POLYNOMIAL OPTIMIZATION AND THE MOMENT PROBLEM

A Thesis

Submitted to the College of Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

in the

Department of Mathematics and Statistics

University of Saskatchewan

Saskatoon

By

Mehdi Ghasemi

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mathematics and Statistics
University of Saskatchewan,
Saskatoon, Saskatchewan, Canada
S7N 5E6

Abstract

There are a wide variety of mathematical problems in different areas which are classified under the title of *Moment Problem*. We are interested in the moment problem with polynomial data and its relation to real algebra and real algebraic geometry. In this direction, we consider two different variants of moment problem.

The first variant is the global polynomial optimization problem, i.e., finding the minimum of a polynomial $f \in \mathbb{R}[\underline{X}] = \mathbb{R}[X_1, \ldots, X_n]$ on $\mathbb{R}^n$. It is known that this problem is NP-hard. One of the most successful approaches to this problem is semidefinite programming (SDP) [35, 47], which gives a polynomial time method to approximate the largest real number r such that $f(\underline{X}) - r$ is a sum of squares of polynomials. But current implementations of SDP are not able to minimize polynomials in rather small number of variables (> 6) of degree > 8. We make use of a result due to Hurwitz [26], to give a criterion in terms of coefficients of a polynomial to be a sum of squares. Then, using this criterion, we introduce a much faster method to approximate a lower bound, in this case using *geometric programs (GP)*.

The second variant is where we wish to determine when a given multi-sequence of reals comes from a Borel measure supported on a given subset K of $\mathbb{R}^n$. This is the so called *K-Moment Problem*. Using Jacobi's Theorem for archimedean quadratic modules [27], we generalize a result of Berg *et al.* which states that the closure of the cone $\sum \mathbb{R}[\underline{X}]^2$ in a weighted ℓ_1-topology consists of nonnegative polynomials on a hypercube [8, 9]. Then we substitute ℓ_1-topology with a locally convex topology τ, $\sum \mathbb{R}[\underline{X}]^2$ by a cone C and the hypercube by a closed set K, and study the relation between $\overline{C}^{\tau}$ and $\text{Psd}(K)$ to solve the moment problem for a τ-continuous linear functional on $\mathbb{R}[\underline{X}]$. We investigate the moment problem for weighted ℓ_p-continuous positive semidefinite (PSD) functionals and coefficientwise convergence topology on $\mathbb{R}[\underline{X}]$. Then, we fix a set K, and find an appropriate locally convex topology τ, such that $\sum \mathbb{R}[\underline{X}]^{2d}$, solves the K-moment problem for τ-continuous linear functionals. In other words, we prove that a τ-continuous linear functional nonnegative on $\sum \mathbb{R}[\underline{X}]^{2d}$ is representable by a Borel measure on K.

Acknowledgements

I would like to thank my supervisors Professors Salma Kuhlmann and Murray Marshall for their encouragement, support and insights patiently they gave me into the area. At the same time they left me free enough to steer my own ideas and leading me on the right direction. I am also grateful for many interesting and helpful discussions with my colleagues, specially Mahmood Alaghmandan and Professor Ebrahim Samei at University of Saskatchewan. I also would like to thank the University of Saskatchewan for financial support during my program. No need to say that preparing this thesis was not possible without patience and endless support of my wife, Bahareh.

Mehdi Ghasemi

Saskatoon, August 2012

To my wife, Bahareh

Contents

List of Tables

Introduction

Hilbert's problems are a set of (originally) unsolved problems in mathematics proposed by Hilbert at the beginning of the 20^{th} century. Hilbert's problems were designed to serve as examples for kinds of problems whose solutions would lead to the development of new disciplines in mathematics [1]. The 17^{th} problem, in its simplest form is as follows:

> *"Given a multivariate polynomial that takes only non-negative values over the reals, can it be represented as a sum of squares of rational functions?"*

This was solved in the affirmative, in 1927, by Emil Artin [2]. A constructive algorithm was found by Delzell in 1984 [12].

The interest of Hilbert in this problem dates back to Minkowski's thesis defence in 1885, where Minkowski made the following assertion:

> *"It is not likely that every nonnegative form can be represented by a sum of squares of forms."*

Hilbert was suspicious about the validity of the above assertion (see [10, Section 6.6] for historical notes). His later works resulted in a complete (non-constructive) characterization of the problem, in terms of number of variables, n, and the degree of the polynomial, d. In 1888, Hilbert proved that *a globally nonnegative real homogeneous polynomial in n variables of degree d is a sum of squares (SOS) if and only if $(n \leq 2)$ or $(d = 2)$ or $(n = 3$ and $d = 4)$* (see [25]).

Hilbert's proof is non-constructive and uses nontrivial facts from the theory of complex projective algebraic curves. The first explicit example of a globally nonnegative real polynomial which is not a sum of squares of polynomials was given in 1967 by Motzkin [46], using the inequality

$$\left(\prod_{i=1}^{m} a_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^{m} a_i,$$

[1]The list consists of 23 problems. Hilbert originally included 24 problems on his list, but decided against including one of them in the published list.

for a list of nonnegative real numbers $a_1 \dots, a_m$ known as arithmetic-geometric mean inequality:

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

It is known that deciding when a polynomial is globally nonnegative is NP-hard [5, Theorem 1.1]. But the method of semidefinite programming (SDP) gives a polynomial time decision procedure to determine whether a given polynomial is a sum of squares or not [35, 47]. The significant difference in the computational complexity of deciding nonnegativity and SOSness, also the similarity of them at the same time is enough to motivate a comprehensive study of the representablity of nonnegative polynomials by sums of squares. Such a result which relates nonnegativity of a polynomial to its representations is usually called a *Positivstellensatz*.

Let $\mathbb{R}[\underline{X}] = \mathbb{R}[X_1, \dots, X_n]$ be the ring of polynomials in n variables $X_1, \dots, X_n$. We denote the set of all finite sums of squares of polynomials by $\sum \mathbb{R}[\underline{X}]^2$. Let $S = \{g_1, \dots, g_s\}$ be a finite set of polynomials in $\mathbb{R}[\underline{X}]$. We consider three objects associated to S:

- The *basic closed semialgebraic set* $\mathcal{K}_S$ associated to S, i.e.,

$$\mathcal{K}_S := \{x \in \mathbb{R}^n \ : \ g_i(x) \geq 0, \ i = 1, \dots, n\};$$

- The *quadratic module* M_S of $\mathbb{R}[\underline{X}]$ generated by S, i.e.,

$$M_S := \left\{ \sum_{i=0}^{s} \sigma_i g_i \ : \ g_0 = 1 \text{ and } \sigma_i \in \sum \mathbb{R}[\underline{X}]^2, \ i = 0, \dots, s \right\},$$

  the set of all SOS combinations of elements of S;

- And the *preordering* T_S of $\mathbb{R}[\underline{X}]$ generated by S, i.e.,

$$T_S := \left\{ \sum_{e \in \{0,1\}^s} \sigma_e \underline{g}^e \ : \ \sigma_e \in \sum \mathbb{R}[\underline{X}]^2, \quad \forall e \in \{0,1\}^s \right\},$$

  where for $e = (e_1, \dots, e_s)$, $\underline{g}^e = \prod_{i=1}^{s} g_i^{e_i}$.

The quadratic module M_S (or the preordering T_S) is said to be *Archimedean* if for every $f \in \mathbb{R}[\underline{X}]$ there exists an integer $N \geq 1$ such that $N \pm f \in M_S$ (or $N \pm f \in T_S$ respectively). The earliest Positivstellensatz was discovered first by Krivine [32] in 1964 and again by Stengle [61] in 1974:

Theorem. *Suppose that* $S \subseteq \mathbb{R}[\underline{X}]$ *is finite and* $f \in \mathbb{R}[\underline{X}]$*. Then* $f \geq 0$ *on* $\mathcal{K}_S$ *if and only if there exists an integer* $m \geq 0$ *and* $p, q \in T_S$ *such that* $pf = f^{2m} + q$*. If* $f > 0$ *on* $\mathcal{K}_S$*, then* m *can be taken to be 0.*

Note that the solution to Hilbert's 17^{th} problem is a corollary of this theorem by taking $S = \varnothing$ and $f = (\frac{1}{p})^2 p(f^{2m} + q) \in \sum \mathbb{R}(\underline{X})^2$ if $f \neq 0$.

In 1991, Schmüdgen proved a rather surprising version of Krivine's Positivstellensatz for compact semialgebraic sets [58] which had a great impact in the real algebraic geometry community and resulted into a series of "Strikt Positivstellensätze". He proved that for a finite S, if $\mathcal{K}_S$ is compact then $f > 0$ on $\mathcal{K}_S$ implies $f \in T_S$. This result was extended by Putinar in 1993 to Archimedean quadratic modules [51] : If M_S is Archimedean and $f > 0$ on $\mathcal{K}_S$ then $f \in M_S$. The compactness of $\mathcal{K}_S$ in this case is a consequence of M_S being Archimedean. In other words, if $f > 0$ on $\mathcal{K}_S$ then f has a denominator-free representation in terms of elements of S and hence the SDP technique can be used to exploit this representation in a reasonable time. Even further improvements were found later by Jacobi in 2001 [27] (see Corollary 1.3.11) and Marshall in 2002 [41] (see Theorem 1.3.10). Note that for a polynomial $f \in \mathbb{R}[\underline{X}]$, if $f \geq 0$ on $\mathcal{K}_S$, then $f + \epsilon > 0$ on $\mathcal{K}_S$ for all $\epsilon > 0$ and hence all of the Strikt Positivstellensätze can be applied to $f + \epsilon$. This simple observation shows that if $f \geq 0$ on $\mathcal{K}_S$, then f can be approximated by elements of T_S and/or M_S, by making small changes in the constant term of f.

One can summarize the main points of the previous paragraph as *SOS representations* and *SOS approximations* of nonnegative polynomials on basic closed semialgebraic sets. This thesis consists of two parts dealing with these two aspects in a more general framework of commutative rings and $\mathbb{R}$-algebras. Chapter 3 describes our contribution on SOS representations of nonnegative even-degree polynomials and some applications to optimization. Chapters 4, 5 and 6 are dedicated to SOS approximations and their applications to the *moment problem*.

In Chapter 1 we introduce some basic notations in real algebra. Most of the proofs are omitted but suitable references are provided to indicate the chronological order of advancements in the area. We begin by introducing preorderings and quadratic modules. Then we explain when each of these rather smaller objects, can be extended to obtain an ordering. We define the real spectrum of a commutative ring A as the space of orderings of A and we explain how it can be equipped with spectral and patch topologies. Then we explain the relation between orderings and ring homomorphisms into real closed fields. For the rest of the thesis we stick to a specific subspace $\mathcal{X}_A$ of the real spectrum, corresponding to $\mathbb{R}$-valued ring homomorphisms. We close the chapter by stating a couple of Positivstellensätze such as Representation Theorem, Jacobi's Representation and Schmüdgen's Strikt Positivstellensatz, which will be used later.

We focus on the *moment problem* in Chapter 2. The term "moment problem" first appeared in the works of Stieljes in 1894. However, some related results can be traced back to earlier works of Chebyshef in 1873. See [1] for an extensive reading on history and applications of the classical moment problem. In Section 2.1, we recall the definitions of (real) topological vector spaces, (real) topological algebras, locally convex topologies, locally multiplicatively convex topologies (*lmc* for short), seminorms and norms. The dual of a topological vector space (V,τ), denoted by $(V,\tau)^*$, is the set of all τ-continuous linear functionals $L : V \longrightarrow \mathbb{R}$. In Section 2.2, we introduce the connection between materials from Chapter 1 and the solution of the moment problem. Assuming $V = \mathbb{R}[\underline{X}]$, the moment problem is the question of when a given linear functional L on $\mathbb{R}[\underline{X}]$ is representable as an integration with respect to a Borel measure μ on a closed subset K of $\mathbb{R}^n$, i.e.,

$$\forall f \in \mathbb{R}[\underline{X}] \quad L(f) = \int_K f \, d\mu. \tag{1}$$

Denote the set of nonnegative polynomials on K by $\mathrm{Psd}(K)$. Then an obvious necessary condition for existence of such a representation (1) is that $L(f) \geq 0$ for all $f \in \mathrm{Psd}(K)$. Haviland in 1935 proved that this necessary condition is also sufficient [23, 24]. The moment problem in its whole generality is difficult to solve. In fact, in 1999 Scheiderer proved that the preordering $\mathrm{Psd}(K)$ is seldom finitely generated [56, Proposition 6.1]. So in general, there is no practical decision procedure for the membership problem for $\mathrm{Psd}(K)$, and hence for $L(\mathrm{Psd}(K)) \subseteq \mathbb{R}^+$. According to Schmüdgen's Positivstellensatz, when S is finite and $\mathcal{K}_S$ is compact, every $f \in \mathrm{Psd}(\mathcal{K}_S)$ has an SOS approximation in terms of elements of S as accurate as we need; i.e., if $f \in \mathrm{Psd}(\mathcal{K}_S)$ then for every $\epsilon > 0$, $f + \epsilon \in T_S$. Therefore, if a functional L is nonnegative on T_S, then the equality $L(f) = \lim_{\epsilon \to 0} L(f + \epsilon)$ shows that it is also nonnegative on $\mathrm{Psd}(\mathcal{K}_S)$. In general $\mathrm{Psd}(\mathcal{K}_S)$ is not finitely generated, but if it is then it is possible to determine if $L(\mathrm{Psd}(\mathcal{K}_S)) \subseteq \mathbb{R}^+$ is true and hence Haviland's result is applicable. We continue with several well-known results on the moment problem for basic closed semialgebraic sets and also continuous linear functionals.

Chapter 3 is devoted to SOS representations. We present a new proof for a classical result of Hurwitz and Reznick [26, 53] and a consequence of this result due to Fidalgo and Kovacec [15]. We apply these results to determine a new sufficient condition for a polynomial to be a SOS, which generalizes the results of [15] and [36]. Then, in Section 3.2, we explain how this condition can be used to determine a lower bound for an even degree polynomial, bounded below. We assign a *geometric program (GP)* (Definition 3.2.5) to any such polynomial and the optimum value of this program, if it exists, results in a lower bound for the polynomial. Although the lower bound found by this method is typically not as good as the lower bound found using SDP, a practical comparison confirms that the computation is faster, and larger problems can be handled with this approach.

In Chapter 4 we return to the moment problem from a different point of view. In Chapter 2, we saw that under certain algebraic conditions on the set K, every linear functional which satisfies some nonnegativity constraints admits a representation as an integral over K. In this chapter we relax the conditions on K and assume the extra constraint of continuity of the linear functionals with respect to a family of seminormed topologies on $\mathbb{R}[\underline{X}]$. In [7] Berg, Christensen and Ressel prove that the closure of the cone $\sum\mathbb{R}[\underline{X}]^2$ in the polynomial ring $\mathbb{R}[\underline{X}]$ in the topology induced by the ℓ_1-norm is equal to $\operatorname{Psd}([-1,1]^n)$. The result is deduced as a corollary of a general result, also established in [7], which is valid for any commutative semigroup. In later work Berg and Maserick [9] and Berg, Christensen and Ressel [8] establish an even more general result for a commutative semigroup with involution, for the closure of the cone of sums of squares of symmetric elements in the weighted ℓ_1-seminorm topology associated to an absolute value. We give new proofs of these results which are based on Jacobi's representation theorem [27]. At the same time, we use Jacobi's representation theorem 1.3.10 to extend these results from sums of squares to sums of $2d$-powers. In particular we show that for any integer $d \geq 1$, the closure of the cone of sums of $2d$-powers $\sum\mathbb{R}[\underline{X}]^{2d}$ in $\mathbb{R}[\underline{X}]$ in the topology induced by the ℓ_1-norm is equal to $\operatorname{Psd}([-1,1]^n)$.

In Chapter 5, we take a locally convex topology τ on $\mathbb{R}[\underline{X}]$, a closed set $K \subseteq \mathbb{R}^n$, and a nonempty cone $C \subseteq \mathbb{R}[\underline{X}]$. We show that the inclusion

$$\operatorname{Psd}(K) \subseteq \overline{C}^{\tau} \tag{2}$$

implies that every τ-continuous linear functional L satisfying $L(C) \subseteq \mathbb{R}^+$ can be represented as an integration with respect to a Borel measure on K, and vice versa. We explain that the version of moment problem introduced in Chapter 2 is the special case of (2), where $C = T_S$ or $C = M_S$, $K = \mathcal{K}_S$ and the topology τ is the finest locally convex topology φ. After that, we fix the cone $C = \sum \mathbb{R}[\underline{X}]^{2d}$ and solve the equation $\mathrm{Psd}(K) = \overline{C}^{\tau}$ for K, where τ is a weighted ℓ_p-norm. Regarding the inclusion (2), for the special case of Chapter 2 the topology is assumed to be φ, and both K and C are related to each other by taking $K = \mathcal{K}_S$ and $C = T_S$ or $C = M_S$. In chapters 4 and 5, we fixed $C = \sum \mathbb{R}[\underline{X}]^{2d}$ and solved (2) in terms of K for various locally convex topologies.

In Chapter 6 by fixing $C = \sum \mathbb{R}[\underline{X}]^{2d}$, for any closed set $K \subseteq \mathbb{R}^n$ we find a locally convex topology $\mathcal{T}_K$, such that $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\mathcal{T}_K} = \mathrm{Psd}(K)$. When S is compact, we show that the sup-norm on K denoted by $\|\cdot\|_K$, induces another topology, strictly finer than $\mathcal{T}_K$, but we still have $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_K} = \mathrm{Psd}(K)$. At the end of the Chapter 6, we make a comparison among all topologies we studied throughout.

Appendix A is a review of the model theory of real closed fields, Tarski's Transfer Principal, and a concrete version of the Abstract Positivstellensatz.

Appendix B connects the two different parts of the thesis. The *generalized moment problem* (GMP for short) is defined as an optimization problem on the set of positive Borel measures $\mathcal{M}_+(K)$ on a locally compact topological space K. It is shown that the K-moment problem and polynomial optimization are variants of the GMP.

Appendix C contains the source code of implementing the computational method introduces in Chapter 3 in SAGE.

Chapter 1

Preliminaries

We always denote by $\mathbb{N}$ the set of nonnegative integers $\{0,1,2,\ldots\}$, $\mathbb{Z}$ the set of integers, $\mathbb{Q}$ the set of rationals, $\mathbb{R}$ the set of reals and $\mathbb{C}$ the set of complex numbers. Throughout A denotes a ring, which is assumed always to be commutative with 1. All ring homomorphisms are considered to be unitary. The ring of polynomials in $X_1,\ldots,X_n$, with coefficients in A, i.e. $A[X_1,\ldots,X_n]$ will be denoted by $A[\underline{X}]$ for short. For $\underline{X}=(X_1,\ldots,X_n)$, $\underline{a}=(a_1,\ldots,a_n)\in A^n$ and $\alpha=(\alpha_1,\ldots,\alpha_n)\in\mathbb{N}^n$, we define $\underline{X}^\alpha := X_1^{\alpha_1}\cdots X_n^{\alpha_n}$, $|\alpha| := \alpha_1+\cdots+\alpha_n$ and $\underline{a}^\alpha := a_1^{\alpha_1}\cdots a_n^{\alpha_n}$ with the convention $0^0=1$. We usually assume that $\frac{1}{2}\in A$ and hence A is an $\mathbb{Z}[\frac{1}{2}]$-algebra, unless otherwise is specified.

1.1 Orderings, Preorderings and Quadratic Modules

For a ring A, we denote by $\sum A^2$, the set of all finite sums of squares of elements of A, i.e.

$$\sum A^2 := \{\sum_{i=1}^m a_i^2 \;:\; m\in\mathbb{N} \text{ and } a_1,\ldots,a_m\in A\}.$$

Definition 1.1.1. A *partially ordered ring* $(A,+,\cdot,\le)$ is a ring $(A,+,\cdot)$ together with a transitive binary relation $\le$ on A, satisfying

1. $\forall x,y\ (x\le y \to (\forall z\ x+z\le y+z))$,
2. $\forall x,y\ (0\le x, 0\le y \to 0\le x\cdot y)$,
3. $\forall x\ (0\le x^2)$.

A subset T of A is called a *preordering* if

$$T+T\subseteq T,\ \sum A^2\subseteq T,\ T\cdot T\subseteq T.$$

Note that for a partially ordered ring $(A,+,\cdot,\le)$ the set $A^+=\{a\in A : 0\le a\}$ forms a preordering of A. We also note that the converse is true.

Proposition 1.1.2. *Suppose that T is a preordering of A. A together with the binary relation on A defined by $a \le_T b \Leftrightarrow b - a \in T$ is a partially ordered ring. Moreover with respect to this order $A^+ = \{a \in A : 0 \le_T a\} = T$.*

Definition 1.1.3. $M \subseteq A$ is called a *quadratic module* if $M + M \subseteq M$, $\sum A^2 \cdot M \subseteq M$ and $1 \in M$. An *ordering* P of A is a preordering satisfying $P \cup -P = A$ and $P \cap -P$ is a prime ideal of A.

The following is straightforward and useful.

Proposition 1.1.4. *Suppose M is a quadratic module of A, and $\frac{1}{2} \in A$ then*

1. *$M \cap -M$ is an ideal of A.*

2. *$-1 \in M$ if and only if $M = A$.*

Proof. 1. Let $I = M \cap -M$. Clearly

$$I + I \subseteq I, \ \ - I = I, \ 0 \in I, \ A^2 I \subseteq I.$$

Using the identity $a = (\frac{a+1}{2})^2 - (\frac{a-1}{2})^2$ this yields also that $AI \subseteq I$ and hence $M \cap -M$ is an ideal of A.

2. If $-1 \in M$ then $1 \in M \cap -M$. Since $M \cap -M$ is an ideal of A, this implies that $M \cap -M = A$ and hence $M = A$. □

A quadratic module (preordering) M is said to be *proper*, if $-1 \notin M$. The ideal $M \cap -M$ is called the *support* of M and is denoted by $\operatorname{supp} M$.

If $\phi : A \longrightarrow B$ is a ring homomorphism and N is a quadratic module of B, then $\phi^{-1}(N)$ is a quadratic module of A called the *contraction* of N to A. If M is a quadratic module of A, then $\sum B^2\phi(M)$, the set of all finite sums $\sum b_i^2 \phi(s_i)$, $b_i \in B$, $s_i \in M$, is a quadratic module of B called the *extension* of M to B.

Proposition 1.1.5. *Suppose that M is a quadratic module of A and S is a multiplicatively closed subset of A such that $(M \cap -M) \cap S = \emptyset$. Then the extension of M to $S^{-1}A$ is proper.*

Proof. Elements of the extension of M to $S^{-1}A$ have the form $\sum_i (\frac{a_i}{b_i})^2 s_i$ where $a_i \in A$, $b_i \in S$, and $s_i \in M$. Since $\sum_i (\frac{a_i}{b_i})^2 s_i = \frac{s}{b^2}$ where $b = \prod_i b_i$ and $s = \sum_i (a_i \prod_{j \neq i} b_j)^2 s_i$, every element of the extension has the form $\frac{s}{b^2}$, $s \in M$, $b \in S$. If $-1 = \frac{s}{b^2}$ for some $s \in M$ and

$b \in S$, then $-b^2 t = st$ for some $t \in S$, hence $-b^2 t^2 = st^2 \in M$. Therefore $-b^2 t^2 \in (M \cap -M) \cap S$ which is a contradiction. $\square$

The following proposition shows that every preordering is extensible to an ordering. This result is originally due to Krivine [32], but the proof given here is due to Prestel [50].

Proposition 1.1.6. *For any preordering T of A with $-1 \notin T$, there exists an ordering P of A containing T.*

Proof. Let P be a preordering of A containing T and maximal subject to $-1 \notin P$. Such a preordering P exists by Zorn's Lemma. We want to show that P is an ordering. We make use of the fact that for any $a \in A$, $P + aP$ is a preordering of A, $P \subseteq P + aP$ and $a \in P + aP$.

If $a \in A \smallsetminus (P \cup -P)$, then $-1 \in P + aP$, $-1 \in P - aP$. Thus

$$-1 = s_1 + at_1, \ -1 = s_2 - at_2, \ s_i, t_i \in P, \ i = 1, 2.$$

Then $-at_1 = 1 + s_1$, $at_2 = 1 + s_2$, so

$$-a^2 t_1 t_2 = (1 + s_1)(1 + s_2) = 1 + s_1 + s_2 + s_1 s_2.$$

Thus $-1 = s_1 + s_2 + s_1 s_2 + a^2 t_1 t_2 \in P$, a contradiction. This proves $P \cup -P = A$.

Let $\mathfrak{p} = P \cap -P$. Clearly $0 \in \mathfrak{p}$, $\mathfrak{p} + \mathfrak{p} \subseteq \mathfrak{p}$, $-\mathfrak{p} = \mathfrak{p}$ and $P\mathfrak{p} \subseteq \mathfrak{p}$. Then

$$A\mathfrak{p} = (P \cup -P)\mathfrak{p} \subseteq P\mathfrak{p} \cup -P\mathfrak{p} \subseteq \mathfrak{p}.$$

This proves $\mathfrak{p}$ is an ideal. Suppose $ab \in \mathfrak{p}$, $a, b \notin \mathfrak{p}$. Replacing a, b by $-a, -b$ if necessary, we can assume $a, b \notin P$. Thus $-1 \in P + aP$, $-1 \in P + bP$, so

$$-1 = s_1 + at_1, \ -1 = s_2 + bt_2, \ s_i, t_i \in P, \ i = 1, 2.$$

Then

$$abt_1 t_2 = (-t_1 a)(-t_2 b) = (1 + s_1)(1 + s_2) = 1 + s_1 + s_2 + s_1 s_2,$$

so $-1 = s_1 + s_2 + s_1 s_2 - abt_1 t_2$. Since $ab \in \mathfrak{p}$ and $\mathfrak{p}$ is an ideal, this is an element of P, a contradiction. This proves $\mathfrak{p}$ is prime and hence P is an ordering of A. $\square$

Definition 1.1.7. A commutative ring A is called *real* if $a_1^2 + \cdots + a_n^2 = 0$ implies that $a_i = 0$ for each $i = 1, \ldots, n$. A field F is called *formally real* if, as a ring, F is real, i.e., $-1 \notin \sum F^2$.

Remark 1.1.8. $\sum A^2$ is the smallest preordering of A. If A is a real ring then $-1 \notin \sum A^2$. Applying Proposition 1.1.6 to A, we deduce that any real ring admits an ordering and consequently, every formally real field admits an ordering.

1.2 Real Spectrum and Positivstellensatz

Let A be a real ring. According to Proposition 1.1.6, A admits at least one nontrivial ordering. In this section, we develop some basics to study the structure of a ring A, by use of all possible orderings on A.

Definition 1.2.1. The set of all orderings of A is denoted by $\mathrm{Sper}(A)$ and is called the *real spectrum* of A. For a preordering T of A, we denote the set of all orderings of A containing T by $\mathrm{Sper}_T(A)$. For $a \in A$, define the set

$$U(a) := \{P \in \mathrm{Sper}(A) \ : \ a \notin P\}.$$

The family $\{U(a) : a \in A\}$ forms a subbasis for a topology on $\mathrm{Sper}(A)$ called the *spectral topology*. The family $\{U(a), U(a)^c : a \in A\}$, again, forms a subbasis for a topology called the *patch topology*. Here, $U(a)^c$ denotes the complement of $U(a)$ in $\mathrm{Sper}(A)$, i.e.,

$$U(a)^c = \mathrm{Sper}(A) \smallsetminus U(a).$$

Remark 1.2.2. For a ring A, the following data are equivalent:

1. An ordering P of the ring A,

2. A couple $(\mathfrak{p}, \leq)$, where $\mathfrak{p}$ is a prime ideal of A, and $\leq$ is an ordering of the field of fractions of $\frac{A}{\mathfrak{p}}$, denoted by $\mathrm{ff}\frac{A}{\mathfrak{p}}$.

3. An equivalence class of ring homomorphisms $\alpha : A \longrightarrow R$ where R is a real closed field (see A.1 for definition), for the smallest equivalence relation, such that α and α' are equivalent if there exists a commutative diagram of ring homomorphisms:

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[dr, "\alpha'"'] & R \arrow[d] \\ & R' \end{tikzcd}$$

One goes from (1) to (2) by taking $(\mathfrak{p}, \leq) = (\text{supp } P, \leq)$, where $\leq$ is the ordering on $\text{ff}\frac{A}{\text{supp } P}$ induced by the ordering P on A, from (2) to (3) by taking $\alpha : A \longrightarrow \text{ff}\frac{A}{\mathfrak{p}} \longrightarrow R$, where R is the real closure of $(\text{ff}\frac{A}{\mathfrak{p}}, \leq)$, and from (3) to (1) by taking $P = \alpha^{-1}(R^+)$ ([10, Proposition 7.1.2]).

