A Positivstellensatz for Non-commutative Polynomials

J.W. Helton* Scott A. McCullough †

July 11, 2003

Abstract

A non-commutative polynomial which is positive on a bounded semi-algebraic set of operators has a weighted sum of squares representation. This Positivstellensatz parallels similar results in the commutative case.

A broader issue is to what extent does real semi-algebraic geometry extend to non-commutative polynomials? Our "strict" Positivstellensatz is positive news, on the opposite extreme from strict positivity would be a Real Nullstellensatz. We give an example which shows that there is no non-commutative Real Nullstellensatz along certain lines. However, we include a successful type of non-commutative Nullstellensatz proved by George Bergman.

1 Introduction

Let $\mathcal{P}$ denote a collection of symmetric polynomials in non-commutative variables $x = \{x_1, \cdots, x_g\}$. The positivity domain $\mathcal{D}_\mathcal{P}$ associated to $\mathcal{P}$ is the set of tuples $X = (X_1, \cdots, X_g)$ of symmetric bounded operators on separable real Hilbert space making $p(X_1, \cdots, X_g)$ a positive semi-definite operator. The domain is bounded if there exists a C so that $C^2 - X_j^T X_j$ is positive semi-definite whenever $X \in \mathcal{D}_\mathcal{P}$

For bounded $\mathcal{D}_\mathcal{P}$, our Positivstellensatz represents a polynomial q which is strictly positive on $\mathcal{D}_\mathcal{P}$ as a weighted sum of squares

$$q = \sum_1^N s_j^T p_j s_j + \sum_1^M r_k^T r_k + \sum t_{m,\ell}^T (C^2 - x_m^2) t_{m,\ell} \tag{1.1}$$

*Partially supported by the the NSF, DARPA and Ford Motor Co.
†Partially supported by NSF grant DMS-0140112

for polynomials $p_j \in \mathcal{P}$ and polynomials $s_j, r_k, t_{m,\ell}$. When $\mathcal{D}_\mathcal{P}$ is a convex set, the Hilbert space used in our positivity hypothesis can be taken to be finite dimensional. Versions of the Positivstellensatz are presented for three classes of matrix-valued non-commutative polynomials. The matrix-valued case is important for future applications and requires little extra effort.

Our proof uses a Hahn Banach separation argument and a GNS type construction much as in the Putinar and Vasilescu [PV] strengthing of Stengle's Positivstellensatz in the commutative case. In the commutative case, dropping the boundedness hypothesis weakens the conclusion in that the decomposition (1.1) must then include an additional weight factor. It is not known for non-commutative situations if boundedness of $\mathcal{D}_\mathcal{P}$ is essential for existence of a weighted sum of squares (SoS) representation. For instance, if $\mathcal{D}_\mathcal{P}$ is all tuples of finite dimensional symmetric matrices, then positivity of q is positivity "everywhere" and we are dealing with a "matrix positive polynomial". Such polynomials are sums of squares [H][M]. This is in distinction to the commutative case where they certainly are not all sums of squares.

In the remainder of the introduction notation and the three classes of non-commutative polynomials, NC polynomials for short, are introduced, definitions of positivity domains and weighted sums of squares representations are given, and a precise statement of the main result is presented. The exposition reflects the authors expectation that the results will appeal to both mathematicians and engineers. Section 2 gives stronger results on "convex" sets. Section 6 addresses the non-commutative Real Nullstellensatz. Sections 3, 4 and 5 give proofs.

1.1 NC Polynomials and Special Classes

Let $\mathcal{F}_g$ denote the free semi-group on the g non-commutative generators $x = \{x_1, \ldots, x_g\}$. In common language, $\mathcal{F}_g$ is the group of words in $x_1, \cdots, x_g$. Note that the empty word $\emptyset$ is the identity in $\mathcal{F}_g$. Below we define several classes of NC polynomials

1.1.1 Polynomials in Symmetric Entries, $\mathcal{N}$

Let $\mathcal{N}$ denote the polynomials, over the field of real numbers $\mathbb{R}$, in the non-commuting generators $x = \{x_1, \cdots, x_g\}$ so that $\mathcal{N}$ consists of real linear combinations of words w from $\mathcal{F}_g$. Concretely, $p \in \mathcal{N}$ is an expression of the form

$$p = \sum_{w \in \mathcal{F}_g} p_w w, \tag{1.2}$$

where the sum is finite and each $p_w \in \mathbb{R}$. The algebra $\mathcal{N}$ has a natural involution ^{T}, which behaves as follows. For a word $w = x_{j_1}x_{j_2}\cdots x_{j_n}$ from $\mathcal{F}_g$ viewed as an element of $\mathcal{N}$,

$$w^T = x_{j_n}\cdots x_{j_2}x_{j_1}.$$

In general, given p as in (1.2), $p^T = \sum p_w w^T$. A polynomial is p **in** $\mathcal{N}$ **symmetric** provided $p^T = p$.

Often we shall be interested in evaluating a polynomial p in $\mathcal{N}$ at a tuple of symmetric operators $X = (X_1, \dots, X_g)$ on a common Hilbert space H. Define $X^\emptyset = I$, the identity operator on H; given a word $w \in \mathcal{F}_g$ different from the empty word, $w = x_{j_1}x_{j_2}\cdots x_{j_n}$, let

$$X^w = X_{j_1}X_{j_2}\cdots X_{j_n};$$

and given p as in (1.2), define $p(X) = \sum p_w X^w$. Note that the involution on $\mathcal{N}$ is compatible with the transpose operation on operators on real Hilbert space,

$$p(X)^T = p^T(X),$$

where $p(X)^T$ denotes the adjoint operator (with respect to the native inner product). Often the Hilbert space is simply $\mathbb{R}^\ell$, and so the operators X_j are real symmetric $\ell\times\ell$ matrices and $p(X)^T$ is just the usual transpose of the $\ell \times \ell$ matrix $p(X)$.

Example 1. Let $p \in \mathcal{N}$ be given by $p = x_3x_2 + 3x_3x_1x_2$ and note $p^T = x_2x_3 + 3x_2x_1x_3$. If $X = (X_1, X_2, X_3)$, where

$$X_1 = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \quad X_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad X_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

then

$$p(X) = \begin{pmatrix} 0 & 0 \\ -5 & 3 \end{pmatrix} \quad \text{and} \quad p^T(X) = \begin{pmatrix} 0 & -5 \\ 0 & 3 \end{pmatrix}. \qquad \bullet\bullet$$

Example 2. If $q \in \mathcal{N}$ is the symmetric polynomial $q = p + p^T = x_3x_2 + 3x_3x_1x_2 + x_2x_3 + 3x_2x_1x_3$, then

$$q(X) = \begin{pmatrix} 0 & -5 \\ -5 & 6 \end{pmatrix}. \qquad \bullet\bullet$$

1.1.2 General Polynomials, $\mathcal{N}_*$

Let $\mathcal{N}_*$ denote the polynomials in the $2g$ non-commutative symbols $\{x_1, \dots, x_g, x_1^T, \dots, x_g^T\}$. The involution on $\mathcal{N}_*$ is most conveniently defined as follows. View the polynomials in $\mathcal{N}_*$ as polynomials in the $2g$

non-commuting variables $\{x_1, \dots, x_g, x_{g+1}, \dots, x_{2g}\}$, by identifying x_{g+j} with x_j^T. In this way, the involution on $\mathcal{N}_*$ is the same as that on $\mathcal{N}$ in $2g$, rather than g, variables.

We now define the rule of substitution of a tuple of (possibly non-symmetric) operators into a polynomial in $\mathcal{N}_*$. Given a word $w \in \mathcal{F}_{2g}$, and a tuple X of (possibly non-symmetric) operators, substitute X_j for x_j and X_j^T for x_{j+g} for $1 \le j \le g$ in each word and extended to all of $\mathcal{N}_*$.

Example 3. Choose $g = 2$ and let $p \in \mathcal{N}_*$ be given by $p = x_1 x_2^T + x_2^T x_1$. Observe $p^T = x_2 x_1^T + x_1^T x_2$ so that p and p^T take on exactly the same values when evaluated at a tuple of symmetric operators. However, with

$$X_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad X_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$$p(X) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \neq \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = p^T(X). \qquad \bullet\bullet$$

The next example shows that polynomials exhibit different properties when evaluated at operators on infinite dimensional spaces than they do when evaluated at matrices, a phenomena which must be considered in our Positivstellensatz. An operator X on a (real) Hilbert space H is positive semidefinite, written $X \succeq 0$, provided $X = X^T$ and $\langle Xx, x \rangle \ge 0$ for all $x \in H$.

