Münster J. of Math. **7** (2014), 671–696
DOI 10.17879/58269756277
urn:nbn:de:hbz:6-58269756512

Münster Journal of Mathematics

Quillen property of real algebraic varieties

Mihai Putinar and Claus Scheiderer

(Communicated by Joachim Cuntz)

Abstract. A conjugation-invariant ideal $I \subseteq \mathbb{C}[z_j, \overline{z}_j \mid j = 1, \dots, n]$ has the Quillen property if every real valued, strictly positive polynomial on the real zero set $V_{\mathbb{R}}(I) \subseteq \mathbb{C}^n$ is a sum of hermitian squares modulo I. We first relate the Quillen property to the archimedean property from real algebra. Using hereditary calculus, we then quantize and show that the Quillen property implies the subnormality of commuting tuples of Hilbert space operators satisfying the identities in I. In the finite rank case we give a complete geometric characterization of when the identities in I imply normality for a commuting tuple of matrices. This geometric interpretation provides simple means to refute Quillen's property of an ideal. We also generalize these notions and results from real algebraic sets to semialgebraic sets in $\mathbb{C}^n$.

1. Introduction

On any (affine) real algebraic variety V there exists a natural source for positivity certificates, namely squares (of regular functions): Any square, and hence any sum of squares, is nonnegative whereever it is defined on the $\mathbb{R}$-points of V. This observation lies at the very basis of real algebra, starting with Hilbert's 17th problem and its solution by Artin. Today the polarity between positivity and sums of squares is the focus of intense research, both from theoretical and applied points of view. See [11] and [19] for recent surveys.

In the present article we consider real algebraic subvarieties V of complex affine space. The embedding in complex space provides V with additional structure and gives the notion of holomorphic (and antiholomorphic) elements in the complexified structural rings of V. Accordingly we get a second, more restricted kind of positivity certificate, namely sums of *hermitian* squares on V, that is, of squared absolute values of holomorphic polynomials restricted to V. Our aim is to study this notion from the points of view of real algebra, geometry and operator theory.

We work with several complex variables $\mathbf{z} = (z_1, \dots, z_n)$ and their conjugates $\overline{\mathbf{z}} = (\overline{z}_1, \dots, \overline{z}_n)$. Let $I \subseteq \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]$ be a conjugation-invariant ideal, and let $V_{\mathbb{R}}(I) \subseteq \mathbb{C}^n$ be its zero set, a real algebraic subset of $\mathbb{C}^n$. Let $p \in \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]$ be a conjugation-invariant polynomial that is nonnegative on $V_{\mathbb{R}}(I)$. We study the question whether p admits an identity

$$(1) \qquad p(\mathsf{z},\overline{\mathsf{z}}) \;=\; |h_1(\mathsf{z})|^2+\cdots+|h_r(\mathsf{z})|^2+g(\mathsf{z},\overline{\mathsf{z}})$$

with $g\in I$, in which $h_1,\dots,h_r\in\mathbb{C}[\mathsf{z}]$ are *holomorphic* polynomials. When such an identity exists we will say that p is a *sum of hermitian squares* modulo I.

A classical instance where this property holds is the case of the unit circle $\mathbb{T}\subseteq\mathbb{C}$ and its vanishing ideal $I=(z\overline{z}-1)$. According to the Riesz–Fejér theorem, any $p\in\mathbb{C}[z,\overline{z}]$ nonnegative on $\mathbb{T}$ is a single hermitian square $p=|h(z)|^2$ modulo I.

The first multivariate example with such a property was discovered almost half a century ago by Quillen [16]. He studied the unit sphere $\mathbb{S}\subseteq\mathbb{C}^n$ and its reduced ideal I, and showed that any p strictly positive on $\mathbb{S}$ is a sum of hermitian squares modulo I.

Quillen's theorem amounts to a Positivstellensatz on the sphere vis-à-vis sums of hermitian squares, rather than ordinary squares. It is our aim to prove this result in greater generality, and to study the algebraic and geometric implications of such a result. Although our approach is basically algebraic, the interlacing with Hilbert space methods and operator theory is a recurrent theme of our study.

Fixing a conjugation-invariant ideal $I\subseteq\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$, we will say that I has the Quillen property if the Positivstellensatz holds for hermitian sums of squares modulo I. Assuming that $V_{\mathbb{R}}(I)$ is compact, an abstract characterization of this property comes from real algebra (Proposition 3.2). This characterization, however, is often not explicit enough. An improvement, on the constructive side, is offered by a known link to operator theory. Specifically, given $p\in\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$, and given a commuting tuple $T=(T_1,\dots,T_n)$ of bounded linear operators on a Hilbert space, define the operator $p(T,T^*)$ using hereditary calculus, thereby putting all adjoints to the left. We consider the following properties of the ideal I:

(A) (*Archimedean property*) $c-\sum_{j=1}^n|z_j|^2$ is a sum of hermitian squares modulo I, for some real number c.

(Q) (*Quillen property*) Every conjugation-invariant polynomial strictly positive on $V_{\mathbb{R}}(I)$ is a sum of hermitian squares modulo I.

(S) (*Subnormality*) Every commuting tuple T of bounded operators on a Hilbert space and satisfying $f(T,T^*)=0$ for all $f\in I$ is subnormal.

(Sf) (*Finite rank subnormality*) Every commuting tuple T of operators acting on a *finite-dimensional* Hilbert space and satisfying $f(T,T^*)=0$ for all $f\in I$ is subnormal (hence normal),

(G) (*Geometric normality*) The ideal I is not contained in the any of the "diamond" ideals

$$I(a,b) \;=\; \Big\{f\in\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}] \;\big|\; f(a,\overline{a})=f(a,\overline{b})=f(b,\overline{a})=f(b,\overline{b})=0\Big\}$$

for $a\neq b$ in $\mathbb{C}^n$, and neither in any of their degenerations $J(a,U)$, see 4.2.

We prove the implications

$$\begin{aligned} &(A) \;\Rightarrow\; (Q) \;\Rightarrow\; (S) \;\Rightarrow\; (Sf) \;\Leftrightarrow\; (G),\\ &(A) \;\Leftarrow\; (Q) \text{ if } V_{\mathbb{R}}(I) \text{ is compact,} \end{aligned}$$

see 3.2, 3.9 and 4.4. We also analyze by means of examples why the missing implications do not hold. For instance, even real conics in $\mathbb{C}$ offer nontrivial features (5.4, 5.8): A circle satisfies (A), an eccentric ellipse has property (S) but not (Q), the nonreduced ideal of a circle with a double point satisfies (Sf) but not (S), and a hyperbola whose asymptotes are perpendicular doesn't satisfy (Sf).

We then extend the study of hermitian Positivstellensätze from real algebraic sets to semialgebraic sets in $\mathbb{C}^n$. To this end we replace the semiring of hermitian sums of squares mod I by a hermitian module M, and the real algebraic set $V_{\mathbb{R}}(I)$ by the semialgebraic set $X_M \subseteq \mathbb{C}^n$ associated with M. Defining properties (Q), (S) and (Sf) for M accordingly, the implications (Q) $\Rightarrow$ (S) $\Rightarrow$ (Sf) remain true. When M is archimedean and satisfies a polynomial convexity property, the reverse (S) $\Rightarrow$ (Q) holds true as well (Theorem 6.16). When M is finitely generated, we prove that the Quillen property is incompatible with X_M containing an analytic disc (Theorem 6.20). In this direction we mention article [6], where a notion of hermitian complexity was introduced for conjugation-invariant ideals with the precise aim of bridging the gap between Quillen's property at one end and the existence of analytic discs in the support at the other.

At the end of the paper we make a few historical comments putting this work into perspective, mentioning some of the analytic roots and applications of hermitian sums of squares.

2. Preliminaries and notation

2.1. Let $\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$ be the polynomial ring in $2n$ independent variables $\mathsf{z} = (z_1,\dots,z_n)$ and $\overline{\mathsf{z}} = (\overline{z}_1,\dots,\overline{z}_n)$. On $\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$ we consider the $\mathbb{C}/\mathbb{R}$-involution $z_j^* = \overline{z}_j$ $(j=1,\dots,n)$. Thus

$$\Big(\sum_{\alpha,\beta} a_{\alpha,\beta}\, \mathsf{z}^\alpha\, \overline{\mathsf{z}}^\beta\Big)^* \;=\; \sum_{\alpha,\beta} \overline{a_{\alpha,\beta}}\, \mathsf{z}^\beta\, \overline{\mathsf{z}}^\alpha$$

for $a_{\alpha,\beta} \in \mathbb{C}$, with the usual multi-index notation $\mathsf{z}^\alpha\overline{\mathsf{z}}^\beta = \prod_{j=1}^n z_j^{\alpha_j}\overline{z}_j^{\beta_j}$. The fixed ring of $*$ is the polynomial ring $\mathbb{R}[\mathsf{x},\mathsf{y}]$ generated by the $2n$ variables $\mathsf{x} = (x_1,\dots,x_n)$ and $\mathsf{y} = (y_1,\dots,y_n)$, where $x_j = \frac{1}{2}(z_j+\overline{z}_j)$ and $y_j = \frac{1}{2i}(z_j-\overline{z}_j)$ (and $i = \sqrt{-1}$). Thus $\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$ is identified with $\mathbb{R}[\mathsf{x},\mathsf{y}]\otimes\mathbb{C}$, and under this identification, the involution $*$ becomes complex conjugation in the second tensor component.

Given $f \in \mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$ and $a,b \in \mathbb{C}^n$, we write $f(a,b) \in \mathbb{C}$ for the result of substituting a for z and b for $\overline{\mathsf{z}}$. We often abbreviate $f(a) := f(a,\overline{a})$.

2.2. There is a one-to-one correspondence between $*$-invariant ideals J of $\mathbb{C}[\mathtt{z},\overline{\mathtt{z}}]$ and arbitrary ideals I of $\mathbb{R}[\mathtt{x},\mathtt{y}]$, given by $J \mapsto I := J \cap \mathbb{R}[\mathtt{x},\mathtt{y}]$. Given an ideal I of $\mathbb{R}[\mathtt{x},\mathtt{y}]$, we denote the zero set of I in $\mathbb{C}^n$ by

$$V_{\mathbb{R}}(I) := \{a \in \mathbb{C}^n \mid \text{ for all } f \in I: \ f(a) = 0\}.$$

This is a real algebraic subset of $\mathbb{C}^n$.

2.3. For every $p \in \mathbb{C}[\mathtt{z},\overline{\mathtt{z}}]$, the hermitian norm $|p|^2 := pp^*$ is a sum of two (usual) squares in $\mathbb{R}[\mathtt{x},\mathtt{y}]$. The convex cone in $\mathbb{R}[\mathtt{x},\mathtt{y}]$ generated by $\{|p|^2 \mid p \in \mathbb{C}[\mathtt{z},\overline{\mathtt{z}}]\}$ will be denoted by Σ; it is the cone of all (usual) *sums of squares* in $\mathbb{R}[\mathtt{x},\mathtt{y}]$. The smaller convex cone in $\mathbb{R}[\mathtt{x},\mathtt{y}]$ generated by $\{|p|^2 \mid p \in \mathbb{C}[\mathtt{z}]\}$ is denoted by Σ_h. Its elements are called the *hermitian sums of squares*.

