**POSITIVE POLYNOMIALS AND SUMS OF SQUARES:
A BEGINNER'S GUIDE**

Victoria Powers

1 Introduction

These notes are based on a series of lectures given by the author as part of the CIMPA school on Combinatorial and Computational Algebraic Geometry at the University of Ibadan, Nigeria, June 12-23, 2017. The material covered during the lectures has been extended with many more details and more background added.

If polynomial f with real number coefficients can be written as a sum of squares of real polynomials, then f takes only nonnegative values on $\mathbb{R}^n$, and an explicit expression of f as a sum of squares gives an immediate proof of this. This idea, and generalizations of it, underlie a large body of theoretical and computational results concerning positive polynomials and sums of squares. In these notes we will present some of these ideas and the history of the subject, much of which originated in work of David Hilbert in the late 19th century. These notes are written for "beginners", by which we mean someone with no background in sums of squares and real algebraic geometry. We assume that the reader knows linear algebra at the undergraduate level along with basic ideas in abstract algebra including fields and polynomials.

2 Preliminaries

In this section we collect some basic facts about the real numbers and polynomials that will be important in what follows.

2.1 The real numbers

Throughout, we work in $\mathbb{R}$, the field of real numbers; $\mathbb{Z}$ denotes the integers, $\mathbb{Z}_{\geq 0}$ the nonnegative integers, $\mathbb{N}$ the natural numbers (positive integers), $\mathbb{Q}$ the rational numbers, and $\mathbb{C}$ the complex numbers. The most important property of $\mathbb{R}$ is that it has an **order**: The statement $a > b$ makes sense for $a, b \in \mathbb{R}$. In classic algebraic geometry one studies **algebraic sets** over $\mathbb{C}$, solution sets to systems of polynomial equations. Algebraic sets defined over $\mathbb{R}$ are studied in real algebraic geometry, but in addition, since $\mathbb{R}$ has an order, we can define **semialgebraic sets**, solution sets to polynomial *inequalities.*

We list some properties of $\mathbb{R}$ that will be important and useful in these notes.

Basic Properties of $\mathbb{R}$.

1. For any $a \in \mathbb{R}$, if $a^2 = 0$, then $a = 0$.
2. For any $a \in \mathbb{R}$, $a^2 \geq 0$.
3. For any $k \in \mathbb{N}$ and $a_1, \ldots, a_k \in \mathbb{R}$, if $\sum_{i=1}^k a_i^2 = 0$, then $a_1 = a_2 = \cdots = a_k = 0$.

Note that the first property holds in any field and the third property follows from the first two.

Finally, we recall the following property of real numbers:

The Arithmetic-Geometric Mean. Suppose $k \geq 2$ is an integer and $a_1, \dots, a_k \in \mathbb{R}$ are nonnegative. Then

$$\frac{a_1 + \cdots + a_k}{k} \geq \sqrt[k]{a_1 a_2 \dots a_k}.$$

2.2 Polynomials

We will be working mainly with polynomials with coefficients from $\mathbb{R}$. Throughout these notes we fix $n \in \mathbb{N}$ and let $\mathbb{R}[\underline{X}]; = \mathbb{R}[X_1, \dots, X_n]$ be the ring of polynomials in n variables over $\mathbb{R}$. Generally we will use capital letters for variables and x, y, etc. for elements of $\mathbb{R}^n$. For convenience, $\mathbb{R}[T]$ will be used in the $n = 1$ (one variable) case, $\mathbb{R}[X, Y]$ will be used in the $n = 2$ case, and $\mathbb{R}[X, Y, Z]$ in the $n = 3$ case.

We use the following monomial notation: For $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{Z}_{\geq 0})^n$, let X^α denote $X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. The **degree** of X^α is $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$ and the degree of $f \in \mathbb{R}[X]$ is the maximum degree of the monomials in f with nonzero coefficients. Thus if $\deg f = d$, we have

$$f = \sum_{|\alpha| \leq d} a_\alpha X^\alpha,$$

where $a_\alpha \in \mathbb{R}$ and $a_\alpha \neq 0$ for some α with $|\alpha| = d$. The **support** of f, $\operatorname{supp}(f)$, is $\{\alpha \mid a_\alpha \neq 0\}$.

Proposition 1. Suppose $f \in \mathbb{R}[\underline{X}]$ and $f(x) = 0$ for all $x \in \mathbb{R}^n$. Then f is the zero polynomial, i.e., $\operatorname{supp}(f) = \emptyset$.

Proof. Assume f is not the zero polynomial, then we prove by induction on n that there is $x \in \mathbb{R}^n$ such that $f(x) \neq 0$. For $n = 1$, this follows from the well-known fact that a non-zero polynomial in one variable has finitely many roots. Now suppose $n > 1$ and let $d \in \mathbb{N}$ be the maximum power of X_n that appears in f. Then we can write

$$f = g_0 + g_1 X_n + g_2 X_n^2 + \cdots + g_d X_n^d,$$

where $g_0, \dots, g_d \in \mathbb{R}[X_1, \dots, X_{n-1}]$. Since f is not the zero polynomial, there is some i such that g_i is not the zero polynomial. Then, by induction, there exists $x = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$ such that $g_i(x) \neq 0$. It follows that

$$f(x_1, \dots, x_{n-1}, X_n) = g_0(x) + g_1(x) X_n + \cdots + g_d(x) X_n^d,$$

is a polynomial in one variable that is not the zero polynomial. Hence there is $x_n \in \mathbb{R}^n$ such that $f(x_1, \dots, x_n) \neq 0$. □

2.3 Polynomials versus Forms

A **form** is a homogeneous polynomial, i.e., one for which all monomials have the same degree. For example, $X^4 + 2X^3Y + X^2Y^2$ is a form while $X^4 + Y^2 + 1$ is not. Given

$f \in \mathbb{R}[\underline{X}]$ of degree d, then the **degree d homogenization** of f, denoted $\bar{f}$, is a form in $\mathbb{R}[X_1, \ldots, X_n, X_{n+1}]$ of degree d, defined by

$$\bar{f} := X_{n+1}^d \cdot f\left(\frac{X_1}{X_{n+1}}, \ldots, \frac{X_n}{X_{n+1}}\right).$$

If $p \in \mathbb{R}[X_1, \ldots, X_{n+1}]$ is a form of degree d, then the **dehomogenization** of p is the polynomial $p(X_1, \ldots, X_n, 1)$, which is a polynomial in $\mathbb{R}[\underline{X}]$ of degree d. It is easy to see that for all $f \in \mathbb{R}[\underline{X}]$, the dehomogenization of $\bar{f}$ is f.