Example 4. Let $p = 2x^T x - 1$ and $q = 2xx^T - 1$. If X is a square matrix and $p(X) \succeq 0$, then we have that X is invertible and $2 - X^{-T} X^{-1} \succeq 0$. Hence, $2I - X^{-1} X^{-T} \succeq 0$ and thus $q(X) \succeq 0$. This property, $p(X) \succeq 0$ implies $q(X) \succeq 0$, does not (necessarily) hold for operators on infinite dimensional spaces. Let S denote the forward shift operator on ℓ^2, $S(a_0, a_1, \dots) = (0, a_0, a_1, \dots)$, so that S^T is the backward shift operator, $S^T(a_0, a_1, \dots) = (a_1, a_2, \dots)$. Straightforward computation gives $p(S) = I \succeq 0$. However, $q(S)(a_0, a_1, \dots) = (-a_0, a_1, a_2, \dots)$ so that $q(S) \not\succeq 0$. $\bullet\bullet$

1.1.3 Hereditary Polynomials, $\mathcal{N}_*\mathcal{N}$

The last type of NC polynomial we consider is a subset of $\mathcal{N}_*$ called the hereditary polynomials [A], denoted by $\mathcal{N}_*\mathcal{N}$. A polynomial $p \in \mathcal{N}_*$ is **hereditary** provided the transposes, if any, always appear on the left and is thus a finite linear combination of terms $v^T w$ where v and w are words in $x = \{x_1, \dots, x_g\}$,

$$p = \sum_{v,w \in \mathcal{F}_g} p_{v,w} v^T w.$$

In particular $p^T = \sum p_{v,w}(v^T w)^T$, so that if p is hereditary, then so is p^T. Given a Hilbert space H and a tuple $X = (X_1, \ldots, X_g)$ of operators on H, the definition of $p(X)$ is induced from that on $\mathcal{N}_*$.

The product of two hereditary polynomials need not be an hereditary polynomial, but if p is hereditary and q and r are polynomials in $x = \{x_1, \ldots, x_g\}$ (no x_j^*'s), then $q^T pr$ is again an hereditary polynomial.

Example 5. Let $p \in \mathcal{N}_*\mathcal{N}$ be given by $p = x_3^T x_2 + 3x_3^T x_1^T x_2$. So $p^T = x_2^T x_3 + 3x_2^T x_1 x_3$. A polynomial $q \in \mathcal{N}_*$ which is both hereditary and symmetric is

$$q = p^T + p = x_3^T x_2 + 3x_3^T x_1^T x_2 + x_2^T x_3 + 3x_2^T x_1 x_3.$$

Note that p, p^T, and q take on the same values as their counterparts in Examples 1 and 2 when evaluated at a tuple of symmetric operators. ●●

Our final example demonstrates for $\mathcal{N}_*\mathcal{N}$ what Example 4 did for $\mathcal{N}_*$, namely that we must consider operators on infinite dimensional Hilbert spaces in our Positivstellensatz for $\mathcal{N}_*\mathcal{N}$.

Example 6. Let $p(x) = -(x^T)^2 x^2 + 2x^T x - 1$ and $q(x) = 1 - x^T x$. If X is a square matrix and $p(X) \succeq 0$, then $q(X) = 0$ (so $q(X) \succeq 0$ as well). However, if X is the Brownian shift operator on $\ell^2(\mathbb{R}) \oplus \mathbb{R}$,

$$X = \begin{pmatrix} S & E \\ 0 & 1 \end{pmatrix},$$

then $p(X) = 0$, but $q(X) \preceq 0$ non-trivially. Here, S denotes the forward shift on $\ell^2(\mathbb{R})$ as in example 4, and $E : \mathbb{R} \to \ell^2(\mathbb{R})$ is given by $Ec = (c, 0, \ldots)$. ●●

1.1.4 Matrix Valued Polynomials, $M_l(\mathcal{N}_*)$ etc

We wish also to consider matrix-valued NC polynomials or, equivalently, NC polynomials with matrix coefficients. Let $M_{a\times b}$ denote the $a \times b$ matrices with entries from $\mathbb{R}$ and let $M_{a\times b}(\mathcal{N}_*)$ denote the $M_{a\times b}$ matrices with entries from $\mathcal{N}_*$. A $p \in M_{a\times b}(\mathcal{N}_*)$ can be expressed as $p = \sum_w p_w w$, where $p_w \in M_{a\times b}$ and $p_w = 0$ except for finitely many w, and may be thought of as $M_{a\times b}$-valued polynomial in the NC indeterminates $\{x_1, \ldots, x_g, x_1^T, \ldots, x_g^T\}$. The involution ^{T} extends to a mapping $M_{a\times b}(\mathcal{N}_*) \longrightarrow M_{b\times a}(\mathcal{N}_*)$ as

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

The substitution rule $p(X)$, for $p \in M_{a\times b}(\mathcal{N}_*)$ and a tuple $X = (X_1, \ldots, X_g)$ on a Hilbert space H, is made entry-wise. That is, writing

$p = (p_{j,\ell})$ where $p_{j,\ell}$ are polynomials in $\mathcal{N}_*$ for $1 \leq j \leq a$ and $1 \leq \ell \leq b$, define $p(X) : \oplus_1^b H \longrightarrow \oplus_1^a H$ as the operator given in block matrix form $p(X) = (p_{j,\ell}(X))_{j,\ell}$.

The substitution rule $p(X)$ is conveniently described using tensor product notation as well. Given Hilbert spaces K and H, define, for elementary tensors $k \otimes h$ and $k' \otimes h'$,

$$\langle k \otimes h, k' \otimes h' \rangle = \langle k, k' \rangle_K \langle h, h' \rangle_H,$$

where $k, k' \in K$ and $h, h' \in H$. Extend the form $\langle \cdot, \cdot \rangle$ to the algebraic tensor product of K and H over $\mathbb{R}$ by linearity and let $K \otimes H$ denote the Hilbert space obtained by forming the completion. If A is an operator on H and X is an operator from K to L, then $X \otimes A$ is the operator from $K \otimes H$ to $L \otimes H$ defined on elementary tensors by

$$X \otimes A \; k \otimes h = Xk \otimes Ah.$$

With these notations,

$$p(X) = \sum p_w \otimes X^w,$$

where each p_w is viewed as an operator $p_w : \mathbb{R}^b \longrightarrow \mathbb{R}^a$.

Definitions for $M_{a\times b}(\mathcal{N})$, the involution ${}^T : M_{a\times b}(\mathcal{N}) \longrightarrow M_{ba}(\mathcal{N})$, and the substitution $p \in M_{a\times b}(\mathcal{N}) \mapsto p(X)$ are made by analogy with the $\mathcal{N}_*$ case. Slightly different notation will be used in the hereditary case. A $M_{a\times b}$-valued hereditary polynomial is a finite sum of the form $p = \sum p_{v,w} v^T w$, where $v, w \in \mathcal{N}$ and $p_{v,w} \in M_{a\times b}$. The involution and substitution are those coming from the inclusion $M_{a\times b}(\mathcal{N}_*\mathcal{N}) \subset M_{a\times b}(\mathcal{N}_*)$.

For notational purposes, let $M_\ell(\mathcal{N}_*) = M_{\ell\times\ell}(\mathcal{N}_*)$. A matrix-valued NC polynomial $p \in M_\ell(\mathcal{N}_*)$ is **symmetric** provided $p^T = p$ which is equivalent to $p_w^T = p_{w^T}$ in the $\mathcal{N}$ and $\mathcal{N}_*$ cases, and $p_{v,w}^T = p_{w,v}$ in the hereditary case. Let $M_\ell^s(\mathcal{N}_*)$ denote the symmetric polynomials in $M_\ell(\mathcal{N}_*)$ and define $M_\ell^s(\mathcal{N})$ and $M_\ell^s(\mathcal{N}_*\mathcal{N})$ analogously. Finally, we use the notation $M_\infty(\mathcal{N}_*) = \cup_{\ell>0} M_\ell(\mathcal{N}_*)$, $M_\infty^s(\mathcal{N}_*) = \cup_{\ell>0} M_\ell^s(\mathcal{N}_*)$, and $M_\infty^s(\mathcal{N}_*\mathcal{N}) = \cup_{\ell>0} M_\ell^s(\mathcal{N}_*\mathcal{N})$.

