

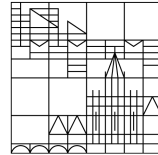
Sum of Squares Representations of Convex Quarternary Quartics and Ternary Sextics

Master Thesis

submitted by
Moritz Schick

at the

Universität
Konstanz



Mathematisch-Naturwissenschaftliche Sektion
Fachbereich Mathematik und Statistik

Supervised by: **Prof. Dr. Salma Kuhlmann**

Konstanz, 2020

Abstract

In 1888 Hilbert managed to prove that the set $P_{n,2d}$ of positive semidefinite (PSD) forms and the set $\Sigma_{n,2d}$ of sum of square (SOS) forms in n variables of degree $2d$ coincide if and only if $n = 2$ or $d = 1$ or $(n, 2d) = (3, 4)$ holds (see [16]). Further, the set $C_{n,2d}$ of convex forms is contained in the set of PSD forms, since convex forms have a global minimum at the origin, where they vanish. For this reason, Parrillo asked in 2007, whether every convex form is SOS. Two years later, Blekherman managed to show that this is not the case, if the number n of variables is sufficiently large (see [3]). However, no explicit example of a convex but not SOS form is known until today. Due to Hilbert's characterization of PSD and SOS forms, the smallest $n, 2d$, where one could expect convex forms that are not SOS are $(n, 2d) = (3, 6)$ or $(n, 2d) = (4, 4)$. In this thesis, we follow the proof of El Khadir in [12] from 2019, where he managed to show that for $(n, 2d) = (4, 4)$ indeed every convex quaternary quartic is SOS. The key steps of his proof are the so-called Generalized Cauchy–Schwarz Inequalities for convex forms as well as a characterization of the SOS cone $\Sigma_{4,4}$ inside the PSD cone $P_{4,4}$, which is based on Blekherman's work in [4]. Further, we see that for $(n, 2d) = (3, 6)$ the same result holds as well, if a conjecture formulated in [4] turns out to be true. This conjecture is related to the Cayley–Bacharach relations from [11].

CONTENTS

1. Introduction	2
2. Problem Formulation	4
3. Preliminaries	9
3.1. Euler's Identity and Positivity of Convex Forms	9
3.2. Tensors	16
3.3. Hilbert Space Representations of Linear Functionals on $H_{n,2d}$	24
3.4. Convex Cones and their Duals	30
4. Generalized Cauchy–Schwarz Inequalities	37
4.1. Formulation and Proof of the Generalized Cauchy–Schwarz Inequalities	38
4.2. Value of the Optimal Constants of the Generalized Cauchy–Schwarz Inequalities	47
5. Characterization of the SOS Cone $\Sigma_{n,2d}$ inside the PSD Cone $P_{n,2d}$	61
5.1. The Projective Space and Chasles' Version of the Cayley Bacharach Theorem	62
5.2. A more general Version of the Cayley–Bacharach Theorem	69
5.3. Structure of Extreme Rays in $\Sigma_{4,4}^*$ and $\Sigma_{3,6}^*$	73
5.4. Sum of Squares Certificates for elements of $P_{4,4}$ and $P_{3,6}$	79
6. Convex Quaternary Quartics are SOS	87
7. Further Extensions and Outlook	90
7.1. Possible Extension to the Case of Ternary Sextics	90
7.2. Outlook	91
References	91
Index	93

1. INTRODUCTION

In 1888, Hilbert proved that the sets $P_{n,2d}$ of positive semidefinite (PSD) real forms, i.e. homogeneous polynomials that are globally nonnegative, and $\Sigma_{n,2d}$ of real forms that can be written as a finite sum of squares (SOS) of real forms coincide if and only if $n = 2, d = 1$ or $(n, 2d) = (3, 4)$ (see [16]). This now famous theorem was inspired by Minkowski's defense of his doctoral dissertation in 1885 at the University of Königsberg. At that time, Minkowski conjectured that there are cases in which polynomials exist that are PSD but not SOS. While Hilbert was an official opponent in the defence, he later said that Minkowski indeed convinced him about that statement, which led him to his proof in 1888.

In 2007, Parrilo was interested in the relation between the set $C_{n,2d}$ of convex forms in n variables of degree $2d$ and the sets $P_{n,2d}$ and $\Sigma_{n,2d}$. Since it can easily be seen that every convex form is PSD, Parrilo asked whether every convex form is SOS. This is interesting from an optimizational point of view. To see this, consider for example the following problem, which is formulated in a non-homogeneous setting. Finding the unconstrained global minimum of a polynomial p is NP-hard (see [27]). If we know that the polynomial p is convex, every local minimum is a global minimum. Hence, a global minimum of a convex polynomial can be found by using even simple descent methods (see e.g. [5] or [35]). Those methods rely completely on the convexity of p and ignore its algebraic properties as a polynomial. On the other hand, we could also minimize p using SOS relaxations as in [26]. In [12] it can be seen that if every convex polynomial is SOS, we could find the global minimum of p by using semidefinite programming, for which efficient implementations are available (see e.g. [6]). In contrary to general descent methods, semidefinite programming highly relies on the algebraic structure of a given problem. In general, SOS relaxations have applications in many different areas of mathematics such as control theory, quantum computation, polynomial games or combinatorial optimization, as remarked in [2]. For this reason, it is interesting to know whether every convex polynomial or - in the homogeneous setting - every convex form is SOS. Furthermore, it is interesting from a pure mathematical point of view to see how Hilbert's characterization of SOS and PSD forms extends if we consider the additional property of convexity.

In 2009, Blekherman showed by using volume arguments that for a large number n of variables, there must be forms of degree as low as 4 which are convex but not SOS. However, no explicit example is known until today.

Due to Hilbert's characterization in 1888, the simplest cases, where convex forms that are not SOS could possibly exist, are for $(n, 2d) = (4, 4)$ or $(n, 2d) = (3, 6)$. In this thesis, we follow the proof of El Khadir, who showed in 2019 that at least in the first case when $(n, 2d) = (4, 4)$ of quaternary quartics, every convex form is SOS, i.e. $C_{4,4} \subseteq \Sigma_{4,4}$. Further, in the latter case the inclusion holds as well, if a conjecture of Blekherman (2012) related to the Cayley-Bacharach relation holds.

Therefore, in Section 2 we formulate again the problem discussed in this thesis but in a more formal way. This involves, for example, the definition of

a directional derivative, since we want to characterize convexity of a polynomial not only via the standard inequality on convex combinations but also by using its Hessian.

Afterwards, we introduce in Section 3 some important concepts from Algebra, Functional Analysis and Convex Geometry. First of all, in Section 3.1, we see two proofs of the inclusion $C_{n,2d} \subseteq P_{n,2d}$. While the first one involves just some basic ideas of optimization, the second one uses the Euler identity which gives a relation between a polynomial and its gradient. Further, we give some examples of SOS polynomials that are not convex, showing that $\Sigma_{n,2d} \not\subseteq C_{n,2d}$. As our own contribution, we present a MATLAB program which can be used to verify that some arbitrary form is not convex. Using this code, we confirm that the famous Motzkin and Robinson forms are not convex.

Further, in Section 3.2, we introduce tensors and give some of their equivalent characterizations. We also prove that there is a bijection between the set $H_{n,k}$ of homogenous polynomials in n variables of degree k and the set $S^k(\mathbb{R}^n)$ of real, symmetric tensors on \mathbb{R}^n of order k . This allows us to associate a form $p \in H_{n,k}$ with a tensor $T_p \in S^k(\mathbb{R}^n)$. From this bijection we can further define a biform $Q_p: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ corresponding to $p \in H_{n,k}$. The objects T_p and Q_p will be essential throughout the work.

Another important idea will be using the Hilbert space structure of $H_{n,2d}$. Therefore, we introduce in Section 3.3 some basic concepts of Functional Analysis involving the dual space and the Riesz Representation Theorem. This allows us to associate linear functionals on $H_{n,2d}$ with elements of $H_{n,2d}$ itself.

In Section 3.4, we cover some notions of Convex Geometry, involving, for example, closed convex cones and their extreme points, directions and rays. We further introduce the dual C^* of a cone C . Together with the Hilbert space theory of the previous section, this allows us to get a convenient representation of the dual $C_{n,2d}^*$, which plays an important role in following section.

Section 4 deals with statements and proofs of two Generalized Cauchy–Schwarz Inequalities on the set $C_{n,2d}$ of convex forms. Among the form $p \in C_{n,2d}$ itself, those inequalities involve the biform Q_p from section 3.2 and some degree dependent constants $A_d, B_d > 0$. The proof of the inequalities uses several simplifications and some techniques from Functional Analysis. We further investigate the optimal constants A_d^* and B_d^* of the inequalities. This leads to some optimization problems, for which we can obtain informations about the values of A_d^* and B_d^* by using both the primal and dual formulation of those problems. It turns out that although the exact value of the constants B_d^* are known for all d , the value A_d^* is only known for $d = 1, 2, 3, 4$.

In Section 5, we aim to characterize the SOS cone $\Sigma_{n,2d}$ inside the PSD cone $P_{n,2d}$. This follows the techniques of Blekherman from 2012. The essential tool in his paper was the Cayley–Bacharach relation, which we therefore want to study in more detail. For this reason, we introduce in Section 5.1 some basic notions from Algebraic Geometry, including e.g. the projective space as well as affine and projective varieties. We further see

one of the first versions of the Cayley–Bacharach relation due to Chasles, who published it in 1865.

In Section 5.2 we see a more general version of the Cayley–Bacharach relation published in 1996 by Eisenbud, Green and Harris. We further see how Blekherman uses that version for his purposes. To this end, we have to introduce some further concepts of Algebraic Geometry including vanishing ideals and the Hilbert function of a projective variety.

Afterwards, we see in Section 5.3 how Blekherman obtains a characterization of the extreme rays of the dual cones $\Sigma_{4,4}^*$ and $\Sigma_{3,6}^*$ by invoking his version of the Cayley–Bacharach relation. Therefore, we use some of the results of Blekherman without proof, while we want to focus on further simplifications of the structure of the extreme rays. This is continued in Section 5.4, where we use techniques from optimization involving the Karush Kuhn Tucker (KKT) conditions to obtain a representation of the extreme rays as in El Khadir’s paper. This way, we get a certificate on the PSD forms in $P_{4,4}$ and $P_{3,6}$ to be SOS.

Finally, we combine in Section 6 the results of Sections 4 and 5 to prove our main result that $C_{4,4} \subseteq \Sigma_{4,4}$ holds. By taking an arbitrary convex form in $C_{4,4}$, we want to show that it is SOS by using the characterization of the SOS cone $\Sigma_{4,4}$ inside the PSD cone $P_{4,4}$ from Section 5. Therefore, we have to show that two certain inequalities coming from the structure of extreme rays in $\Sigma_{4,4}$ hold. It turns out that those inequalities can indeed be deduced from the Generalized Cauchy–Schwarz Inequalities from Section 4. At that point, it is important to know that the optimal constants of the Generalized Cauchy–Schwarz Inequalities for $d = 2$ are indeed $A_d^* = B_d^* = 1$. If a conjecture of Blekherman holds, the same techniques can be applied to show that $C_{3,6} \subseteq \Sigma_{3,6}$ holds as well. The necessity of that conjecture to hold comes from the fact that for $d = 3$ the optimal constant B_d^* is strictly larger than 1, i.e. $B_3^* > 1$.

In Section 7 we give a collection of possible directions of future work related to the topics covered in this thesis. In particular, we see how the main result can be extended to the case of ternary sextics, if a conjecture formulated by Blekherman turns out to be true.

2. PROBLEM FORMULATION

In this section, we formally describe the problem that will be discussed throughout the thesis. Therefore, we introduce homogeneous polynomials (forms) and formalize the notions of positivity, convexity and sum-of-squares. The convexity of a homogeneous polynomial is stated in terms of its Hessian. For this reason, we also give the definition of the directional derivative of a function and introduce some notations including the gradient and Hessian of a multivariate function. Finally, at the end of this section one can find a formulation of the underlying problem of this thesis.

Convention 2.1. We will always write $\mathbb{N} = \{1, 2, 3, \dots\}$ for the strictly positive and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ for the nonnegative integers.

Throughout the thesis, we write $\mathbb{K}[\mathbf{x}] = \mathbb{K}[\mathbf{x}_1, \dots, \mathbf{x}_n]$ for the polynomial ring over \mathbb{K} in $n \in \mathbb{N}$ variables, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Note that we usually

only work with real polynomials, i.e. $\mathbb{K} = \mathbb{R}$. We also write $\mathbf{x}^\alpha := \mathbf{x}_1^{\alpha_1} \cdots \mathbf{x}_n^{\alpha_n}$ and $|\alpha| := \alpha_1 + \dots + \alpha_n$ for $\alpha \in \mathbb{N}_0^n$. Further, we denote by x^\top or A^\top the transpose of a vector $x \in \mathbb{R}^n$ or matrix $A \in \mathbb{R}^{n \times m}$ ($n, m \in \mathbb{N}$), respectively.

Definition 2.2. For $n, k \in \mathbb{N}$,

$$H_{n,k} := \left\{ p \in \mathbb{R}[\mathbf{x}] : p = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k} p_\alpha \mathbf{x}^\alpha \text{ for some } p_\alpha \in \mathbb{R} \right\}$$

is the set of all **homogeneous** real polynomials (**forms**) in n variables of degree k .

Further,

$$P_{n,k} := \{p \in H_{n,k} : p \geq 0 \text{ on } \mathbb{R}^n\}$$

is the set of all **nonnegative** homogeneous polynomials in n variables of degree k . A polynomial $p \in P_{n,k}$ is called **PSD**, which stands for **positive semidefinite**. Note that a PSD form has always even degree, i.e. $P_{n,k} \neq \emptyset$ if and only if $k = 2d$ for some $d \in \mathbb{N}$.

In this case,

$$\Sigma_{n,2d} := \left\{ p \in H_{n,2d} : p = \sum_{j=1}^m \sigma_j^2 \text{ for some } \sigma_j \in H_{n,d} \text{ and } m \in \mathbb{N}_0 \right\}$$

is the set of all homogeneous polynomials in n variables of degree $k = 2d$ which can be represented as **sum-of-squares**. Therefore, a polynomial $p \in \Sigma_{n,k}$ is called **SOS**.

Convention 2.3. Throughout this work, we will always have $n, k, d \in \mathbb{N}$.

Furthermore, a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ will be called multivariate and a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, where $m \in \mathbb{N}$ will be called multidimensional.

Remark 2.4. The set $H_{n,k}$ is an \mathbb{R} -vector space (with the usual addition and scalar multiplication) of dimension $\binom{d+n-1}{n-1}$ [21, Lecture 06]. Note that $0 \in H_{n,k}$ as it is the form with all coefficients being zero.

We are also interested in convex polynomials. Therefore, consider the following classical definition of convexity of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Definition 2.5. A multivariate function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be **convex** if the following holds

$$\forall x, y \in \mathbb{R}^n \forall \lambda \in [0, 1] : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

A polynomial $p \in \mathbb{R}[\mathbf{x}]$ is **convex** if and only if the polynomial function $\mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto p(x)$ is convex.

Convention 2.6. In this work, we do not distinguish between a polynomial $p \in \mathbb{R}[\mathbf{x}]$ and its polynomial function $p : \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto p(x)$. It should always be clear from the context which of the two notions is meant.

However, as in [12], we will work with another equivalent characterization of convexity of a polynomial, which involves the Hessian matrix. For this purpose, we need the definition of a directional derivative as in [9, Section 11]. The directional derivative will again be used in Section 3.3, when we investigate the Hilbert space structure of the vector space $H_{n,2d}$.

Definition 2.7. Let $x, u \in \mathbb{R}^n$, $f: \mathbb{R}^n \rightarrow \mathbb{R}$. If the differential quotient

$$\partial_u f(x) = \lim_{h \rightarrow 0} \frac{f(x + hu) - f(x)}{h}.$$

exists, it is called the **directional derivative** of f at x in the direction u .

Moreover, we write $\partial_{\mathbf{x}_i} f(x) := \partial_{e_i} f(x)$ for the **partial derivative** of f at x with respect to \mathbf{x}_i , where $e_i := (\delta_{i1}, \dots, \delta_{in})^\top \in \mathbb{R}^n$ is the i -th **unit vector**.

The function f is called **partially differentiable**, if all the partial derivatives of f exists at every $x \in \mathbb{R}^n$. In this case, we define the **gradient** of f at $x \in \mathbb{R}^n$ as $\nabla f(x) = (\partial_{\mathbf{x}_1} f(x), \dots, \partial_{\mathbf{x}_n} f(x))^\top \in \mathbb{R}^n$, where $\nabla = (\partial_{\mathbf{x}_1}, \dots, \partial_{\mathbf{x}_n})^\top$ is a differential operator.

The partial derivative can also be extended for multidimensional functions.

Definition 2.8. Let $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a multidimensional function, $m \in \mathbb{N}$. Let $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ ($i = 1, \dots, m$) be such that $f(x) = (f_1(x), \dots, f_m(x))^\top$ ($x \in \mathbb{R}^n$). Then we say that f is **partially differentiable**, if all f_i are. Further, in this case we define the **Jacobian** of f at $x \in \mathbb{R}^n$ as

$$f'(x) = ((\partial_{\mathbf{x}_j} f_i)(x))_{i=1, \dots, m; j=1, \dots, n} \in \mathbb{R}^{m \times n}.$$

Remark 2.9. From [9, Bemerkung 11.7], we know that the i -th partial derivative of a multivariate function $f(\mathbf{x}_1, \dots, \mathbf{x}_n)$ is just the same as differentiating with respect to the i -th variable \mathbf{x}_i . In particular, the partial derivative of a polynomial has its usual algebraic form. More precisely, if $\prod_{i=1}^n \mathbf{x}_i^{\alpha_i} \in \mathbb{R}[\mathbf{x}]$ is a monomial, we have

$$\partial_{\mathbf{x}_j} \left(\prod_{i=1}^n \mathbf{x}_i^{\alpha_i} \right) = \begin{cases} \left(\prod_{i=1, i \neq j}^n \mathbf{x}_i^{\alpha_i} \right) \cdot \alpha_j \mathbf{x}_j^{\alpha_j - 1} & , \alpha_j > 0 \\ 0 & , \text{else} \end{cases}.$$

In fact, there is an easy formula for the directional derivative of a partially differentiable function and hence in particular for polynomials. The following proposition will again be used in Section 3.3.

Proposition 2.10. Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be a partially differentiable function and let $u, x \in \mathbb{R}^n$ be arbitrary. Then

$$\partial_u f(x) = u^\top \nabla f(x) = \sum_{i=1}^n u_i \partial_{\mathbf{x}_i} f(x). \quad (1)$$

Proof. [9, Section 11] □

Remark 2.11. Proposition 2.10 yields that $\partial_u p = \sum_{i=1}^n u_i \partial_{\mathbf{x}_i} p \in \mathbb{R}[\mathbf{x}]$ is again a polynomial for all $p \in \mathbb{R}[\mathbf{x}]$ and $u \in \mathbb{R}^n$. In particular, we can interpret the directional derivative ∂_u as differential operator $\mathbb{R}[\mathbf{x}] \rightarrow \mathbb{R}[\mathbf{x}]$.

We know that we can differentiate a polynomial $p \in \mathbb{R}[\mathbf{x}]$ with respect to any variable an arbitrary number of times. In the following, we define the gradient, Hessian and Laplacian only for polynomials, since this allows us to avoid the definition of being differentiable multiple times for an arbitrary multidimensional function. Moreover, we only deal with polynomials in this work.

Definition 2.12. For $p(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$, $\nabla p(\mathbf{x}) = (\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n)^\top p(\mathbf{x}) \in (\mathbb{R}[\mathbf{x}])^n$ is the **gradient** of p . Further,

$$\nabla^2 p(\mathbf{x}) := \nabla \nabla^\top p(\mathbf{x}) = (\partial \mathbf{x}_i \partial \mathbf{x}_j p(\mathbf{x}))_{i,j=1,\dots,n} \in (\mathbb{R}[\mathbf{x}])^{n \times n}$$

is the **Hessian** and $\Delta p(\mathbf{x}) := \sum_{i=1}^n \partial \mathbf{x}_i^2 p(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ the **Laplacian** of p .

Example 2.13. Consider the following quaternary quarty

$$p = p(\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{w}^4 + \mathbf{x}^2 \mathbf{y}^2 + \mathbf{y}^2 \mathbf{z}^2 + \mathbf{x}^2 \mathbf{z}^2 - 4\mathbf{x}\mathbf{y}\mathbf{z}\mathbf{w} \in H_{4,4}$$

Then the gradient of p is given by

$$\nabla p = \begin{pmatrix} 4\mathbf{w}^3 - 4\mathbf{x}\mathbf{y}\mathbf{z} \\ 2\mathbf{x}\mathbf{y}^2 + 2\mathbf{x}\mathbf{z}^2 - 4\mathbf{w}\mathbf{y}\mathbf{z} \\ 2\mathbf{x}^2 \mathbf{y} + 2\mathbf{y}\mathbf{z}^2 - 4\mathbf{w}\mathbf{x}\mathbf{z} \\ 2\mathbf{y}^2 \mathbf{z} + 2\mathbf{x}^2 \mathbf{z} - 4\mathbf{w}\mathbf{x}\mathbf{y} \end{pmatrix}.$$

Further, the Hessian of p is given by

$$\nabla^2 p = \begin{pmatrix} 12\mathbf{w}^2 & -4\mathbf{y}\mathbf{z} & -4\mathbf{x}\mathbf{z} & -4\mathbf{x}\mathbf{y} \\ -4\mathbf{y}\mathbf{z} & 2\mathbf{y}^2 + 2\mathbf{z}^2 & 4\mathbf{x}\mathbf{y} - 4\mathbf{w}\mathbf{z} & 4\mathbf{x}\mathbf{z} - 4\mathbf{w}\mathbf{y} \\ -4\mathbf{x}\mathbf{z} & 4\mathbf{x}\mathbf{y} - 4\mathbf{w}\mathbf{z} & 2\mathbf{x}^2 + 2\mathbf{y}^2 & 4\mathbf{y}\mathbf{z} - 4\mathbf{w}\mathbf{x} \\ -4\mathbf{x}\mathbf{y} & 4\mathbf{x}\mathbf{z} - 4\mathbf{w}\mathbf{y} & 4\mathbf{y}\mathbf{z} - 4\mathbf{w}\mathbf{x} & 2\mathbf{y}^2 + 2\mathbf{x}^2 \end{pmatrix}.$$

Note that the Laplacian of p is the trace of the Hessian of p , which is in this case

$$\begin{aligned} \Delta p &= 12\mathbf{w}^2 + 2\mathbf{y}^2 + 2\mathbf{z}^2 + 2\mathbf{x}^2 + 2\mathbf{y}^2 + 2\mathbf{y}^2 + 2\mathbf{x}^2 \\ &= 12\mathbf{w}^2 + 4\mathbf{x}^2 + 4\mathbf{y}^2 + 4\mathbf{z}^2. \end{aligned}$$

We can also interpret the polynomial function of p as multidimensional function $p : \mathbb{R}^n \rightarrow \mathbb{R}^m$, where $n = 4$ is the number of variables and $m = 1$. In this case, the Jacobian of p is just the transpose of the gradient, i.e. $p' = (\nabla p)^\top$.

Consider the following standard notation from Linear Algebra.

Notation 2.14. For $A \in \mathbb{R}^{n \times n}$, we write $A \succ 0$ and $A \succeq 0$ if the matrix A is positive definite or positive semidefinite, respectively.

Now, we have the following result which gives an equivalent characterization of the classical convexity of polynomial.

Proposition 2.15. *Let $p \in \mathbb{R}[\mathbf{x}]$. Then p is convex if and only if the Hessian $\nabla^2 p(x)$ is positive semidefinite at every point, i.e. $\nabla^2 p(x) \succeq 0$ for any $x \in \mathbb{R}^n$.*

Proof. Follows directly from [34, Satz 6.4] since polynomials are twice differentiable in the sense of [9, Section 11.3]. \square

This motivates the following definition.

Definition 2.16. The set of all **convex** homogeneous polynomials in n variables of degree k is defined as

$$C_{n,k} = \{p \in H_{n,k} : \nabla^2 p(x) \succeq 0 \text{ for all } x \in \mathbb{R}^n\}.$$

Example 2.17. (a) It holds $p(\mathbf{x}) = \sum_{i=1}^n \mathbf{x}_i^d \in C_{n,2d}$. This can be seen very easily since the Hessian of p is given by

$$\nabla^2 p(\mathbf{x}) = \begin{pmatrix} \mathbf{x}_1^{2d-2} & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{x}_2^{2d-2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathbf{x}_{n-1}^{2d-2} & 0 \\ 0 & \cdots & \cdots & 0 & \mathbf{x}_n^{2d-2} \end{pmatrix}.$$

Hence, the Hessian $\nabla^2 p(x)$ at some $x \in \mathbb{R}^n$ is a diagonal matrix with nonnegative entries, which is clearly positive semidefinite.

(b) More examples of convex forms can be found e.g. in [2]. There the authors define - in addition to convexity and being SOS - the property of being SOS-convex (see [2, Def. 2.4]). Further, they investigate the relation between those three properties.

Now we come to the actual problem that will be investigated throughout this thesis. It is clear that every polynomial which can be represented as sum-of-squares (SOS) is also nonnegative, i.e. $\Sigma_{n,k} \subseteq P_{n,k}$ for all $n, k = 2d \in \mathbb{N}$. The 17th Hilbert problem, which was one of 23 famous problems formulated by David Hilbert, is about whether the converse is also true. The question is, can every nonnegative real polynomial of fixed number n of variables and fixed degree $k = 2d$ be represented as a (finite) number of squares of polynomials? The degree is supposed to be even since the square of a polynomial of course has even degree. Indeed, the following theorem gives a full characterization of the cases in which $P_{n,2d} = \Sigma_{n,2d}$ holds. It was first proven in [16].

Theorem 2.18 (Hilbert (1888)). *We have $P_{n,2d} = \Sigma_{n,2d}$ if and only if at least one of the following holds*

- (a) $n = 2$,
- (b) $2d = 2$,
- (c) $(n, 2d) = (3, 4)$.

Although Theorem 2.18 gives a full characterization of whether the identity $P_{n,2d} = \Sigma_{n,2d}$ holds, it took almost eighty years until the first example of a PSD polynomial that is not SOS was discovered by Motzkin in 1967.

Remark 2.19. The following forms are some of the first and best known examples of polynomials that are PSD but not SOS (for proofs and references to the original papers see e.g. [22, Lecture 23] and [25, Section 1.2]).

(a) Motzkin form:

$$\mathbf{z}^6 + \mathbf{x}^4 \mathbf{y}^2 + \mathbf{x}^2 \mathbf{y}^4 - 3\mathbf{x}^2 \mathbf{y}^2 \mathbf{z}^2 \in P_{3,6} \setminus \Sigma_{3,6}.$$

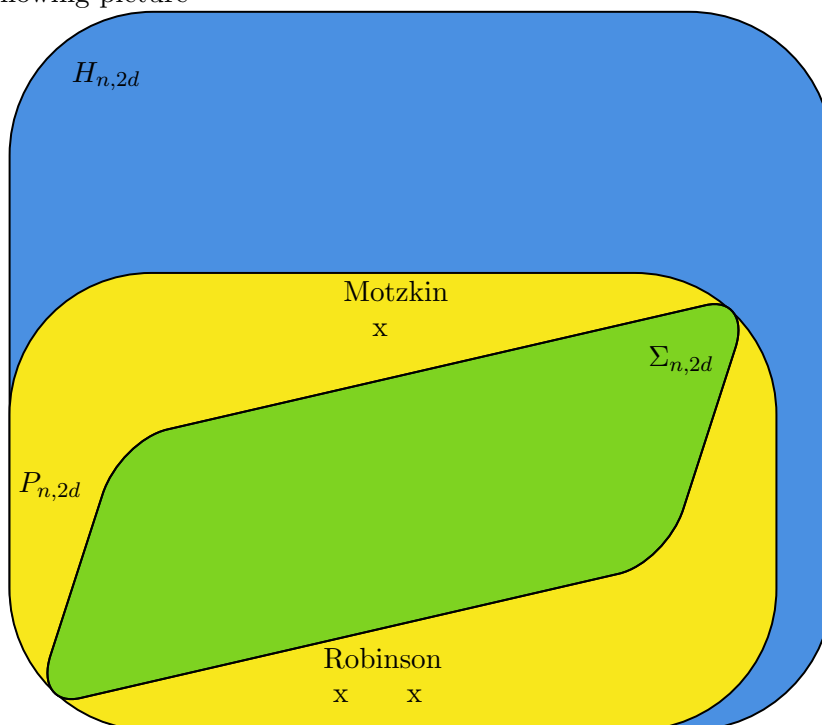
(b) Robinson form:

$$\begin{aligned} \mathbf{x}^6 + \mathbf{y}^6 + \mathbf{z}^6 - (\mathbf{x}^4 \mathbf{y}^2 + \mathbf{x}^4 \mathbf{z}^2 + \mathbf{y}^4 \mathbf{x}^2 + \mathbf{y}^4 \mathbf{z}^2 + \mathbf{z}^4 \mathbf{x}^2 + \mathbf{z}^4 \mathbf{y}^2) \\ + 3\mathbf{x}^2 \mathbf{y}^2 \mathbf{z}^2 \in P_{3,6} \setminus \Sigma_{3,6}. \end{aligned}$$

(c) Robinson form II:

$$\mathbf{w}^4 + \mathbf{x}^2 \mathbf{y}^2 + \mathbf{y}^2 \mathbf{z}^2 + \mathbf{x}^2 \mathbf{z}^2 - 4\mathbf{x}\mathbf{y}\mathbf{z}\mathbf{w} \in P_{4,4} \setminus \Sigma_{4,4}.$$

Hence, for $(n, 2d) \in \mathbb{N}_{\geq 3} \times \mathbb{N}_{\geq 4}$ with $(n, 2d) \neq (3, 4)$, one can think of the following picture



One can now ask if and how the characterization of Hilbert changes if we consider also the set of convex polynomials. Hence, we fix some $n, k = 2d \in \mathbb{N}$ and we want to investigate how the set of convex forms $C_{n,2d}$ is connected to the set $P_{n,2d}$ of PSD forms and the set $\Sigma_{n,2d}$ of SOS forms. We are also interested in a similar characterization as Hilbert's one for convex forms.

3. PRELIMINARIES

The goal of this section is to introduce some concepts that will help us to investigate the problem formulated in Section 2. In Section 3.1, we see our first result about the problem described in Section 2, which is that every convex polynomial is already PSD. Afterwards, we introduce in Section 3.2 the concept of tensors and show an important relation between homogeneous polynomials and symmetric tensors. In Section 3.3, we introduce some basic concepts of Functional Analysis that allow us to deduce that $H_{n,k}$ is a Hilbert space and therefore has some special properties involving its dual. Further, in Section 3.4 we introduce some notions from Convex Geometry to understand the conic structure of the sets $P_{n,2d}, \Sigma_{n,2d}$ and $C_{n,2d}$.

3.1. Euler's Identity and Positivity of Convex Forms. It can be seen that every convex polynomial is already nonnegative, i.e. $C_{n,k} \subseteq P_{n,k}$.

The proof of the following proposition is an alternative way to the proof in [12], which we will see later in Corollary 3.4.

Proposition 3.1. *If $k > 1$, we have $C_{n,k} \subseteq P_{n,k}$.*

Proof. Let $p \in C_{n,k}$ be arbitrary. Obviously, we have $p(0) = 0$. Hence, it would suffice to show that p has a global minimum at the origin.

Assume for a contradiction that there exists a $x \in \mathbb{R}^n \setminus \{0\}$ such that $p(x) < p(0) = 0$. Let $\lambda \in [0, 1]$ be arbitrary. Then by convexity of p

$$p(\lambda x + (1 - \lambda)0) \leq \lambda p(x) + (1 - \lambda)p(0) = \lambda p(x).$$

On the other hand, since p is homogeneous of degree k , we know that $p(\lambda x) = \lambda^k p(x)$. This yields $\lambda^k p(x) \leq \lambda p(x)$. This is a contradiction for $\lambda \in (0, 1)$ (and $k > 1$), since $p(x) < 0$. Hence, p has a global minimum at the origin and is therefore nonnegative. \square

The same result can also be seen differently using Euler's identity as in [12, Section 2]. Therefore, consider the following theorem, which is also stated without proof in [12, Section 2].

Theorem 3.2 (Euler's Identity). *For $p \in H_{n,k}$, it holds $kp(\mathbf{x}) = \mathbf{x}^\top \nabla p(\mathbf{x})$.*

Proof. Let $p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k} p_\alpha \mathbf{x}^\alpha \in H_{n,k}$ be arbitrary. For $j \in \{1, \dots, n\}$, we have

$$(\nabla p(\mathbf{x}))_j = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k, \alpha_j > 0} \alpha_j p_\alpha \mathbf{x}^{\beta_\alpha}, \quad \text{where } (\beta_\alpha)_\ell = \begin{cases} \alpha_\ell - 1, & \ell = j \\ \alpha_\ell, & \text{else} \end{cases}.$$

Hence:

$$\begin{aligned} \mathbf{x}^\top \nabla p(\mathbf{x}) &= \sum_{j=1}^n \mathbf{x}_j \left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k, \alpha_j > 0} \alpha_j p_\alpha \mathbf{x}^{\beta_\alpha} \right) \\ &= \sum_{j=1}^n \left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k, \alpha_j > 0} \alpha_j p_\alpha \mathbf{x}^\alpha \right) \\ &= |\alpha| \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k} p_\alpha \mathbf{x}^\alpha = kp(\mathbf{x}). \end{aligned}$$

\square

Corollary 3.3. *For $p \in H_{n,k}$, we have $k(k-1)p(\mathbf{x}) = \mathbf{x}^\top \nabla^2 p(\mathbf{x}) \mathbf{x}$.*

Proof. Let $p \in H_{n,k}$ be arbitrary. By Euler's identity, we know that $kp(\mathbf{x}) = \mathbf{x}^\top \nabla p(\mathbf{x})$. Now let $j \in \{1, \dots, n\}$ be arbitrary and differentiate both sides with respect to the j -th component. This shows

$$\begin{aligned} k \nabla_j p(\mathbf{x}) &= \nabla_j p(\mathbf{x}) + \sum_{i=1}^n \mathbf{x}_i (\nabla^2 p(\mathbf{x}))_{ij} \\ \iff (k-1) \nabla_j p(\mathbf{x}) &= \sum_{i=1}^n \mathbf{x}_i (\nabla^2 p(\mathbf{x}))_{ij} \end{aligned}$$

Multiplying both sides with \mathbf{x}_j and summing over $j = 1, \dots, n$ shows

$$\begin{aligned} \sum_{j=1}^n (k-1) \nabla_j p(\mathbf{x}) \mathbf{x}_j &= \sum_{j=1}^n \sum_{i=1}^n \mathbf{x}_i (\nabla^2 p(\mathbf{x}))_{ij} \mathbf{x}_j \\ \iff (k-1) \mathbf{x}^\top \nabla p(\mathbf{x}) &= \mathbf{x}^\top \nabla^2 p(\mathbf{x}) \mathbf{x} \\ \text{Euler's identity} \iff & k(k-1)p(\mathbf{x}) = \mathbf{x}^\top \nabla^2 p(\mathbf{x}) \mathbf{x}. \end{aligned}$$

□

From this Corollary, we can deduce that $C_{n,k} \subseteq P_{n,k}$.

Corollary 3.4. *Every convex form is also nonnegative, i.e. $C_{n,k} \subseteq P_{n,k}$.*

Proof. By Corollary 3.3 we know that for every form $p \in H_{n,k}$, it holds $k(k-1)p(\mathbf{x}) = \mathbf{x}^\top \nabla^2 p(\mathbf{x}) \mathbf{x}$. If p is in addition convex, i.e. $p \in C_{n,k}$, then the right hand side evaluated at any $x \in \mathbb{R}^n$ is nonnegative, since the Hessian $\nabla^2 p(x)$ is positive semidefinite. This shows $k(k-1)p(x) \geq 0$ for any $x \in \mathbb{R}^n$ and hence that p is nonnegative. □

On the other hand, not every PSD form is convex. In the following, an easy example is given, showing that not even every SOS form is convex.

Example 3.5. (a) First, consider the easiest case where the number n of variables is two and consider $p(\mathbf{x}, \mathbf{y}) = \mathbf{x}^2 \mathbf{y}^2$. Then we have

$$\nabla p(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} 2\mathbf{x}\mathbf{y}^2 \\ 2\mathbf{x}^2\mathbf{y} \end{pmatrix}, \quad \nabla^2 p(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} 2\mathbf{y}^2 & 4\mathbf{x}\mathbf{y} \\ 4\mathbf{x}\mathbf{y} & 2\mathbf{x}^2 \end{pmatrix}.$$

In particular, it holds

$$\nabla^2 p(1, 1) = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix} \implies \det(\nabla^2 p(1, 1)) = 4 - 16 = -12 < 0,$$

which shows that $\nabla^2 p(1, 1)$ is not positive semidefinite. Hence, p is not convex.

(b) A similar result can be seen for arbitrary even degree $2d \in \mathbb{N}_{>2}$ by considering the form $p(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{2d-2} \mathbf{y}^2$. Here we have

$$\begin{aligned} \nabla p(\mathbf{x}, \mathbf{y}) &= \begin{pmatrix} (2d-2)\mathbf{x}^{2d-3}\mathbf{y}^2 \\ 2\mathbf{x}^{2d-2}\mathbf{y} \end{pmatrix}, \\ \nabla^2 p(\mathbf{x}, \mathbf{y}) &= \begin{pmatrix} (2d-2)(2d-3)\mathbf{x}^{2d-4}\mathbf{y}^2 & 2(2d-2)\mathbf{x}^{2d-3}\mathbf{y} \\ 2(2d-2)\mathbf{x}^{2d-3}\mathbf{y} & 2\mathbf{x}^{2d-2} \end{pmatrix}. \end{aligned}$$

This yields

$$\nabla^2 p(1, 1) = \begin{pmatrix} (2d-2)(2d-3) & 2(2d-2) \\ 2(2d-2) & 2 \end{pmatrix}$$

and hence $\det(\nabla^2 p(1, 1)) = 2(2d-2)(2d-3) - (2(2d-2))^2 < 0$. For this reason, we have $\Sigma_{n,2d} \not\subseteq C_{n,2d}$ and in particular $P_{n,2d} \not\subseteq C_{n,2d}$.

From Example 3.5, one can extract a general scheme on how to prove that a form $p \in H_{n,k}$ is not convex. If p is not convex, there has to be some $x \in \mathbb{R}^n$ such that the Hessian $\nabla^2 p(x)$ of p at x is not positive semidefinite. Equivalently, the Hessian $\nabla^2 p(x)$ has at least one negative eigenvalue. Hence, if p is not convex, evaluating the Hessian at random points and determining its eigenvalues will eventually give a point $x \in \mathbb{R}^n$ where the Hessian is not positive semidefinite. This idea is implemented in Code 3.6, where we use that MATLAB command `rand()` to create a random point $x \in [0, 1]^n$. Note that `rand()` generates points from the uniform distribution, which is important since other probability distributions might accumulate at some portion of $[0, 1]^n$. If p is a form which is not convex, evaluating the Hessian

at uniformly distributed random points in $[0, 1]^n$ and determining its eigenvalues will eventually give a point $x \in \mathbb{R}^n$ where the Hessian is not positive semidefinite. Note that focusing on the unit hypercube is sufficient because of the homogeneity of p .

Code 3.6.

```

1 function [test, x, eigenVal, iter]=
    ConvexityNecessityTest(p, var, maxiter)
2 %%
3 % CONVEXITY NECESSITY TEST
4 % This MATLAB function gives a necessary condition for
5 % the input polynomial p to be convex. The polynomial
6 % p is convex if and only if its Hessian is positive
7 % semidefinite at every point. Therefore, we can try
8 % and determine its Hessian at random points. This
9 % will be done up to maxiter times. If before the
10 % number maxiter of iterations is reached a counter-
11 % example is found where the Hessian is not positive
12 % semidefinite, then we know that the polynomial is
13 % not positive semidefinite. However, this just gives
14 % a necessary condition. Of course, it could be that
15 % the maximum number of iterations is reached without
16 % finding a counterexample, but there might still
17 % exist points, where the polynomial is not positive
18 % semidefinite.
19 %% Input
20 % p          polynomial as symbolic function
21 % var        vector of variables in p
22 % maxiter    maximum number of iterations
23 %% Output
24 % test       boolean. This will be FALSE, if p fails
25 %            the test, which means that p can not be
26 %            convex. Otherwise TRUE and we do not
27 %            know whether p is convex or not
28 % x          If test==false: vector x at which the
29 %            Hessian is not psd
30 %            else:          zero vector
31 % eigenval   If test==false: vector of eigenvalues
32 %            of the Hessian at x
33 %            else:          zero vector
34
35 numbVar=length(var); % Get the number of variables
36 test=true;
37 iter=0;
38 while iter<maxiter && test
39     x=rand(1,numbVar); % Create random vector

```

```

40     iter=iter+1;
41     % Gives the Hessian matrix of p as symbolic
         function
42     hess=hessian(p, var);
43     % Evaluate the Hessian at the random vector x
44     hessEval=subs(hess, var, x);
45     % Convert the evaluated Hessian to double
46     hessEval=double(hessEval);
47     % Get the eigenvalues of the Hessian at x
48     eigenVal=eig(hessEval);
49     % Sort the eigenvalues in ascending order
50     eigenVal=sort(eigenVal);
51     % Since the eigenvalues are sorted, it suffices to
52     % check whether the smallest one is negative
53     if eigenVal(1)<0
54         test=false;
55     end
56 end
57 if iter==maxiter && test
58     x=zeros(numVar,1);
59     eigenVal=zeros(numVar,1);
60 end

```

Now, Code 3.6 can be used to test some well known forms on convexity.

Example 3.7. In Remark 2.19, we have seen some famous homogeneous polynomials which are known to be PSD but not SOS. Those forms are recalled in the following.

(a) Motzkin form:

$$\mathbf{z}^6 + \mathbf{x}^4\mathbf{y}^2 + \mathbf{x}^2\mathbf{y}^4 - 3\mathbf{x}^2\mathbf{y}^2\mathbf{z}^2 \in P_{3,6} \setminus \Sigma_{3,6}.$$

(b) Robinson form I:

$$\mathbf{x}^6 + \mathbf{y}^6 + \mathbf{z}^6 - (\mathbf{x}^4\mathbf{y}^2 + \mathbf{x}^4\mathbf{z}^2 + \mathbf{y}^4\mathbf{x}^2 + \mathbf{y}^4\mathbf{z}^2 + \mathbf{z}^4\mathbf{x}^2 + \mathbf{z}^4\mathbf{y}^2) + 3\mathbf{x}^2\mathbf{y}^2\mathbf{z}^2 \in P_{3,6} \setminus \Sigma_{3,6}.$$

(c) Robinson form II:

$$\mathbf{w}^4 + \mathbf{x}^2\mathbf{y}^2 + \mathbf{y}^2\mathbf{z}^2 + \mathbf{x}^2\mathbf{z}^2 - 4\mathbf{xyzw} \in P_{4,4} \setminus \Sigma_{4,4}.$$

The following code exploits Code 3.6 to verify that the forms in (a)–(c) are not convex.

