Polynomials of almost-normal arguments in C^*-algebras

N. Filonov, I. Kachkovskiy*

Abstract

The functional calculus for normal elements in C^*-algebras is an important tool of Analysis. We suggest an approximate substitute for such calculus for elements a with the small self-commutator norm $\|[a, a^*]\| \leqslant \delta$. We show that many properties of the functional calculus are conserved up to order δ.

1 Introduction

Let a be a *normal* element of a unital C^*-algebra $\mathcal{A}$. The notion of continuous function $f(a)$ of this element is well known. More precisely, there exists a unique C^*-algebra homomorphism

$$C(\sigma(a)) \to \mathcal{A}, \quad f \mapsto f(a)$$

from the algebra of continuous functions on the spectrum $\sigma(a)$ to $\mathcal{A}$ such that the function $f(z) = z$ is mapped to a, $\sigma(f(a)) = f(\sigma(a))$, and

$$\|f(a)\| = \max_{z\in\sigma(a)} |f(z)| \tag{1.1}$$

(see, for example, [4]). This calculus is widely used in solving various problems in Analysis. The aim of the present paper is to introduce an analog of functional calculus for non-normal elements. We restrict the considered class of functions to polynomials (in z and $\overline{z}$). Assume that a is close to a normal element in the sense that the norm of its self-commutator $[a, a^*]$ is small. In what follows, assume that

$$\|a\| \leqslant 1, \quad \|[a, a^*]\| \leqslant \delta. \tag{1.2}$$

We shall show that some properties of the functional calculus hold up to an error of order δ.

Polynomials of a (in the case $aa^* \neq a^*a$) are, in general, not uniquely defined. We fix the following definition. For a polynomial $p(z) = \sum\limits_{k,l} p_{kl} z^k \bar{z}^l$ let

$$p(a) = \sum_{k,l} p_{kl} a^k (a^*)^l.$$

*Steklov Institute, St. Petersburg, and King's College London. The first author was supported by RFBR Grant 11-01-00324-a. The second author was supported by King's Annual Fund Studentship and King's Overseas Research Studentship.

It is clear that the map $p \mapsto p(a)$ is linear and involutive, i. e. $\overline{p}(a) = p(a)^*$, where $\bar{p}(z) = \sum \bar{p}_{lk} z^k \bar{z}^l$. Using the inequality $\|[a, b^m]\| \leqslant m\|b\|^{m-1}\|[a, b]\|$ and (1.2), one can easily show that the map $p \mapsto p(a)$ is "almost multiplicative":

$$\|p(a)q(a) - (pq)(a)\| \leqslant C(p, q)\delta, \tag{1.3}$$

where

$$C(p, q) = \sum_{k,l,s,t} ls\, |p_{kl}|\, |q_{st}|\, .$$

It takes much more effort to obtain an estimate of the norm $\|p(a)\|$. In the case of an analytic polynomial $p(z) = \sum_k p_k z^k$, the von Neumann inequality gives an answer, see, for example, [11, I.9]:

$$\|p(a)\| \leqslant \max_{|z|\leqslant 1} |p(z)| =: p_{\max},$$

where it is only assumed that $\|a\| \leqslant 1$. We prove the following generalization of (1.1), see Theorem 3.2:

$$\|p(a)\| \leqslant p_{\max} + C\delta, \tag{1.4}$$

where the constant C depends on p, but does not depend on a and δ. The second term in the right hand side of (1.4) is essential, see Remark 3.3.

If a is normal and $\mu \notin f(\sigma(a))$, then the usual functional calculus gives that the element $(f(a) - \mu)$ is invertible and

$$\left\|(f(a) - \mu)^{-1}\right\| = \frac{1}{\operatorname{dist}(\mu, f(\sigma(a)))}. \tag{1.5}$$

In Sections 3.3, 3.4, we prove an analogue of this statement (i. e. an estimate of the left hand side of (1.5)) for elements that are close to normal.

The proofs are based on certain representation theorems for positive polynomials. If a real polynomial of x_1, x_2 is non-negative on the unit disk $\{x : x_1^2 + x_2^2 < 1\}$, then, by a result of [9], it admits a representation

$$\sum_j r_j(x)^2 + \left(1 - x_1^2 - x_2^2\right) \sum_j s_j(x)^2 \tag{1.6}$$

with real polynomials r_j and s_j (see Proposition 2.2 below). Such results are usually referred to as *Positivstellensatz*. We also make use of *Positivstellensatz* for polynomials positive on subsets of the real plane bounded by circle arcs. The corresponding results (for the sets bounded by arbitrary algebraic curves) were obtained in [2, 7, 8, 9]. We concentrate on the quantitative versions of these results, where it is possible to find explicit representations similar to (1.6). They were partially obtained in [10, 5]. In our case, their proofs become less complex and completely explicit.

Section 2 is devoted to the necessary results about polynomials, and Section 3 contains applications to polynomial calculus in C^*-algebras.

The authors thank Prof. A. Pushnitski for valuable comments.

2 Representations of non-negative polynomials

2.1 Statements of the results

Let $\lambda_i \in \mathbb{R}^2$, $R_i \in \mathbb{R}$, $i = 1, \dots, m-1$, $R_0 = 1$. Consider the following polynomials $g_0, \dots, g_{m-1} \in \mathbb{R}[x_1, x_2]$:

$$g_0(x) = 1 - |x|^2, \quad g_i(x) = |x - \lambda_i|^2 - R_i^2, \quad i = 1, \dots, m-1, \tag{2.1}$$

where $x = (x_1, x_2)$, $|x|^2 = x_1^2 + x_2^2$. Let

$$S = \{x \in \mathbb{R}^2 : g_i(x) \geqslant 0, \, i = 0, \dots, m-1\}. \tag{2.2}$$

The set S is a unit disk with several "holes" centered in λ_i and of radii R_i.