Given $f \in \mathbb{R}[\underline{X}]$ of degree d, we can collect the monomials of a given degree in f and write

$$f = f_d + f_{d-1} + \cdots + f_0, \tag{1}$$

where f_i is a form in $\mathbb{R}[\underline{X}]$ of degree i. The **leading form** of f is f_d.

Example 1. Let $m = X^4Y^2 + X^2Y^4 - 3X^2Y^2 + 1 \in \mathbb{R}[X, Y]$, the *Motzkin polynomial*. We will encounter this famous polynomial again in the next section. The degree of m is 6 and the decomposition of m into forms as in (1) is $m = m_6 + m_4 + m_0$, where $m_6 = X^4Y^2 + X^2Y^4$, $m_4 = -3X^2Y^2$, and $m_0 = 1$. The degree 6 homogenization of m is $\bar{m} = X^4Y^2 + X^2Y^4 - 3X^2Y^2Z^2 + Z^6$, a form in $\mathbb{R}[X, Y, Z]$.

Many of the ideas and results we discuss in this article originated with the work of David Hilbert in the late 19th century. He worked with forms, however we prefer to work with polynomials; it turns out that for the purposes of results on sums of squares and positive polynomials, it does not matter which setting is used, as we shall see.

3 SOS and PSD polynomials

A polynomial $f \in \mathbb{R}[\underline{X}]$ is a **sum of squares**, **sos** for short, if there exist $g_1, \ldots, g_k \in \mathbb{R}[\underline{X}]$ such that

$$f = g_1^2 + \cdots + g_k^2 \tag{2}$$

An **sos representation** for f is an equation of this type. A polynomial can have more than one sos representation, for example,

$$X^4 + Y^4 + 1 = (X^2)^2 + (Y^2)^2 + 1^2 = (X^2 - Y^2)^2 + (\sqrt{2}XY)^2 + (1)^2.$$

For a commutative ring A, we denote the set of sums of squares of elements in A by $\sum A^2$. In particular, $\sum \mathbb{R}[\underline{X}]^2$ denotes the set of sos polynomials in $\mathbb{R}[\underline{X}]$.

A polynomial $f \in \mathbb{R}[\underline{X}]$ is **positive semidefinite**, **psd** for short, if $f(x) \geq 0$ for all $x \in \mathbb{R}^n$.

We list some useful properties of psd and sos polynomials.

Proposition 2. Given $f \in \mathbb{R}[\underline{X}]$.

1. If f is psd, then the degree of f must be even.
2. If $f \in \mathbb{R}[T]$ is a psd polynomial in one variable, then the leading coefficient of f is positive and all real roots of f appear with even multiplicity.

3. If f is sos and $f = g_1^2 + \cdots + g_k^2$, then $\deg f = 2\max\{\deg g_1, \ldots, \deg g_k\}$.

Proof. 1. This is clear if f is a polynomial in one variable since a polynomial of odd degree has at least one root with odd multiplicity and hence changes sign at this root. Suppose $f \in \mathbb{R}[\underline{X}]$ is psd and has degree d, then we can decompose f into forms as in (1), say $f = f_d + \cdots + f_0$. Suppose d is odd. Fix $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and consider

$$g(T) := f(x_1T, x_2T, \ldots, x_nT) = \sum_{i=0}^{d} f_i(x)T^i,$$

a polynomial in $\mathbb{R}[T]$. Since f is psd, for any $t \in \mathbb{R}$, $g(t) = f(t \cdot x) \geq 0$, hence g is psd. Since $g(T)$ is a polynomial in one variable, it follows that $g(T)$ must have even degree,. Since d is odd, this implies that $f_d(x) = 0$. But this must hold for any $x \in \mathbb{R}^n$, i.e., $f_d(x) = 0$ for all $x \in \mathbb{R}^n$. By Proposition 1, it follows that f_d is the zero polynomial, which contradicts $\deg f = d$. Thus the degree of f must be even.

2. Suppose $f(T) \in \mathbb{R}[T]$ is psd. The degree of f is even hence if the leading coefficient were negative we would have $\lim_{T\to\infty} f(T) = -\infty$, which contradictions f psd. Since f cannot change sign at a root, all roots must appear with even multiplicity.

3. For each i, let $r_i = \deg g_i$ and set $d = \max\{r_i\}$. Then for each i, $\deg g_i \leq d$ which implies that $\deg g_i^2 \leq 2d$. Hence $\deg f \leq 2d$. Now let f_{2d} denote the homogeneous term of degree $2d$ in f, we need to show that $f_{2d} \neq 0$. For each i, let $g_{i,d}$ denote the sum of the terms in g_i of degree d, so that $g_{i,d} = 0$ if $r_i < d$ and $g_{i,d}$ is the leading form of g_i if $r_i = d$. By assumption, there is at least one i such that $g_{i,d} \neq 0$. The only way a term of degree $2d$ can appear in g_i^2 is from $g_{i,d}^2$, hence $f_{2d} = \sum_i g_{i,d}^2$. If $f_{2d} = 0$, then we would have $\sum_i g_{i,d}(x)^2 = 0$ for all $x \in \mathbb{R}^n$. By one of the basic properties of $\mathbb{R}$, this would imply that $g_{i,d}(x) = 0$ for all x and all i, which implies $g_{i,d}$ is the zero polynomial for all i. But this contradicts $g_{i,d} \neq 0$ for at least one i. Hence $f_{2d} \neq 0$ which implies $\deg f = 2d$. □

Given $g \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$, then $g^2(x) = (g(x))^2 \geq 0$ and hence a sum of squares $g_1^2 + \cdots + g_k^2$ take only nonnegative values. In other words

If $f \in \mathbb{R}[\underline{X}]$ is sos, then f is psd.

Question: f sos implies that f is psd. Does f psd imply f is sos?

In general, the answer to this question is "no", however in some special cases the answer is "yes". Two such cases – polynomials in one variable and polynomials of degree 2 – were well-known by the late 19th century.

For $n, d \in \mathbb{N}$, let $\mathrm{PSD}(n, d)$ denote the set of psd polynomials in $\mathbb{R}[\underline{X}]$ of degree d and let $\mathrm{SOS}(n, d)$ denote the set of sos polynomials in $\mathbb{R}[\underline{X}]$ of degree d. By Proposition 2, if d is odd then $\mathrm{PSD}(n, d) = \mathrm{SOS}(n, d) = \emptyset$, thus we do not consider this case further. We have seen that for all n and d, $\mathrm{SOS}(n, 2d) \subseteq \mathrm{PSD}(n, 2d)$.