Theorem 1.2.3. *Sper(A) is compact in the patch topology and hence in the spectral topology.*

Proof. See [44, Theorem 2.4.1]. $\square$

For a given ordering P of A and an element $a \in A$, we define the sign of a with respect to P by

$$\text{Sgn}_P(a) := \begin{cases} 1 & a \notin -P, \\ 0 & a \in P \cap -P, \\ -1 & a \notin P. \end{cases}$$

So $U(a) = \{P \in \text{Sper}(A) : \text{Sgn}_P(a) = -1\}$.

Theorem 1.2.4 (Abstract Positivstellensatz)**.** *Let T be a preordering of A and let $a \in A$. Then*

1. $\forall P \in \text{Sper}_T(A)\ \text{Sgn}_P(a) > 0 \iff \exists p, q \in T\ (pa = 1 + q)$.

2. $\forall P \in \text{Sper}_T(A)\ \text{Sgn}_P(a) \geq 0 \iff \exists p, q \in T\ \exists m \in \mathbb{N}\ (pa = a^{2m} + q)$.

3. $\forall P \in \text{Sper}_T(A)\ \text{Sgn}_P(a) = 0 \iff \exists m \in \mathbb{N}\ (-a^{2m} \in T)$.

Proof. See [44, Theorem 2.5.2]. $\square$

By Proposition 1.1.6, if $-1 \notin T$ then $\text{Sper}_T(A) \neq \varnothing$ and clearly

$$\tilde{T} = \bigcap_{P \in \text{Sper}_T(A)} P$$

is a preordering of A containing T. We refer to $\tilde{T}$ as the *saturation* of T. We say a preordering T is *saturated* if $\tilde{T} = T$. It is easy to show that a preordering T is saturated if and only if it is the intersection of a family of orderings.

We now study a specific subspace of Sper(A).

Definition 1.2.5. The set of all real valued ring homomorphisms will be denoted by $\mathcal{X}_A$, i.e. $\mathcal{X}_A = \text{Hom}(A, \mathbb{R})$.

Remark 1.2.6. It is well-known that the identity map is the only ring homomorphism from $\mathbb{R}$ to $\mathbb{R}$ [44, Proposition 5.4.5]. If we take A to be an $\mathbb{R}$-algebra, then for every ring homomorphism $\alpha : A \longrightarrow \mathbb{R}$, $r \in \mathbb{R}$ and $a \in A$, $\alpha(ra) = \alpha(r)\alpha(a) = r\alpha(a)$ which implies that α is also an $\mathbb{R}$-algebra homomorphism.

To every element $a \in A$, we associate a function $\hat{a} : \mathcal{X}_A \longrightarrow \mathbb{R}$, defined by $\hat{a}(\alpha) = \alpha(a)$. The family of the sets of form

$$U(\hat{a}) = \{\alpha \in \mathcal{X}_A : \hat{a}(\alpha) < 0\},$$

forms a subbasis for the topology on $\mathcal{X}_A$ inherited from the product topology on $\mathbb{R}^A$. With this topology $\mathcal{X}_A$ is a Hausdorff space. By Remark 1.2.2, the map $\Theta : \mathcal{X}_A \longrightarrow \operatorname{Sper}(A)$, defined by $\Theta(\alpha) = \alpha^{-1}(\mathbb{R}^+)$ is obviously a topological embedding, giving $\operatorname{Sper}(A)$ the spectral topology. So $\mathcal{X}_A$ can be considered as a subspace of $\operatorname{Sper}(A)$ with the spectral topology.

Example 1.2.7. Let $A = \mathbb{R}[\underline{X}]$, the ring of polynomials on n variables, $X_1, \ldots, X_n$. Every ring homomorphism $\alpha \in \operatorname{Hom}(\mathbb{R}[\underline{X}], \mathbb{R})$ is completely determined by its values on each X_i, $i = 1, \ldots, X_n$. Conversely, for every choice of real numbers $(x_1, \ldots, x_n) \in \mathbb{R}^n$, the map defined by $X_i \mapsto x_i$, determines a ring homomorphism on $\mathbb{R}[\underline{X}]$ to $\mathbb{R}$. So in this case $\mathcal{X}_{\mathbb{R}[\underline{X}]} = \operatorname{Hom}(\mathbb{R}[\underline{X}], \mathbb{R})$ can be identified by $\mathbb{R}^n$.

Lemma 1.2.8. *$\mathcal{X}_A$ is a closed subset of $\mathbb{R}^A$. If $A = \mathbb{R}[\underline{X}]$, then $\mathcal{X}_A = \mathbb{R}^n$ as a topological space.*

Proof. See [44, Lemma 5.7.1]. □

Proposition 1.2.9. *Let K be a nonempty subset of $\mathcal{X}_A$ and let $\mathrm{C}(K)$ denote the algebra of continuous real valued functions on K, then*

1. *The map $\Phi : A \longrightarrow \mathrm{C}(K)$ defined by $\Phi(a) = \hat{a}|_K$ is a ring homomorphism.*

2. *$Im(\Phi)$ contains a copy of $\mathbb{Z}[\frac{1}{2}]$.*

Proof. (1) This is clear. Let $a, b \in A$, for each $\alpha \in K$ we have

$$\begin{aligned}\Phi(a+b)(\alpha) &= \widehat{(a+b)}|_K(\alpha) \\ &= \alpha(a+b) \\ &= \alpha(a)+\alpha(b) \\ &= \hat{a}|_K(\alpha)+\hat{b}|_K(\alpha) \\ &= \Phi(a)(\alpha)+\Phi(b)(\alpha).\end{aligned}$$

Similarly $\Phi(1) = 1$ and $\Phi(a \cdot b) = \Phi(a) \cdot \Phi(b)$.

(2) Since $\mathcal{X}_A$ consists of unitary homomorphisms, $\hat{1}(\alpha) = \alpha(1) = 1$, so the constant function $1 \in \Phi(A)$. Moreover for each $m \in \mathbb{Z}$ and $n \in \mathbb{N}$, $\frac{m}{2^n} \in A$ and $\widehat{\frac{m}{2^n}}$ is the constant function $\frac{m}{2^n}$ which belongs to $Im(\Phi)$, so $\mathbb{Z}[\frac{1}{2}] \subseteq Im(\Phi)$. □

Note that in general, Φ is not an embedding. A necessary and sufficient condition on K which forces Φ to be injective is given in Theorem 1.2.11.

Definition 1.2.10. To any subset S of A we associate a subset $\mathcal{Z}(S)$ of $\mathcal{X}_A$, called the *zeros of* S or the variety of S by $\mathcal{Z}(S) := \{\alpha \in \mathcal{X}_A : \forall s \in S\ \hat{s}(\alpha) = 0\}$.

Let us denote the ideal in A generated by S with $\langle S \rangle$, then it is easy to check that $\mathcal{Z}(S) = \mathcal{Z}(\langle S \rangle)$, and the family $\{\mathcal{X}_A \setminus \mathcal{Z}(S) : S \subseteq A\}$ are the open sets for a topology on $\mathcal{X}_A$ called the *Zariski topology* on $\mathcal{X}_A$. It is easy to check that the sets

$$D(a) := \{\alpha \in \mathcal{X}_A \ : \ \hat{a}(\alpha) \neq 0\},$$

for $a \in A$ form a basis for the Zariski topology.

Theorem 1.2.11. *Assume that the map* $\hat{}: A \longrightarrow \mathrm{C}(\mathcal{X}_A)$ *is injective and* K *is a subset of* $\mathcal{X}_A$*. Then the ring homomorphism* $\Phi : A \longrightarrow \mathrm{C}(K)$ *defined by* $\Phi(a) = \hat{a}|_K$ *is injective if and only if* K *is dense in Zariski topology.*

Such a subset $K \subseteq \mathcal{X}_A$ is usually called *Zariski dense* in $\mathcal{X}_A$

Proof of 1.2.11. Note that K is not a dense subset of $\mathcal{X}_A$ in Zariski topology if and only if there exists an element $a \in A$ such that $D(a) \neq \varnothing$ and $D(a) \cap K = \varnothing$. Equivalently, there exists $a \in A$ such that $\hat{a} \neq 0$ and so $a \neq 0$ but $\hat{a}|_K = 0$, i.e., $\Phi(a) = 0$. Hence $\ker \Phi \neq \{0\}$ which means Φ is not injective. □

1.3 Representation Theorem

In the previous section we developed a method to switch from an abstract ring to the ring of continuous real valued functions. This approach reveals several facts about the positivity and positive cones on A.

Definition 1.3.1. By a *presemiprime* M in A we mean a subset M of A such that $0, 1 \in M$ and $M + M \subseteq M$.

A *preprime* T in A is a presemiprime which is closed under multiplication, i.e. $0, 1 \in T$, $T + T \subseteq T$ and $T \cdot T \subseteq T$.

T is said to be *generating* if for each $a \in A$ there exists an integer $k \geq 1$ such that $ka \in T - T$.

A presemiprime M is said to be *archimedean*, if for each $a \in A$ there exist integers $n, k \geq 1$ such that $k(n + a) \in M$.

Let T be a preprime and M be a presemiprime in A. M is said to be a T-*module* if $T \cdot M \subseteq M$.

For an integer $d \geq 1$, we denote by $\sum A^d$, the preprime consists of sums of dth powers of elements of A, i.e.

$$\sum A^d := \{\sum_{i=1}^{m} a_i^d \ : \ m \in \mathbb{N} \text{ and } a_1, \ldots, a_m \in A\}.$$

A preprime T which is also a $\sum A^d$-module is called a *preordering of exponent* d.

One can easily see that a $\sum A^2$-module is simply a quadratic module and a preordering of exponent 2 is simply a preordering.

Theorem 1.3.2. *Suppose T is a generating preprime of A and $Q \subseteq A$ is a T-module maximal subject to $-1 \notin Q$. If Q is archimedean, then $Q = \alpha^{-1}(\mathbb{R}^+)$ for some unique unitary ring homomorphism $\alpha : A \longrightarrow \mathbb{R}$.*

Proof. See [44, Theorem 5.2.5]. □

Definition 1.3.3. For any subset S of A, set

$$\mathcal{K}_S := \{\alpha \in \mathcal{X}_A \ : \ \alpha(S) \subseteq \mathbb{R}^+\}.$$

Also for $K \subseteq \mathcal{X}_A$, set

$$\mathrm{Psd}(K) := \{a \in A \ : \ \forall \alpha \in K \ \hat{a}(\alpha) \geq 0\}.$$

Remark 1.3.4. For any $K \subseteq \mathcal{X}_A$, $\operatorname{Psd}(K) = \bigcap_{\alpha \in K} \alpha^{-1}(\mathbb{R}^+)$ which is an intersection of orderings of A. Therefore $\operatorname{Psd}(K)$ is always saturated.

Corollary 1.3.5. *Suppose $M \subseteq A$ is an archimedean T-module, T a generating preprime of A. Then the following are equivalent:*

1. $\mathcal{K}_M \neq \varnothing$.

2. $-1 \notin M$.

Proof. See [44, Corollary 5.4.1]. □

Proposition 1.3.6. *If M is an archimedean presemiprime, then $\mathcal{K}_M$ is compact.*

Proof. This is well known. $\mathcal{K}_M$ is closed because it is intersection of closed sets

$$\mathcal{K}_M = \bigcap_{a \in M} \hat{a}^{-1}(\mathbb{R}^+).$$

Moreover for each $a \in A$ there exist $k, n_a \geq 1$ such that $k(n_a \pm a) \in M$, so $|\hat{a}(\alpha)| \leq n_a$ for all $\alpha \in \mathcal{K}_M$. Therefore

$$\mathcal{K}_M \subseteq \prod_{a \in A} [-n_a, n_a],$$

which is compact by Tychonoff's Theorem. □

Definition 1.3.7. A preprime T is said to be *torsion* if for each $a \in A$, there exists an integer $n \geq 1$ such that $a^n \in T$. T is said to be *weakly torsion* if for each $a \in A$, there exist integers $k, l, m, n \geq 1$ such that $k(l + ma)^n \in T$.

Example 1.3.8.

1. Any archimedean preprime is weakly torsion.

2. Any torsion preprime is weakly torsion.

3. For a compact subset K of $\mathcal{X}_A$ and an integer $n \geq 1$, let T_0 consists of all $\sum a_i^n$, with $\hat{a}_i > 0$ on K together with 0. T_0 is weakly torsion.

Remark 1.3.9.

1. By [41, Lemma 2.4], every weakly torsion preprime is generating.

2. The polynomial identity

$$m!X = \sum_{h=0}^{m-1} (-1)^{m-1-h} \binom{m-1}{h} [(X+h)^m - h^m],$$

shows that any preordering of any exponent is generating(See [21, Page 325].).

3. For any ring homomorphism $\phi : A \longrightarrow B$, if T is a weakly torsion in B then $\phi^{-1}(T)$ is also weakly torsion in A, but this is not true for generating preprimes (See [41, Example 2.2].).

The following theorem is a generalization given by Marshall [41] of a result, usually attributed to Kadison [30] and also Dubois [14]. But other versions of the result were also proved earlier by Krivine [32, 33]. This also sharpens the results of Jacobi and Prestel in [27, 28].

Theorem 1.3.10 (Representation Theorem)**.** *Suppose M is an archimedean T-module, T a weakly torsion preprime in A. Then for each $a \in A$ with $\hat{a} > 0$ on $\mathcal{K}_M$ there exists an integer $k \geq 1$ such that $ka \in M$.*

Proof. See [41, Theorem 2.3]. □

Corollary 1.3.11 (Jacobi's Representation)**.** *Suppose $M \subseteq A$ is an archimedean $\sum A^{2d}$-module for some integer $d \geq 1$. Then, for each $a \in A$,*

$$\hat{a} > 0 \text{ on } \mathcal{K}_M \;\Rightarrow \exists k \geq 1 \; ka \in M.$$

Proof. Combine the Representation Theorem with Example 1.3.8(2) (see [27] for the original proof). □

1.4 Schmüdgen's Positivstellensatz

A finite boolean combination of basic open subsets of $\text{Sper}(A)$ is called a *constructible set*, i.e., every constructible set is a finite union of the sets of the form

$$\{P \in \text{Sper}(A) \;:\; \text{Sgn}_P(a_i) = 1, \text{Sgn}_P(b_j) = -1, \text{Sgn}_P(c_k) = 0\},$$

for a finite number of elements a_i, b_j and c_k of A. The intersection of any constructible subset of $\operatorname{Sper}(A)$ with $\mathcal{X}_A$ is called a *semialgebraic set*. A *basic closed semialgebraic* is a subset K of $\mathcal{X}_A$, for which there exist $a_1, \ldots, a_m \in A$ such that

$$K = \{\alpha \in \mathcal{X}_A \ : \ \hat{a}_i(\alpha) \geq 0, i = 1, \ldots, m\}.$$

Clearly $K = \mathcal{K}_{\{a_1,\ldots,a_m\}}$.

For any subset S of A, we denote by T_S (resp. M_S) the smallest preordering (resp. quadratic module) of A, containing S. We say T (resp. M) is finitely generated, if there exists a finite set $S \subset A$ such that $T = T_S$ (resp. $M = M_S$).

Example 1.4.1. Let $A = \mathbb{R}[\underline{X}]$, then $\mathcal{X}_A = \mathbb{R}^n$ and for any finite number of polynomials $f_1, \ldots, f_m \in \mathbb{R}[\underline{X}]$, the basic closed semialgebraic set defined by $f_1, \ldots, f_m \in \mathbb{R}[\underline{X}]$ is the solution set of the system of inequalities $f_i(x) \geq 0$, i.e. $\{x \in \mathbb{R}^n : f_i(x) \geq 0, i = 1, \ldots, m\}$.

In 1991, Schmüdgen asserted that for $A = \mathbb{R}[\underline{X}]$, if a basic closed semialgebraic set $\mathcal{K}_S$ is compact, then any polynomial strictly positive on $\mathcal{K}_S$ actually belongs to the preordering T_S [58]. This rather surprising result had a big impact in the area. Here we give an algebraic proof of Schmüdgen's result, based on the Concrete Positivstellensatz (Theorem A.3.8) and Jacobi's Representation Theorem (Corollary 1.3.11). This is a modification of the proof due to Wörmann [67].

Lemma 1.4.2. *Suppose that M is a quadratic module of A and $a_1, \ldots, a_n \in M$. If $k - \sum_{i=1}^n a_i^2 \in M$ for some integer $k \geq 1$, then $k \pm a_i \in M$, $i = 1, \ldots, n$.*

Proof. First note that if $k - a^2 \in M$, then

$$k \pm a = \frac{1}{2}((k-1) + (k - a^2) + (a \pm 1)^2) \in M.$$

Therefore, if $k - \sum_{i=1}^n a_i^2 \in M$ then

$$k - a_i^2 = (k - \sum_{j=1}^n a_j^2) + \sum_{j \neq i} a_j^2 \in M,$$

so $k \pm a_i \in M$, $i = 1, \cdots, n$. $\square$

Theorem 1.4.3 (Wörmann)**.** *Let A be a finitely generated $\mathbb{R}$-algebra and $S \subseteq A$ be finite. Then T_S is archimedean if and only if $\mathcal{K}_S$ is compact.*

Proof. Since A is a finitely generated $\mathbb{R}$-algebra, there exist $a_1, \dots, a_n \in A$ such that $A = \mathbb{R}[a_1, \dots, a_n]$. The implication ($\Rightarrow$) is a consequence of Proposition 1.3.6. To prove ($\Leftarrow$), assume $\mathcal{K}_S$ is compact. Then $\widehat{\sum_{i=1}^n a_i^2}$ attains its maximum on $\mathcal{K}_S$. So $k - \widehat{\sum_{i=1}^n a_i^2} > 0$ on $\mathcal{K}_S$, for some integer k sufficiently large. By the Concrete Positivstellensatz A.3.8(1) there exist $p, q \in T_S$ such that

$$p(k - \sum_{i=1}^n a_i^2) = 1 + q,$$

so

$$(1+q)(k - \sum_{i=1}^n a_i^2) = p(k - \sum_{i=1}^n a_i^2)^2 \in T_S.$$

Let $S' = S \cup \{k - \sum_{i=1}^n a_i^2\}$ and $T' = T_{S'}$, i.e.,

$$T' = T_S + (k - \sum_{i=1}^n a_i^2)T_S.$$

According to Lemma 1.4.2, T' is archimedean. Thus for each $a \in A$ there exists an integer $m \geq 1$ such that $m + a \in T'$. Then $m + a = t_1 + (k - \sum_{i=1}^n a_i^2)t_2$ for $t_1, t_2 \in T_S$, so

$$(m+a)(1+q) = t_1(1+q) + p(k - \sum_{i=1}^n a_i^2)^2 t_2 \in T_S.$$

In particular, there exists an integer $m \geq 1$ such that $m - q \in T'$, so $(m-q)(1+q) \in T_S$. It follows that

$$m + \frac{m^2}{4} - q = (m-q)(1+q) + (\frac{m}{2} - q)^2 \in T_S.$$

Multiplying by $k \in T_S$ and adding $(1+q)(k - \sum_{i=1}^n a_i^2) \in T_S$ and $q \sum_{i=1}^n a_i^2 \in T_S$, we have

$$k(\frac{m}{2} + 1)^2 - \sum_{i=1}^n a_i^2 \in T_S.$$

By Lemma 1.4.2, T_S is archimedean. □

Theorem 1.4.4 (Schmüdgen)**.** *Let S be a finite subset of a finitely generated $\mathbb{R}$-algebra A. If $\mathcal{K}_S$ is compact then for any $a \in A$, $\hat{a} > 0$ on $\mathcal{K}_S$ implies $a \in T_S$.*

Proof. Combine 1.3.10 and 1.4.3. □

Chapter 2

The Moment Problem

The classical K-moment problem for a given closed set $K \subseteq \mathbb{R}^n$, is the question of when a linear functional $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ is representable as integration with respect to a positive Borel measure on K. An obvious necessary condition is that $L(f) \geq 0$, for all $f \in \mathrm{Psd}(K)$. In 1935, Haviland proved that this necessary condition is also sufficient [23, 24]. Scheiderer showed that even for semialgebraic sets, except for few cases, $\mathrm{Psd}(K)$ is not finitely generated [55], so in most of cases, there is no practical decision procedure for the membership problem of $\mathrm{Psd}(K)$, and so for $L(\mathrm{Psd}(K)) \subseteq \mathbb{R}^+$. In this chapter we see how ideas from the previous chapter come together with ideas from functional analysis to find practical solutions for the K-moment problem in some special cases of K.

2.1 Topological Vector Spaces

In the following, all vector spaces are over the field of real numbers (unless otherwise specified). A *vector space topology* on a vector space V is a topology τ on V such that every point of V is closed and the vector space operations, i.e., vector addition and scalar multiplication, are τ-continuous. A *topological vector space* is a pair (V, τ) where V is a vector space and τ is a vector space topology on V. A standard argument shows that τ is Hausdorff.

A subset $A \subseteq V$ is said to be *convex* if for every $x, y \in A$ and $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in A$. A *locally convex topology* is a vector space topology which admits a neighbourhood basis of convex open sets at each point. Suppose that in addition V is an $\mathbb{R}$-algebra. A subset $U \subseteq V$ is called an *m-set*, if $U \cdot U \subseteq U$. A locally convex topology τ on V is said to be *locally multiplicatively convex* (or *lmc* for short) if there exists a fundamental system of neighbourhoods for 0 consisting of m-sets. It is immediate from the definition that the algebra multiplication is continuous in a lmc-topology.

Definition 2.1.1. A function $\rho : V \to \mathbb{R}^+$ is called a *seminorm*, if

(i) $\forall x, y \in V \quad \rho(x+y) \leq \rho(x) + \rho(y)$,

(ii) $\forall x \in V \ \forall r \in \mathbb{R} \quad \rho(rx) = |r|\rho(x)$.

If V is an $\mathbb{R}$-algebra, ρ is called a *submultiplicative seminorm*, if in addition ρ satisfies the following

(iii) $\forall x, y \in V \quad \rho(x \cdot y) \leq \rho(x)\rho(y)$.

Let $\mathcal{F}$ be a nonempty family of seminorms on V. The weakest topology on V making all seminorms on $\mathcal{F}$ continuous is a locally convex topology on V. The family of sets of the form

$$U^{\epsilon}_{\rho_1,\ldots,\rho_k}(x) := \{y \in V \ : \ \rho_i(x-y) \leq \epsilon, \ i = 1, \ldots, k\}$$

where $\epsilon > 0$ and $\rho_1 \ldots, \rho_k \in \mathcal{F}$, forms a basis for the topology generated by $\mathcal{F}$ on V. We have the following characterization of locally convex and lmc spaces.

Theorem 2.1.2. *Let V be an algebra and τ a topology on V. Then*

1. *τ is locally convex if and only if it is generated by a family of seminorms on V.*

2. *τ is lmc if and only if it is generated by a family of submultiplicative seminorms on V.*

Proof. See [29, Theorem 6.5.1] for (1) and [4, 4.3-2] for (2). □

A norm on V is a function $\|\cdot\| : V \longrightarrow \mathbb{R}^+$ satisfying

1. $\|v\| = 0 \Leftrightarrow x = 0$,

2. $\forall \lambda \in \mathbb{R}, \ \|\lambda v\| = |\lambda|\|v\|$,

3. $\forall v_1, v_2 \in V, \ \|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$.

A topology τ on V is said to be *normable* (respectively *metrizable*), if there exists a norm (respectively metric) on V which induces the same topology as τ. Every norm induces a locally convex metric topology on V where the induced metric is defined by $d(v_1, v_2) = \|v_1 - v_2\|$.

Theorem 2.1.3. *Let (V, τ) be a topological vector space. If τ is first countable and T_1, then it is metrizable.*

Proof. See [54, Theorem 1.24]. $\square$

We denote the set of all τ-continuous linear functionals $L : V \longrightarrow \mathbb{R}$ by V_τ^* (or simply V^*, if there is no ambiguity about topology).

Remark 2.1.4. For two normed spaces $(X, \|\cdot\|_1)$ and $(Y, \|\cdot\|_2)$, a linear operator $T : X \longrightarrow Y$ is said to be *bounded* if there exists $N \geq 0$ such that for all $x \in X$, $\|T(x)\|_2 \leq N\|x\|_1$. There is a standard result which states that boundedness and continuity in normed spaces are equivalent.

Definition 2.1.5. For $C \subseteq V$, let

$$C_\tau^\vee = \{L \in V_\tau^* : L \geq 0 \text{ on } C\}$$

be the *first dual* of C and define the *second dual* of C by

$$C_\tau^{\vee\vee} = \{a \in V : \forall L \in C_\tau^\vee,\ L(a) \geq 0\}.$$

The following is immediate from the definition:

Corollary 2.1.6. *For a locally convex topological vector space (V, τ) and $C, D \subseteq V$, the following statements hold*

1. $C \subseteq D \Rightarrow D_\tau^\vee \subseteq C_\tau^\vee$.

2. $C \subseteq C_\tau^{\vee\vee}$.

3. $C_\tau^{\vee\vee\vee} = C_\tau^\vee$.

In the special case of our interest, when C is a cone, $C_\tau^{\vee\vee}$ reflects more properties. A subset C of V is called a *cone* if $C + C \subseteq C$ and $\mathbb{R}^+ C \subseteq C$. It is clear that every cone is convex.

Theorem 2.1.7 (Separation)**.** *Suppose that V is a topological vector space and A and B are disjoint nonempty convex sets in V. If A is open, then there exist $L \in V^*$ and $\gamma \in \mathbb{R}$ such that $L(x) < \gamma \leq L(y)$ for every $x \in A$ and $y \in B$. Moreover, if B is a cone, then γ can be taken to be* 0.

Proof. For the first part, see [29, Theorem 7.3.2]. For the rest, suppose that B is a cone and L and γ are given as in the first part. If $\gamma > 0$, then $L(y) > 0$ for all $y \in B$. Therefore $\forall \epsilon > 0$ $\epsilon y \in B$, so

$$0 < \gamma \leq L(\epsilon y) = \epsilon L(y) \xrightarrow{\epsilon \to 0} 0,$$

a contradiction. This implies that $\gamma \leq 0$. Note also that $L \geq 0$ on B. Otherwise, $L(x) < \gamma \leq L(y) < 0$ for any $x \in A$ and some $y \in B$. Then for $r > 0$, $ry \in B$ and

$$L(x) < \gamma \leq L(ry) = rL(y) \xrightarrow{r \to \infty} -\infty,$$

which is impossible. Therefore

$$\forall x \in A \forall y \in B \quad L(x) < \gamma \leq 0 \leq L(y),$$

and hence γ can be chosen to be 0. □

We denote the closure of C with respect to τ by $\overline{C}^{\tau}$.

Corollary 2.1.8 (Duality). *For any nonempty cone C in locally convex space (V, τ), $\overline{C}^{\tau} = C_{\tau}^{\vee\vee}$.*

Proof. Since each $L \in C_{\tau}^{\vee}$ is continuous, for any $a \in \overline{C}^{\tau}$, $L(a) \geq 0$, so $\overline{C}^{\tau} \subseteq C_{\tau}^{\vee\vee}$. Conversely, if $a \notin \overline{C}^{\tau}$, then since τ is locally convex, there exists an open convex set U of V containing a with $U \cap C = \varnothing$. By 2.1.7, there exists $L \in C_{\tau}^{\vee}$ such that $L(a) < 0$, so $a \notin C_{\tau}^{\vee\vee}$. □

2.1.1 Finest Locally Convex Topology on a Vector Space

Let V be any vector space over $\mathbb{R}$ of countable infinite dimension. For any finite dimensional subspace W of V, W has a natural topology making W homeomorphic with $\mathbb{R}^k$ where $\dim(W) = k$. If $W' \subseteq W$, then the natural topology of W' and the subspace topology induced by W are identical. We define the topology φ on V as follows: $U \subseteq V$ is open if and only if $U \cap W$ is open in W for each finite dimensional subspace W of V. That is, our topology φ is just the direct limit topology over all finite dimensional subspaces of V.

Since the dimension of V is countably infinite, we can always fix a sequence of finite dimensional subspaces $V_1 \subseteq V_2 \subseteq V_3 \subseteq \cdots$ such that $V = \bigcup_{i \geq 1} V_i$, e.g., just take $V_i = \mathbb{R}v_1 \oplus \cdots \oplus \mathbb{R}v_i$ where $\{v_1, v_2, \ldots\}$ is some basis for V. In this situation, each finite dimensional subspace of V is contained in some V_i, so $U \subseteq V$ is open if and only if $U \cap V_i$ is open in V_i for each $i \geq 1$.