2.4. We recall a few notions from real algebra. Given an $\mathbb{R}$-algebra A (i.e., a commutative ring containing $\mathbb{R}$), a subset $S \subseteq A$ will be called a *semiring* in A if S contains the nonnegative real numbers and is closed in A under taking sums and products. Given a semiring S, an *S-module* is a subset M of A with $M + M \subseteq M$, $SM \subseteq M$ and $1 \in M$. A particularly important semiring is ΣA^2, the set of all (finite) sums of squares in A. The modules over this semiring are usually referred to as the quadratic modules in A.

The S-module M is said to be *archimedean* if $A = \mathbb{R} + M$, that is, if for every $f \in A$ there exists $c \in \mathbb{R}$ with $c \pm f \in M$.

In this paper we will mostly be concerned with the $\mathbb{R}$-algebra $A = \mathbb{R}[\mathtt{x},\mathtt{y}]$ and with the two semirings $\Sigma_h \subseteq \Sigma$ in $\mathbb{R}[\mathtt{x},\mathtt{y}]$.

2.5. Given a module M over some semiring S in $\mathbb{R}[\mathtt{x},\mathtt{y}]$, we write

$$X_M := \big\{a \in \mathbb{C}^n \mid \text{ for all } g \in M: g(a) \geq 0\big\},$$

which is a closed subset of $\mathbb{C}^n$.

The celebrated archimedean Positivstellensatz from real algebra (see [12] or [19]) implies:

Theorem 2.6. *If M is a module over an archimedean semiring S in $\mathbb{R}[\mathtt{x},\mathtt{y}]$, then M contains any $f \in \mathbb{R}[\mathtt{x},\mathtt{y}]$ that is strictly positive on the set X_M.*

3. Hermitian sums of squares and subnormal tuples of operators

3.1. Let $\mathbb{R}[\mathtt{x},\mathtt{y}] \subseteq \mathbb{C}[\mathtt{z},\overline{\mathtt{z}}]$ be the fixed ring of $*$ (see 2.1). Given any ideal $I \subseteq \mathbb{R}[\mathtt{x},\mathtt{y}]$, we will consider the semiring $S = \Sigma_h + I$ in $\mathbb{R}[\mathtt{x},\mathtt{y}]$. Note that $X_S = V_{\mathbb{R}}(I)$. Consider the following two properties of the ideal I:

(A) *(Archimedean Property)* The semiring $\Sigma_h + I$ in $\mathbb{R}[\mathtt{x},\mathtt{y}]$ is archimedean (see 2.4);

(Q) *(Quillen Property)* $\Sigma_h + I$ contains every $f \in \mathbb{R}[\mathtt{x},\mathtt{y}]$ that is strictly positive on $V_{\mathbb{R}}(I)$.

We will also refer to (A) by saying that Σ_h is archimedean modulo I.

Given any $*$-invariant ideal $J \subseteq \mathbb{C}[\mathtt{z},\overline{\mathtt{z}}]$, we will say that J has property (A) resp. (Q) if the ideal $J \cap \mathbb{R}[\mathtt{x},\mathtt{y}]$ of $\mathbb{R}[\mathtt{x},\mathtt{y}]$ has the respective property (see 2.2).

The following result was proved in [14, Thm. 2.1 and Prop. 2.2]. It is essentially an application of the archimedean Positivstellensatz 2.6:

Proposition 3.2. *For any ideal $I \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$, the following are equivalent:*

(i) *I has the Archimedean property* (A)*;*
(ii) *I has the Quillen property* (Q)*, and $V_{\mathbb{R}}(I)$ is compact;*
(iii) *I contains a polynomial of the form $||\mathtt{z}||^2 + p + a$, where $p \in \Sigma_h$ and $a \in \mathbb{R}$.*

(We are using the shorthand $||\mathtt{z}||^2 := |z_1|^2 + \cdots + |z_n|^2$.)

Remarks 3.3.

1. Quillen's theorem [16], reproved later by Catlin–D'Angelo [3], was mentioned in the introduction. The statement is recovered here in a purely algebraic way, as a very particular instance of Proposition 3.2.

   As observed in [3], Quillen's theorem implies the following classical theorem due to Pólya: Given a homogeneous polynomial $f \in \mathbb{R}[x_1, \dots, x_n]$ strictly positive on $\{a \in \mathbb{R}^n \mid a_1 \geq 0, \dots, a_n \geq 0\} \smallsetminus \{(0, \dots, 0)\}$, the form $(x_1 + \cdots + x_n)^N f$ has positive coefficients for large enough $N \geq 0$.
2. Condition (iii) of 3.2 gives an abstract algebraic characterization of the ideals I with $V_{\mathbb{R}}(I)$ compact and with property (Q). Note that the Positivstellensatz for usual sums of squares holds whenever $V_{\mathbb{R}}(I)$ is compact, by Schmüdgen's theorem [20]. In contrast, "most" ideals with $V_{\mathbb{R}}(I)$ compact do not satisfy property (Q) (see, e.g., 5.4 below).

   The applicability of 3.2(iii) as an algebraic criterion for property (Q) is somewhat limited, since this condition is not sufficiently explicit. In particular, it is usually cumbersome to prove that an ideal I does *not* contain any polynomial of the form given in (iii). Therefore it is desirable to know other conditions on I that are necessary for (A) resp. (Q), and that are more easily checked. In this section and the next we will offer two conditions of very different nature that are both necessary for the Quillen property, one operator-theoretic and one ideal-theoretic.
3. Part of the original motivation for this work came from a question of D'Angelo. Given a compact real algebraic set $X \subseteq \mathbb{C}^n$ which is the boundary of a strictly pseudo-convex region in $\mathbb{C}^n$, D'Angelo had asked whether every strictly positive polynomial on X is a sum of hermitian squares on X. This question was answered in the negative, see [14].

Examples 3.4.

1. The Quillen property (Q) alone does not imply the archimedean property (A), since $V_{\mathbb{R}}(I)$ need not be compact. This is seen by considering a line in $\mathbb{C}$, given (say) by the ideal $I = (y) \subseteq \mathbb{R}[x, y]$. Condition (Q) is satisfied since, in fact, $\Sigma_h + I$ contains every $f \in \mathbb{R}[x, y]$ nonnegative on the line $y = 0$. Indeed, such f is a sum of two usual squares modulo I, from which one sees easily that f is congruent modulo I to a single hermitian square, i.e., $f \equiv |p|^2 \pmod I$ with $p \in \mathbb{C}[z]$.

2. If $n = 1$ and $f \in \mathbb{R}[x, y]$ has degree 2, the principal ideal $I = (f)$ satisfies the Archimedean property (A) if and only if there exist $\alpha \in \mathbb{C}$ and a, $c \in \mathbb{R}$ with $f = a|z - \alpha|^2 + c$. This will be proved in Theorem 5.4 below.

3.5. Let E be a (separable complex) Hilbert space, and let $B(E)$ denote the algebra of bounded linear operators on E. Fix a tuple $T = (T_1, \dots, T_n)$ of operators $T_j \in B(E)$ that commute pairwise. We use hereditary calculus (see [1, Sec. 14.2] for more details). Given a monomial $f = \mathsf{z}^\alpha \overline{\mathsf{z}}^\beta$ (with α, $\beta \in \mathbb{Z}_+^n$) we write

$$f(T, T^*) := T^{*\beta} T^\alpha.$$

We extend this definition $\mathbb{C}$-linearly, thereby putting all adjoints to the left. This defines the $\mathbb{C}$-linear map

$$\psi_T : \mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}] \to B(E),\ f(\mathsf{z}, \overline{\mathsf{z}}) \mapsto \psi_T(f) = f(T, T^*).$$

The map ψ_T commutes with the involution, i.e. $\psi_T(f^*) = \psi_T(f)^*$. In particular, $\psi_T(f)$ is selfadjoint for $f = f^*$. Note that

$$\psi_T\Big(\overline{q(\mathsf{z})} \cdot f(\mathsf{z}, \overline{\mathsf{z}}) \cdot p(\mathsf{z})\Big) = \psi_T(q)^* \, \psi_T(f) \, \psi_T(p)$$

for p, $q \in \mathbb{C}[\mathsf{z}]$ and $f \in \mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}]$. The set

$$M_T := \big\{ f \in \mathbb{R}[\mathsf{x}, \mathsf{y}] \mid \psi_T(f) \geq 0 \big\}$$

(of real polynomials f for which the selfadjoint operator $\psi_T(f)$ is nonnegative) is a Σ_h-module in $\mathbb{R}[\mathsf{x}, \mathsf{y}]$, since $\psi_T(|p|^2 f) = \psi_T(\overline{p(\mathsf{z})} f(\mathsf{z}, \overline{\mathsf{z}}) p(\mathsf{z})) = p(T)^* \, \psi_T(f) \, p(T)$ holds for $f \in \mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}]$ and $p \in \mathbb{C}[\mathsf{z}]$. The support $M_T \cap (-M_T) = \ker(\psi_T)$ of M_T is an ideal in $\mathbb{R}[\mathsf{x}, \mathsf{y}]$. Note that the subset M_T of $\mathbb{R}[\mathsf{x}, \mathsf{y}]$ is closed with respect to the finest locally convex topology on $\mathbb{R}[\mathsf{x}, \mathsf{y}]$.

3.6. Recall that the tuple T is said to be (jointly) *subnormal* if T can be extended to a commuting tuple of normal operators on a larger Hilbert space, i.e., if there is a tuple $T' = (T'_1, \dots, T'_n)$ of commuting normal operators on a Hilbert space E' such that E' contains E and the T'_i leave E invariant and satisfy $T'_i|_E = T_i$ for $i = 1, \dots, n$. Note that subnormal is equivalent to normal when $\dim(E) < \infty$. For details see [4].

According to the Halmos–Bram–Itô criterion (see [9]), the commuting tuple $T = (T_1, \dots, T_n)$ is subnormal if and only if

$$\sum_{\alpha, \beta} \langle T^\alpha \xi_\beta, T^\beta \xi_\alpha \rangle \geq 0$$

for all finitely supported families $\{\xi_\alpha\}_{\alpha \in \mathbb{Z}_+^n}$ in E. Using this criterion we show:

Proposition 3.7. *Let $T = (T_1, \dots, T_n)$ be a commuting tuple in $B(E)$. Then T is subnormal if and only if $\Sigma \subseteq M_T$.*

In other words, the tuple T is subnormal if and only if $\psi_T(|p|^2) \geq 0$ holds for every $p \in \mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}]$.

Proof. Assume $\psi_T(|p|^2) \geq 0$ for every $p \in \mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$. To prove that T is subnormal we can, using a result of Stochel ([21, Cor. 3.2]), assume that there exists a cyclic vector ξ for T, i.e. the linear span of $\{T^\alpha \xi \mid \alpha \in \mathbb{Z}_+^n\}$ is dense in E. It suffices to verify the Halmos–Bram–Itô condition for all finite families $\{\xi_\alpha\}$ lying in the linear span of $\{T^\alpha \xi \mid \alpha \in \mathbb{Z}_+^n\}$. So let $\xi_\alpha = p_\alpha(T)\xi$ where $p_\alpha \in \mathbb{C}[\mathtt{z}]$ for $\alpha \in \mathbb{Z}_+^n$ (and $p_\alpha = 0$ for almost all α), and consider $p := \sum_\alpha p_\alpha(\mathtt{z})\, \overline{\mathtt{z}}^\alpha \in \mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$. Since

$$|p|^2 \;=\; \sum_{\alpha,\beta} p_\alpha(\mathtt{z})\, \overline{p_\beta(\mathtt{z})}\, \mathtt{z}^\beta\, \overline{\mathtt{z}}^\alpha,$$

the assumption $\Sigma \subseteq M_T$ gives

$$0 \leq \langle \psi_T(|p|^2)\xi, \xi\rangle \;=\; \sum_{\alpha,\beta} \big\langle T^\beta p_\alpha(T)\xi,\, T^\alpha p_\beta(T)\xi \big\rangle \;=\; \sum_{\alpha,\beta} \big\langle T^\beta \xi_\alpha,\, T^\alpha \xi_\beta \big\rangle,$$

which shows that T is subnormal. Conversely, the same argument shows that T subnormal implies $\Sigma \subseteq M_T$. □

3.8. We shall consider the following properties of an ideal $I \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$:

(S) *(Subnormality)* Every commuting tuple $T = (T_1, \dots, T_n)$ of bounded linear operators in a Hilbert space satisfying $p(T, T^*) = 0$ for every $p \in I$ is subnormal.