If p is an $M_{b\times c}$-valued polynomial, q is an $M_{b\times a}$-valued polynomial, and r is an $M_{c\times d}$-valued polynomial, then $q^T pr$ is an $M_{a\times d}$-valued polynomial. If p, q, r are all in either $\mathcal{N}$ or $\mathcal{N}_*$, then so is $q^T pr$. In the hereditary case, if p is hereditary and q and r are transpose free, then $q^T pr$ is again hereditary. A special case of particular importance is when p is $M_{b\times b}$-valued, q is $M_{b\times a}$-valued and $r = q^T$.

1.2 Decomposition as Weighted Sums of Squares

Fix a collection of symmetric matrix-valued polynomials $\mathcal{P}$ from either $M_\infty^s(\mathcal{N})$, $M_\infty^s(\mathcal{N}_*)$, or $M_\infty^s(\mathcal{N}_*\mathcal{N})$.

Let $\mathcal{C}_\mathcal{P}^\ell$ denote finite positive linear combinations of s^Tps and r^Tr where $p \in \mathcal{P}$ and the sizes of s and r are such that the products make sense and result in $\ell \times \ell$ matrix-valued polynomials. Thus $q \in \mathcal{C}_\mathcal{P}^\ell$ if there exists non-negative integers M, N, and polynomials $p_j \in \mathcal{P}$ and polynomials s_j and r_k, where say p_j is M_{ℓ_j}-valued, s_j is $M_{\ell_j\times\ell}$-valued and r_j is $M_{1\times\ell}$-valued, and of course all depend upon q, such that

$$q = \sum_1^N s_j^T p_j s_j + \sum_1^M r_k^T r_k. \tag{1.3}$$

We emphasize that, while $\mathcal{P}$ may be an infinite set of polynomials, the decomposition (1.3) is a finite sum and that $\mathcal{C}_\mathcal{P}^\ell$ consists of symmetric polynomials. Let $\mathcal{C}_\mathcal{P}$ denote the union of $\mathcal{C}_\mathcal{P}^\ell$ over ℓ. We call (1.3) a weighted sum of squares representation.

1.3 Domain of Positivity

Fix a subset $\mathcal{P}$ of $M_\infty(\mathcal{N}_*)$, $M_\infty(\mathcal{N})$, or $M_\infty(\mathcal{N}_*\mathcal{N})$. The case that $\mathcal{P}$ consists of symmetric polynomials is of primary interest, but we will have occasion to consider more general collections. Given a real Hilbert space H, let $\mathcal{D}_\mathcal{P}(H)$ denote the tuples $X = (X_1, \dots, X_g)$ such that each X_j is an operator on H and $p(X) \succeq 0$ for each $p \in \mathcal{P}$. The **positivity domain of** $\mathcal{P}$, denoted $\mathcal{D}_\mathcal{P}$, is the collection of tuples X such that $X \in \mathcal{D}_\mathcal{P}(H)$ for some H. The fact that $\mathcal{D}_\mathcal{P}$ is not actually a set presents no logical difficulties and typically it may be assumed that the Hilbert spaces are separable.

A positivity domain $\mathcal{D}_\mathcal{P}$ is called **convex** provided that, if X and Y are both operator tuples on the same Hilbert space H and both X and Y lie in $\mathcal{D}_\mathcal{P}$, then convex combinations $c_1X + c_2Y$ belong to $\mathcal{D}_\mathcal{P}$. Here real numbers $c_1, c_2 \geq 0$ satisfy $c_1 + c_2 = 1$. Thus, $\mathcal{D}_\mathcal{P}$ is convex if $\mathcal{D}_\mathcal{P}(H)$ is convex for each H. A **positivity domain is bounded** provided there is a constant $C > 0$ such that if $X \in \mathcal{D}_\mathcal{P}$, then $\|X_j\| \leq C$ for each $j = 1, 2, \dots, g$.

We now define a special set of polynomials, and state our first lemma. Henceforth set

$$b_j := C^2 - x_j^T x_j \quad \text{and} \quad \tilde{b}_j := C^2 - x_j x_j^T.$$

Lemma 1.1 *For any j between 1 and g, $\mathcal{D}_{b_j} = \mathcal{D}_{\tilde{b}_j}$.*

Proof. Without loss of generality, take $C = 1$. Suppose X is such that $I - X^TX \succeq 0$. Since both $I \succeq 0$ and $I - X^TX \succeq 0$, we have that

$$\begin{pmatrix} Y^T & Z^T \end{pmatrix} \begin{pmatrix} I & X \\ X^T & I \end{pmatrix} \begin{pmatrix} Y \\ Z \end{pmatrix} \succeq 0 \qquad \text{for any } Y, Z.$$

In particular, for $Y = I$ and $Z = -X^T$, we get $I - XX^T \succeq 0$. So, $\mathcal{D}_{b_j} \subset \mathcal{D}_{\tilde{b}_j}$. Replacing X with X^T gives us the reverse containment. ••

The lemma may also be proved by noting that $\|XX^T\| = \|X^TX\|$, but we chose to emphasis the algebraic approach as it illustrates the use of the cones $\mathcal{C}_{\mathcal{P}}$.

1.4 An NC Positivstellensatz

Let b denote the set of polynomials $b := \{b_1, \cdots, b_g\}$. Similarly, let $\tilde{b} := \{\tilde{b}_1, \cdots, \tilde{b}_g\}$. If $\mathcal{D}_{\mathcal{P}}$ is a bounded domain, then for large enough C, we have $\mathcal{D}_{\mathcal{P}\cup\{b\}} = \mathcal{D}_{\mathcal{P}}$. For bounded domains we take the convention:

- In the hereditary case and in the $\mathcal{N}_*$ case , for each i we adjoin the polynomial $C^2 - x_i^T x_i$ to $\mathcal{P}$ and obtain a bigger set $\tilde{\mathcal{P}}$.
- In the $\mathcal{N}$ case, for each i we adjoin the polynomial $C^2 - x_i^2$ to $\mathcal{P}$ and obtain a bigger set $\tilde{\mathcal{P}}$.

Often in what follows, when $\mathcal{D}_{\mathcal{P}}$ is bounded, we will assume $\mathcal{P} = \tilde{\mathcal{P}}$; i.e., that $b \subset \mathcal{P}$.

An operator X on a real Hilbert space H is strictly positive definite if $X = X^T$ and $\langle Xx, x\rangle > 0$ whenever $x \in H$ and $x \neq 0$.

Theorem 1.2 *Suppose $\mathcal{P}$ is a subset of matrices with non-commutative polynomial entries, to be precise $\mathcal{P} \subset M_\infty^s(\mathcal{N}_*)$ (resp. $\mathcal{P} \subset M_\infty^s(\mathcal{N})$, or resp. $\mathcal{P} \subset M_\infty^s(\mathcal{N}_*\mathcal{N})$), and suppose the positivity domain $\mathcal{D}_{\mathcal{P}}$ of $\mathcal{P}$ is bounded. If $q \in M_\ell(\mathcal{N}_*)^s$ (resp. $q \in M_\ell^s(\mathcal{N})$, or resp. $q \in M_\ell^s(\mathcal{N}_*\mathcal{N})$) is strictly positive definite on $\mathcal{D}_{\mathcal{P}}$, that is, if q is symmetric and $q(X)$ is strictly positive definite whenever $X \in \mathcal{D}_{\mathcal{P}}$, then $q \in \mathcal{C}_{\tilde{\mathcal{P}}}^\ell$, that is, q has the representation*

$$q = \sum_1^N s_j^T p_j s_j + \sum_1^M r_k^T r_k. \tag{1.4}$$

in equation (1.3). Here $p_j \in \tilde{\mathcal{P}}$, for $1 \leq j \leq N$.

If in addition $\mathcal{D}_{\mathcal{P}}$ is convex, we need only verify $q(X)$ is a positive definite operator for those $X \in \mathcal{D}_{\mathcal{P}}$ which are defined on a Hilbert space of dimension at most $\ell \sum_0^d (2g)^n$. From this test on finite dimensional matrices we obtain $q \in \mathcal{C}^{\ell}_{\tilde{\mathcal{P}}}$.