Theorem 2.1.9. *The open sets in V which are convex form a basis for the direct limit topology. Moreover (V, φ) is a topological vector space and φ is the finest locally convex topology on V.*

Proof. See [44, Section 3.6]. □

Remark 2.1.10.

1. The vector space (V, φ) is not metrizable. Let U be a neighborhood of 0 in V. From the proof of [44, Theorem 3.6.1], there exist $a_i \in \mathbb{R}^{>0}$, $i = 1, 2, \ldots$, such that $\prod_{i=1}^{\infty} \langle -a_i, a_i \rangle \subseteq U$, where
$$\prod_{i=1}^{\infty} \langle -a_i, a_i \rangle = \left\{ \sum_i t_i e_i : -a_i < t_i < a_i \right\},$$
and $\{e_i\}_{i=1}^{\infty}$ forms a basis for V and all summands are 0 except for finitely many i. If there exists a countable neighborhood basis at 0 then there exist real numbers a_{ij}, $i, j = 1, 2, \ldots$, such that
$$\begin{array}{l} \prod_{i=1}^{\infty} \langle -a_{1i}, a_{1i} \rangle, \\ \prod_{i=1}^{\infty} \langle -a_{2i}, a_{2i} \rangle, \\ \vdots \end{array}$$
forms a neighborhood basis at 0. Take $0 < b_i < a_{ii}$ for each i, then $\prod_{i=1}^{\infty} \langle -b_i, b_i \rangle$ is a neighborhood of 0 which does not contain any of the above basic open sets, a contradiction.

2. Every linear functional is continuous with respect to φ. For the weak topology (induced by the set of all linear functionals), convex sets have the same closure as they have under φ [54, Theorem 3.12].

3. Direct limit topology and finest locally convex topology are defined even when V is uncountably infinite dimensional. But they only coincide when the space is countable dimensional.

2.2 K-Moment Problem

Let X be a topological space. We denote the set of all finite *signed* Borel measures on X by $\mathcal{M}(X)$. Its positive cone consisting of all finite Borel measures on X will be denoted by $\mathcal{M}_+(X)$.

Definition 2.2.1. Given $\gamma_\alpha \in \mathbb{R}$, $\alpha \in \Gamma \subseteq \mathbb{N}^n$ and a set $K \subseteq \mathbb{R}^n$, the *K-moment problem* asks whether there exists a finite Borel measure $\mu \in \mathcal{M}_+(K)$ such that

$$\int_K \underline{X}^\alpha \, d\mu = \gamma_\alpha, \;\; \forall \alpha \in \Gamma.$$

Remark 2.2.2. The most interesting case is when $\Gamma = \mathbb{N}^n$. Since $\{\underline{X}^\alpha : \alpha \in \mathbb{N}^n\}$ is a basis for $\mathbb{R}[\underline{X}]$, to any real multi-sequence $(\gamma_\alpha)_{\alpha\in\mathbb{N}^n}$ one can assign a unique linear functional $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ such that $L(\underline{X}^\alpha) = \gamma_\alpha$, $\alpha \in \mathbb{N}^n$. So the K-moment problem in this case is equivalent to asking whether a given linear functional L on $\mathbb{R}[\underline{X}]$ is representable by a finite Borel measure on K, i.e., there exists $\mu \in \mathcal{M}_+(K)$ such that

$$L(f) = \int_K f \, d\mu, \;\; f \in \mathbb{R}[\underline{X}]. \tag{2.2.1}$$

An obvious necessary condition for (2.2.1) to hold is the following:

$$\forall f \in \mathrm{Psd}(K) \;\; L(f) \geq 0. \tag{2.2.2}$$

Haviland showed that this necessary condition is also sufficient. Here we prove a more general version of what Haviland proved in [23, 24]. We assume the following version of the Riesz Representation Theorem.

Theorem 2.2.3 (Riesz Representation Theorem)**.** *Let X be a locally compact Hausdorff space, and let $L : \mathrm{C}_c(X) \longrightarrow \mathbb{R}$ be a positive linear function. Then there exists a unique $\mu \in \mathcal{M}_+(X)$ such that $L(f) = \int_X f \, d\mu$, for all $f \in \mathrm{C}_c(X)$.*

Here $\mathrm{C}_c(X)$ denotes the $\mathbb{R}$-algebra of all real valued continuous functions f on X with compact support, i.e., the set

$$\overline{\{x \in X : f(x) \neq 0\}}$$

is compact, and L is positive means

$$L(\{f \in \mathrm{C}_c(X) \; : \; f \geq 0 \text{ on } X\}) \subseteq \mathbb{R}^+.$$

Theorem 2.2.4. *Suppose A is an $\mathbb{R}$-algebra, X is a Hausdorff space, and $\hat{\ } : A \longrightarrow \mathrm{C}(X)$ is an $\mathbb{R}$-algebra homomorphism such that for some $p \in Psd(X)$, $\hat{p}^{-1}([0,i])$ is compact for $i = 1, 2, \ldots$. Then for each linear functional $L : A \longrightarrow \mathbb{R}$ satisfying $L(Psd(X)) \subseteq \mathbb{R}^+$, there exists a Borel measure μ on X such that $\forall a \in A$ $L(a) = \int_X \hat{a} \, d\mu$.*

Here $\mathrm{Psd}(X)$ denotes the set $\{a \in A \;:\; \hat{a} \geq 0 \text{ on } X\}$.

Proof of Theorem 2.2.4. (See [42, Theorem 3.1] or [44, Theorem 3.2.2].) Let $A_0 = \{\hat{a} : a \in A\}$. A_0 is a subalgebra of $\mathrm{C}(X)$.

Claim 1. $\bar{L} : A_0 \longrightarrow \mathbb{R}$ defined by $\bar{L}(\hat{a}) = L(a)$ is a well-defined linear map.
Suppose $\hat{a} = 0$. Then $\hat{a} \geq 0$, so $L(a) \geq 0$. Similarly $\widehat{-a} = -\hat{a} \geq 0$, so $-L(a) = L(-a) \geq 0$. This proves $\hat{a} = 0 \Rightarrow L(a) = 0$, which establishes Claim 1.

Define $\mathrm{C}'(X)$ to be the set of all continuous functions $f : X \longrightarrow \mathbb{R}$ for which there exists $a \in A$ such that $|f| \leq |\hat{a}|$ on X. $\mathrm{C}'(X)$ is a subalgebra of $\mathrm{C}(X)$ and $A_0 \subseteq \mathrm{C}'(X)$. If $f \in \mathrm{C}_c(X)$ then $|f| \leq i$ for some integer $i \geq 1$. Since $i \in A$, then $f \in \mathrm{C}'(X)$ and hence $\mathrm{C}_c(X)$ is a subalgebra of $\mathrm{C}'(X)$.

Claim 2. $\bar{L}$ extends to a linear map $\bar{L} : \mathrm{C}'(X) \longrightarrow \mathbb{R}$ such that

$$\bar{L}(\{f \in \mathrm{C}'(X) : f \geq 0 \text{ on } X\}) \subseteq \mathbb{R}^+.$$

By Zorn's Lemma, there exists a pair $(V, \bar{L})$ with V a subspace of $\mathrm{C}'(X)$ containing A_0, $\bar{L} : V \longrightarrow \mathbb{R}$ extending $\bar{L} : A_0 \longrightarrow \mathbb{R}$ and satisfying

$$\bar{L}(\{f \in V : f \geq 0 \text{ on } X\}) \subseteq \mathbb{R}^+, \tag{2.2.3}$$

which is maximal with respect to the partial ordering $\preceq$ defined by

$$(V_1, L'_1) \preceq (V_2, L'_2) \Leftrightarrow V_1 \subseteq V_2 \wedge L'_2|_{V_1} = L'_1. \tag{2.2.4}$$

We show that $V = \mathrm{C}'(X)$. To the contrary, suppose that $g \in \mathrm{C}'(X) \setminus V$. Since $g \in \mathrm{C}'(X)$, $|g| \leq |\hat{a}|$ for some $a \in A$. Also, note that $(\hat{a} \pm 1)^2 \geq 0$ on X and hence $|g| \leq |\hat{a}| \leq \frac{\hat{a}^2+1}{2} \in A_0$. Thus there exists $e \in \mathbb{R}$ such that

$$\sup\{\bar{L}(f_1) : f_1 \in V, f_1 \leq g\} \leq e \leq \inf\{\bar{L}(f_2) : f_2 \in V, g \leq f_2\}.$$

Extend $\bar{L}$ to $V \oplus \mathbb{R}g$ by defining $\bar{L}(g) = e$, i.e., $\bar{L}(f + dg) = \bar{L}(f) + de$, for $f \in V$ and $d \in \mathbb{R}$. Let $h = f + dg \in V \oplus \mathbb{R}g$ with $h \geq 0$ on X.

If $d = 0$ then $\bar{L}(h) = \bar{L}(f)$ which is nonnegative by (2.2.3).

If $d > 0$ then $-\frac{f}{d} \leq g$ on X, so $\bar{L}(-\frac{f}{d}) \leq e$ by (2.2.4), i.e., $\bar{L}(f + dg) \geq 0$.

If $d < 0$ then $-\frac{f}{d} \geq g$ on X, so $\bar{L}(-\frac{f}{d}) \geq e$ by (2.2.4), i.e., $\bar{L}(f + dg) \geq 0$.

This contradicts the maximality of $(V, \bar{L})$, so $V = \mathrm{C}'(X)$.

Since $\mathrm{C}_c(X) \subseteq \mathrm{C}'(X)$, Claim 2 allows us to apply the Riesz Representation Theorem to get a Borel measure $\mu \in \mathcal{M}_+(X)$ such that

$$\forall f \in \mathrm{C}_c(X) \quad \bar{L}(f) = \int_X f \; d\mu.$$

It remains to show that this holds for every $f \in \mathrm{C}'(X)$.

Let $f \in \mathrm{C}'(X)$ and decompose f as $f_+ - f_-$ where $f_+ = \max\{f, 0\}$ and $f_- = -\min\{f, 0\}$. Since $f_+, f_- \in \mathrm{C}'(X)$ are non-negative and $f = f_+ - f_-$, we can reduce to the case where $f \geq 0$ on X. By hypothesis, there exists $p \in A$ such that $\hat{p} \geq 0$ on X and for each $i \in \mathbb{N}$, $\hat{p}^{-1}([0, i])$ is compact. Let $X_i = \{x \in X : \hat{p}(x) \leq i\}$ and $X_i' = \{x \in X : q(x) \leq i\}$ where $q = f + \hat{p}$. Clearly $X_i' \subseteq X_i$, so X_i' is compact and

$$X_i' \subseteq X_{i+1}' \text{ and } \bigcup_{i \geq 1} X_i' = X.$$

Let $Y_i = \{x \in X_{i+1}' : i + \frac{1}{2} \leq q(x)\}$. Using Urysohn's Lemma for each i, there exists a continuous function $g_i : X_{i+1}' \longrightarrow [0, 1]$ such that $g_i|_{Y_i} = 0$ and $g_i|_{X_i'} = 1$. Extend g_i to X by defining $g_i = 0$ on $X \setminus X_{i+1}'$[1]. Take $f_i = f g_i$. Then $f_i \in \mathrm{C}_c(X)$ and

$$0 \leq f_i \leq f, \; f_i = f \text{ on } X_i', \; f_i = 0 \text{ on } X \setminus X_{i+1}'.$$

Claim 3. $\frac{q^2}{i} \geq f - f_i \geq 0$ on X.

The inequality $f - f_i \geq 0$ is clear. Since $f = f_i$ on X_i', the inequality is clear on X_i'. For $x \notin X_i'$, $q(x) > i$, so $q(x)^2 \geq i q(x) = i(f(x) + \hat{p}(x)) \geq i f(x) \geq i(f(x) - f_i(x))$, which shows that the inequality holds off X_i'.

By Claim 3, $\frac{1}{i}\bar{L}(q^2) \geq \bar{L}(f) - \bar{L}(f_i) \geq 0$. So $\bar{L}(f) = \lim_{i\to\infty} \bar{L}(f_i)$. Thus

$$\int_X f \; d\mu = \lim_{i\to\infty} \int_X f_i \; d\mu = \lim_{i\to\infty} \bar{L}(f_i) = \bar{L}(f).$$

□

Theorem 2.2.5 (Haviland)**.** *For a linear functional $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ and a closed set $K \subseteq \mathbb{R}^n$, the following are equivalent:*

1. *L comes from a Borel measure on K, i.e.,*

$$\exists \mu \in \mathcal{M}_+(K) \forall f \in \mathbb{R}[\underline{X}] \; L(f) = \int_K f \; d\mu.$$

[1] This also can be done using Tietze Extension Theorem.

2. $L(Psd(K)) \subseteq \mathbb{R}^+$.

Proof. The polynomial $p(\underline{X}) = \sum_{i=1}^n X_i^2$ satisfies the condition of Theorem 2.2.4. Now the conclusion is trivial. $\square$

Actually, the original result of Haviland in [23, 24] is not stated in terms of Borel measures, but rather in terms of distribution functions. Haviland's Theorem gives a complete solution for the K-moment problem. Scheiderer [55] showed that $\mathrm{Psd}(K)$ usually is not finitely generated. So usually, there is no decision procedure for the membership problem for $\mathrm{Psd}(K)$. Therefore the Haviland's solution is usually impractical, unless, positivity over a decidable subcone of $\mathrm{Psd}(K)$ gives a certificate for positivity over the whole of $\mathrm{Psd}(K)$.

Proposition 2.2.6. *For every closed subset K of $\mathbb{R}^n$, $\overline{Psd(K)}^{\varphi} = Psd(K)$.*

Proof. Each evaluation map $e_a : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$; $f \mapsto f(a)$, is continuous by Remark 2.1.10.(2), so $e_a^{-1}(\mathbb{R}^+)$ is closed. Therefore $\mathrm{Psd}(K) = \bigcap_{a\in K} e_a^{-1}(\mathbb{R}^+)$ is also closed. $\square$

Proposition 2.2.6, shows that $\mathrm{Psd}(K)$ is closed in the finest locally convex topology φ on $\mathbb{R}[\underline{X}]$. So the closure of every subcone C of $\mathrm{Psd}(K)$ with respect to φ, is contained in $\mathrm{Psd}(K)$.

Definition 2.2.7. We say $C \subseteq \mathbb{R}[\underline{X}]$ solves the K-moment problem or C satisfies the K-*strong moment property (SMP)*, if $C_{\varphi}^{\vee\vee} = \mathrm{Psd}(K)$.

Note that by Duality Theorem 2.1.8 and Proposition 2.2.6, C satisfies K-SMP if and only if

$$\overline{C}^{\varphi} = \mathrm{Psd}(K). \tag{2.2.5}$$

Moreover, by Corollary 2.1.6(3), C satisfies K-SMP if and only if $C_{\varphi}^{\vee} = \mathrm{Psd}(K)_{\varphi}^{\vee}$.

The following theorem describes the connection between K-SMP and K-moment problem and also summarizes materials of this section.

Theorem 2.2.8. *Let K be a closed subset of $\mathbb{R}^n$. Then a cone $C \subseteq Psd(K)$ satisfies K-SMP if and only if for each $L \in C_{\varphi}^{\vee}$ there exists $\mu \in \mathcal{M}_+(K)$ such that*

$$\forall f \in \mathbb{R}[\underline{X}] \quad L(f) = \int_K f \, d\mu.$$

In the upcoming section we list several examples of cases where K-SMP holds and also cases where K-SMP fails for a basic closed semialgebraic set K.

2.3 Moment Problem for Semialgebraic sets

In this section we list some result concerning the K-moment problem when K is basic closed semialgebraic set. We begin with several historical 1-dimensional examples for which K-SMP holds.

Example 2.3.1. Recall from Chapter 1 that for a finite set $S \subset \mathbb{R}[\underline{X}]$, T_S and M_S denote the preordering and the quadratic module generated by S, respectively. Also $\mathcal{K}_S$ is the basic closed semialgebraic set $\{a \in \mathbb{R}^n : \forall f \in S \quad f(a) \geq 0\}$.

- (Stieltjes [62]) If $S = \{X\}$ then $T_S = M_S$ solves the $\mathcal{K}_S$-moment problem. Here $\mathcal{K}_S = [0, \infty)$.
- (Hamburger [20]) $\sum \mathbb{R}[X]^2$ satisfies $\mathbb{R}$-SMP.
- (Hausdorff [22]) If $S = \{X, 1-X\}$ then $T_S = M_S$ solves the $\mathcal{K}_S$-moment problem. Here $\mathcal{K}_S = [0,1]$.
- (Švekov [64]) Let $S = \{X^2 - X\}$ so $\mathcal{K}_S = (-\infty, 0] \cup [1, \infty)$. Then $T_S = M_S$ solves the $\mathcal{K}_S$-moment problem.

In general the following well-know results hold for compact semialgebraic sets:

Theorem 2.3.2 (Schmüdgen)**.** *If $\mathcal{K}_S$ is compact, then T_S satisfies the $\mathcal{K}_S$-SMP.*

Proof. Let $f \in \operatorname{Psd}(\mathcal{K}_S)$ and $\epsilon > 0$. Then $f + \epsilon > 0$ on $\mathcal{K}_S$, so $f + \epsilon \in T_S$ by Theorem 1.4.4. So, for each $L \in T_S^{\vee}$, $L(f+\epsilon) \geq 0$, letting $\epsilon \to 0$, we see $L(f) \geq 0$. Thus $f \in T_S^{\vee\vee} = \overline{T_S}^{\varphi}$, so $\overline{T_S}^{\varphi} = \operatorname{Psd}(\mathcal{K}_S)$. □

Theorem 2.3.3. *If M_S is archimedean, then M_S satisfies the $\mathcal{K}_S$-SMP.*

Proof. Follow the argument of the previous theorem, using the Representation Theorem 1.3.10. □

Let $S = \{g_1, \ldots, g_s\} \subseteq \mathbb{R}[\underline{X}]$, $M = M_S$ and $d \geq 0$ be an integer number. Let $\mathbb{R}[\underline{X}]_d$ be the subspace of $\mathbb{R}[\underline{X}]$ consisting of polynomials of degree at most d and $M_d = M \cap \mathbb{R}[\underline{X}]_d$. Each $f \in M$ is expressible as $f = \sum_{i=0}^{s} \sigma_i g_i$ with $\sigma_i \in \sum \mathbb{R}[\underline{X}]^2$ and $g_0 = 1$. Define

$$M[d] := \{\sum_{i=0}^{s} \sigma_i g_i \;:\; \sigma_i \in \sum \mathbb{R}[\underline{X}]^2,\;\; \deg(\sigma_i g_i) \leq d,\;\; i = 0, \ldots, s\}.$$

Clearly $M[d] \subseteq M_d$, but for $f \in M$, of relatively low degree, the degree of each term $\sigma_i g_i$ in any representation of f may be large, so in general $M[d] \subsetneq M_d$. We say M is *stable* if there exists a function $\ell : \mathbb{Z}^+ \longrightarrow \mathbb{Z}^+$ such that for each $d \geq 0$ and $f \in M_d$, f has a representation $f = \sum_{i=0}^{s} \sigma_i g_i$, with $\deg(\sigma_i g_i) \leq \ell(d)$ for each i. The following result, due to Scheiderer connects the concept of stability to the moment problem.

Theorem 2.3.4. *Let $M = M_S$ be a finitely generated quadratic module in $\mathbb{R}[\underline{X}]$. Then*

1. $\epsilon + \sqrt{supp\ M} \subseteq M$ *for all real* $\epsilon > 0$*. In particular* $\sqrt{supp\ M} \subseteq \overline{M}^{\varphi}$.

2. *If M is stable then $\overline{M}^{\varphi} = M + \sqrt{supp\ M}$ and $\overline{M}^{\varphi}$ is also stable.*

Proof. See [56, Theorem 3.17] or [44, Theorem 4.1.2]. □

Lemma 2.3.5. *If $\mathcal{K}_S$ contains an n-dimensional cone, then M_S is stable.*

Proof. (See [44, Example 4.1.5]) We prove that if $f, g \in \mathbb{R}[\underline{X}]$ are nonnegative on an n-dimensional cone C in $\mathbb{R}^n$, then $\deg(f+g) = \max\{\deg(f), \deg(g)\}$. This is clear if $\deg(f) \neq \deg(g)$. Suppose that $\deg(f) = \deg(g) = d$ and $f = f_0 + \cdots + f_d$, $g = g_0 + \cdots + g_d$ are the homogeneous decompositions of f and g. Since f_d and g_d are nonzero, there exists p, an interior point of C such that $f_d(p) \neq 0$ and $g_d(p) \neq 0$. C is a cone, so for any $\lambda > 0$, $\lambda p \in C$, therefore $f(\lambda p), g(\lambda p) \geq 0$. But $f(\lambda p)$ and $g(\lambda p)$ are polynomials of degree d on λ and since $f(\lambda p), g(\lambda p) \geq 0$, we see that $f_d(p), g_d(p) > 0$. Thus $f_d(p) + g_d(p) > 0$ and hence $f_d + g_d \neq 0$. This proves that $\deg(f+g) = d$. Now, one can take ℓ to be the identity function on $\mathbb{Z}^+$. □

For the case $n = 1$ in [34], Kuhlmann and Marshall completely described the relation between T_S and $\mathcal{K}_S$ for finite S. Assume that K is a nonempty, closed semialgebraic set in $\mathbb{R}$, i.e., a finite union of closed intervals and points. The *natural description* of K is a certain finite subset S of $\mathbb{R}[X]$, defined as follows:

- If K has a least element a, then $X - a \in S$.
- If K has a greatest element b, then $b - X \in S$.
- For every $a, b \in K$, $a < b$, if $(a, b) \cap K = \varnothing$, then $(X - a)(b - X) \in S$.
- These are the only elements of S.

Theorem 2.3.6. *Suppose $n = 1$ and $\mathcal{K}_S$ is not compact. Then T_S satisfies $\mathcal{K}_S$-SMP if and only if S contains the natural description of $\mathcal{K}_S$.*

Proof. Note that $\mathcal{K}_S$ contains a 1-dimensional cone, so T_S is stable by Lemma 2.3.5 and is φ-closed by 2.3.4 since $\sqrt{\operatorname{supp} T_S} = \langle 0 \rangle$. So it is enough to check $T_S = \operatorname{Psd}(\mathcal{K}_S)$. For a complete proof see [34, Theorem 2.2]. □

Remark 2.3.7. Theorem 2.3.6 enables us to find an example for which the K-moment problem depends on the description of K. For example $[0, \infty) = \mathcal{K}_{\{X^3\}}$, but $\{X^3\}$ does not contain the natural description of $[0, \infty)$, so $T_{\{X^3\}}$ does not satisfy the $[0, \infty)$-SMP, but $T_{\{X\}}$ does.

In contrast to Theorem 2.3.6, there is no similar result for $n \geq 2$:

Proposition 2.3.8. *Suppose $n \geq 2$ and $\mathcal{K}_S$ contains a 2-dimensional affine cone. Then T_S does not satisfy $\mathcal{K}_S$-SMP.*

Proof. See [34, Corollary 3.10]. □

The *dimension* of a semialgebraic set K in $\mathbb{R}^n$ is defined to be the Krull dimension of the ring $\frac{\mathbb{R}[\underline{X}]}{\mathcal{Z}(K)}$. Scheiderer proved the following result which connects the two concepts of stability and dimension.

Theorem 2.3.9. *If M_S is stable and $\dim(\mathcal{K}_S) \geq 2$ then M_S does not satisfy $\mathcal{K}_S$-SMP.*

Proof. See [56, Theorem 5.4]. □

2.4 Moment Problem for Continuous Linear Functionals

Now, we return to the equation (2.2.5). In Section 2.3, we reviewed several attempts to solve a variation of (2.2.5) for $K = \mathcal{K}_S$ and $C = T_S$ or $C = M_S$, i.e.,

$$\overline{M_S}^{\varphi} = \operatorname{Psd}(\mathcal{K}_S) \text{ or } \overline{T_S}^{\varphi} = \operatorname{Psd}(\mathcal{K}_S).$$

We quote a result due to Berg, Christensen and Jensen [6], concerning the $[-1, 1]^n$-moment problem for ℓ_1-continuous linear functionals. They proved that every positive semidefinite (PSD), ℓ_1-continuous functional on $\mathbb{R}[\underline{X}]$ is representable by a measure on $[-1, 1]^n$, i.e.,

$$\overline{\sum \mathbb{R}[\underline{X}]^2}^{\ell_1} = \operatorname{Psd}([-1, 1]^n).$$

Thus, instead of the whole family $\mathcal{M}_+([-1,1]^n)$, we reduce to the subfamily that induces ℓ_1-continuous functionals. Later, this result has been generalized by Berg and Maserick [9], Lasserre and Netzer [39] and Ghasemi, Marshall and Wagner [19].

These observations suggest that solution of the classical moment problem might be easier for subclasses of functionals, in this case ℓ_1-continuous functionals. We rewrite (2.2.5) in the following format:

$$\overline{C}^{\tau} = \mathrm{Psd}(K), \tag{2.4.1}$$

for a cone C, a closed set K and a locally convex topology τ[2]. Note that by Duality (Corollary 2.1.8), (2.4.1) is equivalent to $C_{\tau}^{\vee\vee} = \mathrm{Psd}(K)$. So if (2.4.1) holds, then every τ-continuous linear functional $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$, with $L(C) \subseteq \mathbb{R}^+$ is representable as an integration with respect to a Borel measure on K, i.e.,

$$L(f) = \int_K f \; d\mu, \quad \forall f \in \mathbb{R}[\underline{X}].$$

We discuss (2.4.1) in chapters 4, 5 and 6.

[2]Since in general we are interested in solving the moment problem for τ-continuous linear functional, it is natural to work with the locally convex topologies. This approach enables us to switch between τ and its weak topology, without losing the closures of convex sets, and for our purpose cones.

Chapter 3

Lower Bounds for a Polynomial in terms of its Coefficients

In this chapter, we make use of a result of Hurwitz and Reznick [26, 53], and a consequence of this result due to Fidalgo and Kovacec [15], to determine a new sufficient condition for a polynomial $f \in \mathbb{R}[X_1, \dots, X_n]$ of even degree to be a sum of squares. We apply this result to obtain a new lower bound f_{gp} for f, and we see how f_{gp} can be computed using geometric programming. The lower bound f_{gp} is generally not as good as the lower bound f_{sos} introduced by Lasserre [35] and Parrilo and Sturmfels [47], which is computed using SDP, but a run time comparison shows that, in practice, the computation of f_{gp} is much faster. The computation is simplest when the highest degree term of f has the form $\sum_{i=1}^{n} a_i X_i^{2d}$, $a_i > 0$, $i = 1, \dots, n$. The material of this chapter is mainly taken from [18].

Fix a non-constant polynomial $f \in \mathbb{R}[\underline{X}] = \mathbb{R}[X_1, \cdots, X_n]$, where $n \geq 1$ is an integer number, and let f_* be the global minimum of f, defined by

$$f_* := \inf\{f(\underline{a}) \ : \ \underline{a} \in \mathbb{R}^n\}.$$

We say f is PSD if $f(\underline{a}) \geq 0 \ \forall \underline{a} \in \mathbb{R}^n$. Clearly

$$\inf\{f(\underline{a}) : \underline{a} \in \mathbb{R}^n\} = \sup\{r \in \mathbb{R} \ : \ f - r \text{ is PSD}\},$$

so finding f_* reduces to determining when $f - r$ is PSD.

Suppose that $\deg(f) = m$ and decompose f as $f = f_0 + \cdots + f_m$ where f_i is a form (a homogeneous polynomial) with $\deg(f_i) = i$, $i = 0, \dots, m$. This decomposition is called the *homogeneous decomposition* of f. A necessary condition for $f_* \neq -\infty$ is that f_m is PSD (hence m is even). A form $g \in \mathbb{R}[\underline{X}]$ is said to be positive definite (PD) if $g(\underline{a}) > 0$ for all $\underline{a} \in \mathbb{R}^n$, $\underline{a} \neq \underline{0}$. A sufficient condition for $f_* \neq -\infty$ is that f_m is PD [43].

It is known that deciding when a polynomial is PSD is NP-hard [5, Theorem 1.1]. Deciding when a polynomial is a sums of squares (SOS) is much easier. Actually, there is a polynomial time method, known as semidefinite programming (SDP), which can be used to decide when a polynomial $f \in \mathbb{R}[\underline{X}]$ is SOS [35, 47]. Note that any SOS polynomial is obviously PSD, so it is natural to ask if the converse is true, i.e. is every PSD polynomial SOS? This question first appeared in Minkowski's thesis and he guessed that in general the answer is NO. Later, in [25], Hilbert gave a complete answer to this question, see [10, Section 6.6]. Let us denote the cone of PSD forms of degree $2d$ in n variables by $P_{2d,n}$ and the cone of SOS forms of degree $2d$ in n variables by $\Sigma_{2d,n}$. Hilbert proved that $P_{2d,n} = \Sigma_{2d,n}$ *if and only if* $(n \le 2)$ *or* $(d = 1)$ *or* $(n = 3$ *and* $d = 2)$.