Code 3.8.

```

1 close all; clear all; clc;
2 % polynomial = 1;
3 % polynomial = 2;
4 polynomial = 3;
5 if polynomial==1

```

```

6     syms x y z ;
7     motzkin=z^6+x^4*y^2+x^2*y^4-3*x^2*y^2*z^2;
8     var=[x,y,z];
9     [test ,x ,eigenval , iter]=ConvexityNecessityTest (
        motzkin , var ,20) ;
10  elseif polynomial==2
11     syms x y z ;
12     robinson=x^6+y^6+z^6-(x^4*y^2+x^4*z^2+y^4*x^2+y^4*
        z^2+z^4*x^2+z^4*y^2)+3*x^2*y^2*z^2;
13     var=[x,y,z];
14     [test ,x ,eigenval , iter]=ConvexityNecessityTest (
        robinson , var ,20) ;
15  elseif polynomial==3
16     syms w x y z ;
17     robinson=w^4+x^2*y^2+y^2*z^2+x^2*z^2-4*x*y*z*w;
18     var=[w,x,y,z];
19     [test ,x ,eigenval , iter]=ConvexityNecessityTest (
        robinson , var ,20) ;
20  end

```

In the following, for each form $p \in H_{n,k}$ in Example 3.7 we give a point $x \in \mathbb{R}^n$, where the Hessian $\nabla^2 p(x)$ is not positive semidefinite together with eigenvalues of the Hessian $\nabla^2 p(x)$ and the number of random points tested in Code 3.6 until that particular point is found.

Form p	point $x \in \mathbb{R}^n$	eigenvalues of $\nabla^2 p(x)$	# points tested
(a)	[0.8147, 0.9058, 0.1270]	-3.3569 -1.0443 16.2957	1
(b)	[0.9134, 0.6324, 0.0975]	-2.3374 0.1958 18.1808	1
(c)	[0.2785, 0.5469, 0.9575, 0.9649]	-2.0291 -0.6403 3.2654 8.9224	1

Remark 3.9. When using code 3.6 to verify that some form $p \in H_{n,k}$ is not convex, a few remarks should be made. First of all, this is a numerical proof and not a formal mathematical one. Since we are using some computer-based algorithm, numerical inaccuracies may occur. This could be a problem, when we obtain a point $x \in \mathbb{R}^n$ in which the negative eigenvalues of the Hessian $\nabla^2 p(x)$ are close to zero. Due to possible rounding errors, we should then not blindly trust the result. However, in Example 3.7, there is always at least one eigenvalue which is significantly smaller than zero. This is why in that case, numerical inaccuracies do not cause any problem.

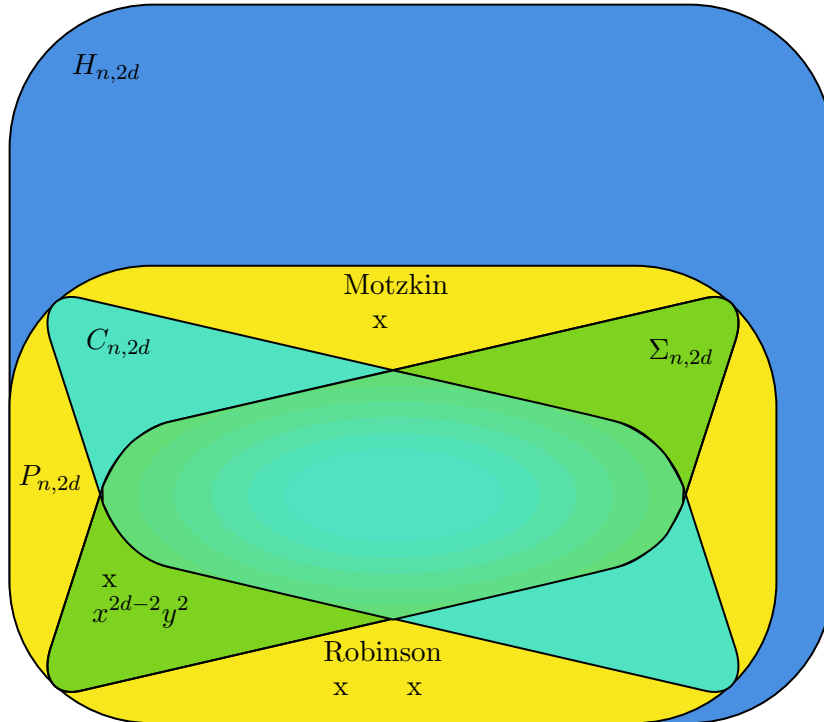
It should also be noted that we could, of course, also use Code 3.6 to try and verify that a non-homogeneous polynomial is convex. However, we should not assume that it works as well as in the homogeneous case. Indeed, in Example 3.7, we always found a point x , where the Hessian $\nabla^2 p(x)$ has a negative eigenvalue after one iteration. The reason is that in Code 3.6, we take random points on the n -dimensional interval $[0, 1]^n \subseteq \mathbb{R}^n$. This works well for forms since by homogeneity, their values on $[0, 1]^n$ already determine the form's values on the whole space \mathbb{R}^n . However, for a non-homogeneous polynomial p this is not true, and even if p was convex on $[0, 1]^n$, this would not allow to make a conclusion about the convexity on the whole space \mathbb{R}^n .

It is therefore reasonable to ask, whether convexity is invariant under homogenization. It turns out that this is not the case. To see this, consider e.g. $p(\mathbf{x}) = \mathbf{x}^2 - 4\mathbf{x} + 1$ and its homogenization (see [7, Chapter 8, Prop. 7])

$$p^h(\mathbf{x}, \mathbf{y}) = \mathbf{x}^2 - 4\mathbf{x}\mathbf{y} + \mathbf{y}^2 = (\mathbf{x} - \mathbf{y})^2 - 2\mathbf{x}\mathbf{y}.$$

Since $p''(\mathbf{x}) = 2 > 0$, clearly p is convex. However, p^h is not convex since it is not even nonnegative, e.g. $p^h(1, 1) = -2 < 0$. By [29, Prop. 4.4], we know that a polynomial p of degree d is convex if and only if the d -th root of its homogenization p^h is convex as well (under the assumption that $p^h(e^1) > 0$, where e^1 is the unit vector with 1 at the new variable).

Remark 3.10. As in Remark 2.19 we have the following picture of the set of PSD forms and the set of convex forms inside the vector space $H_{n,2d}$ of forms in n variables of degree $2d$ ($(n, 2d) \in \mathbb{N}_{\geq 3} \times \mathbb{N}_{\geq 4}$ and $(n, 2d) \neq (3, 4)$)



Whereas we were able to prove that $C_{n,2d} \subseteq P_{n,2d}$ holds even in multiple ways, it is much harder to investigate whether $C_{n,2d} \subseteq \Sigma_{n,2d}$ holds, i.e. whether every convex polynomial is already SOS.

3.2. Tensors. The goal of this subsection is to find a way of describing a homogeneous polynomial p uniquely by a tensor T_p and then to associate a binomial form Q_p to p via using the tensor T_p . To do so, we first have to introduce tensors and investigate their structure. The biform Q_p will be used in Section 4 to formulate generalized Cauchy–Schwarz inequalities for convex forms.

In the following, let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be arbitrary. We introduce tensors as elements of the outer product of finitely many finite dimensional \mathbb{K} -vector spaces. For more details consider e.g. [13, Section 6.4].

Definition 3.11. The **outer product** $\mathbb{K}^{n_1} \otimes \cdots \otimes \mathbb{K}^{n_k}$ of the vector spaces \mathbb{K}^{n_j} , where $n_j \in \mathbb{N}$ ($j = 1, \dots, k$), is defined as

$$\begin{aligned} & \mathbb{K}^{n_1} \otimes \cdots \otimes \mathbb{K}^{n_k} \\ & := \left\{ (x^{j_1 \cdots j_k})_{j_1=1, \dots, n_1; \dots; j_k=1, \dots, n_k} : x^{j_1 \cdots j_k} \in \mathbb{K} \text{ for all } j_1, \dots, j_k \right\}. \end{aligned}$$

If all n_j are equal, i.e. $n_j = n \in \mathbb{N}$ for $j = 1, \dots, k$, we call $T^k := T^k(\mathbb{K}^n) := \mathbb{K}^n \otimes \cdots \otimes \mathbb{K}^n$ the set of all n -dimensional (real or complex) **tensors** of order k .

Remark 3.12. Note that in Definition 3.11 we define tensors not as generally as in [13, Section 6.4] but actually just contravariant tensors. However, this suffices for our purposes.

Example 3.13. (a) For $k = 1$, the set $T^1(\mathbb{K}^n)$ is just the set of (real or complex) vectors, i.e. $T^1(\mathbb{K}^n) = \mathbb{K}^n$.
 (b) For $k = 2$, the set $T^2(\mathbb{K}^n)$ is just the set of (real or complex) $n \times n$ matrices, i.e. $T^2(\mathbb{K}^n) = \mathbb{K}^{n \times n}$.

Next, we define what it means for a tensor to be symmetric.

Definition 3.14. We define S_n as the set of all **permutations** of $\{1, \dots, n\}$, i.e. the set of all bijections $\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$.

Definition 3.15. A tensor $T = (T^{j_1 \cdots j_k})_{j_1, \dots, j_k=1, \dots, n} \in T^k(\mathbb{K}^n)$ is called **symmetric**, if for every permutation $\pi \in S_k$, we have

$$T^{j_1 \cdots j_k} = T^{j_{\pi(1)} \cdots j_{\pi(k)}}$$

for all $j_1, \dots, j_k \in \{1, \dots, n\}$. The set of all (n -dimensional) symmetric tensors (of order k) is denoted by $S^k := S^k(\mathbb{K}^n)$.

The outer product can also be defined for finitely many vectors $x^j \in \mathbb{K}^{n_j}$, $j = 1, \dots, k$.

Definition 3.16. Let $x^j \in \mathbb{K}^{n_j}$, $n_j \in \mathbb{N}$ ($j = 1, \dots, k$). The **outer product** of x^1, \dots, x^k is defined as

$$x^1 \otimes \cdots \otimes x^k = (x_{j_1}^1 \cdots x_{j_k}^k)_{j_1=1, \dots, n_1; \dots; j_k=1, \dots, n_k} \in \mathbb{K}^{n_1} \otimes \cdots \otimes \mathbb{K}^{n_k}.$$

- Remark 3.17.** (a) For $k = 2$, the outer product $x \otimes y$ of two vectors $x \in \mathbb{K}^n, y \in \mathbb{K}^m$ ($m \in \mathbb{N}$) can be thought of as the real or complex $n \times m$ matrix $x \cdot y^\top \in \mathbb{K}^{n \times m}$.
- (b) Obviously, the outer product is not commutative. Consider for example $k = 2$ and the outer product of the two unit vectors $x = e^1 \in \mathbb{R}^m, y = e^2 \in \mathbb{R}^m$, where $n, m \geq 2$. Then $e^1 \otimes e^2$ has a one at $(1, 2)$ and zeros elsewhere, whereas $e^2 \otimes e^1$ has a one at $(2, 1)$ and zeros elsewhere:

$$e^1 \otimes e^2 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} = e^2 \otimes e^1$$

- (c) The outer product is linear, i.e. for $x, y, z \in \mathbb{K}^n$, we have $x \otimes (y + z) = x \otimes y + x \otimes z$ and $(x + y) \otimes z = x \otimes z + y \otimes z$. This follows trivially from the definition of the outer product.

There is also the symmetric outer product of finitely many vectors.

Definition 3.18. Let $x^j \in \mathbb{K}^{n_j}, n_j \in \mathbb{N}$ ($j = 1, \dots, k$). The **symmetric outer product** of x^1, \dots, x^k is defined as

$$x^1 \odot \cdots \odot x^k = \frac{1}{k!} \sum_{\pi \in S_k} x^{\pi(1)} \otimes \cdots \otimes x^{\pi(k)}.$$

Example 3.19. Let $n = 2$ and take $x = (0, 1)^\top, y = (1, 2)^\top \in \mathbb{R}^2$. Then we have

$$x \odot y = \frac{1}{2} (x \otimes y + y \otimes x) = \frac{1}{2} \left(\begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix} \right) = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 2 \end{pmatrix}.$$

- Remark 3.20.** (a) From the definition of the symmetric outer product, it is clear that it is commutative, i.e. for all $x, y \in \mathbb{K}^n$, we have $x \odot y = y \odot x$.
- (b) For $x^1, \dots, x^k \in \mathbb{K}^n$, we have $x^1 \otimes \cdots \otimes x^k \in T^k$ and $x^1 \odot \cdots \odot x^k \in S^k$.
- (c) The outer product of a vector $x \in \mathbb{K}^n$ with itself is the same as the symmetric outer product, i.e. $x \otimes x = x \odot x$.
- (d) For $x \in \mathbb{K}^n$ and $k \in \mathbb{N}_{>0}$ we write $x^k := \underbrace{x \odot \cdots \odot x}_{k \times}$.

Our next goal is to see that the set $T^k = T^k(\mathbb{K}^n)$ of tensors (of order k) can equivalently be defined as the set of multilinear functions $f: (\mathbb{K}^n)^k \rightarrow \mathbb{K}$ as in [20, Section 3.1].

Definition 3.21. A function $f: (\mathbb{K}^n)^k \rightarrow \mathbb{K}$ is called **multilinear** if it is linear in every component, i.e. for all $i \in \{1, \dots, k\}, x^1, \dots, x^k, y^i \in \mathbb{K}^n$ and $\lambda, \mu \in \mathbb{K}$, we have

$$\begin{aligned} & f(x^1, \dots, x^{i-1}, \lambda x^i + \mu y^i, x^{i+1}, \dots, x^k) \\ &= \lambda f(x^1, \dots, x^{i-1}, x^i, x^{i+1}, \dots, x^k) + \mu f(x^1, \dots, x^{i-1}, y^i, x^{i+1}, \dots, x^k). \end{aligned}$$

Further, f is called **symmetric**, if it is invariant under permutations of its arguments, i.e. for all $x^1, \dots, x^k \in \mathbb{K}^n$ and all permutations $\pi \in S_k$, it holds

$$f(x^1, \dots, x^k) = f(x^{\pi(1)}, \dots, x^{\pi(k)}).$$

Definition 3.22. For a tensor $T = (T^{j_1, \dots, j_k})_{j_1, \dots, j_k=1, \dots, n} \in T^k(\mathbb{K})$ of order k , we define the **associated function**

$$f_T: (\mathbb{K}^n)^k \rightarrow \mathbb{K}, (x^1, \dots, x^k) \mapsto \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1 \dots j_k} x_{j_1}^1 \dots x_{j_k}^k. \quad (2)$$

Example 3.23. For $k = 2$, we know that a tensor $T \in T^2(\mathbb{K})$ is just a (real or complex) $n \times n$ matrix, i.e. $T \in \mathbb{K}^{n \times n}$. Hence, we have in this case

$$f_T: \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}, (x, y) \mapsto x^\top T y.$$

Proposition 3.24. *The mapping*

$$T^k(\mathbb{K}) \rightarrow \left\{ f: (\mathbb{K}^n)^k \rightarrow \mathbb{K} : f \text{ multilinear} \right\}, f \mapsto f_T$$

is welldefined and bijective. Further, T is symmetric if and only if f_T is symmetric.

Proof. Let $T = (T^{j_1, \dots, j_k})_{j_1, \dots, j_k=1, \dots, n} \in T^k$ be an arbitrary tensor and consider the associated function $f_T: (\mathbb{K}^n)^k \rightarrow \mathbb{K}$ as in (2). We want to show that f_T is multilinear. Therefore, let $i \in \{1, \dots, k\}$, $x_1, \dots, x_k \in \mathbb{K}^n$, $y^i \in \mathbb{K}^n$ and $\lambda, \mu \in \mathbb{K}$ be arbitrary. Then

$$\begin{aligned} & f_T(x^1, \dots, x^{i-1}, \lambda x^i + \mu y^i, x^{i+1}, \dots, x^k) \\ &= \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1, \dots, j_k} x_{j_1}^1 \dots x_{j_{i-1}}^{i-1} (\lambda x_{j_i}^i + \mu y_{j_i}^i) x_{j_{i+1}}^{i+1} \dots x_{j_k}^k \\ &= \lambda \cdot \left(\sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1, \dots, j_k} x_{j_1}^1 \dots x_{j_k}^k \right) \\ & \quad + \mu \cdot \left(\sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1, \dots, j_k} x_{j_1}^1 \dots x_{j_{i-1}}^{i-1} y_{j_i}^i x_{j_{i+1}}^{i+1} \dots x_{j_k}^k \right) \\ &= \lambda f_T(x^1, \dots, x^k) + \mu f_T(x^1, \dots, x^{i-1}, y^i, x^{i+1}, \dots, x^k). \end{aligned}$$

This shows that f_T is multilinear. On the other hand, let $f: (\mathbb{K}^n)^k \rightarrow \mathbb{K}$ be an arbitrary multilinear mapping. We show that there is a tensor $T \in T^k$ such that $f = f_T$.

Let $e^i \in \mathbb{K}^n$ be the i -th unit vector. For $j_1, \dots, j_k \in \{1, \dots, k\}$ arbitrary, define

$$T^{j_1 \dots j_k} = f(e^{j_1}, \dots, e^{j_k}).$$

Then $T = (T^{j_1 \dots j_k})_{j_1, \dots, j_k=1, \dots, n} \in T^k(\mathbb{K})$ is a tensor of order k .

To show: $f = f_T$.

Let $x^i \in \mathbb{K}^n$ ($i = 1, \dots, k$) be arbitrary. Then the multilinearity property yields

$$\begin{aligned} f(x^1, \dots, x^k) &= f\left(\sum_{i=1}^n x_i^1 e^i, \dots, \sum_{i=1}^n x_i^k e^i\right) \\ &= \sum_{j_1=1}^n \cdots \sum_{j_k=1}^n x_{j_1}^1 \cdots x_{j_k}^k f(e^{j_1}, \dots, e^{j_k}) \\ &= \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1, \dots, j_k} x_{j_1}^1 \cdots x_{j_k}^k = f_T(x^1, \dots, x^k). \end{aligned}$$

Further, since $T^{j_1, \dots, j_k} = f_T(e^{j_1}, \dots, e^{j_k})$, T is the unique tensor with the property $f = f_T$. This shows the first part of the claim. The second part follows easily from the definition of f_T . Therefore, let $x^1, \dots, x^k \in \mathbb{K}^n$ and $\pi \in S_k$ be arbitrary. Then

$$f_T(x^{\pi(1)}, \dots, x^{\pi(k)}) = \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1, \dots, j_k} x_{j_1}^{\pi(1)} \cdots x_{j_k}^{\pi(k)}.$$

Since T is symmetric and $x_{j_1}^{\pi(1)} \cdots x_{j_k}^{\pi(k)} = x_{j_{\pi^{-1}(1)}}^1 \cdots x_{j_{\pi^{-1}(k)}}^k$, we can rewrite this as

$$\begin{aligned} &f_T(x^{\pi(1)}, \dots, x^{\pi(k)}) \\ &= \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_{\pi^{-1}(1)}, \dots, j_{\pi^{-1}(k)}} x_{j_{\pi^{-1}(1)}}^1 \cdots x_{j_{\pi^{-1}(k)}}^k \\ &= \sum_{j_{\pi^{-1}(1)}, \dots, j_{\pi^{-1}(k)}=1, \dots, n} T^{j_{\pi^{-1}(1)}, \dots, j_{\pi^{-1}(k)}} x_{j_{\pi^{-1}(1)}}^1 \cdots x_{j_{\pi^{-1}(k)}}^k \\ &= f(x^1, \dots, x^k), \end{aligned}$$

Hence, f_T is symmetric. On the other hand, if f_T is symmetric, we have by definition of f_T for all $j_1, \dots, j_k \in \{1, \dots, n\}$

$$T^{j_1, \dots, j_k} = f_T(e^{j_1}, \dots, e^{j_k}) = f_T(e^{j_{\pi(1)}}, \dots, e^{j_{\pi(k)}}) = T^{j_{\pi(1)}, \dots, j_{\pi(k)}},$$

i.e. T is symmetric. This shows the claim. \square

In addition, we have another equivalent characterization of tensors in Proposition 3.26, which is stated without proof in [12, Section 2.3]. Therefore, consider the following definition.

Definition 3.25. A function

$$\ell: \mathbb{K}^n \otimes \cdots \otimes \mathbb{K}^n \rightarrow \mathbb{K}, x^1 \otimes \cdots \otimes x^k \mapsto \ell(x^1 \otimes \cdots \otimes x^k),$$

is said to be **linear** in the outer product if for all $x^1, \dots, x^k, y^1, \dots, y^k \in \mathbb{K}^n$

$$\begin{aligned} &\ell\left(\left(x^1 \otimes \cdots \otimes x^k\right) + \left(y^1 \otimes \cdots \otimes y^k\right)\right) \\ &= \ell\left(\left(x^{j_1, \dots, j_k} + y^{j_1, \dots, j_k}\right)_{j_1, \dots, j_k=1, \dots, n}\right) \\ &= \ell\left(x^1 \otimes \cdots \otimes x^k\right) + \ell\left(y^1 \otimes \cdots \otimes y^k\right). \end{aligned}$$

Note that it also holds $x^1 \odot \cdots \odot x^k \in \mathbb{K}^n \otimes \cdots \otimes \mathbb{K}^n$ for all $x^1, \dots, x^k \in \mathbb{R}^n$. Hence, the term $\ell(x^1 \odot \cdots \odot x^k)$ is well defined. For this reason, we say that ℓ is **linear** in the symmetric outer product if for all $x^1, \dots, x^k, y^1, \dots, y^k \in \mathbb{K}^n$

$$\begin{aligned} & \ell\left(\left(x^1 \odot \cdots \odot x^k\right) + \left(y^1 \odot \cdots \odot y^k\right)\right) \\ &= \ell\left(x^1 \odot \cdots \odot x^k\right) + \ell\left(y^1 \odot \cdots \odot y^k\right). \end{aligned}$$

Proposition 3.26. *The mapping $T \mapsto f_T$ ($T \in T^k(\mathbb{K}^n)$) gives a bijection between the set T^k of tensors of order k and the set of functions*

$$\ell: \mathbb{K}^n \otimes \cdots \otimes \mathbb{K}^n \rightarrow \mathbb{K}, \quad x^1 \otimes \cdots \otimes x^k \mapsto \ell(x^1 \otimes \cdots \otimes x^k),$$

that are linear in the outer product. Further, if $T \in S^k$ is symmetric, f_T is linear in the symmetric outer product.

Proof. It is clear that for any tensor $T \in T^k$, the associated function f_T defined in (2) is linear in the components of $x^1 \otimes \cdots \otimes x^k$ ($x^1, \dots, x^k \in \mathbb{R}^n$) and hence linear in the outer product.

On the other hand, let $\ell: \mathbb{K}^n \otimes \cdots \otimes \mathbb{K}^n \rightarrow \mathbb{K}$ be linear in the outer product. But then ℓ must already be of the form

$$\begin{aligned} \ell: \mathbb{K}^n \otimes \cdots \otimes \mathbb{K}^n &\rightarrow \mathbb{K}, \\ x^1 \otimes \cdots \otimes x^k &\mapsto \ell(x^1 \otimes \cdots \otimes x^k) = \sum_{j_1, \dots, j_k=1, \dots, n} \ell^{j_1, \dots, j_k} x_{j_1}^1 \cdots x_{j_k}^k, \end{aligned}$$

where $\ell^{j_1, \dots, j_k} = \ell(e^{j_1} \otimes \cdots \otimes e^{j_k})$ and e^j is the j -th unit vector. This uniquely defines a tensor $T_\ell = (\ell^{j_1, \dots, j_k})_{j_1, \dots, j_k=1, \dots, n} \in T^k$ such that $f_{T_\ell}(x^1, \dots, x^k) = \ell(x^1 \otimes \cdots \otimes x^k)$ for all $x^1, \dots, x^k \in \mathbb{K}^n$.

For the second part of the claim let $T = (T^{j_1, \dots, j_k})_{j_1, \dots, j_k=1, \dots, n} \in S^k$ be symmetric. Then we have

$$\begin{aligned} & f_T(x^1, \dots, x^k) \\ &= \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1 \cdots j_k} x_{j_1}^1 \cdots x_{j_k}^k \\ &= \sum_{j_1, \dots, j_k=1, \dots, n} \left(\frac{1}{k!} \sum_{\pi \in S_k} T^{j_{\pi(1)} \cdots j_{\pi(k)}} \right) x_{j_1}^1 \cdots x_{j_k}^k \\ &\stackrel{T \in S^k}{=} \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1 \cdots j_k} \left(\frac{1}{k!} \sum_{\pi \in S_k} x_{j_{\pi(1)}}^1 \cdots x_{j_{\pi(k)}}^k \right) \\ &= \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1 \cdots j_k} \left(x^1 \odot \cdots \odot x^k \right), \end{aligned}$$

which shows the claim. \square

Note that we can write an arbitrary quadratic form $p(\mathbf{x}) \in H_{n,2}$ always as $p(\mathbf{x}) = \mathbf{x}^\top A_p \mathbf{x}$ for some symmetric matrix $A_p \in \mathbb{R}^{n \times n}$, i.e. $A_p \in S^2(\mathbb{R}^n)$. This can indeed be generalized to forms of arbitrary degree.

Hence, our next goal is to show that there is a one-to-one correspondence between the set $H_{n,k}$ of homogeneous, real polynomials of degree k and the set $S^k(\mathbb{R}^n)$ of symmetric, real tensors. This is stated without a formal proof in [12, Section 2.4]. Before stating that theorem, we need the definition of the generalized binomial coefficient.

Definition 3.27. For $\ell \in \mathbb{N}$ and $\alpha = (\alpha_1, \dots, \alpha_\ell) \in \{1, \dots, n\}^\ell$ such that $\sum_{i=1}^\ell \alpha_i = n$, we define the **generalized binomial coefficient**

$$\binom{n}{\alpha} := \binom{n}{\alpha_1, \dots, \alpha_\ell} := \frac{n!}{\alpha_1! \cdots \alpha_\ell!}.$$

Remark 3.28. For $\ell = 2$ and $\alpha_1, \alpha_2 \in \{1, \dots, n\}$ such that $\alpha_1 + \alpha_2 = n$ we have $\alpha_2 = n - \alpha_1$ and hence $\binom{n}{\alpha_1, \alpha_2} = \binom{n}{\alpha_1} = \binom{n}{\alpha_2}$ is the regular binomial coefficient.

Theorem 3.29. *There is a linear bijection $\varphi: H_{n,k} \rightarrow S^k(\mathbb{R}^n), p \mapsto T_p$ between the set $H_{n,k}$ of homogeneous, real polynomials in n variables of degree k and the set $S^k(\mathbb{R}^n)$ of symmetric, real tensors of order k .*

Proof. The tensor T_p corresponding to a form $p \in H_{n,k}$ is defined in (3)–(5) of this proof via the coefficients of p .

Let $p = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k} p_\alpha \mathbf{x}^\alpha \in H_{n,k}$ be an arbitrary real form. By writing each monomial of degree k as the product of k single variables, the form p can equivalently be written as

$$p = \sum_{j_1, \dots, j_k=1, \dots, n} \tilde{p}^{j_1 \cdots j_k} \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_k} \quad (3)$$

for some $\tilde{p}^{j_1 \cdots j_k} \in \mathbb{R}$, $j_1, \dots, j_k \in \{1, \dots, n\}$. We want to choose the 'new' coefficients $\tilde{p}^{j_1 \cdots j_k}$ such that they are invariant under permutation of indices. Therefore, let $\alpha \in \mathbb{N}_0^n$ s.t. $|\alpha| = k$ be arbitrary. We want to count the number of possible choices of indices $j_1, \dots, j_k \in \{1, \dots, n\}$ such that $\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_k} = \mathbf{x}^\alpha$. There are exactly

$$\begin{aligned} & \binom{k}{\alpha_1} \binom{k - \alpha_1}{\alpha_2} \cdots \binom{k - \alpha_1 - \cdots - \alpha_{n-1}}{\alpha_n} \\ &= \frac{k!}{\alpha_1! \cdots \alpha_n!} = \binom{k}{\alpha_1, \dots, \alpha_n} = \binom{k}{\alpha} \end{aligned}$$

different choices $j_1, \dots, j_k \in \{1, \dots, n\}$ such that for each $i \in \{1, \dots, n\}$, it holds $\alpha_i = \#\{\ell \in \{1, \dots, k\} : j_\ell = i\}$, which means precisely $\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_k} = \mathbf{x}^\alpha$. Hence, we can define

$$\tilde{p}^{j_1 \cdots j_k} = \binom{k}{\alpha_1, \dots, \alpha_n}^{-1} p_\alpha = \binom{k}{\alpha}^{-1} p_\alpha \quad (4)$$

for $j_1, \dots, j_k \in \{1, \dots, k\}$ arbitrary, where $\alpha \in \mathbb{N}_0^n$ is again given by $\alpha_i = \#\{\ell \in \{1, \dots, k\} : j_\ell = i\}$ for all $i \in \{1, \dots, n\}$.

The coefficients $\tilde{p}^{j_1 \cdots j_k}$ are now invariant under permutations of the indices $j_1, \dots, j_k \in \{1, \dots, n\}$. To see this, let $\pi \in S_k$ and $j_1, \dots, j_k \in \{1, \dots, n\}$ be arbitrary. Choose $\alpha \in \mathbb{N}_0^n$ such that for each $i \in \{1, \dots, n\}$, we have

$$\alpha_i = \#\{\ell \in \{1, \dots, k\} : j_\ell = i\}.$$

Since $\pi: \{1, \dots, k\} \rightarrow \{1, \dots, k\}$ is a bijection, this is the same as writing $\alpha_i = \#\{\pi(\ell) \in \{1, \dots, k\} : j_{\pi(\ell)} = i\}$. Again, since π is bijective, this is just the same as counting the number of $\ell \in \{1, \dots, k\}$ such that $j_{\pi(\ell)} = i$, i.e.

$$\alpha_i = \#\{\ell \in \{1, \dots, k\} : j_{\pi(\ell)} = i\}.$$

This shows

$$\tilde{p}^{j_1 \dots j_k} = \binom{k}{\alpha}^{-1} p_\alpha = \tilde{p}^{j_{\pi(1)} \dots j_{\pi(k)}}$$

Further, note that the coefficients $\tilde{p}^{j_1, \dots, j_k}$ are uniquely determined by (3) and the property of being invariant under permutations of indices since for $\alpha \in \mathbb{N}_0^k, |\alpha| = k$ we must have

$$p_\alpha = \sum_{j_1, \dots, j_k \in \{1, \dots, n\} : \mathbf{x}_{j_1} \dots \mathbf{x}_{j_k} = \mathbf{x}^\alpha} \tilde{p}^{j_1, \dots, j_k}.$$

We conclude that

$$\varphi: H_{n,k} \rightarrow S^k(\mathbb{R}^n), p \mapsto T_p = \left(\tilde{p}^{j_1 \dots j_k} \right)_{j_1=1, \dots, j_k=1}^{n, \dots, n} \quad (5)$$

is welldefined. Further, by the choice of the components $\tilde{p}^{j_1, \dots, j_k}$ in (4) as coefficients of the polynomial p in (3), we know that φ is linear.

On the other hand, consider the mapping

$$\begin{aligned} \psi: S^k(\mathbb{R}^n) &\rightarrow H_{n,k}, \\ S = \left(S^{j_1 \dots j_k} \right)_{j_1=1, \dots, j_k=1}^{n, \dots, n} &\mapsto p_S(\mathbf{x}) = \sum_{j_1, \dots, j_k=1, \dots, n} S^{j_1, \dots, j_k} \mathbf{x}_{j_1} \dots \mathbf{x}_{j_k} \end{aligned}$$

Clearly, it holds $\psi(S) = p_S(\mathbf{x}) \in H_{n,k}$ for all $S \in S^k(\mathbb{R}^n)$ and hence the mapping is welldefined. From (3) and the definition of φ and ψ , we also know that $\psi(\varphi(p)) = p$ for all $p \in H_{n,k}$, i.e. $\psi = \varphi^{-1}$.

For $S \in S^k(\mathbb{R}^n)$, we know by the symmetry of S that the components of S are coefficients of $\psi(S) \in H_{n,k}$ as in (3) that are invariant under permutations. As before, coefficients of a form with those properties are unique, which yields that $\varphi(\psi(S)) = S$. Hence, we can deduce that φ is bijective, as desired. \square

Remark 3.30. Analogously as in Theorem 3.29 one can see that there is a bijection between complex forms in n variables of degree k and the set $S^k(\mathbb{C}^n)$ of symmetric, complex tensors of order k . However, Theorem 3.29 is stated in a real setting since we are interested in investigating the sets $P_{n,2d}, \Sigma_{n,2d}$ and $C_{n,2d}$ of real forms.

The one-to-one correspondence from Theorem 3.29 allows us to introduce a biform associated to a polynomial $p \in H_{n,k}$. This biform will be used in the formulation of generalized Cauchy–Schwarz inequalities in Section 4.

Definition 3.31. Let $k = 2d \in \mathbb{N}$ be even. For $p \in H_{n,k}$ we call the tensor $T_p(\mathbb{R}^n) \in S^k$ as in Theorem 3.29 the (symmetric) **tensor associated** to p . Further, we define the **biform** Q_p **associated** to p as

$$Q_p: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, (x, y) \mapsto Q_p(x, y) = f_{T_p} \left(\underbrace{x, \dots, x}_{d \times}, \underbrace{y, \dots, y}_{d \times} \right)$$

For simplicity, we write

$$Q_p(\mathbf{x}, \mathbf{y}) = f_{T_p}(\mathbf{x}, \dots, \mathbf{x}, \mathbf{y}, \dots, \mathbf{y}).$$

Example 3.32. Consider again a quadratic form $p \in H_{n,2}$ and let $A_p \in \mathbb{R}^{n \times n} = T^2(\mathbb{R}^n)$ be such that $p(\mathbf{x}) = \mathbf{x}^\top A_p \mathbf{x}$. In this case, the tensor T_p associated to p is given by $T_p : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $T_p(x^1, x^2) = (x^1)^\top A_p x^2$. Hence, this coincides with the biform associated to p , which is given by $Q_p : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $Q_p(x, y) = x^\top A_p y$.

Remark 3.33. Let $T \in T^k(\mathbb{R}^n)$ be a tensor. Of course, we can also define

$$f_T : (\mathbb{C}^n)^k \rightarrow \mathbb{C}, (x^1, \dots, x^k) \mapsto \sum_{j_1, \dots, j_k=1, \dots, n} T^{j_1 \dots j_k} x_{j_1}^1 \dots x_{j_k}^k$$

for complex vectors. As before, it can be seen that f_T is multilinear. If T is in addition symmetric, i.e. $T \in S^k(\mathbb{R}^n)$, we have $Q_p(z, \bar{z}) \in \mathbb{R}$ for all $z \in \mathbb{C}^n$. This follows directly from the linearity of f_T in every component. The quantity $Q_p(z, \bar{z})$ will come up again in the generalized Cauchy–Schwarz inequalities in Theorem 4.3.

The next goal is to show that the following polarization identity formulated in [12, Section 2.4] holds. It is indeed an easy corollary of what we have seen so far.

Theorem 3.34 (Polarization Identity). *Let $p \in H_{n,k}$ be arbitrary and $T_p \in S^k(\mathbb{R}^n)$ be the associated symmetric tensor. Consider the associated multilinear functional $f_{T_p} : (\mathbb{R}^n)^k \rightarrow \mathbb{R}$. Then we have*

$$\forall x \in \mathbb{R}^n : p(x) = f_{T_p}(x, \dots, x), \quad (6)$$

which can be written as $p(\mathbf{x}) = f_{T_p}(\mathbf{x}, \dots, \mathbf{x})$.

Proof. This follows directly from the definition of the associated tensor T_p in (5), the definition of the associated functional f_{T_p} in (2) and the representation (3) of the form p via the coefficients $\tilde{p}^{j_1, \dots, j_k}$. Indeed, we have

$$f_{T_p}(x, \dots, x) \stackrel{(5) \& (2)}{=} \sum_{j_1, \dots, j_k=1, \dots, n} \tilde{p}^{j_1, \dots, j_k} x_{j_1} \dots x_{j_k} \stackrel{(3)}{=} p(x).$$

□

Corollary 3.35. *Let $p = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=k} p_\alpha \mathbf{x}^\alpha \in H_{n,k}$ be arbitrary. In the situation of Theorem 3.34, we have*

$$\forall \alpha \in \mathbb{N}_0^n, |\alpha| = k : p_\alpha = \binom{k}{\alpha} f_{T_p} \left(\underbrace{e^1, \dots, e^1}_{\alpha_1 \times}, \dots, \underbrace{e^n, \dots, e^n}_{\alpha_n \times} \right),$$

where $e^i \in \mathbb{R}^n$ is the i -th unit vector.

Proof. Let $\alpha \in \mathbb{N}_0^n, |\alpha| = k$ be arbitrary. Let $j_1, \dots, j_k \in \{1, \dots, n\}$ be arbitrary such that for each $i \in \{1, \dots, n\}$, we have $\alpha_i = \#\{\ell \in \mathbb{N} : j_\ell = i\}$. Note that this choice of j_ℓ is not unique!

By definition of the multilinear mapping f_{T_p} in (2) we know that

$$f_{T_p} \left(\underbrace{e^1, \dots, e^1}_{\alpha_1 \times}, \dots, \underbrace{e^n, \dots, e^n}_{\alpha_n \times} \right) = T_p^{1 \dots 1 \dots n \dots n} \stackrel{(5)}{=} \tilde{p}^{1 \dots 1 \dots n \dots n},$$

where at $T_p^{1 \dots 1 \dots n \dots n}$ and $\tilde{p}^{1 \dots 1 \dots n \dots n}$ we have α_i -times the index i ($i = 1, \dots, n$).

Since T_p is symmetric, this is equal to $\tilde{p}^{j_1 \dots j_k}$. Hence, we can deduce by (4) that

$$f_{T_p} (e^1, \dots, e^1, \dots, e^n, \dots, e^n) = \tilde{p}^{j_1 \dots j_k} \stackrel{(4)}{=} \binom{k}{\alpha}^{-1} p_\alpha.$$

Rearranging shows the claim. \square

Remark 3.36. As already mentioned in Remark 3.30, the results of Theorem 3.34 and Corollary 3.35 can also be stated for complex forms. However, for our purposes it suffices to work in the real setting.

Convention 3.37. For simplicity, from now on we no longer distinguish between a tensor T and its associated function f_T .

3.3. Hilbert Space Representations of Linear Functionals on $H_{n,2d}$.

The goal of this section is to show that $H_{n,2d}$ can be equipped with an inner product $\langle \cdot, \cdot \rangle$ called the Fischer inner product, such that $(H_{n,2d}, \langle \cdot, \cdot \rangle)$ is a Hilbert space. In addition, we find a representation of the (algebraic) dual of $H_{n,2d}$ as differential operators.

We start with recalling some basic definition of Functional Analysis. For more details, check e.g. [10, chapters 11–19].

Convention 3.38. For the rest of this work, let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be arbitrary. Furthermore, write \bar{z} for the complex conjugate of a complex number $z \in \mathbb{C}$.

Definition 3.39. Let V be a \mathbb{K} -vector space. An **inner product** on V is a mapping $(\cdot, \cdot) : V \times V \rightarrow \mathbb{K}$ such that the following holds :

- (1) $\forall u, v, w \in V \forall \lambda, \mu \in \mathbb{K} : (\lambda u + \mu v, w) = \lambda(u, w) + \mu(v, w)$ (**linear**),
- (2) If $\mathbb{K} = \mathbb{R} : \forall u, v \in V : (u, v) = (v, u)$ (**symmetric**),
If $\mathbb{K} = \mathbb{C} : \forall u, v \in V : (u, v) = \overline{(v, u)}$ (**hermitian**),
- (3) $\forall v \in V : (v, v) \geq 0 \wedge ((v, v) = 0 \iff v = 0)$ (**positive-definite**).

A **norm** on V is a mapping $\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0}$, which fulfills the following

- (1) $\forall v \in V \forall \lambda \in \mathbb{K} : \|\lambda v\| = |\lambda| \cdot \|v\|$ (**positive homogeneous**),
- (2) $\forall u, v \in V : \|u + v\| \leq \|u\| + \|v\|$ (**triangle inequality**),
- (3) $\forall v \in V : \|v\| = 0 \iff v = 0$ (**positive definite**).

Example 3.40. (a) If $V = \mathbb{K}^n$ and $Q \in \mathbb{R}^{n \times n}$ is positive definite, an inner product on V is given by

$$(\cdot, \cdot)_Q : V \times V \rightarrow \mathbb{K}, (x, y)_Q = \bar{x}^\top Q y.$$

(b) For $Q = I_n$ the identity matrix, we obtain the Euclidean inner product denoted by $(\cdot, \cdot)_2 = (\cdot, \cdot)_{I_n}$.

(c) More general, every finite dimensional \mathbb{K} -vector space V ($n = \dim_{\mathbb{K}} V$) has an inner product. Therefore, take a basis $\mathcal{B} = \{b^1, \dots, b^n\}$ of V and define $(\cdot, \cdot) : V \times V \rightarrow \mathbb{K}$ via $(b^i, b^j) = \delta_{ij}$ for all $i, j \in \{1, \dots, n\}$. Note

that an inner product on V is uniquely defined by its values on a basis of V .

Remark 3.41. Let $(V, (\cdot, \cdot))$ be a \mathbb{K} -vector space equipped with an inner product $(\cdot, \cdot): V \times V \rightarrow \mathbb{K}$. Then we can define a norm on V called the norm **generated** or **induced** by the inner product via

$$\|\cdot\| : V \rightarrow \mathbb{R}, \|v\| = \sqrt{(v, v)}.$$

As the next step, it can be seen how an inner product on $H_{n,2d}$ can be defined.

Definition 3.42. The **Fischer inner product** on $H_{n,2d}$ is defined via

$$\langle \cdot, \cdot \rangle : H_{n,2d} \times H_{n,2d} \rightarrow \mathbb{R}, \langle p, q \rangle = p(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n)q, \quad (7)$$

where $p(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n)$ is the differential operator obtained from p by replacing all variables \mathbf{x}_i with the partial derivatives $\partial \mathbf{x}_i$ ($i = 1, \dots, n$). The expression $p(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n) \cdot q$ is then defined as the differential operator $p(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n)$ applied to q .

Remark 3.43. (a) Let $\alpha, \beta \in \mathbb{N}_0^n$ be arbitrary such that $|\alpha| = |\beta|$. Then we have

$$\underbrace{\partial \mathbf{x}_1 \cdots \partial \mathbf{x}_1}_{\alpha_1 \times} \cdots \underbrace{\partial \mathbf{x}_n \cdots \partial \mathbf{x}_n}_{\alpha_n \times} \mathbf{x}^\beta = \partial \mathbf{x}_1^{\alpha_1} \cdots \partial \mathbf{x}_n^{\alpha_n} \mathbf{x}^\beta = \begin{cases} 0, & \text{if } \alpha \neq \beta \\ \alpha_1! \cdots \alpha_n!, & \text{else} \end{cases}. \quad (8)$$

(b) Using (8), the Fischer inner product $\langle p, q \rangle$ of two forms $p, q \in H_{n,2d}$ may also be expressed in terms of the coefficients of p and q . Therefore, we write $p = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=2d} p_\alpha \mathbf{x}^\alpha$, $q = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=2d} q_\alpha \mathbf{x}^\alpha$, where $p_\alpha, q_\alpha \in \mathbb{R}$. Then we have

$$\begin{aligned} \langle p, q \rangle &= \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=2d} p_\alpha \partial \mathbf{x}_1^{\alpha_1} \cdots \partial \mathbf{x}_n^{\alpha_n} \left(\sum_{\beta \in \mathbb{N}_0^n, |\beta|=2d} q_\beta \mathbf{x}^\beta \right) \\ &\stackrel{(8)}{=} \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=2d} \alpha_1! \cdots \alpha_n! p_\alpha q_\alpha \\ &= (2d)! \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=2d} \binom{2d}{\alpha}^{-1} p_\alpha q_\alpha. \end{aligned} \quad (9)$$

Remark 3.44. Using the expression of the Fischer inner product in (9), it can easily be seen that it is indeed an inner product. Linearity can be deduced by using $(\lambda_1 p_1 + \lambda_2 p_2)_\alpha = \lambda_1 (p_1)_\alpha + \lambda_2 (p_2)_\alpha$ for all $\lambda_1, \lambda_2 \in \mathbb{R}, p_1, p_2 \in H_{n,2d}$ and $\alpha \in \mathbb{N}_0^n, |\alpha| = 2d$. The symmetry follows since (8) is symmetric in the coefficients p_α, q_α of p and q . Further, it is positive definite since $p_\alpha^2 \geq 0$ for all α and $\langle p, p \rangle = 0$ if and only if $p_\alpha^2 = 0$ for all α , which is equivalent to $p = 0$.