Theorem 2.1. *Let $g_0, \dots, g_{m-1}$ be the polynomials* (2.1). *Assume that the set S defined by* (2.2) *is not empty. Let the polynomial $p \in \mathbb{R}[x_1, x_2]$ be positive on S. Then there exists an integer N and polynomials*

$$r_i, r_{ij} \in \mathbb{R}[x_1, x_2], \quad i = 0, \dots, m-1, \quad j = 0, \dots, N,$$

such that

$$p = \sum_{j=0}^{N} r_j^2 + \sum_{i=0}^{m-1} \left(\sum_{j=0}^{N} r_{ij}^2 \right) g_i. \tag{2.3}$$

Theorem 2.1 is proved in Section 2.3. The first result of this type was proved in [2] (for the case $m = 1$ with S being a disk). The proof was not constructive and involved Zorn's Lemma. In [7], the result was generalized to a wider class of sets, including the ones of the type (2.2). In [10] and [5], an alternative proof is presented with its major part being constructive and based on the results of [6]. For the purposes of applications to functional calculus, a special form (2.1) of g_i is interesting. In this case the proof simplifies and becomes completely explicit. We follow the construction of [5] and then apply the results of [6] directly.

If we replace positivity of p with non-negativity, then for $m = 1$ the result still holds. The corresponding theorem was proved in [9]:

Proposition 2.2. [9] *Let $p \in \mathbb{R}[x_1, x_2]$ be non-negative on the unit disk $\{x \in \mathbb{R}^2 : |x| \leqslant 1\}$. Then for some N it admits a representation*

$$p = \sum_{j=0}^{N} r_j^2 + \left(\sum_{j=0}^{N} s_j^2 \right) \left(1 - |x|^2\right),$$

where $r_j, s_j \in \mathbb{R}[x_1, x_2]$, $j = 0, \dots, N$.

The proof in [9] utilizes hard algebraic technique and is not constructive. In [8], an analog of Theorem 2.1 for non-negative polynomials is established for $m > 1$ with some additional assumptions on the zeros of p. The proof is also non-constructive. The next statement shows that, in general, one cannot replace the assumption of positivity in Theorem 2.1 with non-negativity.

Theorem 2.3. *Let g_i be defined by* (2.1), *and assume that $\lambda_i \neq \lambda_j$ for some i and j. Then the polynomial $g_i g_j$ can not be represented in the form* (2.3).

This result is probably well known by the specialists, although we could not find it in the literature. For the convenience of the reader, we give the proof in Section 2.4.

2.2 Lemmas

We need the following particular case of the Lojasiewicz inequality (see, e. g., [1]). Recall that the angle between intersecting circles is the minimal angle between their tangents in the intersection points.

Lemma 2.4. *Let $g_0, \ldots, g_{m-1}$ be the polynomials* (2.1). *Assume that $S \neq \varnothing$ and none of the disks $\{x : |x - \lambda_i| < R_i\}$, $i > 0$, is contained in the union of the others. Then for any $x \in [-1, 1]^2 \setminus S$ the following estimate holds:*

$$\operatorname{dist}(x, S) \leqslant -c_0 \min\{g_0(x), \ldots, g_{m-1}(x)\}.$$

If the circles

$$\{x \in \mathbb{R}^2 : |x| = 1\}, \quad \{x \in \mathbb{R}^2 : |x - \lambda_i| = R_i\},\ i = 1, \ldots, m-1, \tag{2.4}$$

are pairwise disjoint or tangent, then $c_0 = R_{\min}^{-1}$, where $R_{\min} = \min\limits_{i=0,\ldots,m-1} R_i$. Otherwise, c_0 can be chosen as

$$c_0 = \frac{\sqrt{2}+1}{R_{\min}^2 \sin(\varphi_{\min}/2)},$$

where $\varphi_{\min}$ is the minimal angle between the pairs of intersecting non-tangent circles (2.4).

Remark 2.5. In our case the sets $\{x : g_i(x) = 0\}$ are circles, which is the reason why an explicit constant in Lojasiewicz inequality can be written down.

The proof of Lemma 2.4 is elementary, we give it in Section 2.5. For the polynomials

$$q(x) = \sum_{|\alpha| \leqslant d} q_\alpha x^\alpha \in \mathbb{R}[x_1, \ldots, x_n],$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multiindex, consider a norm

$$\|q\| = \max_\alpha |q_\alpha| \frac{\alpha_1! \ldots \alpha_n!}{(\alpha_1 + \ldots + \alpha_n)!}. \tag{2.5}$$

The following proposition is also elementary and is proved in [5]:

Proposition 2.6. *Let $x, y \in [-1, 1]^n$, $q \in \mathbb{R}[x_1, \ldots, x_n]$, $\deg q = d$. Then*

$$|q(x) - q(y)| \leqslant d^2 n^{d-1/2} \|q\| |x - y|.$$

The next fact is proved in [6]:

Proposition 2.7. *Let $f \in \mathbb{R}[y_1, \ldots, y_n]$ be a homogeneous polynomial of degree d. Assume that f is strictly positive on the simplex*

$$\Delta_n = \{y \in \mathbb{R}^n \colon y_i \geqslant 0,\ \sum_i y_i = 1\}. \tag{2.6}$$

Let $f_* = \min\limits_{y \in \Delta_n} f(y) > 0$. Then, for

$$N > \frac{d(d-1)\|f\|}{2 f_*} - d,$$

all the coefficients of the polynomial $(y_1 + \ldots + y_n)^N f(y_1, \ldots, y_n)$ are positive.

2.3 Proof of Theorem 2.1

Without loss of generality, we may assume that $0 \leqslant g_i(x) \leqslant 1$ for all $x \in S$. If not, we can normalize g_i multiplying them by positive constants. This procedure does not affect the statements and the fact that we control the bounds.

Theorem 2.8. *Under the conditions of Theorem* 2.1*, let* $p^* = \min_{x \in S} p(x) > 0$*. Then*

$$p(x) - c_0 d^2 2^{d-1/2} \|p\| \sum_{i=0}^{m-1} (1 - g_i(x))^{2k} g_i(x) \geqslant \frac{p^*}{2}, \quad \forall x \in [-1,1]^2, \tag{2.7}$$

where an integer k *is chosen in such a way that* $(2k+1)p^* \geqslant m c_0 d^2 2^{d+1/2} \|p\|$*, and* c_0 *is the constant from Lemma* 2.4*.*

Proof. Let $x \in S$. Then $p(x) \geqslant p^*$. The elementary inequality

$$(1-t)^{2k} t < \frac{1}{2k+1}, \quad 0 \leqslant t \leqslant 1, \quad k \geqslant 0, \tag{2.8}$$

and our choice of k give that the absolute value of the second term in the left hand side of (2.7) does not exceed $\frac{p^*}{2}$.