Let $\sum \mathbb{R}[\underline{X}]^2$ denote the cone of all SOS polynomials in $\mathbb{R}[\underline{X}]$ and, for $f \in \mathbb{R}[\underline{X}]$, define

$$f_{\text{sos}} := \sup\{r \in \mathbb{R} \ : \ f - r \in \sum \mathbb{R}[\underline{X}]^2\}.$$

Since SOS implies PSD, $f_{\text{sos}} \le f_*$. Moreover, if $f_{\text{sos}} \ne -\infty$ then f_{sos} can be computed numerically in polynomial time, as close as desired, using SDP [35] [47]. We denote by $P^\circ_{2d,n}$ and $\Sigma^\circ_{2d,n}$, the interior of $P_{2d,n}$ and $\Sigma_{2d,n}$ in the vector space of forms of degree $2d$ in $\mathbb{R}[\underline{X}]$, equipped with the euclidean topology. A necessary condition for $f_{\text{sos}} \ne -\infty$ is that $f_{2d} \in \Sigma_{2d,n}$. A sufficient condition for $f_{\text{sos}} \ne -\infty$ is that $f_{2d} \in \Sigma^\circ_{2d,n}$ [45, Proposition. 5.1].

According to our notations, every polynomial $f \in \mathbb{R}[\underline{X}]$ can be written as $f(\underline{X}) = \sum_{\alpha \in \mathbb{N}^n} f_\alpha \underline{X}^\alpha$, where $f_\alpha \in \mathbb{R}$ and $f_\alpha = 0$, except for finitely many α. Assume now that f is non-constant and has even degree. Let $\Omega(f) = \{\alpha \in \mathbb{N}^n \ : \ f_\alpha \ne 0\} \smallsetminus \{\underline{0}, 2d\epsilon_1, \ldots, 2d\epsilon_n\}$, where $2d = \deg(f)$, $\epsilon_i = (\delta_{i1}, \ldots, \delta_{in})$, and

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \ne j. \end{cases}$$

We denote $f_{\underline{0}}$ and $f_{2d\epsilon_i}$ by f_0 and $f_{2d,i}$ for short. Thus f has the form

$$f = f_0 + \sum_{\alpha \in \Omega(f)} f_\alpha \underline{X}^\alpha + \sum_{i=1}^{n} f_{2d,i} X_i^{2d}. \tag{3.0.1}$$

Let $\Delta(f) = \{\alpha \in \Omega(f) \ : \ f_\alpha \underline{X}^\alpha$ is not a square in $\mathbb{R}[\underline{X}]\} = \{\alpha \in \Omega(f) : $ either $f_\alpha < 0$ or α_i is odd for some $1 \le i \le n\}$. Since the polynomial f is usually fixed, we will often denote $\Omega(f)$ and $\Delta(f)$ just by Ω and Δ for short.

Let $\bar{f}(\underline{X},Y) = Y^{2d}f(\frac{X_1}{Y},\dots,\frac{X_n}{Y})$. From (3.0.1) it is clear that

$$\bar{f}(\underline{X},Y) = f_0Y^{2d} + \sum_{\alpha\in\Omega} f_\alpha \underline{X}^\alpha Y^{2d-|\alpha|} + \sum_{i=1}^{n} f_{2d,i}X_i^{2d}$$

is a form of degree $2d$, called the *homogenization* of f. We have the following well-known result:

Proposition 3.0.1. *f is PSD if and only if $\bar{f}$ is PSD. f is SOS if and only if $\bar{f}$ is SOS.*

Proof. See [44, Proposition 1.2.4]. □

3.1 Sufficient conditions for a form to be SOS

We recall the following result, due to Hurwitz and Reznick.

Theorem 3.1.1 (Hurwitz-Reznick). *Suppose $p(\underline{X}) = \sum_{i=1}^{n}\alpha_i X_i^{2d} - 2dX_1^{\alpha_1}\cdots X_n^{\alpha_n}$, where $\alpha = (\alpha_1,\dots,\alpha_n)\in\mathbb{N}^n$, $|\alpha| = 2d$. Then p is sum of binomial squares (SOBS).*

Here, SOBS is shorthand for a sum of binomial squares, i.e., a sum of squares of the form $(a\underline{X}^\alpha - b\underline{X}^\beta)^2$

In his 1891 paper [26], Hurwitz uses symmetric polynomials in $X_1,\dots,X_{2d}$ to give an explicit representation of $\sum_{i=1}^{2d} X_i^{2d} - 2d\prod_{i=1}^{2d} X_i$ as a sum of squares. Theorem 3.1.1 can be deduced from this representation. Theorem 3.1.1 can also be deduced from results in [52, 53], specially, from [53, Theorems 2.2 and 4.4]. Here is another proof.

Proof. By induction on n. If $n = 1$ then $p = 0$ and the result is clear. Assume now that $n \geq 2$. We can assume each α_i is strictly positive, otherwise, we reduce to a case with at most $n-1$ variables.

Case 1: Suppose that there exist $1 \leq i_1, i_2 \leq n$, such that $i_1 \neq i_2$, with $\alpha_{i_1} \leq d$ and $\alpha_{i_2} \leq d$. Decompose $\alpha = (\alpha_1,\dots,\alpha_n)$ as $\alpha = \beta+\gamma$ where $\beta,\gamma\in\mathbb{N}^n$, $\beta_{i_1} = 0$, $\gamma_{i_2} = 0$ and $|\beta| = |\gamma| = d$. Then

$$(\underline{X}^\beta - \underline{X}^\gamma)^2 = \underline{X}^{2\beta} - 2\underline{X}^\beta\underline{X}^\gamma + \underline{X}^{2\gamma} = \underline{X}^{2\beta} - 2\underline{X}^\alpha + \underline{X}^{2\gamma},$$

therefore,

$$\begin{aligned}
p(\underline{X}) &= \sum_{i=1}^{n} \alpha_i X_i^{2d} - 2d\underline{X}^{\alpha} \\
&= \sum_{i=1}^{n} \alpha_i X_i^{2d} - d(\underline{X}^{2\beta} + \underline{X}^{2\gamma} - (\underline{X}^{\beta} - \underline{X}^{\gamma})^2) \\
&= \frac{1}{2}\left(\sum_{i=1}^{n} 2\beta_i X_i^{2d} - 2d\underline{X}^{2\beta}\right) \\
&+ \frac{1}{2}\left(\sum_{i=1}^{n} 2\gamma_i X_i^2 d - 2d\underline{X}^{2\gamma}\right) + d(\underline{X}^{\beta} - \underline{X}^{\gamma})^2.
\end{aligned}$$

Each term is SOBS, by the induction hypothesis.

Case 2: Suppose we are not in Case 1. Since there is at most one i satisfying $\alpha_i > d$, it follows that $n = 2$, so $p(\underline{X}) = \alpha_1 X_1^{2d} + \alpha_2 X_2^{2d} - 2d X_1^{\alpha_1} X_2^{\alpha_2}$. We know that $p \geq 0$ on $\mathbb{R}^2$, by the arithmetic-geometric inequality. Since $n = 2$ and p is homogeneous, it follows that p is SOS (see [25]).

Showing p is SOBS requires more work. Denote by $\mathrm{AGI}(2, d)$ the set of all homogeneous polynomials of the form $p = \alpha_1 X_1^{2d} + \alpha_2 X_2^{2d} - 2d X_1^{\alpha_1} X_2^{\alpha_2}$, $\alpha_1, \alpha_2 \in \mathbb{N}$ and $\alpha_1 + \alpha_2 = 2d$. This set is finite. If $\alpha_1 = 0$ or $\alpha_1 = 2d$ then $p = 0$ which is trivially SOBS. If $\alpha_1 = \alpha_2 = d$ then $p(\underline{X}) = d(X_1^d - X_2^d)^2$, which is also SOBS. Suppose now that $0 < \alpha_1 < 2d$, $\alpha_1 \neq d$ and $\alpha_1 > \alpha_2$ (The argument for $\alpha_1 < \alpha_2$ is similar). Decompose $\alpha = (\alpha_1, \alpha_2)$ as $\alpha = \beta + \gamma$, $\beta = (d, 0)$ and $\gamma = (\alpha_1 - d, \alpha_2)$. Expand p as in the proof of Case 1 to obtain

$$p(\underline{X}) = \frac{1}{2}\left(\sum_{i=1}^{2} 2\beta_i X_i^{2d} - 2d\underline{X}^{2\beta}\right) + \frac{1}{2}\left(\sum_{i=1}^{2} 2\gamma_i X_i^{2d} - 2d\underline{X}^{2\gamma}\right) + d(\underline{X}^{\beta} - \underline{X}^{\gamma})^2.$$

Observe that $\sum_{i=1}^{2} 2\beta_i X_i^{2d} - 2d\underline{X}^{2\beta} = 0$.

Thus $p = \frac{1}{2}p_1 + d(\underline{X}^{\beta} - \underline{X}^{\gamma})^2$, where $p_1 = \sum_{i=1}^{2} 2\gamma_i X_i^{2d} - 2d\underline{X}^{2\gamma}$. If p_1 is SOBS then p is also SOBS. If p_1 is not SOBS then we can repeat to get $p_1 = \frac{1}{2}p_2 + d(\underline{X}^{\beta'} - \underline{X}^{\gamma'})^2$. Continuing in this way we get a sequence $p = p_0, p_1, p_2, \cdots$ with each p_i an element of the finite set $\mathrm{AGI}(2, d)$, so $p_i = p_j$ for some $i < j$. Since $p_i = 2^{i-j} p_j +$ *a sum of binomial squares*, this implies p_i is SOBS and hence that p is SOBS. $\square$

In [15], Fidalgo and Kovacec prove the following result, which is a corollary of the Hurwitz-Reznick result.

Corollary 3.1.2 (Fidalgo-Kovacek)**.** *For a form $p(\underline{X}) = \sum_{i=1}^{n} \beta_i X_i^{2d} - \mu \underline{X}^{\alpha}$ such that $\alpha \in \mathbb{N}^n$, $|\alpha| = 2d$, $\beta_i \geq 0$ for $i = 1, \cdots, n$, and $\mu \geq 0$ if all α_i are even, the following are equivalent:*

1. p *is* PSD.

2. $\mu^{2d} \prod_{i=1}^{n} \alpha_i^{\alpha_i} \le (2d)^{2d} \prod_{i=1}^{n} \beta_i^{\alpha_i}$.

3. p *is* SOBS.

4. p *is* SOS.

Proof. See [15, Theorem 2.3]. (3) $\Rightarrow$ (4) and (4) $\Rightarrow$ (1) are trivial, so it suffices to show (1) $\Rightarrow$ (2) and (2) $\Rightarrow$ (3). If some α_i is odd then, making the change of variables $Y_i = -X_i$, $Y_j = X_j$ for $j \neq i$, μ gets replaced by $-\mu$. In this way, we can assume $\mu \ge 0$. If some α_i is zero, set $X_i = 0$ and proceed by induction on n. In this way, we can assume $\alpha_i > 0$, $i = 1, \dots, n$. If $\mu = 0$ the result is trivially true, so we can assume $\mu > 0$. If some β_i is zero, then (2) fails. Setting $X_j = 1$ for $j \neq i$, and letting $X_i \to \infty$, we see that (1) also fails. Thus the result is trivially true in this case. Thus we can assume $\beta_i > 0$, $i = 1, \dots, n$.

(1) $\Rightarrow$ (2). Assume (1), so $p(x) \ge 0$ for all $x \in \mathbb{R}^n$. Taking

$$x := ((\frac{\alpha_i}{\beta_i})^{1/2d}, \dots, (\frac{\alpha_n}{\beta_n})^{1/2d}),$$

we see that

$$p(x) = \sum_{i=1}^{n} \alpha_i - \mu \prod_{i=1}^{n} (\frac{\alpha_i}{\beta_i})^{\alpha_i/2d} = 2d - \mu \prod_{i=1}^{n} (\frac{\alpha_i}{\beta_i})^{\alpha_i/2d} \ge 0,$$

so $\mu \prod_{i=1}^{n} (\frac{\alpha_i}{\beta_i})^{\alpha_i/2d} \le 2d$. This proves (2).

(2) $\Rightarrow$ (3). Make a change of variables $X_i = (\frac{\alpha_i}{\beta_i})^{1/2d} Y_i$, $i = 1, \dots n$. Let $\mu_1 := \mu \prod_{i=1}^{n} (\frac{\alpha_i}{\beta_i})^{\alpha_i/2d}$ so, by (2), $\mu_1 \le 2d$, i.e., $\frac{2d}{\mu_1} \ge 1$. Then

$$p(\underline{X}) = \sum_{i=1}^{n} \alpha_i Y_i^{2d} - \mu_1 \underline{Y}^{\alpha} = \frac{\mu_1}{2d} [\sum_{i=1}^{n} \alpha_i Y_i^{2d} (\frac{2d}{\mu_1} - 1) + \sum_{i=1}^{n} \alpha_i Y_i^{2d} - 2d \underline{Y}^{\alpha}],$$

which is SOBS, by the Hurwitz-Reznick result. This proves (3). □

Next, we prove our main new result of this section, which gives a sufficient condition on the coefficients for a polynomial to be a sum of squares.

Theorem 3.1.3. *Suppose f is a form of degree $2d$. A sufficient condition for f to be SOBS is that there exist nonnegative real numbers $a_{\alpha,i}$ for $\alpha \in \Delta$, $i = 1, \dots, n$ such that*

1. $\forall \alpha \in \Delta \quad (2d)^{2d} a_{\alpha}^{\alpha} = f_{\alpha}^{2d} \alpha^{\alpha}$.

2. $f_{2d,i} \ge \sum_{\alpha \in \Delta} a_{\alpha,i}$, $i = 1, \dots, n$.

Here, $a_\alpha := (a_{\alpha,1}, \dots, a_{\alpha,n})$.

Proof. Suppose that such real numbers exist. Then condition (1) together with Corollary 3.1.2 implies that $\sum_{i=1}^n a_{\alpha,i} X_i^{2d} + f_\alpha \underline{X}^\alpha$ is SOBS for each $\alpha \in \Delta$, so

$$\sum_{i=1}^n \Big(\sum_{\alpha\in\Delta} a_{\alpha,i}\Big) X_i^{2d} + \sum_{\alpha\in\Delta} f_\alpha \underline{X}^\alpha$$

is SOBS. Combining with (2), it follows that $\sum_{i=1}^n f_{2d,i} X_i^{2d} + \sum_{\alpha\in\Delta} f_\alpha \underline{X}^\alpha$ is SOBS. Since each $f_\alpha \underline{X}^\alpha$ for $\alpha \in \Omega \setminus \Delta$ is a square, this implies $f(\underline{X})$ is SOBS. □

Remark 3.1.4.

1. From condition (1) of Theorem 3.1.3 we see that $a_{\alpha,i} = 0 \Rightarrow \alpha_i = 0$.

2. Let a be an array of real numbers satisfying the conditions of Theorem 3.1.3, and define the array $a^* = (a^*_{\alpha,i})$ by

$$a^*_{\alpha,i} = \begin{cases} a_{\alpha,i} & \text{if } \alpha_i \neq 0 \\ 0 & \text{if } \alpha_i = 0. \end{cases}$$

   Then a^* also satisfies the conditions of Theorem 3.1.3. Thus we are free to require the converse condition $\alpha_i = 0 \Rightarrow a_{\alpha,i} = 0$ too, if we want.

We mention some corollaries of Theorem 3.1.3. Corollaries 3.1.5 and 3.1.6 were known earlier. Corollary 3.1.7 is an improved version of Corollary 3.1.6. Corollary 3.1.9 is a new result.

Corollary 3.1.5. *For any polynomial* $f \in \mathbb{R}[\underline{X}]$ *of degree* $2d$, *if*

$$\text{(L1)} \quad f_0 \geq \sum_{\alpha\in\Delta} |f_\alpha| \frac{2d-|\alpha|}{2d} \quad \textit{and} \quad \text{(L2)} \quad f_{2d,i} \geq \sum_{\alpha\in\Delta} |f_\alpha| \frac{\alpha_i}{2d}, \quad i = 1, \dots, n,$$

then f *is a sum of squares.*

Proof. (See [36, Theorem 3] and [17, Theorem 2.2].) Apply Theorem 3.1.3 to the homogenization $\bar{f}(\underline{X}, Y)$ of f, taking $a_{\alpha,i} = |f_\alpha| \frac{\alpha_i}{2d}$, $i = 1, \dots, n$ and $a_{\alpha,Y} = |f_\alpha| \frac{2d-|\alpha_i|}{2d}$ for each $\alpha \in \Delta$. For $\alpha \in \Delta$,

$$\begin{aligned} (2d)^{2d} a_\alpha^\alpha &= (2d)^{2d} \left(\frac{|f_\alpha|(2d-|\alpha|)}{2d}\right)^{2d-|\alpha|} \textstyle\prod_{i=1}^n \left(\frac{|f_\alpha|\alpha_i}{2d}\right)^{\alpha_i} \\ &= (2d)^{2d} |f_\alpha|^{2d-|\alpha|} (2d-|\alpha|)^{2d-|\alpha|} |f_\alpha|^{|\alpha|} \alpha^\alpha (2d)^{-2d} \\ &= |f_\alpha|^{2d} \alpha^\alpha (2d-|\alpha|)^{2d-|\alpha|}. \end{aligned}$$

So, 3.1.3(1) holds. (L1) and (L2) imply 3.1.3(2), therefore, by Theorem 3.1.3, $\bar{f}$ and hence f is SOBS. □

Corollary 3.1.6. *Suppose $f \in \mathbb{R}[\underline{X}]$ is a form of degree $2d$ and*

$$\min_{i=1,\dots,n} f_{2d,i} \geq \frac{1}{2d} \sum_{\alpha\in\Delta} |f_\alpha|(\alpha^\alpha)^{\frac{1}{2d}}.$$

Then f is SOBS.

Proof. (See [15, Theorem 4.3] and [17, Theorem 2.3].) Apply Theorem 3.1.3 with $a_{\alpha,i} = |f_\alpha|\frac{\alpha^{\alpha/2d}}{2d}$, $\forall \alpha \in \Delta$, $i = 1\dots,n$. □

Corollary 3.1.7. *Suppose f is a form of degree $2d$, $f_{2d,i} > 0$, $i = 1,\dots,n$ and*

$$\sum_{\alpha\in\Delta} \frac{|f_\alpha|\alpha^{\alpha/2d}}{2d\prod_{i=1}^n f_{2d,i}^{\alpha_i/2d}} \leq 1.$$

Then f is SOBS.

Proof. Applying Theorem 3.1.3 with $a_{\alpha,i} = \frac{|f_\alpha|\alpha^{\alpha/2d} f_{2d,i}}{2d\prod_{j=1}^n f_{2d,j}^{\alpha_j/2d}}$, we have we have

$$\begin{aligned}
2d\prod_{\alpha_i\neq0}\left(\frac{a_{\alpha,i}}{\alpha_i}\right)^{\alpha_i/2d} &= 2d\prod_{\alpha_i\neq0}\left(\frac{|f_\alpha| f_{2d,i}\alpha^{\alpha/2d}}{\alpha_i 2d\prod_{j=1}^n f_{2d,j}^{\alpha_j/2d}}\right)^{\alpha_i/2d}\\
&= |f_\alpha|\prod_{\alpha_i\neq0}\left(\frac{f_{2d,i}^{\alpha_i/2d}\alpha^{\alpha_i\alpha/2d}}{\alpha_i^{\alpha_i}\prod_{j=1}^n f_{2d,j}^{\alpha_j^2/2d}}\right)^{1/2d}\\
&= \frac{|f_\alpha|}{\prod_{j=1}^n f_{2d,j}^{\alpha_j/2d}}\prod_{\alpha_i\neq0} f_{2d,i}^{\alpha_i/2d}\\
&= |f_\alpha|.
\end{aligned}$$

Now, by 3.1.3 the conclusion follows. □

Remark 3.1.8. Corollary 3.1.7 is an improved version of Corollary 3.1.6. This requires some explanation. Suppose that $f_{2d,i} \geq \frac{1}{2d}\sum_{\alpha\in\Delta}|f_\alpha|\alpha^{\alpha/2d}$, $i = 1,\dots,n$. Let $f_{2d,i_0} := \min\{f_{2d,i} : i = 1,\dots,n\}$. Then

$$\prod_{i=1}^n f_{2d,i}^{\alpha_i/2d} \geq \prod_{i=1}^n f_{2d,i_0}^{\alpha_i/2d} = f_{2d,i_0},$$

and

$$\begin{aligned}
\sum_{\alpha\in\Delta}\frac{|f_\alpha|\alpha^{\alpha/2d}}{2d\prod_{i=1}^n f_{2d,i}^{\alpha_i/2d}} &= \frac{1}{2d}\sum_{\alpha\in\Delta}\frac{|f_\alpha|\alpha^{\alpha/2d}}{f_{2d,i_0}}\frac{f_{2d,i_0}}{\prod_{i=1}^n f_{2d,i}^{\alpha_i/2d}}\\
&\leq \frac{1}{2d}\sum_{\alpha\in\Delta}\frac{|f_\alpha|\alpha^{\alpha/2d}}{f_{2d,i_0}} \leq 1.
\end{aligned}$$

We note yet another sufficient condition for SOSness.

Corollary 3.1.9. *Let $f \in \mathbb{R}[\underline{X}]$ be a form of degree $2d$. If*

$$f_{2d,i} \geq \sum_{\alpha \in \Delta, \alpha_i \neq 0} \alpha_i \left(\frac{|f_\alpha|}{2d} \right)^{2d/\alpha_i n_\alpha}, \ \ i = 1, \ldots, n$$

then f is SOBS. Here $n_\alpha := |\{i : \alpha_i \neq 0\}|$.

Proof. Defining

$$a_{\alpha,i} = \begin{cases} \alpha_i \left(\frac{|f_\alpha|}{2d}\right)^{2d/\alpha_i n_\alpha} & \text{if } \alpha_i \neq 0 \\ 0 & \text{if } \alpha_i = 0. \end{cases}$$

we have

$$\begin{aligned} 2d \prod_{\alpha_i \neq 0} \left(\frac{a_{\alpha,i}}{\alpha_i} \right)^{\alpha_i/2d} &= 2d \prod_{\alpha_i \neq 0} \left[\left(\frac{|f_\alpha|}{2d} \right)^{2d/n_\alpha \alpha_i} \right]^{\alpha_i/2d} \\ &= 2d \prod_{\alpha_i \neq 0} \left(\frac{|f_\alpha|}{2d} \right)^{1/n_\alpha} \\ &= |f_\alpha|. \end{aligned}$$

Now, Applying Theorem 3.1.3, the conclusion holds. □

The following example shows that the above corollaries are not as strong, either individually or collectively, as Theorem 3.1.3 itself.

Example 3.1.10. Let $f(X,Y,Z) = X^6 + Y^6 + Z^6 - 5X - 4Y - Z + 8$. Corollary 3.1.5 does not apply to f, actually (L1) fails. Also, Corollaries 3.1.6, 3.1.7 and 3.1.9 do not apply to $\bar{f}$, the homogenization of f. We try to apply Theorem 3.1.3. Let $\alpha_1 = (1,0,0,5)$, $\alpha_2 = (0,1,0,5)$ and $\alpha_3 = (0,0,1,5)$, then $\Delta = \{\alpha_1, \alpha_2, \alpha_3\}$. Denote $a_{\alpha_i,j}$ by a_{ij}, we have to find positive reals $a_{11}, a_{22}, a_{33}, a_{14}, a_{24}, a_{34}$ such that the following conditions hold:

$$\begin{aligned} &6^6 a_{11} a_{14}^5 = 5^6 5^5, && 1 \geq a_{11}, \\ &6^6 a_{22} a_{24}^5 = 4^6 5^5, && 1 \geq a_{22}, \\ &6^6 a_{33} a_{34}^5 = 5^5, && 1 \geq a_{33}, \\ &8 \geq a_{14} + a_{24} + a_{34}. \end{aligned}$$

Take $a_{11} = a_{22} = a_{33} = 1$ and solve equations on above set of conditions, we get $a_{14} + a_{24} + a_{34} \approx 7.674 < 8$. This implies that $\bar{f}$ and hence f is SOBS.

3.1.1 Application to PSD linear functionals

As another application of the Hurwitz-Reznick Theorem, we prove that the moments of a PSD linear functional on $\mathbb{R}[\underline{X}]_{2d}$ are bounded by the maximum of the moments of 1 and X_i^{2d}, $i = 1, \ldots, n$. Here, $\mathbb{R}[\underline{X}]_k$ denotes the vector space consisting of polynomials in $\mathbb{R}[\underline{X}]$ of degree $\leq k$. This gives an improvement and simpler proof for the series of lemmas in [39, Section 4].

Corollary 3.1.11. *Suppose* $L : \mathbb{R}[\underline{X}]_{2d} \longrightarrow \mathbb{R}$ *is linear and* $L(p^2) \geq 0 \ \forall \ p \in \mathbb{R}[\underline{X}]_d$. *Then* $|L(\underline{X}^\alpha)| \leq \max\{L(1), L(X_i^{2d}) \ : \ i = 1, \ldots, n\}$, *for all* $\alpha \in \mathbb{N}^n$, $|\alpha| \leq 2d$.

Proof. Suppose $\alpha = (\alpha_1, \cdots, \alpha_n)$, and let $c = \max\{L(1), L(X_i^{2d}) \ : \ i = 1, \ldots, n\}$. Let $f(\underline{X}) = \sum_{i=1}^n \alpha_i X_i^{2d} + (2d - |\alpha|) - 2d\underline{X}^\alpha$ and let

$$\bar{f}(\underline{X}, Y) = \sum_{i=1}^n \alpha_i X_i^{2d} + (2d - |\alpha|)Y^{2d} - 2d\underline{X}^\alpha Y^{2d-|\alpha|},$$

the homogenization of $f(\underline{X})$. By Theorem 3.1.1, $\bar{f}$ is SOBS. Hence f is SOBS, so $L(f) \geq 0$. Note that

$$\begin{array}{rcl} 0 & \leq & L(f) \\ & = & L(\sum_{i=1}^n \alpha_i X_i^{2d} + (2d - |\alpha|) - 2d\underline{X}^\alpha) \\ & = & \sum_{i=1}^n \alpha_i L(X_i^{2d}) + (2d - |\alpha|)L(1) - 2dL(\underline{X}^\alpha) \\ & \leq & \sum_{i=1}^n \alpha_i c + (2d - |\alpha|)c - 2dL(\underline{X}^\alpha) \\ & = & 2dc - 2dL(\underline{X}^\alpha). \end{array}$$

So, $L(\underline{X}^\alpha) \leq c$. If $\alpha = 2\beta$ for some $\beta \in \mathbb{N}^n$ then $\underline{X}^\alpha = (\underline{X}^\beta)^2$, so $L(\underline{X}^\alpha) \geq 0$ and $|L(\underline{X}^\alpha)| \leq c$. Otherwise, α_i is odd for some $1 \leq i \leq n$. In this case, applying the above argument, with f replaced by $\hat{f}(\underline{X}) = f(X_1, \cdots, -X_i, \cdots, X_n)$, we see that $|L(\underline{X}^\alpha)| \leq c$ for each α with $|\alpha| \leq 2d$. □

3.2 Application to global optimization

Let $f \in \mathbb{R}[\underline{X}]$ be a non-constant polynomial of degree $2d$. Recall that f_{sos} denotes the supremum of all real numbers r such that $f - r \in \sum \mathbb{R}[\underline{X}]^2$, f_* denotes the infimum of the set $\{f(\underline{a}) : \underline{a} \in \mathbb{R}^n\}$, and $f_{\text{sos}} \leq f_*$.

Suppose $\underline{f}$ denotes the array of coefficients of non-constant terms of f and f_0 denotes the constant term of f. Suppose $\Phi(\underline{f}, f_0)$ is a formula in terms of coefficients of f such that $\Phi(\underline{f}, f_0)$ implies f is SOS. For such a criterion Φ, we have

$$\forall r \ (\Phi(\underline{f}, f_0 - r) \to \ r \le f_{\text{sos}}),$$

so $f_\Phi := \sup\{r \in \mathbb{R} : \Phi(\underline{f}, f_0 - r)\}$ is a lower bound for f_{sos} and, consequently, for f_*. In this section we develop this idea, using Theorem 3.1.3, to find a new lower bound for f.