Theorem 1. If $f \in \mathbb{R}[T]$ is a psd polynomial in one variable, then f can be written as a sum of two squares of polynomials in $\mathbb{R}[T]$.

Proof. By the Fundamental Theorem of Algebra, f factors into linear factors over $\mathbb{C}$. Since f has real coefficients, any roots in $\mathbb{C} \setminus \mathbb{R}$ must appear in conjugate pairs. By

Proposition 2 the leading coefficient of f is positive, hence equals c^2 for some nonzero $c \in \mathbb{R}$. Suppose the real roots of f are $a_1, \ldots, a_s$ with multiplicities $2u_1, \ldots, 2u_s$ and the non-real roots of f are the conjugate pairs $\{b_1 \pm ic_1, \ldots, b_t \pm ic_t\}$. Then we have

$$f = c^2 \cdot \prod_{j=1}^{s} (T - a_j)^{2u_j} \cdot \prod_{k=1}^{t} (T - (b_k + i\, c_k))(T - (b_k - i\, c_k)).$$

Let $p(T) = c \cdot \prod_{j=1}^{s} (T - a_j)^{u_j}$. Expanding the second product, we can write it as $(q(T) + i\, r(T))(q(T) - i\, r(T))$, where $q(T), r(T) \in \mathbb{R}[T]$. Then we have

$$f = (p(T))^2 \cdot (q(T) + i\, r(T))(q(T) - i\, r(T)) = (p(T)q(T))^2 + (p(T)r(T))^2.$$

□

Example 2. Let $f = T^4 - T^2 + 2T + 2$, which is psd. The roots of f are -1 and the conjugate pair $1 \pm i$ and the factorization of f is

$$f = (T + 1)^2 (T - (1 + i))(T - (1 - i)).$$

In the notation of the proof of the theorem we have $p(T) = T + 1$, $q(T) = T - 1$, and $r(T) = 1$. This yields $f = (T^2 - 1)^2 + (T + 1)^2$.

Theorem 2. If $f \in \mathbb{R}[\underline{X}]$ is psd and $\deg f = 2$, then f is a sum of squares of linear polynomials.

Proof. It is enough to show that $\bar{f}$, the homogenization of f is sos. Since f is psd, $\bar{f}$ is a psd quadratic form. The result follows from the well-known fact that a psd quadratic form can be diagonalized into a sum of squares of linear forms. □

The previous two results show that $\mathrm{SOS}(1, 2d) = \mathrm{PSD}(1, 2d)$ for all d and $\mathrm{SOS}(n, 2) = \mathrm{PSD}(n, 2)$ for all n; these results were well-known by the late 19th century. In 1888, the 26-year old David Hilbert proved two remarkable theorems in one paper [4], settling the question of whether $\mathrm{SOS}(n, 2d) = \mathrm{PSD}(n, 2d)$ in all remaining cases.

Remark 1. As mentioned above, Hilbert worked with forms. Since the process of dehomogenization preserves the properties sos and psd, Hilbert's results are valid in the polynomial setting and we state his results in terms of polynomials.

Theorem 3 (Hilbert)**.**

1. $\mathrm{SOS}(2, 4) = \mathrm{PSD}(2, 4)$. In particular, if f is a polynomial of degree 4 in two variables and f is psd, then f is a sum of three squares of quadratic polynomials.
2. In all other cases, i.e., if $n \geq 2$ and $d \geq 3$ or if $n \geq 3$ and $d \geq 2$, there exists $f \in \mathrm{PSD}(n, 2d)$ such that $f \notin \mathrm{SOS}(n, 2d)$.

Hilbert's proofs were not constructive, in particular, he did not give an explicit example of a psd polynomial that is not sos. The first published example appeared in 1967, due to Motzkin [5]. The Motzkin polynomial is in $\mathrm{PSD}(2, 6) \setminus \mathrm{SOS}(2, 6)$. Recall

$$m(X, Y) = X^4Y^2 + X^2Y^4 - 3X^2Y^2 + 1.$$

Proposition 3. The Motzkin polynomial $m(X, Y)$ is psd.

Proof. Given $(x, y) \in \mathbb{R}^2$, let $a = x^4y^2$, $b = x^2y^4$, and $c = 1$. By the Arithmetic-Geometric Mean with $k = 3$, we have $\frac{1}{3}(x^4y^2 + x^2y^4 + 1) \geq x^2y^2$, hence $x^4y^2 + x^2y^4 - 3x^2y^2 + 1 \geq 0$, which implies $m(x, y) \geq 0$. Thus m is psd. □

Proposition 4. $m(X, Y)$ is not sos.

Proof. Assume m is sos, then by Proposition 2, since $\deg m = 6$, we have $m = \sum g_i^2$ where $\deg g_i \leq 3$ for all i. Thus we have

$$m = \sum_i (a_iX^3 + b_iX^2Y + c_iXY^2 + d_iY^3 + e_iX^2 + f_iXY + g_iY^2 + h_iX + j_iY + k_i)^2, \quad (3)$$

where $a_i, b_i, \ldots, k_i \in \mathbb{R}$. The coefficients on each side of (3) must agree. Consider the coefficient of X^6, which is 0. On the right-hand side, the only way to get an X^6 term is from $(a_iX^3)^2$, hence we have $0 = \sum_i a_i^2$. This implies that $a_i = 0$ for all i, i.e., there are no X^3 terms in the g_i's. A similar argument shows that there are no Y^3 terms in the g_i's. Now look at the coefficient of X^4, which must be 0 on both sides of the equation. On the right-hand side, the coefficient of X^4 is $\sum_i(e_i^2 + a_ih_i)$. Since $a_i = 0$, we have $0 = \sum e_i^2$, which implies $e_i = 0$ for all i. A similar argument, using the coefficient of Y^4, shows that $g_i = 0$ for all i. Finally, consider the coefficient of X^2Y^2. Looking at the coefficients of X^2 and Y^2, a similar argument shows that $h_i = j_i = 0$ for all i. Thus we have

$$m = \sum_i (b_iX^2Y + c_iXY^2 + f_iXY + k_i)^2.$$

Looking at the coefficient of X^2Y^2 on both sides of the equation we have $-3 = \sum_i f_i^2$, a contraction. Thus no such sos representation can exist for m.