Hence, we also have a norm on $H_{n,2d}$ induced by $\langle \cdot, \cdot \rangle$. Next, we recall the concept of completeness of a normed vector space.

Definition 3.45. Let $(V, \|\cdot\|)$ be a normed \mathbb{K} -vector space, $(v_n)_{n \in \mathbb{N}} \subseteq V$ a sequence. Then $(v_n)_{n \in \mathbb{N}} \subseteq V$ is called a **Cauchy sequence** if

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n, m \geq N : \|v_n - v_m\| < \varepsilon.$$

The sequence $(v_n)_{n \in \mathbb{N}} \subseteq V$ **converges** to some $v \in V$ if

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N : \|v_n - v\| < \varepsilon.$$

Further, $(V, \|\cdot\|)$ is said to be **complete** if every Cauchy sequence in V converges to some limit in V . In this case, $(V, \|\cdot\|)$ is called a **Banach space**.

If $(V, (\cdot, \cdot))$ is a vector space with inner product such that V together with the induced norm is complete, it is called a **Hilbert space**.

It is a standard result of functional analysis that all finite dimensional normed vector spaces over \mathbb{K} are complete. This can for example be seen by studying the norms on finite dimensional vector spaces.

Definition 3.46. Let $\|\cdot\|_1, \|\cdot\|_2$ be two norms on a \mathbb{K} -vector space V . Then the norms $\|\cdot\|_1, \|\cdot\|_2$ are said to be **equivalent** if the following holds:

$$\exists m, M > 0 \forall v \in V : m \|v\|_1 \leq \|v\|_2 \leq M \|v\|_1.$$

Remark 3.47. From [10, Bemerkung 13.6] we know that equivalent norms induce the same topology.

Theorem 3.48. *Let V be a finite dimensional \mathbb{K} -vector space. Then all norms on V are equivalent.*

Proof. [10, Satz 13.12]. □

As a consequence, all normed finite dimensional \mathbb{K} -vector spaces are complete.

Corollary 3.49. *Let V be an arbitrary finite dimensional \mathbb{K} -vector space with norm $\|\cdot\|$. Then $(V, \|\cdot\|)$ is complete.*

Proof. From Theorem 3.48, we know that all norms on V are equivalent. Hence, Remark 3.47 yields that V has a unique topology induced by a norm. Since $(\mathbb{K}^n, \|\cdot\|_2)$ is a Banach space (see [10, Beispiel 12.4]) and V is isomorphic to \mathbb{K}^n , this shows that V is complete as well. □

Since every finite dimensional \mathbb{K} -vector space has an inner product and hence an induced norm, we now know that every finite dimensional \mathbb{K} -vector space is complete.

Corollary 3.50. *The set $H_{n,2d}$ of forms in n variables and of degree $2d$ equipped with the Fischer inner product $\langle \cdot, \cdot \rangle$ is a Hilbert space.*

Proof. This follows directly from Corollary 3.49 since $H_{n,2d}$ is finite dimensional. □

Definition 3.51. Let V be a finite dimensional \mathbb{K} -vector space. Then we call the unique topology induced by a norm on V the **Euclidean topology**.

Our next goal is to express linear functionals on $H_{n,2d}$ via differential operators. The key ingredient to do so is the Riesz Representation Theorem for Hilbert spaces. Before we state the theorem, the definition of the (algebraic) dual of a vector space V is recalled. In our setting it suffices to study finite dimensional vector spaces, which is why we do not need to distinguish between the algebraic and topologic dual.

Definition 3.52. Let V be a (finite dimensional) \mathbb{K} -vector space. Then the (algebraic) **dual** V^\vee of V is defined as

$$V^\vee = \{\varphi: V \rightarrow \mathbb{K} : \varphi \text{ is linear}\}.$$

If V is in addition equipped with a norm $\|\cdot\|_V : V \rightarrow \mathbb{R}_{\geq 0}$, we define the **operator norm** (see [10, Def. 13.2]) on V^\vee as

$$\|\cdot\|_{V^\vee} : V^\vee \rightarrow \mathbb{R}_{\geq 0}, \|\ell\|_{V^\vee} = \sup_{x \in V \setminus \{0\}} \frac{|\ell(x)|}{\|x\|_V} = \sup_{x \in V \setminus \{0\}, \|x\|_V=1} |\ell(x)|.$$

Remark 3.53. If V is a \mathbb{K} -vector space, so is the dual V^\vee , where addition and scalar multiplication are defined in a canonical way.

Theorem 3.54 (Riesz Representation Theorem). *Let H be a finite dimensional \mathbb{K} -Hilbert space with inner product $(\cdot, \cdot): H \times H \rightarrow \mathbb{K}$ and $\ell \in H^\vee$ a linear functional. Then there exist an $h_\ell \in H$ such that*

$$\forall x \in H : \ell(x) = (x, h_\ell). \quad (10)$$

Moreover, the mapping $I_{\text{Riesz}}: H^\vee \rightarrow H, \ell \mapsto h_\ell$ is a bijective isometry and conjugated linear, i.e.

$$\begin{aligned} \forall \ell_1, \ell_2 \in H^\vee \quad \forall \lambda_1, \lambda_2 \in \mathbb{K} : \quad I_{\text{Riesz}}(\lambda_1 \ell_1 + \lambda_2 \ell_2) &= \overline{\lambda_1} I_{\text{Riesz}}(\ell_1) + \overline{\lambda_2} I_{\text{Riesz}}(\ell_2), \\ \forall \ell \in H^\vee : \quad \|\ell\|_{H^\vee} &= \|I_{\text{Riesz}}(\ell)\|_H. \end{aligned}$$

Proof. [10, Satz 12.24]. □

The Riesz Representation Theorem yields in particular the following.

Corollary 3.55. *If V is a finite dimensional \mathbb{K} -vector space, then so is the dual V^\vee and the dimensions coincide. In other words, in the finite dimensional case a vector space and its dual are isomorphic.*

Corollary 3.56. *Let $\ell \in H_{n,2d}^\vee$ be an arbitrary linear functional on $H_{n,2d}$. Then there exists a $p_\ell \in H_{n,2d}$ such that*

$$\forall q \in H_{n,2d} : \ell(q) = p_\ell(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n)q.$$

Further, the mapping $H_{n,2d}^\vee \rightarrow H_{n,2d}, \ell \mapsto p_\ell$ is one-to-one.

Proof. This follows directly from Theorem 3.54 using the symmetry of the Fischer inner product. □

An important class of linear functionals on $H_{n,2d}$ are tensor evaluations, which are studied in the following lemma.

Lemma 3.57. *Let $x^1, \dots, x^{2d} \in \mathbb{R}^n$ be arbitrary and consider the linear functional $\ell \in H_{n,2d}^\vee$ given by*

$$\ell: H_{n,2d} \rightarrow \mathbb{R}, p \mapsto \ell(p) = T_p(x^1, \dots, x^{2d}).$$

Further, let $q \in H_{n,2d}$ be as in the Riesz Representation Theorem 3.54 such that

$$\forall p \in H_{n,2d} : \ell(p) = \langle q, p \rangle = q(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n) p. \quad (11)$$

Then

$$q(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_n) = \frac{1}{(2d)!} \partial_{x^1} \cdots \partial_{x^{2d}}. \quad (12)$$

Proof. Note that in (12), $\partial \mathbf{x}_i$ is the partial derivative with respect to the variable \mathbf{x}_i , whereas ∂_{x^i} is the directional derivative with respect to x^i ($i = 1, \dots, 2d$) (see Definition 2.7). Further, Theorem 3.29 tells us that ℓ is indeed linear, i.e. $\ell \in H_{n,k}^\vee$.

In the following, we show that q as in (12) fulfills (11). The claim follows then by the uniqueness in the Riesz Representation Theorem. Therefore, let $p \in H_{n,2d}$ be arbitrary and write

$$p = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha|=2d} p_\alpha \mathbf{x}^\alpha = \sum_{j_1, \dots, j_{2d}=1, \dots, n} \tilde{p}^{j_1 \cdots j_{2d}} \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}},$$

where the coefficients $\tilde{p}^{j_1 \cdots j_{2d}} \in \mathbb{R}$ are defined as in (4). The multilinear map

$$T_p = (\tilde{p}^{j_1 \cdots j_{2d}})_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} : (\mathbb{R}^n)^{2d} \rightarrow \mathbb{R}$$

evaluated at x^1, \dots, x^{2d} is given by

$$T_p(x^1, \dots, x^{2d}) = \sum_{j_1, \dots, j_{2d}=1, \dots, n} \tilde{p}^{j_1 \cdots j_{2d}} x_{j_1}^1 \cdots x_{j_{2d}}^{2d}. \quad (13)$$

On the other hand, $q(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_{2d})$ as in the claim applied to p leads to

$$\begin{aligned} & q(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_{2d}) p \\ &= \sum_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} q(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_{2d}) \tilde{p}^{j_1 \cdots j_{2d}} \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}} \\ &= \sum_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} \frac{1}{(2d)!} \partial_{x^1} \cdots \partial_{x^{2d}} \tilde{p}^{j_1 \cdots j_{2d}} \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}}. \end{aligned} \quad (14)$$

Hence, it suffices to evaluate expressions of the form

$$q(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_{j_{2d}}) \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}},$$

where $j_1, \dots, j_{2d} \in \{1, \dots, n\}$. Using the formula (1) for the directional derivative of a polynomial in Proposition 2.10 reveals

$$\begin{aligned} & q(\partial \mathbf{x}_1, \dots, \partial \mathbf{x}_{j_{2d}}) \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}} = \frac{1}{(2d)!} \partial_{x^1} \cdots \partial_{x^{2d}} \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}} \\ & \stackrel{(1)}{=} \frac{1}{(2d)!} \partial_{x^1} \cdots \partial_{x^{2d-1}} \sum_{i=1}^n x_i^{2d} \partial \mathbf{x}_i (\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}}) \\ & = \frac{1}{(2d)!} \sum_{i_1=1, \dots, i_{2d}=1} x_{i_1}^1 \cdots x_{i_{2d}}^{2d} (\partial \mathbf{x}_{i_1} \cdots \partial \mathbf{x}_{i_{2d}}) (\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}}). \end{aligned}$$

The expression $(\partial_{\mathbf{x}_{i_1}} \cdots \partial_{\mathbf{x}_{i_{2d}}})(\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}})$ is nonzero if and only if for some permutation $\sigma \in S_{2d}$, we have $(i_{\sigma(1)}, \dots, i_{\sigma(2d)}) = (j_1, \dots, j_{2d})$. In this case, $(\partial_{\mathbf{x}_{i_1}} \cdots \partial_{\mathbf{x}_{i_{2d}}})(\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}}) = 1$. We deduce

$$q(\partial_{\mathbf{x}_1}, \dots, \partial_{\mathbf{x}_{2d}})\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}} = \frac{1}{(2d)!} \sum_{\sigma \in S_{2d}} x_{j_{\sigma(1)}}^1 \cdots x_{j_{\sigma(2d)}}^{2d}. \quad (15)$$

Using (15) in (14) shows

$$\begin{aligned} q(\partial_{\mathbf{x}_1}, \dots, \partial_{\mathbf{x}_{2d}})p &\stackrel{(14)}{=} \sum_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} q(\partial_{\mathbf{x}_1}, \dots, \partial_{\mathbf{x}_{2d}})\tilde{p}^{j_1 \cdots j_{2d}}\mathbf{x}_{j_1} \cdots \mathbf{x}_{j_{2d}} \\ &\stackrel{(15)}{=} \sum_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} \left(\tilde{p}^{j_1 \cdots j_{2d}} \frac{1}{(2d)!} \sum_{\sigma \in S_{2d}} x_{j_{\sigma(1)}}^1 \cdots x_{j_{\sigma(2d)}}^{2d} \right) \\ &= \frac{1}{(2d)!} \sum_{\sigma \in S_{2d}} \sum_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} \tilde{p}^{j_1 \cdots j_{2d}} x_{j_{\sigma(1)}}^1 \cdots x_{j_{\sigma(2d)}}^{2d}. \end{aligned}$$

The coefficients $\tilde{p}^{j_1 \cdots j_{2d}}$ are invariant under permutations of the indices, which means in particular $\tilde{p}^{j_1 \cdots j_{2d}} = \tilde{p}^{j_{\sigma(1)} \cdots j_{\sigma(2d)}}$ for all $\sigma \in S_{2d}$. Hence, we have

$$\begin{aligned} q(\partial_{\mathbf{x}_1}, \dots, \partial_{\mathbf{x}_{2d}})p &= \frac{1}{(2d)!} \sum_{\sigma \in S_{2d}} \sum_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} \tilde{p}^{j_{\sigma(1)} \cdots j_{\sigma(2d)}} x_{j_{\sigma(1)}}^1 \cdots x_{j_{\sigma(2d)}}^{2d} \\ &= \frac{1}{(2d)!} \sum_{\sigma \in S_{2d}} \sum_{j_{\sigma(1)}=1, \dots, j_{\sigma(2d)}=1}^{n, \dots, n} \tilde{p}^{j_{\sigma(1)} \cdots j_{\sigma(2d)}} x_{j_{\sigma(1)}}^1 \cdots x_{j_{\sigma(2d)}}^{2d} \\ &= \frac{1}{(2d)!} (2d)! \sum_{j_1=1, \dots, j_{2d}=1}^{n, \dots, n} \tilde{p}^{j_1 \cdots j_{2d}} x_{j_1}^1 \cdots x_{j_{2d}}^{2d} \\ &\stackrel{(13)}{=} T_p(x^1, \dots, x^{2d}) = \ell(p). \end{aligned}$$

The claim follows by the uniqueness in the Riesz Representation Theorem (Theorem 3.54). \square

Remark 3.58. Lemma 3.57 can also be formulated in an analogous way for complex vectors $z^1, \dots, z^{2d} \in \mathbb{C}^n$, where we consider linear mappings $\ell: H_{n,2d} \rightarrow \mathbb{C}$. Note that if $z = x + iy \in \mathbb{C}^n$ is complex, we have for the directional derivative of a polynomial p

$$\partial_z p \stackrel{(1)}{=} \nabla p \cdot z = \nabla p \cdot x + i \nabla p \cdot y = \partial_x p + i \partial_y p$$

and hence $\partial_z = \partial_x + i \partial_y$.

There are some special tensor evaluations, which are of particular interest.

Example 3.59. (a) Point evaluations: Let $x \in \mathbb{R}^n$ be arbitrary and consider the point evaluation at x given by $\text{ev}_x: H_{n,2d} \rightarrow \mathbb{R}, p \mapsto p(x)$. Clearly, $\text{ev}_x \in H_{n,2d}^\vee$ is linear. Further, by polarization identity (6) we know that for all $p \in H_{n,2d}$, we have $\text{ev}_x(p) = p(x) = T_p(x, \dots, x)$. Hence, by using Lemma 3.57, we deduce that the linear operator associated to ev_x is $\frac{1}{(2d)!} \partial_x^{2d}$.

- (b) Complex point evaluations: As mentioned in Remark 3.58, it can be deduced analogously as in (a) that if $z = x + iy \in \mathbb{C}^n$ is complex with $x, y \in \mathbb{R}^n$, the evaluation at z , ev_z corresponds to $\frac{1}{(2d)!} \partial_z^{2d}$. Since $\partial_z = \partial_x + i\partial_y$, this can be written as $\frac{1}{(2d)!} \partial_z^{2d} = \frac{1}{(2d)!} (\partial_x + i\partial_y)^{2d}$.
- (c) Evaluations of bivariate form Q_p : Let $x, y \in \mathbb{R}^n$ be arbitrary and consider the mapping

$$\text{ev}_{x,y}: H_{n,2d} \rightarrow \mathbb{R}, p \mapsto Q_p(x, y).$$

We know by Theorem 3.29 that this is linear, i.e. $\text{ev}_{x,y} \in H_{n,2d}^\vee$. Further, since $\text{ev}_{x,y}(p) = Q_p(x, y) = T_p(\underbrace{x, \dots, x}_{d \times}, \underbrace{y, \dots, y}_{d \times})$, Lemma 3.57 tells us

that the operator associated with $\text{ev}_{x,y}$ is given by $\frac{1}{(2d)!} \partial_x^d \partial_y^d$.

- (d) Complex evaluations of bivariate form Q_p : As in (c) consider the linear functional

$$\text{ev}_{z,\bar{z}}: H_{n,2d} \rightarrow \mathbb{R}, p \mapsto Q_p(z, \bar{z}).$$

Note that by definition of $Q_p(\cdot, \cdot)$ and the multilinearity of T_p , we have $Q_p(z, \bar{z}) \in \mathbb{R}$ for all $p \in H_{n,2d}$, i.e. $\text{ev}_{z,\bar{z}} \in H_{n,2d}^\vee$. Analogously as in (c), the differential operator corresponding to $\text{ev}_{z,\bar{z}}$ is $\frac{1}{(2d)!} \partial_z^d \partial_{\bar{z}}^d$. Now write $z = x + iy$, where $x, y \in \mathbb{R}^n$. Since $\partial_z = \partial_x + i\partial_y$, $\partial_{\bar{z}} = \partial_x - i\partial_y$ (compare (b)), the operator corresponding to $\text{ev}_{z,\bar{z}}$ is $\frac{1}{(2d)!} (\partial_x^2 + \partial_y^2)^d$.

The following two examples will play a role in Section 4.2.

- (e) Combining (a) and (c), we know that the differential operator corresponding to $\mathcal{L}_A \in H_{n,2d}^\vee$, $A > 0$ defined as

$$\mathcal{L}_A: H_{n,2d} \rightarrow \mathbb{R}, \mathcal{L}_A(p) = A(p(e^1) + p(e^2)) - 2Q_p(e^1, e^2)$$

is given by

$$\mathcal{L}_A = \frac{1}{(2d)!} \left(A \left(\partial_{\mathbf{x}^{2d}} + \partial_{\mathbf{y}^{2d}} \right) - 2\partial_{\mathbf{x}^d} \partial_{\mathbf{y}^d} \right).$$

- (f) Consider the differential operator $\mathcal{L}_B \in H_{n,2d}^\vee$ given by

$$\mathcal{L}_B: H_{n,2d} \rightarrow \mathbb{R}, \mathcal{L}_B(p) = BQ_p(e^1 + ie^2, e^1 - ie^2) - \text{Re}(p(e^1 + ie^2)).$$

Analogously as in (d), \mathcal{L}_B can be written as a differential operator as

$$\mathcal{L}_B = \frac{1}{(2d)!} \left(B \left(\partial_{\mathbf{x}^2} + \partial_{\mathbf{y}^2} \right)^d - \text{Re} \left((\partial_{\mathbf{x}} + i\partial_{\mathbf{y}})^{2d} \right) \right).$$

3.4. Convex Cones and their Duals. For studying the sets $C_{n,2d}$ of convex and $\Sigma_{n,2d}$ of SOS forms, it is important to understand their structure as convex cones. Therefore, in this section we first introduce some basic definitions including convex cones as well as convex and conical hulls of sets of points. Those definitions among further details can for example be found in [8, Section 2]. Afterwards, we see an important result that goes back to [31], which helps us to write convex cones as conical hull of some of its points. This will be used in Section 5. We further introduce the duals of convex cones, which will help us in Section 4 to obtain constants for some generalized Cauchy–Schwarz inequalities.

Convention 3.60. In this section, V is always a finite dimensional \mathbb{R} -vector space.

Definition 3.61. A subset $C \subseteq V$ is **convex** if it is closed under **convex combinations**, i.e.

$$\forall x, y \in C \forall \lambda \in [0, 1] : \quad \lambda x + (1 - \lambda)y \in C.$$

Definition 3.62. A subset $C \subseteq V$ of V is a **cone** in V , if it is closed under the multiplication with nonnegative scalars, i.e. for all $c \in C$ and all $\alpha \geq 0$, we have $\alpha c \in C$.

Remark 3.63. A cone $C \subseteq V$ in V is convex if and only if for all $c, d \in C$ and $\alpha, \beta \geq 0$, it holds $\alpha c + \beta d \in C$.

Definition 3.64. A (convex) cone $C \subseteq V$ is **closed** if it is closed with respect to the Euclidean topology (see Definition 3.51).

Remark 3.65. We know that the Euclidean topology on $H_{n,2d}$ is the topology generated by any norm on $H_{n,2d}$. Hence, topological properties like being closed can always be verified by using the norm generated by the Fischer inner product.

Definition 3.66. The **convex hull** of a subset $S \subseteq V$ of V is given by

$$\text{conv}(S) = \left\{ \sum_{i=1}^n \lambda_i x^i : n \in \mathbb{N}, \lambda_i \geq 0, x_i \in S \ (i = 1, \dots, n), \sum_{i=1}^n \lambda_i = 1 \right\}.$$

Equivalently (see e.g. [31, Cor. 2.3.1]), the convex hull of S can be defined as the smallest convex set containing S , i.e.

$$\text{conv}(S) = \bigcap_{C \supseteq S \text{ convex}} C.$$

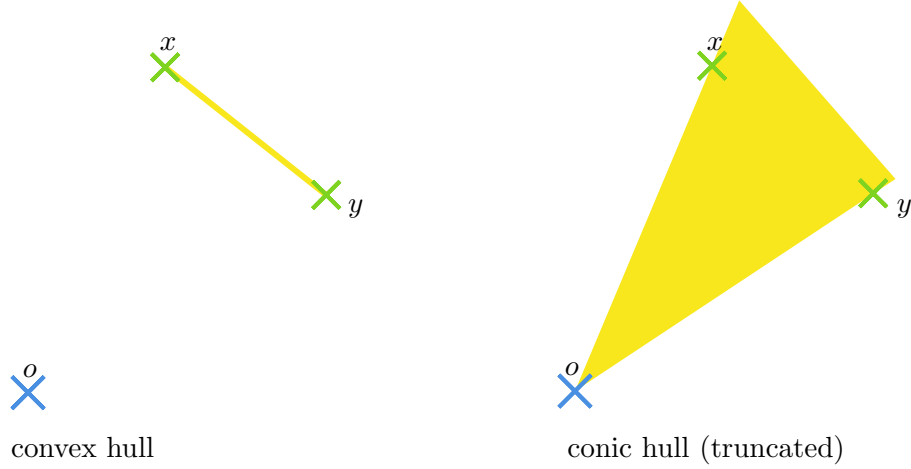
Further, the **conical hull** of S is

$$\text{cone}(S) = \left\{ \sum_{i=1}^n \lambda_i x^i : n \in \mathbb{N}, \lambda_i \geq 0, x_i \in S \ (i = 1, \dots, n) \right\}$$

Equivalently (see e.g. [31, Cor. 2.6.1]), the conical hull of S can be defined as the smallest convex cone containing S , i.e.

$$\text{cone}(S) = \bigcap_{C \supseteq S \text{ convex cone}} C.$$

Remark 3.67. From Definition 3.66 it is clear that the convex hull of a set S is always contained in the conical hull of S , i.e. $\text{conv}(S) \subseteq \text{cone}(S)$. However, the conical hull is in general strictly larger than the convex hull. Consider for example the convex and conical hull of two points x, y in \mathbb{R}^2 :



Remark 3.68. In [29, Lemma 3.7 and Theorem 3.10] Reznick shows that the sets $P_{n,2d}$, $\Sigma_{n,2d}$ and $C_{n,2d}$ are all closed convex cones in the finite dimensional \mathbb{R} -vector space $H_{n,2d}$.

Next, we introduce the notions of extreme points, directions and rays of convex cones. Those are all defined in a more general sense for convex sets instead of just convex cones. For more details, the reader is referred to [8, Section 2.6, 2.8.1] and [31, Section 18].

Definition 3.69. Let $C \subseteq V$ be a convex set. A point $x \in C$ is an **extreme point** of C if it cannot be written as a strict convex combination of two distinct points in C , i.e.

$$\forall y, z \in C \forall \lambda \in (0, 1) : \quad x = \lambda y + (1 - \lambda)z \implies x = y = z.$$

Proposition 3.70. Let $C \subseteq V$ be a convex cone. Then the only extreme point of C is the origin.

Proof. This is trivial since we can write every $0 \neq x \in C$ as a strict convex combination $x = \frac{1}{2}y + \frac{1}{2}z$, where $y := \frac{1}{2}x, z := \frac{3}{2}x \in C$ since C is a cone. \square

Since extreme points are not of much use in a convex cone, one starts to look at half lines emanating from the origin that again may not be written as strict convex combination of two other half lines. This leads to extreme rays as in [31, Section 18] or [8, Section 2.8.1].

Definition 3.71. Let $C \subseteq V$ be a convex set. A **ray** of C is a subset $\mathcal{R} \subseteq C$ of the form $\mathcal{R} = \{\mu\delta_{\mathcal{R}} : \mu \geq 0\}$, where $\delta_{\mathcal{R}} \in V \setminus \{o\}$. A ray \mathcal{R} of C is an **extreme ray** if it cannot be written as a strict convex combination of two other rays, i.e.

$$\begin{aligned} & \forall \delta_1, \delta_2 \in V \setminus \{o\} \forall \lambda \in (0, 1) : \\ & \mathcal{R} = \lambda\{\mu\delta_1 : \mu \geq 0\} + (1 - \lambda)\{\mu\delta_2 : \mu \geq 0\} \\ & \implies \mathcal{R} = \{\mu\delta_1 : \mu \geq 0\} = \{\mu\delta_2 : \mu \geq 0\} \end{aligned}$$

Instead of a whole ray, it can be useful to only consider the directions of a cone C . In that context, a direction being extreme leads to conical combinations instead of convex combinations since directions are geometrically invariant under scaling with positive scalars.

Definition 3.72. Let $C \subseteq V$ be a convex set. A vector $\delta \in V \setminus \{o\}$ is a **direction** of C if $\mathcal{R}_\delta = \{\mu\delta : \mu \geq 0\} \subseteq C$. If δ is a direction of C and \mathcal{R} a ray of C such that $\mathcal{R} = \mathcal{R}_\delta$, we say that δ **spans** \mathcal{R} .

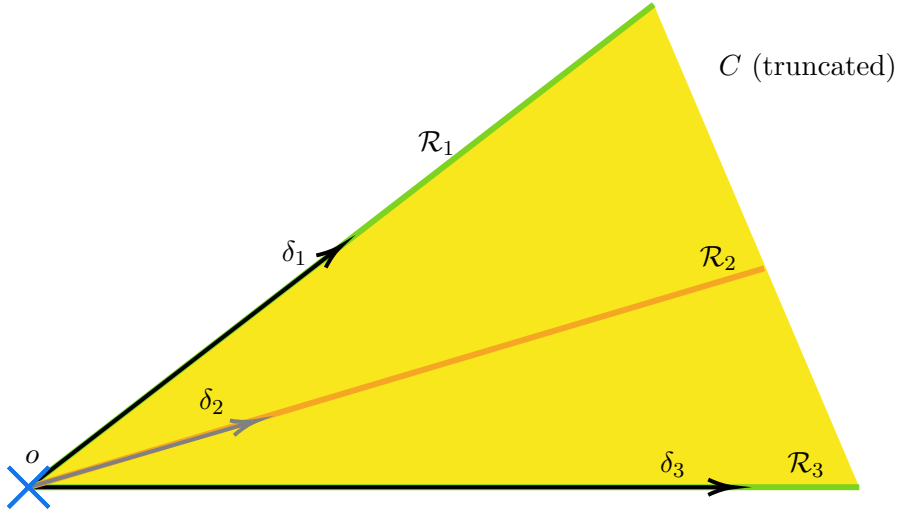
A direction δ of C is an **extreme direction** if it cannot be written as strict conical combination of elements in C , i.e.

$$\forall \delta_1, \delta_2 \in C \setminus \{o\} \quad \forall \mu_1, \mu_2 \geq 0 : \quad \delta = \mu_1\delta_1 + \mu_2\delta_2 \implies \delta_1, \delta_2 \in \{\mu\delta : \mu \geq 0\}.$$

Remark 3.73. A direction $\delta \in C \setminus \{o\}$ of a convex set C is extreme if and only if the whole ray \mathcal{R}_δ spanned by δ is extreme.

Example 3.74. Consider the following convex cone C in \mathbb{R}^2 , where

- the origin o is the only extreme point of C ,
- $\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3$ are rays of C spanned by $\delta_1, \delta_2, \delta_3$, respectively,
- $\mathcal{R}_1, \mathcal{R}_3$ are extreme rays and δ_1, δ_3 extreme directions, respectively,
- \mathcal{R}_2 is not extreme since it can be written as convex combination of \mathcal{R}_1 and \mathcal{R}_3 ,
- δ_2 is not extreme since it can be written as conical combination of δ_1 and δ_3 .



Further, we have the following canonical representation of convex sets as convex hull of some of their points that can be found in [31, Theorem 18.5].

Theorem 3.75. Let $C \subseteq V$ be a closed convex set containing no lines (i.e. subspaces of V of dimension 1). Let $\mathcal{E} \subseteq C$ be the set of all extreme points and extreme directions of C .

Then $C = \text{conv}(\mathcal{E})$ is the convex hull of \mathcal{E} .

Proof. [31, Theorem 18.5]. □

For convex cones, we have the following corollary, which will be used in Section 5 to characterize the cone of convex polynomials $C_{n,2d}$ inside the SOS cone $\Sigma_{n,2d}$ for $(n, 2d) = (4, 4)$.

Corollary 3.76. Let $C \subseteq V$ be a closed convex cone. Then C is the convex hull of its extreme directions.

Proof. This follows directly from Theorem 3.75 since convex cones have no extreme points beyond the origin by Proposition 3.70. \square

Remark 3.77. From what we know, we can also deduce that a closed convex cone C is the convex hull of its extreme rays or the conical hull of its extreme directions with a fixed length, i.e.

$$\begin{aligned} C &= \text{conv}\{\delta \in C \setminus \{o\} : \delta \text{ is an extreme direction}\} \\ &= \text{conv}\{\delta \in \mathcal{R} : \mathcal{R} \subseteq C \text{ is an extreme ray}\} \\ &= \text{cone}\{\delta \in C \setminus \{o\} : \delta \text{ is an extreme direction, } \|\delta\| = c\}, \end{aligned}$$

where $c > 0$ is arbitrary.

Next, we introduce the dual of convex cones, which will be important in Section 4 for optimization purposes when we try to obtain optimal constants in the generalized Cauchy–Schwarz inequalities from Theorem 4.3.

Definition 3.78. Let $C \subseteq V$ be a cone in V . Then its **dual cone** $C^* \subseteq V^\vee$ is defined as

$$C^* = \{\ell \in V^\vee \mid \forall c \in C : \ell(c) \geq 0\}. \quad (16)$$

Remark 3.79. Let V be equipped with an inner product $(\cdot, \cdot) : V \times V \rightarrow \mathbb{R}$ and $C \subseteq V$ a cone. Using the Riesz Representation Theorem 3.54 it can easily be seen that as in [5, Equation (2.19)] or [8, Equation (304)] the dual C^* can equivalently be defined as

$$C^* = \{v \in V \mid \forall u \in C : (u, v) \geq 0\}.$$

In Proposition 3.80 and 3.82, some well known result in the theory of dual cones are collected without proof. For more details the reader is referred to [5, Chapter 2].

Proposition 3.80. *Let V be a finite dimensional \mathbb{R} -vector space, $C \subseteq V$ a cone. Then C^* is a closed and convex cone in V^\vee with respect to the Euclidean topology, as V^\vee is finite dimensional by Corollary 3.55.*

Hence, it makes sense to define the bidual of a cone.

Definition 3.81. Let V be a finite dimensional \mathbb{R} -vector space and $C \subseteq V$ a cone. Then we define the **bidual cone** $C^{**} \subseteq V^\vee$ as

$$C^{**} = \{v \in V \mid \forall \ell \in C^* : \ell(v) \geq 0\}. \quad (17)$$

Proposition 3.82. *Let V be a finite dimensional \mathbb{R} -vector space, $C \subseteq V$ a cone. Then $C^{**} \subseteq V$ is the closure of the convex hull of C . Hence, if C is closed and convex, then $C^{**} = C$.*

We already know that the conical hull $\text{cone}(S)$ of a set S is always convex. Further, by Proposition 3.82, we know $\text{cone}(S)^{**} = \text{cone}(S)$ holds, whenever $\text{cone}(S)$ is closed. For this reason, consider the following criterion on $\text{cone}(S)$ to be closed.

Proposition 3.83. *Let $S \subseteq V$ be a nonempty, compact (i.e. closed and bounded) subset of V not containing the origin. Then the conical hull $\text{cone}(S)$ of S is closed.*

Proof. [18, Proposition 1.4.7] \square

In Remark 3.68, we have already seen that $P_{n,2d}$, $\Sigma_{n,2d}$ and $C_{n,2d}$ are all closed convex cones in $H_{n,2d}$ (see e.g. [29, Thm. 3.10]). Hence, we have in particular $C_{n,2d}^{**} = C_{n,2d}$. Using this fact, we can get in Lemma 3.84 a convenient representation of the dual $C_{n,2d}^*$, which is stated without proof in [12, Section 2.6].

Lemma 3.84. *Let $n, d \in \mathbb{N}$ be arbitrary. Then it holds*

$$C_{n,2d}^* = \text{cone}(\{\ell_{xy} : x, y \in \mathbb{R}^n\})$$

where the linear function $\ell_{xy} \in H_{n,2d}^\vee$, $x, y \in \mathbb{R}^n$ is defined via

$$\ell_{xy} : H_{n,2d} \rightarrow \mathbb{R}, p \mapsto \ell_{xy}(p) = y^\top \nabla^2 p(x) y.$$

Proof. First, let $x, y \in \mathbb{R}^n$ be arbitrary and take $\tilde{x}, \tilde{y} \in \mathbb{R}^n$ such that it holds $x = \|x\|_2 \tilde{x}$, $y = \|y\|_2 \tilde{y}$. For $p \in H_{n,2d}$ arbitrary, we have

$$\ell_{xy}(p) = y^\top \nabla^2 p(x) y = \|y\|_2^2 \|x\|_2^{2d-2} \tilde{y}^\top \nabla^2 p(\tilde{x}) \tilde{y} = \underbrace{\|y\|_2^2 \|x\|_2^{2d-2}}_{\geq 0} \ell_{\tilde{x}\tilde{y}}(p).$$

This shows

$$\text{cone}(\{\ell_{xy} : x, y \in \mathbb{R}^n\}) = \text{cone}(\{\ell_{xy} : x, y \in \mathbb{R}^n, \|x\|_2 = \|y\|_2 = 1\}).$$

Hence, it follows from Proposition 3.83 that $\text{cone}(\{\ell_{xy} : x, y \in \mathbb{R}^n\})$ is a closed and convex cone.

For this reason, it suffices to show that $\text{cone}(\{\ell_{xy} : x, y \in \mathbb{R}^n\})^* = C_{n,2d}$. The claim then follows by taking the dual and applying Proposition 3.82. Indeed, we have

$$\begin{aligned} & \text{cone}(\{\ell_{xy} : x, y \in \mathbb{R}^n\})^* \\ &= \left\{ p \in H_{n,2d} \mid \forall m \in \mathbb{N} \forall_{i=1}^m \lambda_i \geq 0, x^i, y^i \in \mathbb{R}^n : \sum_{i=1}^m \lambda_i \ell_{x^i y^i}(p) \geq 0 \right\} \\ &= \{p \in H_{n,2d} \mid \forall x, y \in \mathbb{R}^n : \ell_{xy}(p) \geq 0\} \\ &= C_{n,2d}, \end{aligned}$$

where the last step follows by Definition 2.16 of $C_{n,2d}$. \square

In Section 3.3 we have seen that the elements of $H_{n,2d}^\vee$ can be interpreted as differential operators. Therefore, it is reasonable to look for a representation of $C_{n,2d}^*$ in terms of differential operators. To do so, consider the following result from [29, Cor. 2.10].

Proposition 3.85. *For $p \in H_{n,2d}$ and $x, y \in \mathbb{R}^n$ arbitrary, we have*

$$y^\top \nabla^2 p(x) y = \ell_{xy}(p) = \frac{1}{(2d-2)!} \cdot \left\langle \left(\sum_{i=1}^n x_i \mathbf{x}_i \right)^{2d-2} \left(\sum_{i=1}^n y_i \mathbf{x}_i \right)^2, p \right\rangle. \quad (18)$$

Proof. [29, Cor. 2.10]. \square

Remark 3.86. Note that the factor on the right hand side of (18) differs from the one in [29] since the author there uses a 'scaled' Fischer inner product (see e.g [29, Prop. 2.9]). However, since in Lemma 3.84, we are

taking the conical hull of the linear functionals ℓ_{xy} , a positive scalar does not change the hull.

Using Proposition 3.85, we can get a useful representations of $C_{n,2d}^*$ and in particular $C_{2,2d}^*$ which are stated without proof in [12, Section 2.6].

Corollary 3.87. *For $n, d \in \mathbb{N}$, it holds*

$$C_{n,2d}^* = \text{cone} \left\{ \partial_x^{2d-2} \partial_y^2 : x, y \in \mathbb{R}^n \right\}.$$

Proof. This follows from Lemma 3.84 using the formula in (18), since for $x, y \in \mathbb{R}^n$ and $p \in H_{n,2d}$, we have

$$\begin{aligned} & \left\langle \left(\sum_{i=1}^n x_i \mathbf{x}_i \right)^{2d-2} \left(\sum_{i=1}^n y_i \mathbf{x}_i \right)^2, p \right\rangle \\ &= \left[\left(\sum_{i=1}^n x_i \partial_{\mathbf{x}_i} \right)^{2d-2} \cdot \left(\sum_{i=1}^n y_i \partial_{\mathbf{x}_i} \right)^2 \right] p \end{aligned}$$

Now write $p = \sum_{i_1, \dots, i_{2d}=1, \dots, n} \tilde{p}^{i_1, \dots, i_{2d}} \mathbf{x}_{i_1} \cdots \mathbf{x}_{i_{2d}}$, where the coefficients $\tilde{p}^{j_1, \dots, j_{2d}} \in \mathbb{R}$ are the coefficients from the associated tensor T_p as in Theorem 3.29. Using (13) from the proof of Lemma 3.57, we can deduce that

$$\begin{aligned} & \left[\left(\sum_{i=1}^n x_i \partial_{\mathbf{x}_i} \right)^{2d-2} \cdot \left(\sum_{i=1}^n y_i \partial_{\mathbf{x}_i} \right)^2 \right] p \\ &= \sum_{j_1, \dots, j_{2d}=1, \dots, n} \left(x_{j_1} \cdots x_{j_{2d-2}} y_{j_{2d-1}} y_{j_{2d}} \right. \\ & \quad \left. \underbrace{\partial_{\mathbf{x}_{j_1}} \cdots \partial_{\mathbf{x}_{j_{2d}}} \sum_{i_1, \dots, i_{2d}=1, \dots, n} \tilde{p}^{i_1, \dots, i_{2d}} \mathbf{x}_{i_1} \cdots \mathbf{x}_{i_{2d}}}_{=\tilde{p}^{j_1, \dots, j_{2d}}} \right) \end{aligned}$$

This yields

$$\begin{aligned} & \left\langle \left(\sum_{i=1}^n x_i \mathbf{x}_i \right)^{2d-2} \left(\sum_{i=1}^n y_i \mathbf{x}_i \right)^2, p \right\rangle \\ &= \sum_{j_1, \dots, j_{2d}=1, \dots, n} \tilde{p}^{j_1, \dots, j_{2d}} x_{j_1} \cdots x_{j_{2d-2}} y_{j_{2d-1}} y_{j_{2d}} \stackrel{(13)}{=} T_p \left(\underbrace{x, \dots, x}_{(2d-2) \times}, y, y \right). \end{aligned}$$

Finally, using Lemma 3.57, we obtain

$$\left\langle \left(\sum_{i=1}^n x_i \mathbf{x}_i \right)^{2d-2} \left(\sum_{i=1}^n y_i \mathbf{x}_i \right)^2, p \right\rangle = T_p \left(\underbrace{x, \dots, x}_{(2d-2) \times}, y, y \right) = \frac{1}{(2d)!} \partial_x^{2d-2} \partial_y^2 p,$$

which shows the claim. \square

Corollary 3.88. *For $d \in \mathbb{N}$, it holds*

$$\begin{aligned} C_{2,2d}^* &= \text{cone} \left\{ \partial_x^{2d-2} \partial_y^2 : x, y \in \mathbb{R}^2 \right\} \\ &= \left\{ \sum_{i=1}^m (\alpha_i \partial_{\mathbf{x}} + \beta_i \partial_{\mathbf{y}})^2 (\gamma_i \partial_{\mathbf{x}} + \delta_i \partial_{\mathbf{y}})^{2d-2} : \alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R} \right\} \end{aligned} \quad (19)$$

Proof. This follows directly from Corollary 3.87 using the definition of the conical hull and formula (1) for the directional derivative of a polynomial. \square

4. GENERALIZED CAUCHY–SCHWARZ INEQUALITIES

The Cauchy–Schwarz Inequality is usually stated in terms of inner products or equivalently by using positive definite matrices as in Theorem 4.1. In Theorem 4.2, there is another, more general version of the Cauchy–Schwarz Inequality, which is stated in terms of positive semidefinite matrices.

Further, in Section 4.1 we state at first in Theorem 4.3 a third version of the Cauchy–Schwarz Inequality, which is formulated for convex forms. It turns out that Theorem 4.2 is a special case of Theorem 4.3, which justifies that we call Theorem 4.3 the Generalized Cauchy–Schwarz Inequalities.

Theorem 4.1 (Cauchy–Schwarz Inequality I). *Let V be a \mathbb{K} -vector space with inner product (\cdot, \cdot) , $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. For $x, y \in V$ arbitrary it holds*

$$|(x, y)| \leq \|x\| \cdot \|y\|.$$

Proof. [9, Satz 5.49]. \square

In fact, we can also state Theorem 4.1 in a more general sense in terms of positive semidefinite matrices, which is stated without proof in [12, Section 3].