The proof has two parts which dictates the organization of the rest of the paper. The first is a Hahn-Banach result which separates $\mathcal{C}^{\ell}_{\mathcal{P}}$ from any polynomial q outside it with a linear functional λ. The second represents such linear functionals λ using a matrix tuple X. That q is outside $\mathcal{C}^{\ell}_{\mathcal{P}}$ forces $q(X)$ to be not strictly positive definite. Before launching into all of this the next section presents properties of convex positivity domains, since it is a pleasant topic, and in this section we prove the last assertion of Theorem 1.2, see Proposition 2.3.

The following example explains the strict positive definite hypothesis on $q(X)$ even in one variable. Let $q = 1 - x^T x$ and $p = q^3$. The positivity domain determined $\mathcal{D}_{\{p\}}$ consists matrices X with $\|X\| \leq 1$. Certainly, $q(X) \geq 0$ whenever $X \in \mathcal{D}_{\{p\}}$. In this case a natural choice for C in Theorem 1.2 is $C = 1$. However, if $C > 1$ is chosen, then q cannot have a representation as in equation (1.4). To see this consider the 1×1 matrices x, for $0 < x < 1$. If the representation holds, then there are polynomials s_j, t_j, r_j so that

$$1 - x^2 = \sum s_j(x)^2 (1 - x^2)^3 + \sum t_j(x)^2 (C^2 - x^2) + \sum r_j(x).$$

The left hand side has a zero of order one at $x = 1$. It follows that each t_j and r_j has a zero at 1. Thus, as $(1 - x^2)^3$ has a zero of order three at 1, the right hand side has a zero of order at least two at 1, a contradiction.

1.4.1 Related Results

We are aware of several variations of Theorem 1.2 with proofs much like the one given here. In fact, the proof here borrows heavily from Putinar and Vasilescu [PV], although the results there are for the commutative $\mathcal{N}$ scalar case and much of the significance of their work is for the case of unbounded positivity regions.

Agler, in his seminal work on Schur class functions on the polydisc [A], begins with the collection $\mathcal{P} = \{1 - x_j^* x_j : j = 1, 2, \ldots, g\}$ (here * is the complex transpose, rather than just transpose) and shows that a matrix-valued analytic function W, on the g-fold polydisc, such that

$W(X)$ is a contraction for each tuple of commuting strict contractions $X = (X_1, \ldots, X_g)$ can be written as

$$I - W(z)W(w)^* = \sum H_j(z)(1 - z_j\overline{w_j})H_j(w)^*. \tag{1.5}$$

Agler and McCarthy [AM] prove a generalization involving a finite collection of scalar polynomials in place of $1 - x_j^* x_j$. Also Joe Ball and Malakorn [BMprep] have in preperation an extension of the representation (1.5) to noncommutative variables which will interest the serious reader. It uses formal power series to represent the functions H_j, so exactly when they are polynomials depends on circumstances.

Also related to the NC Positivstellensatz is the characterization of polynomials which are actually sums of squares. In this case $\mathcal{D}$ is all tuples of operators and $\mathcal{P}$ is empty. For the $\mathcal{N}_*$ case see [H] and for the $\mathcal{N}$ case see [M] which also treats the case when the variables are unitary.

2 Convex Positivity Domains

In this section we specialize our non-commutative Positivstellensatz to convex domains and find that the structure is very rigid. We also prove the finite dimensionality assertion of Theorem 1.2. As background we introduce several properties of positivity domains and two natural notions of convexity.

2.1 Properties of Positivity Domains

Proposition 2.1 *Given a set $\mathcal{P}$ of polynomials in $M_l(\mathcal{N}_*)$, $M_l(\mathcal{N})$, or in $M_l(\mathcal{N}_*\mathcal{N})$. The positivity domain $\mathcal{D}_\mathcal{P}$ has the properties that it is closed with respect to the operations:*

1. **Restriction to Reducing Subspaces**. *Suppose $X = (X_1, \ldots, X_g)$ is a tuple of operators on Hilbert space H and $X \in \mathcal{D}_\mathcal{P}$. If K is a subspace of H which is invariant for X_j and X_j^T for each $1 \leq j \leq d$, then the tuple of operators $V^TXV = (V^TX_1V, \ldots, V^TX_gV)$ on K is in $\mathcal{D}_\mathcal{P}$, where V is the inclusion isometry from K into H.*

2. **Direct Sums**. *Suppose $X = (X_1, \ldots, X_g)$ is a tuple of operators on Hilbert space H, and suppose $Y = (Y_1, \ldots, Y_g)$ is a tuple of operators on Hilbert space K. If both X and Y are in $\mathcal{D}_\mathcal{P}$, then the direct sum $X \oplus Y = (X_1 \oplus Y_1, \ldots, X_g \oplus Y_g)$, which is a tuple of operators on $H \oplus K$, is in $\mathcal{D}_\mathcal{P}$.*

3. **Unitary Conjugation**. *Suppose $X = (X_1, \ldots, X_g)$ is a tuple of operators on Hilbert space H, K is a Hilbert space, and $U : K \longrightarrow H$ is unitary. Then the tuple $U^T XU = (U^T X_1 U, \ldots, U^T X_g U)$ is in $\mathcal{D}_\mathcal{P}$. More generally, if $\pi : \mathcal{B}(H) \longrightarrow \mathcal{B}(K)$ is a unital representation which preserves ^{T} and if $X \in \mathcal{D}_\mathcal{P}$, then $\pi(X) = (\pi(X_1), \ldots, \pi(X_g))$ is in $\mathcal{D}_\mathcal{P}$.*

Proof. This is an immediate consequence of

(a) $p(X \oplus Y) = p(X) \oplus p(Y)$;

(b) $p(V^T XV) = V^T p(X)V$ if either $K \subset H$ is reducing or $V : K \longrightarrow H$ is unitary; and

(c) $p(\pi(X)) = \pi(p(X))$. ••

2.2 Convexity

A positivity domain $\mathcal{D}_\mathcal{P}$ is **closed with respect to compressions** if for each tuple $X = (X_1, \ldots, X_g)$ of operators on Hilbert space H which lies in $\mathcal{D}_\mathcal{P}$ and isometry V from a Hilbert space K into H, the tuple of operators $V^T XV = (V^T X_1 V, \ldots, V^T X_g V)$ on K is in $\mathcal{D}_\mathcal{P}$. Closed with respect to compression is a stronger property than item (1) of Proposition 2.1 in that VK need not be a reducing subspace for X and certainly is not enjoyed by all positivity domains. However, convex positivity domains are closed with respect to compressions.

Lemma 2.2 *Every convex domain which is closed with respect to restriction to reducing subspaces and unitary conjugation is closed with respect to compression. Conversely, a domain $\mathcal{D}$ which is closed with respect to direct sums and compression is convex. In particular, a positivity domain $\mathcal{D}_\mathcal{P}$ is closed with respect to compressions if and only if it is convex.*

Proof. Suppose the convex domain $\mathcal{D}$ is closed with respect to restriction to reducing subspaces and unitary conjugation as described in items (1) and (3) of Proposition 2.1. To see that $\mathcal{D}$ is closed with respect to compressions, let $X \in \mathcal{D}$, a tuple form $\mathcal{B}(H)$, and K, a subspace of H, be given. Let M denote the orthogonal complement of K in H and, write, with respect to the orthogonal direct sum $H = K \oplus M$,

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

The matrix

$$\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} = \frac{1}{2}\begin{pmatrix} A & B \\ C & D \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

belongs to $\mathcal{D}_{\mathcal{P}}$, because $\mathcal{D}_{\mathcal{P}}$ is convex and because $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is unitary. Since K reduces $\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ we get that $A \in \mathcal{D}_{\mathcal{P}}$.

To prove the converse, let tuples $X, Y \in \mathcal{D}$ acting on the common Hilbert space H and real numbers a_1, a_2 satisfying $a_1^2 + a_2^2 = 1$ be given. Let $V : H \oplus H \to H$ denote the isometry

$$V := \begin{pmatrix} a_1^T & a_2^T \end{pmatrix}.$$

Since $X, Y \in \mathcal{D}$, their direct sum $X \oplus Y$ is in $\mathcal{D}$, and hence

$$a_1^2 X + a_2^2 Y = a_1^T X a_1 + a_2^T Y a_2 = V^T (X \oplus Y) V$$

is in $\mathcal{D}$. ••

2.3 Proof of Finite Dimensionality

For a convex positivity domain $\mathcal{D}$ Theorem 1.2 gives a bound on the dimension of Hilbert space H needed in the Positivstellensatz.