(Sf) *(Finite rank subnormality)* Every commuting tuple $T = (T_1, \dots, T_n)$ of complex matrices satisfying $p(T, T^*) = 0$ for every $p \in I$ is normal.

Trivially (S) implies (Sf). Condition (Sf) will be considered in the next section. Here we first show that condition (S) is necessary for the Quillen property (Q). (This fact was announced without proof in [15, Cor. 2.2]).

Proposition 3.9. *For any ideal $I \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$, Quillen property* (Q) *implies the subnormality condition* (S)*.*

Proof. Assume (Q) holds for I. Given a commuting tuple T of bounded operators with $I \subseteq \ker(\psi_T)$, we have $\Sigma_h + I \subseteq M_T$. Since M_T is closed with respect to the finest locally convex topology of $\mathbb{R}[\mathtt{x}, \mathtt{y}]$ (3.5), it follows from (Q) that M_T contains every polynomial that is nonnegative on $V_\mathbb{R}(I)$. In particular we have $\Sigma \subseteq M_T$, which implies that T is subnormal (Proposition 3.7). □

Remark 3.10. In the case when $V_\mathbb{R}(I)$ is compact, we can give a very short proof of Proposition 3.9, using Athavale's theorem [2]. Indeed, assume that $V_\mathbb{R}(I)$ is compact and (Q) holds for I. After suitably scaling the variables we can assume $|\xi_j| < 1$ for every $\xi = (\xi_1, \dots, \xi_n) \in V_\mathbb{R}(I)$. Let T be a commuting tuple of bounded operators satisfying $I \subseteq \ker(\psi_T)$. In order to show that T is subnormal it suffices, by [2, Thm. 4.1], to show for any tuple $\alpha = (\alpha_1, \dots, \alpha_n)$ of nonnegative integers that

$$f \;:=\; \prod_{j=1}^{n} (1 - |z_j|^2)^{\alpha_j} \;\in\; M_T.$$

Now $f > 0$ on $V_{\mathbb{R}}(I)$, so the assumption on I implies $f \in \Sigma_h + I$, from which $\psi_T(f) \geq 0$ is obvious.

Remark 3.11. The subnormality property (S) on an ideal $I \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$ is strictly weaker than the Quillen property (Q). An immediate example to show this is given by the ideal $I = (x_1, \ldots, x_n) = (z_j + \overline{z}_j \mid j = 1, \ldots, n)$ in $\mathbb{R}[\mathtt{x}, \mathtt{y}]$: Every commuting tuple T of operators with $I \subseteq \ker(\psi_T)$ consists clearly of normal operators. On the other hand, for any $n \geq 2$ there exist strictly positive polynomials on $V_{\mathbb{R}}(I) \cong \mathbb{R}^n$ that are not even sums of usual squares, and *a fortiori* not of hermitian squares. For instance, adding a positive constant to the well-known Motzkin polynomial $y_1^4 y_2^2 + y_1^2 y_2^4 - 3y_1^2 y_2^2 + 1$ gives such an example.

It is less straightforward to find an ideal I satisfying (S) but not (Q), for which $V_{\mathbb{R}}(I)$ is compact. Let $f(z, \overline{z}) = 0$ be the equation of an ellipse that is not a circle. Then every bounded operator T satisfying $f(T, T^*) = 0$ is subnormal, that is, the principal ideal $I = (f)$ satisfies (S). But I does not have the Quillen property, see Theorem 5.4 below, and also [15].

4. Normal tuples of matrices

In this section we will provide a complete geometric characterization of the ideals satisfying finite rank subnormality (Sf). More specifically, we will explicitly list those ideals that are maximal with respect to not satisfying (Sf).

First we need some preparation. It seems more natural here to work with $*$-invariant ideals of $\mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$, rather than with ideals of $\mathbb{R}[\mathtt{x}, \mathtt{y}]$.

4.1. Given $a \neq b$ in $\mathbb{C}^n$, let $I(a, b) \subseteq \mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$ be the ideal consisting of all polynomials $f(\mathtt{z}, \overline{\mathtt{z}})$ with

$$f(a, \overline{a}) \;=\; f(b, \overline{b}) \;=\; f(a, \overline{b}) \;=\; f(b, \overline{a}) \;=\; 0.$$

Clearly, $I(a, b)$ is $*$-invariant. As an ideal in $\mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$, note that $I(a, b)$ is generated by the polynomials $p(\mathtt{z})$ and $p(\mathtt{z})^*$, where $p(\mathtt{z}) \in \mathbb{C}[\mathtt{z}]$ is a holomorphic polynomial satisfying $p(a) = p(b) = 0$.

These ideals were introduced in [14], where $I(a, b)$ was denoted by $J_{a,b}$.

4.2. The usual inner product on the space of hermitian $n \times n$ matrices will be denoted by $\langle S, T \rangle := \mathrm{tr}(ST)$. Given $a \in \mathbb{C}^n$ and a complex hermitian $n \times n$ matrix $U \neq 0$, let $J(a, U)$ be the set of all $f \in \mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$ such that

(1) $f(a, \overline{a}) = 0$,
(2) $U \cdot \nabla_{\mathtt{z}} f(a, \overline{a}) \;=\; \overline{U} \cdot \nabla_{\overline{\mathtt{z}}} f(a, \overline{a}) \;=\; 0$,
(3) $\big\langle U, \nabla^2_{\mathtt{z}\overline{\mathtt{z}}} f(a, \overline{a}) \big\rangle \;=\; 0$.

Here we denote the holomorphic resp. antiholomorphic gradient by

$$\nabla_{\mathtt{z}} f = \Big(\frac{\partial f}{\partial z_j}\Big)_{j=1,\ldots,n}, \;\; \nabla_{\overline{\mathtt{z}}} f = \Big(\frac{\partial f}{\partial \overline{z}_j}\Big)_{j=1,\ldots,n}$$

(regarded as column vectors), and the mixed Hessian (Levi form) by

$$\nabla^2_{\mathsf{z}\overline{\mathsf{z}}} f = \Big(\frac{\partial^2 f}{\partial z_j\, \partial \overline{z}_k}\Big)_{j,k=1,\dots,n}.$$

It is easy to see that $J(a,U)$ is a $*$-invariant ideal in $\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$.

Example 4.3. With a view toward the proof of Theorem 4.4 below, let us consider the following example. Fix an integer $r \geq 1$ and column vectors $w_1,\dots,w_n \in \mathbb{C}^r$, not all of them zero. Moreover, let $a = (a_1,\dots,a_n) \in \mathbb{C}^n$, and let

$$T_j \;=\; \begin{pmatrix} a_j & 0 \\ w_j & a_j I_r \end{pmatrix} \;\in\; \mathrm{M}_{r+1}(\mathbb{C})$$

(we are using a $(1,r)$ block matrix notation). Clearly, $T = (T_1,\dots,T_n)$ is a commuting tuple of matrices, and is not normal since $w_j \neq 0$ for at least one j. A straightforward calculation shows $\ker(\psi_T) = J(a,U)$, where U is the nonnegative hermitian $n\times n$-matrix

$$U \;=\; \big(w_j^* w_k\big)_{1\leq j,k\leq n}.$$

Note that the rank of U is the dimension of the linear span of $w_1,\dots,w_n$ in $\mathbb{C}^r$.

Next comes the main result of this section. It gives a complete ideal-theoretic characterization of condition (Sf):

Theorem 4.4. *Let $I \subseteq \mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$ be an ideal. The following are equivalent:*

(Sf) *Every commuting tuple $T = (T_1,\dots,T_n)$ of complex matrices satisfying $I \subseteq \ker(\psi_T)$ is normal;*

(G) *I is not contained in $I(a,b)$ for any pair $a \neq b$ in $\mathbb{C}^n$, and neither in $J(a,U)$ for any $a \in \mathbb{C}^n$ and any nonnegative hermitian $n\times n$ matrix $U \neq 0$.*

4.5. We prove the implication (Sf) $\Rightarrow$ (G) by contraposition. More precisely, we will show:

(a) For any $a \neq b$ in $\mathbb{C}^n$, there exists a commuting non-normal n-tuple T of 2×2 matrices with $\ker(\psi_T) = I(a,b)$.

(b) For any $a \in \mathbb{C}^n$ and any nonnegative hermitian $n\times n$ matrix $U \neq 0$, there exists a commuting non-normal n-tuple T of $m\times m$ matrices with $\ker(\psi_T) = J(a,U)$. (We can take $m = \mathrm{rk}(U)+1$ here.)

In fact, (b) has already been proved by Example 4.3. (The last assertion comes from the fact that a nonnegative hermitian matrix U of rank $r \geq 1$ can be written $U = W^*W$ with $W \in \mathrm{M}_{r\times n}(\mathbb{C})$.) Assertion (a) will be proved in 4.6 and 4.7. The reverse implication (G) $\Rightarrow$ (Sf) will be proved in 4.8.

4.6. Let $a \neq b$ in $\mathbb{C}^n$. Fix two linearly independent vectors u, v in $\mathbb{C}^2$ that are not perpendicular. Let $T_j \in M_2(\mathbb{C})$ be the matrix satisfying $T_j u = a_j u$ and $T_j v = b_j v$ $(j = 1,\dots,n)$. Then $T = (T_1,\dots,T_n)$ is a commuting tuple of matrices. Clearly, the matrix T_j fails to be normal for any index j with $a_j \neq b_j$, and in particular, the tuple T is not normal. We claim $\ker(\psi_T) = I(a,b)$. The inclusion $I(a,b) \subseteq \ker(\psi_T)$ is obvious. Since the vector space $\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]/I(a,b)$ has dimension 4, we have to show that the linear map $\psi_T : \mathbb{C}[\mathsf{z},\overline{\mathsf{z}}] \to M_2(\mathbb{C})$ is surjective. This in turn follows immediately from the following lemma.

Lemma 4.7. *Let $S \in \mathrm{M}_2(\mathbb{C})$. The matrices I, S, S^*, S^*S are linearly dependent if and only if S is normal.*

Proof. Let W_S be the linear span of I, S, S^* and S^*S in $\mathrm{M}_2(\mathbb{C})$. We have $W_S = W_{S-\lambda I}$ for every $\lambda \in \mathbb{C}$. Since $S-\lambda I$ is normal if and only if S is normal, we can replace S by $S - \lambda I$ for any $\lambda \in \mathbb{C}$. In particular, we may do this for λ an eigenvalue of S. After changing to a suitable orthonormal basis we can therefore assume $S = \left(\begin{smallmatrix} 0 & a \\ 0 & b \end{smallmatrix}\right)$ where a, $b \in \mathbb{C}$. For this matrix it is immediate that $W_S \neq \mathrm{M}_2(\mathbb{C})$ if and only if $a = 0$, if and only if S is normal. □

4.8. We now show that (G) implies (Sf) in Theorem 4.4, again by contraposition. To this end let E be a finite-dimensional Hilbert space, and let $T = (T_1, \dots, T_n)$ be a commuting tuple of endomorphisms of E such that at least one T_j is not normal. We will show that the ideal $\ker(\psi_T)$ of $\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}]$ is contained in one of the ideals $I(a,b)$ or $J(a,U)$, as in (G).

