Non-negative hereditary polynomials in a free $*-$algebra

J.William Helton * Scott A. McCullough † Mihai Putinar ‡

December 8, 2003

Abstract

We prove a non-negative-stellensatz and a null-stellensatz for a class of polynomials called hereditary polynomials in a free $*$-algebra.

Mathematics Subject Classification: 46A55, 06F25, 41A63

1 Introduction and main result

This note is concerned with sums of squares decompositions in a free $*$-algebra. Already quite a few facts about such decompositions are known, see [3, 4, 5] and the references cited there. In particular, the article [4] contains a simple example showing that no analogue of the real nullstellensatz, as known in real algebraic geometry, exist. Another divergence between the commutative and the free $*$ theories was singled out in [5], where a stronger than expected non-negative-stellensatz with supports was proved. We consider below another fortunate free case when both a simple non-negative-stellensatz and a nullstellensatz hold. The proofs are based on the same principles as before: Caratheodory's theorem about convex hulls, Minkovski's separation theorem and a general Gelfand-Naimark-Segal construction.

The notation of the article [5] will be used throughout the note. Specifically, $\mathcal{P}$ is the space of polynomials in the non-commuting variables $(x, x^T) =$

*Partially supported by NSF, DARPA and Ford Motor Co.
†Partially supported by NSF grant DMS-0140112
‡Partially supported by NSF grant DMS-0100367

$\{x_1, x_2, \ldots, x_n, x_1^T, \ldots, x_n^T\}$. $\mathcal{P}_v$ is the finite dimensional subspace consisting of all polynomials of degree less than or equal to v.

Call a polynomial **analytic** provided it contains no transposes. The corresponding space is denoted $\mathcal{A}$. Similarly, call a polynomial **hereditary** provided all transposes x_k^T appear to the left of x_j's in any monomial. Denote the corresponding space by $\mathcal{H}$. We introduce notation for various types of polynomials of degree v:

$$\mathcal{A}_v = \mathcal{A} \cap \mathcal{P}_v$$

and let $\mathcal{H}_{2v}$ be the vector space generated by $\mathcal{A}_v^T \mathcal{A}_v$. Thus, $\mathcal{H}_{2v}$ is the vector space of polynomials of the form

$$q = \sum_{u,w} q_{u,w} u^T w$$

where the sum is over words of length at most v is the variables x.

In the following $p_1, ..., p_m \in \mathcal{A}$ will always denote analytic polynomials. The **left ideal** generated by them is

$$(p) := \{\sum r_j p_j : r_j \in \mathcal{A}\};$$

the associated **left symmetrized ideal** is:

$$sym(p) := \{\sum r_j^T p_j + p_j^T r_j : r_j \in \mathcal{A}\}.$$

The main result of this paper is the following theorem. A few extensions are given in Section 3.

Theorem 1.1 *Let $p_1, ..., p_m \in \mathcal{A}$ be analytic polynomials. If a symmmetric hereditary polynomial $q \in \mathcal{H}$ has*

$$\langle q(X)v, v\rangle \geq 0$$

on all pairs (X, v) satisfying $p_\ell(X)v = 0$ $(1 \leq \ell \leq m)$, then:

$$(a) \qquad q = \sum_{j=1}^{k} f_j^T f_j + g, \tag{1}$$

where $g \in sym(p)$ and each f_j is analytic.

(b) If instead, $\langle q(X)v, v\rangle = 0$, for all (X, v) satisfying $p_\ell(X)v = 0$ for all ℓ, then $q \in sym(p)$.

2 Proofs

2.1 Zeroes and Ideals

Denote by V_p the (X, v) zero set of the p_j's, that is the set of all pairs of tuples of (finite) matrices and vectors (X, v) satisfying $p_\ell(X)v = 0$ for all $\ell, 1 \leq \ell \leq m$. We stress that the rank of X or the dimension of the vector space where v belongs are finite, but free otherwise.