□

3.1 Hilbert's 17th Problem

In 1900, in Hilbert's address at the International Congress of Mathematicians in Paris, he posed a generalization of his results as the 17th Problem: Must every psd form p be a sum of squares of quotients of forms? Dehomogenizing, this question becomes the following: Given $f \in \mathbb{R}[\underline{X}]$ such that f is psd, do there exist $g_1, \ldots, g_k, h_1, \ldots h_k \in \mathbb{R}[\underline{X}]$ such that

$$f = \frac{g_1^2}{h_1^2} + \cdots + \frac{g_k^2}{h_k^2}?$$

Clearing denominators, we get an equivalent formulation: Do there exist $g_1, \ldots, g_k, h \in \mathbb{R}[\underline{X}]$ such that

$$h^2 f = g_1^2 + \cdots + g_k^2? \quad (4)$$

In 1927, Emil Artin [1] used the Artin-Schreier theory of real closed fields to answer HIlbert's 17th problem:

Theorem 4 (Artin)**.** If $f \in \mathbb{R}[\underline{X}]$ is psd, then there exists $h \in \mathbb{R}[\underline{X}]$ such that $h^2 f$ is sos.

Example 3. Here is a representation of the Motzkin polynomial as a sum of squares of rational functions:

$$X^4Y^2 + X^2Y^4 - 3X^2Y^2 + 1 = \frac{X^2Y^2(X^2+Y^2+1)(X^2+Y^2-2)^2 + (X^2-Y^2)^2}{(X^2+Y^2)^2}$$

Suppose we have an equation of the form (4) for f, then since h^2 and $g_1^2 + \cdots + g_k^2$ are psd, we can see immediately that f is psd. In other words, we have a **certificate of non-negativity** for f, an immediate proof that f is psd. Artin's Theorem says that such certificates always exist as long as don't insist that the certificate be in the polynomial ring $\mathbb{R}[\underline{X}]$; we must allow denominators. In general we will use the word *certificate* (of positivity/nonnegativity on a set in $\mathbb{R}^n$) to denote an equation which yields an immediate proof of the condition.

Question: What can we say if we replace the condition $f(x) \geq 0$ for all $x \in \mathbb{R}^n$ by a positivity condition $f(x) \geq 0$ for all $x \in S$, where S is a semialgebraic set in $\mathbb{R}^n$?

We will discuss this question in the next section.

4 Representation Theorems

Writing a psd polynomial f as a sum of squares of polynomials or rational functions is a certificate of nonnegativity for f, a proof that f is psd. In this section we generalize this idea to polynomials that take only nonnegative (or only positive) values on semialgebraic sets in $\mathbb{R}^n$. For a subset $S \subseteq \mathbb{R}^n$ and $f \in \mathbb{R}[\underline{X}]$, we write $f \geq 0$ on S ($f > 0$ on S) to denote that $f(x) \geq 0$ ($f(x) > 0$) for all $x \in S$. A theorem about the existence of certificates for $f \geq 0$ or $f > 0$ on a semialgebraic set S is often called a **representation theorem** and an equation which provides a certificate for f a **representation** of f.

Definition 1. For $g_1, \ldots g_k \in \mathbb{R}[\underline{X}]$, the **basic closed semialgebraic set** generated by $g_1, \ldots, g_k$, denoted $S(g_1, \ldots, g_k)$, is $\{x \in \mathbb{R}^n \mid g_i(x) \geq 0 \text{ for } i = 1, \ldots k\}$. In other words, $S(g_1, \ldots, g_k)$ is the solution set in $\mathbb{R}$ to the finite set of inequalities $\{g_1 \geq 0, \ldots, g_k \geq 0\}$.

We define an algebraic object in $\mathbb{R}[\underline{X}]$ associated to $S(g_1, \ldots, g_k)$: The **preorder** generated by $g_1, \ldots, g_k$, denoted $P(g_1, \ldots, g_k)$, is the set of $f \in \mathbb{R}[\underline{X}]$ which can be written as a linear combination over $\sum \mathbb{R}[\underline{X}]^2$ of products of the g_i's. In other words, $P(g_1, \ldots, g_k)$ consists of all elements in $\mathbb{R}[\underline{X}]$ of the form

$$\sum_{\epsilon=(\epsilon_1,\ldots,\epsilon_k)\in\{0,1\}^k} \sigma_\epsilon \, g_1^{\epsilon_1} \ldots g_k^{\epsilon_k}, \tag{5}$$

where $\sigma_\epsilon \in \sum \mathbb{R}[\underline{X}]^2$ for all ϵ. Notice that for $g \in \mathbb{R}[\underline{X}]$ and any $m \in \mathbb{N}$, $g^{2m+1} = (g^m)^2 \cdot g$, which is of the form $\sigma \cdot g$ with $\sigma \in \sum \mathbb{R}[\underline{X}]^2$. Thus in order to obtain all possible products of the g_i's we need only use exponents 0 or 1. For example, for $g, h \in \mathbb{R}[\underline{X}]$,

$$P(g) = \{\sigma_0 + \sigma_1 g \mid \sigma_0, \sigma_1 \in \sum \mathbb{R}[\underline{X}]^2\},$$

$$P(g,h) = \{\sigma_0 + \sigma_1 g + \sigma_2 h + \sigma_3 gh \mid \sigma_0, \sigma_1, \sigma_2, \sigma_3 \in \sum \mathbb{R}[\underline{X}]^2\}.$$

An explicit expression for $f \in \mathbb{R}[\underline{X}]$ of the form (5) is called a representation of f in $P(g_1, \ldots, g_k)$. For example, in $\mathbb{R}[T]$,

$$1 - T = \frac{1}{2}(1-T)^2 + \frac{1}{2}(1-T^2)$$

is a representation of $1 - T$ in $P(1 - T^2)$.

Fix $g_1, \ldots, g_k \in \mathbb{R}[\underline{X}]$ and set $S := S(g_1, \ldots, g_k)$ and $P := P(g_1, \ldots, g_k)$. For any $x \in S$ and $\sigma \in \sum \mathbb{R}[\underline{X}]^2$, $\sigma(x) \geq 0$ since σ is psd and $g_i(x) \geq 0$ by definition of S. This implies that for any $f \in P$, $f \geq 0$ on S. Furthermore, a representation of f in P is a certificate for $f \geq 0$ on S.

Question: Suppose $f \geq 0$ on $S(g_1, \ldots, g_k)$, does this imply that $f \in P(g_1, \ldots, g_k)$?

