

# NON-COMMUTATIVE SUMS OF SQUARES

SCOTT McCULLOUGH<sup>1</sup> AND MIHAI PUTINAR<sup>2</sup>

**A proof of Helton's SoS theorem based upon a theorem of Carathéodory and a Hahn-Banach separation argument is presented.**

## 1. Introduction

Fix a positive integer  $g$ , let  $\mathcal{F}$  denote the free semi-group on the  $2g$  letters of the alphabet  $A = \{x_1, x_2, \dots, x_g, y_1, y_2, \dots, y_g\}$ , and let  $\mathcal{A}$  denote the free semi-group  $\mathbb{C}$ -algebra on  $A$ . Concretely, an element  $p$  of  $\mathcal{A}$  is a linear combination of elements from  $\mathcal{F}$ ,

$$(1) \quad p = \sum p_w w,$$

where the sum is finite and  $p_w \in \mathbb{C}$ , and is referred to as a polynomial in  $A$ . Note that the empty word  $\emptyset$  is the multiplicative identity (and  $0$ , the empty sum, is the additive identity).

Equip  $\mathcal{A}$  with the involution  $*$  determined as follows. On letters,  $x_j^* = y_j$ ,  $y_j^* = x_j$ ; on a word  $w = w_1 \cdots w_n \in \mathcal{F}$ ,

$$(2) \quad w^* = w_n^* \cdots w_2^* w_1^*;$$

and finally, on a polynomial  $p$  in  $A$  as in (1)

$$(3) \quad p^* = \sum p_w^* w^*,$$

where  $p_w^*$  is the complex conjugate of the complex number  $p_w$ .

Let  $\mathcal{B}(\mathcal{H})$  denote the bounded operators on the complex Hilbert space  $\mathcal{H}$ . Evaluation at a tuple  $X = (X_1, \dots, X_g)$  from  $\mathcal{B}(\mathcal{H})$  determines a representation  $\pi_X : \mathcal{A} \rightarrow \mathcal{B}(\mathcal{H})$  which respects the involution. Explicitly, define  $\pi_X(x_j) = X_j$ ,  $\pi_X(y_j) = X_j^*$ , and extend  $\pi_X$  to an algebra homomorphism. It is evident that  $\pi_X(p^*) = \pi_X(p)^*$ . We will write  $p(X)$  instead of  $\pi_X(p)$ .

The purpose of this note is to give a proof of Helton's sum of squares (SoS) theorem [H] based upon a theorem of Carathéodory and a Hahn-Banach separation argument. If  $r_j \in \mathcal{A}$  for  $j = 1, \dots, m$ ,  $p = \sum r_j^* r_j$ , and  $X$  is a tuple of operators, then  $p(X) = \sum r_j(X)^* r_j(X) \geq 0$ , where the notation  $T \geq 0$  is used to indicate the Hilbert space operator  $T$  is positive semi-definite. Given a non-negative integer  $d$ , let  $\mathcal{A}_d$  denote the elements of  $\mathcal{A}$  of degree at most  $d$ . That is  $\mathcal{A}_d$  consists of those  $p$  of the form (1) where

the sum is over words of length at most  $d$ . Let  $N(d)$  denote the dimension of  $\mathcal{A}_d$  (as a  $\mathbb{C}$  vector space).

**Theorem 1.1.** (*Helton*) *Let  $\mathcal{H}$  be a Hilbert space of dimension  $N(d)$  and  $p \in \mathcal{A}_d$  satisfy  $p(X) \geq 0$  for all tuples  $X = (X_1, \dots, X_g)$  from  $\mathcal{B}(\mathcal{H})$ .*

*Then there exists  $r_j \in \mathcal{A}$ ,  $j = 1, 2, \dots, N(d)$ , such that  $p = \sum r_j^* r_j$ .*

REMARK 1.2. Helton states and proves his theorem over  $\mathbb{R}$ , rather than  $\mathbb{C}$ . The interested reader should have no difficulty making the necessary modifications to both the statement and proof of (1.1) to accommodate real scalars. The approach here can also be adapted to deal with the case of self-adjoint variables where  $x_j^* = x_j$  [M].

The authors thank Bill Helton for his generous help and encouragement in the work detailed in this paper and acknowledge borrowing freely from the ideas in [H].

## 2. Carathéodory's Theorem

Let  $\mathcal{C}_d$  denote the convex hull of  $\{r^* r : r \in \mathcal{A}_d\}$ , and observe that  $\mathcal{C}_d$  is a subset of  $\mathcal{A}_{2d}$ .