Lemma 2.1 *If $a \in \mathcal{A}$ and $V_p \subset V_a$, then $a = \sum r_j p_j$ for some analytic polynomials $r_j \in \mathcal{A}$. If q and the p_j have degree no greater than d, then $V_p \subset V_a$ can be taken to be g-tuples in $R^{n\times n}$ with $n = (d+1)^2$.*

Proof: The proof of this lemma, due to G. Bergman, is contained in [4]. ■

Lemma 2.2 *If the $f_j \in \mathcal{A}$ are analytic polynomials and*

$$\sum_{j=1}^{k} f_j^T f_j + sym(p) = 0 \quad on\ V_p,$$

then $f_j \in (p)$.

Proof: The conditions $p_j(X)v = 0$ imply

$$\sum \langle r_j(X) p_j(X) v, v\rangle = 0 \quad \text{and} \quad \sum \langle p_j(X)^T r_j(X)^T v, v\rangle = 0$$

which imply $\sum_{j=1}^{k} \langle f_j(X)^T f_j(X) v, v\rangle = 0$. Thus $f_j(X)v = 0$. Now Lemma 2.1 yields the conclusion. ■

Proof of Theorem 1.1 part (b). Assume Theorem 1.1 part (a) is true. By Lemma 2.2 we get $f_j \in\ (p)$. Thus

$$q \in \sum_{j=1}^{k} f_j^T f_j + \ sym(p)$$

becomes $q \in sym(p)$. ■

Lemma 2.3 *There exist N and points (X_k, v_k), $1 \leq k \leq N$, in V_p that is, $p(X_k)v_k = 0$, such that the function $S : \mathcal{H}_{2d} \to R$ defined on h of the form*

$$h = \sum r_j^T q_j \quad with \quad r_j, q_j \in \mathcal{A}$$

by

$$S(h) = \sum_{j=1}^{n} \sum_{k=1}^{N} \langle r_j(X_k)v_k, q_j(X_k)v_k \rangle, \tag{2}$$

has the property

$$S(f^T f) = 0 \text{ implies } f \in (p)$$

for any $f \in \mathcal{A}_d$.

Proof: Let W_1 denote the closed unit sphere in the finite dimensional space $W = \mathcal{A}_d/(p)_d$, endowed with an Euclidean norm. Given $h \in W_1$, since h is not in (p), by Lemma 2.1 we can pick $(X_h, v_h) \in V_p$, so that

$$\langle r(X_h)^T r(X_h)v_h, v_h \rangle_{\mathcal{R}_{2d}} > 0$$

for all r in an open neighborhood $\mathcal{O}$ of h in W_1. By the compactness of W_1 there exists a finite cover of such open sets $\mathcal{O}_k$ for $k = 1, \cdots, N$, which in turn is associated with a finite set $(X_k, v_k) \in V_p$. Use these points in definition (2) of S. Then for any given h not in (p), a positive multiple of it lies in some open set $\mathcal{O}_k$, so $S(h^T h) \geq \langle h(X_k)v_k, h(X_k)v_k \rangle_{\mathcal{R}_{2d}} > 0$. ■

2.2 Proof of Theorem 1.1

Fix a large enough degree d. For instance the assumption

$$d > 4 \max(\deg(p_1), ..., \deg(p_n), \deg(q))$$

would be sufficient for the following proof.

Denote by $\mathcal{R}_{2d}$ the subset of $\mathcal{P}_{2d}$ consisting of sums of the form

$$\sum_{j=1}^{k} f_j^T f_j + sym(p)_{2d}, \tag{3}$$

with $f_j \in \mathcal{A}_d$ and the integer k arbitrary.

2.2.1 Seperation

If the hereditary polynomial q in the statement of Theorem 1.1 admits a decomposition (1), then there exists a sufficiently large d with $q \in \mathcal{R}_{2d}$.

Henceforth we assume by contradiction that $q \notin \mathcal{R}_{2d}$.