Assume now that $x \in [-1,1]^2 \setminus S$. Let $y \in S$, $\operatorname{dist}(x,y) = \operatorname{dist}(x,S)$. Then Proposition (2.6) and Lemma 2.4 give

$$p(x) \geqslant p(y) - |p(x) - p(y)| \geqslant p^* - d^2 2^{d-1/2} \|p\| \operatorname{dist}(x,S) \geqslant p^* + c_0 d^2 2^{d-1/2} \|p\| g_{\min}(x), \tag{2.9}$$

where $g_{\min}(x)$ is the (negative) minimum of the values of $g_i(x)$. Note that $(1 - g_{\min}(x))^{2k} > 1$. From (2.9), we get

$$p(x) - c_0 d^2 2^{d-1/2} \|p\| (1 - g_{\min}(x))^{2k} g_{\min}(x) \geqslant p(x) - c_0 d^2 2^{d-1/2} \|p\| g_{\min}(x) \geqslant p^*.$$

On the other hand, (2.8) and the choice of k imply that the terms of (2.7) with $g_i(x) > 0$ contribute to the sum with no more than

$$\frac{(m-1) c_0 d^2 2^{d-1/2} \|p\|}{2k+1} \leqslant \frac{p^*}{2}.$$

Finally, the remaining terms in (2.7) with $g_i(x) < 0$ may only increase the left hand side. ∎

Theorem 2.9. *Let* $p \in \mathbb{R}[x_1, x_2]$*,* $p_* = \min_{x \in [-1;1]^2} p(x) > 0$*. Then, for some* $M \in \mathbb{N}$*,*

$$p = \sum_{|\alpha| \leqslant M} b_\alpha \gamma_1^{\alpha_1} \gamma_2^{\alpha_2} \gamma_3^{\alpha_3} \gamma_4^{\alpha_4}, \tag{2.10}$$

where $b_\alpha \geqslant 0$*,*

$$\gamma_1(x) = \frac{1}{4}(1 + x_1), \quad \gamma_2(x) = \frac{1}{4}(1 - x_1), \quad \gamma_3(x) = \frac{1}{4}(1 + x_2), \quad \gamma_4(x) = \frac{1}{4}(1 - x_2). \tag{2.11}$$

This theorem was obtained in [6] for arbitrary convex polyhedra. We reproduce its proof for the square $[-1,1]^2$, because the explicit form of the results is considerably simpler.

Proof. Consider the following $\mathbb{R}$-algebra homomorphism

$$\varphi\colon \mathbb{R}[y_1,y_2,y_3,y_4] \to \mathbb{R}[x_1,x_2], \quad y_i \mapsto \gamma_i(x).$$

In order to prove the theorem, it suffices to find a polynomial $\tilde{p} \in \mathbb{R}[y_1,y_2,y_3,y_4]$ with positive coefficients such that $\varphi(\tilde{p}) = p$. Let $p = \sum\limits_{i+j\leqslant d} p_{ij}x_1^i x_2^j$. Consider the following homogeneous polynomial

$$\tilde{p}_1(y) = \sum_{i+j\leqslant d} 2^{i+j} p_{ij}(y_1-y_2)^i(y_3-y_4)^j(y_1+y_2+y_3+y_4)^{d-i-j}.$$

Note that $\varphi(\tilde{p}_1) = p$, because

$$\varphi(y_1+y_2+y_3+y_4) = 1, \quad 2\varphi(y_1-y_2) = x_1, \quad 2\varphi(y_3-y_4) = x_2.$$

Let

$$V = \{y \in \Delta_4 \colon 2y_1+2y_2 = 2y_3+2y_4 = 1\},$$

where Δ_4 is the simplex (2.6). If $y \in V$, then $\tilde{p}_1(y) = p(4y_1-1, 4y_3-1) \geqslant p_*$, as $(4y_1-1, 4y_3-1) \in [-1,1]^2$. For an arbitrary y let $y_0 \in V$, $\operatorname{dist}(y,y_0) = \operatorname{dist}(y,V)$. Then, from Proposition 2.6,

$$\tilde{p}_1(y) \geqslant \tilde{p}_1(y_0) - |\tilde{p}_1(y) - \tilde{p}_1(y_0)| \geqslant p_* - d^2 2^{2d-1}\|\tilde{p}_1\| \operatorname{dist}(y,V). \tag{2.12}$$

Let

$$r(y) = 2(y_1+y_2-y_3-y_4)^2.$$

It is easy to see that $\varphi(r) = 0$ and

$$r(y) = (2y_1+2y_2-1)^2 + (2y_3+2y_4-1)^2 \quad \text{if} \quad y \in \Delta_4.$$

If we rewrite the last expression in the coordinates $\frac{y_1+y_2}{\sqrt{2}}$, $\frac{y_1-y_2}{\sqrt{2}}$, $\frac{y_3+y_4}{\sqrt{2}}$, $\frac{y_3-y_4}{\sqrt{2}}$ (having made two rotations over $\pi/4$), we get

$$r(y) \geqslant 8 \operatorname{dist}(y,V)^2, \quad y \in \Delta_4. \tag{2.13}$$

Let

$$\tilde{p}_2(y) = \tilde{p}_1(y) + \frac{2^{4d-6}d^4\|\tilde{p}_1\|^2}{p_*}(y_1+y_2+y_3+y_4)^{d-2} r(y).$$

We still have $\varphi(\tilde{p}_2) = p$. Let us apply the inequalities (2.12) and (2.13):

$$\tilde{p}_2(y) \geqslant p_* - d^2 2^{2d-1}\|\tilde{p}_1\| \operatorname{dist}(y,V) + \frac{2^{4d-3}d^4\|\tilde{p}_1\|^2}{p_*}\operatorname{dist}(y,V)^2$$
$$= \frac{2^{4d-3}d^4\|\tilde{p}_1\|^2}{p_*}\left(\operatorname{dist}(y,V) - \frac{p_*}{d^2 2^{2d-1}\|\tilde{p}_1\|}\right)^2 + \frac{p_*}{2} \geqslant \frac{p_*}{2}, \quad y \in \Delta_4.$$