Carathéodory's theorem asserts that a vector belonging to the convex hull of a system of points in an  $n$  dimensional real vector space can be written as a convex combination of at most  $n + 1$  of these points. The proof is elementary and can be found for instance in [5]. In the same spirit we prove the following more precise decomposition result for the cone  $\mathcal{C}_d$ .

**Theorem 2.1.** *If  $p \in \mathcal{C}_d$ , then there exists an  $m \leq N(d)$  and  $r_j \in \mathcal{A}_d$ ,  $j = 1, 2, \dots, m$ , such that  $p = \sum r_j^* r_j$ .*

Sketch of proof. Let  $N = N(d)$ . Index  $\mathbb{C}^N$  and  $\mathcal{A}_d^N$  - the algebraic direct sum of  $\mathcal{A}_d$  with itself  $N$  times - by  $\mathcal{F}_d$ , the elements of  $\mathcal{F}$  of length at most  $d$ . Let  $V \in \mathcal{A}_d^N$  denote the border vector [H]; i.e., the vector whose  $w \in \mathcal{F}_d$  entry is  $w$  (thought of as a column). Given  $r = \sum r_w w \in \mathcal{A}_d$ , let  $R$  denote the (column) vector with  $w$  entry  $r_w^*$ . Then

$$(4) \quad r = R^* V.$$

Since  $p \in \mathcal{C}_d$ , there exists an  $M$  and  $r_1, \dots, r_M \in \mathcal{A}_d$  such that  $p = \sum r_j^* r_j$ . Let  $R_j$  denote the corresponding vectors from  $\mathbb{C}^N$  and let  $Q = \sum R_j R_j^*$ . Then  $Q \geq 0$  and

$$(5) \quad p = V^* Q V.$$

Since  $Q$  is positive semi-definite and  $N \times N$ , there exists vectors  $Q_j \in \mathbb{C}^N$  such that  $Q = \sum_{j=1}^N Q_j Q_j^*$ . Let  $q_j = Q_j^* V$  and verify

$$(6) \quad p = V^* Q V = \sum V^* Q_j Q_j^* V = \sum q_j^* q_j.$$

□

### 3. Positive Linear Functionals

In this section we consider positive linear functionals on  $\mathcal{A}_{2d}$ ; i.e., functionals  $\lambda : \mathcal{A}_{2d} \rightarrow \mathbb{C}$  such that  $\lambda(p^*p) > 0$  for all nonzero  $p \in \mathcal{A}_d$ .

**Lemma 3.1.** *Suppose  $\lambda : \mathcal{A}_{2d+2} \rightarrow \mathbb{C}$  is a linear. If  $\lambda(p^*p) > 0$  whenever  $p \in \mathcal{A}_{d+1}$  is nonzero, then there exists a Hilbert space  $\mathcal{H}$  of dimension  $N(d)$ , a vector  $\gamma \in \mathcal{H}$ , and tuple  $X$  from  $\mathcal{B}(\mathcal{H})$  such that  $\langle p(X)\gamma, q(X)\gamma \rangle = \lambda(q^*p)$  for all  $p, q \in \mathcal{A}_d$ .*

Proof. Let  $\mathcal{K}$  denote the Hilbert space obtained by defining the inner product  $\langle p, q \rangle = \lambda(q^*p)$  on  $\mathcal{A}_{d+1}$  as in the GNS construction. The hypothesis on  $\lambda$  guarantees that there are no null vectors. In particular there is no difficulty in defining the following operators. Let  $\mathcal{H}$  denote the span of  $\mathcal{A}_d$  in  $\mathcal{K}$  and let  $\mathcal{N}$  denote the orthogonal complement of  $\mathcal{H}$ . Define  $S_j$  and  $T_j$  by  $S_j p = x_j p$  and  $T_j p = y_j p$  if  $p \in \mathcal{H}$  and  $S_j p = 0 = T_j p$  if  $p \in \mathcal{N}$ . Let  $P$  denote the orthogonal projection of  $\mathcal{K}$  onto  $\mathcal{H}$  and let  $X_j = PS_jP$  and  $Y_j = PT_jP$ .