Lemma 2.4 will show that $\mathcal{R}_{2d}$ is closed. By Minkowski's separation theorem there exists a linear functional L_1 on $\mathcal{P}_{2d}$ satisfying:

$$L_1(q) < 0 \leq L_1(c), \quad c \in \mathcal{R}_{2d}. \tag{4}$$

Note that $L_1(sym(p)) = 0$, since $L_1(sym(p)) \geq 0$ and $sym(p)$ is a vector subspace. We modify L_1 to $L := L_1 + \epsilon S$ with S the function defined in Lemma 2.3 and with $\epsilon > 0$, chosen small enough to make

$$L(q) < 0 \leq L(c), \quad c \in \mathcal{R}_{2d}, \tag{5}$$

still hold true. In addition, L has the critical property: if $h \in \mathcal{A}_d$, then

$$L(h^T h) = 0 \text{ implies } h \in (p).$$

2.2.2 The Hilbert Space

We consider again the vector space

$$W = \mathcal{A}_d/(p)_d$$

and denote by $[f]$ or simply f the class of $f \in \mathcal{A}_d$ in W.

Next we define a symmetric positive semi-definite form on W by

$$\langle [a], [b] \rangle = \frac{1}{2} L(a^T b + b^T a).$$

We must check that the definition is independent of choice of representatives a, b:

$$L((a + rp)^T (b + sp) + (b + sp)^T (a + rp)) =$$
$$= L(a^T b + b^T a) + L(a^T sp + (sp)^T a) + L((rp)^T b + b^T rp) + L((rp)^T sp + (sp)^T rp).$$

The last three terms have the form $L(sym(p))$, so are 0; thus we have a well defined inner product on W. The inner product is strictly positive definite by the choice of L. Thus W is a Hilbert space.

2.2.3 Matrices

Let W' be the image in W of $\mathcal{A}_{d-1}$. Let $X_j : W' \longrightarrow W$ be the left multiplication by the variable x_j. This is well defined due to the degrees assumptions and the fact that $[m] \in (p)$ implies $[x_j m] \in (p)$. Extend arbitrarily X_j to a linear transformation from W into W, and denote the extension by the same letter. Then the adjoint X_j^T with respect to the hilbertian structure of W is unambiguously defined.

Suppose that our symmetric hereditary polynomial has the form $q = \sum_k (g_k^T h_k + h_k^T g_k)/2$ with $\deg g_k, \deg h_k < d - 1$ for all k. Then:

$$L(q) = \sum_k \langle g_k, h_k \rangle = \sum_k \langle g_k(X)1, h_k(X)1 \rangle =$$

$$1/2\sum_k \langle (h_k(X)^T g_k(X) + g_k(X)^T h_k(X))1, 1\rangle = \langle q(X^T, X)1, 1\rangle.$$

Thus $\langle q(X^T, X)1, 1\rangle = L(q) < 0$ and at the same time

$$\langle p_k(X)1, 1\rangle = \langle [p_k], 1\rangle = 0,$$

for all k.

In conclusion the pair $(X, 1)$ contradicts the assumption in the statement of Theorem 1.1 and the proof is complete.■

2.2.4 The Cone is Closed

Lemma 2.4 *For every $d \geq 0$ the cone $\mathcal{R}_{2d}$ is closed in $\mathcal{P}_{2d}$.*

Proof: This is an application of Carathéodory's theorem, see for instance [9] for a similar derivation in the commutative case. Namely, if $k-1$ denotes the dimension of $\mathcal{H}_{2d}$, then Carathéodory's theorem says every element h in the convex cone $\mathcal{R}_{2d}$ can be written as a combination of at most k elements from the set $\{f^T f + g : f \in \mathcal{P}_d,\ g \in sym(p)\}$ which generates it. Thus, there exist $f_1, \ldots, f_k \in \mathcal{P}_d$ such that

$$h = f_1^T f_1 + f_2^T f_2 + \ldots + f_k^T f_k + w.$$

with $w \in sym(p)$.