Proposition 2.3 *Let $\mathcal{P}$ be a collection of symmetric polynomials from $M_\ell^s(\mathcal{N}_*)$ (resp. $M_\ell^s(\mathcal{N})$, resp. $M_\ell^s(\mathcal{N}_*\mathcal{N})$) and suppose $\mathcal{D}_{\mathcal{P}}$ is a bounded convex positivity domain. If q is a symmetric $M_l(\mathcal{N}_*)$-valued polynomial of degree d and if $q \notin \mathcal{C}_{\mathcal{P}}$, then there exists a Hilbert space H of dimension at most $\ell \sum_0^d (2g)^n$, a non-zero vector $\gamma \in H$, and a tuple $X = (X_1, \ldots, X_g)$ of operators on H such that $X \in \mathcal{D}_{\mathcal{P}}$, but $\langle q(X)\gamma, \gamma\rangle \leq 0$.*

Proof. From the Positivstellensatz, there exists a tuple $Z = (Z_1, \ldots, Z_g)$ acting on a Hilbert space H and a non-zero vector $\gamma = \oplus\gamma_j \in \oplus_1^\ell H$ such that $\langle q(Z)\gamma, \gamma\rangle \leq 0$. Here $q \in M_l(\mathcal{N}_*)$ and has degree d. Let

$$K_d = \operatorname{span}\{Z^w \gamma_j : \ w \text{ is a word of length at most } d, \ j = 1, \cdots, \ell\} \subset H.$$

Then K has dimension at most equal to the total number of words w of length at most d times ℓ. Since there are $2g$ generators for the words in $\mathcal{N}_*$, we get K_d has dimension at most $\ell \sum_0^d (2g)^n$.

If P is the projection of H onto K_d, then

$$\langle q(PZP)\gamma, \gamma\rangle = \langle q(Z)\gamma, \gamma\rangle \ \leq 0.$$

On the other hand, convexity implies $PZP \in \mathcal{D}_{\mathcal{P}}$, since by Lemma 2.2 the domain $\mathcal{D}_{\mathcal{P}}$ is closed under compression. This proves the theorem by taking $X := PZP$. ●●

A tighter bound on dimension is immediate for $\mathcal{N}$ and $\mathcal{N}_*\mathcal{N}$, since there are fewer words. The bound for $\mathcal{N}$ is $\ell \sum_0^d g^n$, and for $\mathcal{N}_*\mathcal{N}$ is $2\ell \sum_0^d g^n$.

3 Separating $\mathcal{C}_{\mathcal{P}}$ from Outsiders with a Linear Functional

In this section a fairly general version of the Hahn-Banach theorem is used to construct, given $q \in M_\ell^s(\mathcal{N})$ but $q \notin \mathcal{C}_{\mathcal{P}}^\ell$, a linear functional $\lambda : M_\ell^s(\mathcal{N}) \longrightarrow \mathbb{R}$ such that $\lambda(p) \geq 0$ for $p \in \mathcal{C}_{\mathcal{P}}^\ell$, but $\lambda(q) \leq 0$. The argument depends upon the boundedness of $\mathcal{D}_{\mathcal{P}}$.

3.1 Properties of $\mathcal{C}_{\mathcal{P}}$ for Bounded $\mathcal{D}_{\mathcal{P}}$

We first focus on the structure related to boundedness of $\mathcal{D}_{\mathcal{P}}$. Recall $b_j = C^2 - x_j^T x_j$ and b denotes the set of polynomials $b = \{b_1, \cdots, b_g\}$; likewise $\tilde{b}$ corresponds to $\tilde{b}_j = C^2 - x_j x_j^T$.

Lemma 3.1 *If w is a word of length n, then*

1. $C^{2n} - w^T w \in \mathcal{C}_{\{b,\tilde{b}\}}^\ell$ *for w in $\mathcal{N}_*$;*
2. $C^{2n} - w^T w \in \mathcal{C}_b^\ell$ *for w in $\mathcal{N}$; and*
3. $C^{2n} - w^T w \in \mathcal{C}_b^\ell$ *for w in $\mathcal{N}_*\mathcal{N}$.*

Proof. First consider the case of hereditary words w in which case w is a word in $x = \{x_1, \ldots, x_g\}$ and w^T is a word in $x^T = \{x_1^T, \ldots, x_g^T\}$. We use $(C^2 - x_j^T x_j) \subset \mathcal{C}_b^\ell$ and argue by induction. Accordingly suppose the result is true for the word v of length n and consider $w = x_{j_0} v$. We have,

$$(C^{2n+2} - w^T w) = C^2(C^{2n} - v^T v) + v^T(C^2 - x_{j_0}^T x_{j_0})v. \tag{3.1}$$

This implies

$$(C^{2n+2} - w^T w) \in C^2(C^{2n} - v^T v) + \mathcal{C}_b^\ell$$

which yields the induction step for going from hereditary word v to hereditary word w.

When $w \in \mathcal{N}_*$ is a word in x_j, x_k^T which is not necessarily hereditary we proceed as before but also consider $w = x_{j_0}^T v$, and obtain

$$(C^{2n+2} - w^T w) = C^2(C^{2n} - v^T v) + v^T(C^2 - x_{j_0} x_{j_0}^T)v. \tag{3.2}$$

Combine (3.1) and (3.2) to obtain the induction step for going from words v in $\mathcal{N}_*$ to words w in $\mathcal{N}_*$.

For words in symmetric variables, that is in $\mathcal{N}$, the same argument prevails. ••

A set $\mathcal{A}$ in a vector space $\mathcal{V}$ is called **absorbing** provided that, for each $v \in \mathcal{V}$, there is an $t > 0$ such that $v \in t\mathcal{A}$. Note that if $\mathcal{A}$ is absorbing, then $0 \in \mathcal{A}$ and if $\mathcal{A}$ is also convex, then for each $v \in \mathcal{V}$, there exists an $\epsilon > 0$ such that $v \in t\mathcal{A}$ for $t > \epsilon$.

Lemma 3.2 *Let M denote either $M_\ell^s(\mathcal{N}_*)$, $M_\ell^s(\mathcal{N}_*\mathcal{N})$, or $M_\ell^s(\mathcal{N})$. The set $\mathcal{C}_b^\ell - I$ is absorbing in M. In fact, if $p \in M$, then there exists a real number $t \geq 0$ and an $s \in \mathcal{C}_b^\ell$ such that $p = s - tI$.*

Proof. We do the hereditary case first, since the calculations involved include the calculations for the other cases. For the hereditary case, let $g, h \in \mathbb{R}^\ell$ and $v, w \in \mathcal{F}_g$ be given. Choose d so that $|w|, |v| \leq d$, where $|w|$ denotes the length of the word w, and assume that $C \geq 1$ so that $C^{2d} \geq C^{2|w|}$ and $C^{2d} \geq C^{2|v|}$. We have,

$$\begin{aligned} &gh^T w^T v \ + \ hg^T v^T w \\ &\qquad = \ (g^T w + h^T v)^T (g^T w + h^T v) - gg^T v^T v - hh^T w^T w. \end{aligned} \tag{3.3}$$

Observe that

$$\begin{aligned} C^{2d}(g^T g + h^T h)I \ &- \ gg^T v^T v - hh^T w^T w \\ = gg^T(C^{2d} \ &- \ v^T v) + hh^T(C^{2d} - w^T w) \\ &+ \ C^{2d}(g^T g I - gg^T)1 + C^{2d}(h^T h I - hh^T)1 \end{aligned} \tag{3.4}$$

and

$$(g^T w + h^T v)^T (g^T w + h^T v) \tag{3.5}$$

are both in $\mathcal{C}_\mathcal{P}^\ell$ and

$$\begin{aligned} &- \ (gg^T v^T v + hh^T w^T w) \\ &= \ (C^{2d}(g^T g + h^T h)I - gg^T v^T v - hh^T w^T w) - C^{2d}(g^T g + h^T h). \end{aligned} \tag{3.6}$$

Combining (3.3), (3.4), (3.5), (3.6), shows that

$$gh^T w^T v + hg^T v^T w = (\text{poly in } \mathcal{C}_{\{b\}}) - tI$$

for some $t \geq 0$. To complete the proof of the lemma in the hereditary case, observe, if p (in $M_l(\mathcal{N}_*\mathcal{N})$) is a given hereditary polynomial, then there is a d such that p is a linear combination of terms of the form $gh^T w^T v + hg^T v^T w$ where all the words have length at most d. Thus, there exists $t \geq 0$ and $s \in \mathcal{C}_{\{b\}}$ so that $p = s - tI$. Since $s + I \in \mathcal{C}_{\{b\}}$, we have $p = (s + I) - (t + 1)I$ and thus, $p \in (t + 1)(\mathcal{C}_{\{b\}} - I)$.