For  $p, q \in \mathcal{H}$ ,

$$\begin{aligned}
 (7) \quad \langle X_j p, q \rangle &= \langle S_j p, q \rangle \\
 &= \langle x_j p, q \rangle \\
 &= \langle p, y_j q \rangle \\
 &= \langle p, Y_j q \rangle,
 \end{aligned}$$

where the third equality results from

$$\begin{aligned}
 (8) \quad \langle x_j p, q \rangle &= \lambda(q^* x_j p) \\
 &= \lambda((y_j q)^* p) \\
 &= \langle p, y_j q \rangle.
 \end{aligned}$$

Thus,  $Y_j = X_j^*$  and therefore, if  $p \in \mathcal{A}_d$ , then  $p(X)\emptyset = p$ . Further, if  $q$  is also in  $\mathcal{A}_d$ , then  $\langle p(X)\emptyset, q(X)\emptyset \rangle = \langle p, q \rangle$ .

Finally, since, as sets  $\mathcal{H} = \mathcal{A}_d$ , the dimension of  $\mathcal{H}$  is  $N(d)$ .  $\square$

**Lemma 3.2.** *There exists a linear functional  $\mu : \mathcal{A}_{2d} \rightarrow \mathbb{C}$  such that  $\mu(p^*p) > 0$  for all nonzero  $p \in \mathcal{A}_d$ .*

Proof. Our construction proceeds by induction. Suppose  $\mu_d : \mathcal{A}_{2d} \rightarrow \mathbb{C}$  is a linear functional satisfying  $\mu_d(p^*p) > 0$  for all nonzero  $p \in \mathcal{A}_d$ . Define an extension  $\mu_{d+1} : \mathcal{A}_{2d+2} \rightarrow \mathbb{C}$  of  $\mu_d$  depending on a choice of  $C > 0$  as follows. A word  $v$  is a square if there is a word  $w$  such that  $v = w^*w$ . Let  $\mu_{d+1}(v) = \mu_d(v)$  if  $v$  is a word of length at most  $2d$ ; let  $\mu_{d+1}(v) = 0$  if  $v$  is a word of length  $2d+1$  or  $v$  is a word of length  $2d+2$  but is not a square; and let  $\mu_{d+1}(v) = C$  if  $v$  is a word of length  $2d+2$  which is a square. Since the form  $\langle p, q \rangle_d = \mu_d(q^*p)$  is (strictly) positive definite on  $\mathcal{A}_d$ , there is a choice of  $C > 0$  such that  $\langle p, q \rangle_{d+1} = \mu_{d+1}(q^*p)$  is positive definite on  $\mathcal{A}_{d+1}$ .

To complete the induction argument, simply observe that  $\mu_0 : \mathcal{A}_0 \rightarrow \mathbb{C}$ ,  $\mu_0(c\emptyset) = c$ , gets the induction started.  $\square$

**Lemma 3.3.** *There exists a tuple  $X = (X_1, \dots, X_g)$  from  $\mathcal{B}(\mathbb{C}^{N(d)})$  such that if  $p \in \mathcal{A}_d$  and  $p(X) = 0$ , then  $p = 0$ .*

Proof. Combine (3.1) and (3.2).  $\square$

Key to the proof given here of (1.1) is the fact that  $\mathcal{C}_d$  is closed in  $\mathcal{A}_{2d}$ . Here we mean closed in some, and hence any, norm of  $\mathcal{A}_{2d}$ .

**Proposition 3.4.** *The cone  $\mathcal{C}_d$  is closed in  $\mathcal{A}_{2d}$ .*

Proof. Let  $X$  denote the tuple from (3.3) corresponding to  $2d$ . Then  $\|p\|_X = \|p(X)\|$  defines a norm on  $\mathcal{A}_{2d}$ . For  $p \in \mathcal{A}_{2d}$  expressed as in (1), the formula  $\|p\|_2^2 = \sum |p_w|^2$  also defines a norm on  $\mathcal{A}_{2d}$ . Since  $\mathcal{A}_{2d}$  is finite dimensional, these norms are equivalent.

Suppose  $p_n \in \mathcal{C}_d$  converges to  $p \in \mathcal{A}_{2d}$ . Then  $p_n(X)$  converges to  $p(X)$  so that the sequence  $\{p_n(X)\}$  is bounded. In view of (2.1), for each  $n$ , there exists  $r_{j,n} \in \mathcal{A}_d$ ,  $1 \leq j \leq N(d)$ , such that

$$(9) \quad p_n = \sum_{j=1}^{N(d)} r_{j,n}^* r_{j,n}.$$

Evaluating (9) at  $X$  it follows that each sequence  $\{r_{j,n}(X)\}_n$  is bounded and thus, by passing to a subsequence, there exists  $r_j \in \mathcal{A}_d$  such that  $r_{n,j}(X)$  converges to  $r_j(X)$ ,  $j = 1, 2, \dots, N(d)$  and therefore the sequence  $\{p_n(X)\}$  converges to  $\sum_j r_j(X)^* r_j(X)$  and the lemma follows.  $\square$

#### 4. Helton's Theorem

It is now possible to prove (1.1) by arguing the contrapositive. Accordingly, let  $\mathcal{H}$  be a Hilbert space of dimension  $N(d)$  and suppose  $q \in \mathcal{A}_d$  satisfies  $q(X) \geq 0$  for all tuples  $X$  of matrices acting on  $\mathcal{H}$ , but  $q \notin \mathcal{C}_d$ .