The answer to this question is "no" since it includes the question of whether a psd polynomial is sos. As in the psd/sos case, there are some special cases where the answer is "yes" and we will see that we can obtain a general result if we relax the requirement that the certificates are in the polynomial ring $\mathbb{R}[\underline{X}]$. Recall that Artin's Theorem says that if f is psd then $h^2 f$ is sos for some polynomial h. We refer to the h^2 term as a denominator and more generally we will talk about representations with denominators when we obtain a representation of f in the preorder (or some other certificate in $\mathbb{R}[\underline{X}]$) after multiplying by a polynomial that is obviously nonnegative on the semialgebraic set.

In 1974, Stengle [9] proved his celebrated Positivstellensatz, which showed that for any basic closed semialgebraic set S there always exist certificates for $f > 0$ and $f \geq 0$ on S, provided we allow denominators. In some sense, this theorem marks the beginning of modern real algebraic geoemtry. There are several versions of this famous theorem, here is one.

Positivstellensatz. With S and P as above, we have

1. $S \neq 0$ if and only if $-1 \in P$.
2. If $f \geq 0$ on S, then there exist $p, q \in P$ and an integer $m \geq 1$ such that $pf = f^{2m} + q$.
3. If $f > 0$ on S, then there exist $p, q \in P$ such that $pf = 1 + q$.

Suppose $x \in S$ and we have $pf = f^{2m} + q$ for $p, q \in P$. Then $p(x), q(x)$, and $f^{2m}(x)$ are all nonnegative, hence $f(x)$ must be nonnegative. Similarly, if we have $pf = 1 + q$, then plugging x into both sides of this equation implies $f(x) > 0$. Thus the Positivstellensatz implies the existence of certificates for $f \geq 0$ and $f > 0$ for all basic closed semialgebraic sets S if we allow representations with denominators.

If the generators $g_1, \ldots, g_k$ are linear, then we can obtain representation theorems without denominators in some cases. In 1928, Pólya's [6] proved a representation theorem for forms positive on the standard simplex.

Let $\Delta_n = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \geq 0 \text{ for all } i \text{ and } \sum_{i=1}^n x_i = 1\}$, the standard n-simplex. Notice that Δ_n is a basic closed semialgebraic set:

$$\Delta_n = S(X_1, \ldots, X_n, 1 - \sum X_i).$$

Theorem 5 (Pólya)**.** Suppose $f \in \mathbb{R}[\underline{X}]$ such that $f > 0$ on Δ_n. Then there exists $N \in \mathbb{N}$ such that $(X_1 + \cdots + X_n)^N f$ has only positive coefficients.

Suppose $f > 0$ on Δ_n and $\deg f = d$. Let N be as in the theorem, then we have

$$(X_1 + \cdots + X_n)f = \sum_{|\alpha| \leq d+N} a_\alpha X^\alpha, \tag{6}$$

where $a_\alpha > 0$ for all α. Let $x \in \Delta_n$, then plugging x into both sides of equation (6) we see immediately that $f(x) > 0$. Hence (6) is a certificate for $f > 0$ on Δ_n.

Handelman's Theorem from 1988 [3] applies to polynomials positive on any compact polyhedron. Fix $g_1 \ldots, g_k \in \mathbb{R}[\underline{X}]$. We use the following notation for products of the g_i's: For $\beta = (\beta_1, \ldots, \beta_k) \in \mathbb{Z}^k_{\geq 0}$, let g^β denote $g_1^{\beta_1} \ldots g_k^{\beta_k}$.

Theorem 6 (Handelman)**.** Suppose $g_1, \ldots, g_k \in \mathbb{R}[\underline{X}]$ are linear and the polyhedron $S = S(g_1, \ldots, g_k)$ is nonempty and bounded. Given $f \in \mathbb{R}[\underline{X}]$, if $f > 0$ on S, then f can be written as a positive linear combination of products of the g_i's. In other words, there exist $\beta(1), \ldots, \beta(r) \in \mathbb{Z}^k_{\geq 0}$ and $a_1, \ldots, a_r \in \mathbb{R}^+$ such that

$$f = \sum_{i=1}^{r} a_i g^{\beta(i)}.$$

In 1991, Schmüdgen [8] proved a beautiful and unexpected theorem that gives a denominator-free version of the Positivstellensatz in the case where the basic closed semialgebraic set is compact.

Schmüdgen's Theorem. Given $g_1, \ldots, g_k \in \mathbb{R}[\underline{X}]$, let $S = S(g_1, \ldots, g_k)$ and $P = P(g_1, \ldots, g_k)$. Suppose S is compact, then for all f such that $f > 0$ on S, $f \in P$.

Note that the theorem is not true in general if we drop the assumption that S is compact or if we relax $f > 0$ on S to $f \geq 0$ on S.

One remarkable aspect of Schmüdgen's Theorem is that it does not depend on the choice of generators $g_1, \ldots, g_k$ for S. Consider a very simple case when $n = 1$: $S = [-1, 1]$. Then $S = S(1 - T^2)$ and the theorem says that for any $f \in \mathbb{R}[T]$ such that $f > 0$ on $[-1, 1]$, there exist $\sigma_0, \sigma_1 \in \sum \mathbb{R}[T]^2$ such that $f = \sigma_0 + \sigma_1(1 - T)^2$. This seems reasonable, in fact, given such f it is not too difficult to obtain an explicit representation of this type. But we also have $[-1, 1] = S((1 - T)^{2017})$ and the theorem says that for any $f > 0$ on $[-1, 1]$ there exist $\sigma_0, \sigma_1 \in \sum \mathbb{R}[T]^2$ such that $f = \sigma_0 + \sigma_1(1 - T^2)^{2017}$. This is surprising, and for specific f it is not clear how we could find a representation of this type.

Schmüdgen's Theorem is very general, but as we increase the number of generators, the number of possible terms in a representation in the preorder increases exponentially since there are 2^k different products of k generators. In 1993, M. Putinar proved a representation theorem that gives simpler representations and applies if we have an additional condition on our set of generators.

The **quadratic module** generated by $g_1, \ldots, g_k \in \mathbb{R}[\underline{X}]$, denoted $\mathrm{QM}(g_1, \ldots, g_k)$, is the set of linear combinations of the g_i's over $\sum \mathbb{R}[\underline{X}]^2$, in other words, all elements in $\mathbb{R}[\underline{X}]$ of the form

$$\sigma_0 + \sigma_1 g_1 + \cdots + \sigma_k g_k,$$

where $\sigma_0, \ldots, \sigma_k \in \sum \mathbb{R}[\underline{X}]^2$. Notice that the preorder $P(g_1, \ldots, g_k)$ is a quadratic module, in fact, the quadratic module generated by all products $g_1^{\epsilon_1} \ldots g_k^{\epsilon_k}$, where $\epsilon_i \in \{0, 1\}$ for all i.