The argument for the $\mathcal{N}$ case is nearly identical to that given above for the hereditary case. Simply note that if $p \in M_\ell^s(\mathcal{N})$, then there is a d such that p is a linear combination of terms of the form $gh^T w^T v + hg^T v^T w$ where $v, w \in \mathcal{F}_g$ are words of length at most d and where now v^T is still a word in $\{x_1, \ldots, x_g\}$, rather than in the variables $\{x_1^T, \ldots, x_g^T\}$ as was the case above.

Finally, for the $\mathcal{N}_*$ case observe that $p \in M_\ell^s(\mathcal{N}_*)$ is a linear combination of terms of the form $gh^T w^T v + hg^T v^T w$ where $v, w \in \mathcal{F}_{2g}$ are words of length at most d, for some d, in the variables $\{x_1, \ldots, x_g, x_1^T, \ldots, x_g^T\}$.
••

3.2 Separating a Polynomial from $\mathcal{C}_\mathcal{P}$

Recall the convention for bounded positivity domains set forth in Section 1.4. Henceforth when working with bounded positivity domains we assume $b = \{b_1, \ldots, b_g\} \subset \mathcal{P}$, equivalently $\mathcal{P} = \tilde{\mathcal{P}}$.

Now we give the main result of Section 3.

Proposition 3.3 *We are given $\mathcal{P}$ with a bounded positivity domain $\mathcal{D}_\mathcal{P}$. If q is in $M_\ell^s(\mathcal{N}_*)$ (resp. $M_\ell^s(\mathcal{N})$, resp. $M_\ell^s(\mathcal{N}_*\mathcal{N})$), but not in the corresponding $\mathcal{C}_\mathcal{P}^\ell$, then there is a non-zero linear functional $\lambda : M_\ell^s(\mathcal{N}_*) \longrightarrow \mathbb{R}$, (resp. $\lambda : M_\ell^s(\mathcal{N}) \longrightarrow \mathbb{R}$, resp. $\lambda : M_\ell^s(\mathcal{N}_*\mathcal{N}) \longrightarrow \mathbb{R}$) such that $\lambda(\mathcal{C}_\mathcal{P}^\ell) \geq 0$, $\lambda(q) \leq 0$, and $\lambda(I) > 0$. Here I is the polynomial constantly equal to the $\ell \times \ell$ identity matrix.*

Proof. Let M denote $M_\ell^s(\mathcal{N}_*)$, $M_\ell^s(\mathcal{N})$, or $M_\ell(\mathcal{N}_*\mathcal{N})$ as the case may be and let $\mathcal{C}_\mathcal{P}^\ell$ denote the corresponding cone. Suppose $q \notin \mathcal{C}_\mathcal{P}^\ell$. Since the set $\mathcal{A} = \mathcal{C}_\mathcal{P}^\ell - I$ is convex and since it contains $\mathcal{C}_b^\ell - I$, it is absorbing by Lemma 3.2. Thus $\mathcal{A}$ has a Minkowski functional $\mu_\mathcal{A} : M \mapsto \mathbb{R}$,

$$\mu_\mathcal{A}(p) := \inf\{t > 0 : p \in t\mathcal{A}\}. \tag{3.7}$$

As $q - I \notin \mathcal{A}$, $\mu_{\mathcal{A}}(q - I) \geq 1$. Let L denote the span of $q - I$ in M and define $f : L \mapsto \mathbb{R}$ by $f(t(q - I)) = t$. Verify that $f \leq \mu_{\mathcal{A}}$ on L so that by a version (see [R]) of the Hahn-Banach theorem, f extends to a linear functional F on M satisfying $F(p) \leq \mu_{\mathcal{A}}(p)$ for $p \in M$. (Note this does not assume any topology.) In particular, if $s \in \mathcal{C}^{\ell}_{\mathcal{P}}$, then

$$F(s) - F(q) + 1 = F(s - q + (q - I)) = F(s - I) \leq \mu_{\mathcal{A}}(s - I) \leq 1.$$

Thus, $F(s) \leq F(q)$. Since $s \in \mathcal{C}^{\ell}_{\mathcal{P}}$ and $t > 0$ implies $ts \in \mathcal{C}^{\ell}_{\mathcal{P}}$, it follows that $tF(s) = F(ts) \leq F(q)$ for all $t > 0$. Thus, $F(s) \leq 0$ and $F(q) \geq 0$. Let $\lambda = -F$. Then $\lambda(s) \geq 0$ for all $s \in \mathcal{C}^{\ell}_{\mathcal{P}}$ and $\lambda(q) \leq 0$.

To see that $\lambda(I) > 0$, suppose to the contrary that $\lambda(I) = 0$ (since $I \in \mathcal{C}^{\ell}_{\mathcal{P}}$, $\lambda(I) \geq 0$). If $p \in \mathcal{C}^{\ell}_{\mathcal{P}}$, then, by Lemma 3.2 applied to $-p$, there exists $t > 0$ and $s \in \mathcal{C}^{\ell}_{\mathcal{P}}$ such that $-p = s - tI$. Thus, $tI = s + p$. Now, $\lambda(s)$ and $\lambda(p)$ are nonnegative, but $\lambda(tI) = 0$ and therefore, $\lambda(p) = 0$. Thus, $\lambda(\mathcal{C}^{\ell}_{\mathcal{P}}) = 0$. Finally, if $p \in M$, then there exists $r, s \in \mathcal{C}^{\ell}_{\mathcal{P}}$ so that $p = r - s$ and therefore $\lambda(p) = 0$, a contradiction to $\lambda(I - q) = 1$. ••

4 Representing Linear Functionals

This section is devoted to a representation which will soon be applied to λ of the previous section.

Proposition 4.1 *We are given $\mathcal{P}$ with a bounded positivity domain $\mathcal{D}_{\mathcal{P}}$. If $\lambda : M^s_{\ell}(\mathcal{N}_*) \longrightarrow \mathbb{R}$ (resp $\lambda : M^s_{\ell}(\mathcal{N}) \longrightarrow \mathbb{R}$, resp. $\lambda : M^s_{\ell}(\mathcal{N}_*\mathcal{N}) \longrightarrow \mathbb{R}$) is non-negative on $\mathcal{C}^{\ell}_{\mathcal{P}}$ and $\lambda(I) > 0$, then there exists a real Hilbert space H, a tuple $X = (X_1, \ldots, X_g)$ of bounded operators (resp. symmetric operators, resp. operators) on H, and a non-zero vector $\gamma \in \oplus^{\ell}_1 H$, the ℓ fold direct sum of H, such that $p(X) \succeq 0$ for all $p \in \mathcal{P}$ and for any symmetric $s \in M^s_{\ell}(\mathcal{N}_*)$ (resp $s \in M^s_{\ell\times\ell}(\mathcal{N})$, resp $s \in M^s_{\ell\times\ell}(\mathcal{N}_*\mathcal{N})$),*

$$\langle s(X)\gamma, \gamma\rangle = \lambda(s).$$

Proof for $M_l(\mathcal{N}_*)$ Case. Define a positive semi-definite symmetric bilinear form on $M_{1\times\ell}(\mathcal{N}_*)$ as follows. Given $1 \times \ell$ matrix-valued polynomials s, t with entries in $\mathcal{N}_*$, define

$$\langle s, t\rangle = \frac{1}{2}\lambda(t^T s + s^T t) \tag{4.1}$$

and verify that $\langle s, t\rangle$ is indeed symmetric and bilinear. It is positive semi-definite as $s^T s \in \mathcal{C}_{\mathcal{P}}$ and λ is positive. Let H be the Hilbert space formed by moding out $\langle\cdot,\cdot\rangle$-null vectors and forming the completion. Note H may be infinite dimensional. In what follows r will denote both $r \in M_{1\times\ell}(\mathcal{N}_*)$ and its equivalence class $[r] \in H$.