Note that, for all positive integers  $k, m$ ,  $\mathcal{C}_{m+k} \cap \mathcal{A}_m = \mathcal{C}_m \cap \mathcal{A}_m$ , due to the impossibility of cancelling the top degree terms. Thus we can regard  $q \in \mathcal{A}_{2d}$  and assume that  $q \notin \mathcal{C}_{d+1}$ .

Let  $\mathcal{A}_k^{\text{sa}}$  denote the self-adjoint elements of  $\mathcal{A}_k$ , where  $p \in \mathcal{A}_k$  is self-adjoint provided  $p^* = p$ . Observe that  $\mathcal{C}_k \subset \mathcal{A}_{2k}^{\text{sa}}$  and if  $p \in \mathcal{A}_k$ , then  $p = \text{re}(p) + i\text{im}(p)$ , where both  $\text{re}(p) = \frac{1}{2}(p + p^*)$  and  $\text{im}(p) = \frac{1}{2i}(p - p^*)$  are in  $\mathcal{A}_k^{\text{sa}}$ . If  $q^* \neq q$ , then, by (3.3), there is a tuple  $X$  of operators on the Hilbert space  $\mathcal{H}$  so that  $q(X)^* \neq q(X)$ . Thus, it may be assumed that  $q$  is self-adjoint.

By (3.4) and Minkowski's separation theorem (see for instance [RS]) there exists a real linear functional  $\nu : \mathcal{A}_{2d+2}^{\text{sa}} \rightarrow \mathbb{R}$  and a real number  $c$  such that  $\nu(q) < c \leq \nu(p)$  for all  $p \in \mathcal{C}_{d+1}$ . Since  $\mathcal{C}_{d+1}$  is a cone,  $c = 0$ . Define

$\Lambda : \mathcal{A}_{2d+2} \rightarrow \mathbb{C}$  by  $\Lambda(p) = \nu(\operatorname{re}(p)) + i\nu(\operatorname{im}(p))$  and verify that  $\Lambda$  is indeed a (complex) linear functional and  $\Lambda(p^*p) = \nu(p^*p) \geq 0$  for  $p \in \mathcal{A}_{d+1}$ .

Let  $\mu$  denote the linear functional from (3.2). There exists  $k > 0$  such that  $(\Lambda + k\mu)(q) < 0$ . Let  $\lambda = \Lambda + k\mu$ . Then  $\lambda(p^*p) > 0$  for all  $p \in \mathcal{A}_{d+1}$ . Hence, by (3.1), there exists a Hilbert space  $\mathcal{H}$  of dimension  $N(d)$ , a vector  $\gamma \in \mathcal{H}$ , and a tuple  $X$  from  $\mathcal{B}(\mathcal{H})$  such that  $\langle q(X)\gamma, \gamma \rangle = \lambda(q) < 0$ . A contradiction.

## References

- [H] J.W. Helton, *Positive non commutative polynomials are sums of squares* Ann. Math (to appear).
- [M] S. McCullough, *Factorization of operator-valued polynomials in several non-commuting variables*, Linear Algebra Appl. **326** (2001) no. 1-3, 193–203.
- [PV] M. Putinar and F.-H. Vasilescu *Solving moment problems by dimensional extension*, Ann. Math. **149** (1999), 1087-1107.
- [RS] M.Reed and B.Simon *Methods of Modern Mathematical Physics Vol. 1: Functional Analysis*, Academic Press, San Diego, 1980.
- [R] B. Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc. **96** (1992) No. **463**, mer. Math. Soc., Providence, R.I.

Received Received date / Revised version date

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF FLORIDA  
GAINESVILLE, FL 32611-8105  
*E-mail address:* sam@math.ufl.edu

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF CALIFORNIA  
SANTA BARBARA, CA 93106  
*E-mail address:* mputinar@mail.math.ucsb.edu

<sup>1</sup> Research supported by NSF grant DMS-0140112.

<sup>2</sup> Research supported by NSF grant DMS-0100367.