Theorem 4.2 (Cauchy–Schwarz Inequalities II). *Let $Q \in \mathbb{R}^{n \times n}$ be positive semidefinite. Then*

$$\forall x, y \in \mathbb{R}^n : x^\top Q y \leq \sqrt{x^\top Q x \cdot y^\top Q y}. \quad (20)$$

Further, if x, y is a pair of complex conjugated vectors, the inequality reverses, i.e.

$$\forall z \in \mathbb{C}^n : z^\top Q \bar{z} \geq \sqrt{z^\top Q z \cdot \bar{z}^\top Q \bar{z}}. \quad (21)$$

Proof. First, let $x, y \in \mathbb{R}^n$ be arbitrary. Since Q is positive semidefinite there is a Cholesky decomposition $Q = L^\top L$, where $L \in \mathbb{R}^{n \times n}$ (see [32, Satz 4.30]). Hence, using the regular Cauchy–Schwarz Inequality from Theorem 4.1 shows

$$\begin{aligned} (x^\top Q y)^2 &= ((Lx)^\top \cdot (Ly))^2 = (Lx, Ly)_2^2 \leq \|Lx\|_2^2 \|Ly\|_2^2 \\ &= (Lx)^\top Lx (Ly)^\top Ly = x^\top Q x y^\top Q y. \end{aligned}$$

Taking the root on both sides shows (20).

Now, let $z \in \mathbb{C}^n$ be arbitrary. Write $z = x + iy$, where $x, y \in \mathbb{R}^n$. We have

$$z^\top Q \bar{z} = (x + iy)^\top Q (x - iy) = x^\top Q x + y^\top Q y. \quad (22)$$

Now take a look at the right hand side of (21). First, one can see that $\sqrt{z^\top Qz \cdot \bar{z}^\top Q\bar{z}} = |z^\top Qz|$. Further, it holds

$$z^\top Qz = (x + iy)^\top Q(x + iy) = \underbrace{x^\top Qx - y^\top Qy}_{=\text{Re}(z^\top Qz)} + i \cdot \underbrace{2x^\top Qy}_{=\text{Im}(z^\top Qz)}.$$

This leads to

$$\begin{aligned} z^\top Qz \cdot \bar{z}^\top Q\bar{z} &= \left| z^\top Qz \right|^2 = \text{Re}(z^\top Qz)^2 + \text{Im}(z^\top Qz)^2 \\ &= (x^\top Qx - y^\top Qy)^2 + 4(x^\top Qy)^2 \\ &= (x^\top Qx)^2 - 2x^\top Qxy^\top Qy + (y^\top Qy)^2 + 4(x^\top Qy)^2 \\ &\stackrel{(20)}{\leq} (x^\top Qx)^2 - 2x^\top Qxy^\top Qy + (y^\top Qy)^2 + 4x^\top Qxy^\top Qy \\ &= (x^\top Qx + y^\top Qy)^2 \\ &\stackrel{(22)}{=} (z^\top Q\bar{z})^2. \end{aligned}$$

Again, taking the root on both sides shows (21). \square

Theorem 4.2 should be kept in mind, while studying the Generalized Cauchy–Schwarz Inequalities in the following section.

4.1. Formulation and Proof of the Generalized Cauchy–Schwarz Inequalities.

The goal of this section is to prove the following theorem.

Theorem 4.3 (Generalized Cauchy–Schwarz Inequalities). *Let $p \in C_{n,2d}$ be an arbitrary convex form in n variables of degree $2d$ and $Q_p: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ the associated biform. Then there are constants $A_d, B_d > 0$ that depend only on the degree $2d$ such that the following holds*

$$\forall x, y \in \mathbb{R}^n : Q_p(x, y) \leq A_d \cdot \sqrt{p(x) \cdot p(y)}. \quad (23)$$

Further, in the complex case, we have

$$\forall z \in \mathbb{C}^n : |p(z)| \leq B_d \cdot Q_p(z, \bar{z}). \quad (24)$$

At first, we want to check that Theorem 4.3 is indeed a generalization of Theorem 4.2. Therefore, consider the following proposition.

Proposition 4.4. *Let $Q \in \mathbb{R}^{n \times n}$ be symmetric and consider the corresponding quadratic form $p_Q(\mathbf{x}) = \mathbf{x}^\top Q\mathbf{x} \in H_{n,2}$. Then Q is positive semidefinite if and only if p_Q is convex, i.e. $p_Q \in C_{n,2}$.*

Proof. This is trivial, since by the symmetry of Q we have for the gradient $\nabla p_Q(\mathbf{x}) = 2Q\mathbf{x}$ and hence for the Hessian $\nabla^2 p_Q(\mathbf{x}) = 2Q$. Hence, p_Q being convex coincides with Q being positive semidefinite. \square

Remark 4.5. Note that the tensor associated to p_Q defined as in Proposition 4.4 is given by $T_{p_Q}: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, T_{p_Q}(x^1, x^2) = (x^1)^\top Qx^2$ (see Theorem 3.29). Hence, the biform associated to p_Q is given by $Q_{p_Q}(x, y) = x^\top Qy$ ($x, y \in \mathbb{R}^n$). For this reason, by using Proposition 4.4 it is clear that Theorem 4.2 is just a special case of Theorem 4.3 where $2d = 2, p = p_Q$ and the constants A_d, B_d are equal to 1. In Proposition 4.12 and 4.29 of Section 4.2 we will see that the smallest possible constants A_1, B_1 in Theorem 4.3 can indeed be chosen to be 1.

Our goal is now to prove the Generalized Cauchy–Schwarz Inequalities. Therefore, we first linearize inequalities (23) and (24) and then show that it suffices to consider the inequalities in just two variables, i.e. for bivariate convex forms.

For the first step, consider the following lemma, whose proof can be found in [12, Section 3.1].

Lemma 4.6. *Let $p \in C_{n,2d}$ be an arbitrary convex form in n variables of degree $2d$ and $Q_p: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ the associated biform. Further, let $A_d, B_d > 0$ be arbitrary. Then p satisfies (23) with constant $A_d > 0$ if and only if*

$$\forall x, y \in \mathbb{R}^n : 2Q_p(x, y) \leq A_d(p(x) + p(y)). \quad (25)$$

Further, in the complex case p satisfies (24) with constant $B_d > 0$ if and only if

$$\forall z \in \mathbb{C}^n : \operatorname{Re}(p(z)) \leq B_d Q_p(z, \bar{z}). \quad (26)$$

Proof. We have to show that (23) is equivalent to (25) and (24) is equivalent to (26), respectively.

(i) Claim: (23) \iff (25).

" \implies ": Assume that (23) holds. Let $x, y \in \mathbb{R}^n$ be arbitrary. By Young's inequality (see e.g. [9, Satz 10.2]) we know that $\sqrt{ab} \leq \frac{a+b}{2}$ for $a, b \in \mathbb{R}_{\geq 0}$. This shows

$$\begin{aligned} 2Q_p(x, y) &\stackrel{(23)}{\leq} 2A_d \sqrt{p(x)p(y)} \stackrel{p(x), p(y) \geq 0}{\leq} 2A_d \left(\frac{p(x) + p(y)}{2} \right) \\ &= A_d(p(x) + p(y)), \end{aligned}$$

as desired.

" \impliedby ": Assume that (25) holds and let $x, y \in \mathbb{R}^n$ be arbitrary. If $Q_p(x, y)$ is lower than or equal to zero, it is clear that (23) holds. If $p(y)$ is zero, then we obtain from (25) for an arbitrary $\varepsilon > 0$

$$2\varepsilon Q_p(x, y) = 2Q_p(\varepsilon^{1/d}x, y) \leq A_d \left(p\left(\varepsilon^{1/d}x\right) + \underbrace{p(y)}_{=0} \right) \stackrel{\deg(p)=2d}{=} A_d \varepsilon^2 p(x).$$

Dividing both sides by $\varepsilon > 0$ shows

$$2Q_p(x, y) \leq A_d \varepsilon p(x).$$

Since $\varepsilon > 0$ was arbitrary, it can be deduced that $Q_p(x, y) \leq 0$ and therefore that (23) holds. An analogous observation can be made if $p(x)$ was zero. Hence, using $p \in P_{n,2d}$ we can without loss of generality assume that $p(x), p(y) > 0$ and $Q_p(x, y) > 0$.

By (25) we know that for all $\lambda \in \mathbb{R}_{\geq 0}$, it holds

$$\begin{aligned} 2Q_p(x, \lambda^{1/d}y) &\stackrel{(25)}{\leq} A_d(p(x) + p(\lambda^{1/d}y)) \\ \Leftrightarrow 2\lambda Q_p(x, y) &\leq A_d(p(x) + \lambda^2 p(y)) \\ \Leftrightarrow f(\lambda) := A_d(p(x) + \lambda^2 p(y)) - 2\lambda Q_p(x, y) &\text{ is nonnegative on } [0, \infty). \end{aligned}$$

Clearly, $f(\lambda) = A_d p(y) \lambda^2 - 2Q_p(x, y) \lambda + A_d p(x)$ is a univariate polynomial of degree 2. Hence, it has two possibly complex conjugated roots $\lambda_1, \lambda_2 \in \mathbb{C}$. Further, we know that $\lambda_{1,2}$ can be computed via

$$\lambda_{1,2} = \underbrace{\frac{Q_p(x, y)}{A_d p(y)}}_{>0} \pm \left(\underbrace{\left(\frac{Q_p(x, y)}{A_d p(y)} \right)^2 - \frac{p(x)}{p(y)}}_{=: \delta} \right)^{1/2}.$$

If the discriminant δ was greater than zero, $f(\lambda)$ would have two distinct real roots and since $\delta < \left(\frac{Q_p(x, y)}{A_d p(y)} \right)^2$ in particular two changes of sign on $\mathbb{R}_{\geq 0}$. This would contradict the fact that $f(\lambda)$ is nonnegative on $[0, \infty)$. Hence, we must have $\delta \leq 0$. Rearranging then shows that (23) holds.

(ii) Claim: (24) \iff (26).

" \implies ": This is trivial since for $\alpha \in \mathbb{C}$ arbitrary, it always holds

$$\operatorname{Re}(\alpha) \leq \sqrt{\operatorname{Re}(\alpha)^2 + \operatorname{Im}(\alpha)^2} = |\alpha|.$$

Hence, if (24) holds, then we have in particular for all $z \in \mathbb{C}$

$$\operatorname{Re}(p(z)) \leq |p(z)| \stackrel{(24)}{\leq} B_d Q_p(z, \bar{z}),$$

which is (26).

" \impliedby ": Assume that (26) holds and let $z \in \mathbb{C}^n$ be arbitrary. We know that we can write $p(z) = |p(z)| \cdot e^{i\varphi}$, where $\varphi = \arg(p(z)) \in [0, \pi)$.

We now want to choose a point $\tilde{z} = e^{i\psi} z$ where ψ is chosen such that $|p(z)| = p(\tilde{z})$ and hence in particular $|p(z)| = \operatorname{Re}(p(\tilde{z}))$.

This leads to $\psi = -\frac{\varphi}{2d} = -\frac{\arg(p(z))}{2d}$. Now we have

$$p(\tilde{z}) = p\left(e^{-i\frac{\varphi}{2d}} z\right) \stackrel{p \in H_{n,2d}}{=} \frac{1}{e^{i\varphi}} \cdot p(z) = |p(z)|.$$

Hence, we have in particular $\operatorname{Re}(p(\tilde{z})) = \operatorname{Re}(|p(z)|) = |p(z)|$. Further, one can see that $e^{i\psi} = \cos \psi - i \sin \psi = \cos(-\psi) + i \sin(-\psi) = e^{-i\psi}$. This shows

$$\begin{aligned} Q_p(\tilde{z}, \tilde{\bar{z}}) &= Q_p\left(e^{i\psi} z, e^{i\psi} \bar{z}\right) = Q_p\left(e^{i\psi} z, e^{-i\psi} \bar{z}\right) \\ &= \left(e^{i\psi}\right)^d \left(e^{-i\psi}\right)^d Q_p(z, \bar{z}) = Q_p(z, \bar{z}). \end{aligned}$$

For this reason, using (26) for \tilde{z} leads to

$$|p(z)| = \operatorname{Re}(p(\tilde{z})) \stackrel{(26)}{\leq} B_d \sqrt{Q_p(\tilde{z}, \tilde{\bar{z}})} = B_d \sqrt{Q_p(z, \bar{z})},$$

which is (24). □

Hence, to prove Theorem 4.3, it suffices to prove the two conditions (25) and (26) from Lemma 4.6. We call those conditions linearized, since both sides of the inequalities in (25) and (26) are linear in p . Recall that the mapping $H_{n,k} \rightarrow S^k(\mathbb{R}^n), p \mapsto T_p$ is linear by Theorem 3.29 and so is $p \mapsto Q_p$.

The next step is to show that it suffices to show the Generalized Cauchy-Schwarz Inequalities in just two variables. A sketch of the proof of the

following lemma can be found in [12, Section 3.1], while details are added in the following.

Lemma 4.7. *Let $A_d, B_d > 0$ be arbitrary. Then condition (25) holds for all convex $p \in C_{n,2d}$ and all $n \in \mathbb{N}$ with constant $A_d > 0$ if and only if for all bivariate, convex $q \in C_{2,2d}$ it holds*

$$2Q_q(e^1, e^2) \leq A_d (q(e^1) + q(e^2)). \quad (27)$$

Further, condition (26) holds for all convex $p \in C_{n,2d}$ and all $n \in \mathbb{N}$ with constant $B_d > 0$ if and only if for all bivariate, convex $q \in C_{2,2d}$, it holds

$$\operatorname{Re}(q(e^1 + ie^2)) \leq B_d Q_q(e^1 + ie^2, e^1 - ie^2). \quad (28)$$

Proof. Obviously, the conditions (27) and (28) are just special cases of (25) and (26), respectively, where the number n of variables is two. Hence, if (25) or (26) hold for all $n \in \mathbb{N}$, it is clear that (27) or (28) hold as well, respectively. In the following, we show that the converse is true as well.

- (i) Assume that (27) holds for all bivariate, convex forms. Let $n \in \mathbb{N}$ and $p \in C_{n,2d}$ be arbitrary.

To show: (25) holds with constant $A_d > 0$.

Let $x, y \in \mathbb{R}^n$ be arbitrary. Define a bivariate form $q \in \mathbb{R}[\alpha, \beta]$ as $q(\alpha, \beta) := p(\alpha x + \beta y)$. Clearly, q evaluated at the unit vectors leads to $q(e^1) = p(x)$, $q(e^2) = p(y)$.

Claim: It also holds $Q_p(x, y) = Q_q(e^1, e^2)$.

On the one hand, we have

$$\begin{aligned} Q_p(x, y) &= T_p(x, \dots, x, y, \dots, y) \\ &= T_p \left(\underbrace{\sum_{i=1}^n x_i e^i, \dots, \sum_{i=1}^n x_i e^i, \sum_{i=1}^n y_i e^i, \dots, \sum_{i=1}^n y_i e^i}_{2d \times} \right) \\ &= \sum_{i_1, \dots, i_{2d}=1, \dots, n} x_{i_1} \cdots x_{i_d} y_{i_{d+1}} \cdots y_{i_{2d}} \cdot T_p(e^{i_1}, \dots, e^{i_{2d}}) \\ &= \sum_{i_1, \dots, i_{2d}=1, \dots, n} x_{i_1} \cdots x_{i_d} y_{i_{d+1}} \cdots y_{i_{2d}} \cdot \tilde{p}^{i_1, \dots, i_{2d}}, \end{aligned}$$

where the last step follows by using Corollary 3.35 and formula (4) for the permutation-invariant coefficients $\tilde{p}^{i_1, \dots, i_{2d}}$ of p . On the other hand, it holds

$$Q_q(e^1, e^2) = T_q \left(\underbrace{e^1, \dots, e^1}_{d \times}, \underbrace{e^2, \dots, e^2}_{d \times} \right) \stackrel{\text{Cor. 3.35}}{=} \binom{2d}{d}^{-1} q_{(d,d)},$$

where $q_{(d,d)}$ is the coefficient of q that belongs to $\alpha^d \beta^d$. By definition of q , we have

$$q(\alpha, \beta) = p(\alpha x + \beta y) = \sum_{i_1, \dots, i_{2d}=1, \dots, n} \tilde{p}^{i_1, \dots, i_{2d}} (\alpha x_{i_1} + \beta y_{i_1}) \cdots (\alpha x_{i_{2d}} + \beta y_{i_{2d}}).$$

This leads to

$$q_{(d,d)} = \sum_{i_1, \dots, i_{2d}=1, \dots, n} \left(\tilde{p}^{i_1, \dots, i_{2d}} \cdot \sum_{S \subseteq \{1, \dots, 2d\}, |S|=d} \left[\prod_{s \in S} x_{i_s} \prod_{r \in \{1, \dots, 2d\} \setminus S} y_{i_r} \right] \right).$$

Let $S \subseteq \{1, \dots, 2d\}, |S| = d$ be arbitrary and write $S = \{s_1, \dots, s_d\}$, $\{r_1, \dots, r_d\} = \{1, \dots, 2d\} \setminus S$. Define a permutation $\pi_S \in S_{2d}$ via $\pi_S(1) = s_1, \dots, \pi_S(d) = s_d, \pi_S(d+1) = r_1, \dots, \pi_S(2d) = r_d$. Using this, we obtain

$$q_{(d,d)} = \sum_{i_1, \dots, i_{2d}=1, \dots, n} \left(\sum_{S \subseteq \{1, \dots, 2d\}, |S|=d} \tilde{p}^{i_1, \dots, i_{2d}} \cdot \prod_{k=1}^d x_{i_{\pi_S(k)}} \prod_{\ell=d+1}^{2d} y_{i_{\pi_S(\ell)}} \right).$$

Since the coefficients $\tilde{p}^{i_1, \dots, i_{2d}}$ are invariant under permutations of indices, we further have

$$\begin{aligned} q_{(d,d)} &= \sum_{i_1, \dots, i_{2d}=1, \dots, n} \left(\sum_{S \subseteq \{1, \dots, 2d\}, |S|=d} \tilde{p}^{i_{\pi_S(1)}, \dots, i_{\pi_S(2d)}} \cdot \prod_{k=1}^d x_{i_{\pi_S(k)}} \prod_{\ell=d+1}^{2d} y_{i_{\pi_S(\ell)}} \right) \\ &= \sum_{S \subseteq \{1, \dots, 2d\}, |S|=d} \left(\sum_{i_1, \dots, i_{2d}=1, \dots, n} \tilde{p}^{i_{\pi_S(1)}, \dots, i_{\pi_S(2d)}} \cdot \prod_{k=1}^d x_{i_{\pi_S(k)}} \prod_{\ell=d+1}^{2d} y_{i_{\pi_S(\ell)}} \right) \end{aligned}$$

Finally, for $S \subseteq \{1, \dots, 2d\}, |S| = d$ arbitrary, we have

$$\begin{aligned} & \sum_{i_1, \dots, i_{2d}=1, \dots, n} \tilde{p}^{i_{\pi_S(1)}, \dots, i_{\pi_S(2d)}} \cdot \prod_{k=1}^d x_{i_{\pi_S(k)}} \prod_{\ell=d+1}^{2d} y_{i_{\pi_S(\ell)}} \\ &= \sum_{i_1=1}^n \cdots \sum_{i_{2d}=1}^n \tilde{p}^{i_{\pi_S(1)}, \dots, i_{\pi_S(2d)}} \cdot \prod_{k=1}^d x_{i_{\pi_S(k)}} \prod_{\ell=d+1}^{2d} y_{i_{\pi_S(\ell)}} \\ &= \sum_{i_{\pi_S(1)}=1}^n \cdots \sum_{i_{\pi_S(2d)}=1}^n \tilde{p}^{i_{\pi_S(1)}, \dots, i_{\pi_S(2d)}} \cdot \prod_{k=1}^d x_{i_{\pi_S(k)}} \prod_{\ell=d+1}^{2d} y_{i_{\pi_S(\ell)}} \\ &= \sum_{i_{\pi_S(1)}=1, \dots, i_{\pi_S(2d)}=1}^{n, \dots, n} \tilde{p}^{i_{\pi_S(1)}, \dots, i_{\pi_S(2d)}} \cdot \prod_{k=1}^d x_{i_{\pi_S(k)}} \prod_{\ell=d+1}^{2d} y_{i_{\pi_S(\ell)}}. \end{aligned}$$

Note that the last term is independent of the choice of S . Hence, since there are $\binom{2d}{d}$ possible choices of $S \subseteq \{1, \dots, 2d\}$ such that $|S| = d$, we obtain after relabeling the indices

$$q_{(d,d)} = \binom{2d}{d} \cdot \left(\sum_{i_1, \dots, i_{2d}=1, \dots, n} \tilde{p}^{i_1, \dots, i_{2d}} \cdot \prod_{k=1}^d x_{i_k} \prod_{\ell=d+1}^{2d} y_{i_\ell} \right)$$

and therefore

$$\begin{aligned} Q_q(e^1, e^2) &= \binom{2d}{d}^{-1} q_{(d,d)} \\ &= \sum_{i_1, \dots, i_{2d}=1, \dots, n} \tilde{p}^{i_1, \dots, i_{2d}} \cdot x_{i_1} \cdots x_{i_d} y_{i_{d+1}} \cdots y_{i_{2d}} = Q_p(x, y), \end{aligned}$$

as desired.

We have now shown that $p(x), p(y)$ and $Q_p(x, y)$ only depend on p through its 2-dimensional restriction q . Further, q is convex since for all $(\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{R}^2$ and $\lambda \in [0, 1]$, it holds

$$\begin{aligned} q(\lambda(\alpha_1, \beta_1) + (1 - \lambda)(\alpha_2, \beta_2)) &= p([\lambda\alpha_1 + (1 - \lambda)\alpha_2]x + [\lambda\beta_1 + (1 - \lambda)\beta_2]y) \\ &= p(\lambda[\alpha_1x + \beta_1y] + (1 - \lambda)[\alpha_2x + \beta_2y]) \\ &\leq \lambda p(\alpha_1x + \beta_1y) + (1 - \lambda)p(\alpha_2x + \beta_2y) \\ &= \lambda q(\alpha_1, \beta_1) + (1 - \lambda)q(\alpha_2, \beta_2). \end{aligned}$$

This shows $q \in C_{2,2d}$. Hence, using (27) yields

$$2Q_p(x, y) = 2Q_q(e^1, e^2) \stackrel{(27)}{\leq} A_d(q(e^1) + q(e^2)) = A_d(p(x) + p(y)),$$

which shows that (25) holds with constant $A_d > 0$, as desired.

- (ii) Assume that (28) holds for all bivariate, convex forms. Let $n \in \mathbb{N}$ and $p \in C_{n,2d}$ be arbitrary.

To show: (26) holds with constant $B_d > 0$.

Let $z \in \mathbb{C}^n$ be arbitrary and write $z = x + iy$, where $x, y \in \mathbb{R}^n$. As in (i), define $q \in \mathbb{R}[\alpha, \beta]$ via $q(\alpha, \beta) = p(\alpha x + \beta y)$. Further, one can define the complex polynomial $\hat{q}(\gamma, \delta) \in \mathbb{C}[\gamma, \delta]$ via

$$\hat{q}(\gamma, \delta) = p(\gamma z + \delta \bar{z}).$$

Note that in Remark 3.30 we have seen that complex forms are in bijection to complex, symmetric tensors. Therefore, it makes sense to use the quantity $Q_{\hat{q}}(e^1, e^2)$. Since $\hat{q}(e^1) = p(z)$ and $\hat{q}(e^2) = p(\bar{z})$ one can see analogously as in (i) that

$$Q_p(z, \bar{z}) = Q_{\hat{q}}(e^1, e^2).$$

Further, note that $\hat{q}(e^1) = q(e^1 + ie^2)$ and $\hat{q}(e^2) = q(e^1 - ie^2)$. Hence, again as in (i), one can see that

$$Q_q(e^1 + ie^2, e^1 - ie^2) = Q_{\hat{q}}(e^1, e^2).$$

This yields $Q_p(z, \bar{z}) = Q_q(e^1 + ie^2, e^1 - ie^2)$. Hence, (26) follows also directly from (28) as

$$\operatorname{Re}(p(z)) = \operatorname{Re}(q(e^1 + ie^2)) \stackrel{(26)}{\leq} B_d Q_q(e^1 + ie^2, e^1 - ie^2) = Q_p(z, \bar{z}).$$

□

Hence, to prove Theorem 4.3, it suffices to prove the two conditions (27) and (28) on bivariate, complex forms. To do so, let us first state and prove a lemma, which can be found in [12, Lemma 3.7].

Lemma 4.8. *Let $C \subseteq H_{n,2d}$ be a closed cone and $\ell \in C^*$ such that*

$$\forall p \in C : \ell(p) = 0 \implies p = 0.$$

Then the closed unit ball $C_\ell = \{p \in C : \ell(p) \leq 1\} \subseteq H_{n,2d}$ defined by ℓ is compact.

Proof. We show compactness in $H_{n,2d}$ with respect to the topology generated by the norm $\|\cdot\|$ induced by the Fischer inner product. It is trivial that C_ℓ is closed since it is the intersection of the closed cone C with the closed half-space $\{p \in H_{n,2d} : \ell(p) \leq 1\}$. Since in the finite dimensional case being compact coincides with being bounded and closed and since $H_{n,2d}$ is finite dimensional, it suffices to show that C_ℓ is bounded.

Assume that C_ℓ is not bounded, i.e. there exists a sequence $(q^{(k)})_{k \in \mathbb{N}} \subseteq C_\ell$ such that $\lim_{k \rightarrow \infty} \|q^{(k)}\| = \infty$. Without loss of generality, we can assume $q^{(k)} \neq 0$ for all k . Since C is a cone, it further holds $\left(\frac{q^{(k)}}{\|q^{(k)}\|}\right)_{k \in \mathbb{N}} \subseteq C$. Since C is closed, there is a converging subsequence of $\left(\frac{q^{(k)}}{\|q^{(k)}\|}\right)_{k \in \mathbb{N}}$ with limit $q_{\text{lim}} \in C$. For simplicity, we denote the subsequence again by $\left(\frac{q^{(k)}}{\|q^{(k)}\|}\right)_{k \in \mathbb{N}}$.

Since $q^{(k)} \in C_\ell$ for all $k \in \mathbb{N}$, we know that $\ell(q^{(k)}) \leq 1$ for all k . Further, ℓ is nonnegative on C . This shows

$$0 \leq \ell\left(\frac{q^{(k)}}{\|q^{(k)}\|}\right) = \frac{1}{\|q^{(k)}\|} \ell(q^{(k)}) \leq \frac{1}{\|q^{(k)}\|} \rightarrow 0 \quad (k \rightarrow \infty),$$

since $\|q^{(k)}\| \rightarrow \infty$ ($k \rightarrow \infty$). We also know that ℓ is continuous since on a finite dimensional vector space, every linear functional is continuous. This shows that $\ell(q_{\text{lim}}) = \lim_{k \rightarrow \infty} \ell\left(\frac{q^{(k)}}{\|q^{(k)}\|}\right) = 0$. Therefore, the assumption on ℓ tells us directly that $q_{\text{lim}} = 0$. But on the other hand, by the continuity of the norm we have

$$\|q_{\text{lim}}\| = \lim_{k \rightarrow \infty} \left\| \frac{q^{(k)}}{\|q^{(k)}\|} \right\| = 1,$$

which is a contradiction. Hence, C_ℓ must be bounded and therefore compact. \square

Using Lemma 4.8, we can now prove the Generalized Cauchy–Schwarz Inequalities (Theorem 4.3). A sketch of the proof can be found in [12, Section 3.1] while details are added in the following.

Proof of Theorem 4.3. From Lemma 4.6 and Lemma 4.7, we know that to prove Theorem 4.3, it suffices to prove that the two conditions (27) and (28) hold for some constants $A_d, B_d > 0$. To do so, we use Lemma 4.8. We already know that $C_{2,2d} \subseteq H_{2,2d}$ is a closed cone. To show (27) and (28), we define linear functionals on $C_{2,2d}$ in a suitable way such that we can apply Lemma 4.8.

- (i) We show that there is a constant $A_d > 0$ such that (27) holds for all bivariate, convex $q \in C_{2,2d}$, i.e.

$$\forall q \in C_{2,2d} : \quad 2Q_q(e^1, e^2) \leq A_d(q(e^1) + q(e^2)).$$

For this purpose, define

$$\ell : H_{2,2d} \rightarrow \mathbb{R}, \ell(q) = q(e^1) + q(e^2).$$

Clearly, ℓ is linear in q , which means $\ell \in H_{2,2d}^\vee$. We show that it also satisfies the assumptions of Lemma 4.8.

First of all, it is clear that $\ell \in C_{2,2d}^*$ since convex functions are in particular nonnegative. Now assume that $\ell(q) = 0$ for some $q \in C_{2,2d}$.

To show: $q = 0$.

Since $\ell(q) = 0$ and $q(e^1), q(e^2) \geq 0$, we know that $q(e^1) = q(e^2) = 0$ and hence by homogeneity that $q(\lambda e^1) = q(\lambda e^2) = 0$ for all $\lambda \in \mathbb{R}$.

Now let $x = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$ be arbitrary. Then

$$\begin{aligned} 0 \leq q \begin{pmatrix} a \\ b \end{pmatrix} &= q(ae^1 + be^2) \\ &= q(|a| \operatorname{sgn}(a)e^1 + |b| \operatorname{sgn}(b)e^2) \\ &= q\left((|a| + |b|) \cdot \left(\frac{|a|}{|a| + |b|} \operatorname{sgn}(a)e^1 + \frac{|b|}{|a| + |b|} \operatorname{sgn}(b)e^2\right)\right) \\ &= (|a| + |b|)^{2d} \cdot q\left(\frac{|a|}{|a| + |b|} \operatorname{sgn}(a)e^1 + \frac{|b|}{|a| + |b|} \operatorname{sgn}(b)e^2\right). \end{aligned}$$

Hence, using convexity of q yields

$$\begin{aligned} 0 \leq q \begin{pmatrix} a \\ b \end{pmatrix} &\leq (|a| + |b|)^{2d} \cdot \left(\frac{|a|}{|a| + |b|} q(\operatorname{sgn}(a)e^1) + \frac{|b|}{|a| + |b|} q(\operatorname{sgn}(b)e^2)\right) \\ &= (|a| + |b|)^{2d} \cdot 0 = 0, \end{aligned}$$

which means $q = 0$. For this reason, ℓ fulfills the assumption of Lemma 4.8, which tells us that

$$L := \{q \in C_{2,2d} : q(e^1) + q(e^2) \leq 1\}$$

is a compact subset of $H_{2,2d}$. Hence, $\lambda := \sup_{q \in L} \|q\| > 0$ is finite as supremum of a linear functional on a compact set, where $\|\cdot\|$ is again the norm induced by the Fischer inner product. Further, $\mu := \sup_{q \in H_{2,2d} \setminus \{0\}} \frac{Q_q(e^1, e^2)}{\|q\|} > 0$ is finite as it is bounded from above by the operator norm of the linear functional $q \mapsto Q_q(e^1, e^2)$ on the finite dimensional vector space $H_{2,2d}$ (see [10, Chapter 13]). Note that μ is positive, since we can take for example $q(\mathbf{x}, \mathbf{y}) = \mathbf{x}^d \mathbf{y}^d$, where by Corollary 3.35 we know that

$$Q_q(e^1, e^2) = T_q(e^1, \dots, e^1, e^2, \dots, e^2) = q_{(d,d)} \binom{2d}{d}^{-1} = \binom{2d}{d}^{-1} > 0.$$

This yields $\mu \geq \frac{Q_q(e^1, e^2)}{\|q\|} > 0$.

To show: (27) holds for all $q \in C_{2,2d}$ with constant $A_d := \lambda \cdot \mu > 0$.

Let $q \in C_{2,2d}$ be arbitrary. Without loss of generality, we assume that $q \neq 0$, otherwise the claim is trivial. As above, one can see that either $q(e^1)$ or $q(e^2)$ must be strictly positive. This shows $\ell(q) > 0$.

Further, since $\frac{q}{\ell(q)} \in L$, we know by the definition of λ, μ that

$$\begin{aligned} \lambda &\geq \left\| \frac{q}{\ell(q)} \right\| \quad \text{and} \quad \mu \geq \frac{Q_q(e^1, e^2)}{\|q\|} \\ \iff \|q\| &\leq \lambda \ell(q) \quad \text{and} \quad Q_q(e^1, e^2) \leq \mu \|q\| \end{aligned}$$

Finally, we conclude that

$$Q_q(e^1, e^2) \leq \lambda \mu \ell(q) = A_d (q(e^1) + q(e^2)),$$

which means that (27) holds.

- (ii) We show that there is a constant $B_d > 0$ such that (27) holds for all bivariate, convex $q \in C_{2,2d}$, i.e.

$$\forall q \in C_{2,2d} : \quad \operatorname{Re}(q(e^1 + ie^2)) \leq B_d Q_q(e^1 + ie^2, e^1 - ie^2).$$

For this purpose, define

$$s : H_{2,2d} \rightarrow \mathbb{R}, s(q) = Q_q(z, \bar{z}), \quad \text{where } z = e^1 + ie^2.$$

Clearly, s is linear in q , which means that $s \in H_{2,2d}^\vee$. We show that it also satisfies the assumptions of Lemma 4.8.

Therefore, consider the following identity, which is proven in [12, Appendix A]:

$$\forall p \in H_{n,2d} \quad \forall z = x + iy \in \mathbb{C}^n :$$

$$Q_p(z, \bar{z}) = \frac{4^d (d+1)}{\pi} \binom{2d}{d}^{-1} \int \int_{\alpha^2 + \beta^2 \leq 1} p(\alpha x + \beta y) d\alpha d\beta. \quad (29)$$

Identity (29) means that $Q_p(z, \bar{z})$ is proportional to the average of p on the ellipse $\{\alpha x + \beta y : \alpha, \beta \in \mathbb{R}, \alpha^2 + \beta^2 \leq 1\}$. For $n = 2$, this means that $s(q) \geq 0$ for all $q \in P_{2,2d}$ and hence in particular $s \geq 0$ on $C_{2,2d}$. Further, let $s(q) = 0$ for some $q \in C_{2,2d}$.

To show: $q = 0$.

Again, using identity (29) for $n = 2$ and $z = e^1 + ie^2$ shows that the average of q on $\{\alpha e^1 + \beta e^2 : \alpha, \beta \in \mathbb{R}, \alpha^2 + \beta^2 \leq 1\}$ is zero. Note that

$$\{\alpha e^1 + \beta e^2 : \alpha, \beta \in \mathbb{R}, \alpha^2 + \beta^2 \leq 1\} = \{(\alpha, \beta)^\top \in \mathbb{R}^2 : \|(\alpha, \beta)^\top\|_2 \leq 1\}$$

is the unit disk in \mathbb{R}^2 . Since q is in particular nonnegative, this yields $q = 0$ on the unit disk. By homogeneity, this means that q must be zero.

As in (i), using Lemma 4.8, we know that

$$S = \{q \in C_{2,2d} : Q_q(e^1 + ie^2, e^1 - ie^2) \leq 1\}$$

is compact. Define $\xi = \sup_{q \in S} \|q\| > 0$, which is finite as in (i). Further, let $\zeta = \sup_{q \in H_{2,2d}} \frac{\operatorname{Re}(q(e^1 + ie^2))}{\|q\|} > 0$, which is also finite as in (i).

To show: (28) holds for all $q \in C_{2,2d}$ with constant $B_d := \xi \zeta > 0$.

Let $q \in C_{2,2d}$ be arbitrary. Without loss of generality, assume that $q \neq 0$, otherwise the claim is trivial. As before, using (29) shows that

$s(q) > 0$. This means $\frac{q}{s(q)} \in S$ and hence $\xi \geq \frac{\|q\|}{s(q)}$. Further, it is clear that $\zeta \geq \frac{\operatorname{Re}(q(e^1 + ie^2))}{\|q\|}$. Hence, it can be deduced

$$\operatorname{Re}(q(e^1 + ie^2)) \leq \zeta \|q\| \leq \zeta \xi s(q) = B_d s(q),$$

which shows that (28) holds. \square

4.2. Value of the Optimal Constants of the Generalized Cauchy–Schwarz Inequalities. For the purpose of this thesis, it is crucial to know the value of the optimal constants $A_d^*, B_d^* > 0$ in the Generalized Cauchy–Schwarz Inequalities in Theorem 4.3, i.e. the smallest possible constants $A_d, B_d > 0$ such that (23) and (24) hold, respectively. This is formalized in the following definition.

Definition 4.9. For $d \in \mathbb{N}$, the **optimal constants** $A_d^*, B_d^* > 0$ of the inequalities (23) and (24), which only depend on the degree $2d$ are defined as

$$A_d^* = \inf_{A>0} A \text{ s.t. } \forall n \in \mathbb{N} \forall p \in C_{n,2d} \forall x, y \in \mathbb{R}^n : Q_p(x, y) \leq A \sqrt{p(x)p(y)}$$

$$B_d^* = \inf_{B>0} B \text{ s.t. } \forall n \in \mathbb{N} \forall p \in C_{n,2d} \forall z \in \mathbb{C}^n : |p(z)| \leq B Q_p(z, \bar{z}).$$

It turns out that it is rather easy to obtain a general lower bound for A_d^* and B_d^* . This is stated without a formal proof in [12, Section 3].

Proposition 4.10. For all $d \in \mathbb{N}$, we have $A_d^*, B_d^* \geq 1$.

Proof. Let $n \in \mathbb{N}$ arbitrary and consider the homogeneous polynomial $p(\mathbf{x}) = \mathbf{x}_1^{2d} \in H_{n,2d}$. The Hessian of p has only one nonzero entry and it is given by

$$\nabla^2 p(\mathbf{x}) = \begin{pmatrix} 2d(2d-1)\mathbf{x}_1^{2d-2} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$

Since $2d(2d-1)\mathbf{x}_1^{2d-2} = 2d(2d-1)(x_1^2)^{d-1} \geq 0$ for all $x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$, $\nabla^2 p(x)$ is clearly positive semidefinite for all $x \in \mathbb{R}^n$. Hence $p(\mathbf{x}) \in C_{n,2d}$. Further, the tensor associated to p is given by

$$T_p : (\mathbb{R}^n)^{2d} \rightarrow \mathbb{R}, T_p(x^1, \dots, x^{2d}) = x_1^1 \dots x_1^{2d},$$

which yields that for $x = (x_1, \dots, x_n)^\top, y = (y_1, \dots, y_n)^\top \in \mathbb{R}^n$, we have $Q_p(x, y) = (x_1)^d \cdot (y_1)^d$. Now, taking $x = y = e^1 \in \mathbb{R}^n, z = e^1 \in \mathbb{C}^n$ as the first unit vector shows

$$A_d^* = A_d^* \cdot \underbrace{\sqrt{p(x)p(y)}}_{=1} \stackrel{(23)}{\geq} Q_p(x, y) = 1$$

$$\text{and } B_d^* = B_d^* \cdot \underbrace{Q_p(z, \bar{z})}_{=1} \stackrel{(24)}{\geq} |p(z)| = 1,$$

and hence $A_d^*, B_d^* \geq 1$. \square

Remark 4.11. (a) The infima in Definition 4.9 are actually minima, since A_d^*, B_d^* are bounded from below by 1 and we minimize over a closed set since the inequalities in the conditions are not strict.

(b) In Section 4.1 we have seen that A_d^*, B_d^* can equivalently be computed via the optimal constants of the inequalities (27) and (28) in Lemma 4.7. Hence, we have

$$A_d^* = \inf_{A>0} A \quad \text{s.t.} \quad \forall q \in C_{2,2d} : \quad 2Q_q(e^1, e^2) \leq A(q(e^1) + q(e^2)) \quad (30)$$

$$B_d^* = \inf_{B>0} B \quad \text{s.t.} \quad \forall q \in C_{2,2d} : \quad \begin{aligned} &\text{Re}(q(e^1 + ie^2)) \\ &\leq BQ_q(e^1 + ie^2, e^1 - ie^2). \end{aligned} \quad (31)$$

Of course, Proposition 4.10 can analogously be proven using (30) and (31). On the other hand, it is hard to obtain upper bounds on A_d^*, B_d^* this way. Therefore, we use a dual approach, where we rewrite (30) and (31) using the linear functionals

$$\mathcal{L}_A: H_{2,2d} \rightarrow \mathbb{R}, \ell_A(q) = A(q(e^1) + q(e^2)) - 2Q_q(e^1, e^2)$$

$$\mathcal{L}_B: H_{2,2d} \rightarrow \mathbb{R}, \ell_B(q) = BQ_q(e^1 + ie^2, e^1 - ie^2) - \text{Re}(q(e^1 + ie^2)).$$

By definition of the dual cone $C_{2,2d}^*$, we can rewrite (30) and (31) as

$$A_d^* = \inf_{A>0} A \quad \text{s.t.} \quad \mathcal{L}_A \in C_{2,2d}^* \quad (32)$$

$$B_d^* = \inf_{B>0} B \quad \text{s.t.} \quad \mathcal{L}_B \in C_{2,2d}^*. \quad (33)$$

In Example 3.59(e) and 3.59(f), we have already seen that $\mathcal{L}_A, \mathcal{L}_B$ can be written as differential operators in the form

$$\begin{aligned} \mathcal{L}_A &= \frac{1}{(2d)!} \left(A \left(\partial_{\mathbf{x}^{2d}} + \partial_{\mathbf{y}^{2d}} \right) - 2\partial_{\mathbf{x}^d} \partial_{\mathbf{y}^d} \right) \\ \mathcal{L}_B &= \frac{1}{(2d)!} \left(B \left(\partial_{\mathbf{x}^2} + \partial_{\mathbf{y}^2} \right)^d - \text{Re} \left((\partial_{\mathbf{x}} + i\partial_{\mathbf{y}})^{2d} \right) \right). \end{aligned}$$

Since the factor $\frac{1}{(2d)!} > 0$ plays no role in verifying whether $\mathcal{L}_A, \mathcal{L}_B \in C_{2,2d}^*$, we reduce (32)-(33) to

$$A_d^* = \inf_{A>0} A \quad \text{s.t.} \quad \ell_A = A \left(\partial_{\mathbf{x}^{2d}} + \partial_{\mathbf{y}^{2d}} \right) - 2\partial_{\mathbf{x}^d} \partial_{\mathbf{y}^d} \in C_{2,2d}^* \quad (34)$$

$$B_d^* = \inf_{B>0} B \quad \text{s.t.} \quad \ell_B = B \left(\partial_{\mathbf{x}^2} + \partial_{\mathbf{y}^2} \right)^d - \text{Re} \left((\partial_{\mathbf{x}} + i\partial_{\mathbf{y}})^{2d} \right) \in C_{2,2d}^* \quad (35)$$

Further, recalling Corollary 3.88, we know that

$$C_{2,2d}^* = \left\{ \sum_{i=1}^m (\alpha_i \partial_{\mathbf{x}} + \beta_i \partial_{\mathbf{y}})^2 (\gamma_i \partial_{\mathbf{x}} + \delta_i \partial_{\mathbf{y}})^{2d-2} : \alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R} \right\}.$$

Hence, to show that $\ell_A \in C_{2,2d}^*$ for some $A > 0$, we need to find such a representation of ℓ_A . An analogous statement can be made for B .

4.2.1. *Values of A_d^* .* Using the dual approach (34) and (35) in Remark 4.11, one can see the following.