Finally, Proposition 2.7 with $N > \frac{d(d-1)\|\tilde{p}_2\|}{p_*} - d$ gives us that all the coefficients of

$$\tilde{p}(y) = (y_1+y_2+y_3+y_4)^N \tilde{p}_2(y)$$

are positive. Applying the homomorphism φ to $\tilde{p}$, we get the desired representation of p. ∎

Proof of Theorem 2.1. Let us apply Theorem 2.8 to p. It is enough to find a representation of the first term in the left hand side of (2.7), because the second term, if moved to the right hand side, already has the desired form. By Theorem 2.9, the left hand side of (2.7) can be represented in the form (2.10). Note that γ_i can be rewritten as

$$\frac{1}{4}(1 \pm x_{1,2}) = \frac{1}{8}\left((1 \pm x_{1,2})^2 + g_0(x) + x_{2,1}^2\right). \tag{2.14}$$

Substituting the last equality into (2.10), we get the desired representation for (2.7) and therefore for p. ∎

Remark 2.10. The proof of Theorem 2.1 is constructive. We have a polynomial p such that $p(x) \geqslant p^* > 0$, $x \in S$. Then

$$p(x) = \hat{p}(x) + c_0 d^2 2^{d-1/2} \|p\| \sum_{i=0}^{m-1} (1 - g_i(x))^{2k} g_i(x), \tag{2.15}$$

where k is chosen in such a way that $(2k+1)p^* \geqslant mc_0 d^2 2^{d+1/2}\|p\|$. The second term in the right hand side of (2.15) is an explicit expression of the form (2.3), and the coefficients of $\hat{p}$ can be found from (2.15). From Theorem 2.8, we know that $\hat{p}(x) \geqslant p^*/2$, $x \in [-1;1]^2$. It now suffices to represent

$$\hat{p}(x) = \sum_{k+l \leqslant \hat{d}} \hat{p}_{kl}\, x_1^k x_2^l$$

in the form (2.3). Consider the following polynomials in $\mathbb{R}[y_1, y_2, y_3, y_4]$:

$$\tilde{p}_1(y) = \sum_{i+j \leqslant \hat{d}} 2^{i+j} \hat{p}_{ij} (y_1 - y_2)^i (y_3 - y_4)^j (y_1 + y_2 + y_3 + y_4)^{\hat{d}-i-j},$$

$$\tilde{p}_2(y) = \tilde{p}_1(y) + \frac{2^{4\hat{d}-5}\hat{d}^4 \|\tilde{p}_1\|^2}{p_*} (y_1 + y_2 + y_3 + y_4)^{\hat{d}-2} (y_1 + y_2 - y_3 - y_4)^2,$$

and

$$\tilde{p}(y) = (y_1 + y_2 + y_3 + y_4)^N \tilde{p}_2(y), \quad \text{where} \quad N > \frac{\hat{d}(\hat{d}-1)\|\tilde{p}_2\|}{p_*} - \hat{d}.$$

From the proof of Theorem 2.9, if we replace in the last expression y_i with $\gamma_i(x)$ defined by (2.11), we will get $\hat{p}(x)$. The coefficients of $\tilde{p}$ are positive. Therefore, if we substitute y_i with γ_i and then apply (2.14), we will get an expression of the form (2.3) for $\hat{p}(x)$. Combining it with (2.15), we get the desired expression for p.

2.4 Proof of Theorem 2.3

Let g_i be defined by (2.1). Let us denote

$$S_i = \{x \in \mathbb{R}^2 : g_i(x) = 0\}, \quad S_i(\mathbb{C}) = \{x \in \mathbb{C}^2 : g_i(x) = 0\}. \tag{2.16}$$

Lemma 2.11. *Let $q \in \mathbb{R}[x_1, x_2]$, $q(x) = 0$ on some arc of S_i. Then $g_i \mid q$.*

Proof. Consider q as an analytic function on $S_i(\mathbb{C})$. The set $S_i(\mathbb{C})$ is connected, so $q \equiv 0$ on the whole $S_i(\mathbb{C})$. Hilbert's Nullstellensatz (see, for example, [12, Section 16.3]) gives that $g_i \mid q^k$ for some integer k (in $\mathbb{C}[x_1, x_2]$ and, consequently, in $\mathbb{R}[x_1, x_2]$). The polynomial g_i is irreducible, so $g_i \mid q$. ∎

Lemma 2.12. *Let $\lambda_i \neq \lambda_j$. Then $S_i(\mathbb{C}) \cap S_j(\mathbb{C}) \neq \varnothing$.*

Proof. Let the circles S_i and S_j be the solution sets of the equations

$$(x_1 - a_1)^2 + (x_2 - a_2)^2 = R_1^2, \quad (x_1 - b_1)^2 + (x_2 - b_2)^2 = R_2^2.$$

Subtracting one from another, we get a system of a linear and a quadratic equation. The linear one is solvable (because $\lambda_i \neq \lambda_j$), and after substituting the solution, we get a non-degenerate quadratic equation in one complex variable, which also has a solution. ∎

Proof of Theorem 2.3. Assume that $p = g_i g_j$ satisfies (2.3). On the set $S_i \cap \partial S$, the left hand side of (2.3) equals zero. All the terms r_k^2 and $r_{kl}^2 g_k$ in the right hand side of (2.3) are non-negative on $S_i \cap \partial S$, and therefore are equal to zero on this set. By Lemma 2.11, they all are multiples of g_i. Similarly, all the terms in the right hand side are multiples of g_j. Then $g_i \mid r_k$, $g_j \mid r_k$, and $g_i^2 g_j^2 \mid r_k^2$.