Let F be any T-invariant subspace of E (that is, $T_jF \subseteq F$ holds for each j), and let $T|F$ denote the restriction of T to F. So $T|F$ is a commuting tuple of endomorphisms of F. Let $i : F \to E$ be the inclusion map and $\pi : E \to F$ the orthogonal projection onto F, and let

$$\rho : \mathrm{End}(E) \to \mathrm{End}(F),\ \rho(S) = \pi \circ S \circ i.$$

For $S \in \mathrm{End}(E)$ we have $(S|F)^* = \rho(S^*)$. Moreover, if S leaves F invariant, then $\rho(S'S) = \rho(S')\rho(S)$, $\rho(S^*S') = \rho(S)^*\rho(S')$ hold for any $S' \in \mathrm{End}(E)$. As maps $\mathbb{C}[\mathsf{z},\overline{\mathsf{z}}] \to \mathrm{End}(F)$, we therefore have $\psi_{T|F} = \rho \circ \psi_T$. In particular, $\ker(\psi_T) \subseteq \ker(\psi_{T|F})$. In order to prove what we want, we can therefore replace E and T by F and $T|F$ whenever F is a T-invariant subspace of E for which $T|F$ is not normal.

For any tuple $a = (a_1, \dots, a_n) \in \mathbb{C}^n$, denote by

$$E(T,a) \ = \ \{\xi \in E \mid \ (T_j - a_j)\xi = 0 \text{ for } j = 1, \dots, n\}$$

resp. by

$$E_\infty(T,a) \ = \ \{\xi \in E \mid \ (T_j - a_j)^{\dim(E)}\xi = 0 \text{ for } j = 1, \dots, n\}$$

the a-eigenspace resp. the generalized a-eigenspace of T. These are T-invariant subspaces of E, and $E = \bigoplus_{a \in \mathbb{C}^n} E_\infty(T,a)$. Since T is not normal, one of the following two situations occurs:

(1) One of the T_j is not diagonalizable;
(2) each T_j is diagonalizable, but for at least one index j there are two eigenspaces of T_j that are not perpendicular.

Let us first discuss case (2). By assumption we have $E = \bigoplus_{a \in \mathbb{C}^n} E(T,a)$, and there exist $a \neq b$ in $\mathbb{C}^n$ such that $E(T,a)$ and $E(T,b)$ are not perpendicular. Pick vectors $x \in E(T,a)$ and $y \in E(T,b)$ that are not perpendicular. The two-dimensional subspace F spanned by x and y is T-invariant, and $T|F$ is not normal. By the argument used in 4.6, we see that $\ker(\psi_{T|F}) = I(a,b)$. So we are finished with case (2).

Now we discuss case (1) and assume that one of the T_j cannot be diagonalized. Then there exists $a \in \mathbb{C}^n$ with $E(T,a) \neq E_\infty(T,a)$. Replacing E by $E_\infty(T,a)$ and T_j by $T_j - a_j$ for each j (the latter corresponding to a change of variables $z_j \to z_j - a_j$ in the polynomial ring), we can assume that each T_j is nilpotent and $T_j \neq 0$ for at least one j. Let $c \geq 2$ be the highest order of nilpotency among the T_j, that is, assume $T_j^c = 0$ for all j and $T_{j_0}^{c-1} \neq 0$ for one index j_0. Replacing E by $\ker(T_{j_0}^{c-2})$ we can assume $T_j^2 = 0$ for all j.

Let $V_j = \ker(T_j)$ for $j = 1,\dots,n$. Whenever there are two indices j, k with $V_j \not\subseteq V_k$, we can replace E by V_j. Iterating this step we arrive at the case where all nonzero operators among $T_1,\dots,T_n$ have the same kernel $V \neq E$. Thus, for each j, we have either $T_j = 0$ or $\operatorname{im}(T_j) \subseteq \ker(T_j) = V$, and the latter occurs for at least one index j.

Choose a nonzero vector $x \in V^\perp$. The subspace $F := \mathbb{C}x \oplus V$ of E is T-invariant, and we can replace E with F. Put $y_j = T_j x$ ($j = 1,\dots,n$), and let $W \subseteq V$ be the linear span of $y_1,\dots,y_n$. We can replace E by $\mathbb{C}x \oplus W$, and have now arrived at a minimal non-normal tuple of operators.

Let $r = \dim(W)$, so $1 \leq r \leq n$. Fixing an orthonormal linear basis of W, we represent the operators T_j by $(r+1)\times(r+1)$ matrices as

$$T_j \;=\; \begin{pmatrix} 0 & 0 & \cdots & 0 \\ y_{1j} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ y_{rj} & 0 & \cdots & 0 \end{pmatrix}.$$

Let $w_j = (y_{1j},\dots,y_{rj})$, regarded as a column vector ($j = 1,\dots,n$), and let

$$U \;=\; \big(w_j^* w_k\big)_{1\leq j,k\leq n}$$

a psd hermitian matrix of rank r. From Example 4.3 we see $\ker(\psi_T) = J(a,U)$.

This completes the proof of Theorem 4.4. $\square$

Remark 4.9. Ideals of the form $I(a,b)$ or $J(a,U)$, as in 4.4, are pairwise incomparable with respect to inclusion, except that $J(a,U) = J(a,cU)$ for every real number $c > 0$. To see that $J(a,U) \subseteq J(a,U')$ implies $U' = cU$ with $c > 0$, observe that the mixed Hessians of elements of $J(a,U)$ are precisely the matrices that are orthogonal to U (condition (3) of 4.2). Therefore U is determined by $J(a,U)$ up to (positive) scaling.

So we see that the ideals $I(a,b)$ and $J(a,U)$ are precisely the maximal ones among the ideals of relations between non-normal commuting tuples of matrices (in the sense of hereditary calculus).

Remark 4.10. A complex square matrix may be non-normal for two reasons: It may fail to be diagonalizable, or it may have two nonperpendicular eigenvectors for different eigenvalues. The ideals $J(a,U)$ and $I(a,b)$ in Theorem 4.4

correspond to these two possibilities. More precisely, if $T = (T_1, \dots, T_n)$ is a commuting tuple of matrices, and if one of the T_j is not diagonalizable, then $\ker(\psi_T) \subseteq J(a, U)$ for some pair (a, U). On the other hand, if the T_j are diagonalizable but one of them has two nonperpendicular eigenspaces, then $\ker(\psi_T) \subseteq I(a, b)$ for some pair (a, b). Both assertions are clear from the proof in 4.8.

Remark 4.11. Consider commuting tuples $T = (T_1, \dots, T_n)$ in $B(E)$ where E is a complex Hilbert space, together with the associated maps $\psi_T : \mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}] \to B(E)$ given by hereditary calculus (3.5). It is a consequence of Fuglede's theorem that the tuple T is normal if and only if ψ_T is a ring homomorphism. As a consequence of Theorem 4.4, we can add another characterization, as long as E has finite dimension. It shows that the normality of a commuting tuple T of matrices can be decided from its ideal $\ker(\psi_T)$ of relations:

Corollary 4.12. *A commuting tuple T of matrices is normal if and only if the ideal $\ker(\psi_T)$ is not contained in $I(a, b)$ for any $a \neq b$ in $\mathbb{C}^n$, and neither in $J(a, U)$ for any $a \in \mathbb{C}^n$ and any nonnegative hermitian $n \times n$-matrix $U \neq 0$.*

Proof. Indeed, if T is normal, there is an orthogonal basis of simultaneous eigenvectors. This implies that $\ker(\psi_T)$ is an intersection of finitely many ideals $\mathfrak{m}_a = \{f \in \mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}] \mid f(a, \overline{a}) = 0\}$, $a \in \mathbb{C}^n$. Such an intersection is never contained in any of the ideals $I(a, b)$ or $J(a, U)$. □

Remarks 4.13.

1. Up to holomorphic linear coordinate changes there exist precisely n essentially different ideals $J(a, U)$ in $\mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}]$. Indeed, we can assume that $a = 0$ and that
$$U = U_r := \operatorname{diag}(1, \dots, 1, 0, \dots, 0)$$
is the diagonal matrix of rank r, where $1 \leq r \leq n$ can be arbitrary. In this case, $J(0, U_r)$ consists of all $f \in \mathbb{C}[\mathsf{z}, \overline{\mathsf{z}}]$ which are modulo $(z_1, \dots, z_n)^2 + (\overline{z}_1, \dots, \overline{z}_n)^2$ congruent to
$$\sum_{j=r+1}^{n} (b_j z_j + b'_j \overline{z}_j) + \sum_{j,k=1}^{n} c_{jk}\, z_j \overline{z}_k$$
with b_j, b'_j, $c_{jk} \in \mathbb{C}$ and
$$c_{11} + \cdots + c_{rr} = 0.$$
A system of generators for the ideal $J(0, U_r)$ is therefore given by the following list of polynomials:
$$\begin{array}{rl} z_j z_k,\ \overline{z}_j \overline{z}_k & 1 \leq j \leq k \leq r, \\ z_j \overline{z}_k,\ z_k \overline{z}_j & 1 \leq j < k \leq r, \\ |z_j|^2 - |z_{j+1}|^2 & 1 \leq j < r, \\ z_j,\ \overline{z}_j & r+1 \leq j \leq n. \end{array}$$

2. In [14], the ideals

$$J_{a,a} \;=\; (z_1 - a_1, \dots, z_n - a_n)^2 + (\overline{z}_1 - \overline{a}_1, \dots, \overline{z}_n - \overline{a}_n)^2 \;\; (a \in \mathbb{C}^n)$$

of $\mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$ were used. They ideals relate to the ideals $J(a, U)$ studied here via

$$J_{a,a} \;=\; \bigcap_U J(a, U),$$

intersection over all nonnegative hermitian matrices $U \neq 0$. In particular, in the one variable case ($n = 1$) we have $J_{a,a} = J(a, 1)$.

As a consequence of Theorem 4.4 and Proposition 3.9, we obtain:

Corollary 4.14. *Let $I \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$ be an ideal, and assume that $I \subseteq I(a, b)$ for some $a \neq b$ in $\mathbb{C}^n$, or that $I \subseteq J(a, U)$ for some $a \in \mathbb{C}^n$ and some nonnegative hermitian $n \times n$ matrix $U \neq 0$. Then there exists $f \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$ such that $f > 0$ on $V_{\mathbb{R}}(I)$, but f is not a hermitian sum of squares modulo I.* □

In the first case of Corollary 4.14, the assertion was already proved in [14, Prop. 3.1], by a different argument.

5. Examples

We start by identifying some classes of (principal) ideals that satisfy the subnormality condition (S).

Proposition 5.1. *Let $f = f^* \in \mathbb{C}[\mathtt{z}, \overline{\mathtt{z}}]$ be of the form*

$$f \;=\; \operatorname{Re} g(\mathtt{z}) - \sum_{k=1}^{r} |q_k(\mathtt{z})|^2$$

where $g,\, q_1, \dots, q_r \in \mathbb{C}[\mathtt{z}]$. Assume for every $j = 1, \dots, n$ that z_j is a polynomial in $g,\, q_1, \dots, q_r$, that is,

$$\mathbb{C}[g,\, q_1, \dots, q_r] \;=\; \mathbb{C}[z_1, \dots, z_n].$$

Then every commuting tuple T satisfying $f(T, T^*) = 0$ is subnormal. In other words, the principal ideal $I = (f)$ has property (S).