Proposition 4.12. *It holds $A_1^* = A_2^* = A_3^* = 1$.*

Proof. We already know by Proposition 4.10 that $A_1^*, A_2^*, A_3^* \geq 1$. Now the claim follows since for $d = 1, 2, 3$, ℓ_1 has a representation as in Corollary 3.88:

$$\begin{aligned} d=1: \ell_1 &\stackrel{(34)}{=} \partial \mathbf{x}^2 + \partial \mathbf{y}^2 - 2\partial \mathbf{x} \partial \mathbf{y} = (\partial \mathbf{x} - \partial \mathbf{y})^2 \in C_{2,2}^* \\ d=2: \ell_1 &\stackrel{(34)}{=} \partial \mathbf{x}^4 + \partial \mathbf{y}^4 - 2\partial \mathbf{x}^2 \partial \mathbf{y}^2 = (\partial \mathbf{x} - \partial \mathbf{y})^2 (\partial \mathbf{x} + \partial \mathbf{y})^2 \in C_{2,4}^* \\ d=3: \ell_1 &\stackrel{(34)}{=} \partial \mathbf{x}^6 + \partial \mathbf{y}^6 - 2\partial \mathbf{x}^3 \partial \mathbf{y}^3 \\ &= \frac{1}{2} (\partial \mathbf{x} - \partial \mathbf{y})^2 \left(\partial \mathbf{x}^4 + \partial \mathbf{y}^4 + (\partial \mathbf{x} + \partial \mathbf{y})^2 \right) \in C_{2,6}^*. \end{aligned}$$

□

The exact value of A_d^* is not known in general for all $d \in \mathbb{N}$. However, in [12, Appendix B] it is proven that $A_d^* > 1$ for all even $d \in \mathbb{N}_{>2}$. For the proof we define similar as in Proposition 4.10 a polynomial $p_d \in C_{2,2d}$ in a clever way such that $p_d(e^1) = p_d(e^2) = 1$ but $Q_{p_d}(e^1, e^2) > 1$. Since $2 \underbrace{Q_{p_d}(e^1, e^2)}_{>1} \leq A_d^* \underbrace{(p_d(e^1) + p_d(e^2))}_{=2}$ by (30), this then allows us to deduce

that $A_d^* > 1$. The polynomial p_d is the linear combination of two other polynomials s and q . Those polynomials are chosen such that the Hessian $\nabla^2 s(x, y)$ is positive definite everywhere except for $x = \pm y$, where it is only positive semidefinite. The Hessian $\nabla^2 q(x, y)$ on the other hand is precisely positive definite for $x = \pm y$. This allows us to take a linear combination, more precisely a conical combination of s and q such that the resulting function is convex.

Proposition 4.13. *Let $d \in \mathbb{N}_{>2}$ be even and consider the bivariate form $s(\mathbf{x}, \mathbf{y}) := \frac{1}{2} ((\mathbf{x} + \mathbf{y})^{2d} + (\mathbf{x} - \mathbf{y})^{2d}) \in H_{n,2d}$. Then $s(\mathbf{x}, \mathbf{y}) \in C_{2,2d}$ is convex and the Hessian $\nabla^2 s(x, y)$ ($(x, y)^\top \in \mathbb{R}^2$) is positive definite everywhere except for $x = \pm y$, where it is only positive semidefinite. Further, it holds $s(e^1) = s(e^2) = Q_s(e^1, e^2) = 1$.*

Proof. It is clear that $(\mathbf{x} + \mathbf{y})^{2d}, (\mathbf{x} - \mathbf{y})^{2d} \in H_{2,2d}$ are convex since their Hessians are given by

$$\begin{aligned} &2d(2d-1) \cdot \begin{pmatrix} (\mathbf{x} + \mathbf{y})^{2d-2} & (\mathbf{x} + \mathbf{y})^{2d-2} \\ (\mathbf{x} + \mathbf{y})^{2d-2} & (\mathbf{x} + \mathbf{y})^{2d-2} \end{pmatrix} \\ \text{and } &2d(2d-1) \cdot \begin{pmatrix} (\mathbf{x} - \mathbf{y})^{2d-2} & -(\mathbf{x} - \mathbf{y})^{2d-2} \\ -(\mathbf{x} - \mathbf{y})^{2d-2} & (\mathbf{x} - \mathbf{y})^{2d-2} \end{pmatrix}, \end{aligned}$$

which are clearly positive semidefinite for all $(x, y)^\top \in \mathbb{R}^2$.

Hence, $s(\mathbf{x}, \mathbf{y}) = \frac{1}{2} ((\mathbf{x} + \mathbf{y})^{2d} + (\mathbf{x} - \mathbf{y})^{2d})$ is convex as a convex combination of convex functions. More precisely, for $(x, y)^\top \in \mathbb{R}^2$ arbitrary, we

have

$$\begin{aligned} & \nabla^2 s(x, y) \\ &= \frac{1}{2} 2d(2d-1) \begin{pmatrix} (x+y)^{2d-2} + (x-y)^{2d-2} & (x+y)^{2d-2} - (x-y)^{2d-2} \\ (x+y)^{2d-2} - (x-y)^{2d-2} & (x+y)^{2d-2} + (x-y)^{2d-2} \end{pmatrix} \end{aligned}$$

and hence

$$\begin{aligned} \det(\nabla^2 s(x, y)) &= \frac{1}{2} 2d(2d-1) \cdot 4(x+y)^{2d-2}(x-y)^{2d-2} \\ & \begin{cases} > 0, & \text{if } y \neq \pm x \\ = 0, & \text{if } y \in \{\pm x\} \end{cases}. \end{aligned}$$

This means that the Hessian of s is positive definite everywhere except for $x = \pm y$, where it is only positive semidefinite. Moreover, note that $s(e^1) = s(e^2) = 1$. Further, by Corollary 3.35 we know that

$$\begin{aligned} Q_s(e^1, e^2) &= T_s(e^1, \dots, e^1, e^2, \dots, e^2) = s_{(d,d)} \binom{2d}{d}^{-1}, \\ s_{(d,d)} &= \frac{1}{2} \left(\binom{2d}{d} + \binom{2d}{d} (-1)^d \right) \stackrel{d \text{ even}}{=} \binom{2d}{d} \end{aligned}$$

and hence $Q_s(e^1, e^2) = 1$ whenever d is even. \square

For the second proposition consider the following notation.

Notation 4.14. Write $(\pm 1, \pm 1)$ or $(\pm 1, \mp 1)$ for any of the two elements in the set $\{(1, 1), (-1, -1)\}$ or $\{(1, -1), (-1, 1)\}$, respectively.

Proposition 4.15. *Let $d \in \mathbb{N}$ be even and consider the bivariate form $q(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{d-1} \mathbf{x}^{2k} \mathbf{y}^{2d-2k} \in H_{2,2d}$. If $d > 7/2$, the Hessian $\nabla^2 q(x, y)$ is positive definite whenever $x = \pm y \in \mathbb{R} \setminus \{0\}$. Further, it holds $q(e^1) = q(e^2) = 0$ and $Q_q(e^1, e^2) > 0$.*

Proof. We have for the gradient

$$\nabla q(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \sum_{k=1}^{d-1} 2k \mathbf{x}^{2k-1} \mathbf{y}^{2d-2k} \\ \sum_{k=1}^{d-1} (2d-2k) \mathbf{x}^{2k} \mathbf{y}^{2d-2k-1} \end{pmatrix}$$

and further for the Hessian

$$\begin{aligned} & \nabla^2 q(\mathbf{x}, \mathbf{y}) \\ &= \begin{pmatrix} \sum_{k=1}^{d-1} 2k(2k-1) \mathbf{x}^{2k-2} \mathbf{y}^{2d-2k} & \sum_{k=1}^{d-1} 2k(2d-2k) \mathbf{x}^{2k-1} \mathbf{y}^{2d-2k-1} \\ \sum_{k=1}^{d-1} 2k(2d-2k) \mathbf{x}^{2k-1} \mathbf{y}^{2d-2k-1} & \sum_{k=1}^{d-1} (2d-2k)(2d-2k-1) \mathbf{x}^{2k} \mathbf{y}^{2d-2k-2} \end{pmatrix}. \end{aligned}$$

It is clear that $(\nabla^2 q(\pm 1, \pm 1))_{1,1}, (\nabla^2 q(\pm 1, \pm 1))_{2,2} \geq 0$ and $(\nabla^2 q(\pm 1, \mp 1))_{1,1}, (\nabla^2 q(\pm 1, \mp 1))_{2,2} > 0$ hold, since every summand is strictly positive. Further, we have for the determinant

$$\begin{aligned}
& \det(\nabla^2 q(\pm 1, \pm 1)) = \det(\nabla^2 q(\pm 1, \mp 1)) \\
&= \begin{vmatrix} \sum_{k=1}^{d-1} 2k(2k-1) & \pm \left(\sum_{k=1}^{d-1} 2k(2d-2k) \right) \\ \pm \left(\sum_{k=1}^{d-1} 2k(2d-2k) \right) & \sum_{k=1}^{d-1} (2d-2k)(2d-2k-1) \end{vmatrix} \\
&= \begin{vmatrix} \sum_{k=1}^{d-1} 2k(2k-1) & \pm \left(\sum_{k=1}^{d-1} 2k(2d-2k) \right) \\ \pm \left(\sum_{k=1}^{d-1} 2k(2d-2k) \right) & \sum_{k=1}^{d-1} 2k(2k-1) \end{vmatrix} \\
&= \begin{vmatrix} 4 \frac{(d-1)d(2d-1)}{6} - 2 \frac{(d-1)d}{2} & \pm \left(4d \frac{(d-1)d}{2} - 4 \frac{(d-1)d(2d-1)}{6} \right) \\ \pm \left(4d \frac{(d-1)d}{2} - 4 \frac{(d-1)d(2d-1)}{6} \right) & 4 \frac{(d-1)d(2d-1)}{6} - 2 \frac{(d-1)d}{2} \end{vmatrix} \\
&= \frac{d(d-1)}{3} \begin{vmatrix} 4d-5 & \pm(2d+2) \\ \pm(2d+2) & 4d-5 \end{vmatrix} \\
&= d(d-1)(2d-7)(2d-1).
\end{aligned}$$

This shows that $\det(\nabla^2 q(\pm 1, \pm 1)) > 0$ if $d > 7/2$. Hence, it holds $\nabla^2 q(\pm 1, \pm 1), \nabla^2 q(\pm 1, \mp 1) \succ 0$ for $d > 7/2$. By homogeneity, this allows us to deduce that $\nabla^2 q(x, y)$ is strictly positive definite whenever $x = \pm y \in \mathbb{R} \setminus \{0\}$.

Moreover, it is clear that $q(e^1) = q(e^2) = 0$ holds and again by using Corollary 3.35, one can see that

$$Q_q(e^1, e^2) = q_{(d,d)} \binom{2d}{d}^{-1} = 1 \cdot \binom{2d}{d}^{-1} > 0,$$

where $q_{(d,d)} = 1$ since d is even. \square

In the following proposition, we show how the two bivariate forms $s(\mathbf{x}, \mathbf{y})$ and $q(\mathbf{x}, \mathbf{y})$ can be combined in a clever way to obtain a form p_d that has the desired properties to show that $A_d^* > 1$ whenever $d \in \mathbb{N}_{>2}$ is even. A sketch of the proof can be found in [12, Appendix B].

Proposition 4.16. *Let $d \in \mathbb{N}_{>2}$ be even and let $s, q \in H_{2,2d}$ be defined as in the Propositions 4.13 and 4.15. Then there is a constant $\alpha_d > 0$ such that the bivariate form $p_d(\mathbf{x}, \mathbf{y}) := s(\mathbf{x}, \mathbf{y}) + \alpha_d q(\mathbf{x}, \mathbf{y})$ is convex, i.e. $p_d \in C_{2,2d}$. Further, in this case it holds that $p_d(e^1) = p_d(e^2) = 1$ and $Q_{p_d}(e^1, e^2) > 1$.*

Proof. By homogeneity, it suffices to show that there is an $\alpha_d > 0$ such that the Hessian of $p_d := s + \alpha_d q$ is positive semidefinite on the circle $\mathcal{C} := \{(x, y)^\top \in \mathbb{R}^2 : x^2 + y^2 = 2\}$ of radius $\sqrt{2}$. Since by Proposition 4.15 $q(\mathbf{x}, \mathbf{y})$ is positive definite on the set $X := \{(\pm 1, \pm 1)^\top, (\pm 1, \mp 1)^\top\} \subseteq \mathbb{R}^2$ for all even $d \in \mathbb{N}_{>2}$, we can take an open subset $U \subseteq \mathcal{C}$ of \mathcal{C} , which contains

X , such that q is positive definite on U . This allows to define

$$\begin{aligned}\alpha_d &:= \inf_{v \in \mathbb{R}^2, \|v\|_2=1, (x,y)^\top \in \mathcal{C} \setminus U} \frac{v^\top \nabla^2 s(x,y)v}{|v^\top \nabla^2 q(x,y)v|} \\ &= \min_{v \in \mathbb{R}^2, \|v\|_2=1, (x,y)^\top \in \mathcal{C} \setminus U} \frac{v^\top \nabla^2 s(x,y)v}{|v^\top \nabla^2 q(x,y)v|}.\end{aligned}$$

The minimum exists, since we minimize over a compact set. Further, by Proposition 4.13 we know that the numerator is always strictly positive. Again, since we minimize over a compact set, this allows us to deduce that $\alpha_d > 0$.

Remains to show that p_d is convex. To do so, let $v \in \mathbb{R}^2$ such that $\|v\|_2 = 1$ and $(x,y)^\top \in \mathcal{C}$ be arbitrary. If $(x,y)^\top \in \mathcal{C} \setminus U$, we know by construction of α_d that

$$v^\top \nabla^2 p_d(x,y)v = \underbrace{v^\top \nabla^2 s(x,y)v}_{>0} + \underbrace{\alpha_d v^\top \nabla^2 q(x,y)v}_{\in \{\pm v^\top \nabla^2 s(x,y)v\}} \geq 0.$$

On the other hand, if $(x,y)^\top \in U$, it is clear by the choice of U that we have $v^\top \nabla^2 q(x,y)v > 0$. Further, since s is convex by Proposition 4.13 we know that $v^\top \nabla^2 s(x,y)v \geq 0$.

This shows that the Hessian $\nabla^2 p_d(\mathbf{x}, \mathbf{y})$ is positive semidefinite on \mathcal{C} . By homogeneity it must also be globally positive semidefinite and hence p_d is convex.

For the second part of the claim, we know by Proposition 4.13 and 4.15 that $s(e^1) = s(e^2) = 1, q(e^1) = q(e^2) = 0$ and $Q_s(e^1, e^2) = 1, Q_q(e^1, e^2) > 0$. Since $p_d = s + \alpha_d q$ and $\alpha_d > 0$, this yields

$$\begin{aligned}p_d(e^1) &= s(e^1) + \alpha_d q(e^1) = 1 + 0 = 1 \stackrel{\text{analogously}}{=} p_d(e^2), \\ Q_{p_d}(e^1, e^2) &= Q_s(e^1, e^2) + \alpha_d Q_q(e^1, e^2) = 1 + \underbrace{\alpha_d}_{>0} \binom{2d}{d}^{-1} > 1.\end{aligned}$$

□

Using Proposition 4.16, we obtain the following.

Lemma 4.17. *For all even $d \in \mathbb{N}_{>2}$, it holds $A_d^* > 1$.*

Proof. Note that without loss of generality, it suffices to consider $n = 2$, since the optimal constant A_d^* does not depend on the number n of variables. For $p_d \in C_{2,2d}$ as in Proposition 4.16 it holds $p_d(e^1) = p_d(e^2) = 1$ and $Q_{p_d}(e^1, e^2) > 1$. Hence, using $\underbrace{2Q_{p_d}(e^1, e^2)}_{>2} \stackrel{(30)}{\leq} A_d^* \underbrace{(p_d(e^1) + p_d(e^2))}_{=2} = 2A_d^*$, it can be deduced that $A_d^* > 1$ must hold. □

Remark 4.18. In [12, Section 3.2.1] the author presented an exact value of $A_4^* > 1$ where he used the SDP (semidefinite programming) structure of the minimization problem (34) in dual and (30) in primal form, respectively. More precisely, he showed by solving the dual problem (34) exactly, that A_4^*

is given by

$$\frac{1}{70}\omega^{1/3} + \frac{128}{15}\omega^{-1/3} + \frac{11}{35}, \text{ where } \omega := 14336 + i\frac{14336\sqrt{3}}{9}.$$

In addition, he was able to numerically approximate some constants A_d^* for $d > 4$. In [12, Table 1], he presented the following values

d	1	2	3	4	5	6	7	8
A_d^*	1.000	1.000	1.000	1.011	1.000	1.061	1.000	1.048

Conjecture 4.19. In [12, Section 3.2.1], the author conjectured that $A_d^* = 1$ whenever d is odd. However, although in Proposition 4.12 we are able to prove that $A_1^* = A_3^* = 1$, no further results are known that cover any odd $d \geq 5$.

4.2.2. *Values of B_d^* .* Whereas we know the exact value of A_d^* only for $d = 1, 2, 3$ and $d = 4$ by Remark 4.18, we see in this section that there is a general formula for B_d^* . More precisely, we prove that $B_d^* = \frac{\binom{2(d-1)}{d-1}}{d}$. This can be done by showing that $\frac{\binom{2(d-1)}{d-1}}{d}$ is at the same time a lower and upper bound for B_d^* .

Step 1: Show that B_d^* is bounded from above by $\frac{\binom{2(d-1)}{d-1}}{d}$, i.e.

$$B_d^* \leq \frac{\binom{2(d-1)}{d-1}}{d}.$$

To show this, we use a dual approach, where we consider the dual problem given in (35). Therefore, it suffices to show that for $B_d = \frac{\binom{2(d-1)}{d-1}}{d}$, it holds $\ell_{B_d} = B_d(\partial\mathbf{x}^2 + \partial\mathbf{y}^2)^d - \text{Re}((\partial\mathbf{x} + i\partial\mathbf{y})^{2d}) \in C_{2,2d}^*$. For this purpose, we use the representation of $C_{2,2d}^*$ as in Corollary 3.88.

First, we need some identity involving the trigonometric functions \sin, \cos .

Proposition 4.20. *For all $\varphi \in \mathbb{R}$, it holds*

$$\frac{\binom{2(d-1)}{d-1}}{d} - \cos(2d\varphi) = \frac{4^d}{2d} \sum_{k=0}^{d-1} \sin^2\left(\varphi - \frac{k\pi}{2d}\right) \cos^{2d-2}\left(\frac{k\pi}{2d} - \varphi\right).$$

Proof. [12, Appendix C]. □

Now we can prove that $\ell_{B_d} \in C_{2,2d}^*$ for $B_d = \frac{\binom{2(d-1)}{d-1}}{d}$.

Proposition 4.21. *For all $x, y \in \mathbb{R}$, it holds*

$$\frac{\binom{2(d-1)}{d-1}}{d} (x^2 + y^2)^d - \text{Re}\left((x + iy)^{2d}\right) = \frac{4^d}{2d} \sum_{k=0}^{d-1} (-s_k x + c_k y)^2 (c_k x + s_k y)^{2d-2}, \quad (36)$$

where $c_k = \cos\left(\frac{k\pi}{2d}\right), s_k = \sin\left(\frac{k\pi}{2d}\right)$ ($k = 0, \dots, d-1$). In particular, it follows from Corollary 3.88 that $\ell_{B_d} \in C_{2,2d}^*$ for $B_d = \frac{\binom{2(d-1)}{d-1}}{d}$ and hence $B_d^* \leq \frac{\binom{2(d-1)}{d-1}}{d}$.

Proof. Let $d \in \mathbb{N}$ and $x, y \in \mathbb{R}$ be arbitrary. Since equation (36) is homogeneous in $(x, y)^\top \in \mathbb{R}^2$ on both sides, it suffices to show that equality holds on the unit sphere. Hence, we can without loss of generality assume that $\|(x, y)^\top\|_2 = 1$. This allows to write $(x, y)^\top = (\cos \varphi, \sin \varphi)^\top$ for some $\varphi \in \mathbb{R}$ in polar coordinates. For this reason, one can see that

$$\begin{aligned}
& \frac{\binom{2(d-1)}{d-1}}{d} (x^2 + y^2)^d - \operatorname{Re} \left((x + iy)^{2d} \right) \\
&= \frac{\binom{2(d-1)}{d-1}}{d} \underbrace{(\cos^2 \varphi + \sin^2 \varphi)^d}_{=1} - \operatorname{Re} \left(\underbrace{(\cos \varphi + i \sin \varphi)^{2d}}_{=e^{i2d\varphi}} \right) \\
&= \frac{\binom{2(d-1)}{d-1}}{d} - \operatorname{Re} \left(e^{i2d\varphi} \right) \\
&= \frac{\binom{2(d-1)}{d-1}}{d} - \cos(2d\varphi) \\
\text{Prop. 4.20} &= \frac{4^d}{2d} \sum_{k=0}^{d-1} \sin^2 \left(\varphi - \frac{k\pi}{2d} \right) \cos^{2d-2} \left(\frac{k\pi}{2d} - \varphi \right)
\end{aligned}$$

Using

$$\begin{aligned}
\sin^2 \left(\varphi - \frac{k\pi}{2d} \right) &= \left(\underbrace{\sin \varphi}_{=y} \cos \left(\frac{k\pi}{2d} \right) - \underbrace{\cos \varphi}_{=x} \sin \left(\frac{k\pi}{2d} \right) \right)^2 \\
&= \left(- \left(\sin \left(\frac{k\pi}{2d} \right) \right) x + \left(\cos \left(\frac{k\pi}{2d} \right) \right) y \right)^2
\end{aligned}$$

and

$$\begin{aligned}
\cos^{2d-2} \left(\frac{k\pi}{2d} - \varphi \right) &= \left(\cos \left(\frac{k\pi}{2d} \right) \underbrace{\cos \varphi}_{=x} + \sin \left(\frac{k\pi}{2d} \right) \underbrace{\sin \varphi}_{=y} \right)^{2d-2} \\
&= \left(\left(\cos \left(\frac{k\pi}{2d} \right) \right) x + \left(\sin \left(\frac{k\pi}{2d} \right) \right) y \right)^{2d-2},
\end{aligned}$$

yields that (36) holds. Hence, the representation of $C_{2,2d}^*$ in Corollary 3.88

yields that for $B_d = \frac{\binom{2(d-1)}{d-1}}{d}$, it holds

$$\begin{aligned}
\ell_{B_d} &= \frac{\binom{2(d-1)}{d-1}}{d} (\partial \mathbf{x}^2 + \partial \mathbf{y}^2)^d - \operatorname{Re} \left((\partial \mathbf{x} + i \partial \mathbf{y})^{2d} \right) \\
&\stackrel{(36)}{=} \frac{4^d}{2d} \sum_{k=0}^{d-1} (-s_k \partial \mathbf{x} + c_k \partial \mathbf{y})^2 (c_k \partial \mathbf{x} + s_k \partial \mathbf{y})^{2d-2} \in C_{2,2d}^*.
\end{aligned}$$

For this reason, it follows from the definition of B_d^* as in the dual problem

$$(35) \quad B_d^* \leq \frac{\binom{2(d-1)}{d-1}}{d}. \quad \square$$

Step 2: Show that B_d^* is bounded from below by $\frac{\binom{2(d-1)}{d-1}}{d}$, i.e.

$$B_d^* \geq \frac{\binom{2(d-1)}{d-1}}{d}.$$

For this purpose, we find for each $d \in \mathbb{N}$ some bivariate convex polynomial $p_d \in \mathcal{C}_{2,2d}$ such that in the inequality (28) from Lemma 4.7, equality holds for $B_d = \frac{\binom{2(d-1)}{d-1}}{d}$, i.e.

$$\operatorname{Re}(p_d(z)) = \frac{\binom{2(d-1)}{d-1}}{d} Q_{p_d}(z, \bar{z}), \quad (37)$$

where $z = e^1 + ie^2 \in \mathbb{C}^2$.

Definition 4.22. For a complex polynomial $p \in \mathbb{C}[\mathbf{x}]$ of degree k in n variables such that $p = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq k} p_\alpha \mathbf{x}^\alpha$, we define the **real part** of f as

$$\operatorname{Re}(p)(\mathbf{x}) = \frac{1}{2} \left(\sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq k} f_\alpha \mathbf{x}^\alpha + \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq k} \overline{f_\alpha} \mathbf{x}^\alpha \right) = \sum_{\alpha \in \mathbb{N}_0^n, |\alpha| \leq k} \operatorname{Re}(p_\alpha) \mathbf{x}^\alpha.$$

This defines a function $\operatorname{Re}: \mathbb{C}[\mathbf{x}] \rightarrow \mathbb{R}[\mathbf{x}]$. In particular, if p is homogeneous of degree k , we have $\operatorname{Re}(p) \in H_{n,k}$.

Remark 4.23. Let $p \in \mathbb{C}[\mathbf{x}]$ be arbitrary. Note that it does not hold $\operatorname{Re}(p)(x) = \operatorname{Re}(p(x))$ as functions, where $x \in \mathbb{C}^n$. To see so, consider for example $p(\mathbf{x}, \mathbf{y}) = (\mathbf{x} + i\mathbf{y})^{2d} \in \mathbb{C}[\mathbf{x}, \mathbf{y}]$. Then we have

$$\operatorname{Re}(p)(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \left((\mathbf{x} + i\mathbf{y})^{2d} + (\mathbf{x} - i\mathbf{y})^{2d} \right) = \sum_{k=0}^d \binom{2d}{2k} \mathbf{x}^{2d-2k} (i\mathbf{y})^{2k}.$$

Hence, it holds for $x = (1, i)^\top \in \mathbb{C}^2$

$$\operatorname{Re}(p)(1, i) = 2^{2d-1} \neq 0 = \operatorname{Re}(0) = \operatorname{Re}(p(1, i)).$$

For proving $B_d^* \geq \frac{\binom{2(d-1)}{d-1}}{d}$, we need the following lemma, which gives the Hessian and Laplacian in polar coordinates. It can be found in [12, Lemma 3.11].

Lemma 4.24. *Let $p \in H_{n,k}$ be a form with an expression in polar coordinates as $p(x, y) = r^k q(\varphi)$, where $(x, y)^\top = r(\cos \varphi, \sin \varphi) \in \mathbb{R}^2$ and $q(\varphi)$ is a polynomial expression in $\cos \varphi, \sin \varphi$. Then the Hessian and Laplacian of p evaluated at $(x, y)^\top$ are given by*

$$\begin{aligned} \nabla^2 p(x, y) &= r^{k-2} \left(k(k-1)q(\varphi)E_{rr} + (k-1)q'(\varphi)E_{r\varphi} + (kq(\varphi) + q''(\varphi))E_{\varphi\varphi} \right) \\ \Delta p(x, y) &= r^{k-2} \left(k^2 q(\varphi) + q''(\varphi) \right), \end{aligned}$$

where $e_r = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, e_\varphi = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \in \mathbb{R}^2$ and

$$\begin{aligned} E_{rr} &= e_r e_r^\top = \begin{pmatrix} \cos^2 \varphi & \cos \varphi \sin \varphi \\ \cos \varphi \sin \varphi & \sin^2 \varphi \end{pmatrix}, \\ E_{r\varphi} &= e_r e_\varphi^\top + e_\varphi e_r^\top = \begin{pmatrix} -2 \sin \varphi \cos \varphi & \cos^2 \varphi - \sin^2 \varphi \\ \cos^2 \varphi - \sin^2 \varphi & 2 \sin \varphi \cos \varphi \end{pmatrix}, \\ E_{\varphi\varphi} &= e_\varphi e_\varphi^\top = \begin{pmatrix} \sin^2 \varphi & -\sin \varphi \cos \varphi \\ -\sin \varphi \cos \varphi & \cos^2 \varphi \end{pmatrix} \end{aligned}$$

Proof. We follow the proof in [12, Appendix D]. However, there is an error in the formulation of the corresponding lemma [12, Lemma 3.11]. More precisely, the last term of the Hessian $\nabla^2 p(x, y)$ in Lemma 4.24 has to be $(kq(\varphi) + q''(\varphi))E_{\varphi\varphi}$ instead of $(k + q''(\varphi))E_{\varphi\varphi}$ as stated in [12, Lemma 3.11]. For this reason, we follow the proof in [12, Appendix D] as long as it is correct.

As in [12, Appendix D], the Hessian operator in polar coordinates is given by

$$\nabla^2 = \frac{\partial^2}{\partial r^2} E_{rr} + \frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial}{\partial \varphi} \right) E_{r\varphi} + \left(\frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} \right) E_{\varphi\varphi}. \quad (38)$$

The next step is where the error in [12] occurred. Applying the Hessian as above to a twice differentiable function $p(x, y) = r^k q(\varphi)$, where $(x, y)^\top = r(\cos \varphi, \sin \varphi) \in \mathbb{R}^2$ and $q(\varphi)$ is a polynomial expression in $\cos \varphi, \sin \varphi$ leads to

$$\begin{aligned} \nabla^2 r^k q(\varphi) &\stackrel{(38)}{=} k(k-1)r^{k-2}q(\varphi)E_{rr} + \frac{\partial}{\partial r} \left(\frac{1}{r} r^k q'(\varphi) \right) E_{r\varphi} \\ &\quad + \left(\frac{1}{r} k r^{k-1} q(\varphi) + \frac{1}{r^2} r^k q''(\varphi) \right) E_{\varphi\varphi}. \end{aligned}$$

This yields

$$\begin{aligned} &\nabla^2 r^k q(\varphi) \\ &= k(k-1)r^{k-2}q(\varphi)E_{rr} + (k-1)r^{k-2}q'(\varphi)E_{r\varphi} + r^{k-2}(kq(\varphi) + q''(\varphi))E_{\varphi\varphi} \\ &= r^{k-2} (k(k-1)q(\varphi)E_{rr} + (k-1)q'(\varphi)E_{r\varphi} + (kq(\varphi) + q''(\varphi))E_{\varphi\varphi}), \end{aligned}$$

as stated in Lemma 4.24. The rest of the proof can now again be done as in [12, Appendix D]. The Laplacian $\Delta p(x, y)$ is the trace of the matrix $\nabla^2 p(x, y)$. Since the trace of E_{rr} and $E_{\varphi\varphi}$ is one and the trace of $E_{r\varphi}$ is zero, we conclude

$$\begin{aligned} \Delta p(x, y) &= r^{k-2} \left(k(k-1)q(\varphi) \underbrace{\text{tr}(E_{rr})}_{=1} + (k-1)q'(\varphi) \underbrace{\text{tr}(E_{r\varphi})}_{=0} \right. \\ &\quad \left. + (kq(\varphi) + q''(\varphi)) \underbrace{\text{tr}(E_{\varphi\varphi})}_{=1} \right). \end{aligned}$$

This yields

$$\begin{aligned} \Delta p(x, y) &= r^{k-2} (k(k-1)q(\varphi) + (kq(\varphi) + q''(\varphi))) \\ &= r^{k-2} (k^2 q(\varphi) + q''(\varphi)), \end{aligned}$$

as desired. \square

In the following proposition, one can find the polynomial $p_d \in H_{2,2d}$, for which equality (37) holds. Before showing this, we prove at first that p_d is indeed convex. The proof is done independently of the work in [12] because of the error mentioned at the beginning of the proof of Lemma 4.24.

Proposition 4.25. *Let $p_d \in H_{2,2d}$ be defined as*

$$p_d(\mathbf{x}, \mathbf{y}) = \text{Re} \left((\mathbf{x} + i\mathbf{y})^{2d} \right) + (2d-1) (\mathbf{x}^2 + \mathbf{y}^2)^d. \quad (39)$$

Then p_d is convex.

Proof. For the proof, we write the form p_d in polar coordinates. Therefore, let $(x, y)^\top \in \mathbb{R}^2$ be arbitrary. One can write $(x, y)^\top = r(\cos \varphi, \sin \varphi)^\top$, where $r = \|(x, y)^\top\|_2 \in \mathbb{R}_{\geq 0}$, $\varphi = \arg((x, y)^\top) \in \mathbb{R}$. For $p \in H_{2,2d}$ arbitrary, this leads to $p(x, y) = r^{2d}q(\varphi)$, where, $q(\varphi)$ is a polynomial expression in $\cos \varphi, \sin \varphi$. In particular, $q: \mathbb{R} \rightarrow \mathbb{R}, \varphi \mapsto q(\varphi)$ is differentiable infinitely many times. In our case, we have

$$\begin{aligned} p_d(x, y) &= p_d(r, \varphi) = r^{2d} \left(\operatorname{Re}((\cos \varphi + i \sin \varphi)^{2d}) + 2d - 1 \right) \\ &= r^{2d} \left(\operatorname{Re} \left(e^{i2d\varphi} \right) + 2d - 1 \right). \end{aligned} \quad (40)$$

We can now apply Lemma 4.24 two times. First, for $k = 2d$ and $q(\varphi) = 1$, one can see that

$$\begin{aligned} \nabla^2 \left((2d - 1)r^{2d} \right) &= (2d - 1)\nabla^2 r^{2d} \\ &= (2d - 1)r^{2d-2} (2d \cdot (2d - 1) \cdot 1 \cdot E_{rr} + 0 + 2d \cdot 1 E_{\varphi\varphi}) \\ &= r^{2d-2} 2d(2d - 1) ((2d - 1)E_{rr} + E_{\varphi\varphi}). \end{aligned} \quad (41)$$

Further, using Lemma 4.24 for $k = 2d$ and $q(\varphi) = e^{i2d\varphi}$ reveals that

$$\begin{aligned} \nabla^2 \left(r^{2d} e^{i2d\varphi} \right) &= r^{2d-2} \left(2d(2d - 1) e^{i2d\varphi} E_{rr} + (2d - 1) 2di e^{i2d\varphi} E_{r\varphi} \right. \\ &\quad \left. + \left(2de^{i2d\varphi} - (2d)^2 e^{i2d\varphi} \right) E_{\varphi\varphi} \right) \\ &= r^{2d-2} 2d(2d - 1) e^{i2d\varphi} (E_{rr} + iE_{r\varphi} - E_{\varphi\varphi}). \end{aligned} \quad (42)$$

Summing both terms and taking the real part leads to the Hessian of p_d . Before we do so, let us make some observations, that help us in taking the real part

$$\begin{aligned} e^{i2d\varphi} - e^{-i2d\varphi} &= \cos(2d\varphi) + i \sin(2d\varphi) - (\cos(-2d\varphi) + i \sin(-2d\varphi)) \\ &= \cos(2d\varphi) + i \sin(2d\varphi) - (\cos(2d\varphi) - i \sin(2d\varphi)) = 2i \sin(2d\varphi) \end{aligned}$$

and analogously

$$\begin{aligned} e^{i2d\varphi} + e^{-i2d\varphi} &= \cos(2d\varphi) + i \sin(2d\varphi) + \cos(-2d\varphi) + i \sin(-2d\varphi) \\ &= \cos(2d\varphi) + i \sin(2d\varphi) + \cos(2d\varphi) - i \sin(2d\varphi) = 2 \cos(2d\varphi). \end{aligned}$$

This shows

$$\begin{aligned} &\operatorname{Re} \left(e^{i2d\varphi} (E_{rr} + iE_{r\varphi} - E_{\varphi\varphi}) \right) \\ &= \frac{1}{2} \left(e^{i2d\varphi} (E_{rr} + iE_{r\varphi} - E_{\varphi\varphi}) + \overline{e^{i2d\varphi} (E_{rr} + iE_{r\varphi} - E_{\varphi\varphi})} \right) \\ &= \frac{1}{2} \left(\left(e^{i2d\varphi} + e^{-i2d\varphi} \right) (E_{rr} - E_{\varphi\varphi}) + i \left(e^{i2d\varphi} - e^{-i2d\varphi} \right) E_{r\varphi} \right) \\ &= \frac{1}{2} (2 \cos(2d\varphi) (E_{rr} - E_{\varphi\varphi}) + i 2i \sin(2d\varphi) E_{r\varphi}) \\ &= \cos(2d\varphi) (E_{rr} - E_{\varphi\varphi}) - \sin(2d\varphi) E_{r\varphi}. \end{aligned} \quad (43)$$

For the Hessian $\nabla^2 p_d(r, \varphi)$ we obtain from (40), (41) and (42)

$$\begin{aligned}
& \nabla^2 p_d(r, \varphi) \\
&= \operatorname{Re} \left(\nabla^2 \left(r^{2d} e^{i2d\varphi} \right) + \nabla^2 \left((2d-1)r^{2d} \right) \right) \\
&= \operatorname{Re} \left[r^{2d-2} 2d(2d-1) e^{i2d\varphi} (E_{rr} + iE_{r\varphi} - E_{\varphi\varphi}) \right. \\
&\quad \left. + r^{2d-2} 2d(2d-1) ((2d-1)E_{rr} + E_{\varphi\varphi}) \right] \\
&= r^{2d-2} 2d(2d-1) \cdot \left[\operatorname{Re} \left(e^{i2d\varphi} (E_{rr} + iE_{r\varphi} - E_{\varphi\varphi}) \right) + (2d-1)E_{rr} + E_{\varphi\varphi} \right].
\end{aligned}$$

Using (43) this can further be simplified to

$$\begin{aligned}
& \nabla^2 p_d(r, \varphi) \\
&= r^{2d-2} 2d(2d-1) \cdot [\cos(2d\varphi)(E_{rr} - E_{\varphi\varphi}) - \sin(2d\varphi)E_{r\varphi} \\
&\quad + (2d-1)E_{rr} + E_{\varphi\varphi}].
\end{aligned}$$

Define

$$\begin{aligned}
A_{r\varphi} &:= [\cos(2d\varphi)(E_{rr} - E_{\varphi\varphi}) - \sin(2d\varphi)E_{r\varphi} + (2d-1)E_{rr} + E_{\varphi\varphi}] \\
&= [(\cos(2d\varphi) + (2d-1))E_{rr} - \sin(2d\varphi)E_{r\varphi} + (1 - \cos(2d\varphi))E_{\varphi\varphi}].
\end{aligned}$$

Then we have $\nabla^2 p_d(r, \varphi) = r^{2d-2} 2d(2d-1) \cdot A_{r\varphi}$. Since $r^{2d-2} 2d(2d-1) \geq 0$ is nonnegative, for p_d being convex it suffices to show that $A_{r\varphi} \succeq 0$ for all $(r, \varphi) \in \mathbb{R}_{\geq 0} \times \mathbb{R}$.

First, take a look at the entries on the diagonal of $A_{r\varphi}$. Using the formulas for E_{rr} , $E_{r\varphi}$ and $E_{\varphi\varphi}$ in Lemma 4.24, we obtain

$$\begin{aligned}
& (A_{r\varphi})_{1,1} \\
&= (\cos(2d\varphi) + (2d-1)) \cos^2 \varphi + 2 \sin(2d\varphi) \sin \varphi \cos \varphi + (1 - \cos(2d\varphi)) \sin^2 \varphi \\
&= \cos(2d\varphi) \underbrace{(\cos^2 \varphi - \sin^2 \varphi)}_{=\cos(2\varphi)} + \sin(2d\varphi) \underbrace{2 \sin \varphi \cos \varphi}_{=\sin(2\varphi)} + (2d-1) \underbrace{\cos^2 \varphi + \sin^2 \varphi}_{=1} \\
&= \underbrace{\cos(2d\varphi) \cos(2\varphi) + \sin(2d\varphi) \sin(2\varphi)}_{=\cos(2d\varphi-2\varphi)} + (2d-1) - (2d-2) \sin^2 \varphi \\
&= 2d-1 + \underbrace{\cos(2d\varphi-2\varphi) - (2d-2) \sin^2 \varphi}_{\geq -(2d-1)} \geq 0.
\end{aligned}$$

Analogously, one can see that

$$\begin{aligned}
& (A_{r\varphi})_{2,2} \\
&= (\cos(2d\varphi) + (2d-1)) \sin^2 \varphi - 2 \sin(2d\varphi) \sin \varphi \cos \varphi + (1 - \cos(2d\varphi)) \cos^2 \varphi \\
&= \cos(2d\varphi) \underbrace{(\sin^2 \varphi - \cos^2 \varphi)}_{=-\cos(2\varphi)} - \sin(2d\varphi) \underbrace{2 \sin \varphi \cos \varphi}_{=\sin(2\varphi)} + (2d-1) \underbrace{\sin^2 \varphi + \cos^2 \varphi}_{=1} \\
&= \underbrace{-\cos(2d\varphi) \cos(2\varphi) - \sin(2d\varphi) \sin(2\varphi)}_{=-\cos(2d\varphi-2\varphi)} + (2d-1) - (2d-2) \cos^2 \varphi \\
&= 2d-1 - \underbrace{\cos(2d\varphi-2\varphi) - (2d-2) \cos^2 \varphi}_{\geq -(2d-1)} \geq 0.
\end{aligned}$$

Further:

$$\begin{aligned}
& (A_{r\varphi})_{1,2} = (A_{r\varphi})_{2,1} \\
& = (\cos(2d\varphi) + (2d - 1)) \cos \varphi \sin \varphi - \sin(2d\varphi)(\cos^2 \varphi - \sin^2 \varphi) \\
& \quad - (1 - \cos(2d\varphi)) \sin \varphi \cos \varphi \\
& = \underbrace{2 \cos(2d\varphi) \sin \varphi \cos \varphi}_{=\cos(2d\varphi) \sin(2\varphi)} + \underbrace{(2d - 2) \cos \varphi \sin \varphi}_{=(d-1) \sin(2\varphi)} - \sin(2d\varphi) \underbrace{(\cos^2 \varphi - \sin^2 \varphi)}_{\cos(2\varphi)} \\
& = \underbrace{\cos(2d\varphi) \sin(2\varphi) - \sin(2d\varphi) \cos(2\varphi)}_{=\sin(2\varphi-2d\varphi)} + (d - 1) \sin(2\varphi) \\
& = -\sin(2d\varphi - 2\varphi) + (d - 1) \sin(2\varphi).
\end{aligned}$$

Hence, we have for the determinant

$$\begin{aligned}
\det(A_{r\varphi}) &= (A_{r\varphi})_{11}(A_{r\varphi})_{22} - (A_{r\varphi})_{21}(A_{r\varphi})_{12} \\
&= ([2d - 1 + \cos(2d\varphi - 2\varphi)] - (2d - 2) \sin^2 \varphi) \\
& \quad \cdot ([2d - 1 - \cos(2d\varphi - 2\varphi)] - (2d - 2) \cos^2 \varphi) \\
& \quad - (-\sin(2d\varphi - 2\varphi) + (d - 1) \sin(2\varphi))^2 \\
&= (2d - 1)^2 - \cos^2(2d\varphi - 2\varphi) - ((2d - 1) + \cos(2d\varphi - 2\varphi))(2d - 2) \cos^2 \varphi \\
& \quad - ((2d - 1) - \cos(2d\varphi - 2\varphi))(2d - 2) \sin^2 \varphi + (2d - 2)^2 \sin^2 \varphi \cos^2 \varphi \\
& \quad - \sin^2(2d\varphi - 2\varphi) + 2(d - 1) \sin(2d\varphi - 2\varphi) \sin(2\varphi) - (d - 1)^2 \sin^2(2\varphi).
\end{aligned}$$

Rearranging the terms leads to

$$\begin{aligned}
& \det(A_{r\varphi}) \\
&= (2d - 1)^2 - \underbrace{\cos^2(2d\varphi - 2\varphi) - \sin^2(2d\varphi - 2\varphi)}_{=-1} \\
& \quad - (2d - 1)(2d - 2) \underbrace{(\cos^2 \varphi + \sin^2 \varphi)}_{=1} \\
& \quad + (2d - 2) \cos(2d\varphi - 2\varphi) \underbrace{(\sin^2 \varphi - \cos^2 \varphi)}_{-\cos(2\varphi)} + 2(d - 1) \sin(2d\varphi - 2\varphi) \sin(2\varphi) \\
& \quad \underbrace{\hspace{10em}}_{=-(2d-2) \cos(2d\varphi)} \\
& \quad + \underbrace{(2d - 2)^2 \sin^2 \varphi \cos^2 \varphi - (d - 1)^2 \sin^2(2\varphi)}_{=(d-1)^2((2 \sin \varphi \cos \varphi)^2 - \sin^2(2\varphi))=0} \\
&= (2d - 1)^2 - 1 - (2d - 1)(2d - 2) - (2d - 2) \cos(2d\varphi) \\
&= 2d - 2 - \underbrace{(2d - 2) \cos(2d\varphi)}_{\geq -(2d-2)} \geq 0.
\end{aligned}$$

Finally, since $(A_{r\varphi})_{1,1}, (A_{r\varphi})_{2,2} \geq 0$ and $\det(A_{r\varphi}) \geq 0$, we know that $A_{r\varphi} \succeq 0$ and hence the Hessian $\nabla^2 p_d(r, \varphi)$ is positive semidefinite. Therefore, p_d is convex. \square

For proving identity (37), we also need the Laplacian of p_d .