A quadratic module $Q = QM(g_1, \ldots, g_k) \subseteq \mathbb{R}[\underline{X}]$ is **Archimedean** if for every $h \in \mathbb{R}[\underline{X}]$ there exists $N \in \mathbb{N}$ such that $N \pm h \in Q$.

Theorem 7. The following are equivalent for a quadratic module $QQM(g_1, \ldots g_k)$.

1. Q is Archimedean.
2. There exists $N \in \mathbb{N}$ such that $N - \sum X_i^2 \in Q$.
3. There exists $h \in Q$ such that $S(h)$ is compact.

It is important to note that the Archimedean property can depend on the choice of generators $g_1, \ldots, g_k$.

Putinar's Theorem. Given $g_1, \ldots, g_k \in \mathbb{R}[\underline{X}]$, let $S = S(g_1, \ldots, g_k)$ and $Q = QM(g_1, \ldots, g_k)$. Suppose Q is Archimedean. For $f \in \mathbb{R}[\underline{X}]$, if $f > 0$ on S, then $f \in Q$.

Remark 2. Given $S = S(g_1, \ldots, g_k)$, then S compact does not in general imply the $Q(g_1, \ldots, g_k)$ is Archimedean. On the other hand, if we know or can compute some $N \in \mathbb{N}$ such that S is contained in the ball $\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid N - \sum x_i^2\}$, then it follows that $N - \sum X_i^2 \geq 0$ on S. We can add $N - \sum X_i^2$ to the set of generators for S without changing S and we have immediately that $Q(g_1, \ldots, g_k, N - \sum X_i^2)$ is Archimedean. Thus by adding just one extra generator we are in a situation where Putinar's Theorem applies.

5 The Gram matrix method

The Gram matrix method, developed by Choi, Lam, and Reznick [2], is a method for finding and counting sos representations of a polynomial. We first recall some linear algebra that will be needed. In this section we will view $x \in \mathbb{R}^n$ as a row vector, i.e., a $1 \times n$ matrix.

Definition 2. Let A be a $k \times k$ matrix over $\mathbb{R}$.

1. A is **positive semidefinite**, or psd, if for all $x \in \mathbb{R}^k$,
$$x \cdot A \cdot x^T \geq 0.$$
2. For $1 \leq j \leq k$, the $j \times j$ **leading principal minor** of A is the determinant of the $j \times j$ submatrix of A in the upper left corner of A.
3. A is **orthogonal** if $A \cdot A^T = A^T \cdot A = I$, where I is the $k \times k$ identity matrix. This means that all rows and columns are unit vectors and if $x, y \in \mathbb{R}^k$ are two different rows or two different columns of A, then $x \cdot y^T = 0$.

Lemma 1. For any matrix B over $\mathbb{R}$, $B \cdot B^T$ is psd.

Proof. Suppose $B = (b_{ij})$ is a $k \times m$ matrix. Then $B \cdot B^T$ is a $k \times k$ matrix and for any $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$,

$$x \cdot B \cdot B^T \cdot x^T = \sum_{i=1}^{m} (\sum_{j=1}^{k} x_j b_{ji})^2 \geq 0.$$

□

Theorem 8. The following are equivalent for a $k \times k$ matrix A over $\mathbb{R}$.

1. A is psd.
2. All eigenvalues of A are nonnegative.
3. All leading principal minors of A are nonnegative.

As we have seen, a polynomial f can have more than one sos representation. We would like to count the number of sos representations for a given f, but we first need to determine what it means for two sos representations to be (essentially) different. For example, given $f, g \in \mathbb{R}[\underline{X}]$, it seems reasonable to consider the two representations

$$f^2 + g^2 = \frac{1}{2}(f+g)^2 + \frac{1}{2}(f-g)^2,$$

as essentially the same representation. More generally, if $A = (a_{ij})$ is a real $k \times k$ orthogonal matrix and $g_1, \ldots, g_k \in \mathbb{R}[\underline{X}]$, then

$$g_1^2 + \cdots + g_k^2 = \sum_{i=1}^{k} (\sum_{j=1}^{k} a_{ij} g_j)^2. \tag{7}$$

We say that the two sos representations in (7) are **orthogonally equivalent** and we count sos representations up to orthogonal equivalence.

Given $f \in \mathbb{R}[X]$ of degree $2d$, write $f = \sum_{|\alpha| \leq 2d} a_\alpha X^\alpha$. Suppose f is sos, say

$$f = g_1^2 + \cdots + g_k^2. \tag{8}$$

We want to write (8) in matrix form. Since $\deg f = 2d$, by Proposition 2 each g_i has degree at most d. Assume $g_i = \sum_{|\beta| \leq d} b_\beta^{(i)} X^\beta$.

For each $\beta \in \mathbb{N}^n$ with $|\beta| \leq d$, set $U_\beta = (b_\beta^{(1)}, \ldots, b_\beta^{(t)})$, the vector of coefficients of X^β in the g_i's. Then (8) becomes

$$f = \sum_{\beta, \beta'} U_\beta \cdot U_{\beta'} X^{\beta + \beta'}. \tag{9}$$

For each α with $|\alpha| \leq 2d$, the coefficients of X^α on both sides of (9) must agree, hence, for each such α,

$$a_\alpha = \sum_{\beta + \beta' = \alpha} U_\beta \cdot U_{\beta'}. \tag{10}$$

The matrix $V := [U_\beta \cdot U_{\beta'}]$, where the rows and columns are indexed by $\beta \in \mathbb{N}^n$ with $|\beta| = d$, is the **Gram matrix** of p associated to (8). Let $V = (v_{\beta,\beta'})$. Then V is clearly symmetric and is psd by Lemma 1. Furthermore, the entries satisfy the equations

$$a_\alpha = \sum_{\beta+\beta'=\alpha} v_{\beta,\beta'}. \tag{11}$$

Here is another way to construct the Gram matrix associated to (8). Let $N = \binom{n+d}{d}$, the number of monomials in $\mathbb{R}[X]$ with degree at most d, and let $\{\beta_1, \ldots, \beta_N\}$ be these monomials. Let $M = (X^{\beta_1}, \ldots, X^{\beta_N})$ and let B be the $N \times k$ matrix with i, j entry the coefficient of X^{β_i} in g_j. In other words, the j-th column of B consists of the coefficients of g_j, in the order given by M. Then (8) becomes

$$f = M \cdot B \cdot B^T \cdot M^T,$$

and $B \cdot B^T$ is the Gram matrix of f associated to (8).