Proof. Choose a real number $c > 0$ so large that the operator $A := g(T) + c\,\mathrm{id}$ is invertible. From $2c\,\operatorname{Re}(g) = |g + c|^2 - c^2 - |g|^2$ we get

$$2cf \;=\; |g + c|^2 - c^2 - |g|^2 - 2c \sum_k |q_k|^2.$$

This implies

$$A^* A \;=\; c^2\,\mathrm{id} + g(T)^* g(T) + 2c \sum_k q_k(T)^* q_k(T),$$

hence suitable scalings of the commuting operators

$$A^{-1},\; g(T)A^{-1},\; q_1(T)A^{-1},\; \dots,\; q_r(T)A^{-1}$$

satisfy the identity of the sphere. Therefore the tuple consisting of these operators is subnormal, by Athavale's theorem [2]. Using rational functional calculus in conjunction with the spectral inclusion theorem [13], we conclude that the tuple $\big(g(T), q_1(T), \dots, q_k(T)\big)$ commutes and is subnormal. Now the hypothesis implies that the tuple $T = (T_1, \dots, T_n)$ is subnormal. □

5.2. We discuss yet another class of identities that entail the subnormality condition (S), this time in one variable ($n = 1$). Let $f \in \mathbb{R}[x, y]$ have the form

$$f \;=\; |g(z)|^2 - a - |l(z)|^2 - \sum_{k=1}^{r} |q_k(z)|^2$$

where $a > 0$ is a real number, g, $q_1, \dots, q_r \in \mathbb{C}[z]$ are arbitrary polynomials and $l \in \mathbb{C}[z]$ has degree one. The identity $f(T, T^*) = 0$ implies $g(T)^* g(T) \geq aI$, and we conclude that $g(T)$ is invertible. Inverting $g(T)$ we again arrive at a sphere identity, and arguing as in 5.1 we conclude in particular that $l(T)$ is subnormal, whence T is subnormal.

This construction can also be performed in any number of variables.

5.3. In certain cases we can prove that $V_{\mathbb{R}}(f)$ is compact and Σ_h is not archimedean modulo f, for f as in 5.2. Indeed, let

$$f \;=\; |z|^{2m} - \sum_{j=0}^{m-1} a_j |z|^{2j}$$

with $m \geq 2$ and real coefficients $a_0, \dots, a_{m-1} \geq 0$. Then $\Sigma_h + (f)$ is not archimedean. Indeed, assume $c - |z|^2 + fg \in \Sigma_h$, with $c \in \mathbb{R}$ and $g = g^* \in \mathbb{C}[z, \overline{z}]$. Let b_j be the coefficient of $|z|^{2j}$ in g. For any $j \geq 0$, the coefficient of $|z|^{2j}$ in $c - |z|^2 + fg$ is ≥ 0. For $j = 1$ this gives

$$-1 - a_1 b_0 - a_0 b_1 \geq 0, \tag{2}$$

while for $j = m + k$ with $k \geq 0$ it gives

$$b_k - a_{m-1} b_{k+1} - \cdots - a_0 b_{k+m} \;\geq\; 0 \;(k \geq 0). \tag{3}$$

Let $l \geq 0$ be the largest index for which $b_l \neq 0$ (by (2), there has to be such l). From (3) for $k = l$ we get $b_l > 0$. By a downward induction, repeatedly using (3), we conclude that $b_k \geq 0$ holds for all $k \geq 0$. But this contradicts (2). On the other hand, property (2) holds as soon as $a_0 > 0$ and $a_1 > 0$, see 5.2. The zero set $V_{\mathbb{R}}(f)$ is the union of (at least one, at most $m - 1$) concentric circles around 0.

Next, we look at the simplest case, which is plane conics. The following theorem shows that we can completely decide in which cases the various properties discussed so far are satisfied. In particular, it turns out that properties (S) and (Sf) are equivalent for plane conics:

Theorem 5.4. *Consider a nonconstant polynomial*

$$f \;=\; az\overline{z} + \alpha z^2 + \overline{\alpha z}^2 + \beta z + \overline{\beta z} + c$$

with a, $c \in \mathbb{R}$ and α, $\beta \in \mathbb{C}$, and let (f) be the principal ideal generated by f in $\mathbb{R}[x, y]$.

(a) (f) *has the Archimedean property* (A) *if, and only if,* $\alpha = 0$ *and* $a \neq 0$.
(b) (f) *has the Quillen property* (Q) *if, and only if,* $\alpha = 0$.
(c) *Properties* (S), (Sf) *and* (G) *for the ideal* (f) *are equivalent among each other, and are also equivalent to* $(a \neq 0 \ \vee \ a = \alpha = 0)$.

Note that the result [15, Thm. 3.3] on the ellipse is contained in (c) as a particular case.

For the proof of the theorem we need the following simple observation. (A similar argument was used in [15], proof of Proposition 3.1.) The leading form $\mathrm{lf}(f)$ of $0 \neq f \in \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]$ is the nonvanishing homogeneous part of f of highest degree. Clearly $\mathrm{lf}(fg) = \mathrm{lf}(f)\mathrm{lf}(g)$ and $\mathrm{lf}(f^*) = \mathrm{lf}(f)^*$.

Lemma 5.5. *(n arbitrary) For* $0 \neq f = f^* \in \Sigma_h$ *we have* $\mathrm{lf}(f) \in \Sigma_h$. *In particular, when* $n = 1$, *this implies* $\mathrm{lf}(f) = a(z\overline{z})^m$ *where* $\deg(f) = 2m$ *and* $0 < a \in \mathbb{R}$.

The lemma is obvious since in a sum $\sum_j |q_j(\mathbf{z})|^2$, no cancellation of leading forms can occur.

Proof of Theorem 5.4. Assume $\alpha = 0$ and $a \neq 0$. Then the identity

$$af \ = \ \left|az + \overline{\beta}\right|^2 + (ac - |\beta|^2),$$

combined with Proposition 3.2, shows that $\Sigma_h + (f)$ is archimedean and (hence) contains every polynomial g with $g > 0$ on $V_{\mathbb{R}}(f)$. If $\alpha = a = 0$ then f is linear, and after a holomorphic change of variables we may assume $f = \frac{1}{2i}(z - \overline{z}) = y$. By Remark 3.4.1 we see that $\Sigma_h + (f)$ contains every $g \in \mathbb{R}[x, y]$ with $g \geq 0$ on $V_{\mathbb{R}}(f)$.

Conversely, we show that $\Sigma_h + (f)$ cannot contain all polynomials strictly positive on $V_{\mathbb{R}}(f)$ when $\alpha \neq 0$. Indeed, assume $\alpha \neq 0$ and choose $\gamma \in \mathbb{C}$ with $\gamma \notin V_{\mathbb{R}}(f)$. For sufficiently small real $r > 0$, the polynomial $g = |z - \gamma|^2 - r^2$ is strictly positive on $V_{\mathbb{R}}(f)$. Assuming $g \in \Sigma_h + (f)$ would mean $g + fh \in \Sigma_h$ for some $h \in \mathbb{C}[z, \overline{z}]$, and necessarily $h \neq 0$. When h is constant then $l(g + fh)$ contains λz^2 for some $\lambda \neq 0$, contradicting Lemma 5.5. Otherwise $\deg(h) > 0$, and then $\mathrm{lf}(f)$ divides $\mathrm{lf}(fh) = \mathrm{lf}(g + fh)$, again contradicting 5.5.

We have thus proved (a) and (b). For the proof of (c) we easily dispense with the linear case $a = \alpha = 0$, and can assume $\deg(f) = 2$. If $a = 0$ then $f = g + g^*$ with a quadratic holomorphic polynomial $g \in \mathbb{C}[z]$. For generic choice of $t \in \mathbb{R}$ there are two different numbers $\alpha \neq \beta$ in $\mathbb{C}$ with $g(\alpha) = g(\beta) = it$. For any such pair we have $f(\alpha) = f(\beta) = f(\alpha, \overline{\beta}) = 0$, and hence $f \in I(\alpha, \beta)$. For $a = 0$, therefore, the ideal (f) does not satisfy condition (G), and *a fortiori* does not satisfy conditions (S) and (Sf), by 4.4.

On the other hand, assume $a \neq 0$. Then f has the form

$$f \ = \ \mathrm{Re}\, g(z) + a|z|^2$$

with $g \in \mathbb{C}[z]$. According to Proposition 5.1, every bounded operator T satisfying $f(T, T^*) = 0$ is subnormal. So (f) satisfies conditions (S), (Sf) and (G) in this case. □

Remarks 5.6.

1. We rephrase part of Theorem 5.4 in geometric terms, and assume $\deg(f) = 2$ and $V_{\mathbb{R}}(f) \neq \varnothing$ for simplicity. Then $\Sigma_h + (f)$ contains all polynomials positive on $V_{\mathbb{R}}(f)$ if and only if $V_{\mathbb{R}}(f)$ is a circle. On the other hand, the identity $f(T, T^*) = 0$ implies subnormality for a bounded operator T if and only if $V_{\mathbb{R}}(f)$ is *not* a hyperbola with perpendicular asymptotes (and neither a union of two perpendicular lines).
2. From Theorem 5.4 we see in particular that there exist ideals $I \subseteq \mathbb{R}[x, y]$ with $V_{\mathbb{R}}(I) = \varnothing$ for which $-1 \notin \Sigma_h + I$. This is in striking contrast to the case of usual sums of squares, where it is well known that $V_{\mathbb{R}}(I) = \varnothing$ implies $-1 \in \Sigma + I$. Such ideals may well have the subnormality property (S). For example, this is so for $I = (ax^2 + by^2 + c)$ with a, b, $c > 0$ and $a \neq b$.

5.7. For (reduced) plane conics, the normality condition (Sf) for finite-dimensional Hilbert spaces already implies the subnormality condition (S) for arbitrary Hilbert spaces, as shown in 5.4. We now show that this ceases to hold when we take a suitable nonreduced version of a conic.

To this end consider the $*$-invariant ideal

$$J \;=\; \Big((z-1)(z\overline{z}-1),\; (\overline{z}-1)(z\overline{z}-1)\Big)$$

in $\mathbb{C}[z, \overline{z}]$, respectively its real version

$$I \;=\; J \cap \mathbb{R}[x, y] \;=\; (x^2 + y^2 - 1) \cdot (x - 1, y).$$

The ideal corresponds to the unit circle with nilpotents added at one point. We will see that hermitian sums of squares modulo I behave quite different than modulo $\sqrt{I}$.

Let $T \in B(E)$ satisfy $I \subseteq \ker(\psi_T)$, that is,

$$(T^* - \mathrm{id})(T^*T - \mathrm{id}) \;=\; 0. \tag{4}$$

We decompose the Hilbert space as $E = \ker(T - \mathrm{id}) \oplus \ker(T - \mathrm{id})^{\perp}$. With respect to this decomposition, T has a block matrix representation

$$T \;=\; \begin{pmatrix} \mathrm{id} & A \\ 0 & B \end{pmatrix}$$

with $\ker(A) \cap \ker(B - \mathrm{id}) = \{0\}$. From (4) we deduce $A(B - \mathrm{id}) = 0$ and $A^*A + (B^*B - \mathrm{id})(B - \mathrm{id}) = 0$. The second identity implies that $B - \mathrm{id}$ is actually injective. If $\dim(E) < \infty$, then $B - \mathrm{id}$ is invertible, and we get $A = 0$ and $B^*B - \mathrm{id} = 0$. In short, T is unitary. Every (finite-dimensional) matrix annihilated by the ideal J is therefore unitary, and hence normal.