Remark 4.26. We use without proof that $\Delta((x+iy)^{2d}) = 0$. This follows from the standard result of Functional Analysis that the Laplacian of a holomorphic function is zero, see e.g. [10, Section 5].

Proposition 4.27. Let $p_d \in C_{2,2d}$ be defined as in Proposition 4.25. Then the Laplacian of p_d in polar coordinates is given by

$$\Delta^d p_d = \Delta^d \operatorname{Re}((x+iy)^{2d}) + \Delta^d (2d-1)r^{2d} = (2d-1)2^{2d}(d!)^2.$$

Proof. Let $(x, y)^\top = r(\cos \varphi, \sin \varphi)^\top \in \mathbb{R}^2$ be arbitrary. Using (40) and the definition of p_d in Proposition (4.28), we know

$$p_d = \operatorname{Re}((x+iy)^{2d}) + r^{2d}(2d-1).$$

Further, Remark 4.26 tells us that $\Delta((x+iy)^{2d}) = 0$. This yields that $\Delta \operatorname{Re}((x+iy)^{2d}) = 0$ and hence by induction $\Delta^k \operatorname{Re}((x+iy)^{2d}) = 0$ for all $k \in \mathbb{N}$. On the other hand, by Lemma 4.24, we know

$$\Delta(2d-1)r^{2d} = (2d-1)\Delta r^{2d} = (2d-1)r^{2d-2}(2d)^2 = (2d-1)2^2 d^2 r^{2(d-1)}.$$

Inductively, this shows

$$\begin{aligned} \Delta^d(2d-1)r^{2d} &= (2d-1)\Delta^d r^{2d} = (2d-1) \left(2^2 d^2 \Delta^{d-1} r^{2(d-1)} \right) \\ &= (2d-1) \left(2^{2+2} (d(d-1))^2 \Delta^{d-2} r^{2(d-2)} \right) \\ &= (2d-1)2^{2d}(d!)^2. \end{aligned}$$

This shows $\Delta^d p_d = \Delta^d \operatorname{Re}((x+iy)^{2d}) + \Delta^d (2d-1)r^{2d} = (2d-1)2^{2d}(d!)^2$. \square

Proposition 4.28. Let $p_d \in C_{2,2d}$ be defined as in Proposition 4.25. Further, take $z := \begin{pmatrix} 1 \\ i \end{pmatrix} = e^1 + ie^2 \in \mathbb{C}^2$. Then equality (37) holds, i.e.

$$\operatorname{Re}(p_d(z)) = \frac{\binom{2(d-1)}{d-1}}{d} Q_{p_d}(z, \bar{z}).$$

In particular, we have $B_d^* \geq \frac{\binom{2(d-1)}{d-1}}{d}$.

Proof. For showing that (37) holds at $z = (1, i)^\top$, we have to compute $p_d(z)$ and $Q_{p_d}(z, \bar{z})$. recall that $p_d = \operatorname{Re}((\mathbf{x}+i\mathbf{y})^{2d}) + (2d-1)(\mathbf{x}^2 + \mathbf{y}^2)^d$. As in Remark 4.23 we see that

$$\begin{aligned} p_d(1, i) &= \operatorname{Re}((\mathbf{x}+i\mathbf{y})^{2d})(1, i) + \underbrace{(2d-1)(\mathbf{x}^2 + \mathbf{y}^2)^d}_{=0}(1, i) \\ &= \frac{1}{2} \left((\mathbf{x}+i\mathbf{y})^{2d} + (\mathbf{x}-i\mathbf{y})^{2d} \right) (1, i) = 2^{2d-1} \end{aligned}$$

Now we have to compute $Q_{p_d}(z, \bar{z})$. Using the differential operator representation of $p \mapsto Q_p(z, \bar{z})$ in Example 3.59(d), we know that

$$Q_{p_d}(z, \bar{z}) = \frac{1}{(2d)!} (\partial_{e^1}^2 + \partial_{e^2}^2) p_d = \frac{1}{(2d)!} (\partial^2 \mathbf{x} + \partial^2 \mathbf{y}) p_d = \frac{1}{(2d)!} \Delta^d p_d.$$

In Proposition 4.27, we have seen that $\Delta^d p_d = (2d-1)2^{2d}(d!)^2$. This yields

$$\begin{aligned} Q_{p_d}(z, \bar{z}) &= \frac{1}{(2d)!} (2d-1)2^{2d}(d!)^2 \\ &= \frac{1}{2d} \frac{1}{(2d-2)!} 2^{2d} d^2 ((d-1)!)^2 \\ &= 2^{2d-1} d \binom{2d-2}{d-1}^{-1} = 2^{2d-1} d \binom{2(d-1)}{d-1}^{-1} \\ \implies \operatorname{Re}(p_d(z)) &= \operatorname{Re}(p_d(1, i)) = 2^{2d-1} = \frac{\binom{2(d-1)}{d-1}}{d} Q_{p_d}(z, \bar{z}), \end{aligned}$$

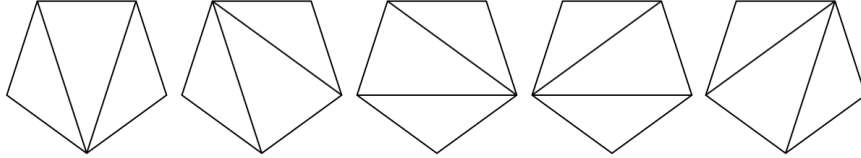
which shows that equality (37) holds. Therefore, inequality (28) of Lemma 4.7 holds with equality for the form $p_d \in C_{2,2d}$ as in the claim and constant $B_d = \frac{\binom{2(d-1)}{d-1}}{d}$. This yields that $B_d^* \geq \frac{\binom{2(d-1)}{d-1}}{d}$. \square

Combining Propositions 4.28 and 4.21, we obtain the following.

Proposition 4.29. *Let $d \in \mathbb{N}$ be arbitrary. The optimal constant $B_d^* > 0$ of the Generalized Cauchy–Schwarz Inequalities (24) is given by $B_d^* = \frac{\binom{2(d-1)}{d-1}}{d}$.*

Proof. This is clear now, since by Proposition 4.21, $\frac{\binom{2(d-1)}{d-1}}{d}$ is an upper bound for B_d^* and by Proposition 4.28 a lower bound for B_d^* . \square

Remark 4.30. For $d \in \mathbb{N}$, the value $c_d := \frac{1}{d+1} \cdot \binom{2d}{d} = B_{d+1}^*$ is known as the d -th **Catalan Number** (see [17]). For $d \in \mathbb{N}_{\geq 3}$, the number c_{d-2}^* is for example the number triangulations of a convex polygon with d sides. For $d = 5$ for example, there are $c_3 = \frac{1}{4} \cdot \binom{6}{3} = 5$ possible triangulations of a regular pentagon, as one can see in the following picture



5. CHARACTERIZATION OF THE SOS CONE $\Sigma_{n,2d}$ INSIDE THE PSD CONE $P_{n,2d}$

In this section, we are interested in characterizing the SOS cone $\Sigma_{n,2d}$ inside the PSD cone $P_{n,2d}$ for the cases $(n, 2d) \in \{(4, 4), (3, 6)\}$. As in [12, Section 4], this characterization uses the structure of extreme rays of $\Sigma_{n,2d}^*$ (further details can be found in [4]). The essential tool for doing so is the Cayley–Bacharach relation, which is studied in detail in the Sections 5.1 and 5.2. For further informations about the Cayley–Bacharach relation, the reader is referred to [11]. From there we proceed as in [4] to obtain in Section 5.3 a characterization of the extreme rays of $\Sigma_{n,2d}^*$ that separate $\Sigma_{n,2d}$ from $P_{n,2d}$. The structure of those extreme rays is further analyzed in Section 5.4

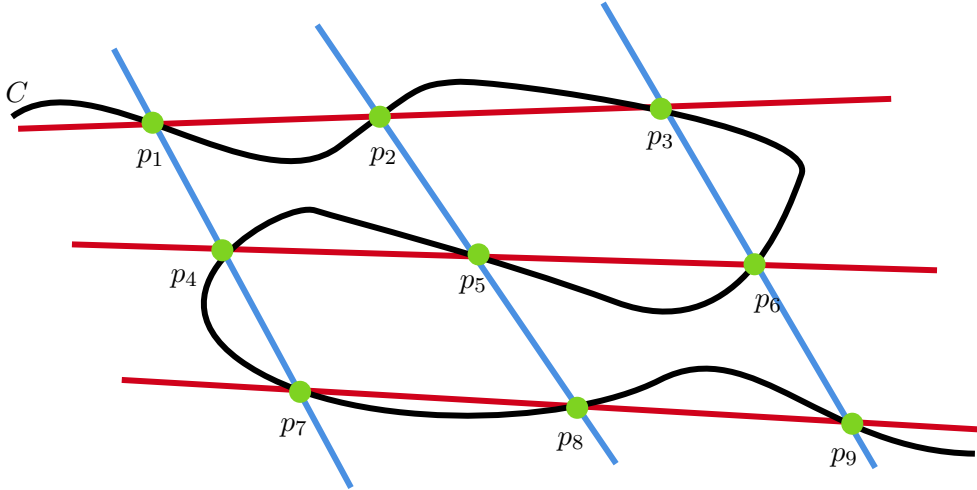
by using techniques from optimization including the Karush Kuhn Tucker (KKT) conditions.

In Section 6 we will use this characterization to prove that every convex quaternary quartic is indeed SOS. The same holds for ternary sextics, if a conjecture formulated in [4] turns out to be true.

5.1. The Projective Space and Chasles' Version of the Cayley

Bacharach Theorem. The goal of this section is to get familiar with the Cayley–Bacharach relation by seeing one of its first versions introduced by Chasles in 1865.

Suppose you have a set of points $\{p_1, \dots, p_s\}$ ($s \in \mathbb{N}$) in a vector space V and a curve C that passes through all but one point. One could then ask under which assumptions, the curve also needs to pass through the remaining point. In the following, one can see an illustration of this problem on the plane where the nine points p_1, \dots, p_9 come from the intersection of six lines. A formal answer to this problem can be found in Theorem 5.20.



Before we can state Theorem 5.20, we need some definitions from Algebraic Geometry including the projective space as well as affine and projective varieties. For more details the reader is referred to e.g. [7, Chapter 1., Chapter 8.]. We further introduce the notion of transversal intersection, which is a concept of Differential Geometry and can e.g. be found in [15, Chapter 2, §5].

Recall that by Convention 3.38, we have $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 5.1. Let V be a vector space over \mathbb{K} . Define an equivalence relation on V via

$$x \sim y \Leftrightarrow \exists \lambda \in \mathbb{K} \setminus \{0\} : x = \lambda y.$$

The (real or complex) **projective space** $\mathbb{P}(V)$ is defined as the quotient space of V with respect \sim , i.e. $\mathbb{P}(V) = V/\sim$. If $V = \mathbb{K}^{n+1}$ for some $n \in \mathbb{N}$, $\mathbb{P}^n \mathbb{K} := \mathbb{P}(\mathbb{K}^{n+1})$ is called the **projective space of dimension n over \mathbb{K}** . In this case, we denote the elements of the equivalence classes, i.e. the elements of $\mathbb{P}^n \mathbb{K}$ as $(x) = (x_0 : \dots : x_n)$, where $x = (x_0, \dots, x_n)^\top \in$

$\mathbb{K}^{n+1} \setminus \{o\}$. We call $y \in \mathbb{K}^{n+1} \setminus \{o\}$ **homogeneous coordinates** of (x) if $y \sim x$.

Next, we define (algebraic) varieties, which are basically zero sets of polynomials. Note that a homogeneous polynomial $p \in H_{n,k}$ has a zero at some point $x \in \mathbb{K}^n$ if and only if it vanishes on the whole line $\{\lambda x : \lambda \in \mathbb{K}\} \subseteq \mathbb{K}^n$, i.e. at every homogeneous coordinate of the projective point $(x) \in \mathbb{P}^{n-1}\mathbb{K}$. Hence, it makes sense to distinguish between affine and projective varieties, where projective varieties are generated by homogeneous polynomials.

Definition 5.2. A subset $V \subseteq \mathbb{K}^n$ is an **affine variety**, if it is the zero set of a finite number of polynomials, i.e. there are $p_1, \dots, p_m \in \mathbb{R}[\mathbf{x}]$ ($m \in \mathbb{N}$) such that

$$\begin{aligned} V &= V(p_1, \dots, p_m) := V_{\mathbb{K}}(p_1, \dots, p_m) \\ &:= \{x \in \mathbb{K}^n \mid \forall i = 1, \dots, m : p_i(x) = 0\}. \end{aligned}$$

Analogously, a subset $V^h \subseteq \mathbb{P}^{n-1}\mathbb{K}$ is a **projective variety** if it is the zero set of a finite set of forms, i.e. there are $p_1, \dots, p_m \in \bigcup_{k \in \mathbb{N}} H_{n,k}$ ($m \in \mathbb{N}$) such that

$$\begin{aligned} V^h &= V^h(p_1, \dots, p_m) := V_{\mathbb{K}}^h(p_1, \dots, p_m) \\ &:= \{(x) \in \mathbb{P}^{n-1}\mathbb{K} \mid \forall i = 1, \dots, m : p_i(x) = 0\}. \end{aligned}$$

Remark 5.3. A finite set $\Gamma = \{\gamma^1, \dots, \gamma^d\} \subseteq \mathbb{P}^{n-1}\mathbb{K}$ containing d projective points is always a projective variety. More precisely, it can be generated by a single form with degree at most d [28, Proposition 1.2]. We will use this result in the proof of Lemma 5.27.

We have particular interest in some special kinds of varieties, which are defined in the following (see [14, Chapter 1] and [7, Chapter 8, §2]).

Definition 5.4. An affine variety generated by a single polynomial in two variables is called an **affine plane curve**. Analogously, a projective variety generated by a single form in three variables is called a **projective plane curve**.

In general, an affine or projective variety generated by a single polynomial in n (affine case) or $n + 1$ (projective case) variables is called an affine or projective **hypersurface**.

Further, we need the notion of the degree of a hypersurface.

Definition 5.5. If H is a (affine or projective) hypersurface (or curve) generated by the polynomial $p \in \mathbb{R}[\mathbf{x}]$ of degree d , we say that V is a hypersurface (or curve) of **degree** d .

Starting with a variety V , we can also define the set of polynomials vanishing on it. Therefore, we recall the definition of an ideal.

Definition 5.6. A subset $I \subseteq \mathbb{K}[\mathbf{x}]$ is called an **ideal**, if the following holds

- (1) $0 \in I$,
- (2) $\forall f, g \in I : f + g \in I$,
- (3) $\forall f \in I \forall h \in \mathbb{K}[\mathbf{x}] : hf \in I$.

In this case, we also write $I \triangleleft \mathbb{K}[\mathbf{x}]$.

It can easily be seen that the polynomials vanishing on a (affine or projective) variety form an ideal. This leads to the following definition.

Definition 5.7. Let $n \in \mathbb{N}$ and $V \subseteq \mathbb{K}^n$ be an affine variety. Then the **vanishing ideal** of V is the set of all polynomials vanishing on V , i.e.

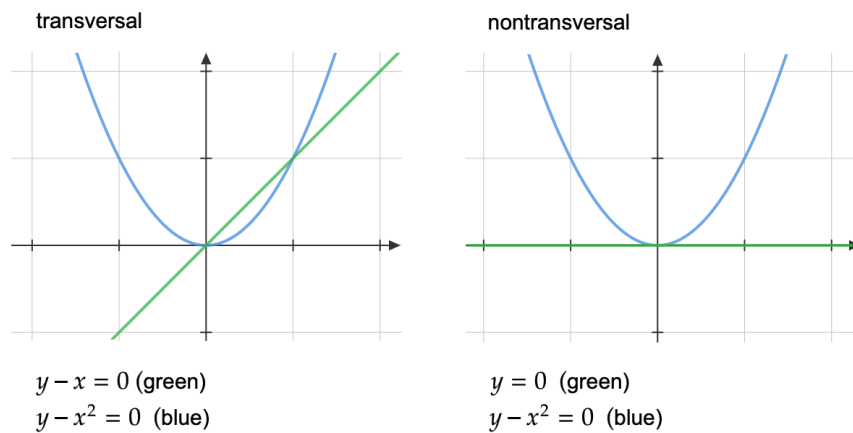
$$I(V) = \{p \in \mathbb{K}[\mathbf{x}] \mid \forall v \in V : p(v) = 0\}.$$

If $V^h \subseteq \mathbb{P}^{n-1}\mathbb{K}$ is a projective variety, the **vanishing ideal** of V^h is the set of polynomials vanishing on all homogeneous coordinates of elements of V^h , i.e.

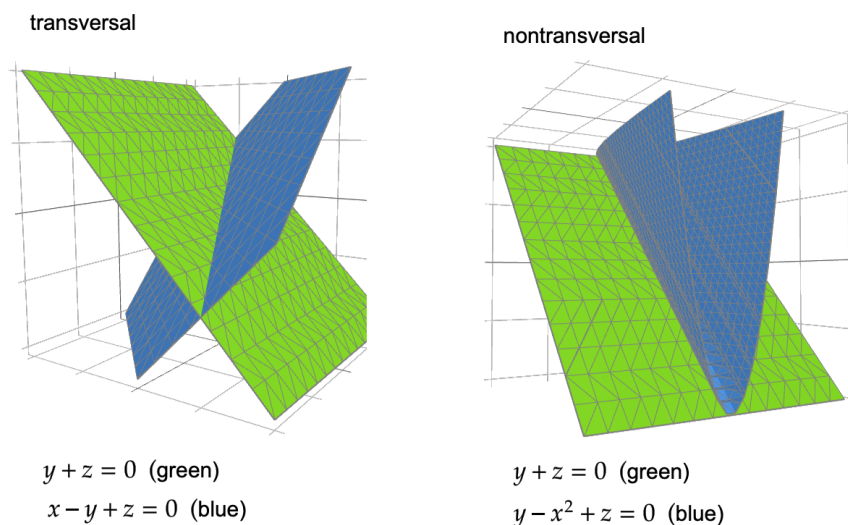
$$I(V^h) = \left\{ p \in \mathbb{K}[\mathbf{x}] \mid \forall (v) \in V^h, v \in \mathbb{K}^n : p(v) = 0 \right\}.$$

We are not only interested in the zero set of a polynomial, but also in the intersection of two or more polynomials or - equivalently - the intersection of hypersurfaces. More precisely, we want to distinguish between transversal and nontransversal intersections. It should be mentioned that transversal intersection is a wellknown notion in Differential Geometry, which is usually stated in terms of manifolds as in [15, Chapter 2, §5]. However, it can also be defined for hypersurfaces as they can be interpreted as submanifolds in the Euclidean space \mathbb{K}^n (see [24, Chapter 4]). To get an intuition of transversal and nontransversal intersections, consider the following examples.

Example 5.8. (a) Curves in \mathbb{R}^2 :



(b) Surfaces in \mathbb{R}^3 :



When we take a look at the examples above, we can get an intuitive idea on how to define transversal intersections. Two hypersurfaces $X_1, X_2 \subseteq \mathbb{K}^n$ intersect transversally in a common point $z \in \mathbb{K}^n$ if their tangents at z span the whole \mathbb{K}^n . Therefore, we need the notion of tangent spaces of an affine curve or hypersurface.

The following definition can be found in [7, Chapter 9, §6].

Definition 5.9. Let $V \subseteq \mathbb{K}^n$ be an affine variety, $p \in \mathbb{K}[\mathbf{x}]$ a polynomial and $z = (z_1, \dots, z_n)^\top \in \mathbb{K}^n$ a point. Then the **linear part** of p at z is defined as

$$d_z(p) = (\partial_{\mathbf{x}_1} p)(z)(\mathbf{x}_1 - z_1) + \dots + (\partial_{\mathbf{x}_n} p)(z)(\mathbf{x}_n - z_n) \in \mathbb{K}[\mathbf{x}].$$

Note that $\deg(d_z(p)) \leq 1$.

Further, the **tangent space** of V at z is the variety

$$T_z(V) = V(\{d_z(p) : p \in I(V)\}).$$

Remark 5.10. Note that in the following, we are mainly interested in the case where V is a hypersurface, i.e. $V = V(p)$ for some polynomial $p \in \mathbb{K}[\mathbf{x}]$. From the Hilbert Nullstellensatz (see [7, Chapter 4, §2, Theorem 2]) we know that for an arbitrary ideal $I \triangleleft \mathbb{K}[\mathbf{x}]$ it holds

$$I(V(I)) = \sqrt{I},$$

where $\sqrt{I} = \{p \in \mathbb{K}[\mathbf{x}] \mid \exists m \in \mathbb{N} : p^m \in I\}$ is the **radical** of I (see [7, Chapter 4, §2]). Further, if $I = \langle p \rangle$ is generated by a single polynomial and p has a factorization of the form $p = cp_1^{\alpha_1} \cdots p_r^{\alpha_r}$ in distinct irreducible factors $p_i \in \mathbb{K}[\mathbf{x}]$, where $c \neq 0$ and $\alpha_i \in \mathbb{N}_{>0}$ ($i = 1, \dots, r$), it holds by [7, Chapter 4, §2, Proposition 9] that

$$\sqrt{I} = \langle p_1, \dots, p_r \rangle.$$

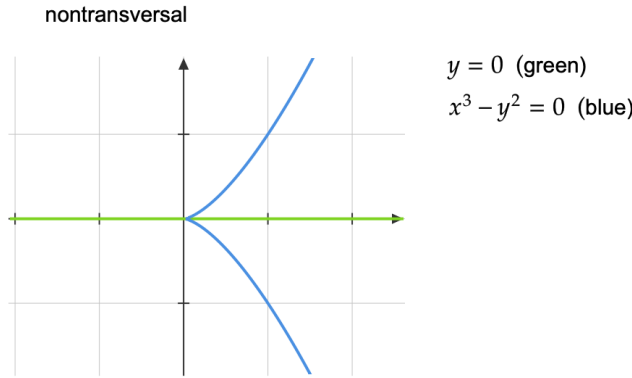
Moreover, note that if we define $p_{\text{red}} = p_1 \cdots p_r$, then $V(p) = V(p_{\text{red}})$.

This means in particular that if p is already irreducible, the tangent space of $V = V(p)$ at z is determined by

$$T_z(V(p)) = V(d_z(p)),$$

i.e. the variety of the linear part of p at z .

Remark 5.11. There is another thing, which has to be considered when we want to define transversal intersections. Therefore, consider the following intersection in \mathbb{R}^2 :



From an intuitive standpoint, the intersection is nontransversal. However, since for $p(\mathbf{x}, \mathbf{y}) = \mathbf{x}^3 - \mathbf{y}^2$ we have $\partial_{\mathbf{x}}p(\mathbf{x}, \mathbf{y}) = 3\mathbf{x}^2$, $\partial_{\mathbf{y}}p(\mathbf{x}, \mathbf{y}) = -2\mathbf{y}$ and hence, for $z = (0, 0)^\top$ it holds $d_z(p) = 0$. This means that the tangent space of $V(p)$ at z is already the whole \mathbb{R}^2 , i.e. $T_z(V(p)) = \mathbb{R}^2$.

Hence, it does not suffice to say that two hypersurfaces $V(p), V(q)$ intersect transversally at z if their corresponding tangent spaces span the whole space. Instead, we have to say in addition that the partial derivatives must not all vanish at z , which leads to the following definition of nonsingular points.

Definition 5.12. Let $V = V(p) \subseteq \mathbb{K}^n$ be an affine hypersurface generated by p such that p is, as in Remark 5.10, without loss of generality the product of single, distinct, irreducible factors, i.e. $p = p_1 \cdots p_r$, where $p_i \in \mathbb{K}[\mathbf{x}]$ are irreducible. Let $z \in \mathbb{C}^n$ be arbitrary. Then we call z a **singular** point of V if all partial derivatives of p vanish at z , i.e. $(\partial_{\mathbf{x}_i}p)(z) = 0$ ($i = 1, \dots, n$). Otherwise, z is called a **nonsingular** point of V .

Remark 5.13. Note that the general definition of a singular point of a variety V can be found in [7, Chapter 9] and involves the dimension of a variety. However, since for our purposes it suffices to define singular points only for hypersurfaces, we use the results of [7, Chapter 9, §6, Example 4] to give in Definition 5.12 an equivalent characterization of singular points of hypersurfaces. This allows us to avoid the formal definition of the dimension of a variety.

Now we can finally define the transversal intersection of two (affine) hypersurfaces as in [15, Chapter 2].

Definition 5.14. Let $X_1, X_2 \subseteq \mathbb{K}^n$ be two affine hypersurfaces and let $z \in X_1 \cap X_2$ be such that z is a nonsingular point of both X_1 and X_2 .

Then we say that X_1 and X_2 intersect **transversally** at $z \in \mathbb{K}^n$ if their corresponding tangent spaces span the whole \mathbb{K}^n , i.e.

$$T_z(X_1) + T_z(X_2) = \mathbb{K}^n.$$

Example 5.15. In the following, we take again a look at the curves given in Example 5.8(a).

- Consider $p_1 = \mathbf{y} - \mathbf{x}^2, p_2 = \mathbf{y} - \mathbf{x} \in \mathbb{R}[\mathbf{x}, \mathbf{y}]$. We want to show that $X_1 = V(p_1), X_2 = V(p_2) \subseteq \mathbb{R}^2$ intersect transversally at the point $z = (0, 0)^\top \in \mathbb{R}^2$. First of all, it can easily be seen that both p_1 and p_2 are irreducible over $\mathbb{R}[\mathbf{x}, \mathbf{y}]$. Further, z is a nonsingular point of both X_1 and X_2 since $(\partial \mathbf{y} p_1)(z) = 1 = (\partial \mathbf{y} p_2)(z)$. Hence, we know by Remark 5.10 that

$$I(V(p_1)) = \langle p_1 \rangle, \quad I(V(p_2)) = \langle p_2 \rangle.$$

Further, since $(\partial \mathbf{x} p_1)(z) = 0, (\partial \mathbf{y} p_2)(z) = -1$, we have

$$d_z(p_1) = \mathbf{y}, \quad d_z(p_2) = -\mathbf{x} + \mathbf{y}.$$

and hence $T_z(X_1) = \text{span}_{\mathbb{R}}((0, 1)^\top), T_z(X_2) = \text{span}_{\mathbb{R}}((1, 1)^\top)$. This shows that $T_z(X_1) + T_z(X_2) = \mathbb{R}^2$, which means that X_1 and X_2 intersect transversally at z , as desired.

- Now consider $p_1 = \mathbf{y} - \mathbf{x}^2, p_2 = \mathbf{y} \in \mathbb{R}[\mathbf{x}, \mathbf{y}]$. As above, we know that $z = (0, 0)^\top$ is a nonsingular point of both $X_1 = V(p_1)$ and $X_2 = V(p_2)$. Further, one can see that $d_z(p_1) = \mathbf{y} = d_z(p_2)$. For this reason, the tangent spaces at z , coincide, more precisely

$$T_z(X_1) = T_z(X_2) = \text{span}_{\mathbb{R}}((0, 1)^\top).$$

This shows that the intersection of X_1 and X_2 at z is not transversal.

An analogous observation can be made for the hypersurfaces in Example 5.8(b).

Remark 5.16. Let $p \in \mathbb{K}[\mathbf{x}]$ be an arbitrary polynomial. If $z \in \mathbb{K}^n$ is a nonsingular point of p , the tangent space $T_z(V(p))$ has dimension $n - 1$ as a \mathbb{K} -vector space (see [7, Chapter 9, §6, Example 4]). In particular, this means for $n = 2$ that two affine curves $V(p_1), V(p_2) \subseteq \mathbb{K}^n$ ($p_1, p_2 \in \mathbb{K}[\mathbf{x}, \mathbf{y}]$) intersect transversally at $z \in \mathbb{K}^2$ if and only if z is a nonsingular point of both $V(p_1)$ and $V(p_2)$ and p_1 and p_2 have distinct tangent directions at z . This is precisely the definition of transversal curves given in [33, Chapter A.3] or [23, Chapter 7].

The projective case can be traced back to the affine case using dehomogenization as in [33, Chapter A.2].

Definition 5.17. Let $X = V^h(p) \subseteq \mathbb{P}^{n-1}\mathbb{K}$ be a projective hypersurface such that $p \in \mathbb{K}[\mathbf{x}]$ is without loss of generality the product of its single, distinct, irreducible factors $p = p_1 \cdots p_r$. Let $k \in \{1, \dots, n\}$ be arbitrary. Then we call

$$\tilde{p} = p(\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, 1, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n) \in \mathbb{K}[\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n]$$

the **dehomogenization** of p with respect to the k -th variable and $\tilde{X} = V(\tilde{p}) \subseteq \mathbb{K}^n$ the **affine part** of X with respect to \mathbf{x}_k .

Further, if $z \in \mathbb{K}^n$ such that $z_k \neq 0$, we call

$$\tilde{z} = \left(\frac{z_1}{z_k}, \dots, \frac{z_{k-1}}{z_k}, \frac{z_{k+1}}{z_k}, \dots, \frac{z_n}{z_k} \right) \in \mathbb{K}^{n-1}$$

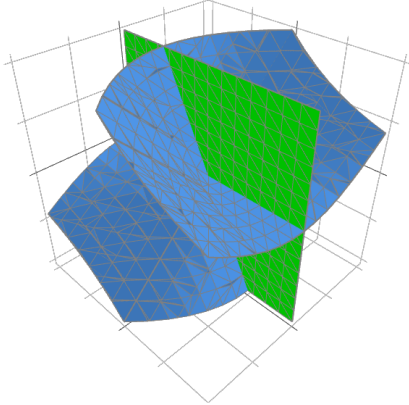
the $(n-1)$ -dimensional **projection** of z with respect to z_k .

Now we can define the transversal intersection of projective hypersurfaces as in [33, Chapter A.2].

Definition 5.18. Let $X_1, X_2 \subseteq \mathbb{P}^{n-1}\mathbb{K}$ be projective hypersurfaces, $\gamma \in X_1 \cap X_2$ and $z \in \mathbb{K}^n \setminus \{0\}$ homogeneous coordinates of γ . Let $z_k \neq 0$ for some $k \in \{1, \dots, n\}$. Further, let $\tilde{X}_1, \tilde{X}_2 \subseteq \mathbb{K}^{n-1}$ be the affine part of X_1, X_2 , respectively with respect to \mathbf{x}_k and $\tilde{z} \in \mathbb{K}^{n-1}$ the $(n-1)$ -dimensional projection of z with respect to z_k . Continue this process if necessary until \tilde{X}_1, \tilde{X}_2 are generated by nonhomogeneous polynomials. Then we say that X_1, X_2 intersect **transversally** at γ if the affine hypersurfaces \tilde{X}_1, \tilde{X}_2 intersect transversally at \tilde{z} .

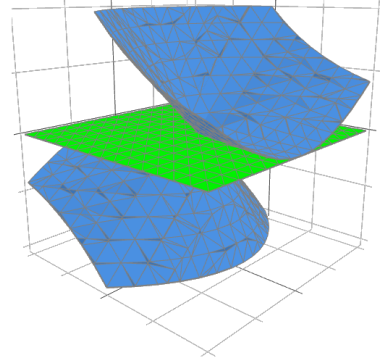
Example 5.19. Consider the following examples of intersections of projective hypersurfaces in $n = 3$ variables:

transversal



$$\begin{aligned} p_1 &= z^2 + 2zy - x^2 \quad (\text{blue}) \\ p_2 &= x \quad (\text{green}) \end{aligned}$$

nontransversal



$$\begin{aligned} p_1 &= z^2 + 2zy - x^2 \quad (\text{blue}) \\ p_2 &= z \quad (\text{green}) \end{aligned}$$

- On the left hand side, we have $p_1(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{z}^2 + 2\mathbf{z}\mathbf{y} - \mathbf{x}^2$, $p_2(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{x} \in \mathbb{R}[\mathbf{x}, \mathbf{y}, \mathbf{z}]$. Clearly, the projective hypersurfaces $V^h(p_1), V^h(p_2)$ intersect at $\gamma = (0 : 1 : 0) \in \mathbb{P}^2\mathbb{R}$. Take $z = (0, 1, 0)^\top \in \mathbb{R}^3$. Dehomogenization with respect to \mathbf{y} leads to

$$\tilde{p}_1(\mathbf{x}, \mathbf{z}) = \mathbf{z}^2 + 2\mathbf{z} - \mathbf{x}^2, \quad \tilde{p}_2(\mathbf{x}, \mathbf{z}) = \mathbf{x} \in \mathbb{R}[\mathbf{x}, \mathbf{z}],$$

which are both irreducible over $\mathbb{R}[\mathbf{x}, \mathbf{z}]$ and intersect at $\tilde{z} = (0, 0)^\top \in \mathbb{R}^2$. Further, we have

$$\begin{aligned} (\partial_{\mathbf{x}}\tilde{p}_1)(\tilde{z}) &= 0, \quad (\partial_{\mathbf{z}}\tilde{p}_1)(\tilde{z}) = 2, \\ (\partial_{\mathbf{x}}\tilde{p}_2)(\tilde{z}) &= 1, \quad (\partial_{\mathbf{z}}\tilde{p}_2)(\tilde{z}) = 0 \end{aligned}$$

and hence $d_{\tilde{z}}(\tilde{p}_1) = 2\mathbf{z}$, $d_{\tilde{z}}(\tilde{p}_2) = \mathbf{x}$. This shows that

$$T_{\tilde{z}}(V(\tilde{p}_1)) = \text{span}_{\mathbb{R}}((0, 1)^\top), \quad T_{\tilde{z}}(V(\tilde{p}_2)) = \text{span}_{\mathbb{R}}((1, 0)^\top),$$

which yields that the intersection is transversal.

- On the right hand side, we consider $p_1(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{z}^2 + 2\mathbf{z}\mathbf{y} - \mathbf{x}^2$, $p_2(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{z} \in \mathbb{R}[\mathbf{x}, \mathbf{y}, \mathbf{z}]$, which also intersect at $\gamma = (0 : 1 : 0) \in \mathbb{P}^2\mathbb{R}$. Dehomogenization with respect to \mathbf{y} leads to

$$\tilde{p}_1(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{z}^2 + 2\mathbf{z} - \mathbf{x}^2, \quad \tilde{p}_2(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \mathbf{z} \in \mathbb{R}[\mathbf{x}, \mathbf{z}].$$

Further, for $\tilde{z} = (0, 0)^\top \in \mathbb{R}^2$ we have for the linear parts $d_{\tilde{z}}(\tilde{p}_1) = 2\mathbf{z}$ and $d_{\tilde{z}}(\tilde{p}_2) = \mathbf{z}$. For this reason, the tangent spaces at \tilde{z} coincide and we have in particular

$$T_{\tilde{z}}(V(\tilde{p}_1)) = \text{span}_{\mathbb{R}}((0, 1)^\top) = T_{\tilde{z}}(V(\tilde{p}_2)),$$

showing that the intersection is nontransversal.

Finally, we can formally state the version of the Cayley–Bacharach Theorem, which covers the example illustrated at the beginning of the section. This version goes back to the French Mathematician Chasles and was published in 1865. A proof can be found in [11, Theorem (CB3)].

Theorem 5.20 (Chasles). *Let $X_1, X_2 \subseteq \mathbb{P}^2\mathbb{C}$ be cubic plane curves, i.e. curves generated by forms of degree 3 meeting transversally in 9 points $\gamma^1, \dots, \gamma^9 \in \mathbb{P}^2\mathbb{C}$. If $p \in H_{3,3}$ is a cubic form vanishing on $\gamma^1, \dots, \gamma^8$, it must also vanish on γ^9 .*

5.2. A more general Version of the Cayley–Bacharach Theorem. For our purposes, we need a stronger version of the Cayley–Bacharach Theorem, which can also be found in [11]. Before we state it, we formalize and generalize the idea of Theorem 5.20 that given a set $\Gamma \subseteq \mathbb{P}^n\mathbb{C}$ of projective points, vanishing on a subset of Γ can already be sufficient for a form of certain degree k to vanish on the whole set Γ . The following definition is taken from [11].

Definition 5.21. Let $s \in \mathbb{N}$ and $\Gamma = \{\gamma_1, \dots, \gamma_s\} \subseteq \mathbb{P}^{n-1}\mathbb{C}$ be a finite set of projective points. Fix a degree k . Then the **failure** of Γ to impose independent conditions on forms of degree k is the quantity $s - r$, where $0 \leq r \leq s$ is the least number such that there are r distinct points $\gamma^{i_1}, \dots, \gamma^{i_r}$, $i_1, \dots, i_r \in \{1, \dots, s\}$ such that every form $p \in H_{n,k}$ that vanishes on $\gamma^{i_1}, \dots, \gamma^{i_r}$ already vanishes on the whole set Γ .

Example 5.22. Consider for example $n = 3$ and $k = 1$, i.e. the set of ternary linear forms and a set of $s = 3$ projective points $\Gamma = \{\gamma^1, \gamma^2, \gamma^3\}$ given by the homogeneous coordinates $z^1 = (1, 1, 0)^\top$, $z^2 = (1, 0, 1)^\top$, $z^3 = (2, 1, 1) \in \mathbb{C}^3$, respectively. Let $p = p_1\mathbf{x} + p_2\mathbf{y} + p_3\mathbf{z} \in H_{3,1}$ be arbitrary, where $p_1, p_2, p_3 \in \mathbb{R}$. Clearly, if p vanishes on two of the points, say γ^1, γ^2 it also vanishes on any linear combination $\alpha z^1 + \beta z^2$ ($\alpha, \beta \in \mathbb{R}$) of the homogeneous coordinates. In our case, we have $z^3 = z^1 + z^2$. Hence, if p vanishes on γ^1 and γ^2 , it must also vanish on γ^3 . This shows that the failure of Γ to impose independent conditions on forms of degree $k = 1$ is at least $1=3-2$. Clearly it can also not be bigger. For example, $\mathbf{x} - \mathbf{y} \in H_{3,1}$ vanishes only on γ^1 and neither on γ^2 , nor on γ^3 . A similar observation can be made for the other two points. Hence, the failure is exactly 1.

It will later be important to know an equivalent characterization of the failure in terms of linear independency of point evaluations. Further, there is also a nice characterization in terms of the Hilbert function. Before we can state this equivalence of characterizations, we need some more definitions, which can be found in [28, Lecture 7].

Definition 5.23. Let $I \triangleleft \mathbb{K}[\mathbf{x}]$ be an ideal. Then

$$I_d := I \cap \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\} = \{p \in I : \deg(p) = d\}$$

is the d -th homogeneous part of I .

Definition 5.24. Let $n \in \mathbb{N}$ and $V \subseteq \mathbb{P}^{n-1}\mathbb{K}$ be a projective variety. The **homogeneous coordinate ring** of V is the quotient ring $S(V) = \mathbb{K}[\mathbf{x}]/I(V)$. Further, let $I_d(V) := (I(V))_d$ be the d -th homogeneous part of the vanishing ideal on V . Then we define the quotient space $S_d(V) = \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\}/I_d(V)$.

Remark 5.25. We write $\{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\}$ and not $H_{n,d}$ since $H_{n,d}$ was only the set of real forms in n variables of degree d , whereas the forms in the first set might have complex coefficients.

Note that $S_d(V)$ is a subspace of $S(V)$ for all d and $S(V) = \bigoplus_{d \in \mathbb{N}} S_d(V)$. Further, we can define the following.

Definition 5.26. Let $V \subseteq \mathbb{P}^{n-1}\mathbb{K}$ a projective variety. The **Hilbert function** of V is defined as

$$h_V : \mathbb{N} \rightarrow \mathbb{N}, h_V(d) = \dim_{\mathbb{K}}(S_d(V)).$$

Now, there is the following equivalence, which is stated and proven in [28, Lemma 8.7].

Lemma 5.27. Let $\Gamma = \{\gamma^1, \dots, \gamma^s\} \subseteq \mathbb{P}^{n-1}\mathbb{K}$ be a set of $s \in \mathbb{N}$ projective points with homogeneous coordinates $z^1, \dots, z^s \in \mathbb{C}^n$. Then the following are equivalent:

- (i) The Hilbert function satisfies $h_{\Gamma}(d) = s$.
- (ii) The failure of Γ to impose independent conditions on forms of degree d is 0.
- (iii) The set of point evaluations $\{ev_{z^i} : i = 1, \dots, s\}$ is a linearly independent subset of the vector space of all linear functionals of the form

$$\{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\} \longrightarrow \mathbb{K}.$$

Proof. First, note that by Remark 5.3, Γ is indeed a projective variety. Further, (ii) is equivalent to

$$\begin{aligned} \forall i \in \{1, \dots, s\} \exists p \in \mathbb{K}[\mathbf{x}], \deg(p) = d : \\ [\forall j \in \{1, \dots, s\} \setminus \{i\} : p(z^j) = 0] \wedge p(z^i) \neq 0, \end{aligned} \quad (44)$$

i.e. for each point $\gamma^i \in \Gamma$, there is a form vanishing at all points of Γ except γ^i . We can now prove the equivalence. Therefore, consider the mapping

$$\varphi : \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\} \rightarrow \mathbb{K}^s, \varphi(p) = (p(z^1), \dots, p(z^s))^{\top}.$$

Clearly, we have $I_d(\Gamma) = \text{Ker}(\varphi)$ and hence for the dimension of the image of φ

$$\begin{aligned} \dim_{\mathbb{K}} \text{Im}(\varphi) &= \dim_{\mathbb{K}} \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\} - \dim \text{Ker}(\varphi) \\ &= \dim_{\mathbb{K}} \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\} - \dim I_d(\Gamma) \\ &= \dim_{\mathbb{K}} (\{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\} / I_d(\Gamma)) = h_{\Gamma}(d). \end{aligned}$$

For this reason, we know "(i) \implies (ii)", since (i) is equivalent to saying that φ is surjective. In particular, for all $j \in \{1, \dots, s\}$ there is a $p_j \in \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\}$ such $\varphi(p_j) = e^j$.

Further, this also shows that "(i) \implies (iii)". Therefore, let $\lambda_i \in \mathbb{C}$ be arbitrary such that $\sum_{i=1}^s \lambda_i \text{ev}_{z^i} = 0$. In particular, if we choose p_j as above, we have $0 = \sum_{i=1}^s \lambda_i \text{ev}_{z^i}(p_j) = \lambda_j$ for all $j \in \{1, \dots, s\}$, showing the linear independency of the point evaluations.

"(ii) \implies (i)": If (ii) holds, we know by (44) that for each $j \in \{1, \dots, s\}$, there are $q_j \in \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\}$ and $\lambda_j \in \mathbb{C} \setminus \{0\}$ such that $\varphi(q_j) = \lambda_j e^j$. By homogeneity, this means $\varphi(\lambda_j^{-1/d} q_j) = e^j$, which shows that $e^j \in \text{Im}(\varphi)$ for all j and hence the surjectivity of φ .