Theorem 3.2.1. *Let f be a non-constant polynomial of degree $2d$ and $r \in \mathbb{R}$. Suppose there exist nonnegative real numbers $a_{\alpha,i}$, $\alpha \in \Delta$, $i = 1, \ldots, n$, $a_{\alpha,i} = 0$ if and only if $\alpha_i = 0$, such that*

1. $(2d)^{2d} a_\alpha^\alpha = |f_\alpha|^{2d} \alpha^\alpha$ *for each* $\alpha \in \Delta$ *such that* $|\alpha| = 2d$,
2. $f_{2d,i} \ge \sum_{\alpha\in\Delta} a_{\alpha,i}$ *for* $i = 1, \ldots, n$, *and*
3. $f_0 - r \ge \sum_{\alpha\in\Delta^{<2d}} (2d - |\alpha|) \left[\frac{|f_\alpha|^{2d}\alpha^\alpha}{(2d)^{2d} a_\alpha^\alpha}\right]^{\frac{1}{2d-|\alpha|}}$.

Then $f - r$ is SOBS. Here $\Delta^{<2d} := \{\alpha \in \Delta : |\alpha| < 2d\}$.

Proof. Apply Theorem 3.1.3 to $g := \overline{f - r}$, the homogenization of $f - r$. Since $f = f_0 + \sum_{i=1}^n f_{2d,i} X_i^{2d} + \sum_{\alpha\in\Omega} f_\alpha \underline{X}^\alpha$, it follows that $g = (f_0 - r)Y^{2d} + \sum_{i=1}^n X_i^{2d} + \sum_{\alpha\in\Omega} f_\alpha \underline{X}^\alpha Y^{2d-|\alpha|}$. We know $f - r$ is SOBS if and only if g is SOBS. The sufficient condition for g to be SOBS given by Theorem 3.1.3 is that there exist nonnegative real numbers $a_{\alpha,i}$ and $a_{\alpha,Y}$, $a_{\alpha,i} = 0$ if and only if $\alpha_i = 0$, $a_{\alpha,Y} = 0$ if and only if $|\alpha| = 2d$ such that

$(1)'$ $\forall\ \alpha \in \Delta$ $(2d)^{2d} a_\alpha^\alpha a_{\alpha,Y}^{2d-|\alpha|} = |f_\alpha|^{2d} \alpha^\alpha (2d - |\alpha|)^{2d-|\alpha|}$, and

$(2)'$ $f_{2d,i} \ge \sum_{\alpha\in\Delta} a_{\alpha,i}$, $i = 1, \ldots, n$ and $f_0 - r \ge \sum_{\alpha\in\Delta} a_{\alpha,Y}$.

Solving $(1)'$ for $a_{\alpha,Y}$ yields

$$a_{\alpha,Y} = (2d - |\alpha|) \left[\frac{|f_\alpha|^{2d}\alpha^\alpha}{(2d)^{2d} a_\alpha^\alpha}\right]^{\frac{1}{2d-|\alpha|}},$$

if $|\alpha| < 2d$. Take $a_{\alpha,Y} = 0$ if $|\alpha| = 2d$. Conversely, defining $a_{\alpha,Y}$ in this way, for each $\alpha \in \Delta$, it is easy to see that (1), (2), and (3) imply $(1)'$ and $(2)'$. □

Definition 3.2.2. For a non-constant polynomial f of degree $2d$ we define

$$f_{\text{gp}} := \sup\{r \in \mathbb{R} : \exists a_{\alpha,i} \in \mathbb{R}^+,\ \alpha \in \Delta, i = 1, \ldots, n,\ a_{\alpha,i} = 0 \text{ if and only if } \alpha_i = 0$$
$$\text{satisfying conditions (1), (2) and (3) of Theorem 3.2.1}\}.$$

It follows, as a consequence of Theorem 3.2.1, that $f_{\mathrm{gp}} \le f_{\mathrm{sos}}$.

Example 3.2.3. Let $f(X,Y) = X^4 + Y^4 - X^2Y^2 + X + Y$. Here, $\Delta = \{\alpha_1, \alpha_2, \alpha_3\}$, where $\alpha_1 = (1,0)$, $\alpha_2 = (0,1)$ and $\alpha_3 = (2,2)$. We are looking for nonnegative reals $a_{i,j}$, $i = 1,2,3$, $j = 1,2$ satisfying $a_{11} + a_{21} + a_{31} \le 1$, $a_{12} + a_{22} + a_{32} \le 1$, $a_{31}a_{32} = \frac{1}{4}$. Taking $a_{11} = a_{22} = a_{31} = a_{32} = \frac{1}{2}$, $a_{12} = a_{21} = 0$, we see that $f_{\mathrm{gp}} \ge -\frac{3}{2^{4/3}}$. Taking $X = Y = -\frac{1}{2^{1/3}}$ we see that $f_* \le f(-\frac{1}{2^{1/3}}, -\frac{1}{2^{1/3}}) = -\frac{3}{2^{4/3}}$. Since $f_{\mathrm{gp}} \le f_{\mathrm{sos}} \le f_*$, it follows that $f_{\mathrm{gp}} = f_{\mathrm{sos}} = f_* = -\frac{3}{2^{4/3}}$.

Remark 3.2.4. If $|\Omega| = 1$ then $f_* = f_{\mathrm{sos}} = f_{\mathrm{gp}}$.

Proof. Say $\Omega = \{\alpha\}$, so $f = \sum_{i=0}^n f_{2d,i}X_i^{2d} + f_0 + f_\alpha \underline{X}^\alpha$. We know $f_{\mathrm{gp}} \le f_{\mathrm{sos}} \le f_*$, so it suffices to show that, for each real number r, $f_* \ge r \Rightarrow f_{\mathrm{gp}} \ge r$. Fix r and assume $f_* \ge r$. We want to show $f_{\mathrm{gp}} \ge r$, i.e., that r satisfies the constrains of Theorem 3.2.1. Let g denote the homogenization of $f - r$, i.e., $g = \sum_{i=1}^n f_{2d,i}X_i^{2d} + (f_0 - r)Y^{2d} + f_\alpha \underline{X}^\alpha Y^{2d-|\alpha|}$. Thus g is PSD. This implies, in particular, that $f_{2d,i} \ge 0$, $i = 1, \dots, n$ and $f_0 \ge r$. There are two cases to consider.

Case 1. Suppose $f_\alpha > 0$ and all α_i are even. Then $\alpha \notin \Delta$, so $\Delta = \varnothing$. In this case r satisfies trivially the constraints of Theorem 3.2.1, so $f_{\mathrm{gp}} \ge r$.

Case 2. Suppose either $f_\alpha < 0$ or not all of the α_i are even. Then $\alpha \in \Delta$, i.e., $\Delta = \Omega = \{\alpha\}$. In this case, applying Corollary 3.1.2, we deduce that

$$f_\alpha^{2d} \alpha^\alpha (2d - |\alpha|)^{2d-|\alpha|} \le (2d)^{2d} \prod_{i=1}^n f_{2d,i}^{\alpha_i} (f_0 - r)^{2d-|\alpha|}. \tag{3.2.1}$$

There are two subcases to consider. If $|\alpha| < 2d$ then r satisfies the constraints of Theorem 3.2.1, taking

$$a_{\alpha,i} = \begin{cases} f_{2d,i} & \text{if } \alpha_i \ne 0 \\ 0 & \text{if } \alpha_i = 0. \end{cases}$$

If $|\alpha| = 2d$ then (3.2.1) reduces to $f_\alpha^{2d}\alpha^\alpha \le (2d)^{2d} \prod_{i=1}^n f_{2d,i}^{\alpha_i}$. In this case, r satisfies the constraints of Theorem 3.2.1, taking

$$a_{\alpha,i} = \begin{cases} s f_{2d,i} & \text{if } \alpha_i \ne 0 \\ 0 & \text{if } \alpha_i = 0. \end{cases}$$

where

$$s = \left[\frac{|f_\alpha|^{2d} \alpha^\alpha}{(2d)^{2d} \prod_{i=1}^n f_{2d,i}^{\alpha_i}} \right]^{\frac{1}{|\alpha|}}.$$

□

If $f_{2d,i} > 0$, $i = 1, \dots, n$ then computation of f_{gp} is a geometric programming problem. We explain this now.

Definition 3.2.5. (geometric program)

(1) A function $\phi : \mathbb{R}^n_{>0} \to \mathbb{R}$ of the form

$$\phi(\underline{x}) = c x_1^{a_1} \cdots x_n^{a_n},$$

where $c > 0$, $a_i \in \mathbb{R}$ and $\underline{x} = (x_1, \dots, x_n)$ is called a *monomial function*. A sum of monomial functions, i.e., a function of the form

$$\phi(\underline{x}) = \sum_{i=1}^{k} c_i x_1^{a_{1i}} \cdots x_n^{a_{ni}}$$

where $c_i > 0$ for $i = 1, \dots, k$, is called a *posynomial function*.

(2) An optimization problem of the form

$$\begin{cases} \text{Minimize} & \phi_0(\underline{x}) \\ \text{Subject to} & \phi_i(\underline{x}) \le 1,\ i = 1, \dots, m \text{ and } \ \psi_i(\underline{x}) = 1,\ i = 1, \dots, p \end{cases}$$

where $\phi_0, \dots, \phi_m$ are posynomials and $\psi_1, \dots, \psi_p$ are monomial functions, is called a *geometric program*.

See [11, Section 4.5] or [48, Section 5.3] for detail on geometric programs.

Corollary 3.2.6. *Let f be a non-constant polynomial of degree $2d$ with $f_{2d,i} > 0$, $i = 1, \dots, n$. Then $f_{gp} = f_0 - m^*$ where m^* is the output of the geometric program*

$$\begin{cases} \textit{Minimize} & \sum_{\alpha \in \Delta^{<2d}} (2d - |\alpha|) \left[\left(\frac{f_\alpha}{2d} \right)^{2d} \alpha^\alpha a_\alpha^{-\alpha} \right]^{\frac{1}{2d-|\alpha|}} \\ \textit{Subject to} & \sum_{\alpha \in \Delta} \frac{a_{\alpha,i}}{f_{2d,i}} \le 1,\ i = 1, \cdots, n \ \textit{ and } \ \frac{(2d)^{2d} a_\alpha^\alpha}{|f_\alpha|^{2d} \alpha^\alpha} = 1,\ \alpha \in \Delta,\ |\alpha| = 2d. \end{cases}$$

The variables in the program are the $a_{\alpha,i}$, $\alpha \in \Delta$, $i = 1, \dots, n$, $\alpha_i \ne 0$, the understanding being that $a_{\alpha,i} = 0$ if and only if $\alpha_i = 0$.

Proof. $f_{\mathrm{gp}} = f_0 - m^*$ is immediate from the definition of f_{gp}. Observe that

$$\phi_0(a) := \sum_{\alpha \in \Delta, |\alpha| < 2d} (2d - |\alpha|) \left[\left(\frac{f_\alpha}{2d} \right)^{2d} \alpha^\alpha a_\alpha^{-\alpha} \right]^{\frac{1}{2d-|\alpha|}}$$

and $\phi_i(a) := \sum_{\alpha \in \Delta} \frac{a_{\alpha,i}}{f_{2d,i}}$, $i = 1, \dots, n$ are posynomials in the variables $a_{\alpha,i}$, and $\psi_\alpha(a) := \frac{(2d)^{2d} a_\alpha^\alpha}{|f_\alpha|^{2d} \alpha^\alpha}$, $\alpha \in \Delta$, $|\alpha| = 2d$ are monomial functions in the variables $a_{\alpha,i}$. $\square$

Remark 3.2.7. If either $f_{2d,i} < 0$ for some i or $f_{2d,i} = 0$ and $\alpha_i \neq 0$ for some i and some α then $f_{\text{gp}} = -\infty$. In all remaining cases, after deleting the columns of the array $(a_{\alpha,i})$ corresponding to the indices i such that $f_{2d,i} = 0$, we are reduced to the case where $f_{2d,i} > 0$ for all i, i.e., we can apply geometric programming to compute f_{gp}.

A special case occurs when $f_{2d,i} > 0$, for $i = 1, \dots, n$ and $\{\alpha \in \Delta : |\alpha| = 2d\} = \varnothing$. In this case, the equality constraints in the computation of m^* are vacuous and the feasibility set is always non-empty, so $f_{\text{gp}} \neq -\infty$.

Corollary 3.2.8. *If $|\alpha| < 2d$ for each $\alpha \in \Delta$ and $f_{2d,i} > 0$ for $i = 1, \dots, n$, then $f_{gp} \neq -\infty$ and $f_{gp} = f_0 - m^*$ where m^* is the output of the geometric program*

$$\begin{cases} \textit{Minimize} & \sum_{\alpha\in\Delta}(2d-|\alpha|)\left[\left(\frac{f_\alpha}{2d}\right)^{2d}\alpha^\alpha a_\alpha^{-\alpha}\right]^{\frac{1}{2d-|\alpha|}} \\ \textit{Subject to} & \sum_{\alpha\in\Delta} a_{\alpha,i} \le f_{2d,i}, \quad i = 1, \cdots, n. \end{cases}$$

Proof. Immediate from Corollary 3.2.6. □

Example 3.2.9.

1. Let f be the polynomial of Example 3.1.10. Then $f_{\text{gp}} = f_{\text{sos}} = f_* \approx 0.3265$.

2. For $g(X,Y,Z) = X^6 + Y^6 + Z^6 + X^2YZ^2 - X^4 - Y^4 - Z^4 - YZ^3 - XY^2 + 2$, $g_* \approx 0.667$, and $g_{\text{gp}} = g_{\text{sos}} \approx -1.6728$.

3. For $h(X,Y,Z) = g(X,Y,Z) + X^2$, we have $h_{\text{gp}} \approx -1.6728 < h_{\text{sos}} \approx -0.5028$ and $h_* \approx 0.839$.

To compare the running time efficiency of computation of f_{sos} using SDP with computation of f_{gp} using geometric programming, we set up a test to keep track of the running times. All the polynomials were taken randomly of the form $X_1^{2d} + \cdots + X_n^{2d} + g(\underline{X})$ where $g \in \mathbb{R}[\underline{X}]$ is of degree $\le 2d-1$. In each case the computation is done for 50 polynomial (for the case $n = 6$, $2d = 12$ the algorithm we used to generate random coefficients was taking too long to run, so we used just 10 polynomials instead of 50) with coefficients uniformly distributed in the interval $[-10, 10]$, using SosTools and GPposy for Matlab. The result is shown in Tables 3.1 and 3.2 [1].

[1] **Hardware and Software specifications.** Processor: Intel® Core™2 Duo CPU P8400 @ 2.26GHz, Memory: 2 GB, OS: Ubuntu 11.10-32 bit, Matlab: 7.9.0.529 (R2009b), Sage: 4.6.2

Although, sometimes there is a large gap between f_{sos} and f_{gp}, the running time tables show that computation of f_{gp} is much faster than f_{sos}.

Table 3.1: Average, Minimum and Maximum of $f_{\text{sos}} - f_{\text{gp}}$

n	$2d$	4	6	8	10	12
3	avg	12.4	82.6	204.7	592	1096.3
	min	0	23.5	109.5	311.2	808.6
	max	27.1	141.6	334.5	851.5	1492.8
4	avg	27.5	205.5	730.9	2663.0	6206.1
	min	6.9	100.8	298.3	2098.1	5003.9
	max	51.1	333.1	1044.7	3254.9	7306.3
5	avg	47.9	539.0	2369.0	9599.7	-
	min	19.9	336.6	1823.8	8001.0	-
	max	100.2	763.9	2942.2	11129.7	-
6	avg	84.4	1125.9	5963.1	-	-
	min	36.1	780.3	4637.3	-	-
	max	146.3	1424.1	7421.3	-	-

Table 3.2: Average running time (seconds)

n	$2d$	4	6	8	10	12
3	f_{gp}	0.08	0.09	0.11	0.23	0.33
	f_{sos}	0.73	1.00	1.64	2.81	6.27
4	f_{gp}	0.09	0.13	0.27	0.78	2.16
	f_{sos}	0.96	1.76	5.6	26.14	176.45
5	f_{gp}	0.10	0.23	0.76	3.44	15.41
	f_{sos}	1.42	4.13	45.18	673.63	-
6	f_{gp}	0.11	0.35	2.17	16.54	105.5
	f_{sos}	1.56	13.31	574.9	-	-

Example 3.2.10. Let $f(X,Y,Z) = X^{40} + Y^{40} + Z^{40} - XYZ$. According to Remark 3.2.4, $f_* = f_{\text{sos}} = f_{\text{gp}}$. The running time for computing $f_{\text{gp}} \approx -0.686$ using geometric programming

was 0.18 seconds, but when we attempted to compute f_{sos} directly, using SDP, the machine ran out of memory and halted, after about 4 hours.

Table 3.3 shows the running time for computation of f_{gp} for larger values of n and $2d$ in cases where $|\Omega|$ is relatively small. This computation was done in SAGE [60] using the CVXOPT package, on the same computer, for one (randomly chosen) polynomial in each case.

Table 3.3: Computation time for f_{gp} (seconds) for various sizes of Ω

n	$2d\backslash\|\Omega\|$	10	20	30	40	50	60	70	80	90	100
10	20	0.24	1.4	2.9	10.8	13	31.8	45.6	67.7	121	152
	40	0.28	1.5	4.3	14.8	29	43.5	70.3	133	170	220
	60	0.42	1.6	5.8	15	26	53.5	86.4	129	180	343
20	20	0.69	6.1	13.1	36	71.3	151	180	348	432	659
	40	1.0	7.4	33.6	78.1	154	255	512	749	1033	1461
	60	1.4	12.1	41.3	104	205	451	778	1101	1551	2130
30	20	1.5	9.1	37.6	80.6	153.4	290	462	717	984	1491
	40	4.4	31.3	82.1	175	416	683	1286	2024	3015	3999

3.3 Explicit lower bounds

We explain how the lower bounds for f established in [17, Section 3] can be obtained by evaluating the objective function of the geometric program in Corollary 3.2.8 at suitably chosen feasible points.

Recall that for a (univariate) polynomial of the form $p(t) = t^n - \sum_{i=0}^{n-1} a_i t^i$, where each a_i is nonnegative and at least one a_i is nonzero, $C(p)$ denotes the unique positive root of p [49, Theorem 1.1.3]. See [13], [31, Ex. 4.6.2: 20] or [17, Proposition 1.2] for more details.

Corollary 3.3.1. *If $|\alpha| < 2d$ for each $\alpha \in \Delta$ and $f_{2d,i} > 0$ for $i = 1, \dots, n$, then $f_{gp} \geq r_L$, where*

$$\begin{aligned} r_L &:= f_0 - \tfrac{1}{2d} \textstyle\sum_{\alpha\in\Delta} (2d - |\alpha|) |f_\alpha| k^{|\alpha|} (f_{2d}^{-\alpha})^{\frac{1}{2d}} \\ k &\geq \max_{i=1,\cdots,n} C(t^{2d} - \tfrac{1}{2d} \textstyle\sum_{\alpha\in\Delta} \alpha_i |f_\alpha| f_{2d,i}^{-\frac{|\alpha|}{2d}} t^{|\alpha|}). \end{aligned}$$

Here, $f_{2d}^{-\alpha} := \prod_{i=1}^{n} f_{2d,i}^{-\alpha_i}$.

Proof. For each $\alpha \in \Delta$ and $i = 1, \cdots, n$. Let

$$a_{\alpha,i} = \frac{\alpha_i}{2dk^{2d-|\alpha|}} |f_\alpha| (f_{2d,i})^{1-\frac{|\alpha|}{2d}}.$$

By definition of k, for each i, $\frac{1}{2d} \sum_{\alpha \in \Delta} \alpha_i |f_\alpha| (f_{2d,i})^{-\frac{|\alpha|}{2d}} k^{|\alpha|} \le k^{2d}$, hence

$$\sum_{\alpha \in \Delta} a_{\alpha,i} = \sum_{\alpha \in \Delta} \frac{\alpha_i}{2dk^{2d-|\alpha|}} |f_\alpha| (f_{2d,i})^{1-\frac{|\alpha|}{2d}} \le f_{2d,i}.$$

This shows that the array $(a_{\alpha,i} : \alpha \in \Delta, i = 1, \cdots, n)$ is a feasible point for the geometric program in the statement of Corollary 3.2.8. Plugging this into the objective function of the program yields

$$\begin{aligned}
& \textstyle\sum_{\alpha \in \Delta} (2d - |\alpha|) \left[\left(\frac{f_\alpha}{2d} \right)^{2d} \prod_{\alpha_i \neq 0} \left(\frac{\alpha_i}{a_{\alpha,i}} \right)^{\alpha_i} \right]^{\frac{1}{2d-|\alpha|}} \\
= \;& \textstyle\sum_{\alpha \in \Delta} (2d - |\alpha|) \left[\left(\frac{f_\alpha}{2d} \right)^{2d} \prod_{\alpha_i \neq 0} \left(\frac{2d\alpha_i}{\alpha_i |f_\alpha| k^{|\alpha|-2d}} (f_{2d,i})^{\frac{|\alpha|}{2d}-1} \right)^{\alpha_i} \right]^{\frac{1}{2d-|\alpha|}} \\
= \;& \textstyle\sum_{\alpha \in \Delta} (2d - |\alpha|) \left[\left(\frac{f_\alpha}{2d} \right)^{2d} \prod_{\alpha_i \neq 0} \left(\frac{2d}{|f_\alpha|} k^{2d-|\alpha|} (f_{2d,i})^{\frac{|\alpha|-2d}{2d}} \right)^{\alpha_i} \right]^{\frac{1}{2d-|\alpha|}} \\
= \;& \textstyle\frac{1}{2d} \sum_{\alpha \in \Delta} (2d - |\alpha|) |f_\alpha| k^{|\alpha|} (f_{2d}^{-\alpha})^{\frac{1}{2d}},
\end{aligned}$$

so $r_L = f_0 - \frac{1}{2d} \sum_{\alpha \in \Delta} (2d - |\alpha|) |f_\alpha| k^{|\alpha|} (f_{2d}^{-\alpha})^{\frac{1}{2d}} \le f_{\text{gp}}$. □

Corollary 3.3.2. *If $|\alpha| < 2d$ for each $\alpha \in \Delta$ and $f_{2d,i} > 0$ for $i = 1, \ldots, n$, then $f_{gp} \ge r_{FK}$, where $r_{FK} := f_0 - k^{2d}$, $k \ge C(t^{2d} - \sum_{i=1}^{2d-1} b_i t^i)$,*

$$b_i := \frac{1}{2d} (2d - i)^{\frac{2d-i}{2d}} \sum_{\alpha \in \Delta, |\alpha| = i} |f_\alpha| (\alpha^\alpha f_{2d}^{-\alpha})^{\frac{1}{2d}}, \; i = 1, \ldots, 2d - 1.$$

Proof. Define

$$a_{\alpha,i} := (2d - |\alpha|)^{\frac{2d-|\alpha|}{2d}} \frac{|f_\alpha|}{2d} (\alpha^\alpha f_{2d}^{-\alpha})^{1/2d} f_{2d,i} k^{|\alpha|-2d}.$$

Note that $\sum_{i=1}^{2d-1} b_i k^i \le k^{2d}$ and, for each $i = 1, \ldots, n$,

$$\begin{aligned}
\sum_{\alpha \in \Delta} a_{\alpha,i} &= \sum_{\alpha \in \Delta} (2d - |\alpha|)^{\frac{2d-|\alpha|}{2d}} \frac{|f_\alpha|}{2d} (\alpha^\alpha f_{2d}^{-\alpha})^{\frac{1}{2d}} f_{2d,i} k^{|\alpha|-2d} \\
&= \sum_{j=1}^{2d-1} \sum_{\alpha \in \Delta, |\alpha| = j} (2d - j)^{\frac{2d-j}{2d}} \frac{|f_\alpha|}{2d} (\alpha^\alpha f_{2d}^{-\alpha})^{\frac{1}{2d}} f_{2d,i} k^{j-2d} \\
&= f_{2d,i} \sum_{j=1}^{2d-1} \frac{1}{2d} k^{-2d} k^j (2d - j)^{\frac{2d-j}{2d}} \sum_{\alpha \in \Delta, |\alpha| = j} |f_\alpha| (\alpha^\alpha f_{2d}^{-\alpha})^{\frac{1}{2d}} \\
&= f_{2d,i} k^{-2d} \sum_{j=1}^{2d-1} b_j k^j \\
&\le f_{2d,i}.
\end{aligned}$$

Hence, $(a_{\alpha,i} : \alpha \in \Delta, i = 1, \cdots, n)$ belongs to the feasible set of the geometric program in Corollary 3.2.8. Plugging into the objective function, one sees after some effort that

$$\sum_{\alpha\in\Delta} (2d - |\alpha|) \left[\left(\frac{f_\alpha}{2d} \right)^{2d} \alpha^\alpha a_\alpha^{-\alpha} \right]^{\frac{1}{2d-|\alpha|}} = \sum_{j=1}^{2d-1} b_j k^j \le k^n,$$

so $r_{FK} \le f_{\text{sos}}$. □

Corollary 3.3.3. *If $|\alpha| < 2d$ for each $\alpha \in \Delta$ and $f_{2d,i} > 0$ for $i = 1, \ldots, n$, then*

$$f_{gp} \ge r_{dmt} := f_0 - \sum_{\alpha\in\Delta} (2d - |\alpha|) \left[\left(\frac{f_\alpha}{2d} \right)^{2d} t^{|\alpha|} \alpha^\alpha f_{2d}^{-\alpha} \right]^{\frac{1}{2d-|\alpha|}},$$

where $t := |\Delta|$.

Proof. Take $a_{\alpha,i} = \frac{f_{2d,i}}{t}$ and apply Corollary 3.2.8. □

Let C be a cone in a finite dimensional real vector space V. Let C° denote the interior of C. If $a \in C^\circ$ and $b \in V$ then $b \in C^\circ$ if and only if $b - \epsilon a \in C$ for some real $\epsilon > 0$ (See [44, Lemma 6.1.3] or [17, Remark 2.6]). Since $\sum_{i=1}^n X_i^{2d} \in \Sigma_{2d,n}^\circ$ [17, Corollary 2.5], for a polynomial f of degree $2d$, with $f_{2d} \in \Sigma_{2d,n}^\circ$, there exists an $\epsilon > 0$ such that $g = f_{2d} - \epsilon(\sum_{i=1}^n X_i^{2d}) \in \Sigma_{2d,n}$. The hypothesis of Corollary 3.2.8 holds for $f - g$. In this way, corollaries 3.2.8, 3.3.1, 3.3.2 and 3.3.3, provide lower bounds for f_{sos}. Moreover, the lower bounds obtained in this way, using corollaries 3.3.1, 3.3.2 and 3.3.3, are exactly the lower bounds obtained in [17]. Here we give some approximation for ϵ.

Corollary 3.3.4. *If f is a form of degree $2d$ and $\epsilon := \max\{\epsilon_1, \epsilon_2\} > 0$ where*

$$\epsilon_1 := \min_{i=1,\ldots,n} (f_{2d,i} - \sum_{\alpha\in\Delta} |f_\alpha| \frac{\alpha_i}{2d}), \quad \epsilon_2 := \min_{i=1,\ldots,n} f_{2d,i} - \frac{1}{2d} \sum_{\alpha\in\Delta} |f_\alpha| (\alpha^\alpha)^{\frac{1}{2d}},$$

then $f \in \Sigma_{2d,n}^\circ$ and $f - \epsilon \sum_{i=1}^n X_i^{2d} \in \Sigma_{2d,n}$.

Proof. Applying Theorem 3.1.5 or Theorem 3.1.6 (depending on whether $\epsilon = \epsilon_1$ or $\epsilon = \epsilon_2$) to the form $f - \epsilon \sum_{i=1}^n X_i^{2d}$, we see that $f - \epsilon \sum_{i=1}^n X_i^{2d}$ is SOS. □

Applying Corollary. 3.3.4 to the form f_{2d} allows us to compute ϵ in certain cases: If $\epsilon := \max\{\epsilon_1, \epsilon_2\} > 0$ where

$$\epsilon_1 := \min_{i=1,\ldots,n} (f_{2d,i} - \sum_{\alpha\in\Delta, |\alpha|=2d} |f_\alpha| \frac{\alpha_i}{2d}), \quad \epsilon_2 := \min_{i=1,\ldots,n} f_{2d,i} - \frac{1}{2d} \sum_{\alpha\in\Delta, |\alpha|=2d} |f_\alpha| (\alpha^\alpha)^{\frac{1}{2d}},$$

then $f_{2d} \in \Sigma^{\circ}_{2d,n}$ and $f_{2d} - \epsilon \sum_{i=1}^{n} X_i^{2d} \in \Sigma_{2d,n}$.

