nu+1}{n} < 1$,
(ii) $X \in \mathbb{S}_n(\mathbb{R}^g)$ *and* $v \in \mathbb{R}^n$, *and*
(iii) there exists a subspace $\mathcal{H}$ of $\mathcal{T}(X, v)$ of codimension at most one such that for each $H \in \mathcal{H}$ (4.1) *holds, and*
(iv) there does not exist a nonzero polynomial q of degree at most $d-1$ satisfying $q(X)v = 0$,

then $\mathcal{Z}$ is positive semidefinite.

To prove Theorem 1.1 simply observe that the existence of an n that satisfies conditions (i) - (iv) of Lemma 5.4 is guaranteed by Lemma 5.2. The conclusion $\mathcal{Z}$ is positive semidefinite, combined with Lemma 5.3, now completes the proof.

It remains to prove Theorem 1.3. The equivalence of conditions (i) and (ii) is the main result of [10]. That (ii) implies (iii) is easily checked. If (iii) holds, then, by definition, $p(x) - p(0)$ is quasi-convex. Moreover, by the second part of Theorem 1.1, condition (iii) implies $p(0) - p(x)$ is not a nonzero sum of squares. Hence (iii) implies (iv). If (iv) holds, then Theorem 1.1 implies $p(x) - p(0)$ is a linear term plus a (hermitian) sum of squares of linear polynomials, and thus (i) holds.

Acknowledgment

The authors thank the referee for many corrections, suggestions and references related to this article.

References

[1] Joseph A. Ball, Gilbert Groenewald, and Tanit Malakorn, *Structured noncommutative multidimensional linear systems*, SIAM J. Control Optim. **44** (2005), no. 4, 1474–1528, DOI 10.1137/S0363012904443750. MR2178040 (2006j:93035)

[2] Jakob Cimprič, *Noncommutative Positivstellensätze for pairs representation-vector*, Positivity **15** (2011), no. 3, 481–495, DOI 10.1007/s11117-010-0098-0. MR2832601 (2012k:16099)

[3] Harry Dym, William Helton, and Scott McCullough, *Irreducible noncommutative defining polynomials for convex sets have degree four or less*, Indiana Univ. Math. J. **56** (2007), no. 3, 1189–1231, DOI 10.1512/iumj.2007.56.2904. MR2333470 (2008d:47033)

[4] Harry Dym, J. William Helton, Scott McCullough, *Non-commutative Varieties with Curvature having Bounded Signature*. Illinois J. Math. **55** (2011), no. 2, 427–464. MR3020690

[5] Harry Dym, J. William Helton, and Scott McCullough, *The Hessian of a noncommutative polynomial has numerous negative eigenvalues*, J. Anal. Math. **102** (2007), 29–76, DOI 10.1007/s11854-007-0016-y. MR2346553 (2008k:47010)

[6] Edward G. Effros and Soren Winkler, *Matrix convexity: operator analogues of the bipolar and Hahn-Banach theorems*, J. Funct. Anal. **144** (1997), no. 1, 117–152, DOI 10.1006/jfan.1996.2958. MR1430718 (98e:46066)

[7] J. William Helton, *"Positive" noncommutative polynomials are sums of squares*, Ann. of Math. (2) **156** (2002), no. 2, 675–694, DOI 10.2307/3597203. MR1933721 (2003k:12002)

[8] Damon M. Hay, J. William Helton, Adrian Lim, and Scott McCullough, *Non-commutative partial matrix convexity*, Indiana Univ. Math. J. **57** (2008), no. 6, 2815–2842, DOI 10.1512/iumj.2008.57.3638. MR2483003 (2010b:47036)

[9] J. William Helton and Scott McCullough, *Every convex free basic semi-algebraic set has an LMI representation*, Ann. of Math. (2) **176** (2012), no. 2, 979–1013. MR2950768

[10] J. William Helton and Scott McCullough, *Convex noncommutative polynomials have degree two or less*, SIAM J. Matrix Anal. Appl. **25** (2004), no. 4, 1124–1139, DOI 10.1137/S0895479803421999. MR2081134 (2005e:47043)

[11] J. William Helton, Scott McCullough, and Mihai Putinar, *Strong majorization in a free $*$-algebra*, Math. Z. **255** (2007), no. 3, 579–596, DOI 10.1007/s00209-006-0032-0. MR2270289 (2008k:47011)

[12] J. William Helton, Scott McCullough, Mihai Putinar, and Victor Vinnikov, *Convex matrix inequalities versus linear matrix inequalities*, IEEE Trans. Automat. Control **54** (2009), no. 5, 952–964, DOI 10.1109/TAC.2009.2017087. MR2518108 (2010f:52013)

[13] Dmitry Kaliuzhnyi-Verbovetskyi and Victor Vinnikov, work in progress.

[14] Mauricio C. de Oliveira, J. William Helton, Scott A. McCullough, and Mihai Putinar, *Engineering systems and free semi-algebraic geometry*, Emerging applications of algebraic geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 17–61, DOI 10.1007/978-0-387-09686-5_2. MR2500463 (2010d:93045)

[15] Igor Klep and Markus Schweighofer, *A nichtnegativstellensatz for polynomials in noncommuting variables*, Israel J. Math. **161** (2007), 17–27, DOI 10.1007/s11856-007-0070-2. MR2350154 (2008j:47063)

[16] Igor Klep and Markus Schweighofer, *Sums of Hermitian squares and the BMV conjecture*, J. Stat. Phys. **133** (2008), no. 4, 739–760, DOI 10.1007/s10955-008-9632-x. MR2456943 (2009j:82003)

[17] Scott McCullough, *Factorization of operator-valued polynomials in several non-commuting variables*, Linear Algebra Appl. **326** (2001), no. 1-3, 193–203, DOI 10.1016/S0024-3795(00)00285-8. MR1815959 (2002f:47035)

[18] Paul S. Muhly and Baruch Solel, *Schur class operator functions and automorphisms of Hardy algebras*, Doc. Math. **13** (2008), 365–411. MR2520475 (2010g:46098)

[19] Gelu Popescu, *Free holomorphic functions on the unit ball of $B(\mathcal{H})^n$. II*, J. Funct. Anal. **258** (2010), no. 5, 1513–1578, DOI 10.1016/j.jfa.2009.10.014. MR2566311 (2011b:47020)

[20] M. P. Schützenberger, *On the definition of a family of automata*, Information and Control **4** (1961), 245–270. MR0135680 (24 #B1725)
[21] Konrad Schmüdgen, *Unbounded operator algebras and representation theory*, Operator Theory: Advances and Applications, vol. 37, Birkhäuser Verlag, Basel, 1990. MR1056697 (91f:47062)
[22] Konrad Schmüdgen, *Noncommutative real algebraic geometry—some basic concepts and first ideas*, Emerging applications of algebraic geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 325–350, DOI 10.1007/978-0-387-09686-5_9. MR2500470 (2010m:13036)
[23] Dan-Virgil Voiculescu, *Free analysis questions II: the Grassmannian completion and the series expansions at the origin*, J. Reine Angew. Math. **645** (2010), 155–236, DOI 10.1515/CRELLE.2010.063. MR2673426 (2012b:46144)

Department of Mathematics and Statistics, Indian Institute of Science Education and Research (IISER) – Kolkata, Mohanpur Campus, Nadia District, Pin: 741246, West Bengal, India
E-mail address: bsriram@iiserkol.ac.in

Department of Mathematics, The University of Florida, Box 118105, Gainesville, Florida 32611-8105
E-mail address: sam@ufl.edu