On the other hand, we will produce an operator T acting on $E = \ell^2(\mathbb{N})$ such that T is annihilated by the ideal J and T is not subnormal. Let E have Hilbert basis e_k $(k \geq 0)$, and let $S : e_k \mapsto e_{k+1}$ $(k \geq 0)$ be the unilateral shift. Let π be the orthogonal projection onto the space generated by e_0, and define $T = S + \pi$. A direct computation, supported by the relations

$$S^*S = \mathrm{id},\; SS^* = \mathrm{id} - \pi,\; \pi S = 0,$$

yields $(T^*-\mathrm{id})(T^*T-\mathrm{id})=0$. But T is not subnormal. Indeed, any subnormal operator X satisfies the hyponormality inequality $[X^*,X]\geq 0$, evident from the matrix decomposition of a normal extension $N=\left(\begin{smallmatrix}X & Y\\ 0 & Z\end{smallmatrix}\right)$ and the equation $[N^*,N]=0$. The commutator $[T^*,T]=\pi-S\pi-\pi S^*$ acts on the span of e_0 and e_1 as $\begin{pmatrix}1 & -1\\ -1 & 0\end{pmatrix}$, and this is not a nonnegative operator. In summary, this shows:

Proposition 5.8. *Let $I\subseteq\mathbb{R}[x,y]$ be the above ideal, corresponding to the unit circle with a thickened point. Then I satisfies condition* (Sf)*, but not* (S) *(and a fortiori, not* (A)*). In contrast, its reduced version $\sqrt{I}$ satisfies* (A) *(and therefore also* (S) *and* (Sf)*).* □

6. Semialgebraic sets

6.1. In Section 3 we studied the question whether every polynomial strictly positive on a real algebraic set $X\subseteq\mathbb{C}^n$ is a hermitian sum of squares on X. We now extend this question to (real) semialgebraic subsets of $\mathbb{C}^n$. Algebraically, this means that instead of an ideal $I\subseteq\mathbb{R}[\mathtt{x},\mathtt{y}]$ and the semiring Σ_h+I we consider *hermitian modules*, that is, modules over the semiring Σ_h (see 2.4). This means that the real algebraic set $X=V_{\mathbb{R}}(I)$ is replaced by the closed set

$$X_M \;=\; \{a\in\mathbb{C}^n \mid \text{ for all } f\in M\ f(a)\geq 0\}$$

(see 2.6). If the hermitian module M is finitely generated (or, more generally, if the quadratic module generated by M is finitely generated), the closed set X_M is basic closed, i.e., there are finitely many $f_1,\dots,f_k\in\mathbb{R}[\mathtt{x},\mathtt{y}]$ with $X_M=\{a\in\mathbb{C}^n \mid f_1(a)\geq 0,\dots,f_k(a)\geq 0\}$.

6.2. Each of the four properties of an ideal $I\subseteq\mathbb{R}[\mathtt{x},\mathtt{y}]$ labelled (A), (Q), (S), (Sf) that were discussed in the first part of this paper is in fact a property of the semiring $S=\Sigma_h+I$, i.e., can be expressed in terms of S. We now extend these properties to arbitrary hermitian modules $M\subseteq\mathbb{R}[\mathtt{x},\mathtt{y}]$:

- (A) M is archimedean;
- (Q) M contains every $f\in\mathbb{R}[\mathtt{x},\mathtt{y}]$ with $f>0$ on X_M;
- (S) every commuting tuple T of bounded operators in a Hilbert space satisfying $p(T,T^*)\geq 0$ for every $p\in M$ is subnormal;
- (Sf) every commuting tuple T of complex matrices satisfying $p(T,T^*)\geq 0$ for every $p\in M$ is normal.

When $M=\Sigma_h+I$ for some ideal I, the above properties agree with the respective properties of the ideal I, as defined in 3.1 and 3.8.

We start by the following characterization of archimedean hermitian modules, thereby generalizing part of Proposition 3.2:

Lemma 6.3. *A hermitian module $M\subseteq\mathbb{R}[\mathtt{x},\mathtt{y}]$ is archimedean if and only if $c-||\mathtt{z}||^2\in M$ for some real number c.*

Proof. Assuming $c - ||\mathtt{z}||^2 \in M$ we have to show $\mathbb{R} + M = \mathbb{R}[\mathtt{x}, \mathtt{y}]$. For this let $A := \{p \in \mathbb{C}[\mathtt{z}] \mid -|p|^2 \in \mathbb{R} + M\}$. It suffices to prove $A = \mathbb{C}[\mathtt{z}]$. Indeed, any $f \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$ can be written $f = \sum_j |p_j|^2 - \sum_k |q_k|^2$ with p_j, $q_k \in \mathbb{C}[\mathtt{z}]$; if $q_k \in A$ for every k, then $f \in \mathbb{R} + M$.

From $c - |z_j|^2 = (c - ||\mathtt{z}||^2) + \sum_{k \neq j} |z_k|^2$ we see $z_j \in A$ for $j = 1, \ldots, n$. Therefore (and since $\mathbb{C} \subseteq A$) it is enough to prove that A is a ring. From $a - |f|^2$, $b - |g|^2 \in M$ with a, $b \geq 0$ we get

$$ab - |fg|^2 = a(b - |g|^2) + |g|^2(a - |f|^2) \in M,$$

so A is closed under products. From $|f + g|^2 + |f - g|^2 = 2(|f|^2 + |g|^2)$ we see that A is also closed under sums. The lemma is proved. □

Before we start discussing a Positivstellensatz for hermitian modules, we need to mention a subtle point. The archimedean Positivstellensatz 2.6 holds for modules over archimedean semirings, but *not* in general for archimedean modules over semirings. This distinction is relevant for hermitian modules, as the following example shows:

Example 6.4. The hermitian module $M = \Sigma_h + \Sigma_h(1 - ||\mathtt{z}||^2)$ is archimedean by Lemma 6.3. But there exist polynomials that are strictly positive on the closed unit ball X_M and are not contained in M. In fact, $\epsilon + (1 - ||\mathtt{z}||^2)^2$ is such a polynomial for $0 < \epsilon < 1$. To see this, assume

$$\epsilon + (1 - ||\mathtt{z}||^2)^2 = p + q(1 - ||\mathtt{z}||^2)$$

with p, $q \in \Sigma_h$. Comparing constant coefficients gives $q(0) \leq 1 + \epsilon$, while comparing coefficients of $z_1\overline{z}_1$ gives $-2 \geq -q(0)$, i.e. $q(0) \geq 2$, since the coefficient of $z_1\overline{z}_1$ in any hermitian sum of squares is nonnegative.

The point is that, although the hermitian module M is archimedean, M is not a module over any archimedean semiring.

The proper "quantization" of the module M is a linear operator T acting on a Hilbert space, subject to the contractivity condition

$$I - T^*T \geq 0.$$

It is clear that not every contractive operator T is subnormal.

Example 6.4 has shown (c.f. also Proposition 3.2):

Lemma 6.5. *Consider the following two properties of a hermitian module M in $\mathbb{R}[\mathtt{x}, \mathtt{y}]$:*

(i) *X_M is compact, and M has the Quillen property* (Q)*;*
(ii) *M is archimedean* (A).

Then (i) implies (ii), but the converse fails in general. □

Here we are mainly interested in a Positivstellensatz, that is, in the Quillen property (Q), in the case when X_M is compact. Therefore, we will often *assume* that M is archimedean (which implies that X_M is compact), and try to find additional properties for M that will imply the Positivstellensatz. Verifying the archimedean property of a concretely given hermitian module M is usually easy, using the criterion of Lemma 6.3.

One instance where we get the Positivstellensatz for free is the following:

Proposition 6.6. *Let $I \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$ be any ideal with the archimedean property* (A)*. Then for any $p_1, \dots, p_r \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$, the hermitian module*

$$M = I + \Sigma_h + p_1\Sigma_h + \cdots + p_r\Sigma_h$$

has the Quillen property (Q)*.*

Proof. M is a module over the archimedean semiring $\Sigma_h + I$, so the assertion follows from the archimedean Positivstellensatz 2.6. □

Remarks 6.7.

1. For M as in 6.6, the associated semialgebraic set is

$$X_M = V_{\mathbb{R}}(I) \cap \Big\{ a \in \mathbb{C}^n \;\Big|\; p_1(a) \geq 0, \dots, p_r(a) \geq 0 \Big\}.$$

2. In the particular case $I = (1 - ||\mathtt{z}||^2)$, Proposition 6.6 was proved by D'Angelo and Putinar ([5, Thm. 3.1]).

Generalizing Proposition 3.9, the Positivstellensatz for M implies the following subnormality property for commuting tuples T of bounded operators.

Corollary 6.8. *For any hermitian module $M \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$, we have* (Q) $\Rightarrow$ (S)*.*

Proof. The proof of Proposition 3.9 carries over (essentially) verbatim. □

We next discuss a refinement of the last corollary, in which we are going to weaken condition (Q) and strengthen condition (S).

6.9. Let $M \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$ be a hermitian module. Recall [20] that M is said to have the *Strong Moment Property* if the following holds:

(SMP) Every linear functional $L : \mathbb{R}[\mathtt{x}, \mathtt{y}] \to \mathbb{R}$ with $L|_M \geq 0$ is integration with respect to some positive Borel measure on X_M.

Lemma 6.10. *Let $M \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$ be a hermitian module. The Quillen property* (Q) *for M implies the strong moment property* (SMP) *for M. The converse is true if M is archimedean.*

Proof. Assume that M has property (Q). Then any $f \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$ nonnegative on X_M lies in the closure $\overline{M}$ with respect to the finest locally convex topology on $\mathbb{R}[\mathtt{x}, \mathtt{y}]$. So $L|_M \geq 0$ implies $L(f) \geq 0$, and therefore L is integration with respect to some positive Borel measure on X_M, according to the Riesz–Haviland theorem [8]. Conversely, let M be archimedean and satisfy (SMP), and assume that there is $f \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$ with $f > 0$ on X_M but $f \notin M$. By Eidelheit's separation theorem (e.g., [10, §17.1]) there is a linear functional $L : \mathbb{R}[\mathtt{x}, \mathtt{y}] \to \mathbb{R}$ with $L(1) = 1$, $L|_M \geq 0$ and $L(f) \leq 0$. By assumption, L is integration with respect to a probability measure μ on X_M. We conclude $L(f) = \int_{X_M} f \, d\mu > 0$, a contradiction. □

Lemma 6.11. *For M a hermitian module in $\mathbb{R}[\mathtt{x},\mathtt{y}]$, consider the following property:*

(SOS) $\Sigma \subseteq \overline{M}$ *(closure with respect to the finest locally convex topology on $\mathbb{R}[\mathtt{x},\mathtt{y}]$).*

Then property (SOS) *for M implies property* (S) *for M. The converse is true if M is archimedean.*

Proof. (SOS) $\Rightarrow$ (S): Let T be a commuting tuple of operators with $M \subseteq M_T$. Since M_T is closed in $\mathbb{R}[\mathtt{x},\mathtt{y}]$ (3.5) we have $\overline{M} \subseteq M_T$. So assumption (SOS) implies $\Sigma \subseteq M_T$, and hence T is subnormal according to Lemma 3.7.