The bounds r_L, r_{FK}, r_{dmt} provided by corollaries 3.3.1, 3.3.2 and 3.3.3 are typically not as good as the bound f_{gp} provided by Corollary 3.2.8 .

Example 3.3.5. (Compare to [17, Example 4.2])

(a) For $f(X,Y) = X^6 + Y^6 + 7XY - 2X^2 + 7$, we have $r_L \approx -1.124$, $r_{FK} \approx -0.99$, $r_{dmt} \approx -1.67$ and $f_{\text{sos}} = f_{\text{gp}} \approx -0.4464$, so $f_{\text{gp}} > r_{FK} > r_L > r_{dmt}$.

(b) For $f(X,Y) = X^6 + Y^6 + 4XY + 10Y + 13$, $r_L \approx -0.81$, $r_{FK} \approx -0.93$, $r_{dmt} \approx -0.69$ and $f_{\text{gp}} \approx 0.15 \approx f_{\text{sos}}$, so $f_{\text{gp}} > r_{dmt} > r_L > r_{FK}$.

(c) For $f(X,Y) = X^4 + Y^4 + XY - X^2 - Y^2 + 1$, $f_{\text{sos}} = f_{\text{gp}} = r_L = -0.125$, $r_{FK} \approx -0.832$ and $r_{dmt} \approx -0.875$, so $f_{\text{gp}} = r_L > r_{FK} > r_{dmt}$.

Chapter 4

Exponentially Bounded Functionals

As we mentioned earlier in Section 2.4, in [7] Berg, Christensen and Ressel prove that the closure of the cone of sums of squares $\sum \mathbb{R}[\underline{X}]^2$ in the polynomial ring $\mathbb{R}[\underline{X}]$ in the topology induced by the ℓ_1-norm is equal to $\mathrm{Psd}([-1,1]^n)$. In this chapter, the result is deduced as a corollary of a general result, also established in [7], which is valid for any commutative semigroup. In later work Berg and Maserick [9] and Berg, Christensen and Ressel [8] establish an even more general result, for a commutative semigroup with involution, for the closure of the cone of sums of squares of symmetric elements in the weighted ℓ_1-seminorm topology associated to an absolute value. In this chapter, we give a new proof of these results which is based on Jacobi's Representation Theorem [27]. At the same time, we use Jacobi's theorem to extend these results from sums of squares to sums of $2d$-powers.

4.1 Exponentially Bounded Maps over Polynomials

For any function $\phi : \mathbb{N}^n \longrightarrow \mathbb{R}^+$, we define

$$\mathcal{K}_\phi := \{\underline{x} \in \mathbb{R}^n \ : \ \forall s \in \mathbb{N}^n \ |\underline{x}^s| \le \phi(s)\}.$$

Fix an integer $d \ge 1$. We denote by $M_{\phi,2d}$ the $\sum \mathbb{R}[\underline{X}]^{2d}$-module of $\mathbb{R}[\underline{X}]$ (Definition 1.3.1) generated by the elements $\phi(s) \pm \underline{X}^s$, $s \in \mathbb{N}^n$. $M_{\phi,2d}$ is archimedean. This is a consequence of the fact that

$$\sum_s |f_s|\phi(s) + f = \sum_{f_s>0} |f_s|(\phi(s) + \underline{X}^s) + \sum_{f_s<0} |f_s|(\phi(s) - \underline{X}^s) \in M_{\phi,2d},$$

for any $f = \sum_s f_s \underline{X}^s \in \mathbb{R}[\underline{X}]$. Also, $\mathcal{K}_\phi$ is the non-negativity set of $M_{\phi,2d}$ in $\mathbb{R}^n$ so, by Jacobi's theorem 1.3.11, any $f \in \mathbb{R}[\underline{X}]$ strictly positive on $\mathcal{K}_\phi$ belongs to $M_{\phi,2d}$.[1]

[1] If one insists on $\mathcal{K}_\phi \ne \varnothing$ (equivalently, $-1 \notin M_{\phi,2d}$) it is necessary to assume $\phi(0) \ge 1$.

Definition 4.1.1. A function $\phi : \mathbb{N}^n \longrightarrow \mathbb{R}^+$ is called an *absolute value* if

1. $\phi(0) \geq 1$,

2. $\phi(s+t) \leq \phi(s)\phi(t) \quad \forall s, t \in \mathbb{N}^n$.

Suppose now that ϕ is an absolute value. Denote by $\mathbb{R}[\![\underline{X}]\!]$ the ring of formal power series in $X_1, \ldots, X_n$ with coefficients in $\mathbb{R}$. For $f = \sum_s f_s \underline{X}^s \in \mathbb{R}[\![\underline{X}]\!]$ define the *ϕ-seminorm* of f to be $\|f\|_\phi := \sum_s |f_s|\phi(s)$ and denote by $\mathbb{R}[\![\underline{X}]\!]_\phi$ the subset of $\mathbb{R}[\![\underline{X}]\!]$ consisting of all $f \in \mathbb{R}[\![\underline{X}]\!]$ having finite ϕ-seminorm.

Lemma 4.1.2.

1. $\|f+g\|_\phi \leq \|f\|_\phi + \|g\|_\phi$,

2. $\|rf\|_\phi = |r|\|f\|_\phi$ *and,*

3. $\|fg\|_\phi \leq \|f\|_\phi \|g\|_\phi$.

Proof. (1) and (2) are trivial. To see (3) suppose that $f = \sum_s f_s \underline{X}^s$ and $g = \sum_t g_t \underline{X}^t$ are given. Then $fg = \sum_{s,t} f_s g_t \underline{X}^{s+t}$ and

$$\begin{aligned}
\|f \cdot g\|_\phi &= \|\textstyle\sum_{s,t} f_s g_t \underline{X}^{s+t}\|_\phi \\
&= \|\textstyle\sum_u (\sum_{s+t=u} f_s g_t)\underline{X}^u\|_\phi \\
&= \textstyle\sum_u |\sum_{s+t=u} f_s g_t|\phi(u) \\
&\leq \textstyle\sum_u (\sum_{s+t=u} |f_s g_t|)\phi(s+t) \\
&\leq \textstyle\sum_{s,t} |f_s| \cdot |g_t|\phi(s)\phi(t) \\
&= \|f\|_\phi \|g\|_\phi.
\end{aligned}$$

□

Lemma 4.1.2 implies that $\mathbb{R}[\![\underline{X}]\!]_\phi$ is a subalgebra of the $\mathbb{R}$-algebra $\mathbb{R}[\![\underline{X}]\!]$. It is the closure of $\mathbb{R}[\underline{X}]$ in the topology induced by the ϕ-seminorm.

Lemma 4.1.3. *Suppose* $r \in \mathbb{R}$, $s \in \mathbb{N}^n$, $r > \phi(s)$. *Then* $(r \pm \underline{X}^s)^{1/2d} \in \mathbb{R}[\![\underline{X}]\!]_\phi$.

Proof. We may assume $s \neq 0$. Denote by $\sum_{i=0}^\infty a_i t^i$ the power series expansion of $f(t) = (r \pm t)^{1/2d}$ about $t = 0$, i.e., $a_i = \frac{f^{(i)}(0)}{i!}$. This has radius of convergence r so it converges absolutely for $|t| < r$. In particular, it converges absolutely for $t = \phi(s)$, i.e., $\sum_{i=0}^\infty |a_i|\phi(s)^i < \infty$. Since $\phi(is) \leq \phi(s)^i$ for $i \geq 1$, this implies $\sum_{i=0}^\infty |a_i|\phi(is) < \infty$, i.e., $(r \pm \underline{X}^s)^{1/2d} = \sum_{i=0}^\infty a_i \underline{X}^{is} \in \mathbb{R}[\![\underline{X}]\!]_\phi$. □

An important example of an absolute value, perhaps the most important one, is the constant function 1. If $\phi = 1$ then $\mathcal{K}_\phi = [-1,1]^n$ and the ϕ-seminorm is the standard ℓ_1-norm $\|f\|_1 := \sum_s |f_s|$.

Theorem 4.1.4. *Suppose ϕ is an absolute value on $\mathbb{N}^n$ and $f \in \mathbb{R}[\underline{X}]$, $f > 0$ on $\mathcal{K}_\phi$. Then $f \in \sum \mathbb{R}[\![\underline{X}]\!]_\phi^{2d}$.*

Proof. For each real $\delta > 0$ consider the function $\phi + \delta : \mathbb{N}^n \longrightarrow \mathbb{R}^+$ defined by

$$(\phi + \delta)(s) := \phi(s) + \delta.$$

Since $\bigcap_{\delta>0} \mathcal{K}_{\phi+\delta} = \mathcal{K}_\phi$, each $\mathcal{K}_{\phi+\delta}$ is compact and for any $f > 0$ on $\mathcal{K}_\phi$, $\exists \delta > 0$ such that $f > 0$ on $\mathcal{K}_{\phi+\delta}$. The $\sum \mathbb{R}[\underline{X}]^{2d}$-module $M_{\phi+\delta,2d}$ of $\mathbb{R}[\underline{X}]$ generated by the elements $\phi(s) + \delta \pm \underline{X}^s$, $s \in \mathbb{N}^n$ is archimedean. By Jacobi's theorem 1.3.11, $f \in M_{\phi+\delta,2d}$. By Lemma 4.1.3, $(\phi(s) + \delta \pm \underline{X}^s)^{1/2d} \in \mathbb{R}[\![\underline{X}]\!]_\phi$ for each $s \in \mathbb{N}^n$. Thus $f \in \sum \mathbb{R}[\![\underline{X}]\!]_\phi^{2d}$. □

Corollary 4.1.5. *For any absolute value ϕ on $\mathbb{N}^n$ the closure of the cone $\sum \mathbb{R}[\underline{X}]^{2d}$ in $\mathbb{R}[\underline{X}]$ in the topology induced by the ϕ-seminorm (i.e., $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_\phi}$) is Psd$(\mathcal{K}_\phi)$.*

Proof. The inclusion $(\subseteq)$ follows from continuity of the evaluation map $f \mapsto f(\underline{x})$, for $\underline{x} \in \mathcal{K}_\phi$, which follows in turn from the fact that $|f(\underline{x}) - g(\underline{x})| \le \|f - g\|_\phi$, for $\underline{x} \in \mathcal{K}_\phi$. To prove $(\supseteq)$, suppose $f \in \mathbb{R}[\underline{X}]$, $f \ge 0$ on $\mathcal{K}_\phi$ and $\epsilon > 0$. Then $f + \frac{\epsilon}{2} > 0$ on $\mathcal{K}_\phi$ so $\exists f_1, \dots, f_m \in \mathbb{R}[\![\underline{X}]\!]_\phi$ such that $f + \frac{\epsilon}{2} = f_1^{2d} + \cdots + f_m^{2d}$, by Theorem 4.1.4. Take $g = g_1^{2d} + \cdots + g_m^{2d}$ where $g_i \in \mathbb{R}[\underline{X}]$ is such that $\|f_i^{2d} - g_i^{2d}\|_\phi \le \frac{\epsilon}{2m}$, $i = 1, \dots, m$. Then $g \in \sum \mathbb{R}[\underline{X}]^{2d}$, $\|f - g\|_\phi \le \epsilon$. □

Definition 4.1.6. A linear functional $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ is said to be *exponentially bounded* if there exists an absolute value ϕ, and a constant $C > 0$ such that $\forall s \in \mathbb{N}^n$ $|L(\underline{X}^s)| \le C\phi(s)$.

Corollary 4.1.7. *Suppose $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ is an exponentially bounded linear functional with respect to an absolute value ϕ, such that for all $p \in \mathbb{R}[\underline{X}]$, $L(p^{2d}) \ge 0$. Then there is a unique positive Borel measure $\mu \in \mathcal{M}_+(\mathcal{K}_\phi)$ such that*

$$L(f) = \int_{\mathcal{K}_\phi} f \, d\mu, \quad \forall f \in \mathbb{R}[\underline{X}].$$

Proof. The hypothesis implies that $|L(f) - L(g)| \le C\|f - g\|_\phi$ for some constant $C > 0$, so L is $\|\cdot\|_\phi$-continuous. Fix $f \in \text{Psd}(\mathcal{K}_\phi)$. Fix $\epsilon > 0$. By Corollary 4.1.5, $\exists g \in \sum \mathbb{R}[\underline{X}]^{2d}$ such that $\|f - g\|_\phi \le \epsilon$, so $|L(f) - L(g)| \le C\epsilon$. Since $L(g) \ge 0$, this implies $L(f) \ge -C\epsilon$. Since

$\epsilon > 0$ is arbitrary, this implies $L(f) \geq 0$. The existence follows from Haviland's Theorem 2.2.5. The uniqueness is a result of the fact that $\mathbb{R}[\underline{X}]$ is dense in $\mathrm{C}(\mathcal{K}_\phi)$ and hence L has a unique continuous extension to $\mathrm{C}(\mathcal{K}_\phi)$ by Hahn-Banach Theorem. Therefore by the Riesz Representation Theorem 2.2.3 the representing measure is unique. □

Remark 4.1.8.

1. In the case $d = 1$ Corollary 4.1.7 is well-known. It can be obtained by applying [8, Theorem 4.2.5] to the semigroup $(\mathbb{N}^n, +)$ equipped with the identity involution; see [39, Theorem 2.2]. At the same time, the proof given here is new, even in the case $d = 1$.

2. The converse of Corollary 4.1.7 holds: If $L(f) = \int f \, d\mu$ where $\mu \in \mathcal{M}_+(\mathcal{K}_\phi)$ then $L(p^{2d}) \geq 0$ for all $p \in \mathbb{R}[\underline{X}]$ and $|L(\underline{X}^s)| \leq C\phi(s)$ where $C := \mu(\mathcal{K}_\phi)$. This is clear.

3. We have proved Corollary 4.1.7 from Corollary 4.1.5 using Theorem 2.2.5. One can also prove Corollary 4.1.5 from Corollary 4.1.7 using Corollary 2.1.8. In this way, Corollary 4.1.5 and Corollary 4.1.7 can be seen to carry exactly the same information.

4. Corollary 4.1.5 extends [7, Theorem 9.1].

5. In [39], Lasserre and Netzer use [8, Theorem 4.2.5] to prove that for ϕ equal to the constant function 1 and for any $f \in \mathrm{Psd}(\mathcal{K}_\phi)$ and any real $\epsilon > 0$, and any integer $k \geq 1$ sufficiently large (depending on ϵ and f),

$$f + \epsilon(1 + \sum_{i=1}^{n} X_i^{2k}) \in \sum \mathbb{R}[\underline{X}]^2.$$

   It is not clear how to extend this result with $\sum \mathbb{R}[\underline{X}]^2$ replaced by $\sum \mathbb{R}[\underline{X}]^{2d}$.

4.2 Exponentially Bounded Maps over Semigrouprings

Our goal in this section is to extend Corollary 4.1.5 and Corollary 4.1.7 to arbitrary commutative semigroups with involution; see Theorem 4.2.6 and Corollary 4.2.7.

As in [8, 9], we work with a commutative $*$-semigroup $S = (S, \cdot, 1, *)$ with neutral element 1 and involution $*$. The involution $* : S \longrightarrow S$ satisfies

$$(st)^* = s^* t^*, \ (s^*)^* = s \text{ and } 1^* = 1.$$

We denote by $\mathbb{C}[S]$ the semigroupring of S with coefficients in $\mathbb{C}$. Elements of $\mathbb{C}[S]$ have the form $f = \sum_{s\in S} f_s s$ (finite sum), $f_s \in \mathbb{C}$. $\mathbb{C}[S]$ has the structure of a $\mathbb{C}$-algebra with involution. Addition, scalar multiplication and multiplication are defined by

$$f + g = \sum (f_s + g_s)s,\ \ zf = \sum (zf_s)s,\ \ fg = \sum_{s,t}(f_s g_t)st = \sum_u (\sum_{st=u} f_s g_t)u.$$

The involution is defined by

$$f^* = \sum \overline{f_s} s^*.$$

An element $f \in \mathbb{C}[S]$ is said to be *symmetric* if $f^* = f$, i.e., if $f_{s^*} = \overline{f_s}$ for all $s \in S$. We denote the $\mathbb{R}$-algebra consisting of all symmetric elements of $\mathbb{C}[S]$ by A_S. Clearly

$$\mathbb{C}[S] = A_S \oplus iA_S.$$

As an $\mathbb{R}$-vector space A_S is generated by the elements $s + s^*$ and $i(s - s^*)$, $s \in S$. If the involution on S is the identity, i.e., $s^* = s$ for all $s \in S$, then $A_S = \mathbb{R}[S]$, the semigroupring of S with coefficients in $\mathbb{R}$.

Definition 4.2.1. A *semicharacter* of S is a function $\alpha : S \longrightarrow \mathbb{C}$ satisfying

1. $\alpha(1) = 1$;

2. $\forall s,t \in S\ \alpha(st) = \alpha(s)\alpha(t)$;

3. $\forall s \in S\ \alpha(s^*) = \overline{\alpha(s)}$.

We denote by $\hat{S}$ the set of all semicharacters of S. Semicharacters α of S correspond bijectively to $*$-algebra homomorphisms $\alpha : \mathbb{C}[S] \longrightarrow \mathbb{C}$ via $\alpha(f) := \sum_{s\in S} f_s \alpha(s)$. In turn, $*$-algebra homomorphisms $\alpha : \mathbb{C}[S] \longrightarrow \mathbb{C}$ correspond bijectively to ring homomorphisms $\alpha : A_S \longrightarrow \mathbb{R}$ via $\alpha(f + gi) = \alpha(f) + \alpha(g)i$. In this way, $\hat{S}$ and $\mathcal{X}_{A_S}$ are naturally identified.

For any function $\phi : S \longrightarrow \mathbb{R}^+$ define

$$\mathcal{K}_\phi := \{\alpha \in \hat{S}\ :\ \forall s \in S\ |\alpha(s)| \le \phi(s)\}.$$

Fix an integer $d \ge 1$. Denote by $M_{\phi,2d}$ the $\sum A_S^{2d}$-module of A_S generated by the elements

$$\phi(s)^2 - ss^*,\ 2\phi(s) \pm (s + s^*) \text{ and } 2\phi(s) \pm i(s - s^*),\ s \in S.$$

The following holds for $M_{\phi,2d}$:

Lemma 4.2.2.

1. *$M_{\phi,2d}$ is archimedean.*

2. *The non-negativity set of $M_{\phi,2d}$ in $\hat{S}$ is $\mathcal{K}_\phi$.*

Proof. (1) The elements $s+s^*$, $i(s-s^*)$ generate A_S as an $\mathbb{R}$-vector space and

$$2\phi(s) \pm (s+s^*),\ 2\phi(s) \pm i(s-s^*) \in M_{\phi,2d},$$

so $M_{\phi,2d}$ is archimedean.

(2) For $\alpha \in \hat{S}$,

$$|\alpha(s)| \le \phi(s) \iff \alpha(s)\overline{\alpha(s)} \le \phi(s)^2 \iff \phi(s)^2 - ss^* \ge 0 \text{ at } \alpha.$$

Also, using the inequality $\sqrt{a^2+b^2} \ge \max\{|a|,|b|\}$,

$$|\alpha(s)| \le \phi(s) \Rightarrow |\frac{\alpha(s)+\overline{\alpha(s)}}{2}| \le \phi(s) \iff 2\phi(s) \pm (s+s^*) \ge 0 \text{ at } \alpha,$$

and

$$|\alpha(s)| \le \phi(s) \Rightarrow |\frac{\alpha(s)-\overline{\alpha(s)}}{2i}| \le \phi(s) \iff 2\phi(s) \pm i(s-s^*) \ge 0 \text{ at } \alpha.$$

□

Definition 4.2.3. A function $\phi : S \longrightarrow \mathbb{R}^+$ is called an *absolute value* if

1. $\phi(1) \ge 1$;

2. $\forall s,t \in S\ \phi(st) \le \phi(s)\phi(t)$;

3. $\forall s \in S\ \phi(s^*) = \phi(s)$.

A linear functional $L : \mathbb{C}[S] \longrightarrow \mathbb{C}$ is said to be exponentially bounded with respect to ϕ, if $\forall s \in S\ |L(s)| \le C\phi(s)$ for some $C > 0$.

Suppose that ϕ is an absolute value on S. For $f = \sum_s f_s s \in \mathbb{C}[S]$ define the *ϕ-seminorm* of f to be $\|f\|_\phi := \sum_s |f_s|\phi(s)$.

Lemma 4.2.4.

1. *$\|f+g\|_\phi \le \|f\|_\phi + \|g\|_\phi$,*

2. $\|zf\|_\phi = |z|\|f\|_\phi$,

3. $\|fg\|_\phi \le \|f\|_\phi \|g\|_\phi$,

4. $\|f^*\|_\phi = \|f\|_\phi$.

So the addition, scalar multiplication, multiplication and conjugation in the semigroup algebra $\mathbb{C}[S]$ are continuous in the topology induced by the ϕ-seminorm.

Lemma 4.2.5. *Let $r \in \mathbb{R}$, $f \in A_S$, $r > \|f\|_\phi$. Then, for each real $\epsilon > 0$, there exists $g \in A_S$ such that $\|(r + f) - g^{2d}\|_\phi < \epsilon$.*

Proof. Consider the $\mathbb{R}$-algebra homomorphism $\tau : \mathbb{R}[X] \longrightarrow A_S$ defined by $X \mapsto f$ and consider the absolute value ϕ' on $(\mathbb{N}, +)$ defined by $\phi'(i) = \|f^i\|_\phi$. Applying Lemma 4.1.3 we see that $(r + X)^{1/2d} \in \mathbb{R}[\![X]\!]_{\phi'}$. Combining this with the density of $\mathbb{R}[X]$ in $\mathbb{R}[\![X]\!]_{\phi'}$ and the continuity of the multiplication in the topology induced by the ϕ'-seminorm, there exists $h \in \mathbb{R}[X]$ such that $\|r+X-h^{2d}\|_{\phi'} < \epsilon$. Take $g = \tau(h)$. Since $\tau(r+X-h^{2d}) = r+f-g^{2d}$ and $\|\tau(p)\|_\phi \le \|p\|_{\phi'}$, for all $p \in \mathbb{R}[X]$, this completes the proof. $\square$

Theorem 4.2.6. *Suppose ϕ is an absolute value on a commutative semigroup S with involution and d is any positive integer. Then $\overline{\sum A_S^{2d}}^{\|\cdot\|_\phi} = Psd(\mathcal{K}_\phi)$ in A_S.*

Proof. Since $\sum A_S^{2d} \subseteq \mathrm{Psd}(\mathcal{K}_\phi)$ and $\mathrm{Psd}(\mathcal{K}_\phi)$ is closed, one inclusion in clear. The fact that $\mathrm{Psd}(\mathcal{K}_\phi)$ is closed comes from the fact that each $\alpha \in \mathcal{K}_\phi$, viewed as a ring homomorphism $\alpha : A_S \longrightarrow \mathbb{R}$ in the standard way, satisfies $|\alpha(f)| \le \|f\|_\phi$ for all $f \in A_S$, so α is continuous for each $\alpha \in \mathcal{K}_\phi$, and $\mathrm{Psd}(\mathcal{K}_\phi) = \bigcap_{\alpha \in \mathcal{K}_\phi} \alpha^{-1}(\mathbb{R}^+)$.

For the other inclusion we must show if $f \in \mathrm{Psd}(\mathcal{K}_\phi)$ and $\epsilon > 0$, there exists $g \in \sum A_S^{2d}$ such that $\|f - g\|_\phi \le \epsilon$. Note that $f + \frac{\epsilon}{2}$ is strictly positive at each $\alpha \in \mathcal{K}_\phi$ so, by Lemma 4.2.2 and Jacobi's theorem 1.3.11,

$$f + \frac{\epsilon}{2} = \sum_{i=0}^{k} g_i m_i$$

where $g_i \in \sum A_S^{2d}$, $i = 0, \ldots, k$, $m_0 = 1$, and $m_i \in \{\phi(s)^2 - ss^*,\ 2\phi(s) \pm (s+s^*),\ 2\phi(s) \pm i(s-s^*) : s \in S\}$, $i = 1, \ldots, k$. Choose $\delta > 0$ so that $(\sum_{i=1}^{k} \|g_i\|_\phi)\delta \le \frac{\epsilon}{2}$. By Lemma, 4.2.5 there exists $h_i \in A_S$ such that $\|\frac{\delta}{2} + m_i - h_i^{2d}\|_\phi \le \frac{\delta}{2}$, i.e., $\|m_i - h_i^{2d}\|_\phi \le \delta$, $i = 1, \ldots, k$. Take $g = g_0 + \sum_{i=1}^{k} g_i h_i^{2d}$. Then $g \in \sum A_S^{2d}$, and

$$\|f - g\|_\phi = \|\sum_{i=1}^{k} g_i m_i - \sum_{i=1}^{k} g_i h_i^{2d} - \frac{\epsilon}{2}\|_\phi \le \sum_{i=1}^{k} \|g_i\|_\phi \|m_i - h_i^{2d}\|_\phi + \frac{\epsilon}{2} \le \epsilon.$$

□

Corollary 4.2.7. *Let S be a commutative semigroup with involution and let d be a positive integer. Let $L : \mathbb{C}[S] \longrightarrow \mathbb{C}$ be an exponentially bounded $*$-linear mapping with respect to an absolute value ϕ such that $L(p^{2d}) \geq 0$ for all $p \in A_S$. Then*

$$\exists \mu \in \mathcal{M}_+(\mathcal{K}_\phi) \forall f \in \mathbb{C}[S] \quad L(f) = \int_{\mathcal{K}_\phi} \hat{f}\ d\mu.$$

Here, $\hat{f} : \hat{S} \longrightarrow \mathbb{C}$ is defined by $\hat{f}(\alpha) := \alpha(f)$ for all $\alpha \in \hat{S}$, equivalently, if $f = g + ih$, $g, h \in A_S$, then $\hat{f} := \hat{g} + i\hat{h}$.

Proof. $*$-linear mappings $L : \mathbb{C}[S] \longrightarrow \mathbb{C}$ correspond bijectively to $\mathbb{R}$-linear mappings $L : A_S \longrightarrow \mathbb{R}$, the correspondence being given by $L(f + gi) = L(f) + L(g)i$. The hypothesis implies that $|L(f) - L(g)| \leq C\|f - g\|_\phi$, so L is continuous. Fix $f \in \operatorname{Psd}(\mathcal{K}_\phi)$. Fix $\epsilon > 0$. By Theorem 4.2.6, $\exists g \in \sum A_S^{2d}$ such that $\|f - g\|_\phi \leq \epsilon$, so $|L(f) - L(g)| \leq C\epsilon$. Since $L(g) \geq 0$, this implies $L(f) \geq -C\epsilon$. Since $\epsilon > 0$ is arbitrary, this implies $L(f) \geq 0$. The conclusion follows, by Theorem 2.2.4. □

Remark 4.2.8. For $p \in \mathbb{C}[S]$, $p = q + ir$, $q, r \in A_S$, $pp^* = (q + ir)(q - ir) = q^2 + r^2$. Thus, for $L : \mathbb{C}[S] \longrightarrow \mathbb{C}$, $*$-linear, $L(p^2) \geq 0$ for each $p \in A_S$ if and only if $L(pp^*) \geq 0$ for each $p \in \mathbb{C}[S]$ if and only if L is positive (semi)definite, terminology as in [7], [8] and [9]. Consequently, Corollary 4.2.7 generalizes and provides another proof of what is proved in [7, Corollary 2.5] and [8, Theorem 4.2.5].

4.3 Berg-Maserick Result

In this section we relax the requirement that an absolute value satisfies $\phi(1) \geq 1$. If $\phi(1) < 1$ then, since $\phi(s) = \phi(s1) \leq \phi(s)\phi(1)$ for all $s \in S$, ϕ is identically zero. Then $\|\cdot\|_\phi$ is also identically zero, so the topology on $\mathbb{C}[S]$ is the trivial one and the closure of $\sum A_S^{2d}$ in A_S is A_S. At the same time, $\mathcal{K}_\phi = \varnothing$ so $\operatorname{Psd}(K\phi) = A_S$. Consequently, Theorem 4.2.6 and Corollary 4.2.7 continue to hold in this more general situation.

We explain how the Berg-Maserick result [9, Theorem 2.1] can be deduced as a consequence of Corollary 4.2.7. See Corollary 4.3.4.

Definition 4.3.1. A *weak absolute value* on S is a function $\phi : S \longrightarrow \mathbb{R}^+$ satisfying

$$\forall \ s \in S \ \phi(ss^*) \leq \phi(s)^2.$$

Replacing s by s^*, we see that $\phi(ss^*) \leq \phi(s^*)^2$, so

$$\forall s \in S \ \phi(ss^*) \leq \min\{\phi(s)^2, \phi(s^*)^2\},$$

for any weak absolute value ϕ on S.