Next, for $k \neq i$ the polynomials g_k and g_i are coprime. So, $g_i^2 \mid r_{kl}^2$ for $k \neq i$, and $g_j^2 \mid r_{kl}^2$ for $k \neq j$. Then any term in the right hand side of (2.3) is a multiple of either $g_i^2 g_j$ or $g_i g_j^2$. If we divide (2.3) by $g_i g_j$, we get that the left hand side is identically 1, and the right hand side equals zero at least on $S_i(\mathbb{C}) \cap S_j(\mathbb{C})$. This contradiction proves the theorem. ∎

2.5 Proof of Lemma 2.4

Lemma 2.13. *Let S_1, S_2 be a pair of intersecting circles with centers at λ_1, λ_2 and of radii R_1, R_2. Let y, y' be the intersection points of S_1 and S_2, and let $\varphi = \angle(S_1, S_2)$ be the angle between the circles S_1 and S_2. Assume that x lies inside of the first circle, so that $|x - \lambda_1| < R_1$, and suppose also that the points x and λ_2 are in the same half-plane with respect to the line $\lambda_1 y$. Finally, let $|x - y| \leqslant \min(R_1, R_2) \sin \varphi/2$. Then*

$$|x - y| \leqslant \frac{2}{\sin \varphi/2} \max_{i=1,2} \left(R_i - |x - \lambda_i|\right). \tag{2.17}$$

Proof. It is easy to see that

$$\angle y\lambda_1\lambda_2 + \angle y\lambda_2\lambda_1 = \varphi \text{ or } \pi - \varphi.$$

So, $\max\left(\angle y\lambda_1\lambda_2, \angle y\lambda_2\lambda_1\right) \geqslant \varphi/2$, which gives

$$\frac{|yy'|}{2} = R_1 \sin \angle y\lambda_1\lambda_2 = R_2 \sin \angle y\lambda_2\lambda_1 \geqslant \min(R_1, R_2) \sin \varphi/2 \geqslant |x - y|. \tag{2.18}$$

Denote the intersection points of the line $\lambda_1\lambda_2$ with the circles S_1 and S_2 by z' and z respectively (the distance between z and z' is chosen to be smallest possible). From (2.18) it follows that x lies inside the sector $\lambda_1 y z'$.

Let us show that at least one of the following conditions holds:

1) $\angle(xy, S_1) \geqslant \varphi/2$;

2) $|x - \lambda_2| < R_2$ and $\angle(xy, S_2) \geqslant \varphi/2$.

Indeed, $\angle zyz' = \varphi/2$ or $(\pi - \varphi)/2$. If x does not belong to the intersection of the disks, then $\angle(xy, S_1) \geqslant \angle zyz' \geqslant \varphi/2$, and the first condition holds. If x belongs to the intersection, then $\max\left(\angle(xy, S_1), \angle(xy, S_2)\right) \geqslant \varphi/2$, and either 1) or 2) is true.

The cases 1) and 2) can be treated in a similar way. Let us restrict ourselves to the first one. Denote $\psi = \angle(xy, S_1)$. By the cosine theorem for the triangle $xy\lambda_1$, we have

$$|x-\lambda_1| = \sqrt{R_1^2 + |x-y|^2 - 2R_1|x-y|\sin\psi} \leqslant \sqrt{R_1^2 - R_1|x-y|\sin\psi},$$

because, by assumption, $|x-y| \leqslant R_1 \sin\varphi/2 \leqslant R_1 \sin\psi$. Consequently,

$$R_1 - |x-\lambda_1| \geqslant R_1\left(1 - \sqrt{1 - \frac{|x-y|\sin\psi}{R_1}}\right) \geqslant \frac{|x-y|\sin\psi}{2} \geqslant \frac{|x-y|\sin\varphi/2}{2},$$

and this implies (2.17). ∎

Proof of Lemma 2.4. Let $x \notin S$. Then there exists i such that $g_i(x) < 0$. Let y be the closest to x point of S, $\operatorname{dist}(x, S) = |x-y|$. It is clear that $y \in S_i$ (see (2.16)). If y belongs to S_i only for a single i, or if it is a tangent point of S_i and S_j (but not an intersection point), then

$$\operatorname{dist}(x,S) = |x-y| = R_i - |x-\lambda_i| = \frac{R_i^2 - |x-\lambda_i|^2}{R_i + |x-\lambda_i|} \leqslant \frac{-g_i(x)}{R_{\min}}, \quad i \neq 0, \tag{2.19}$$

$$\operatorname{dist}(x,S) = |x-y| = \frac{|x|^2-1}{|x|+1} \leqslant \frac{-g_0(x)}{R_0} \text{ for } i = 0, \tag{2.20}$$

and there is nothing more to prove.

Let $\varepsilon = R_{\min}\sin(\varphi_{\min}/2)$, and consider the case $|x-y| \geqslant \varepsilon$. Then $-g_i(x) \geqslant R_{\min}\varepsilon$ (see (2.19), (2.20)). However, $\operatorname{dist}(x,S) \leqslant \sqrt{2}+1$ for all $x \in [-1,1]^2$. Then,

$$\operatorname{dist}(x,S) \leqslant -\frac{\sqrt{2}+1}{\varepsilon R_{\min}}\, g_i(x),$$

which also completes the proof in this case.

Suppose now that $|x-y| < \varepsilon$ and y is an intersection point of multiple circles. First assume that none of these circles is S_0. Then there exists S_j such that it contains y and its center λ_j lies in the same half-plane as x with respect to $\lambda_i y$ (otherwise, the point y would not be the closest to x point of S). By Lemma 2.13,

$$|x-y| \leqslant \frac{-2\min g_i(x)}{R_{\min}\sin(\varphi_{\min}/2)} \leqslant \frac{-(\sqrt{2}+1)\min g_i(x)}{R_{\min}^2\sin(\varphi_{\min}/2)}.$$

The case when one of the circles is S_0 can be treated in a similar way, there are several options. There may exists a pair of circles S_i, S_j, $i,j > 0$, satisfying the conditions of Lemma 2.13. Or, alternatively, one of the circles may satisfy Condition 1) from the proof of Lemma 2.13. These two cases were already considered. The third alternative is that the point x lies outside of S_0 and the angle between xy and S_0 is not less that $\varphi/2$. Here a similar cosine-theorem computation can be made. We omit further details. ∎

3 Polynomials of almost-normal elements

3.1 Positive elements of C^*-algebras

Let $\mathcal{A}$ be a unital C^*-algebra with the unit $\mathbf{1}$. Recall that a Hermitian element $b \in \mathcal{A}$ is called *positive* ($b \geqslant 0$) if one of the following two equivalent conditions holds (see, for example, [4, §1.6]):

1. $\sigma(b) \subset [0, +\infty)$.
2. $b = h^* h$ for some $h \in \mathcal{A}$.