Example. In $\mathbb{R}[T]$, let $f(T) = T^4 + 5T^2 + 4 = (T^2 + 2)^2 + T^2$. Here $d = 1$, $k = 2$, $N = 3$, and we set $M = (T^2, T, 1)$. With $g_1 = T^2 + 2$ and $g_2 = T$, we have

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{bmatrix},$$

so that

$$B \cdot B^T = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 4 \end{bmatrix}$$

is the Gram matrix associated to the given sos representation.

The following theorem is due to Choi, Lam, and Reznick [2].

Theorem 9. Suppose $f \in \mathbb{R}[\underline{X}]$ has degree $2d$, say $f = \sum_{|\alpha| \le 2d} a_\alpha X^\alpha$. Let $N = \binom{n+d}{d}$ and let $M = (X^{\beta_1}, \ldots X^{\beta_N})$ be the vector of monomials of degree d or less. Suppose $V = (v_{ij})$ is a symmetric $N \times N$ matrix over $\mathbb{R}$.

1. f is a sum of squares and V is the Gram matrix associated to f with respect to some sos representation if and only if V is psd and for all $\alpha \in \mathbb{Z}_{\ge 0}$ such that $|\alpha| \le 2d$, the entries of V satisfy the equation

$$\sum_{\beta_i+\beta_j=\alpha} v_{ij} = a_\alpha \tag{12}$$

2. If V is the Gram matrix of f associated to $f = \sum_{i=1}^k g_i^2$, then $rank(V) = k$.

3. Two sos representations of f are orthogonally equivalent if and only if they have the same Gram matrix.

How can we use the theorem to decide if $f \in \mathbb{R}[\underline{X}]$ is sos or not and count the number of different sos representations? According the the theorem, this amounts to finding and counting real, symmetric, psd matrices whose entries satisfy the equations (12). Here is one method for finding and counting such matrices.

Given $f \in \mathbb{R}[\underline{X}]$ with $\deg f = 2d$, let $V = (v_{ij})$ be an $N \times N$ symmetric matrix with variable entries, where $N = \binom{n+d}{d}$. Since V is symmetric, there are $N(N+1)/2$ variables in V. The equations (12) define a linear system in these $N(N+1)/2$ variables with $\binom{n+2d}{2d}$ equations. This linear system is underdetermined in general and is trivial to solve since each variable v_{ij} appears in only one equation! In general, the solution will be in terms of a number of parameters.

Once we have solved the linear system coming from (12), we obtain the general form of Gram matrix for f, the **general Gram matrix** for f. This matrix is symmetric with entries linear in the parameters. Theorem 9 says that f is sos if and only if we can find values for the parameters which make the matrix V psd. We can use one of the equivalent conditions for a matrix to be psd from Theorem 1.

Example 4. Consider again $f(T) = T^4 + 5T^2 + 4$ and suppose we want to find all possible sos representations. The general Gram matrix is 3×3, suppose it is $V = (v_{i,j})$ with $v_{j,i} = v_{i,j}$. There are 5 monomials of degree 4 or less and there are 6 variables in V. The linear system coming from (12) in this case is

$$\begin{aligned} v_{1,1} &= 1 \\ 2v_{1,2} &= 0 \\ 2v_{1,3} + v_{2,2} &= 5 \\ 2v_{2,3} &= 0 \\ v_{3,3} &= 4 \end{aligned}$$

The solution contains one parameter, let's call it r, and the general form of a Gram matrix for f is

$$V(r) := \begin{bmatrix} 1 & 0 & r \\ 0 & 5-2r & 0 \\ r & 0 & 4 \end{bmatrix}.$$

We need to determine for which values of r this matrix is psd. One way to do this is to find the eigenvalues of this matrix (by hand or using computer algebra software); they are $5 - 2r$ and $1/2(5 \pm \sqrt{4r^2 + 9})$. These are all nonnegative if and only if $-2 \leq r \leq 2$. It follows that hence f has infinitely many Gram matrices and therefore infinitely many different sos representations.

It is easy to check that the rank of $V(\pm 2)$ is 2 and for $-2 < r < 2$ the rank of $V(r)$ is 3. In fact, all sos representations of f are given by

$$f = (T^2 + r)^2 + (\sqrt{5-2r}\,T)^2 + (\sqrt{4-r^2})^2, -2 \leq r \leq 2.$$

The alert reader will wonder how we went from the Gram matrix to the sos representation. Here is one way to do it, which can be done "by hand" for small examples: Suppose $V = (v_{i,j})$ is an $N \times N$ Gram matrix for f, then associated to V is a psd quadratic form defined by

$$q(X_1, \ldots, X_N) = \sum_{i,j} v_{i,j} X_i X_j.$$

Because q is psd, it can be written as a sum of squares of linear forms. Once we obtain this sos representation of q, which can be done by completing the square, if we plug in X^{β_i} for X_i we will obtain the sos representation for f corresponding to V.

In Example 4, the quadratic form corresponding to $V(r)$ is

$$X_1^4 + 2rX_1X_3 + (5-2r)X_2^2 + 4X_3^2 = (X_1+r)^2 + (\sqrt{5-2r})X_2)^2 + (\sqrt{5-2r}\,X_3)^2.$$

If $f \in SOS(n, 2d)$ then, in general, the size of the Gram matrix will be $N \times N$, where $N = \binom{n+d}{d}$. This number N grows very fast as n and d increase. However, in some cases the size of the general Gram matrix can be reduced since if we know that a certain monomial X^β of degree d cannot appear in the g_i's whenever $f = \sum g_i^2$, then we can omit this monomial from the vector M. This was the case for the Motzkin polynomial.

Example 5. Let $m = X^4Y^2 + X^2Y^4 - 3X^2Y^2 + 1$. We will find the general Gram matrix of m and use this to show that m is not sos. In our proof that m is not sos, we showed that if $m = \sum_i g_i^2$, then the only monomials that can occur in the g_i's are X^2Y, XY^2, XY, and 1. Thus the general Gram matrix of m is 4×4 and it is not to hard to show that (with the given order of the monomials) it is

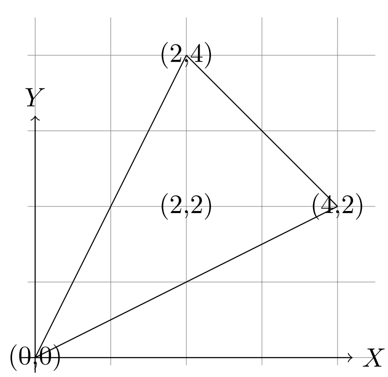
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This matrix is not psd, hence m is not sos.