For any weak absolute value ϕ on S, define $\phi' : S \longrightarrow \mathbb{R}^+$ by

$$\phi'(s) = \inf\{\prod_{i=1}^{k} \min\{\phi(s_i), \phi(s_i^*)\} \ : \ k \geq 1, \ s_1, \dots, s_k \in S, \ s = s_1 \dots s_k\}.$$

Lemma 4.3.2. *Let ϕ be a weak absolute value on S. Then*

1. *ϕ' is an absolute value (possibly $\phi' \equiv 0$).*

2. *If $L : \mathbb{C}[S] \longrightarrow \mathbb{C}$ is $*$-linear and positive semidefinite and there exists $C > 0$ such that $|L(s)| \leq C\phi(s)$ holds for all $s \in S$, then $\forall s \in S \quad |L(s)| \leq C\phi'(s)$.*

3. *$\mathcal{K}_\phi = \mathcal{K}_{\phi'}$.*

Proof. (1) This is clear. (2) It suffices to show

$$\forall s_1, \dots, s_k \in S \ |L(s_1 \dots s_k)| \leq C \prod_{i=1}^{k} \min\{\phi(s_i), \phi(s_i^*)\}.$$

Since $|L(s)| \leq C\phi(s)$ and $|L(s)| = |\overline{L(s)}| = |L(s^*)| \leq C\phi(s^*)$, the result is clear when $k = 1$. Suppose now that $k \geq 2$. We make use of the Cauchy-Schwartz inequality for the inner product

$$\langle f, g \rangle := L(fg^*), \ f, g \in \mathbb{C}[S].$$

This implies, in particular, that

$$\forall s, t \in S \ |L(st^*)|^2 \leq L(ss^*)L(tt^*).$$

Using this we obtain

$$\begin{array}{rcl} |L(s_1 \dots s_k)|^2 & \leq & L(s_1 s_1^*)L(s_2 s_2^* \dots s_k s_k^*) \\ & \leq & C\phi(s_1 s_1^*)C \prod_{i=2}^{k} \phi(s_i s_i^*) \\ & = & C^2 \prod_{i=1}^{k} \phi(s_i s_i^*) \\ & \leq & C^2 \prod_{i=1}^{k} \min\{\phi(s_i)^2, \phi(s_i^*)^2\} \end{array}$$

(the second inequality by induction on k). The result follows, by taking square roots.

(3) Since $\phi'(s) \leq \phi(s)$ for all $s \in S$, the inclusion $\mathcal{K}_{\phi'} \subseteq \mathcal{K}_\phi$ is clear. For the other inclusion, note that each $\alpha \in \hat{S}$ is positive semidefinite, so $\mathcal{K}_\phi \subseteq \mathcal{K}_{\phi'}$, by (2). □

Corollary 4.3.3. *Suppose ϕ is a weak absolute value on S. Then the closure of $\sum A_S^2$ in A_S in the topology induced by the ϕ-seminorm $\|f\|_\phi := \sum |f_s|\phi(s)$ is equal to Psd($\mathcal{K}_\phi$).*

Proof. Denote the closure of $\sum A_S^2$ in A_S in the topology induced by the ϕ-seminorm by $\overline{\sum A_S^2}^{\|\cdot\|_\phi}$. By Lemma 4.3.2(2), an $\mathbb{R}$-linear map $L : A_S \longrightarrow \mathbb{R}$ non-negative on $\sum A_S^2$ is continuous in the topology induced by $\|\cdot\|_\phi$ if and only if it is continuous in the topology induced by $\|\cdot\|_{\phi'}$. It follows, using Corollary 2.1.8, that $\overline{\sum A_S^2}^{\|\cdot\|_\phi} = \overline{\sum A_S^2}^{\|\cdot\|_{\phi'}}$. By Lemma 4.3.2(3), $\mathcal{K}_\phi = \mathcal{K}_{\phi'}$ so $\mathrm{Psd}(\mathcal{K}_\phi) = \mathrm{Psd}(\mathcal{K}_{\phi'})$. The result follows now using Lemma 4.3.2(1) and Theorem 4.2.6. □

Corollary 4.3.4. *Suppose $L : \mathbb{C}[S] \longrightarrow \mathbb{C}$ is $\star$-linear and positive semidefinite and there exists a weak absolute value ϕ on S and a constant $C > 0$ such that $\forall s \in S$ $|L(s)| \leq C\phi(s)$. Then there exists a unique positive Borel measure $\mu \in \mathcal{M}_+(\mathcal{K}_\phi)$ such that $L(f) = \int_{\mathcal{K}_\phi} \hat{f}\, d\mu$ for each $f \in \mathbb{C}[S]$.*

Proof. In view of Lemma 4.3.2, this is immediate from Corollary 4.2.7. □

Since the argument in Lemma 4.3.2(2) makes essential use of the Cauchy-Schwartz inequality, it seems unlikely that Corollaries 4.3.3 and 4.3.4 extend to the case $d > 1$.

Chapter 5

Closures in Weighted ℓ_p-Norms

In this chapter, we compute the closure of the cone $\sum \mathbb{R}[\underline{X}]^{2d}$ in $\mathbb{R}[\underline{X}]$ under certain norm topologies, $d \geq 1$. Results of this chapter are a slight generalization of results of [16]. We start by reviewing some basic facts about $\|\cdot\|_p$-norms.

5.1 Norm-p topologies

Recall that a real n-sequence, is a function $s : \mathbb{N}^n \longrightarrow \mathbb{R}$. Taking $\delta_\alpha : \mathbb{N}^n \longrightarrow \mathbb{R}$ for each $\alpha \in \mathbb{N}^n$ to be the map such that $\delta_\alpha(\beta) = 1$ if $\beta = \alpha$ and $\delta_\alpha(\beta) = 0$ otherwise, we can assume that each s is equal to the formal sum $s = \sum_{\alpha\in\mathbb{N}^n} s(\alpha)\delta_\alpha$. Let $1 \leq p < \infty$, and define the mapping

$$\|\cdot\|_p : \mathbb{R}^{\mathbb{N}^n} \longrightarrow \mathbb{R} \cup \{\infty\}$$

for each n-sequence s, as follows:

$$\|s\|_p = (\sum_{\alpha\in\mathbb{N}^n} |s(\alpha)|^p)^{\frac{1}{p}} = (\sum_{d=0}^{\infty} \sum_{|\alpha|=d} |s(\alpha)|^p)^{\frac{1}{p}}.$$

For $p = \infty$, define

$$\|s\|_\infty = \sup_{\alpha\in\mathbb{N}^n} |s(\alpha)|.$$

For $1 \leq p \leq \infty$, we let

$$\ell_p(\mathbb{N}^n) = \{s \in \mathbb{R}^{\mathbb{N}^n} : \|s\|_p < \infty\},$$

and

$$c_0(\mathbb{N}^n) = \{s \in \mathbb{R}^{\mathbb{N}^n} : \lim_{\alpha\in\mathbb{N}^n} |s(\alpha)| = 0\},$$

where $\lim_{\alpha\in\mathbb{N}^n} |s(\alpha)| = 0$ means

$$\forall \epsilon > 0\ \exists N > 0\ (|\alpha| > N \Rightarrow |s(\alpha)| < \epsilon).$$

It is well-known that $\|\cdot\|_p$ is a norm on $\ell_p(\mathbb{N}^n)$ and $(\ell_p(\mathbb{N}^n), \|\cdot\|_p)$ forms a Banach space. Moreover, if $1 \leq p < q \leq \infty$ then $\ell_p(\mathbb{N}^n) \subsetneq \ell_q(\mathbb{N}^n)$ and $\ell_p(\mathbb{N}^n) \subsetneq c_0(\mathbb{N}^n) \subsetneq \ell_\infty(\mathbb{N}^n)$. Let V_p be the set of all finite support real n-sequences, equipped with $\|\cdot\|_p$.

Remark 5.1.1. Fixing the monomial basis $\{\underline{X}^\alpha : \alpha \in \mathbb{N}^n\}$, we can identify the space of real polynomials $\mathbb{R}[\underline{X}]$, endowed by $\|\cdot\|_p$-norm, with V_p using the identification $\delta_\alpha \mapsto \underline{X}^\alpha$.

For $1 \leq p \leq \infty$, define the conjugate q of p as follows:

- If $p = 1$, let $q = \infty$,
- If $p = \infty$, let $q = 1$,
- if $1 < p < \infty$, let q be the real number satisfying $\frac{1}{p} + \frac{1}{q} = 1$.

Lemma 5.1.2 (Hölder's inequality)**.** *Let $1 \leq p \leq \infty$ and q be the conjugate of p. Let $a \in \ell_p(\mathbb{N}^n)$ and $b \in \ell_q(\mathbb{N}^n)$. Then $ab \in \ell_1(\mathbb{N}^n)$ and*

$$\|a \cdot b\|_1 \leq \|a\|_p \|b\|_q,$$

where we define $(ab)(\alpha) := a(\alpha)b(\alpha)$ for every $\alpha \in \mathbb{N}^n$.

Proof. See [11, Section 3.1.9] or [29, 2.10.6]. □

Proposition 5.1.3. *Let $1 \leq p \leq \infty$, then*

1. *For $p \neq \infty$, V_p is a dense subspace of $\ell_p(\mathbb{N}^n)$, and V_∞ is dense in $(c_0(\mathbb{N}^n), \|\cdot\|_\infty)$.*
2. *A linear functional L on V_p, $p \neq \infty$, is continuous if and only if $\|(L(\delta_\alpha))_{\alpha \in \mathbb{N}^n}\|_q < \infty$ where q is the conjugate of p.*

Proof. (1) Take an element $s \in \ell_p(\mathbb{N}^n)$, $1 \leq p \leq \infty$ and for each $k \geq 1$, let s_k to be the element of $\ell_p(\mathbb{N}^n)$ defined by $s_k(\alpha) = s(\alpha)$ if $|\alpha| < k$ and $s_k(\alpha) = 0$ for $|\alpha| \geq k$. Take an arbitrary $\epsilon > 0$. For $1 \leq p < \infty$, since $\|s\|_p < \infty$, there exists $N > 0$ such that for $k \geq N$, $\sum_{|\alpha|=k}^{\infty} |s(\alpha)|^p < \epsilon^p$. Therefore $\|s - s_k\|_p < \epsilon$ for each $k > N$ which proves the density of V_p in $\ell_p(\mathbb{N}^n)$ for $1 \leq p < \infty$. For $p = \infty$, by definition, there exists $N > 0$ such that $|s(\alpha)| < \epsilon$, for $|\alpha| > N$. Hence, for every $k > N$, $\|s - s_k\|_\infty < \epsilon$. This completes the proof of (1).

(2) Since V_p is dense in $\ell_p(\mathbb{N}^n)$, every continuous linear functional L on V_p extends continuously to $\ell_p(\mathbb{N}^n)$ in a unique way. So V_p^* corresponds bijectively with $\ell_p(\mathbb{N}^n)^*$. Representing $s \in \ell_p(\mathbb{N}^n)$, as $s = \sum_{\alpha\in\mathbb{N}^n} s(\alpha)\delta_\alpha$, we have

$$\begin{array}{rcl} |L(s)| & = & |\sum_{\alpha\in\mathbb{N}^n} s(\alpha)L(\delta_\alpha)| \\ \text{(by Hölder's inequality)} & \le & \|s\|_p \cdot \|(L(\delta_\alpha))_\alpha\|_q. \end{array}$$

So, if $(L(\delta_\alpha))_\alpha \in \ell_q(\mathbb{N}^n)$, then $L \in \ell_p(\mathbb{N}^n)^*$. This shows that $\ell_q(\mathbb{N}^n) \subseteq \ell_p(\mathbb{N}^n)^*$. It remains to show that if $L \in \ell_p(\mathbb{N}^n)^*$ then $(L(\delta_\alpha))_\alpha \in \ell_q(\mathbb{N}^n)$. Since L is continuous, there exists $c > 0$ such that $|L(s)| \le c \cdot \|s\|_p$. If $p = 1$ then $|L(\delta_\alpha)| \le c \cdot \|\delta_\alpha\|_p = c$, so $(L(\delta_\alpha))_\alpha \in \ell_\infty(\mathbb{N}^n)$. For $1 < p < \infty$ let $s(\alpha) = L(\delta_\alpha)|L(\delta_\alpha)|^{q-2}$ if $L(\delta_\alpha) \ne 0$ and 0 otherwise. Then $|s(\alpha)|^p = |L(\delta_\alpha)|^q = s(\alpha)L(\delta_\alpha)$ for each $\alpha \in \mathbb{N}^n$. Moreover for an integer $d \ge 1$

$$\begin{array}{rcl} \sum_{|\alpha|<d} |L(\delta_\alpha)|^q & = & \sum_{|\alpha|<d} s(\alpha)L(\delta_\alpha) \\ & = & L(\sum_{|\alpha|<d} s(\alpha)\delta_\alpha) \\ & \le & c \cdot (\sum_{|\alpha|<d} |s(\alpha)|^p)^{1/p} \\ & = & c \cdot (\sum_{|\alpha|<d} |L(\delta_\alpha)|^q)^{1/p}. \end{array}$$

Hence for each $d \ge 1$

$$(\sum_{|\alpha|<d} |L(\delta_\alpha)|^q)/(\sum_{|\alpha|<d} |L(\delta_\alpha)|^q)^{1/p} \le c < \infty,$$

which as $d \to \infty$, becomes $\|(L(\delta_\alpha))_\alpha\|_q^q / \|(L(\delta_\alpha))_\alpha\|_q^{q/p} = \|(L(\delta_\alpha))_\alpha\|_q \le c < \infty$, implying $(L(\delta_\alpha))_\alpha \in \ell_q(\mathbb{N}^n)$ as desired. □

Lemma 5.1.4. *For $1 \le p \le q \le \infty$, the identity map $id_{pq} : V_p \longrightarrow V_q$ is continuous.*

Proof. Let $s \in V_p$ with $\|s\|_p = 1$, so $|s(\alpha)|^p \le 1$ for $\alpha \in \mathbb{N}^n$, since $1 \le p \le q$, $|s(\alpha)|^q \le |s(\alpha)|^p$ and hence

$$\sum_{\alpha\in\mathbb{N}^n} |s(\alpha)|^q \le \sum_{\alpha\in\mathbb{N}^n} |s(\alpha)|^p = 1.$$

Therefore $\|s\|_q \le 1$. This proves that id_{pq} is bounded:

$$\|id_{pq}\| = \sup_{\|s\|_p=1} \frac{\|s\|_q}{\|s\|_p} = \sup_{\|s\|_p=1} \|s\|_q \le 1.$$

□

Note that Lemma 5.1.4 implies that for $1 \le p \le q \le \infty$, $\|\cdot\|_p$ induces a finer topology than that induced by $\|\cdot\|_q$ on $\mathbb{R}[\underline{X}]$. Therefore, for $C \subseteq \mathbb{R}[\underline{X}]$, $id_{pq}^{-1}(\overline{C}^{\|\cdot\|_q})$ is closed in $\|\cdot\|_p$-topology. Thus, we have $\overline{C}^{\|\cdot\|_p} \subseteq \overline{C}^{\|\cdot\|_q}$.

In [7, Theorem 9.1], Berg, Christensen and Ressel showed that the closure of $\sum \mathbb{R}[\underline{X}]^2$ in the $\|\cdot\|_1$- topology is $\operatorname{Psd}([-1,1]^n)$. Recently, Lasserre and Netzer [39] revisited this result with a different approach. The proof given by Berg, Christensen and Ressel in [7, 8] is based on techniques from harmonic analysis on semigroups, whereas in [39], Lasserre and Netzer gave a concrete approximation to construct a sequence in $\sum \mathbb{R}[\underline{X}]^2$ for every limit point in $\|\cdot\|_1$. We also gave a different proof of a more general form of this result in Chapter 4, specifically in Corollary 4.1.5. In the following, we determine the closure of sums of $2d$-powers of polynomials in all $\|\cdot\|_p$-topologies which extends the result of Berg, Christensen and Ressel in two directions:

1. Extending the result for $\|\cdot\|_1$ to $\|\cdot\|_p$, for any $1 \le p \le \infty$,

2. Replacing $\sum \mathbb{R}[\underline{X}]^2$ by $\sum \mathbb{R}[\underline{X}]^{2d}$.

We begin by determining which evaluation linear functionals are continuous.

Theorem 5.1.5. *Let $1 \le p \le \infty$ and $\underline{x} \in \mathbb{R}^n$, and let $e_{\underline{x}} : V_p \longrightarrow \mathbb{R}$ be the evaluation homomorphism on V_p defined by $e_{\underline{x}}(f) := f(\underline{x})$. Then the following statements are equivalent:*

1. *$e_{\underline{x}}$ is continuous;*

2. *$\|(\underline{x}^\alpha)_{\alpha\in\mathbb{N}^n}\|_q < \infty$, where q is the conjugate of p;*

3. *$\underline{x} \in (-1,1)^n$ if $1 \le q < \infty$, and $\underline{x} \in [-1,1]^n$ if $q = \infty$.*

Proof. (2)⇔(3) First assume that $1 \le q < \infty$. Let $\underline{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Then

$$\begin{aligned}\|(\underline{x}^\alpha)_{\alpha\in\mathbb{N}^n}\|_q &= (\textstyle\sum_{\alpha\in\mathbb{N}^n} |\underline{x}^\alpha|^q)^{1/q} \\ &= (\textstyle\sum_{\alpha_1,\cdots,\alpha_n=0}^\infty |x_1|^{q\alpha_1} \ldots |x_n|^{q\alpha_n})^{1/q} \\ &= (\textstyle\sum_{\alpha_1=0}^\infty |x_1|^{q\alpha_1})^{1/q} \cdots (\textstyle\sum_{\alpha_n=0}^\infty |x_n|^{q\alpha_n})^{1/q},\end{aligned}$$

where the latter term is a product of geometric series which is finite if and only if $|x_i| < 1$ for $i = 1, \ldots n$. For $q = \infty$,

$$\|(\underline{x}^\alpha)_{\alpha\in\mathbb{N}^n}\|_\infty = \sup_{\alpha\in\mathbb{N}^n} |\underline{x}^\alpha|.$$

Hence $\|(\underline{x}^\alpha)_{\alpha\in\mathbb{N}^n}\|_\infty < \infty$ if and only if $|x_i| \le 1$, for each $1 \le i \le n$.

(1)⇔(2) For $1 \le p < \infty$ it follows from Lemma 5.1.3.2. For $p = \infty$,

$$\begin{aligned}\|e_{\underline{x}}\| = \sup_{\|f\|_\infty=1}|f(\underline{x})| &= \sup_{\|f\|_\infty=1}|\textstyle\sum_{\alpha\in\mathbb{N}^n} f_\alpha \underline{x}^\alpha| \\ &\le \sup_{\|f\|_\infty=1}\textstyle\sum_{\alpha\in\mathbb{N}^n}|f_\alpha|\cdot|\underline{x}^\alpha| \\ &\le \textstyle\sum_{\alpha\in\mathbb{N}^n}|\underline{x}^\alpha| \\ &= \|(\underline{x}^\alpha)_{\alpha\in\mathbb{N}^n}\|_1.\end{aligned}$$

So, if $\|(\underline{x}^\alpha)_{\alpha\in\mathbb{N}^n}\|_1 < \infty$, then $e_{\underline{x}}$ is continuous. Conversely, if $p = \infty$ and $\|(\underline{x}^\alpha)_{\alpha\in\mathbb{N}^n}\|_1 = \infty$, by equivalence of part (2) and part (3), for some $1 \le i \le n$, $|x_i| \ge 1$. For any $k \in \mathbb{N}$, $k \ge 1$, let $f_k(\underline{X}) = \frac{1}{k}(1 + X_i + X_i^2 + \cdots + X_i^k)$ and $g_k(\underline{X}) = \frac{1}{k}(1 - X_i + X_i^2 - \cdots + (-X_i)^k)$. Clearly $f_k, g_k \to 0$ in $\|\cdot\|_\infty$, but

$$\begin{aligned}|e_{\underline{x}}(f_k)| &\ge \tfrac{k+1}{k}, \quad \text{if } x_i \ge 1, \\ |e_{\underline{x}}(g_k)| &\ge \tfrac{k+1}{k}, \quad \text{if } x_i \le -1,\end{aligned}$$

Therefore in either cases at least one of $(e_{\underline{x}}(f_k))$ or $(e_{\underline{x}}(g_k))$ does not converge to 0. Hence, for $\underline{x} \notin (-1,1)^n$, $e_{\underline{x}}$ is not continuous. This completes the proof. □

Theorem 5.1.6. *Let $1 \le p \le \infty$. Then $Psd([-1,1]^n)$ is a closed subset of V_p.*

Proof. We first note that

$$\mathrm{Psd}([-1,1]^n) = \mathrm{Psd}((-1,1)^n) = \bigcap_{\underline{x}\in(-1,1)^n} e_{\underline{x}}^{-1}([0,+\infty)).$$

However, by Theorem 5.1.5(3), for every $\underline{x} \in (-1,1)^n$, $e_{\underline{x}}$ is continuous on V_p. Hence the result follows. □

Theorem 5.1.7. *For $1 \le p \le \infty$ and $d \ge 1$, $\overline{\sum\mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_p} = Psd([-1,1]^n)$.*

Proof. First note that by Theorem 5.1.6, $\mathrm{Psd}([-1,1]^n)$ is closed in V_p, and so,

$$\overline{\sum\mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_p} \subseteq \mathrm{Psd}([-1,1]^n).$$

On the other hand, by Lemma 5.1.4, $id_{1p}^{-1}(\overline{\sum\mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_p})$ is closed in V_1 and contains $\sum\mathbb{R}[\underline{X}]^{2d}$. Hence, by Corollary 4.1.5, it contains $\mathrm{Psd}([-1,1]^n)$. Therefore

$$\mathrm{Psd}([-1,1]^n) = id_{1p}(\mathrm{Psd}([-1,1]^n)) \subseteq \overline{\sum\mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_p}.$$

Thus $\overline{\sum\mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_p} = \mathrm{Psd}([-1,1]^n)$. □

Corollary 2.1.8 and Theorem 5.1.7 have an important consequence for the moment problem: it implies that $\sum \mathbb{R}[\underline{X}]^{2d}$ satisfies the K-moment property for $K = [-1,1]^n$, and $\|\cdot\|_p$-continuous functionals on $\mathbb{R}[\underline{X}]$.

Corollary 5.1.8. *Let $1 \leq p \leq \infty$, $d \geq 1$ an integer, and let $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ be a linear functional on $\mathbb{R}[\underline{X}]$ such that $\|(L(\underline{X}^\alpha))_{\alpha\in\mathbb{N}^n}\|_q < \infty$ where q is the conjugate of p. If $L(h^{2d}) \geq 0$ for all $h \in \mathbb{R}[\underline{X}]$, then there exists a positive Borel measure $\mu \in \mathcal{M}_+([-1,1]^n)$ such that*

$$\forall f \in \mathbb{R}[\underline{X}] \;\; L(f) = \int_{[-1,1]^n} f \; d\mu.$$

Remark 5.1.9. A description for $[-1,1]^n$ is $\{1 - X_i^2 \;:\; i = 1,\ldots,n\}$. Using Schmüdgen's result (Theorem 2.3.2), to verify the moment problem for a linear functional L, one should check the following set of inequalities:

$$\begin{array}{ll}
L(h^2) \geq 0 & \forall h \in \mathbb{R}[\underline{X}], \\
L(h^2(1-X_1^2)) \geq 0 & \forall h \in \mathbb{R}[\underline{X}], \\
\vdots & \\
L(h^2(1-X_1^2)(1-X_2^2)) \geq 0 & \forall h \in \mathbb{R}[\underline{X}], \\
\vdots & \\
L(h^2(1-X_1^2)\cdots(1-X_n^2)) \geq 0 & \forall h \in \mathbb{R}[\underline{X}].
\end{array}$$

Even if we apply Jacobi's result (Corollary 1.3.11), we still need to check the following set of inequalities:

$$\begin{array}{ll}
L(h^2) \geq 0 & \forall h \in \mathbb{R}[\underline{X}], \\
L(h^2(1-X_i^2)) \geq 0 & \forall h \in \mathbb{R}[\underline{X}], \;\; i = 1,\cdots,n.
\end{array}$$

By Corollary 5.1.8, for a *continuous* linear functional on V_p, we are reduced to the single condition $L(h^{2d}) \geq 0$ for every $h \in \mathbb{R}[\underline{X}]$ and some integer $d \geq 1$.

5.2 Weighted Norm-p Topologies

We can extend the result of the preceding section to a more general class of norms known as *weighted norm-p topologies*. Let $r = (r_1,\ldots,r_n)$ be a n-tuple of positive real numbers and $1 \leq p < \infty$. It is easy to check that the vector space

$$\ell_{p,r}(\mathbb{N}^n) := \{s \in \mathbb{R}^{\mathbb{N}^n} : \sum_{\alpha\in\mathbb{N}^n} |s(\alpha)|^p r_1^{\alpha_1}\ldots r_n^{\alpha_n} < \infty\}$$

is a Banach space with respect to the norm

$$\|s\|_{p,r} = (\sum_{\alpha \in \mathbb{N}^n} |s(\alpha)|^p r_1^{\alpha_1} \dots r_n^{\alpha_n})^{\frac{1}{p}}.$$

Also the vector space

$$\ell_{\infty,r}(\mathbb{N}^n) := \{s \in \mathbb{R}^{\mathbb{N}^n} : \sup_{\alpha \in \mathbb{N}^n} |s(\alpha)| r_1^{\alpha_1} \dots r_n^{\alpha_n} < \infty\}$$

is a Banach space with respect to the norm

$$\|s\|_{\infty,r} = \sup_{\alpha \in \mathbb{N}^n} |s(\alpha)| r_1^{\alpha_1} \dots r_n^{\alpha_n}.$$

Moreover, if we let

$$c_{0,r}(\mathbb{N}^n) := \{s \in \mathbb{R}^{\mathbb{N}^n} : \lim_{\alpha \in \mathbb{N}^n} |s(\alpha)| r_1^{\alpha_1} \dots r_n^{\alpha_n} = 0\},$$

then it is straightforward to verify that $c_{0,r}(\mathbb{N}^n)$ is a closed subspace of $\ell_{\infty,r}(\mathbb{N}^n)$ with respect to the norm $\|\cdot\|_{\infty,r}$.

Similar to the case of norm-p topologies, it is essential for us to determine what are the continuous linear functionals on $\ell_{p,r}(\mathbb{N}^n)$.

Lemma 5.2.1. *Let $1 < p < \infty$, and let q be the conjugate of p. Then $\ell_{p,r}(\mathbb{N}^n)^* = \ell_{q,r^{-\frac{q}{p}}}(\mathbb{N}^n)$, $\ell_{1,r}(\mathbb{N}^n)^* = \ell_{\infty,r^{-1}}(\mathbb{N}^n)$, and $c_{0,r}(\mathbb{N}^n)^* = \ell_{1,r^{-1}}(\mathbb{N}^n)$. In either of the cases, the duality is densely defined by*

$$\langle t, k \rangle = \sum_{\alpha \in \mathbb{N}^n} t(\alpha) k(\alpha),$$

for every $t, k \in c_{00}(\mathbb{N}^n) := \{s \in \mathbb{R}^{\mathbb{N}^n} : supp\ s \text{ is finite}\}$.