The set of all positive elements in $\mathcal{A}$ is a cone: if $a, b \geqslant 0$, then $\alpha a + \beta b \geqslant 0$ for all real $\alpha, \beta \geqslant 0$. There exists a partial ordering on the set of Hermitian elements of $\mathcal{A}$: $a \leqslant b$ iff $b - a \geqslant 0$. The ordering is consistent with addition. For a Hermitian b, it is true that $-\|b\|\mathbf{1} \leqslant b \leqslant \|b\|\mathbf{1}$ and, moreover, if $0 \leqslant b \leqslant \beta\mathbf{1}$, $\beta \in \mathbb{R}$, then $\|b\| \leqslant \beta$. The following fact is also well known.

Proposition 3.1. *Let $h \in \mathcal{A}$, $\rho > 0$. Then $h^* h \geqslant \rho^2 \mathbf{1}$ if and only if the element h is invertible and $\|h^{-1}\| \leqslant \rho^{-1}$.*

3.2 Estimate of the norm $\|p(a)\|$

Consider a polynomial $p \in \mathbb{C}[x_1, x_2]$. It can be uniquely represented in the form

$$p(z) = \sum_{k,l} p_{kl} z^k \bar{z}^l. \tag{3.1}$$

In this section, we associate z with $x_1 + i x_2$, $\bar{z}$ with $x_1 - i x_2$, and $p(z)$ with $p(x_1, x_2)$. For $a \in \mathcal{A}$, denote

$$p(a) = \sum_{k,l} p_{kl} a^k (a^*)^l. \tag{3.2}$$

Theorem 3.2. *Let $p \in \mathbb{C}[x_1, x_2]$. There exists a constant $C(p)$ such that the estimate*

$$\|p(a)\| \leqslant p_{\max} + C(p)\delta \tag{3.3}$$

holds for all a satisfying (1.2)*. Here $p(a)$ is defined by* (3.2)*, and $p_{\max} = \max\limits_{|z| \leqslant 1} |p(z)|$.*

Proof. Consider the polynomial

$$q(z) = p_{\max}^2 - |p(z)|^2. \tag{3.4}$$

It is non-negative on the disk $\{z \in \mathbb{C} : |z| \leqslant 1\}$ and has real coefficients as a polynomial in x_1, x_2. By Proposition 2.2, it admits a representation

$$q(z) = \sum_{j=0}^{N} r_j(z)^2 + \left(\sum_{j=0}^{N} s_j(z)^2 \right) \left(1 - |z|^2\right).$$

The polynomials q, r_j, s_j are real, their coefficients at $z^k \bar{z}^l$ and $z^l \bar{z}^k$ are mutually conjugate. Therefore, $q(a)$, $r_j(a)$, $s_j(a)$ are Hermitian elements of $\mathcal{A}$. Using (1.3), we get

$$\left\| q(a) - \sum_{j=0}^{N} r_j(a)^2 - \sum_{j=0}^{N} s_j(a) \left(\mathbf{1} - a a^*\right) s_j(a) \right\| \leqslant C_1(p)\delta. \tag{3.5}$$

The element $\mathbf{1} - aa^*$ is positive, and therefore can be represented as h^*h for some $h \in \mathcal{A}$. So, in the last equation, all the terms in sums are positive. In view of the results of Section 3.1, this gives $q(a) \geqslant -C_1(p)\delta\mathbf{1}$, from which, using (3.4) and (1.3), we get

$$p(a)^*p(a) \leqslant \left(p_{\max}^2 + C_2(p)\delta\right)\mathbf{1}$$

and, therefore,

$$\|p(a)\| \leqslant p_{\max} + \frac{C_2(p)\delta}{2p_{\max}}. \blacksquare$$

Remark 3.3. The fact that $C = C(p)$ in (3.3) may depend on p is essential. As an example, consider $\mathcal{A} = M_2(\mathbb{C})$,

$$a = \begin{pmatrix} 0 & \sqrt{\delta} \\ 0 & 0 \end{pmatrix}, \quad 0 < \delta < 1.$$

It is clear that a satisfies (1.2). Let $\varepsilon < 1$. There exists a continuous function f such that $f(z) = -1/z$ for $|z| \geqslant \varepsilon$ and $|f(z)| \leqslant 1/\varepsilon$, $|z| \leqslant 1$. There also exist a polynomial q of the type (3.1) such that $|q(z) - f(z)| \leqslant \varepsilon$, $|z| \leqslant 1$. Now, let

$$p(z) = \frac{1}{\varepsilon}\left(z + z^2 q(z)\right).$$

Then $p_{\max} \leqslant 2 + \varepsilon^2$, but $p(a) = a/\varepsilon$ and $\|p(a)\| = \sqrt{\delta}/\varepsilon$. Taking ε small, we get that (3.3) can not hold for any fixed C.

Theorem 3.4. *Let $R_j > 0$, $j = 1, \dots, m-1$. Consider the set*

$$S = \{z \in \mathbb{C}\colon |z| \leqslant 1,\ |z - \lambda_j| \geqslant R_j,\ j = 1, \dots, m-1\} \tag{3.6}$$

(we assume $S \neq \varnothing$). Let $p \in \mathbb{C}[x_1, x_2]$. For each $\varepsilon > 0$ there exists a constant $C(p, \varepsilon)$ such that the estimate

$$\|p(a)\| \leqslant \max_{z \in S} |p(z)| + \varepsilon + C(p, \varepsilon)\delta$$

holds for all $a \in \mathcal{A}$ satisfying

$$\|a\| \leqslant 1, \quad \|[a, a^*]\| \leqslant \delta, \quad \|(a - \lambda_j)^{-1}\| \leqslant R_j^{-1},\ j = 1, \dots, m-1. \tag{3.7}$$