(S) $\Rightarrow$ (SOS): Let M be archimedean and have property (S). Assume that there is $f \in \Sigma$ with $f \notin \overline{M}$. There exists a linear functional $L : \mathbb{R}[\mathtt{x},\mathtt{y}] \to \mathbb{R}$ satisfying $L|_M \geq 0$ and $L(f) < 0$. We extend L to a complex linear functional on $\mathbb{R}[\mathtt{x},\mathtt{y}] \otimes_{\mathbb{R}} \mathbb{C} = \mathbb{C}[\mathtt{z},\overline{\mathtt{z}}]$ and perform a GNS construction: Consider the positive semidefinite inner product $\langle p, q\rangle := L(pq^*)$ on $\mathbb{C}[\mathtt{z}]$, and let E be the corresponding Hilbert space completion of $\mathbb{C}[\mathtt{z}]$. Since M is archimedean, there is a real number $c > 0$ with $c - |z_j|^2 \in M$ for $j = 1, \dots, n$, and it follows for any $p \in \mathbb{C}[\mathtt{z}]$ that

$$L(|z_j|^2\,|p(\mathtt{z})|^2) \;\leq\; c \cdot L(|p(\mathtt{z})|^2).$$

Hence multiplication by z_j induces a bounded linear operator T_j on E ($j = 1, \dots, n$), and $T = (T_1, \dots, T_n)$ is a commuting tuple in $B(E)$. For $g \in \mathbb{C}[\mathtt{z},\overline{\mathtt{z}}]$ and $q \in \mathbb{C}[\mathtt{z}]$ we have $\langle \psi_T(g)q,\, q\rangle = L(g|q|^2)$. So for $g \in M$ the operator $\psi_T(g)$ is nonnegative, which means $M \subseteq M_T$. On the other hand, $f \notin M_T$ since $\langle \psi_T(f)1, 1\rangle = L(f) < 0$. Therefore, the tuple T is not subnormal, according to Proposition 3.7. This contradicts property (S). □

Lemma 6.12. *For any hermitian module M we have* (SMP) $\Rightarrow$ (SOS).

Proof. By hypothesis, any linear functional $L : \mathbb{R}[\mathtt{x},\mathtt{y}] \to \mathbb{R}$ with $L|_M \geq 0$ is integration with respect to a measure on X_M. In particular, $L(f) \geq 0$ for any $f \in \Sigma$. This implies $f \in \overline{M}$. □

Remarks 6.13.

1. The implication (S) $\Rightarrow$ (SOS) for archimedean M (Lemma 6.11) is uninteresting if M is a module over an archimedean semiring. Indeed, in this case we know anyway that M contains all polynomials strictly positive on X_M, and hence $\Sigma \subseteq \overline{M}$ is clear. But in the other cases, the equivalence of (S) and (SOS) is a new information.
2. Altogether we have now obtained the chain of implications

$$\text{(Q)} \;\Rightarrow\; \text{(SMP)} \;\Rightarrow\; \text{(SOS)} \;\Rightarrow\; \text{(S)}$$

for any hermitian module M. When M is archimedean, the first and the last implication can be reversed.

6.14. Under a stronger condition on M, we are now going to prove the implication (SOS) $\Rightarrow$ (SMP), and hence the equivalence of (Q) and (S), when M is archimedean. Recall that a closed subset $K \subseteq \mathbb{C}^n$ is said to be *polynomially* *convex* if the following holds: For every $\xi \in \mathbb{C}^n$ with $\xi \notin K$, there exists a polynomial $p \in \mathbb{C}[\mathtt{z}]$ with $|p| \le 1$ on K and $|p(\xi)| > 1$. We shall consider the following property for a hermitian module $M \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$:

(PC) For every $\xi \in \mathbb{C}^n \setminus X_M$, there exist $f \in M$ and $q \in \Sigma_h$ such that $q \le 1$ on $\{a \in \mathbb{C}^n \mid f(a) \ge 0\}$, and such that $q(\xi) > 1$.

Remarks 6.15.

1. If M satisfies condition (PC), then the set X_M is polynomially convex in $\mathbb{C}^n$. Indeed, X_M has the form
$$X_M = \bigcap_\nu \{a \in \mathbb{C}^n \mid q_\nu(a) \le 1\}$$
for some family of polynomials $q_\nu \in \Sigma_h$. For each ν, the set $\{a \in \mathbb{C}^n \mid q_\nu(a) \le 1\}$ is polynomially convex, since it is the preimage of the closed unit ball in some $\mathbb{C}^m$ under a polynomial map $\mathbb{C}^n \to \mathbb{C}^m$. Therefore X_M is polynomially convex.
2. Condition (PC) is satisfied when the hermitian module M is generated by polynomials of the form $1 - q_\nu$ with $q_\nu \in \Sigma_h$. More generally, (PC) holds when M contains a family $\{f_\nu\}$ of polynomials with $X_M = \bigcap_\nu \{a \in \mathbb{C}^n \mid f_\nu(a) \ge 0\}$ such that, for every ν, the set $\{f_\nu \ge 0\}$ is polynomially convex.

Theorem 6.16. *Let M be an archimedean hermitian module in $\mathbb{R}[\mathtt{x}, \mathtt{y}]$ which satisfies condition* (PC). *Then the subnormality property* (S) *implies the Quillen property* (Q) *for M.*

Proof. By Lemma 6.10 it suffices to show that M has the strong moment property (SMP). Given a linear functional $L : \mathbb{R}[\mathtt{x}, \mathtt{y}] \to \mathbb{R}$ with $L|_M \ge 0$, we need to show that L is integration with respect to a positive Borel measure supported on X_M. Similar to the proof of Lemma 6.11, we use a GNS construction to get a Hilbert space E together with a commuting tuple $T = (T_1, \dots, T_n)$ in $B(E)$ and a cyclic vector ξ, such that $L(pq^*) = \langle p(T)\xi, q(T)\xi \rangle$ for all $p, q \in \mathbb{C}[\mathtt{z}]$. Since M has property (S), the tuple T is subnormal. The spectral measure of a commuting normal extension S of T gives a Borel measure μ on $\mathbb{C}^n$ with
$$L(f) = \langle \psi_S(f)\xi, \xi \rangle = \int f \, d\mu$$
for all $f \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$. It remains to prove $\operatorname{supp}(\mu) \subseteq X_M$, which follows from the next lemma. □

Lemma 6.17. *Let M be a hermitian module, and let μ be a positive measure on $\mathbb{C}^n$ all of whose moments exist, satisfying $\int f \, d\mu \ge 0$ for every $f \in M$. If M satisfies condition* (PC), *then* $\operatorname{supp}(\mu) \subseteq X_M$.

Proof. Assume there exists $\xi \in \operatorname{supp}(\mu)$ with $\xi \notin X_M$. By (PC) we find $f \in M$ and $q \in \Sigma_h$ such that $q(\xi) > 1$ and $q < 1$ on $\{f \ge 0\}$. We have
$$\int f q^m \, d\mu = \int_{\{f \ge 0\}} f q^m \, d\mu + \int_{\{f < 0\}} f q^m \, d\mu.$$

For $m \to \infty$, the first summand on the right tends to zero by the dominated convergence theorem. On the other hand, the second summand tends to $-\infty$: Consider a small ball B around ξ on which $q \geq a > 1$ and $f \leq b < 0$, and note that B has positive μ-measure. So there is $m \in \mathbb{N}$ for which the integral on the left is negative, a contradiction since $fq^m \in M$. □

Summarizing Lemmas 6.10–6.12 and Proposition 6.16, we obtain:

Theorem 6.18. *Let M be a hermitian module in $\mathbb{R}[\mathtt{x}, \mathtt{y}]$ which is archimedean and satisfies condition* (PC)*. Then for M we have*

$$(\mathrm{Q}) \;\Leftrightarrow\; (\mathrm{SMP}) \;\Leftrightarrow\; (\mathrm{SOS}) \;\Leftrightarrow\; (\mathrm{S}).$$ □

Example 6.19. Without any hypothesis like (PC), the implication (S) ⇒ (Q) is false, even if we assume that M is archimedean. Indeed, let $n = 1$ and $0 < r < R$, and consider the Σ_h-module

$$M \;=\; \Sigma + \Sigma_h(R^2 - |z|^2) + \Sigma_h(|z|^2 - r^2).$$

Here X_M is the annulus around the origin with radii $r < R$. Clearly, M is archimedean (6.3) and satisfies condition (S) (Lemma 6.11). But there exists a compactly supported measure μ on $\mathbb{C}$ with $\mathrm{supp}(\mu) \not\subseteq X_M$ and with

$$\int f(z)\,\mu(dz) \;\geq\; 0$$

for every $f \in M$. Namely, let $r < \rho < R$, and let

$$\int f\,d\mu \;:=\; \epsilon f(0) + \int_{-\pi}^{\pi} f(\rho e^{it})\,dt.$$

When $\epsilon > 0$ is sufficiently small, we have $\int (|z|^2 - r^2)|p(z)|^2\,d\mu \geq 0$ for every $p \in \mathbb{C}[z]$, and hence $\int f(z)\,\mu(dz) \geq 0$ for every $f \in M$. Namely, the integral is

$$(\rho^2 - r^2)\int_{-\pi}^{\pi} |p(\rho e^{it})|^2\,dt - \epsilon r^2|p(0)|^2 \;\geq\; \big(2\pi(\rho^2 - r^2) - \epsilon r^2\big)|p(0)|^2.$$

Note that the annulus X_M is not polynomially convex, and so the hypotheses of Theorem 6.16 are not satisfied.

Theorem 6.20. *Let $M \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$ be a finitely generated hermitian module. If M satisfies* (Sf)*, then the semialgebraic set X_M in $\mathbb{C}^n$ does not contain an analytic disc.*

Proof. By an analytic disc we mean the image of a nonconstant holomorphic map $\varphi : \mathbb{D} \to \mathbb{C}^n$. We can assume $\varphi(0) = 0$, and since we work locally, we can assume that there exists an analytic function $F : \varphi(\mathbb{D}) \to \mathbb{D}$ such that $F(\varphi(\zeta)) = \zeta$ ($\zeta \in \mathbb{D}$). By assumption, $\varphi(\mathbb{D}) \subseteq X_M$. Let M be generated by nonzero polynomials $f_1, \dots, f_r \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$. We reorder the generators so that $f_1, \dots, f_s$ vanish identically on $\varphi(\mathbb{D})$, and $f_{s+1}, \dots, f_r$ have only isolated zeros on $\varphi(\mathbb{D})$. By passing to an appropriate subdisc on $\mathbb{D}$ we can assume $f_{s+1}(0) > 0, \dots, f_r(0) > 0$.

Let $\epsilon > 0$ be sufficiently small, and choose a non-normal matrix A of norm $||A|| \leq \epsilon$. Then the commuting tuple of matrices $\varphi(A)$ is not normal, as $A = F(\varphi(A))$ by the superposition property of the analytic functional calculus. In addition, $M \subseteq M_{\varphi(A)}$. Indeed, fixing $1 \leq j \leq s$, the composite function $f_j(\varphi(z), \overline{\varphi(z)})$ is identically zero. If

$$f_j(\mathtt{z}, \overline{\mathtt{z}}) \;=\; \sum_{\alpha,\beta} c_{\alpha,\beta}\, \mathtt{z}^\alpha \overline{\mathtt{z}}^\beta,$$

this means that the power series

$$\sum_{\alpha,\beta} c_{\alpha,\beta}\, \varphi_1(z)^{\alpha_1} \cdots \varphi_n(z)^{\alpha_n}\, \overline{\varphi_1(z)}^{\beta_1} \cdots \overline{\varphi_n(z)}^{\beta_n}$$

in z and $\overline{z}$ is identically zero, from which we see $f_j(\varphi(A)) = 0$. On the other hand, $f_{s+1}(\varphi(A)) > 0, \ldots, f_r(\varphi(A)) > 0$ by the continuity of the functional calculus.