"(iii) \implies (i)": We prove this by contraposition. Assume that φ is not surjective. Hence, $\text{Im}(\varphi)$ is contained in a subspace of \mathbb{K}^s of codimension 1, i.e. a hyperplane H in \mathbb{K}^s . From Linear Algebra, we know that this hyperplane can be written as

$$H = \left\{ z \in \mathbb{K}^s : \sum_{i=1}^s \lambda_i z_i = 0 \right\}$$

for some $\lambda = (\lambda_1, \dots, \lambda_s) \in \mathbb{K}^s \setminus \{0\}$. Since $\text{Im}(\varphi) \subseteq H$, we know in particular

$$\forall p \in \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = d\} : \sum_{i=1}^s \lambda_i \underbrace{(\varphi(p))_i}_{=p(z^i)} = \sum_{i=1}^s \lambda_i \text{ev}_{z^i}(p) = 0,$$

which shows that $\{\text{ev}_{z^i} : i = 1, \dots, s\}$ is linearly dependent. Hence, we obtain the claim. \square

Now, consider the following version of the Cayley–Bacharach Theorem.

Theorem 5.28. *Let $X_1, \dots, X_n \subseteq \mathbb{P}^n \mathbb{K}$ be hypersurfaces of degree d_1, \dots, d_n respectively, meeting transversally in $\Gamma = X_1 \cap \dots \cap X_n \subseteq \mathbb{P}^n \mathbb{K}$. Let $\Gamma = \Gamma' \dot{\cup} \Gamma''$ be the distinct union of two sets $\Gamma', \Gamma'' \subseteq \mathbb{P}^n \mathbb{K}$. Set $r = (\sum_{i=1}^n d_i) - n - 1$. If $0 \leq k \leq r$ is a nonnegative integer, then the dimension of the family $\mathcal{F} \subseteq \{p \in \mathbb{K}[\mathbf{x}] : \deg(p) = k\}$ of forms of degree k in $n + 1$ variables vanishing on Γ' is equal to the failure of Γ'' to impose independent conditions on forms of 'complementary' degree $r - k$.*

Proof. [11, Theorem CB6]. \square

For our purposes, we need a special case of Theorem 5.20, which is stated without proof in [4, Lemma 2.9]. This will be the key tool in characterizing the SOS cone $\Sigma_{n,2d}$ inside the PSD cone $P_{n,2d}$ for $(n, 2d) \in \{(4, 4), (3, 6)\}$ in Section 5.3.

Lemma 5.29. *Let $(n, d) \in \{(4, 2), (3, 3)\}$ and $p_1, \dots, p_{n-1} \in H_{n,d}$ be forms intersecting transversally in $s = d^{n-1}$ projective points $\gamma^1, \dots, \gamma^s \in \mathbb{P}^{n-1}\mathbb{C}$. Further, let $z^i \in \mathbb{C}^n \setminus \{0\}$ be homogeneous coordinates of the γ^i .*

Then there is a up to scaling unique linear relation on the values of any form $p \in H_{n,d}$ on z^i , i.e. there are $a_i \in \mathbb{C} \setminus \{0\}$ such that

$$\forall p \in H_{n,d} : \sum_{i=1}^s a_i p(z^i) = 0. \quad (45)$$

Proof. In this proof, we want to apply Theorem 5.28. For this reason, let $X_i = \{(z) \in \mathbb{P}^{n-1}\mathbb{C} : p_i(z) = 0\}$ be the hypersurface of degree d defined by p_i ($i = 1, \dots, n-1$). In the sense of Theorem 5.28 define $\Gamma = \{\gamma^1, \dots, \gamma^s\} = \Gamma' \dot{\cup} \Gamma''$, where $\Gamma' = \emptyset$ and $\Gamma'' = \Gamma$. Further, we have $r = \left(\sum_{i=1}^{n-1} d\right) - (n-1) - 1 = (n-1)(d-1) - 1$.

Now let $k = 0$. Clearly, the dimension of the family $\mathcal{F} \subseteq H_{n,k}$ of forms of degree $k = 0$ vanishing on $\Gamma' = \emptyset$ is 1, since it is the set of all constants. Further, note that since $(n, d) \in \{(4, 2), (3, 3)\}$, we have

$$r = (n-1)(d-1) - 1 = \begin{cases} 3 \cdot 1 - 1 = 2 = d, & \text{if } (n, d) = (4, 2) \\ 2 \cdot 2 - 1 = 3 = d, & \text{if } (n, d) = (3, 3) \end{cases}.$$

This shows $r - k = r - 0 = r = d$ and Theorem 5.20 tells us that the failure of $\Gamma'' = \Gamma$ to impose independent conditions on forms of (complementary) degree d is 1 as well. Hence, Lemma 5.27 tells us that the family of point evaluations $\{\text{ev}_{z^i} : i \in \{1, \dots, s\}\}$ is linearly dependent, i.e. there are $a_i \in \mathbb{C}$ not all $a_i = 0$ such that

$$\forall p \in H_{n,d} : \sum_{i=1}^s a_i p(z^i) = 0. \quad (46)$$

This is almost the claim. It remains to show that the a_i can also be chosen such that they are all nonzero. To do so, apply again Theorem 5.20 but this time on $\Gamma' = \{\gamma^j\}, \Gamma'' = \Gamma \setminus \{\gamma^j\}$, where $j \in \{1, \dots, s\}$ is arbitrary. Clearly, the dimension of the family of forms of degree $k = 0$ vanishing on $\Gamma' = \{\gamma^j\}$ is 0. Hence, the failure of Γ'' to impose independent conditions on forms $p \in H_{n,d}$ is 0 as well. This time, Lemma 5.27 shows that the set $\{\text{ev}_{z^i} : i \in \{1, \dots, s\} \setminus \{j\}\}$ is linearly independent. In particular, we must have $a_j \neq 0$, since otherwise $\sum_{i=1, i \neq j}^s a_i \text{ev}_{z^i} = 0$ would be a nontrivial linear combination. Since $j \in \{1, \dots, s\}$ was arbitrary, we obtain $a_j \in \mathbb{C} \setminus \{0\}$ for all j .

The linear combination is also unique up to scaling since otherwise, we could find a linear combination $\sum_{i=1}^s \tilde{a}_i \text{ev}_{z^i} = 0$ ($\tilde{a}_i \in \mathbb{C}$) where some $\tilde{a}_j = 0$. This shows the claim. \square

Convention 5.30. From now on, we call the coefficients $a_i \in \mathbb{C} \setminus \{0\}$ of (45) the **Cayley–Bacharach coefficients** and property (45) itself the **Cayley–Bacharach relation**.

5.3. Structure of Extreme Rays in $\Sigma_{4,4}^*$ and $\Sigma_{3,6}^*$. This section follows [4, Section 4]. Therefore, results from [4, Section 2-3] are used without proof. However, at the beginning of this section, for our purposes crucial results from [4, Section 2-3] are recalled. On this basis, we gain some knowledge about the structure of extreme rays of the closed convex cones $\Sigma_{4,4}^*$ and $\Sigma_{3,6}^*$. This is the key in understanding what separates the SOS cone $\Sigma_{n,2d}$ from the non SOS forms inside $P_{n,2d}$ ($(n, d) \in \{(4, 2), (3, 3)\}$).

Convention 5.31. In this section, we always consider $(n, d) \in \{(4, 2), (3, 3)\}$, i.e. $H_{n,2d}$ is either the set of quaternary quartics or the set of ternary sextics.

For a linear functional $\ell \in H_{n,2d}^\vee$, we can define a quadratic form Q_ℓ on $H_{n,d}$ as follows.

Definition 5.32. Let $\ell \in H_{n,2d}^\vee$. We define a **quadratic form** Q_ℓ on $H_{n,d}$ via

$$Q_\ell: H_{n,d} \rightarrow \mathbb{R}, Q_\ell(q) = \ell(q^2).$$

Further, the **kernel** of Q_ℓ is denoted by $W_\ell := \text{Ker } Q_\ell$. The quadratic form is said to be **positive semidefinite** if $Q_\ell(q) \geq 0$ for all $q \in H_{n,d}$.

Now, the setting we are working on in this section is specified:

Let $\ell \in \Sigma_{n,2d}^*$ span an extreme ray of the closed convex cone $\Sigma_{n,2d}^*$ such that $\ell \neq \text{ev}_z$ for all $z \in \mathbb{R}^n$, i.e. ℓ does not correspond to a point evaluation. By [4, Section 2], there are $p_1, \dots, p_n \in W_\ell \subseteq H_{n,d}$ that do not have common zeros real or complex, i.e.

$$V_{\mathbb{C}}^h(p_1, \dots, p_n) = \emptyset.$$

Further, by [4, Theorem 2.8] there are $q_1, \dots, q_{n-1} \in \text{span}_{\mathbb{R}}\{p_1, \dots, p_n\} \subseteq W_\ell$ that intersect transversally in $s = d^{n-1}$ possibly complex projective points, i.e.

$$V_{\mathbb{C}}^h(q_1, \dots, q_{n-1}) = \Gamma = \{\gamma^1, \dots, \gamma^s\} \subseteq \mathbb{P}^{n-1}\mathbb{C}.$$

Further, the q_i can be assumed to be linearly independent. Let $z^i \in \mathbb{C}^n \setminus \{0\}$ be homogeneous coordinates of the γ^i ($i = 1, \dots, s$). By the Cayley–Bacharach relation from Lemma 5.29, there are $a_i \in \mathbb{C} \setminus \{0\}$ that are up to scaling uniquely determined such that

$$\forall q \in H_{n,d}: \sum_{i=1}^s a_i p(z^i) = \sum_{i=1}^s a_i \text{ev}_{z^i}(p) = 0.$$

Definition 5.33. For a set $S \subseteq \mathbb{K}[\mathbf{x}]$, we write $\langle S \rangle \subseteq \mathbb{K}[\mathbf{x}]$ for the **ideal** generated by S (see [7, Chapter 2, §4]). For $k \in \mathbb{N}$, we define $\langle S \rangle_k = \{p \in \langle S \rangle : \deg(p) = k\}$ as the **k -th homogeneous part** of $\langle S \rangle$.

By [4, Theorem 2.7], we know that $\dim_{\mathbb{R}} W_\ell = n$. Further, the n -dimensional space W_ℓ defines the linear functional ℓ uniquely as linear functional vanishing on $\langle W_\ell \rangle_{2d}$. In particular, since $q_1, \dots, q_{n-1} \in W_\ell$ are in the kernel of ℓ , the functional ℓ has to vanish on $\langle q_1, \dots, q_{n-1} \rangle_{2d} \subseteq \langle W_\ell \rangle_{2d}$.

Further, it is stated in [4, Section 2] that since the q_i intersect transversally, ℓ can be written in the form

$$\ell = \sum_{i=1}^s \alpha_i \text{ev}_{z^i} \quad (47)$$

for some $\alpha_i \in \mathbb{C}$.

Goal: The first aim of this section is to understand, where the coefficients α_i in (47) come from and what properties they have. Second, we want to choose the homogeneous coordinates z^i in such a way that as many as possible of them are real. Later it can be seen that the SOS cone $\Sigma_{n,2d}$ can be separated from the non SOS forms in $P_{n,2d}$ precisely by extreme rays of $\Sigma_{n,2d}^*$ of the form (47).

Therefore, start by choosing a form $q_n \in W_\ell \subseteq H_{n,d}$ such that $\{q_1, \dots, q_n\}$ forms a basis of W_ℓ . The following observation is stated without proof in [4, Section 4].

Proposition 5.34. *The values $q_n(z^i)$ ($i = 1, \dots, s$) are up to scaling independent of the choice of $q_n \in W_\ell$. Further, it holds $q_n(z^i) \neq 0$ for all $i = 1, \dots, s$.*

Proof. Let $\tilde{q}_n \in W_\ell$ be a second form such that $\{q_1, \dots, q_{n-1}, \tilde{q}_n\}$ is also a basis of W_ℓ . Then there must be $\lambda_j \in \mathbb{R}$ ($j = 1, \dots, n$), $\lambda_n \neq 0$ such that $\tilde{q}_n = \sum_{j=1}^n \lambda_j q_j$. Since $\Gamma = V^h(q_1, \dots, q_{n-1})$ and the z^i are homogeneous coordinates of the elements in Γ , we know

$$\tilde{q}_n(z^i) = \sum_{j=1}^n \lambda_j \underbrace{q_j(z^i)}_{=0 \text{ for } j \neq n} = q_n(z^i) \quad (i = 1, \dots, s), \quad (48)$$

which shows the first part of the claim. For the second part, assume that $q_n(z^i) = 0$ for some $i \in \{1, \dots, s\}$. Recall that $V^h(p_1, \dots, p_n) = \emptyset$ and $q_j \in \text{span}_{\mathbb{R}}\{p_1, \dots, p_n\}$ for $j = 1, \dots, n-1$. Hence, there must be some $j \in \{1, \dots, n-1\}$ such that $p_j(z^i) \neq 0$. Now take $\tilde{q}_n = q_n + p_j$. But then it holds $\lambda \underbrace{q_n(z^i)}_{=0} = 0 \neq p_j(z^i) = q_n(z^i) + p_j(z^i) = \tilde{q}_n(z^i)$ for all $\lambda \in \mathbb{R} \setminus \{0\}$, which contradicts (48). This shows the claim. \square

The following lemma describes how the coefficients $\alpha_i \in \mathbb{C}$ from (47) arise from the coefficients a_i of the Cayley–Bacharach relation (45) and the values $q_n(z^i)$. The lemma is stated and proven in [4, Lemma 4.1].

Lemma 5.35. *Up to scaling, it holds $\alpha_i = \frac{a_i}{q_n(z^i)}$ for $i = 1, \dots, s$. Hence, the extreme ray ℓ is given by $\ell = \sum_{i=1}^s \frac{a_i}{q_n(z^i)} \text{ev}_{z^i}$.*

Proof. Since ℓ is defined as a linear combination of point evaluations at zeros of q_1, \dots, q_{n-1} , it follows $\ell(pq_j) = \ell(p) \cdot \ell(q_j) = 0$ for $j = 1, \dots, n-1$ and all $p \in H_{n,d}$. Further, $\alpha_i \in \mathbb{C}$ defined as in the claim yields

$$\ell(pq_n) = \sum_{i=1}^s \alpha_i \text{ev}_{z^i}(pq_n) = \sum_{i=1}^s \frac{a_i}{q_n(z^i)} p(z^i) q_n(z^i) = \sum_{i=1}^s a_i p(z^i) = 0$$

by the Cayley–Bacharach relation (45). Further, $\{q_1, \dots, q_n\}$ forms a basis of W_ℓ . Hence, we obtain the claim as ℓ is the up to scaling unique linear functional which vanishes on $\langle W_\ell \rangle_{2d} = \langle q_1, \dots, q_n \rangle_{2d}$. \square

Next, we are going to investigate the number of complex points in Γ . Since the polynomials q_1, \dots, q_{n-1} are real and intersect in $\Gamma \subseteq \mathbb{P}^{n-1}\mathbb{C}$, Γ must be closed under complex conjugation, i.e. if $(z) = (z_1 : \dots : z_n) \in \Gamma$, $z = (z_1, \dots, z_n)^\top \in \mathbb{C}^n$, so $(\bar{z}) \in \Gamma$. In particular, the homogeneous coordinates of γ^i ($i = 1, \dots, s$) can be chosen such that they are closed under complex conjugation. Hence, we can do the following.

Convention 5.36. Denote by r_i the elements of z^i that can be chosen to be real homogeneous coordinates. For simplicity, we denote the purely complex z^i again by z^i . Hence, the homogeneous coordinates of the elements in $\Gamma = V_{\mathbb{C}}^h(q_1, \dots, q_{n-1})$ can without loss of generality be chosen as

$$r^1, \dots, r^k \in \mathbb{R}^n, \quad z^1, \dots, z^m, \bar{z}^1, \dots, \bar{z}^m \in \mathbb{C}^n \quad \text{s.t.} \quad z^j \neq \bar{z}^j \quad (j = 1, \dots, m).$$

In particular, this means $s = k + 2m$.

Goal: We now determine the number m of complex conjugated pairs (z^i, \bar{z}^i) ($i = 1, \dots, m$).

Convention 5.37. By (47) we know that with the z^i as in Convention 5.36, ℓ is of the form

$$\ell = \sum_{i=1}^k \beta_i \text{ev}_{r^i} + \sum_{i=1}^m (\alpha_i \text{ev}_{z^i} + \gamma_i \text{ev}_{\bar{z}^i}),$$

where $\beta_i, \alpha_j, \gamma_j \in \mathbb{C} \setminus \{0\}$ for all $i = 1, \dots, k, j = 1, \dots, m$. Since $\ell(p) \in \mathbb{R}$ for all $p \in H_{n,2d}$, it can without loss of generality be assumed that $\beta_i \in \mathbb{R}$ ($i = 1, \dots, k$) and $\gamma_j = \bar{\alpha}_j$ ($j = 1, \dots, m$). Further, $\beta_i, \alpha_j \neq 0$ for all i, j . More precisely ℓ can be written as

$$\ell = \sum_{i=1}^k \beta_i \text{ev}_{r^i} + \sum_{i=1}^m (\alpha_i \text{ev}_{z^i} + \bar{\alpha}_i \text{ev}_{\bar{z}^i}). \quad (49)$$

The following lemma is stated and proven in [4, Lemma 4.2].

Lemma 5.38. *Suppose that the quadratic form Q_ℓ is positive semidefinite. Define the evaluation map*

$$E_{\mathbb{R}}: H_{n,d} \rightarrow \mathbb{R}^s = \mathbb{R}^{k+2m},$$

$$E_{\mathbb{R}}(p) = \left(p(r^1), \dots, p(r^k), \text{Re } p(z^1), \text{Im } p(z^1), \dots, \text{Re } p(z^m), \text{Im } p(z^m) \right)$$

and let $c := \dim_{\mathbb{R}} E_{\mathbb{R}}(H_{n,d}) = \text{rk}(E_{\mathbb{R}})$, where $\text{rk}(E_{\mathbb{R}})$ is the rank of the linear map $E_{\mathbb{R}}$. Then it holds $c \leq k + m$.

Proof. For $q \in H_{n,d}$, it holds

$$Q_\ell(q) = \ell(q^2) = \sum_{i=1}^k \beta_i q^2(r^i) + \sum_{i=1}^m (\alpha_i q^2(z^i) + \bar{\alpha}_i q^2(\bar{z}^i)).$$

Further, let $\tilde{Q}_\ell: \mathbb{R}^s \rightarrow \mathbb{R}$ be the quadratic form on $\mathbb{R}^s = \mathbb{R}^{k+2m}$ given by

$$\tilde{Q}_\ell(x) = \sum_{i=1}^k \beta_i x_i^2 + \sum_{i=1}^m \left(\alpha_i (x_{2i-1} + \sqrt{-1} x_{2i})^2 + \bar{\alpha}_i (x_{2i-1} - \sqrt{-1} x_{2i})^2 \right).$$

for $x = (x_1, \dots, x_s)^\top \in \mathbb{R}^s$. Clearly, it holds $Q_\ell = \tilde{Q}_\ell \circ E_\mathbb{R}$. Note that $E_\mathbb{R}$ is welldefined since for all $i = 1, \dots, m$ and $x = (x_1, \dots, x_s)^\top \in \mathbb{R}^s$, we have

$$\begin{aligned} & \alpha_i (x_{2i-1} + \sqrt{-1} x_{2i})^2 + \bar{\alpha}_i (x_{2i-1} - \sqrt{-1} x_{2i})^2 \\ = & \operatorname{Re}(\alpha_i) \cdot (x_{2i-1}^2 + 2\sqrt{-1} x_{2i-1}x_{2i} - x_{2i}^2 + x_{2i-1}^2 - 2\sqrt{-1} x_{2i-1}x_{2i} - x_{2i}^2) \\ & + \operatorname{Im}(\alpha_i) \cdot (x_{2i-1}^2 + 2\sqrt{-1} x_{2i-1}x_{2i} - x_{2i}^2 - (x_{2i-1}^2 - 2\sqrt{-1} x_{2i-1}x_{2i} - x_{2i}^2)) \\ = & 2\operatorname{Re}(\alpha_i)x_{2i-1}^2 + 4\operatorname{Im}(\alpha_i)x_{2i-1}x_{2i} - 2\operatorname{Re}(\alpha_i)x_{2i}^2 \\ = & \begin{pmatrix} x_{2i-1} & x_{2i} \end{pmatrix} \cdot \begin{pmatrix} 2\operatorname{Re}(\alpha_i) & -2\operatorname{Im}(\alpha_i) \\ -2\operatorname{Im}(\alpha_i) & -2\operatorname{Re}(\alpha_i) \end{pmatrix} \cdot \begin{pmatrix} x_{2i-1} \\ x_{2i} \end{pmatrix} \in \mathbb{R}. \end{aligned}$$

Further, the 2×2 matrix $T_i := \begin{pmatrix} 2\operatorname{Re}(\alpha_i) & -2\operatorname{Im}(\alpha_i) \\ -2\operatorname{Im}(\alpha_i) & -2\operatorname{Re}(\alpha_i) \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ has the characteristic polynomial

$$\begin{aligned} & (2\operatorname{Re}(\alpha_i) - \lambda)(-2\operatorname{Re}(\alpha_i) - \lambda) - 4\operatorname{Im}(\alpha_i)^2 \\ = & \lambda^2 - 4\operatorname{Re}(\alpha_i)^2 - 4\operatorname{Im}(\alpha_i)^2 = \lambda^2 - 4|\alpha_i|^2 \end{aligned}$$

Since $\alpha_i \neq 0$, the matrix T_i has exactly one positive and one negative eigenvalue. For this reason, it holds

$$\tilde{Q}_\ell(x) = x^\top A_{\tilde{Q}_\ell} x \quad (x \in \mathbb{R}^s),$$

where

$$A_{\tilde{Q}_\ell} = \begin{pmatrix} \operatorname{diag}(\beta_1, \dots, \beta_k) & 0_{k \times 2m} \\ 0_{2m \times k} & \operatorname{diag}(T_1, \dots, T_m) \in \mathbb{R}^{s \times s} \end{pmatrix}$$

has at least m negative eigenvalues. Equivalently, \tilde{Q}_ℓ is strictly negative on a subspace $U \subseteq \mathbb{R}^s$ with $\dim_{\mathbb{R}}(U) \geq m$. Since Q_ℓ on the other hand is positive semidefinite and $Q_\ell = \tilde{Q}_\ell \circ E_\mathbb{R}$, it can be deduced that $c = \dim_{\mathbb{R}} E_\mathbb{R}(H_{n,d}) \leq \operatorname{codim}_{\mathbb{R}}(U) \leq k + m$, which shows the claim. \square

The following proposition is stated and proven independent of [4].

Proposition 5.39. *Let Q_ℓ be positive semidefinite and $E_\mathbb{R}$ be defined as in Lemma 5.38. Then $c = \dim_{\mathbb{R}} E_\mathbb{R}(H_{n,d})$ is the maximum number of points $s^1, \dots, s^c \in S := \{r^1, \dots, r^k, z^1, \dots, z_m, \bar{z}_1, \dots, \bar{z}_m\}$ such that the point evaluations $ev_{s^i}: H_{n,d} \rightarrow \mathbb{C}$ are \mathbb{C} -linearly independent.*

Proof. Let $N = N(n, d) = \binom{n-1+d}{n-1}$ (see Remark 2.4) be the dimension of the \mathbb{R} -vector space $H_{n,d}$ and $\mathcal{F} = \{f_1, \dots, f_N\} \subseteq H_{n,d}$ be a basis of $H_{n,d}$ (e.g. the basis of monomials). Consider

$$\mathcal{E} := \begin{pmatrix} E_\mathbb{R}(f_1) \\ \vdots \\ E_\mathbb{R}(f_N) \end{pmatrix} \in \mathbb{R}^{N \times s} = \mathbb{R}^{N \times (k+2m)}.$$

By definition of c , we know that $c = \text{rk}(\mathcal{E})$. Note that $\text{rk}(\mathcal{E})$ is the same if we interpret \mathcal{E} as a complex $N \times s$ matrix. For $z \in \mathbb{C}^n$, define

$$\text{ev}_z(\mathcal{F}) := \begin{pmatrix} \text{ev}_z(f_1) \\ \vdots \\ \text{ev}_z(f_N) \end{pmatrix} \in \mathbb{R}^N.$$

Hence, using the definition of $E_{\mathbb{R}}$, \mathcal{E} can be rewritten as

$$\mathcal{E} = (\text{ev}_{r^1}(\mathcal{F}), \dots, \text{ev}_{r^k}(\mathcal{F}), \text{Re}(\text{ev}_{z^1}(\mathcal{F})), \text{Im}(\text{ev}_{z^1}(\mathcal{F})), \dots, \text{Re}(\text{ev}_{z^m}(\mathcal{F})), \text{Im}(\text{ev}_{z^m}(\mathcal{F}))).$$

From Linear Algebra, we know that $c = \text{rk}(\mathcal{E})$ is also the dimension of the column space of \mathcal{E} as \mathbb{R} - or \mathbb{C} -vector space, since \mathcal{E} is a real matrix. Hence, it is in particular the dimension of the set of all \mathbb{C} -linear combinations of the functionals

$$\text{ev}_{r^1}, \dots, \text{ev}_{r^k}, \text{Re}(\text{ev}_{z^1}), \text{Im}(\text{ev}_{z^1}), \dots, \text{Re}(\text{ev}_{z^m}), \text{Im}(\text{ev}_{z^m}).$$

as a subspace of $\{\varphi: H_{n,d} \rightarrow \mathbb{C} \mid \varphi \text{ linear}\}$. Since $\text{ev}_{z^j} = \text{Re}(\text{ev}_{z^j}) + i \text{Im}(\text{ev}_{z^j})$ ($j = 1, \dots, m$), this is the \mathbb{C} -vector space

$$\text{span}_{\mathbb{C}} \{\text{ev}_{r^1}, \dots, \text{ev}_{r^k}, \text{ev}_{z^1}, \text{ev}_{\bar{z}^1}, \dots, \text{ev}_{z^m}, \text{ev}_{\bar{z}^m}\}$$

Hence, the claim follows. \square

As a consequence, we have the following corollary as in [4, Corollary 4.4], which is there stated without proof.

Corollary 5.40. *Suppose that $\ell \in H_{n,2d}^{\vee}$ is an extreme ray of $\Sigma_{n,2d}^*$ that does not correspond to a point evaluation and let q_1, \dots, q_{n-1} be forms in the kernel W_{ℓ} of Q_{ℓ} that intersect transversally in $s = d^{n-1}$ points $\Gamma = \{\gamma^1, \dots, \gamma^s\}$. Let $r_i \in \mathbb{R} \setminus \{0\}$, $z_j \in \mathbb{C} \setminus \{0\}$, $z_j \neq \bar{z}_j$ ($i = 1, \dots, k$, $j = 1, \dots, m$) be homogeneous coordinates of the elements of Γ as in Convention 5.36. Then the set Γ includes at most 1 complex conjugated pair, i.e. $m \leq 1$.*

Proof. Note that since ℓ is nonnegative on $\Sigma_{n,2d}$, it follows instantly that the quadratic form Q_{ℓ} is positive semidefinite as $Q_{\ell}(p) = \ell(p^2)$ for all $p \in H_{n,d}$. By the Cayley–Bacharach relation of Lemma 5.29, we know that

$$\sum_{i=1}^k \beta_i \text{ev}_{r^i} + \sum_{i=1}^n (\alpha_i \text{ev}_{z^i} + \bar{\alpha}_i \text{ev}_{\bar{z}^i}) = 0 \text{ on } H_{n,d},$$

where the coefficients $\beta_i \in \mathbb{R} \setminus \{0\}$, $\alpha_j \in \mathbb{C} \setminus \{0\}$ ($i = 1, \dots, k$, $j = 1, \dots, m$) are uniquely determined up to scaling. For this reason, the maximum number of \mathbb{C} -linearly independent functions $\text{ev}_{r^i}, \text{ev}_{z^j}, \text{ev}_{\bar{z}^j}: H_{n,d} \rightarrow \mathbb{C}$ must be $s - 1 = k + 2m - 1$. Further, by Proposition 5.39 we know that this is $c = \dim_{\mathbb{R}} E_{\mathbb{R}}(H_{n,d})$, where $E_{\mathbb{R}}$ is defined as in Lemma 5.38, i.e. $c = s - 1 = k + 2m - 1$. In addition, Lemma 5.38 tells that $c \leq k + m$. This shows $m \leq 1$. \square

This allows to prove the two main theorems [4, Theorem 1.1, Theorem 1.2] of Blekherman, which are stated in the following.

Theorem 5.41. *Suppose that $p \in P_{3,6}$ is not SOS, i.e. $p \notin \Sigma_{3,6}$. Then there exist two real cubics $q_1, q_2 \in H_{3,3}$ intersecting in $9 = 3^2$ (possibly complex) projective points $\gamma^1, \dots, \gamma^9 \in \mathbb{P}^2\mathbb{C}$ such that the values of p on γ^i certify that p is not SOS. More precisely, let $z^1, \dots, z^9 \in \mathbb{C}^3 \setminus \{0\}$ be homogenous coordinates of the γ^i such that they are closed under complex conjugation. Then there exists a real linear functional $\ell: H_{3,6} \rightarrow \mathbb{R}$ given by*

$$\ell(f) = \sum_{i=1}^9 \alpha_i f(z^i) \quad (f \in H_{3,6})$$

for some $\alpha_i \in \mathbb{C} \setminus \{0\}$ such that $\ell(\Sigma_{3,6}) \geq 0$ but $\ell(p) < 0$. Furthermore, there is at most one complex conjugated pair (z_i, \bar{z}_i) in the homogeneous coordinates.

Theorem 5.42. *Suppose that $p \in P_{4,4}$ is not SOS, i.e. $p \notin \Sigma_{4,4}$. Then there exist three real quadrics $q_1, q_2, q_3 \in H_{4,2}$ intersecting in $8 = 2^3$ (possibly complex) projective points $\gamma^1, \dots, \gamma^8 \in \mathbb{P}^3\mathbb{C}$ such that the values of p on γ^i certify that p is not SOS. More precisely, let $z^1, \dots, z^8 \in \mathbb{C}^4 \setminus \{0\}$ be homogenous coordinates of the γ^i such that they are closed under complex conjugation. Then there exists a real linear functional $\ell: H_{4,4} \rightarrow \mathbb{R}$ given by*

$$\ell(f) = \sum_{i=1}^8 \alpha_i f(z^i) \quad (f \in H_{4,4})$$

for some $\alpha_i \in \mathbb{C} \setminus \{0\}$ such that $\ell(\Sigma_{4,4}) \geq 0$ but $\ell(p) < 0$. Furthermore, there is at most one complex conjugated pair (z_i, \bar{z}_i) in the homogeneous coordinates.

The following proof can be found in [4, Section 5].

Proof of Theorem 5.41 and 5.42. As remarked in Convention 5.31, we have $(n, d) \in \{(4, 2), (3, 3)\}$. Let $p \in P_{n,2d} \setminus \Sigma_{n,2d}$ be arbitrary. By Proposition 3.80 we know that $\Sigma_{n,2d}^*$ is a closed convex cone. Hence, by Theorem 3.75 we know that $\Sigma_{n,2d}^*$ is the convex hull of its extreme directions. Since $\Sigma_{n,2d}^{**} = \Sigma_{n,2d}$ by Proposition 3.82 and Remark 3.68, there must be an extreme direction $\ell \in \Sigma_{n,2d}^*$ such that $\ell(p) < 0$. Since $\ell(p) < 0$ but p is nonnegative, $\ell \in H_{n,2d}^\vee$ spans an extreme ray of $\Sigma_{n,2d}^*$ that does not correspond to point evaluation. By Lemma 5.35, ℓ is of the desired form and by Corollary 5.40 we know that there is at most one complex conjugated pair in the z^i . This shows the claim. \square

We close this subsection by giving the following conjecture, which can be found in [4, Conjecture 7.3].

Conjecture 5.43. Let $(n, d) \in \{(4, 2), (3, 3)\}$, $W \subseteq H_{n,d}$ an n -dimensional subspace such that $V_{\mathbb{C}}^h(W) = \emptyset$. Further, assume that any collection of forms $q_1, \dots, q_{n-1} \in W$ intersecting transversally in s points $\gamma^1, \dots, \gamma^s \in \mathbb{P}^{n-1}\mathbb{C}$ has at most one pair complex conjugated projective zeros. Then there exist $q_1, \dots, q_{n-1} \in W$ intersecting transversally in only real points.

Remark 5.44. In our setting, Conjecture 5.43 means that an extreme direction $\ell \in \Sigma_{n,2d}^*$ as in the Theorems 5.41 and 5.42, which does not correspond

to a point evaluation, can always be written in the form $\ell = \sum_{i=1}^s \alpha_i \text{ev}_{z^i}$, where $s = d^{n-1}$ and $\alpha_i \in \mathbb{R} \setminus \{0\}$, $z^i \in \mathbb{R}^n \setminus \{0\}$ are all real.

5.4. Sum of Squares Certificates for elements of $P_{4,4}$ and $P_{3,6}$. From the Theorems 5.41 and 5.42 and Convention 5.37, we know that the SOS cone $\Sigma_{n,2d} ((n, d) \in \{(4, 2), (3, 3)\})$ can be separated from the non SOS forms inside the cone $P_{n,2d}$ by extreme rays that are either of the form

$$\ell = \sum_{i=1}^s \alpha_i \text{ev}_{r^i} \quad (50)$$

with real $r^i \in \mathbb{R}^n \setminus \{0\}$ and $\alpha_i \in \mathbb{R} \setminus \{0\}$ or of the form

$$\ell = \alpha \text{ev}_z + \bar{\alpha} \text{ev}_{\bar{z}} + \sum_{i=3}^s \alpha_i \text{ev}_{r^i}, \quad (51)$$

where $s = d^{n-1}$, $r^i \in \mathbb{R}^n \setminus \{0\}$, $z \in \mathbb{C}^n \setminus \{0\}$, $z \neq \bar{z}$ and $\alpha_i \in \mathbb{R} \setminus \{0\}$, $\alpha \in \mathbb{C} \setminus \{0\}$. Further, by Lemma 5.35 we know that the coefficient α_i and α depend from the coefficients $a_i \in \mathbb{R} \setminus \{0\}$ and $a \in \mathbb{C} \setminus \{0\}$ of the Cayley–Bacharach relation (45). This relation can be written as

$$a_1 \text{ev}_{r^1} + \dots + a_s \text{ev}_{r^s} = 0 \quad \text{on } H_{n,d}$$

in the case of (50) or

$$\alpha \text{ev}_z + \bar{\alpha} \text{ev}_{\bar{z}} + a_3 \text{ev}_{r^3} + \dots + a_s \text{ev}_{r^s} = 0 \quad \text{on } H_{n,d}$$

in the case of (51), respectively.

Convention 5.45. After scaling the r^i and z if necessary and using homogeneity of the forms in $H_{n,d}$, we can without loss of generality assume that the Cayley–Bacharach coefficients $a_i \in \mathbb{R} \setminus \{0\}$ and $a \in \mathbb{C} \setminus \{0\}$ have norm 1. In particular, it holds $\alpha_i \in \{-1, 1\}$ if α_i is real. Further, we can even assume that $a = \bar{a} = 1$ since those are the coefficients of complex point evaluations.

The goal is now to further specify the choice of coefficients α_i of (50) and (51). Therefore, consider the following result without proof, which can be deduced from [4, Theorem 6.1, Theorem 7.1].

Theorem 5.46. *Let $(n, d) \in \{(4, 2), (3, 3)\}$ and $\ell \in \Sigma_{n,2d}^*$ span an extreme ray that does not correspond to a point evaluation, $Q_\ell: H_{n,d} \rightarrow \mathbb{R}$ its quadratic form with kernel $W_\ell \subseteq H_{n,d}$ of dimension n . Further, let $q_1, \dots, q_{n-1} \in H_{n,d}$ be forms intersecting transversally in $\Gamma = \{\gamma^1, \dots, \gamma^s\} = V^h(q_1, \dots, q_{n-1}) \subseteq \mathbb{P}^{n-1}\mathbb{C}$, where $s = d^{n-1}$. Let z^i be homogeneous coordinates of the γ_i that are closed under complex conjugation, $a_i \in \mathbb{C} \setminus \{0\}$ the Cayley–Bacharach coefficients as in (45) such that $a_1 \text{ev}_{z^1} + \dots + a_s \text{ev}_{z^s} = 0$ on $H_{n,d}$. Then one of the following two cases must occur:*

Case 1: *All z^i are real, i.e. $z^i \in \mathbb{R}^n$ and ℓ is of the form (50) where the coefficients $\alpha_i \in \mathbb{R} \setminus \{0\}$ satisfy*

$$\sum_{i=1}^s \frac{1}{\alpha_i} = 0$$

with a single negative α_i , without loss of generality $\alpha_1 < 0$ and $\alpha_i > 0$ for $i > 1$.

Case 2: The z^i contain exactly one complex conjugated pair, without loss of generality $\bar{z}^1 = z^2$ and ℓ is of the form (51) where the coefficients α_i and α satisfy $\alpha_i > 0$ for $i = 3, \dots, s$ and

$$\frac{1}{\alpha} + \frac{1}{\bar{\alpha}} + \sum_{i=3}^s \frac{1}{\alpha_i} = 0.$$

Moreover, any linear functional $\ell \in H_{n,2d}^\vee$ of the form (50) or (51) with coefficients as in Case 1 or Case 2 is an extreme direction of $\Sigma_{n,2d}^*$.

In [12, Section 4], the author develops a certificate on a PSD form $p \in P_{n,2d}$ to be SOS ($(n, d) \in \{(4, 2), (3, 3)\}$). If $p \in P_{n,2d}$ was not SOS, the Theorems 5.41 and 5.42 tell us that there would be an $\ell \in \Sigma_{n,2d}^*$ that spans an extreme ray such that $\ell(\Sigma_{n,2d}) \geq 0$ but $\ell(p) < 0$. Further, Theorem 5.46 tells us precisely how such an extreme ray has to look like. For this reason, in order to deduce that p is SOS, it suffices to show $\ell(p) \geq 0$ for all extreme rays as in Theorem 5.46.

Remark 5.47. Consider for example the first case of extreme rays as in Case 1 of Theorem 5.46. We would have to show that

$$\sum_{i=1}^s \alpha_i p(z^i) \geq 0$$

for all $\alpha_i \in \mathbb{R}, z^i \in \mathbb{R}^n$ such that $\alpha_1 < 0, \alpha_i > 0$ ($i = 1, \dots, s$), $s = d^{n-1}$ and $\sum_{i=1}^s \frac{1}{\alpha_i} = 0$. This is equivalent to

$$\min \left\{ \sum_{i=1}^s \alpha_i p(z^i) : \alpha_1 < 0, \alpha_i > 0 \ (i = 1, \dots, s), \sum_{i=1}^s \frac{1}{\alpha_i} = 0, z^i \in \mathbb{R}^n \right\} \geq 0.$$

A similar observation can be made for extreme rays as in Case 2 of Theorem 5.46. In order to solve such a minimization problem, the common tool of Karush Kuhn Tucker (KKT) conditions (see [34, Satz 16.26]) from optimization can be used.

We start with the following definition, which can be found in [35, Section 10].

Definition 5.48. Let $m, p \in \mathbb{N}$ and $J: \mathbb{R}^n \rightarrow \mathbb{R}, e: \mathbb{R}^n \rightarrow \mathbb{R}^m, g: \mathbb{R}^n \rightarrow \mathbb{R}^p$ be such that e and g are twice (partially) differentiable and consider the minimization problem

$$\min J(x) \quad \text{s.t.} \quad e(x) = 0, \quad g(x) \leq 0, \quad x \in \mathbb{R}^n. \quad (52)$$

A point $x^* \in \mathbb{R}^n$ is called **admissible point** for (52) if $e(x^*) = 0$ and $g(x^*) \leq 0$. The set of **active indices** at x^* is the set $\mathcal{A} = \{j \in \{1, \dots, p\} : g_j(x^*) = 0\}$. The point x^* is called **local minimum** of (52) if there is some $\varepsilon > 0$ such that $J(x^*) < J(x)$ for all admissible $x \in B_\varepsilon(x^*)$, where $B_\varepsilon(x^*)$ is the ball at x^* with radius ε with respect to the Euclidean norm. Further, we say that x^* is a **regular point** for (52) if the set of vectors $\nabla e_i(x^*)$ ($i = 1, \dots, m$) and $\nabla g_j(x^*)$ ($j \in \mathcal{A}$) is linearly independent.

The following theorem is a very important tool in nonlinear optimization. It gives first order necessary optimality conditions - the KKT conditions -

for a minimization problem of the form (52). It can be found in [34, Satz 16.26] or [35, Satz 10.18].

Theorem 5.49 (Karush Kuhn Tucker). *Let $n, m, p \in \mathbb{N}$ and $J: \mathbb{R}^n \rightarrow \mathbb{R}$, $e: \mathbb{R}^n \rightarrow \mathbb{R}^m, g: \mathbb{R}^n \rightarrow \mathbb{R}^p$ be such that e and g are twice (partially) differentiable and consider the minimization problem (52). If $x^* \in \mathbb{R}^n$ is a local minimum and a regular point of (52) then there are $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that $\mu^* \geq 0$ and*

$$\nabla J(x^*) + e'(x^*)^\top \lambda^* + g'(x^*)^\top \mu^* = 0, \quad (53)$$

$$(\mu^*)^\top g(x^*) = 0. \quad (54)$$

Proof. [35, Satz 10.18]. \square

Using the KKT conditions (53)-(54), the following two optimization problems can be solved. They will later help us in verifying that a given $p \in C_{n,2d}$ satisfies $\ell(p) \geq 0$ for all extreme directions $\ell \in \Sigma_{n,2d}$. A proof of the following lemma can be found in [12, Appendix E].