Recall our constant $C > 0$ such that $C^2 - x_j^T x_j \in \mathcal{C}_\mathcal{P}$ and $C^2 - x_j x_j^T \in \mathcal{C}_\mathcal{P}$. Since

$$C^2\langle s, s\rangle - \langle x_j s, x_j s\rangle = \lambda(s^T(C^2 - x_j^T x_j)s)$$

and since $s^T(C^2 - x_j^T x_j)s \in \mathcal{C}_\mathcal{P}$, it follows that multiplication by x_j on $M_{1\times\ell}(\mathcal{N}_*)$ defines a bounded operator X_j on H. Likewise

$$C^2\langle s, s\rangle - \langle x_j^T s, x_j^T s\rangle = \lambda(s^T(C^2 - x_j x_j^T)s)$$

implies multiplication by x_j^T on $M_{1\times\ell}(\mathcal{N}_*)$ defines a bounded operator Y_j on H. Since

$$\langle x_j s, t\rangle = \frac{1}{2}\lambda(t^T x_j s + s^T x_j^T t) = \frac{1}{2}\lambda((x_j^T t)^T s + s^T(x_j^T t)) = \langle s, x_j^T t\rangle$$

we have $Y_j = X_j^T$ and so the notation X_j^T for multiplication by x_j^T is unambiguous.

Now suppose $p \in \mathcal{C}_\mathcal{P}$ is $m\times m$ and symmetric. We shall be substituting X_j for x_j in $p(x)$ and this forces the substitution X_j^T for x_j^T in $p(x)$, since X_j^T is the adjoint of X_j^T. If r is an m vector where each entry is a $1\times\ell$ matrix-valued polynomial, then

$$\langle p(X)r, r\rangle = \sum_{a,b}\langle p_{a,b}(X)r_b, r_a\rangle = \lambda(\sum_{a,b} r_a^T p_{a,b} r_b) = \lambda(r^T p r) \geq 0,$$

where r is also canonically identified with an $m\times\ell$ matrix-valued polynomial and where the inequality results from $r^T p r \in \mathcal{C}_\mathcal{P}$. Hence $p(X) \succeq 0$.

Let γ_j denote the (equivalence class of the) constant polynomial $e_j^T \in \mathbb{R}^\ell$, where $\{e_1,\dots,e_\ell\}$ is the standard basis for $\mathbb{R}^\ell$. Note that $\sum \gamma_j^T\gamma_j = I$. Thus, in view of the hypothesis $\lambda(I) > 0$, there is a j_0 such that $\langle\gamma_{j_0},\gamma_{j_0}\rangle = \lambda(\gamma_{j_0}^T\gamma_{j_0}) > 0$. Hence the vector $\gamma = \oplus\gamma_j$ is non-zero. Finally, if s is a symmetric M_ℓ-valued polynomial, then

$$\langle s(X)\gamma, \gamma\rangle = \lambda(\sum \gamma_a^T s_{a,b}\gamma_b) = \lambda(s)$$

where the last equality takes into account that s is symmetric and that $(\gamma_a^T s_{a,b}\gamma_b)$ is the $\ell\times\ell$ matrix with $s_{a,b}$ in the (a,b) entry. This completes the proof for $M_\ell^s(\mathcal{N}_*)$.

Proof for $M_l(\mathcal{N})$ Case . The construction for the $M_l(\mathcal{N})$ case is very similar to that for the $M_l(\mathcal{N}_*)$ case. Given $1\times\ell$ matrix-valued polynomials s, t with entries in $\mathcal{N}$, define $\langle s, t\rangle$ as in (4.1) and verify that $\langle s, t\rangle$

is indeed symmetric and bilinear. It is positive semi-definite as $s^2 \in \mathcal{C}_{\mathcal{P}}$ and λ is positive on $\mathcal{C}_{\mathcal{P}}$. Let H be the Hilbert space formed by moding out $\langle \cdot, \cdot \rangle$-null vectors and forming the completion. Note H may be infinite dimensional. Once again the operator of multiplication by x_j on $M_{1\times\ell}(\mathcal{N})$ determines a well defined bounded operator X_j on H which is readily seen to be symmetric. The proof now proceeds as in the $\mathcal{N}_*$ case.

Proof for Hereditary $M_l(\mathcal{N}_*\mathcal{N})$ Case . Now we turn to the hereditary case. Given $1 \times \ell$ matrix-valued polynomials s, t with entries in $\mathcal{N}$ (so s, t contain no transposes and consequently $s^T r$ and $r^T s$ are hereditary), define $\langle s, t \rangle$ as in (4.1) and verify that $\langle s, t \rangle$ is indeed symmetric and bilinear. It is positive semi-definite as $s^T s \in \mathcal{C}_{\mathcal{P}}$ and λ is positive. Let H be the Hilbert space formed by moding out $\langle \cdot, \cdot \rangle$-null vectors and forming the completion. As before, multiplication by x_j on $M_{1\times\ell}(\mathcal{N})$ determines a well define bounded operator on H. In this case it is not true that X_j^T is multiplication by x_j^T, as $x_j^T s$ is not in $\mathcal{N}$. However, if P is an hereditary polynomial, $P = \sum P_{v,w} v^T w$, $X = (X_1, \ldots, X_g)$ is a tuple, and $\gamma = \oplus \gamma_j$ and $\delta = \oplus \delta_j$ are vectors, then

$$\langle P(X)\gamma, \delta \rangle = \sum P_{v,w} \langle X^w \gamma, X^v \delta \rangle$$

so that it is not actually necessary to have an explicit representation for X_j^T.

From this point the proof is exactly as it was for the $M_l(\mathcal{N}_*)$ case, keeping in mind that all products are hereditary. ••

5 Proof of Theorem 1.2

Let M denote either either $M_\ell^s(\mathcal{N}_*)$, $M_\ell^s(\mathcal{N})$, or $M_\ell^s(\mathcal{N}_*\mathcal{N})$. Suppose $q \in M$ and $q \notin \mathcal{C}_{\mathcal{P}}$. By Proposition 3.3 there is a linear functional $\lambda : M \longrightarrow \mathbb{R}$ such that λ is non-negative on $\mathcal{C}_{\mathcal{P}}$, $\lambda(I) > 0$, and $\lambda(q) \le 0$.

By Proposition 4.1, there exists a Hilbert space H, a non-zero vector γ in $\oplus_1^\ell H$ and a tuple of operators $X = (X_1, \ldots, X_g)$ on H such that $p(X) \succeq 0$ for all $p \in \mathcal{P}$ and for each symmetric M_l-valued polynomial s, we have $\langle s(X)\gamma, \gamma \rangle = \lambda(s)$. In particular, substituting q for s gives, $\langle q(X)\gamma, \gamma \rangle = \lambda(q) \le 0$. Since γ is non-zero, it follows that there is an $X \in \mathcal{D}_{\mathcal{P}}$ such that $q(X)$ is not strictly positive definite and this proves the contrapositive of Theorem 1.2.

6 No NC Real Nullstellensatz

The Real Nullstellensatz for commutative polynomials, cf. Corollary 4.1.8 to Theorem 4.1.4 [BCR], says that if p and q are polynomials on $\mathbb{R}$ satisfying

$$p(x) = 0 \quad \textit{implies} \quad q(x) = 0,$$

then there is an integer m and a polynomial S which is a sum of squares such that $q^{2m} + S$ lies in the ideal generated by p. This paper has focused on non-commutative polynomials in x and x^T. In this section we show that a rather weak non-commutative Real Nullstellensatz for such polynomials is false. This is a bit of a surprise, since in non-commutative situations analyzed so far, for example, sums of squares [H] [M] [MP] the non-commutative case is better behaved than the commutative case. Also there is a strong Nullstellensatz for polynomials in non-commutative variables x without transposes, whose proof was generously provided to us by George Bergman. Details are in subsection 6.2.

As a summary we compare the NC situation to a standard commutative version of the Real Positivstellensatz in Corollary 4.4.3 (i) (ii) (iii)in [BCR]. The counter example below refutes a natural non-commutative extension of (i) and similarly (ii). While Theorem 1.2 is an NC variant on (iii).

6.1 Polynomials in x and x^T

One possible weak version of a non-commutative Real Nullstellensatz goes like this. Given p and q symmetric NC polynomials from $\mathcal{N}_*$, if for any tuple $X = (X_1, \dots, X_g)$ acting on the Hilbert space H and vector $v \in H$ the condition $p(X)v = 0$ implies $q(X)v = 0$, does it follow that

$$q^{2m} + S = pr + r^T p$$

for some positive integer m, polynomial $r \in \mathcal{N}_*$, and S a sum of squares so that $S = \sum r_j^T r_j$ for $r_j \in \mathcal{N}_*$?