Suppose $h^\nu \in \mathcal{R}_{2d}$ is a Cauchy sequence in the topology of $\mathcal{P}_{2d}$. For each (positive integer) ν pick any $f_1^\nu, \ldots, f_k^\nu$ (some of them possibly equal to zero) in $\mathcal{A}_d$ and w^ν in $sym(p)$ such that

$$h^\nu = \sum_{j=1}^{k} f_j^{\nu T} f_j^\nu + w^\nu. \tag{6}$$

Now h^ν is bounded, and if the f_j^ν are also bounded we can choose convergent subsequences and the proof finishes.

Here we use the fact that $\mathcal{A}_d$ is finite dimensional so that the bilinear mapping $\mathcal{A}_d \times \mathcal{A}_d \to \mathcal{H}_{2d}$ given by $(f, g) \mapsto g^T f$ is continuous.

To prove the lemma without assuming that the f_j^ν and w^ν are bounded sequences, first recall the linear functional $S : \mathcal{H}_{2d} \to \mathbb{R}$ from Lemma 2.3. Since S is a linear map on the finite dimensional vector space $\mathcal{H}_{2d}$ and since the sequence $\{h^\nu\}$ is bounded, it follows that

$$S(h^\nu) = \sum_{j=1}^{k} \sum_{\ell} \|f_j^\nu(X_\ell) v_\ell\|^2 \tag{7}$$

is bounded. Recall also that

$$\|[f]\|^2 = \sum_k \|f(X_\ell)v_\ell\|^2$$

defines a norm on the quotient $\mathcal{W}_d = \mathcal{A}_d/(p)$. Thus, equation (7) implies that each of the sequences $[f_j^\nu]$, $j = 1, 2, \ldots, k$, is bounded in $\mathcal{W}_d$ (with respect to any norm) and therefore we may assume, by passing to a subsequence if necessary, that each of these sequences is Cauchy. Since $\mathcal{W}_d$ is finite dimensional it is complete and thus each of the sequences $\{[f_j^\nu]\}$ converges. Let $[f_1], \ldots, [f_k]$ denote the limits which are independent of the norm on $\mathcal{W}_d$. Implicitly we have chosen representatives of the equivalence classes.

The mapping $\mathcal{A}_d \to \mathcal{W}_d$ given by $f \mapsto [f]$ is linear and onto and therefore has a right inverse T. Consequently, the sequences $\{g_j^\nu = Tf_j^\nu\}_\nu$ are bounded for each $j = 1, \ldots, k$. Since $f_j^\nu = g_j^\nu + r_j^\nu$ for $r_j^\nu \in (p)$,

$$h^\nu = \sum (g_j^\nu)^T g_j^\nu + u^\nu$$

for some u^ν in $sym(p)$. ■

3 Ramifications of the main result

This last part of the note contains a couple of generalizations of Theorem 1.1.

3.1 Analytic modules

The main result above can easily be adapted to more general polynomial sets. For instance consider a left $\mathcal{A}$-module $\mathcal{M} \subset \mathcal{P}$ which contains 1, hence contains $\mathcal{A}$.

Take elements $p_1, ..., p_m \in \mathcal{M}$ and consider the left $\mathcal{P}$-submodule $I = \mathcal{A}p_1 + \mathcal{A}p_2 + ... + \mathcal{A}p_m$ and its symmetrized form:

$$sym(I) = \{\sum_j (r_j^T p_j + p_j^T r_j; \quad r_j \in \mathcal{P}\}.$$

Let h be a $\mathcal{M}$-hereditary element, that is:

$$h = \sum_j f_j^T g_j, \quad f_j, g_j \in \mathcal{M}.$$

Instead of Lemma 2.2 we assume that,

A. *If a sum of squares*

$$f = \sum_k f_k^T f_k, \quad f_k \in \mathcal{M}$$

vanishes on all pairs (X, v) *which themselves satisfy the condition* $p_j(X)v = 0, \quad 1 \le j \le m$, *then each* $f_k \in I \cap \mathcal{M}$.