Proof. The proof is similar to the proof of Theorem 3.2, with using Theorem 2.1 instead of Proposition 2.2. Consider

$$q(z) = p_{\max}^2 + \varepsilon p_{\max} - |p(z)|^2,$$

where now $p_{\max} = \max_{x \in S} |p(x)|$. Note that (3.7) implies $g_i(a) \geqslant 0$. Similarly to the proof of Theorem 3.2, we get

$$q(a) \geqslant -C_1\delta\mathbf{1},$$
$$p(a)p(a)^* \leqslant \left(p_{\max}^2 + \varepsilon p_{\max} + C_2\delta\right)\mathbf{1},$$

and

$$\|p(a)\| \leqslant p_{\max}\sqrt{1 + \frac{\varepsilon}{p_{\max}} + \frac{C_2(p, \varepsilon)\delta}{p_{\max}^2}} \leqslant p_{\max} + \varepsilon + \frac{C_2(p, \varepsilon)\delta}{p_{\max}^2}. \blacksquare$$

Corollary 3.5. *Under the conditions of Theorem* 3.4*, there exist a constant* $C(p,\varepsilon)$ *such that*

$$\|\operatorname{Im} p(a)\| \leqslant \max_{z\in S} |\operatorname{Im} p(z)| + \varepsilon + C(p,\varepsilon)\delta.$$

Proof. It suffices to apply Theorem 3.4 to the polynomial $q(z) = \frac{p(z)-\overline{p(z)}}{2i}$. ∎

Remark 3.6. In other words, if the values of p on S are "almost real", then the element $p(a)$ itself is "almost self-adjoint".

Corollary 3.7. *Under the assumptions of Theorem* 3.4*, there exists a constant* $C(p,\varepsilon)$ *such that*

$$\|p(a)p(a)^* - \mathbf{1}\| \leqslant \max_{z\in S} \left||p(z)|^2 - 1\right| + \varepsilon + C(p,\varepsilon)\delta, \tag{3.8}$$

$$\|p(a)^*p(a) - \mathbf{1}\| \leqslant \max_{z\in S} \left||p(z)|^2 - 1\right| + \varepsilon + C(p,\varepsilon)\delta. \tag{3.9}$$

Proof. It is sufficient to apply Theorem 3.4 to the polynomial $q(z) = |p(z)|^2 - 1$ and use (1.3).∎

Remark 3.8. Denote the right hand side of (3.8), (3.9) by γ. If $\gamma < 1$, then

$$(1-\gamma)\mathbf{1} \leqslant p(a)^*p(a) \leqslant (1+\gamma)\mathbf{1}, \quad (1-\gamma)\mathbf{1} \leqslant p(a)p(a)^* \leqslant (1+\gamma)\mathbf{1},$$

which implies that $p(a)$ and $p(a)^*p(a)$ are invertible. The element $u = p(a)\left(p(a)^*p(a)\right)^{-1/2}$ is unitary (because it is also invertible and $uu^* = 1$) and close to u:

$$\|p(a) - u\| \leqslant \sqrt{1+\gamma}\left(\frac{1}{\sqrt{1-\gamma}} - 1\right) \to 0 \quad \text{as} \quad \gamma \to 0.$$

This means that if the absolute values of p on S are close to 1, then the element $p(a)$ is close to a unitary one.

3.3 Resolvent estimates for $p(a)$

Theorem 3.9. *Let* $R_j > 0$*,* $j = 1, \dots, m-1$*, let* S *be defined by* (3.6)*. Let also* $p \in \mathbb{C}[x_1, x_2]$*. Then for each* $\varepsilon > 0$ *and* $\varkappa > 0$ *there exist constants* $C(p,\varkappa,\varepsilon)$*,* $\delta_0(p,\varkappa,\varepsilon)$ *such that for all* $\delta < \delta_0(p,\varkappa,\varepsilon)$ *and for all* $\mu \in \mathbb{C}$ *satisfying* $\operatorname{dist}(\mu, p(S)) \geqslant \varkappa$ *the estimate*

$$\|(p(a) - \mu\mathbf{1})^{-1}\| \leqslant \varkappa^{-1} + \varepsilon + C(p,\varkappa,\varepsilon)\delta$$

holds for all $a \in \mathcal{A}$ *satisfying* (3.7)*.*

Proof. Let g_i, $i = 0, \dots, m-1$, be the polynomials from Theorem 2.1:

$$g_0(z) = 1 - |z|^2, \quad g_i(z) = |z-\lambda_i|^2 - R_i^2, \quad i = 1, \dots, m-1.$$

Fix $\gamma > 0$ and consider

$$q(z) = |p(z) - \mu|^2 - \varkappa^2 + \gamma.$$

The coefficients of q considered as a polynomial in x_1 and x_2 are real. Moreover, $q(z) \geqslant \gamma > 0$ for $z \in S$. By Theorem 2.1, there exists a representation

$$q(z) = \sum_{j=1}^{N} r_j(z)^2 + \sum_{i=0}^{m-1}\left(\sum_{j=0}^{N} r_{ij}(z)^2\right) g_i(z).$$

The polynomials q, r_j, r_{ij}, g_i are real, so the elements $q(a)$, $r_j(a)$, $r_{ij}(a)$, $g_i(a)$ are Hermitian. Similarly to the proof of Theorem 3.2, we obtain

$$q(a) \geqslant \sum_{i=0}^{m-1} \left(\sum_{j=0}^{N} r_{ij}(a) g_i(a) r_{ij}(a) \right) - C'\delta \mathbf{1}.$$

By Proposition 3.1, the inequality $\|(a - \lambda_i \mathbf{1})^{-1}\| \leqslant R_i^{-1}$ yields $(a - \lambda_i \mathbf{1})(a - \lambda_i \mathbf{1})^* \geqslant R_i^2$. Therefore, $g_i(a) \geqslant 0$ and $g_i(a) = h_i^* h_i$ for some $h_i \in \mathcal{A}$. Then

$$q(a) \geqslant \sum_{i=0}^{m-1} \sum_{j=0}^{N} (h_i r_{ij}(a))^* \, (h_i r_{ij}(a)) - C'\delta \mathbf{1} \geqslant -C'\delta \mathbf{1}.$$