We want to give a general result about the monomials that can occur in the g_i's if $f = \sum g_i^2$. The criterion involves the Newton polytope of f.

Definition 3. The **Newton polytope** of $f \in \mathbb{R}[\underline{X}]$, denoted $\text{NP}(f)$, is the convex hull of $\text{supp}(f)$ in $\mathbb{R}^n$, i.e., the smallest convex set in $\mathbb{R}^n$ that contains $\text{supp}(f)$.

Example 6. Consider the Motzkin polynomial $m = X^4Y^2 + X^2Y^4 - 3X^2Y^2 + 1$, then $\text{supp}(M) = \{(2,2), (2,0), (0,2), (0,0)\}$ and $\text{NP}(f)$ is the triangle with vertices $(4,2)$, $(2,4)$, and $(0,0)$:



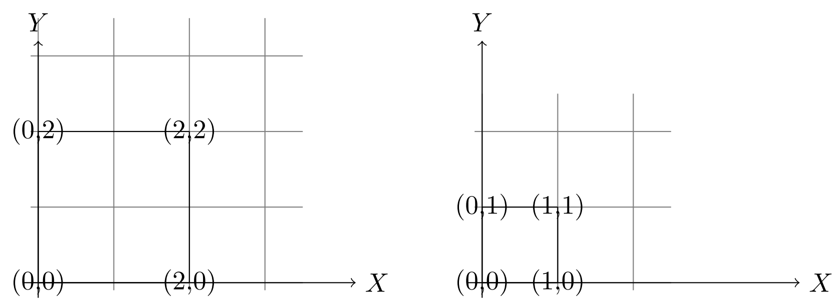
The following useful theorem can be found in [7].

Theorem 10. Suppose $f = g_1^2 + \cdots + g_k^2$ is sos. Then $2\,\mathrm{NP}(g_i) \subseteq \mathrm{NP}(f)$ for all i.

Let m be the Motzkin polynomial, then we saw that $\mathrm{NP}(m)$ is the triangle with vertices $(4, 2)$, $(2, 4)$, and $(0, 0)$. The theorem says that if $m = \sum g_i^2$, then for all i, $\mathrm{NP}(g_i)$ is contained in the triangle with vertices $(2, 1)$, $(1, 2)$, and $(0, 0)$. We proved this in Section 3 in the course of proving that m is not sos.

We end with another example of the Gram matrix method.

Example 7. Let $f(X, Y) = X^2Y^2 + X^2 + Y^2 + 1$, then f is visibly a sum of squares: $f = (XY)^2 + X^2 + Y^2 + 1^2$. We want to find all sos representations of f. In particular, we might want to know whether there is an sos representation for f with fewer than four squares. $\mathrm{NP}(f)$ is a rectangle with vertices $(2, 2)$, $(2, 0)$, $(0, 2)$, and $(0, 0)$ and hence, by Theorem 10, if $f = \sum g_i^2$, the Newton polytope of any g_i must be contained in the rectangle with vertices $(1, 1)$, $(1, 0)$, $(0, 1)$, and $(0, 0)$.



This means that the only monomials that can occur in the g_i's are XY, X, Y, and 1. Let $M = (XY, X, Y, 1)$ and $V = (v_{i,j})$ be a symmetric 4×4 matrix, the the linear system arising from (12) is

$$\begin{aligned}
&v_{1,1} = 1,\ 2v_{1,2} = 0,\ 2v_{1,3} = 0,\ 2v_{1,4} + 2v_{2,3} = 0\\
&v_{2,2} = 1,\ 2v_{2,4} = 0\\
&v_{3,3} = 1,\ 2v_{3,4} = 0\\
&v_{4,4} = 1
\end{aligned}$$

There is one parameter in the solution and the general Gram matrix for f is

$$V(r) := \begin{bmatrix} 1 & 0 & 0 & r \\ 0 & 1 & -r & 0 \\ 0 & -r & 1 & 0 \\ 0 & -r & 1 & 0 \\ r & 0 & 0 & 1 \end{bmatrix}$$

The eigenvalues of $V(r)$ are $1 \pm r$, hence $V(r)$ is psd if and only if $-1 \le r \le 1$. It is easy to see that $r = 0$ corresponds to the original sos representation $f = (XY)^2 + X^2 + Y^2 + 1^2$. Furthermore, $V(r)$ has rank 2 is $r = \pm$ and rank 4 for all other values of r. The sos representation corresponding to $V(r)$ is

$$f = (XY + r)^2 + (X - rY)^2 + (\sqrt{1 - r^2}\,Y)^2 + (\sqrt{1 - r^2})^2.$$

The two sos representations with two squares are $f = (XY \pm 1)^2 + (X \mp Y)^2$.

References

[1] Emil Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, Abh. Math. Sem. Univ. Hamburg **5** (1927), no. 1, 100–115.

[2] M. D. Choi, T. Y. Lam, and B. Reznick, *Sums of squares of real polynomials*, K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992), Proc. Sympos. Pure Math., vol. 58, Amer. Math. Soc., Providence, RI, 1995, pp. 103–126.

[3] David Handelman, *Representing polynomials by positive linear functions on compact convex polyhedra*, Pacific J. Math. **132** (1988), no. 1, 35–62.

[4] David Hilbert, *Ueber die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. **32** (1888), no. 3, 342–350.

[5] T. S. Motzkin, *The arithmetic-geometric inequality*, Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965), Academic Press, New York, 1967, pp. 205–224.

[6] G. Pólya, *Über gewisse notwendige Determinantenkriterien für die Fortsetzbarkeit einer Potenzreihe*, Math. Ann. **99** (1928), no. 1, 687–706 (In Collected Papers 2 (1974), MIT Press, 309–313).

[7] Bruce Reznick, *Extremal PSD forms with few terms*, Duke Math. J. **45** (1978), no. 2, 363–374.

[8] Konrad Schmüdgen, *The K-moment problem for compact semi-algebraic sets*, Math. Ann. **289** (1991), no. 2, 203–206.

[9] Gilbert Stengle, *A nullstellensatz and a positivstellensatz in semialgebraic geometry*, Math. Ann. **207** (1974), 87–97.