Since $\varphi(A)$ is not (sub-) normal, this implies that M does not satisfy (Sf). ☐

Remark 6.21. In the case where $M = \Sigma_h + I$ for some ideal $I \subseteq \mathbb{R}[\mathtt{x}, \mathtt{y}]$, Theorem 6.20 also follows from Theorem 4.4. Indeed, assume that $\varphi : \mathbb{D} \to \mathbb{C}^n$ is a holomorphic map with $\varphi(\mathbb{D}) \subseteq V_{\mathbb{R}}(I)$ and with $u := \nabla_z \varphi(0) \neq 0$. A direct calculation shows that $I \subseteq J(a, U)$ (see 4.2) with $a := \varphi(0)$ and $U := u^* u$ (we consider u as a row vector). Therefore, the ideal I does not satisfy condition (G), and by Theorem 4.4, it neither satisfies (Sf).

By Corollary 6.8, Theorem 6.20 implies:

Corollary 6.22. *Let M be a finitely generated hermitian module. If X_M contains an analytic disc, then M does not satisfy Quillen's property* (Q). ☐

Proof. Let M be generated by nonzero polynomials $f_1, \ldots, f_r \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$. Since X_M contains an analytic disc, the subnormality condition (S) does not hold for M, by Theorem 6.20 above. According to Corollary 6.8, the Positivstellensatz does not hold either. ☐

7. Historical comments

We feel that leaving aside the analytic roots of the questions encountered in this article would deprave the reader of some essential insight. We briefly describe below some of the old sources and applications of the decomposition of a real polynomial in a sum of hermitian squares.

Start with Riesz–Fejér Theorem asserting that a polynomial $p(z, \overline{z})$ which is nonnegative on the unit circle can be decomposed as

$$p(z, \overline{z}) \;=\; |h(z)|^2 + (1 - |z|^2) g(z, \overline{z}), \tag{5}$$

where $h \in \mathbb{C}[z]$ and $g \in \mathbb{C}[z, \overline{z}]$.

Next we "quantize" the above setting, that is we replace the complex variable by a linear transformation. Let T be a bounded linear operator acting on a Hilbert space E and denote by T^* its adjoint. The simple operator identity

$$T^*T = \mathrm{id}$$

defines an isometric transformation, with the known consequences: spectral picture, functional model and classification, see [4]. In particular, the operator T has in this case spectrum contained in the closed unit disc $\overline{\mathbb{D}}$, and for every real valued polynomial $p(z,\overline{z})$ the estimate

$$p(T,T^*) \le \max_{\lambda\in\mathbb{T}} p(\lambda,\overline{\lambda})\,\mathrm{id} \tag{6}$$

holds true. Recall that here we adopt the hereditary calculus convention, putting the powers of T^* to the left of the powers of T, in every monomial appearing in p.

Inequality (6) is a simple consequence of (5): If $p(z,\overline{z}) \le M$ on $\mathbb{T}$, then

$$M - p(T,T^*) = h(T)^*h(T) + [(1-|z|^2)g(z,\overline{z})](T,T^*) = h(T)^*h(T) \ge 0. \tag{7}$$

As a matter of fact, estimate (6) implies that the linear functional calculus $p(z,\overline{z}) \mapsto p(T,T^*)$ possesses an additional positivity property. The latter implies, essentially repeating F. Riesz construction of the representing measure for a positive functional, that the operator T is subnormal, that is, there exists a larger Hilbert space $E \subseteq K$ and a normal operator U acting on K, such that $U(E) \subseteq E$ and $U|_E = T$. By choosing U minimal with this property we can also assume that the spectrum of U is contained in the torus $\mathbb{T}$, hence U is unitary, see for instance [4]. In particular, if in addition $TT^* = \mathrm{id}$, that is T is unitary from the beginning, we obtain in this manner a proof of the spectral theorem, as advocated by F. Riesz from the dawn of functional analysis [17, 18].

Turning now to several complex variables, or their quantized form, commuting tuples of linear operators, we encounter Quillen's idea [16]. Let $P(z,\overline{z})$ be a conjugation-invariant polynomial, bihomogeneous of the same degree in the variables z and $\overline{z}$. Assume that $P(z,\overline{z}) > 0$ whenever $z \ne 0$. Denote by $M = (M_{z_1},\dots,M_{z_n})$ the n-tuple of commuting multipliers by the complex variables, on the Bargmann–Fock space of entire functions (square integrable in $\mathbb{C}^n$ with respect to the Gaussian weight). Using analytical tools (elliptic estimates and Fredholm theory), Quillen analyzes the positivity of the operator $P(M,M^*)$ inherited from the positivity of the symbol P. He reaches the purely algebraic conclusion that there exists a positive integer N and homogeneous complex analytic polynomials $h_1,\dots,h_k$ such that

$$\|z\|^{2N}P(z,\overline{z}) = |h_1(z)|^2 + \dots + |h_k(z)|^2. \tag{8}$$

Very recently Drout and Zworski [7] have obtained, using the same Bargmann–Fock space representation, degree bounds in Quillen's decomposition above.

An elementary dehomogenization argument shows that (8) implies that every positive polynomial on the unit sphere of $\mathbb{C}^n$ is equal, on the sphere, to a sum of hermitian squares, as stated by condition (Q) in our article.

On the abstract operator theory side, we mention the 1987 discovery of Athavale [2] stating that every commuting tuple of bounded operators $T = (T_1, \dots, T_n)$ subject to the sphere identity

$$T_1^* T_1 + \cdots + T_n^* T_n \; = \; \mathrm{id}$$

is subnormal, and hence possesses a functional calculus with a positivity property of type (6). Athavale's work belongs to a framework advocated for several dozen years by now by Conway [4], Agler and McCarthy [1] and their followers.

Quillen's theorem was rediscovered in 1996, generalized and put into the context of Cauchy–Riemann geometry and function theory of several complex variables by Catlin and D'Angelo [3]. Their proof also uses analysis, this time employing analytic Toeplitz operators acting on the Bergman space of the unit ball. One of the main themes of research in Cauchy–Riemann geometry is the (local) classification up to biholomorphic transformations of real algebraic subvarieties of $\mathbb{C}^n$. There is no surprise that Quillen property, or better its algebro-geometric consequences (Sf) and (G) are relevant for CR manifold theory. A modest step into this direction was taken in [6].

References

[1] J. Agler and J. E. McCarthy, *Pick interpolation and Hilbert function spaces*, Graduate Studies in Mathematics, 44, Amer. Math. Soc., Providence, RI, 2002. MR1882259 (2003b:47001)

[2] A. Athavale, Holomorphic kernels and commuting operators, Trans. Amer. Math. Soc. **304** (1987), no. 1, 101–110. MR0906808 (88m:47039)

[3] D. W. Catlin and J. P. D'Angelo, A stabilization theorem for Hermitian forms and applications to holomorphic mappings, Math. Res. Lett. **3** (1996), no. 2, 149–166. MR1386836 (97f:32025)

[4] J. B. Conway, *The theory of subnormal operators*, Mathematical Surveys and Monographs, 36, Amer. Math. Soc., Providence, RI, 1991. MR1112128 (92h:47026)

[5] J. P. D'Angelo and M. Putinar, Polynomial optimization on odd-dimensional spheres, in *Emerging applications of algebraic geometry*, 1–15, IMA Vol. Math. Appl., 149, Springer, New York. MR2500462 (2010m:14074)

[6] J. P. D'Angelo and M. Putinar, Hermitian complexity of real polynomial ideals, Internat. J. Math. **23** (2012), no. 6, 1250026, 14 pp. MR2925474

[7] A. Drouot and M. Zworski, A quantitative version of Catlin-D'Angelo–Quillen theorem, Anal. Math. Phys. **3** (2013), no. 1, 1–19. MR3015627

[8] E. K. Haviland, On the momentum problem for distributions in more than one dimension, II. Amer. J. Math. 58 (1936), no. 1, 164–168. MR1507139

[9] T. Itô, On the commutative family of subnormal operators, J. Fac. Sci. Hokkaido Univ. Ser. I **14** (1958), 1–15. MR0107177 (21 #5902)

[10] G. Köthe, *Topological vector spaces. I*, Translated from the German by D. J. H. Garling. Die Grundlehren der mathematischen Wissenschaften, Band 159, Springer-Verlag New York Inc., New York, 1969. MR0248498 (40 #1750)

[11] M. Laurent, Sums of squares, moment matrices and optimization over polynomials, in *Emerging applications of algebraic geometry*, 157–270, IMA Vol. Math. Appl., 149, Springer, New York. MR2500468 (2010j:13054)

[12] A. Prestel and C. N. Delzell, *Positive polynomials*, Springer Monographs in Mathematics, Springer, Berlin, 2001. MR1829790 (2002k:13044)
[13] M. Putinar, Spectral inclusion for subnormal n-tuples, Proc. Amer. Math. Soc. **90** (1984), no. 3, 405–406. MR0728357 (85f:47029)
[14] M. Putinar and C. Scheiderer, Sums of Hermitian squares on pseudoconvex boundaries, Math. Res. Lett. **17** (2010), no. 6, 1047–1053. MR2729629 (2011h:12002)
[15] M. Putinar and C. Scheiderer, Hermitian algebra on the ellipse, Illinois J. Math. **56** (2012), no. 1, 213–220 (2013). MR3117026
[16] D. G. Quillen, On the representation of hermitian forms as sums of squares, Invent. Math. **5** (1968), 237–242. MR0233770 (38 #2091)
[17] F. Riesz: Les systèmes d'équations linéaires à une infinité d'inconnues. Gauthier Villars, Paris, 1913.
[18] F. Riesz and B. Sz.-Nagy, *Functional analysis*, Translated from the second French edition by Leo F. Boron. Reprint of the 1955 original. Dover Books on Advanced Mathematics. Dover Publications, Inc., New York, 1990. MR1068530 (91g:00002)
[19] C. Scheiderer, Positivity and sums of squares: a guide to recent results, in *Emerging applications of algebraic geometry*, 271–324, IMA Vol. Math. Appl., 149, Springer, New York. MR2500469 (2010h:14092)
[20] K. Schmüdgen, On the moment problem of closed semi-algebraic sets, J. Reine Angew. Math. **558** (2003), 225–234. MR1979186 (2004e:47019)
[21] J. Stochel, Characterizations of subnormal operators, Studia Math. **97** (1991), no. 3, 227–238. MR1100688 (92m:47047)

Received January 29, 2014; accepted February 7, 2014

Mihai Putinar
Department of Mathematics, University of California at Santa Barbara
Santa Barbara, CA 93106-3080, USA
E-mail: `mputinar@math.ucsb.edu`

Mathematics Department, Newcastle University
Newcastle upon Tyne, UK
E-mail: `mihai.putinar@ncl.ac.uk`

Claus Scheiderer
Fachbereich Mathematik and Statistik, Universität Konstanz
78457 Konstanz, Germany
E-mail: `claus.scheiderer@uni-konstanz.de`
URL: `http://www.math.uni-konstanz.de/~scheider`