Using the definition of q, we get

$$(p(a) - \mu \mathbf{1})^* (p(a) - \mu \mathbf{1}) \geqslant \left(\varkappa^2 - \gamma - C''\delta \right) \mathbf{1}. \tag{3.10}$$

The constant C'', in general, depends on $p, \varkappa, \gamma$, and μ. Let us show that it can be chosen to be independent of μ. For $|\mu| \geqslant \|p(a)\| + \varkappa$ the statement of the theorem becomes trivial, as

$$\left\|(p(a) - \mu \mathbf{1})^{-1}\right\| \leqslant \frac{1}{|\mu| - \|p(a)\|} \leqslant \varkappa^{-1}.$$

So, we can restrict ourselves to a compact set

$$M = \{\mu \in \mathbb{C} \colon |\mu| \leqslant \|p(a)\| + \varkappa, \ \mathrm{dist}(\mu, p(S)) \geqslant \varkappa\}.$$

The condition $q(z) \geqslant \gamma$ holds there. By Remark 2.10, for the coefficients r_j and r_{ij} we have explicit formulas, which depend on μ continuously. Therefore, the constant C'' may be chosen independent on $\mu \in M$.

Let us choose γ and δ_0 such that $\gamma + C''\delta \leqslant \varkappa^2/2$. Now, (3.10) and Proposition 3.1 give

$$\|(p(a) - \mu \mathbf{1})^{-1}\| \leqslant \left(\varkappa^2 - \gamma - C''\delta \right)^{-1/2} \leqslant \varkappa^{-1} + \frac{\gamma}{\varkappa^2} + \frac{C''\delta}{\varkappa^2}.$$

The choice $\gamma \leqslant \varepsilon \varkappa^2$ completes the proof. ∎

Remark 3.10. In the case $[a, a^*] = 0$, the standard functional calculus gives an implication

$$\sigma(a) \subset S \quad \Rightarrow \quad \|(p(a) - \mu \mathbf{1})^{-1}\| \leqslant \frac{1}{\mathrm{dist}(\mu, p(S))}.$$

Theorem 3.9 is an analog of this statement for non-normal elements. We obtain a weaker estimate of $(p(a) - \mu \mathbf{1})^{-1}$, while assuming a stronger condition

$$\|(a - \lambda_j \mathbf{1})^{-1}\| \leqslant R_j^{-1}, \ j = 1, \ldots, m-1,$$

instead of $\sigma(a) \subset S$.

3.4 Estimates of pseudospectra

Definition 3.11. Let $\mathcal{A}$ be a unital Banach algebra, $a \in \mathcal{A}$. The set

$$\sigma_\varepsilon(a) = \{\lambda \in \mathbb{C}\colon \|(a - \lambda\mathbf{1})^{-1}\| > 1/\varepsilon\} \cup \sigma(a)$$

is called the *ε-pseudospectrum* of the element a.

Its main properties can be found, for example, in [3, Ch. 9]. Note that, under the assumptions of Theorem 3.9, $\sigma_\varepsilon(a) \subset \mathcal{O}_\varepsilon(S)$ for all $\varepsilon > 0$, where $\mathcal{O}_\varepsilon(S)$ is the ε-neighbourhood of S. In the case of normal a the following equality holds:

$$\sigma_\varkappa(p(a)) = \mathcal{O}_\varkappa\left(p(\sigma(a))\right), \quad \varkappa > 0.$$

Theorem 3.9 is an analogue of the last statement.

Corollary 3.12. *Under the assumptions of Theorem* 3.9*, for all $\varepsilon > 0$ and $\varkappa > 0$ there exist $C(p, \varkappa, \varepsilon)$ and $\delta_0(p, \varkappa, \varepsilon)$ such that*

$$\sigma_{\varkappa'}(p(a)) \subset \mathcal{O}_\varkappa(p(S)), \quad (\varkappa')^{-1} = \varkappa^{-1} + \varepsilon + C(p, \chi, \varepsilon)\delta, \quad \delta < \delta_0(p, \varkappa, \varepsilon).$$

Proof. Assume that $\operatorname{dist}(\mu, p(S)) \geqslant \varkappa$. By Theorem 3.9, $\|(p(a) - \mu\mathbf{1})^{-1}\| \leqslant (\varkappa')^{-1}$ and so $\mu \notin \sigma_{\varkappa'}(p(a))$. ∎

References

[1] Bochnak J., Coste M., Roy M.-F., *Real Algebraic Geometry*, Erg. Math. Grenzgeb. (3) 36, Springer, Berlin, 1998.

[2] Cassier G., *Problème des moments sur un compact de $\mathbb{R}^n$ et décomposition de polynômes à plusieurs variables*, J. Funct. Anal. 58 (1984), 254–266.

[3] Davies E. B., *Linear Operators and their Spectra*, Cambridge Studies in Advanced Mathematics, No. 106, 2007.

[4] Dixmier J., *C^*-algebras*, North-Holland Mathematical Library, Vol. 15. North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977.

[5] Nie J., Schweighofer M., *On the complexity of Putinar's Positivstellensatz*, Journal of Complexity, vol. 23, 1 (2007), 135–150.

[6] Powers V., Reznick B., *A new bound for Pòlya's theorem with applications to polynomials positive on polyhedra*, J. Pure Applied Algebra 164 (2001), 221–229.

[7] Putinar M., *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J. 42 (1993), 969-984.

[8] Scheiderer C., *Distinguished representations of non-negative polynomials*, Journal of Algebra, vol. 289, 2 (2005), 558–573.

[9] Scheiderer C., *Sums of squares on real algebraic surfaces*, Manuscripta mathematica, vol. 119, 4 (2006), 395–410.

[10] Schweighofer M., *On the complexity of Schmudgen's Positivstellensatz*, Journal of Complexity, vol. 20, 4 (2004), 529–543.

[11] Sz-Nagy B., Foias C., Bercovici H., Kérchy L., *Harmonic Analysis of Operators on Hilbert Space*, Springer, 2nd ed., 2010.

[12] Van der Waerden B. L., *Algebra*, Vol II, Springer, 2003.