Lemma 5.50. *Let $x_1, \dots, x_k \geq 0$ be arbitrary. Then it holds*

$$\min \left\{ \sum_{i=1}^k \alpha_i x_i : \alpha_i > 0, \sum_{i=1}^k \frac{1}{\alpha_i} = 1 \right\} = \left(\sum_{i=1}^k \sqrt{x_i} \right)^2. \quad (55)$$

Further, for $z \in \mathbb{C}^n$ arbitrary, we have

$$\max \left\{ \alpha z + \bar{\alpha} \bar{z} : \alpha \in \mathbb{C} \setminus \{0\}, \frac{1}{\alpha} + \frac{1}{\bar{\alpha}} = 1 \right\} = 2(|z| + \operatorname{Re}(z)) \quad (56)$$

Proof. Without loss of generality, we can assume that at least one x_i is nonzero, otherwise the claim is trivial. Problem (55) can be rewritten in the standard form (52) as

$$\begin{aligned} \min J(\alpha) = \sum_{i=1}^k \alpha_i x_i \quad \text{s.t.} \quad & e(\alpha) = \left(\sum_{i=1}^k \frac{1}{\alpha_i} \right) - 1 = 0 \\ & g(\alpha) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} > 0. \end{aligned} \quad (57)$$

Assume that $\alpha^* \in \mathbb{R}^n$ is a local minimum and a regular point of (57). Note that we can use the Karush Kuhn Tucker conditions although we have a strict inequality, since an admissible $\alpha \in \mathbb{R}^k$ must necessarily be nonzero in every component. Otherwise $e(\alpha)$ was not welldefined. Further, every admissible $\alpha \in \mathbb{R}^n$ is already a regular point for (57). Hence, Theorem 5.49 yields that there are $\lambda^* \in \mathbb{R}$ and $\mu^* \in \mathbb{R}^k$ such that

$$\begin{aligned} \nabla J(\alpha^*) + e'(\alpha^*)^\top \lambda^* + g'(\alpha^*)^\top \mu^* &= 0 \\ (\mu^*)^\top g(\alpha^*) &= 0 \end{aligned} \quad (58)$$

for some $\lambda^* \in \mathbb{R}$ and $\mu^* \in \mathbb{R}^k, \mu^* \geq 0$. In our case, we have

$$\nabla J(\alpha) = \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \in \mathbb{R}^k, \quad e'(\alpha)^\top = \begin{pmatrix} -\frac{1}{\alpha_1^2} \\ \vdots \\ -\frac{1}{\alpha_k^2} \end{pmatrix} \in \mathbb{R}^k, \quad g'(\alpha)^\top = I_k \in \mathbb{R}^{k \times k}$$

for $\alpha \in \mathbb{R}^n$. Hence, (58) can be rewritten as

$$\begin{aligned} x_i - \frac{1}{(\alpha_i^*)^2} \lambda^* + \mu_i^* &= 0 \quad (i = 1, \dots, k) \\ \sum_{i=1}^k \mu_i^* \alpha_i^* &= 0. \end{aligned} \quad (59)$$

Since $g(\alpha^*) = \alpha^* > 0$ and $\mu^* \geq 0$, it follows from the second line of (59) that $\mu^* = 0$. For this reason, (59) can be reduced to

$$x_i - \frac{1}{(\alpha_i^*)^2} \lambda^* = 0 \quad (i = 1, \dots, k). \quad (60)$$

Multiplying each equation of (60) by α_i^* and summing them together shows

$$\begin{aligned} \sum_{i=1}^k \alpha_i^* x_i - \underbrace{\left(\sum_{i=1}^k \frac{1}{\alpha_i^*} \right)}_{=1} \lambda^* &= 0 \\ \iff \lambda^* &= \sum_{i=1}^k \alpha_i^* x_i > 0. \end{aligned} \quad (61)$$

Further, we can deduce from (60) that $\frac{1}{\alpha_i^*} = \left(\frac{x_i}{\lambda^*}\right)^{1/2}$ for all $i = 1, \dots, k$.

Inserting this into $e(\alpha^*) = \left(\sum_{i=1}^k \frac{1}{\alpha_i^*}\right) - 1 = 0$ shows

$$1 = \sum_{i=1}^k \left(\frac{x_i}{\lambda^*}\right)^{1/2} = \frac{1}{\sqrt{\lambda^*}} \sum_{i=1}^k \sqrt{x_i}$$

and hence $\lambda^* = \left(\sum_{i=1}^k \sqrt{x_i}\right)^2$. Therefore, equation (61) yields

$$J(\alpha^*) = \sum_{i=1}^k \alpha_i^* x_i = \lambda^* = \left(\sum_{i=1}^k \sqrt{x_i}\right)^2,$$

as desired. On the other hand, this minimum can actually be obtained if we take

$$\alpha_i^* = \begin{cases} \left(\sum_{i=1}^k \sqrt{x_i}\right) \cdot x_i^{-\frac{1}{2}}, & \text{if } x_i \neq 0 \\ 0, & \text{else} \end{cases} \quad (i = 1, \dots, k).$$

This shows (55).

For the second part, let $\alpha \in \mathbb{C}$ such that $\frac{1}{\alpha} + \frac{1}{\bar{\alpha}} = 1$ be arbitrary. Let $\varphi \in \mathbb{R}$ be such that $\alpha = |\alpha| e^{i\varphi} = |\alpha| (\cos \varphi + i \sin \varphi)$. Then it holds $\bar{\alpha} =$

$|\alpha| (\cos \varphi - i \sin \varphi)$ and hence

$$\frac{1}{\alpha} + \frac{1}{\bar{\alpha}} = \frac{1}{|\alpha|} \cdot \left(\frac{1}{\cos \varphi + i \sin \varphi} + \frac{1}{\cos \varphi - i \sin \varphi} \right) = \frac{1}{|\alpha|} \cdot 2 \cos \varphi.$$

This shows that $\frac{1}{\alpha} + \frac{1}{\bar{\alpha}} = 1$ if and only if $|\alpha| = 2 \cos \varphi$, i.e. α is of the form $\alpha = 2 \cos(\varphi) e^{i\varphi}$ for some $\varphi \in \mathbb{R}$ such that $\cos(\varphi) > 0$. This yields

$$\begin{aligned} & \max \left\{ \underbrace{\alpha z + \bar{\alpha} \bar{z}}_{=2 \operatorname{Re}(\alpha z)} : \alpha \in \mathbb{C} \setminus \{0\}, \frac{1}{\alpha} + \frac{1}{\bar{\alpha}} = 1 \right\} \\ &= \max_{\varphi \in \mathbb{R}, \cos \varphi > 0} 2 \operatorname{Re} (2 \cos(\varphi) e^{i\varphi} z). \end{aligned}$$

Further, writing $z = |z| e^{i\vartheta}$ for some $\vartheta \in \mathbb{R}$ leads to

$$\begin{aligned} & \max_{\varphi \in \mathbb{R}, \cos \varphi > 0} 2 \operatorname{Re} (2 \cos(\varphi) e^{i\varphi} z) \\ &= \max_{\varphi \in \mathbb{R}, \cos \varphi > 0} 4 |z| \operatorname{Re} (\cos(\varphi) e^{i(\varphi+\vartheta)}) \\ &= 4 |z| \cdot \max_{\varphi \in \mathbb{R}, \cos \varphi > 0} \cos(\varphi) \cdot \cos(\varphi + \vartheta). \end{aligned}$$

Note that $\cos x \cdot \cos y = \frac{1}{2} (\cos(x - y) + \cos(x + y))$ for $x, y \in \mathbb{R}$. Using this formula for $x = \varphi + \vartheta, y = \varphi$ leads to

$$\begin{aligned} & 4 |z| \cdot \max_{\varphi \in \mathbb{R}, \cos \varphi > 0} \cos(\varphi) \cdot \cos(\varphi + \vartheta) \\ &= 2 |z| \cdot \max_{\varphi \in \mathbb{R}, \cos \varphi > 0} (\cos(\vartheta) + \cos(2\varphi + \vartheta)) \\ &= 2 |z| \cdot (\cos(\vartheta) + 1) \\ &= 2(|z| \cos(\vartheta) + |z|) \\ &= 2(|z| + \operatorname{Re}(z)), \end{aligned}$$

as desired. \square

Finally, the following theorem can be proven, which can be found [12, Theorem 4.2].

Theorem 5.51. *A nonnegative quaternary quartic $p \in P_{4,4}$ is SOS if and only if both of the following two conditions hold:*

(1.) *For every $r^1, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}$ and every $a_2, \dots, a_8 \in \{-1, 1\}$ such that $r^1 (r^1)^\top = \sum_{i=2}^8 a_i r^i (r^i)^\top$, it holds*

$$p(r^1) \leq \left(\sum_{i=2}^8 \sqrt{p(r^i)} \right)^2. \quad (62)$$

(2.) *For every $r^3, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}$ and $z \in \mathbb{C}^4 \setminus \{0\}$, such that $\bar{z} \neq z$ and every $a_3, \dots, a_8 \in \{-1, 1\}$ such that $z z^\top + \bar{z} \bar{z}^\top = \sum_{i=3}^8 a_i r^i (r^i)^\top$, it holds*

$$2(|p(z)| + \operatorname{Re}(p(z))) \leq \left(\sum_{i=3}^8 \sqrt{p(r^i)} \right)^2. \quad (63)$$

Proof. By Theorem 5.42 we know that a form $p \in P_{4,4}$ is SOS if for every $\ell \in \Sigma_{4,4}^*$ of the form $\ell = \sum_{i=1}^8 \alpha_i \text{ev}_{z^i}$, where $\alpha_i \in \mathbb{C} \setminus \{0\}$ and $z^i \in \mathbb{C}^4 \setminus \{0\}$, it holds

$$\ell(p) = \sum_{i=1}^8 \alpha_i p(z^i) \geq 0. \quad (64)$$

Since $\Sigma_{4,4}^*$ is a closed convex cone by Proposition 3.80, we know by Theorem 3.75 that $\Sigma_{4,4}^*$ is the convex hull of its extreme directions. Hence, $p \in \Sigma_{4,4}$ is SOS if and only if (64) holds for all extreme directions $\ell \in \Sigma_{4,4}^*$.

Further, the points z^i in (64) must satisfy the Cayley–Bacharach relation $\sum_{i=1}^8 a_i \text{ev}_{z^i} = 0$ on $H_{n,d} = H_{4,2}$ for some $a_i \in \mathbb{C} \setminus \{0\}$. We want to rewrite the linear relation of point evaluations as linear relation on the symmetric $n \times n$ matrices $z^i(z^i)^\top$ ($i = 1, \dots, 8$) as in the claim. Clearly, it holds

$$\begin{aligned} \sum_{i=1}^8 a_i \text{ev}_{z^i} &= a_1 \text{ev}_{z^1} + \dots + a_8 \text{ev}_{z^8} = 0 \quad \text{on } H_{n,d} = H_{4,2} \\ \iff \forall q \in H_{n,d} &: a_1 q(z^1) + \dots + a_8 q(z^8) = 0 \\ \stackrel{(*)}{\iff} \forall i, j = 1, \dots, 4 &: a_1 z_i^1 z_j^1 + \dots + a_8 z_i^8 z_j^8 = 0 \\ \iff a_1 z^1(z^1)^\top &+ \dots + a_8 z^8(z^8)^\top = 0, \end{aligned}$$

where (*) follows since $\{\mathbf{x}_i \mathbf{x}_j : i, j = 1, \dots, 4\}$ forms a basis of $H_{4,2}$.

In addition, as in Convention 5.45, we can assume that either all z^i are real and it holds $a_i \in \{-1, 1\}$ for all $i = 1, \dots, 8$ or they contain exactly one pair of complex conjugated vectors. In the second case, we can without loss of generality assume that $\bar{z}^1 = z^2 \neq z^1$, $z^i \in \mathbb{R}^4$ for $i = 3, \dots, 8$ and $a_1 = a_2 = 1$, $a_i \in \{-1, 1\}$ ($i = 3, \dots, 8$). This yields that we have one of the following two cases:

Case 1: It holds $z^i =: r^i \in \mathbb{R}^4 \setminus \{0\}$ for all $i = 2, \dots, 8$ and $\sum_{i=1}^8 a_i r^i (r^i)^\top = 0$ for some $a_i \in \{-1, 1\}$ ($i = 1, \dots, 8$). After dividing by a_1 and rearranging, this can be written as $r^1 (r^1)^\top = \sum_{i=2}^8 \tilde{a}_i r^i (r^i)^\top$ where $\tilde{a}_i \in \{-1, 1\}$ ($i = 1, \dots, 8$).

Case 2: It holds $z^1 =: z, z^2 = \bar{z} \neq z \in \mathbb{C}^4 \setminus \{0\}$, $z^i =: r^i \in \mathbb{R}^4 \setminus \{0\}$ for $i = 3, \dots, 8$ and $z z^\top + \bar{z} \bar{z}^\top + \sum_{i=3}^8 a_i r^i (r^i)^\top = 0$ for some $a_i \in \{-1, 1\}$ ($i = 3, \dots, 8$). After dividing by -1 and rearranging, this can be written as $z z^\top + \bar{z} \bar{z}^\top = \sum_{i=3}^8 \tilde{a}_i r^i (r^i)^\top$, where $\tilde{a}_i \in \{-1, 1\}$ ($i = 3, \dots, 8$).

For simplicity, we denote the 'new' Cayley–Bacharach coefficients $\tilde{a}_i \in \{-1, 1\}$ again by a_i .

Further, Theorem 5.46 tells us that in Case 1 and Case 2 as above, the coefficients $\alpha_i \in \mathbb{C} \setminus \{0\}$, that together with the points z^i determine an extreme ray ℓ as in (64), must satisfy the following

Case 1: All α_i are nonzero reals, i.e. $\alpha_i \in \mathbb{R} \setminus \{0\}$ ($i = 1, \dots, 8$) and satisfy

$$\sum_{i=1}^8 \frac{1}{\alpha_i} = 0.$$

Further, it holds $\alpha_1 < 0$ and $\alpha_i > 0$ for $i = 2, \dots, 8$. After scaling the coefficients $\alpha_i \in \mathbb{R} \setminus \{0\}$, we can without loss of generality assume that

$$\sum_{i=2}^8 \frac{1}{\alpha_i} = -\frac{1}{\alpha_1} = 1. \quad (65)$$

In particular, this means $\alpha_1 = -1$.

Case 2: It holds $\alpha_1 =: \alpha, \alpha_2 = \bar{\alpha} \in \mathbb{C} \setminus \{0\}$ and $\alpha_i \in \mathbb{R} \setminus \{0\}$ ($i = 3, \dots, 8$) such that

$$\frac{1}{\alpha} + \frac{1}{\bar{\alpha}} + \sum_{i=3}^8 \frac{1}{\alpha_i} = 0.$$

Again, after scaling the coefficients $\alpha_i \in \mathbb{R} \setminus \{0\}$ and $\alpha \in \mathbb{C} \setminus \{0\}$, we can without loss of generality assume that

$$\sum_{i=3}^8 \frac{1}{\alpha_i} = -\left(\frac{1}{\alpha} + \frac{1}{\bar{\alpha}}\right) = 1. \quad (66)$$

Hence, it can be deduced that (64) holds for all extreme $\ell \in \Sigma_{4,4}^*$ if and only if the following two conditions hold

(1.) For all $r^1, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}, a_2, \dots, a_8 \in \{-1, 1\}$ and $\alpha_2, \dots, \alpha_8 \in \mathbb{R} \setminus \{0\}$ such that $r^1(r^1)^\top = \sum_{i=2}^8 a_i r^i (r^i)^\top$ and $\sum_{i=2}^8 \frac{1}{\alpha_i} = 1$, it holds

$$p(r^1) \leq \sum_{i=2}^8 \alpha_i p(r^i).$$

(2.) For all $z \in \mathbb{C}^4 \setminus \{0\}, z \neq \bar{z}, r^3, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}, a_3, \dots, a_8 \in \{-1, 1\}$ and $\alpha \in \mathbb{C} \setminus \{0\}, \alpha_3, \dots, \alpha_8 \in \mathbb{R} \setminus \{0\}$ such that $zz^\top + \bar{z}\bar{z}^\top = \sum_{i=3}^8 a_i r^i (r^i)^\top$ and $\sum_{i=3}^8 \frac{1}{\alpha_i} = -\left(\frac{1}{\alpha} + \frac{1}{\bar{\alpha}}\right) = 1$, it holds

$$-(\alpha p(z) + \bar{\alpha} p(\bar{z})) \leq \sum_{i=3}^8 \alpha_i p(r^i).$$

Writing those conditions as optimization problems shows that $p \in P_{4,4}$ is SOS if and only if both of the following two conditions hold.

(I) For all $r^1, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}, a_2, \dots, a_8 \in \{-1, 1\}$ such that $r^1(r^1)^\top = \sum_{i=2}^8 a_i r^i (r^i)^\top$, it holds

$$p(r^1) \leq \min \left\{ \sum_{i=2}^8 \alpha_i p(r^i) : \alpha_i > 0, \sum_{i=2}^8 \frac{1}{\alpha_i} = 1 \right\}.$$

(II) For all $z \in \mathbb{C}^4 \setminus \{0\}, \bar{z} \neq z, r^3, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}, a_3, \dots, a_8 \in \{-1, 1\}$ such that $zz^\top + \bar{z}\bar{z}^\top = \sum_{i=3}^8 a_i r^i (r^i)^\top$, it holds

$$\begin{aligned} & \max \left\{ (-\alpha)p(z) + (-\bar{\alpha})p(\bar{z}) : \frac{1}{-\alpha} + \frac{1}{-\bar{\alpha}} = 1 \right\} \\ & \leq \min \left\{ \sum_{i=3}^8 \alpha_i p(r^i) : \alpha_i > 0, \sum_{i=3}^8 \frac{1}{\alpha_i} = 1 \right\}. \end{aligned}$$

Finally, by Lemma 5.50, we know the exact value of each of the optimization problems. Therefore, the two conditions (I) and (II) can be rewritten as follows.

(I) For all $r^1, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}$ and $a_2, \dots, a_8 \in \{-1, 1\}$ such that $r^1(r^1)^\top = \sum_{i=2}^8 a_i r^i (r^i)^\top$, it holds

$$p(r^1) \leq \min \left\{ \sum_{i=2}^8 \alpha_i p(r^i) : \alpha_i > 0, \sum_{i=2}^8 \frac{1}{\alpha_i} = 1 \right\} = \left(\sum_{i=2}^8 \sqrt{p(r^i)} \right)^2.$$

(II) For all $z \in \mathbb{C}^4 \setminus \{0\}$, $\bar{z} \neq z$, $r^3, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}$ and $a_3, \dots, a_8 \in \{-1, 1\}$ such that $zz^\top + \bar{z}\bar{z}^\top = \sum_{i=3}^8 a_i r^i (r^i)^\top$, it holds

$$\begin{aligned} 2(|p(z)| + \operatorname{Re}(p(z))) &= \max \left\{ (-\alpha)p(z) + (-\bar{\alpha})p(\bar{z}) : \frac{1}{-\alpha} + \frac{1}{-\bar{\alpha}} = 1 \right\} \\ &\leq \min \left\{ \sum_{i=3}^8 \alpha_i p(r^i) : \alpha_i > 0, \sum_{i=3}^8 \frac{1}{\alpha_i} = 1 \right\} = \left(\sum_{i=3}^8 \sqrt{p(r^i)} \right)^2. \end{aligned}$$

This shows the claim. \square

Remark 5.52. If Conjecture 5.43 turns out to be true, it would suffice to consider Case 1 of Theorem 5.46. This would mean that for $p \in P_{4,4}$ to be SOS, it would suffice to show that property (1.) of Theorem 5.51 holds.

In the following, one can see how Theorem 5.51 can be used to show that a certain PSD form $p \in P_{4,4}$ is not SOS.

Example 5.53. As mentioned in Remark 2.19, the following quaternary quartic, which goes back to Robinson, is known to be PSD but not SOS:

$$p(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) = \mathbf{w}^4 + \mathbf{x}^2 \mathbf{y}^2 + \mathbf{y}^2 \mathbf{z}^2 + \mathbf{x}^2 \mathbf{z}^2 - 4\mathbf{x}\mathbf{y}\mathbf{z}\mathbf{w} \in P_{4,4} \setminus \Sigma_{4,4}.$$

Assuming the nonnegativity of p , one can use the certificate from Theorem 5.51 to show that p is not SOS. Therefore, take the following points $r^2, \dots, r^8 \in \mathbb{R}^4$:

$$\begin{aligned} r^2 &= e^1, r^3 = e^2, r^4 = e^3 \\ r^5 &= (1 \ 1 \ 1 \ 1)^\top, \quad r^6 = (1 \ 1 \ -1 \ -1)^\top, \\ r^7 &= (1 \ -1 \ 1 \ -1)^\top, \quad r^8 = (1 \ -1 \ -1 \ 1)^\top. \end{aligned}$$

It can easily be seen that $p(r^2) = \dots = p(r^8) = 0$. Further, for $r^1 = e^4$ it holds $p(r^1) = 1 > 0$. We also have the following linear relation between the points r^1, \dots, r^8

$$\begin{aligned} 4r^1(r^1)^\top &= -4 \cdot \left(r^2(r^2)^\top + r^3(r^3)^\top + r^4(r^4)^\top \right) \\ &\quad + r^5(r^5)^\top + r^6(r^6)^\top + r^7(r^7)^\top + r^8(r^8)^\top. \end{aligned} \tag{67}$$

Note that the points r^1, \dots, r^8 are indeed common zeros of the three real quadratic forms $\{\mathbf{x}\mathbf{y} - \mathbf{z}\mathbf{w}, \mathbf{x}\mathbf{z} - \mathbf{y}\mathbf{w}, \mathbf{x}\mathbf{w} - \mathbf{y}\mathbf{z}\} \subseteq H_{4,2}$, see [30, Section 1].

In Theorem 5.51, we want the coefficients of the linear relation in (67) to be in $\{-1, 1\}$ therefore, define

$$\tilde{r}^i := \begin{cases} 2r^i, & i = 1, \dots, 4 \\ r^i, & i = 5, \dots, 8 \end{cases}.$$

Equation (67) then yields

$$\tilde{r}^1(\tilde{r}^1)^\top = \sum_{i=2}^4 (-1) \cdot \tilde{r}^i(\tilde{r}^i)^\top + \sum_{i=5}^8 \tilde{r}^i(\tilde{r}^i)^\top.$$

Further, since $p(\tilde{r}^1) = p(2e^1) = 16 > 0 = \left(\sum_{i=2}^8 \sqrt{p(\tilde{r}^i)}\right)^2$ it can be deduced from Theorem 5.51 that p is not SOS.

6. CONVEX QUARternary QUARTICS ARE SOS

In this section, we prove our main result, which is that every convex quarternary quartic is already SOS, i.e. $C_{4,4} \subseteq \Sigma_{4,4}$. This can be shown based on the previous results from Section 4 and 5. Precisely, we follow the proof in [12, Section 5], which uses the Generalized Cauchy–Schwarz Inequalities from Theorem 4.3 to show that any convex quarternary quartic $p \in C_{4,4}$ satisfies the conditions from Theorem 5.51, which then yields that p is SOS. Therefore, it is crucial to know that the optimal constants $A_2^*, B_2^* > 0$ from the Generalized Cauchy–Schwarz Inequalities (23) and (24) are both equal to 1.

Theorem 6.1. *Every convex quarternary quartic is a SOS, i.e. $C_{4,4} \subseteq \Sigma_{4,4}$.*

Proof. Let $p \in C_{4,4}$ be arbitrary. By Theorem 4.3, we know that p satisfies the following Generalized Cauchy–Schwarz Inequalities

$$\forall x, y \in \mathbb{R}^4 : Q_p(x, y) \leq A_2 \sqrt{p(x)p(y)} \quad (68)$$

$$\forall z \in \mathbb{C}^4 : |p(z)| \leq B_2 Q_p(z, \bar{z}) \quad (69)$$

for some $A_2, B_2 > 0$, where Q_p is the biform associated to p . Further, by Proposition 4.12 and 4.29, we know that the optimal constants $A_d^*, B_d^* > 0$ from (68) and (69) for $d = 2$ are both equal to 1. This yields

$$\forall x, y \in \mathbb{R}^4 : Q_p(x, y) \leq \sqrt{p(x)p(y)} \quad (70)$$

$$\forall z \in \mathbb{C}^4 : |p(z)| \leq Q_p(z, \bar{z}). \quad (71)$$

To deduce that p is SOS, it suffices to show that p fulfills the two conditions from Theorem 5.51.

(1.) Let $r^1, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}$ and $a_2, \dots, a_8 \in \{-1, 1\}$ be arbitrary such that $r^1(r^1)^\top = \sum_{i=2}^8 a_i r^i(r^i)^\top$.

$$\text{To show: } p(r^1) \leq \left(\sum_{i=2}^8 \sqrt{p(r^i)}\right)^2.$$

Note that the matrix product of a vector $x \in \mathbb{R}^n$ with its transpose is the same as the symmetric outer product of x viewed as a tensor with itself, i.e. $xx^\top = x \odot x = x^2$. Hence, using those notations from Section

3.2, we have $(r^1)^2 = \sum_{i=2}^8 a_i (r^i)^2$. Squaring both sides, i.e. taking again the symmetric outer product of each side with itself yields

$$(r^1)^4 = \sum_{i,j=2,\dots,8} a_i a_j (r^i)^2 (r^j)^2, \quad (72)$$

where we used the commutativity of the symmetric outer product. Moreover, write $p = \sum_{j_1,\dots,j_4=1,\dots,4} \tilde{p}^{j_1 \dots j_4} \mathbf{x}_{j_1} \cdots \mathbf{x}_{j_4}$ with coefficients chosen as in the proof of Theorem 3.29. Note that by definition of the quadratic form Q_p via the tensor T_p associated to p , it holds as in (5) that

$$\begin{aligned} Q_p(x, y) &= \sum_{j_1,\dots,j_4=1,\dots,4} \tilde{p}^{j_1 \dots j_4} x_{j_1} x_{j_2} y_{j_3} y_{j_4} \\ &= \sum_{j_1,\dots,j_4=1,\dots,4} \tilde{p}^{j_1 \dots j_4} (x^2 y^2)^{j_1 \dots j_4}, \end{aligned} \quad (73)$$

where the last step follows as in the proof of Proposition 3.26. For this reason, it holds

$$\begin{aligned} p(r^1) &= \sum_{j_1,\dots,j_4=1,\dots,4} \tilde{p}^{j_1 \dots j_4} ((r^1)^4)^{j_1 \dots j_4} \\ &\stackrel{(72)}{=} \sum_{i,j=2,\dots,8} a_i a_j \sum_{j_1,\dots,j_4=1,\dots,4} \tilde{p}^{j_1 \dots j_4} ((r^i)^2 (r^j)^2)^{j_1 \dots j_4} \\ &\stackrel{(73)}{=} \sum_{i,j=2,\dots,8} \underbrace{a_i a_j Q_p(r^i, r^j)}_{\in \{\pm Q_p(r^i, r^j)\}}. \end{aligned} \quad (74)$$

It remains to show:

$$\forall x, y \in \mathbb{R}^4 : |Q_p(x, y)| \leq \sqrt{p(x)p(y)}. \quad (75)$$

Note that this is a priori not clear, since the Generalized Cauchy–Schwarz Inequality (68) is stated without an absolute value on the left hand side.

Recall that by Example 3.59(c), we know that $Q_p(x, y) = \frac{1}{(2d)!} \partial_x^d \partial_y^d p$. Further, Corollary 3.87 tells us that $C_{n,2d}^* = \{\partial_x^{2d-2} \partial_y^2 : x, y \in \mathbb{R}^n\}$ for all n, d . For $(n, 2d) = (4, 4)$, i.e. $d = 2$, this means in particular that the mapping

$$H_{n,2d} \rightarrow \mathbb{R}, q \mapsto Q_q(x, y)$$

belongs to $C_{4,4}^*$ for all $x, y \in \mathbb{R}^4$. Since $p \in C_{4,4}$ is convex, this yields $Q_p(x, y) \geq 0$ for all $x, y \in \mathbb{R}^4$ and hence

$$|Q_p(x, y)| = Q_p(x, y) \stackrel{(70)}{\leq} \sqrt{p(x)p(y)},$$

which is (75). Finally, one can see that

$$\begin{aligned} p(r^1) &\stackrel{(74)}{=} \sum_{i,j=2}^8 a_i a_j Q_p(r^i, r^j) \leq \sum_{i,j=2,\dots,8} \underbrace{|a_i a_j|}_{=1} |Q_p(r^i, r^j)| \\ &\stackrel{(75)}{\leq} \sum_{i,j=2,\dots,8} \sqrt{p(r^i)p(r^j)} = \left(\sum_{i=2}^8 \sqrt{p(r^i)} \right)^2. \end{aligned}$$

Hence, Condition (1.) from Theorem 5.51 holds.

- (2.) Let $r^3, \dots, r^8 \in \mathbb{R}^4 \setminus \{0\}$, $z \in \mathbb{C}^4 \setminus \{0\}$ be such that $\bar{z} \neq z$ and $a_3, \dots, a_8 \in \{-1, 1\}$ be arbitrary such that $z z^\top + \bar{z} \bar{z}^\top = \sum_{i=3}^8 a_i r^i (r^i)^\top$ holds. As in (1.), using the tensor notation and taking the symmetric outer product of each side with itself, shows

$$z^4 + 2z^2 \bar{z}^2 + \bar{z}^4 = \sum_{i,j=3,\dots,8} a_i a_j (r^i)^2 (r^j)^2. \quad (76)$$

Similarly to (1.), we obtain

$$\begin{aligned} p(z) + 2Q_p(z, \bar{z}) + p(\bar{z}) &= \sum_{j_1, \dots, j_4=1,\dots,4} \tilde{p}^{j_1 \dots j_4} (z^4 + 2z^2 \bar{z}^2 + \bar{z}^4)^{j_1 \dots j_4} \\ &\stackrel{(76)}{=} \sum_{i,j=3,\dots,8} a_i a_j \sum_{j_1, \dots, j_4=1,\dots,4} \tilde{p}^{j_1 \dots j_4} \left((r^i)^2 (r^j)^2 \right)^{j_1 \dots j_4} \\ &\stackrel{(73)}{=} \sum_{i,j=3,\dots,8} a_i a_j Q_p(r^i, r^j). \end{aligned} \quad (77)$$

Further, note that $p(z) + p(\bar{z}) = 2 \operatorname{Re}(p(z))$. Finally, again using the Generalized Cauchy–Schwarz Inequalities yields

$$\begin{aligned} 2(|p(z)| + \operatorname{Re}(p(z))) &\stackrel{(71)}{\leq} 2(Q_p(z, \bar{z}) + \operatorname{Re}(p(z))) \\ &= p(z) + 2Q_p(z, \bar{z}) + p(\bar{z}) \\ &\stackrel{(77)}{=} \sum_{i,j=3,\dots,8} a_i a_j Q_p(r^i, r^j). \\ &\leq \sum_{i,j=3,\dots,8} |Q_p(r^i, r^j)| \\ &\stackrel{(75)}{\leq} \sum_{i,j=3,\dots,8} \sqrt{p(r^i)p(r^j)} = \left(\sum_{i=3}^8 \sqrt{p(r^i)} \right)^2. \end{aligned}$$

Hence, condition (2.) from Theorem 5.51 holds as well, which shows that p is indeed SOS. Therefore, we have shown that $C_{4,4} \subseteq \Sigma_{4,4}$.

□

7. FURTHER EXTENSIONS AND OUTLOOK

In this section, we want to first give some remarks on a possible extension of the result presented in Section 6 to the case of ternary sextics and then an outlook on possible future research directions.

7.1. Possible Extension to the Case of Ternary Sextics. In an analogous manner as in Theorem 5.51, one can get a certificate for a nonnegative ternary sextic $p \in P_{3,6}$ to be SOS. This is stated without proof as the proof is completely analogous to the one of Theorem 5.51. It can also be found in [12, Theorem 6.1].

Theorem 7.1. *A nonnegative ternary sextic $p \in P_{3,6}$ is SOS if and only if both of the following two conditions hold:*

(1.) *For every $r^1, \dots, r^9 \in \mathbb{R}^4 \setminus \{0\}$ and every $a_2, \dots, a_9 \in \{-1, 1\}$ such that $(r^1)^3 = \sum_{i=2}^9 a_i (r^i)^3$, it holds*

$$p(r^1) \leq \left(\sum_{i=2}^9 \sqrt{p(r^i)} \right)^2. \quad (78)$$

(2.) *For every $r^3, \dots, r^9 \in \mathbb{R}^4 \setminus \{0\}$ and $z \in \mathbb{C}^4 \setminus \{0\}$, such that $\bar{z} \neq z$ and every $a_3, \dots, a_9 \in \{-1, 1\}$ such that $z^3 + \bar{z}^3 = \sum_{i=3}^9 a_i (r^i)^3$, it holds*

$$2(|p(z)| + \operatorname{Re}(p(z))) \leq \left(\sum_{i=3}^9 \sqrt{p(r^i)} \right)^2. \quad (79)$$

We could now try to follow the steps in the proof of Theorem 6.1 to show that every convex ternary sextic is SOS as well, i.e. to show that $C_{3,6} \subseteq \Sigma_{3,6}$ holds. Therefore, we would take an arbitrary convex ternary sextic $p \in C_{3,6}$ and we would have to show that it satisfies both conditions of Theorem 7.1. As we remember, it was crucial to know in the proof of Theorem 6.1 that the Generalized Cauchy–Schwarz Inequalities for convex quaternary quartics hold with optimal constants $A_2^* = B_2^* = 1$.

For the case of ternary sextics we know on the one hand by Proposition 4.12 that $A_3^* = 1$ holds as well. Hence, we could proceed as in the proof of Theorem 6.1 to show that ternary sextics fulfill condition (78). The only part, which would be different, is to show that an arbitrary $p \in C_{3,6}$ fulfills

$$\forall x, y \in \mathbb{R}^3 : |Q_p(x, y)| \leq \sqrt{p(x)p(y)},$$

This corresponds to (75) in the proof for quaternary quartics.

For ternary sextics, this can be seen as follows. If $|Q_p(x, y)| = Q_p(x, y)$, the inequality follows from the first Generalized Cauchy–Schwarz Inequality. Further, if $|Q_p(x, y)| = -Q_p(x, y)$, we have

$$\begin{aligned} |Q_p(x, y)| &= -Q_p(x, y) \stackrel{d \text{ odd}}{=} Q_p(-x, y) \leq \sqrt{p(-x)p(y)} \\ &\stackrel{\deg(p)=2d}{=} \sqrt{(-1)^{2d}p(x)p(y)} = \sqrt{p(x)p(y)}, \end{aligned}$$

as desired. This way it can be seen that an arbitrary ternary sextic fulfills the first condition (78) of Theorem 7.1.

But on the other hand, by Proposition 4.29 we know that the second optimal constant $B_3^* = \frac{\binom{2,2}{3}}{3} = 2 > 1$ of the Generalized Cauchy Schwarz inequalities is strictly larger than one. For this reason, we would fail in showing that (79) holds for an arbitrary $p \in C_{3,6}$, if we proceed as in the proof of Theorem 6.1.

However, if Conjecture 5.43 turns out to be true, we could conclude as in Remark 5.52 that it would suffice for a convex form $p \in C_{3,6}$ to be SOS if it fulfills property (78) of Theorem 7.1. Hence, if Conjecture 5.43 was true, we could conclude that every convex ternary sextic is SOS, i.e. $C_{3,6} \subseteq \Sigma_{3,6}$.

7.2. Outlook. We want to close this thesis by giving in the following some ideas on possible future work related on the topics discussed in this thesis:

- Determine the coefficients $A_d^* > 0$ of the first Generalized Cauchy–Schwarz Inequality (23). Therefore, consider for example Conjecture 4.19, which states that $A_d^* = 1$ for all odd $d \in \mathbb{N}$. One could for example start by searching an upper bound on A_d^* similar as in the proof of the optimal constants B_d^* in Section 4.2.2.
- It could also be interesting to go more into Semidefinite Programming (SDP) as this is used in [12, Section 2.3.1] to calculate the exact value of A_4^* as mentioned in Remark 4.18.
- As already mentioned in the previous Section, one could try to prove Conjecture 5.43 in order to extend the Result from Section 6 to the case of ternary sextics.
- Note that it was crucial for proving our Main Theorem in Section 6 to understand what separates the SOS cone $\Sigma_{4,4}$ from the non SOS forms inside the PSD cone $P_{4,4}$. This is based on the results of [4]. Hence, for a possible extension of our results beyond quarternary quartics or ternary sextics, it could be very important, to also get a general description of what separates $\Sigma_{n,2d}$ from $P_{n,2d}$.

REFERENCES

- [1] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, J. N. Tsitsiklis: *NP–Hardness of deciding Convexity of Quartic Polynomials and Related Problems*, Mathematical Programming, DOI: 10.1007/s10107-011-0499-2 (2011)
- [2] A. A. Ahmadi, P. A. Parrilo: *A complete Characterization of the Gap between Convexity and SOS–Convexity*, SIAM Journal on Optimization, 23 (2013), pp. 811–833
- [3] G. Blekherman: *Convex Forms that are not Sums of Squares*, arXiv preprint arXiv:0910.0656 (2009)
- [4] G. Blekherman: *Nonnegative Polynomials and Sums of Squares*, J. Amer. Math. Soc., 25 (2012), pp. 617–635
- [5] S. Boyd, L. Vandenberghe: *Convex Optimization*, Cambridge University Press (2004)
- [6] S. Boyd, L. Vandenberghe: *Semidefinite Programming*, SIAM Review, 38(1):49–95 (1996)
- [7] D. A. Cox, J. Little, D. O’Shea: *Ideals, Varieties and Algorithms - An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer International Publishing Switzerland (2015)
- [8] J. Dattorro: *Convex Optimization - Euclidean Distance Geometry*, Meboo Publishing USA (2019)

- [9] R. Denk, R. Racke: *Kompendium der Analysis, Band 1: Differential- und Integralrechnung, Gewöhnliche Differentialgleichungen*, Springer Spektrum (2011)
- [10] R. Denk, R. Racke: *Kompendium der Analysis, Band 2: Maß- und Integrationstheorie, Funktionentheorie, Funktionalanalysis, Partielle Differentialgleichungen*, Springer Spektrum (2012)
- [11] D. Eisenbud, M. Green, J. Harris: *Cayley-Bacharach Theorems and Conjectures*, Bulletin of the American Mathematical Society, 33 (1996), pp. 295-324
- [12] B. El Khadir: *On Sum of Squares Representation of Convex Forms and Generalized Cauchy-Schwarz Inequalities*, arXiv:1909.07546v1 (2019)
- [13] G. Fischer: *Lineare Algebra: Eine Einführung für Studienanfänger*, Springer Spektrum, 18. Auflage (2013)
- [14] C. G. Gibson: *Elementary Geometry of Algebraic Curves: An Undergraduate Introduction*, Cambridge University Press (1998)
- [15] V. Guillemin, A. Pollack: *Differential Topology*, Prentice-Hall, Englewood Cliffs, New Jersey (1974)
- [16] D. Hilbert: *Über die Darstellung definiter Formen als Summe von Formengquadraten*, Math. Ann., 32: 342-350, 1888; Ges. Abh. 2: 154-161, Springer, Berlin, reprinted by Chelsea, New York (1981)
- [17] P. Hilton, J. Pedersen: *Catalan Numbers, their Generalization, and their Uses*, The Mathematical Intelligencer, 13 (1991), pp. 64-75
- [18] J.-B. Hiriart-Urruty, C. Lemarchal: *Convex Analysis and Minimization Algorithms I: Fundamentals*, Springer Berlin Heidelberg (1993)
- [19] M. Infusino: *Lecture Notes on Positive Polynomials and Moment Problems*, <http://www.math.uni-konstanz.de/infusino/PPMP-SS19/PPMP-Notes-2019.pdf>, Universität Konstanz (2019)
- [20] N. Jeevanjee: *An Introduction to Tensors and Group Theory for Physicists*, Birkhäuser, Springer Science+Business Media (2011)
- [21] S. Kuhlmann: *Lecture Notes on Positive Polynome*, <https://www.math.uni-konstanz.de/kuhlmann/Lehre/SS10-PosPoly/Skripts.htm>, Universität Konstanz (Summer Semester 2010)
- [22] S. Kuhlmann: *Lecture Notes on Real Algebraic Geometry*, <http://www.math.uni-konstanz.de/kuhlmann/Lehre/WS18-19-ReelleAlgGeo/Skripts/Gesamtscript-ReelleAlgGeo1.pdf>, Universität Konstanz (Winter Semester 2018/19)
- [23] E. Kunz: *Introduction to Algebraic Plane Curves*, Birkhäuser Basel (2005)
- [24] J. M. Lee: *Manifolds and Differential Geometry*, American Mathematical Society, Providence, Rhode Island (2009)
- [25] M. Marshall: *Positive Polynomials and Sum of Squares*, 146, Math. Surveys Monographs, AMS (2008)
- [26] P. A. Parrilo: *Semidefinite Programming Relaxations for Semialgebraic Problems*, Mathematical Programming, 96(2, Ser. B):293-320 (2003)
- [27] P. A. Parrilo, B. Sturmfels: *Minimizing Polynomial Functions. Algorithmic and Quantitative Real Algebraic Geometry*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 60:83-99 (2003)
- [28] D. Plaumann: *Lecture Notes on Classical Algebraic Geometry*, Universität Konstanz (Summer Semester 2015)
- [29] B. Reznick: *Blenders*, in Notions of Positivity and the Geometry of Polynomials, Springer (2011), pp. 345-373
- [30] B. Reznick: *On Hilbert's Construction of Positive Polynomials*, arXiv preprint arXiv:0707.2156 (2007)
- [31] R. T. Rockafellar: *Convex Analysis*, Princeton University Press (1970)
- [32] C. Scheiderer: *Skript zur Vorlesung Lineare Algebra II*, Universität Konstanz (2017)
- [33] J. H. Silverman, J. H. Tate: *Rational Points on Elliptic Curves*, Second Edition, Springer (2015)
- [34] M. Ulbrich, S. Ulbrich: *Nichtlineare Optimierung*, Birkhäuser, Springer Basel (2011)
- [35] S. Volkwein: *Skript zur Vorlesung Optimierung*, Universität Konstanz (2019)

INDEX

- active indices, 80
- admissible point, 80
- affine part
 - of projective varieties, 67
- Banach space, 26
- bidual cone, 34
- biform associated to a form, 22
- binomial coefficient, 21
 - generalized, 21
- Catalan number, 61
- Cauchy Schwarz Inequality, 37
 - Generalized, 38
- Cauchy sequence, 26
- Cayley Bacharach Theorem, 69, 71
 - of Chasles, 69
- closedness
 - of a cone, 31
- completeness
 - of a normed vector space, 26
- cone, 31
- conical hull, 31
- convergence
 - of a sequence, 26
- convex
 - combination, 31
 - form, 7
 - function, 5
 - hull, 31
 - set, 31
- dehomogenization, 67
- derivative
 - directional, 6
 - partial, 6
- differentiable
 - partially, 6
- direction, 33
- dual
 - cone, 34
 - space, 27
- equivalence
 - of two norms, 26
- Euclidean topology, 26
- Euler Identity, 10
- extreme
 - direction, 33
 - point, 32
 - ray, 32
- failure, 69
- Fischer inner product, 25
 - convex, 7
 - positive semidefinite (PSD), 5
 - sum of squares (SOS), 5
- gradient, 6, 7
- Hessian, 7
- Hilbert
 - function, 70
 - space, 26
 - Theorem (1888), 8
- homogeneous
 - coordinate ring, 70
 - coordinates, 63
 - part, 70, 73
 - polynomial, 5
- hypersurface, 63
 - degree of, 63
- ideal, 63, 73
 - radical, 65
- inner product, 24
- Jacobian, 6
- Karush Kuhn Tucker conditions, 81
- linear
 - in the outer product, 19
 - in the symmetric outer product, 20
 - part, 65
- local minimum, 80
- Motzkin form, 8, 13
- multilinear function, 17
- nonsingular point, 66
- norm, 24
 - generated by an inner product, 25
 - operator, 27
- optimal constants
 - of Generalized Cauchy Schwarz Inequality, 47
- outer product
 - of vector spaces, 16
 - of vectors, 16
 - symmetric, 17
- permutation, 16
- plane curve
 - affine, 63
 - projective, 63
- polarization identity, 23
- polynomial
 - homogeneous, 5
- positive semidefinite
 - matrix, 7
 - quadratic form, 73

- positive semidefinite
 - form, 5
- projection of a point, 68
- projective space, 62
 - n -dimensional, 62
- quadratic form
 - of a linear functional, 73
- ray, 32
 - spanned by a direction, 33
- real part of a polynomial, 55
- regular point, 80
- Riesz Representation Theorem, 27
- Robinson form, 8, 13, 86
- singular point, 66
- symmetric
 - multilinear function, 18
 - tensor, 16
- tangent space, 65
- tensor, 16
 - associated function, 18
 - associated to a form, 22
 - symmetric, 16
- transversal intersection
 - of affine hypersurfaces, 67
 - of projective hypersurfaces, 68
- vanishing ideal, 64
- variety
 - affine, 63
 - projective, 63