The following example shows this is false.

Example 6.1 *Let $p = (x^T x + x x^T)^2$ and $q = x + x^T$ where x is a single variable, that is, $g = 1$. Then $p(X)v = 0$ implies $q(X)v = 0$; however, there does not exist a positive integer m and $r, r_j \in \mathcal{N}_*$ so that*

$$q^{2m} + \sum r_j^T r_j = r^T p + pr. \tag{6.1}$$

Moreover, we can modify the example to add the condition $p(X)$ is positive semidefinite implies $q(X)$ is positive semidefinite and still not obtain this representation.

Proof. Since $A := XX^T + X^TX$ is self-adjoint, $A^2v = 0$ if and only if $Av = 0$. It now follows that if $p(X)v = 0$, then $Xv = 0 = X^Tv$ and therefore $q(X)v = 0$.

For $\lambda \in \mathbb{R}$, let

$$X = X(\lambda) = \begin{pmatrix} 0 & \lambda & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

viewed as an operator on $\mathbb{R}^3$ and let $v = e_1$, where $\{e_1, e_2, e_3\}$ is the standard basis for $\mathbb{R}^3$.

We begin by calculating the first component of even powers of the matrix $q(X)$. Let $Q = q(X)^2$ and verify,

$$Q = \begin{pmatrix} \lambda^2 & 0 & \lambda \\ 0 & 1+\lambda^2 & 0 \\ \lambda & 0 & 1 \end{pmatrix}. \tag{6.2}$$

For each positive integer m there exist a polynomial q_m so that

$$Q^m e_1 = \begin{pmatrix} \lambda^2(1+\lambda q_m(\lambda)) \\ 0 \\ \lambda(1+\lambda q_m(\lambda)) \end{pmatrix} \tag{6.3}$$

which we now establish by an induction argument. In the case $m = 1$, from equation (6.2), it is evident that $q_1 = 0$. Now suppose equation (6.3) holds for m. Then, computation of $QQ^m e_1$ shows that equation (6.3) holds for $m+1$ with $q_{m+1} = \lambda(q_m + \lambda + \lambda q_m)$. Thus, for any m,

$$\lim_{\lambda \to 0} \frac{1}{\lambda^2} < Q^m e_1, e_1 >= \lim_{\lambda \to 0}(1 + \lambda q_m(\lambda)) = 1. \tag{6.4}$$

Now we look at p and get

$$p(X) = \begin{pmatrix} \lambda^4 & 0 & 0 \\ 0 & (1+\lambda^2)^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus

$$\lim_{\lambda \to 0} \frac{1}{\lambda^2} \left(< r(X)^T p(X) e_1, e_1 > + < p(X) r(X) e_1, e_1 > \right) = 0.$$

If the representation of equation (6.1) holds, then apply $< \cdot\, e_1, e_1 >$ to both sides and take λ to 0. We just saw that the right side is 0, so the left side is 0, which because

$$< \sum r_j(X)^T r_j(X) e_1, e_1 > \geq 0$$

forces

$$\lim_{\lambda \to 0} \frac{1}{\lambda^2} < Q^m e_1, e_1 > \leq 0$$

a contradiction to equation (6.4). Hence the representation of equation (6.1) does not hold.

The last sentence claimed in the example is true when we use the same polynomial p and replace q with q^2. ••

6.2 The Transpose Free Nullstellensatz

Now we look at the case where p and q are polynomials purely in x and the "matrix zero set" of q contains that of p. This gives a satisfying result conjectured by the authors and proved by George Bergman.

Theorem 6.2 *Fix a finite collection $\mathcal{P}$ of polynomials in non-commuting variables $\{x_1, \dots, x_g\}$ and let q be a given polynomial in $\{x_1, \dots, x_g\}$. Let d denote the maximum of the deg(q) and $\{deg(p) : p \in \mathcal{P}\}$. There exists a real Hilbert space $\mathcal{H}$ of dimension $\sum_{j=0}^{d} g^j$, such that, if*

$$q(X)v = 0$$

whenever $X = (X_1, \dots, X_g)$ is a tuple of operators on $\mathcal{H}$, $v \in \mathcal{H}$, and

$$p(X)v = 0 \text{ for all } p \in \mathcal{P},$$

then q is in the left ideal generated by $\mathcal{P}$.

We sketch a slight variant of George Bergman's proof.

Proof. Let P denote the $\mathbb{R}$ algebra of polynomials in the non-commuting variables $\{x_1, \dots, x_g\}$ and let I denote the left ideal generated by $\mathcal{P}$. Define X_j on the vector space P/I by $X_j[p] = [x_j p]$, where $[p]$ denotes the equivalence class of $p \in P$ in the quotient P/I. Thus, X_j is determined by the left regular representation. If P/I is finite dimensional, then $X = (X_1, \dots, X_g)$ can be viewed as a tuple of matrices and we have, for $p \in P$,

$$p(X)[1] = [p].$$

In particular, if $p \in \mathcal{P} \subset I$, then $p(X)[1] = 0$. Hence, if we assume that $p(X)v = 0$ for all $p \in \mathcal{P}$ with X a tuple acting on a Hilbert space of dimension at least as large as the dimension of P/I, then $0 = q(X)[1] = [q]$ and therefore $q \in I$. Minus the precise statement about the dimension of $\mathcal{H}$ this establishes the result when P/I is finite dimensional.

For the general case, let V denote the vector space

$$\{[p] : p \in P, \deg(p) \leq d\}.$$

Note that the dimension of V is at most $\sum_{j=0}^{d} g^j$. Let W denote the subspace $\{[p] : p \in P, \deg(p) < d\}$ of V and choose a basis $\{f_1, \ldots, f_m\}$ for W. Extend this basis to a basis $\{f_1, \ldots, f_m, f_{m+1}, \ldots, f_{n+k}\}$ for V. There exists polynomials $r_1, \ldots, r_m$ of degree at most $d-1$ so that $f_\ell = [r_\ell]$. For $1 \leq j \leq g$, define $X_j[r_\ell] = [x_j r_\ell]$ for $1 \leq \ell \leq m$ and $X_j f_{m+\ell} = 0$ for $1 \leq \ell \leq k$. If p has degree at most $d-1$, then $X_\ell[p] = [x_\ell p]$. Thus, if p has degree at most d, then $p(X)[1] = [p]$. The proof now proceeds just as in the previous paragraph. ••

7 Thanks

We are grateful to Jeff Ovall for his careful reading of the manuscript, for the many comments, and for supplying the examples at the beginning of the paper. Thanks are due to Mihai Putinar for very helpful conversations, and to George Bergman for letting us include his result.

8 References

[A] Agler, Jim On the representation of certain holomorphic functions defined on a polydisc. Topics in operator theory: Ernst D. Hellinger memorial volume, 47–66, Oper. Theory Adv. Appl., 48, Birkhauser, Basel, 1990.

[AM] Agler, Jim and McCarthy John Featured talk by McCarthy at SEAM in Athens GA, 2001.

[BMprep] Ball, Joseph and ?? Malakorn, Conservative Structured Realizations ... in preperation

[BCR] Bochnak, Jacek; Costi, Michel and Roy, Marie- Francoise *Real Algebraic Geometry* Springer 1991, pp 430.

[H] Helton, J. William Positive non commutative polynomials are sums of squares, Annals of Math vol 56 num 2, Sept 2002 pp. 675-694.

[M] McCullough, Scott Factorization of operator-valued polynomials in several non-commuting variables. Linear Algebra Appl. **326** (2001), no. 1-3, 193–203.

[MP] McCullough, Scott and Putinar, Mihai Non-commutative Sums of Squares preprint

[PV] Putinar, Mihai and Vasilescu, Florian-Horia Solving moment problems by dimensional extension. Annals of Math. (2) **149** (1999), no. 3, 1087–1107.

[R] Rudin, Walter Functional analysis, second edition, International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, 1991. xviii+424 pp.

Department of Mathematics
University of California
San Diego, CA 92093, USA
E-mail: helton@osiris.ucsd.edu
WWW URL: http://www.ucsd.edu

Department of Mathematics
University of Florida
Gainesville, Florida 32611-8105
E-mail: sam@math.ufl.edu

NOT FOR PUBLICATION

Contents