With these notations and assumptions the following variant of Theorem 1.1 is valid.

Proposition 3.1 *Let* $\mathcal{M} \subset \mathcal{P}$ *be an* $\mathcal{A}$ *submodule containing* 1 *and let* I *be a left* $\mathcal{P}$ *ideal generated by (finitely many) elements of* $\mathcal{M}$. *Suppose that* $\mathcal{M}$ *satisfies condition A.*

If a symmetric $\mathcal{M}$*-hereditary element* h *satisfies*

$$\langle h(X)v, v\rangle \ge 0, \text{ whenever } \ I(X)v = 0,$$

then there are $f_j \in \mathcal{M}, \ 1 \le j \le N$, *such that*

$$h \equiv \sum_k f_k^T f_k \mod(symI).$$

Note that assumption A on I is parallel to the notion of a real ideal in the commutative case, see [2].

3.2 Complex coefficients

The case of a free complex $*$-algebra with antilinear involution is similar, with one little improvement in the second part of Theorem 1.1.

Namely we consider the $*$-algebra $\mathcal{P} \otimes_{\mathbf{R}} \mathbf{C}$ with involution $f \mapsto f^*$. The definitions of an analytic, respectively hereditary element are the same.

Proposition 3.2 *Let* I *be a left ideal of* $\mathcal{A} \otimes_{\mathbf{R}} \mathbf{C}$ *and let*

$$V(I) = \{(X, v); \ f(X)v = 0, \ for \ all \ f \in I\}$$

be its zero set.

If an element $q \in \mathcal{P} \otimes_{\mathbf{R}} \mathbf{C}$ *vanishes on* $V(I)$, *that is* $\langle q(X)v, v\rangle = 0$ *for all* $(X, v) \in V(I)$, *then* $q \in sym(I)$.

The proof can be reduced to the self-adjoint case treated by the analogue of Theorem 1.1 via the decomposition $2q = (q + q^*) + i[(q - q^*)/i]$.

References

[1] J. Agler, *An abstract approach to model theory*, Surveys of some recent results in operator theory, Vol. II, 1–23, Pitman Res. Notes Math. Ser., **192**, Longman Sci. Tech., Harlow, 1988.

[2] J. Bochnack, M. Coste and J.-F. Roy, *Géométrie Algébrique Réele*, Springer, New York, 1987.

[3] J.W. Helton, *Positive non commutative polynomials are sums of squares* Ann. Math **56**(2002, 675-694.

[4] J.W. Helton and S. McCullough, *A Positivstellensatz for Noncommutative Polynomials*, Trans. Amer. Math. Soc., to appear.

[5] J.W. Helton, S. McCullough, M.Putinar, *A noncommutative Positivstellensatz on isometries*, J. reine angew. Math., to appear.

[6] A.Prestel and C.N.Delzell, *Positive Polynomials*, Springer, Berlin, 2001.

[7] C. Procesi and M. Schacher, *A noncommutative real Nullstellensatz and Hilbert's 17th problem*, Ann. of Math. (2) **104** (1976), no. 3, 395–406.

[8] M. Reed and B. Simon, *Methods of Modern Mathematical Physics Vol. 1: Functional Analysis*, Academic Press, San Diego, 1980.

[9] B. Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc. **96** (1992) No. **463**, Amer. Math. Soc., Providence, R.I.

[10] C. Scheiderer, *Positivity and sums of squares: a guide to some recent results*, preprint 2003.

[11] G. Stengle, *A Nullstellensatz and a Positivstellensatz in semialgebraic geometry*, Math. Ann. **289**(1991), 203-206.

J. W. Helton, Department of Mathematics, University of California at San Diego, La Jolla CA 92093
helton@osiris.ucsd.edu

S.A. McCullough, Department of Mathematics, University of Florida, Gainesville, FL 32611-8105
sam@mail.math.ufl.edu

M. Putinar, Department of Mathematics, University of California, Santa Barbara, CA 93106
mputinar@math.ucsb.edu