Proof. Let $1 \le p \le \infty$. The map defined by

$$\begin{array}{rcl} T_{p,r} : \ell_p(\mathbb{N}^n) & \longrightarrow & \ell_{p,r}(\mathbb{N}^n) \\ (s(\alpha))_{\alpha \in \mathbb{N}^n} & \longmapsto & (s(\alpha) r_1^{-\frac{\alpha_1}{p}} \dots r_n^{-\frac{\alpha_n}{p}})_{\alpha \in \mathbb{N}^n} \end{array}$$

is an isometric isomorphism with the inverse $T_{p,r}^{-1} : \ell_{p,r}(\mathbb{N}^n) \longrightarrow \ell_p(\mathbb{N}^n)$ given by

$$T_{p,r}^{-1}((t(\alpha))_{\alpha \in \mathbb{N}^n}) = (t(\alpha) r_1^{\frac{\alpha_1}{p}} \dots r_n^{\frac{\alpha_n}{p}})_{\alpha \in \mathbb{N}^n}.$$

Now suppose that $f \in \ell_{p,r}(\mathbb{N}^n)^*$. Then $f \circ T_{p,r} \in \ell_p(\mathbb{N}^n)^* = \ell_q(\mathbb{N}^n)$. Hence there exist $t \in \ell_q(\mathbb{N}^n)$ such that

$$t = f \circ T_{p,r}.$$

Define the function $t' : \mathbb{N}^n \longrightarrow \mathbb{R}$ by

$$t'(\alpha) = r_1^{\frac{\alpha_1}{p}} \dots r_n^{\frac{\alpha_n}{p}} t(\alpha) \quad (\alpha \in \mathbb{N}^n).$$

It is straightforward to verify that $t' \in \ell_{q, r^{-\frac{q}{p}}}(\mathbb{N}^n)$ if $1 \le p < \infty$, and $t' \in \ell_{\infty, r^{-1}}(\mathbb{N}^n)$ if $p = 1$. Moreover

$$t'(\alpha) = f(\delta_\alpha),$$

where δ_α is the Kroneker symbol at the point $\alpha \in \mathbb{N}^n$. The proof of $c_{0,r}(\mathbb{N}^n)^* = \ell_{1,r^{-1}}(\mathbb{N}^n)$ is similar to the preceding cases. Here the duality we need to consider is $c_0(\mathbb{N}^n)^* = \ell_1(\mathbb{N}^n)$ which is the classical Riesz Representation Theorem. $\square$

Now suppose that $V_{p,r}$ is the set of all finite support real n-sequences, equipped with $\|\cdot\|_{p,r}$. We can naturally identify the space of real polynomials $\mathbb{R}[\underline{X}]$ with $V_{p,r}$. It is straightforward to verify that $V_{p,r}$ is not a Banach space. In fact, similar to Proposition 5.1.3, we can show that the completion of $V_{p,r}$ is exactly $\ell_{p,r}(\mathbb{N}^n)$ when $1 \le p < \infty$ and $c_{0,r}(\mathbb{N}^n)$ when $p = \infty$. Nonetheless, we have enough information on $V_{p,r}$ so that we can characterize the closure of sums of squares in $V_{p,r}$.

Theorem 5.2.2. *Let $1 \le p \le \infty$ and $d \ge 1$ an integer. Then:*

1. *For $1 \le p < \infty$, $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{p,r}} = Psd(\prod_{i=1}^n [-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}])$;*
2. *$\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{\infty,r}} = Psd(\prod_{i=1}^n [-r_i, r_i])$.*

Proof. (1) Suppose that $f \in \mathbb{R}[\underline{X}]$ and $f \ge 0$ on $\prod_{i=1}^n [-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}]$. Since the polynomial $\tilde{f}(\underline{X}) = f(r_1^{\frac{1}{p}} X_1, \cdots, r_n^{\frac{1}{p}} X_n)$ is a nonnegative polynomial on $[-1,1]^n$, by Theorem 5.1.7, there exist a sequence $(g_i)_{i \in \mathbb{N}}$ in $\sum \mathbb{R}[\underline{X}]^{2d}$ which approaches $\tilde{f}$ in $\|\cdot\|_p$. On the other hand,

$$\begin{array}{rcl} \|g_i - \tilde{f}\|_p^p & = & \sum_{\alpha \in \mathbb{N}^n} |g_{i\alpha} - \tilde{f}_\alpha|^p \\ & = & \displaystyle\sum_{\alpha \in \mathbb{N}^n} |g_{i\alpha} - r_1^{\frac{\alpha_1}{p}} \dots r_n^{\frac{\alpha_n}{p}} f_\alpha|^p \\ & = & \displaystyle\sum_{\alpha \in \mathbb{N}^n} r_1^{\alpha_1} \dots r_n^{\alpha_n} |r_1^{\frac{-\alpha_1}{p}} \dots r_n^{\frac{-\alpha_n}{p}} g_{i\alpha} - f_\alpha|^p \\ & = & \|\tilde{g}_i - f\|_{p,r}^p, \end{array}$$

where

$$\tilde{g}_i(\underline{X}) = g_i(r_1^{\frac{-1}{p}} X_1, \dots, r_n^{\frac{-1}{p}} X_n).$$

However $(\tilde{g}_i)_{i\in\mathbb{N}}$ is a sequence of elements of $\sum \mathbb{R}[\underline{X}]^{2d}$. Thus

$$\mathrm{Psd}(\prod_{i=1}^{n}[-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}]) \subseteq \overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{p,r}}.$$

For the converse, we first note that

$$\begin{aligned} \mathrm{Psd}(\textstyle\prod_{i=1}^{n}[-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}]) &= \mathrm{Psd}(\textstyle\prod_{i=1}^{n}(-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}})) \\ &= \bigcap_{\underline{x}\in\prod_{i=1}^{n}(-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}})} e_{\underline{x}}^{-1}([0,+\infty)), \end{aligned}$$

where $e_{\underline{x}}$ is the evaluation map at $\underline{x}$ defined in Theorem 5.1.5. A routine calculation shows that for every $\underline{x} \in \prod_{i=1}^{n}(-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}})$,

$$(\underline{x}^{\alpha})_{\alpha\in\mathbb{N}^n} \in \ell_{\infty,r^{-1}}(\mathbb{N}^n) \text{ if } p = 1,$$

and

$$(\underline{x}^{\alpha})_{\alpha\in\mathbb{N}^n} \in \ell_{q,r^{\frac{-q}{p}}}(\mathbb{N}^n) \text{ if } 1 < p < \infty,$$

where q is the conjugate of p. Therefore it follows from Lemma 5.2.1 that $e_{\underline{x}}$ is continues on $V_{p,r}$. Hence $\mathrm{Psd}(\prod_{i=1}^{n}[-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}])$ is a closed subset of $V_{p,r}$ containing $\sum \mathbb{R}[\underline{X}]^{2d}$. Thus

$$\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{p,r}} \subseteq \mathrm{Psd}(\prod_{i=1}^{n}[-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}]).$$

This completes the proof.

(2) Similar to the argument presented in part (1), we can show that

$$\mathrm{Psd}(\prod_{i=1}^{n}[-r_i, r_i]) \subseteq \overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{\infty,r}}.$$

On the other hand,

$$\begin{aligned} \mathrm{Psd}(\textstyle\prod_{i=1}^{n}[-r_i, r_i]) &= \mathrm{Psd}(\textstyle\prod_{i=1}^{n}(-r_i, r_i)) \\ &= \bigcap_{\underline{x}\in\prod_{i=1}^{n}(-r_i, r_i)} e_{\underline{x}}^{-1}([0,+\infty)), \end{aligned}$$

where again $e_{\underline{x}}$ is the evaluation map. A routine calculation shows that for every $\underline{x} \in \prod_{i=1}^{n}(-r_i, r_i)$, we have $(\underline{x}^{\alpha})_{\alpha\in\mathbb{N}^n} \in \ell_{1,r^{-1}}(\mathbb{N}^n)$. Therefore it follows from Lemma 5.2.1 that $e_{\underline{x}}$ is continues on $V_{\infty,r}$. Hence $\mathrm{Psd}(\prod_{i=1}^{n}[-r_i, r_i])$ is a closed subset of $V_{\infty,r}$ containing $\sum \mathbb{R}[\underline{X}]^{2d}$. Thus

$$\overline{\sum \mathbb{R}[\underline{X}]^{2}}^{\|\cdot\|_{\infty,r}} \subseteq \mathrm{Psd}(\prod_{i=1}^{n}[-r_i, r_i]).$$

The proof is now complete. $\square$

We can now apply the preceding theorem to obtain the K-moment property for $\sum \mathbb{R}[\underline{X}]^{2d}$ for certain convex compact polyhedron and weighted norm-p topologies as we summarize below in the following three theorems:

Theorem 5.2.3. *Let $d \geq 1$ be an integer, $r = (r_1, \ldots, r_n)$ with $r_i > 0$ for $i = 1, \ldots, n$, and let $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ be a linear functional such that the sequence $s(\alpha) = L(\underline{X}^\alpha)$ satisfies*

$$\sup_{\alpha \in \mathbb{N}^n} |s(\alpha)| r_1^{-\alpha_1} \cdots r_n^{-\alpha_n} < \infty.$$

Then $\forall h \in \mathbb{R}[\underline{X}]$ $L(h^{2d}) \geq 0$ if and only if there exists a positive Borel measure $\mu \in \mathcal{M}_+(\prod_{i=1}^n [-r_i, r_i])$ such that

$$\forall f \in \mathbb{R}[\underline{X}] \quad L(f) = \int f \, d\mu.$$

Theorem 5.2.4. *Let $d \geq 1$ be an integer, $1 < p < \infty$, q the conjugate of p, and $r = (r_1, \ldots, r_n)$ with $r_i > 0$ for $i = 1, \ldots, n$. Suppose that $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ is a linear functional such that the sequence $s(\alpha) = L(\underline{X}^\alpha)$ satisfies*

$$\sum_{\alpha \in \mathbb{N}^n} |s(\alpha)|^q r_1^{-\frac{q}{p}\alpha_1} \cdots r_n^{-\frac{q}{p}\alpha_n} < \infty.$$

Then $\forall h \in \mathbb{R}[\underline{X}]$ $L(h^{2d}) \geq 0$ if and only if there exists a positive Borel measure $\mu \in \mathcal{M}_+(\prod_{i=1}^n [-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}])$ such that

$$\forall f \in \mathbb{R}[\underline{X}] \quad L(f) = \int f \, d\mu.$$

Theorem 5.2.5. *Let $d \geq 1$ be an integer, $r = (r_1, \ldots, r_n)$ with $r_i > 0$ for $i = 1, \ldots, n$, and let $L : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ be a linear functional such that the sequence $s(\alpha) = L(\underline{X}^\alpha)$ satisfies*

$$\sum_{\alpha \in \mathbb{N}^n} |s(\alpha)| r_1^{-\alpha_1} \cdots r_n^{-\alpha_n} < \infty.$$

Then $\forall h \in \mathbb{R}[\underline{X}]$ $L(h^{2d}) \geq 0$ if and only if there exists a positive Borel measure $\mu \in \mathcal{M}_+(\prod_{i=1}^n [-r_i, r_i])$ such that

$$\forall f \in \mathbb{R}[\underline{X}] \quad L(f) = \int f \, d\mu.$$

Note that Theorem 5.2.3 can be proved using Corollary 4.1.7. Also theorems 5.2.4 and 5.2.5 are consequences of 5.2.3.

5.3 Coefficient-wise Convergence Topology

In this final section, we characterize the closure of $\sum \mathbb{R}[\underline{X}]^{2d}$ in the *coefficient-wise convergent topology.* A net $(f_i)_{i\in I} \subset \mathbb{R}[\underline{X}]$ converges in the coefficient-wise convergent topology to $f \in \mathbb{R}[\underline{X}]$ if for every $\alpha \in \mathbb{N}^n$, the coefficients of $\underline{X}^\alpha$ in f_i converges to the coefficient of $\underline{X}^\alpha$ in f. It is straightforward to verify that this topology is exactly the locally convex topology generated by the family of seminorms $\pi_\alpha : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}$ defined by

$$\pi_\alpha(f) = |f_\alpha| \quad (f \in \mathbb{R}[\underline{X}], \alpha \in \mathbb{N}^n).$$

In the following lemma, we actually show that this locally convex topology comes from a certain metric topology. For two polynomials $f, g \in \mathbb{R}[\underline{X}]$, let

$$\rho(f,g) = \sum_{\alpha\in\mathbb{N}^n} \frac{|f_\alpha - g_\alpha|}{2^{|\alpha|}(1 + |f_\alpha - g_\alpha|)}. \tag{5.3.1}$$

It is routine to verify that ρ defines a metric on $\mathbb{R}[\underline{X}]$.

Lemma 5.3.1. *Let $(f_i)_{i\in I} \subset \mathbb{R}[\underline{X}]$ be a net and $f \in \mathbb{R}[\underline{X}]$. Then $f_i \xrightarrow{\rho} f$ if and only if $f_i \to f$ in the coefficient-wise convergent topology.*

Proof. Let $\epsilon > 0$ be given. First suppose that $f_{i\alpha} \to f_\alpha$ for each $\alpha \in \mathbb{N}$. Since for each α, $\dfrac{|f_{i\alpha} - f_\alpha|}{1 + (|f_{i\alpha} - f_\alpha|)} < 1$ and

$$\sum_{\alpha\in\mathbb{N}^n} \frac{1}{2^{|\alpha|}} = \sum_{\alpha_1\in\mathbb{N}} \frac{1}{2^{\alpha_1}} \cdots \sum_{\alpha_n\in\mathbb{N}} \frac{1}{2^{\alpha_n}} = 2^n,$$

there exists $N \in \mathbb{N}$ such that

$$\sum_{\alpha\in\mathbb{N}^n, |\alpha|>N} \frac{|f_{i\alpha} - f_\alpha|}{2^{|\alpha|}(1 + |f_{i\alpha} - f_\alpha|)} \le \sum_{\alpha\in\mathbb{N}^n, |\alpha|>N} \frac{1}{2^{|\alpha|}} < \frac{\epsilon}{2},$$

By assumption, for each α with $|\alpha| \le N$, there exists i_α such that

$$|f_{i_\alpha\alpha} - f_\alpha| < \frac{\epsilon}{2D_{|\alpha|}},$$

where $D_{|\alpha|}$ is the number of monomials of degree $|\alpha|$ in n variables. So, for $i \ge \max\{i_\alpha : |\alpha| \le N\}$, we have $\rho(f_i, f) < \epsilon$, therefore $f_i \xrightarrow{\rho} f$.

For the converse, in contrary, suppose that $f_i \xrightarrow{\rho} f$ but for some $\beta \in \mathbb{N}^n$, $f_{i\beta} \not\to f_\beta$. Then, for each $N > 0$, there is $i > n$ such that $|f_{i\beta} - f_\beta| \ge \epsilon$, hence, $\rho(f_i, f) \ge \dfrac{\epsilon}{2^{|\beta|}(1+\epsilon)}$. Thus $\rho(f_i, f) \not\to 0$ which is a contradiction. So for each α, $f_{i\alpha} \to f_\alpha$. $\square$

We can apply the preceding lemma to obtain the main result of this section.

Theorem 5.3.2. *Let $f \in \mathbb{R}[\underline{X}]$ and $d \geq 1$. Then $f(\underline{0}) \geq 0$ if and only if f is coefficient-wise limit of elements of $\sum \mathbb{R}[\underline{X}]^{2d}$.*

Proof. Suppose that $f(0) \geq 0$ and let $\epsilon > 0$ be given. Then for the polynomial $g = f + \frac{\epsilon}{3}$, there exists $0 < r_\epsilon \leq 1$ such that $g \geq 0$ on $[-r_\epsilon, r_\epsilon]^n$ by the continuity of g. So by Theorem 5.2.2, there is a polynomial sequence $(g_i^{(\epsilon)})_{i \in \mathbb{N}} \subset \sum \mathbb{R}[\underline{X}]^{2d}$ such that $\|g_i^{(\epsilon)} - g\|_{1,r_\epsilon} \xrightarrow{i \to \infty} 0$. For a typical element of the sequence we have

$$\rho(g_i^{(\epsilon)}, g) = \sum_{\alpha \in \mathbb{N}^n} \frac{|g_i^{(\epsilon)} - g_\alpha|}{2^{|\alpha|}(1 + |g_i^{(\epsilon)} - g_\alpha|)}.$$

Regardless to what $g_i^{(\epsilon)}$'s and f are, there is $N > 0$ such that

$$\sum_{\alpha \in \mathbb{N}^n, |\alpha| > N} \frac{|g_i^{(\epsilon)} - g_\alpha|}{2^{|\alpha|}(1 + |g_i^{(\epsilon)} - g_\alpha|)} < \frac{\epsilon}{3}.$$

Since $\|g_i^{(\epsilon)} - g\|_{1,r_\epsilon} \xrightarrow{i \to \infty} 0$, one can find sufficiently large i such that

$$\|g_i^{(\epsilon)} - g\|_{1,r_\epsilon} \leq \frac{\epsilon r_\epsilon^N}{3}.$$

Hence

$$\begin{aligned}
\sum_{\alpha \in \mathbb{N}^n, |s| \leq N} \frac{|g_{i\alpha}^{(\epsilon)} - g_\alpha|}{2^{|\alpha|}(1 + |g_{i\alpha}^{(\epsilon)} - g_\alpha|)} &\leq \sum_{\alpha \in \mathbb{N}^n, |\alpha| \leq N} |g_{i\alpha}^{(\epsilon)} - g_\alpha| \\
&= \sum_{\alpha \in \mathbb{N}^n, |\alpha| \leq N} \frac{|g_{i\alpha}^{(\epsilon)} - g_\alpha| r_\epsilon^{|\alpha|}}{r_\epsilon^{|\alpha|}} \\
\text{(Since } r_\epsilon \leq 1) &\leq \frac{\|g_i^{(\epsilon)} - g\|_{1,r_\epsilon}}{r_\epsilon^N} \\
&\leq \frac{\epsilon}{3}.
\end{aligned}$$

Therefore

$$\rho(g_i^{(\epsilon)}, g) \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} = \frac{2\epsilon}{3}.$$

This implies that

$$\rho(g_i^{(\epsilon)}, f) \leq \rho(g_i^{(\epsilon)}, g) + \rho(g, f) < (\frac{2\epsilon}{3}) + \frac{\epsilon}{3} = \epsilon,$$

and so $f \in \overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\rho}$. Thus $\mathrm{Psd}(\{0\}) \subseteq \overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\rho}$. Since $\sum \mathbb{R}[\underline{X}]^{2d} \subseteq \mathrm{Psd}(\{0\})$, the reverse inclusion is clear. □

Remark 5.3.3. One can give a direct proof for Theorem 5.3.2, without identifying the coefficient-wise topology as the topology induced by the metric given on (5.3.1) in the following way. Assume that $f \in \mathrm{Psd}(\{0\})$ is given and let $f_k = \sqrt[2d]{\frac{1}{k} + f}$ which is an element in $\mathbb{R}[\![\underline{X}]\!]$. So it can be written as $f_k = \sum_{\alpha \in \mathbb{N}^n} f_{k,\alpha} \underline{X}^\alpha$. Let $g_k = \sum_{|\alpha| \le k} f_{k,\alpha} \underline{X}^\alpha$. Then for each $\alpha \in \mathbb{N}^n$ with $|\alpha| \le k$ we have

$$(g_k^{2d})_\alpha = \begin{cases} \frac{1}{k} + f_0 & |\alpha| = 0 \\ f_\alpha & 0 < |\alpha| \le k. \end{cases}$$

Therefore for each α, $(g_k^{2d})_\alpha \xrightarrow{k \to \infty} f_\alpha$. So $g_k^{2d} \to f$ as $k \to \infty$ coefficient-wise.

Chapter 6

Closures of $\sum A^{2d}$

In Chapter 2, specially in Section 2.3 we mention several attempts to solve the moment problem of the form

$$\overline{M_S}^{\varphi} = \mathrm{Psd}(\mathcal{K}_S) \text{ or } \overline{T_S}^{\varphi} = \mathrm{Psd}(\mathcal{K}_S),$$

in terms of S. In Section 2.4 we formulate a variant of the moment problem (2.4.1), involving the continuous linear functionals with respect to a locally convex topology instead of any given functional. In Chapter 4, we discussed the moment problem for continuous functionals with respect to a seminorm topology $\|\cdot\|_\phi$ corresponding to an absolute value ϕ and showed that

$$\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_\phi} = \mathrm{Psd}(\mathcal{K}_\phi),$$

which solves (2.4.1) for $\|\cdot\|_\phi$. In Chapter 5, we extend the results of Chapter 4 to the weighted ℓ_p-norm, for a specific family of absolute values. In this chapter, we relax the condition that K be semialgebraic and solve (2.4.1) for $C = \sum \mathbb{R}[\underline{X}]^{2d}$, $K \subseteq \mathcal{X}_{\mathbb{R}[\underline{X}]}$ by finding a suitable locally convex topology τ.

Convention. *Throughout, we always assume that the map* $\hat{\ }: A \longrightarrow \mathrm{C}(\mathcal{X}_A)$ *is injective. Otherwise we factor by the kernel of '$\hat{\ }$', to reduce to this case.*

6.1 The Compact Case

Fix a commutative ring A with unity and assume that $\frac{1}{2} \in A$ and $\mathcal{X}_A \neq \varnothing$. Let K be a compact subset of $\mathcal{X}_A$ in the spectral topology on $\mathcal{X}_A$. By Proposition 1.2.9, the map $\Phi : A \longrightarrow \mathrm{C}(K)$ is a ring homomorphism. Since K is assumed to be compact, $\mathrm{C}(K)$ carries a natural topology, induced by a norm defined by

$$\|f\|_K := \sup_{\alpha \in K} |f(\alpha)|, \quad \forall f \in \mathrm{C}(K).$$

Lemma 6.1.1. *If $K \subseteq \mathcal{X}_A$ is compact, then $\Phi(A)$ is dense in $(\mathrm{C}(K), \|\cdot\|_K)$.*

Proof. Let $\mathcal{A} = \overline{\Phi(A)}^{\|\cdot\|_K}$ in $\mathrm{C}(K)$. We make use of the Stone-Weierstrass Theorem to show that $\mathcal{A} = \mathrm{C}(K)$. K is compact and Hausdorff, so once we show that $\mathcal{A}$ is an $\mathbb{R}$-algebra which contains all constant functions and separates points of K, we are done (See [66, Theorem 44.7]).

Claim 1. $\mathcal{A}$ contains all constant functions.

This is clear. $\mathbb{Z}[\frac{1}{2}] \subset \Phi(A)$ which is dense in $\mathbb{R}$, so $\mathcal{A}$ contains all constant functions.

Claim 2. $\mathcal{A}$ is closed under addition.

Let $f, g \in \mathcal{A}$ and $\epsilon > 0$ be given. Then there are $a, b \in A$ such that $\|f - \hat{a}\|_K < \frac{\epsilon}{2}$ and $\|g - \hat{b}\|_K < \frac{\epsilon}{2}$. Clearly $\|f + g - \hat{a} - \hat{b}\|_K \leq \|f - \hat{a}\|_K + \|g - \hat{b}\|_K < \epsilon$, hence $f + g \in \mathcal{A}$.

Claim 3. $\mathcal{A}$ is closed under multiplication.

Let $f, g \in \mathcal{A}$ and $0 < \epsilon < 1$ be given. There are $a, b \in A$ such that $\|f - \hat{a}\|_K < \epsilon/(\|f\|_K + \|g\|_K + 1)$ and $\|g - \hat{b}\|_K < \epsilon/(\|f\|_K + \|g\|_K + 1)$, so $\|\hat{b}\|_K = \|(\hat{b} - g) + g\|_K < \|g\|_K + \epsilon/(\|f\|_K + \|g\|_K + 1) < \|g\|_K + 1$. Then

$$
\begin{aligned}
\|fg - \hat{a}\hat{b}\|_K &= \|f(g - \hat{b}) + \hat{b}(f - \hat{a})\|_K \\
&\leq \|f\|_K \|g - \hat{b}\|_K + \|\hat{b}\|_K \|f - \hat{a}\|_K \\
&< \|f\|_K \epsilon/(\|f\|_K + \|g\|_K + 1) + (\|g\|_K + 1)\epsilon/(\|f\|_K + \|g\|_K + 1) \\
&= \epsilon.
\end{aligned}
$$

Therefore $fg \in \mathcal{A}$.

Finally, $\mathcal{A}$ separates points of K, because $\Phi(A)$ does. Now by the Stone-Weierstrass Theorem $\mathcal{A} = \mathrm{C}(K)$. □

Definition 6.1.2. A function $\rho : A \longrightarrow \mathbb{R}^+$ is called a (ring) *seminorm* if (1)-(4) hold for all $a, b \in A$, ρ is called a *norm* if in addition (5) also holds.

1. $\rho(0) = 0$,
2. $\rho(a + b) \leq \rho(a) + \rho(b)$,
3. $\rho(-a) = \rho(a)$,
4. $\rho(ab) \leq \rho(a)\rho(b)$,
5. $\rho(a) = 0$ only if $a = 0$.

See [65] for details on topological rings and normed rings. Note that if in addition A is an $\mathbb{R}$-algebra then a submultiplicative seminorm is in particular a ring seminorm.

$\Phi(A)$ as a subspace of $(\mathrm{C}(K), \|\cdot\|_K)$ inherits a topology. Defining $\|\cdot\|_K$ on A by $\|a\|_K = \|\hat{a}\|_K$ induces a topology which is the weakest topology such that Φ is continuous. But $\|\cdot\|_K$ is not necessarily a norm, unless when Φ is injective.

Corollary 6.1.3. *$\|\cdot\|_K$ induces a norm on A if and only if K is a Zariski dense subset of $\mathcal{X}_A$.*

Proof. Since we are assuming that the map $\hat{}: A \longrightarrow \mathrm{C}(\mathcal{X}_A)$ is one-to-one, by Theorem 1.2.11, Φ is injective if and only if K is Zariski dense. So

$$\begin{aligned}
\|a\|_K = 0 \quad &\Leftrightarrow \quad \|\hat{a}\|_K = 0 \\
&\Leftrightarrow \quad \hat{a} = 0 \\
&\Leftrightarrow \quad a \in \ker \Phi \\
&\Leftrightarrow \quad a = 0,
\end{aligned}$$

which completes the proof. □

For any $\alpha \in K$ we define the evaluation map at α, $e_\alpha : \mathrm{C}(K, \|\cdot\|_K) \longrightarrow \mathbb{R}$ by $e_\alpha(f) = f(\alpha)$. We also denote the set of real valued, nonnegative continuous functions over K by $C^+(K)$, i.e.,

$$C^+(K) := \{f \in \mathrm{C}(K) \ : \ f \geq 0 \text{ on } K\}.$$

Also we note that $\mathrm{Psd}(K) := \{a \in A : \hat{a} \in C^+(K)\} = \Phi^{-1}(C^+(K))$.

Proposition 6.1.4.

1. *For each $\alpha \in K$, e_α is $\|\cdot\|_K$-continuous,*

2. *$C^+(K)$ is closed in $(\mathrm{C}(K), \|\cdot\|_K)$,*

3. *$Psd(K)$ is closed in $(A, \|\cdot\|_K)$.*

Proof. (1) Since e_α is linear, it suffices to show that

$$\|e_\alpha\| = \sup_{f \in \mathrm{C}(K)} \frac{|f(\alpha)|}{\|f\|_K} < \infty.$$

But this is clear. Since $|f(\alpha)| \leq \|f\|_K$ we deduce $\|e_\alpha\| \leq 1$.

(2) For each $\alpha \in K$, $e_\alpha^{-1}([0,\infty))$ is closed, by continuity of e_α. Therefore $C^+(K) = \bigcap_{\alpha\in K} e_\alpha^{-1}([0,\infty))$ is closed.

(3) The conclusion follows from the fact that Φ is continuous, $C^+(K)$ is closed and $\mathrm{Psd}(K) = \Phi^{-1}(C^+(K))$. □

Theorem 6.1.5. *For any compact set $K \subseteq \mathcal{X}_A$ and integer $d \geq 1$, $\overline{\sum A^{2d}}^{\|\cdot\|_K} = Psd(K)$.*

Proof. Since $\sum A^{2d} \subseteq \mathrm{Psd}(K)$ and $\mathrm{Psd}(K)$ is closed, clearly $\overline{\sum A^{2d}}^{\|\cdot\|_K} \subseteq \mathrm{Psd}(K)$. To show the reverse inclusion, let $a \in \mathrm{Psd}(K)$ and $\epsilon > 0$ be given. Since $\hat{a} \geq 0$ on K, $\sqrt[2d]{\hat{a}} \in \mathrm{C}(K)$. Continuity of multiplication implies the continuity of the map $f \mapsto f^{2d}$. Therefore, there exists $\delta > 0$ such that $\|\sqrt[2d]{\hat{a}} - f\|_K < \delta$ implies $\|\hat{a} - f^{2d}\|_K < \epsilon$. Using Lemma 6.1.1, there is $b \in A$ such that $\|\sqrt[2d]{\hat{a}} - \hat{b}\|_K < \delta$ and so $\|\hat{a} - \hat{b}^{2d}\|_K < \epsilon$. By definition, $\hat{a} - \hat{b}^{2d} = \Phi(a - b^{2d})$ and hence $\|a - b^{2d}\|_K < \epsilon$. Therefore, any neighbourhood of a has nonempty intersection with $\sum A^{2d}$ which proves the reverse inclusion $\mathrm{Psd}(K) \subseteq \overline{\sum A^{2d}}^{\|\cdot\|_K}$. □

Corollary 6.1.6. *For any compact, Zariski dense subset K of $\mathcal{X}_A$ and $d \geq 1$, there exists a locally convex and Hausdorff topology τ on A such that $\overline{\sum A^{2d}}^{\tau} = Psd(K)$.*

Proof. This is an immediate consequence of theorems 6.1.5 and 6.1.3. □

Corollary 6.1.7. *If K is a compact subset of $\mathcal{X}_A$ and $L : A \longrightarrow \mathbb{R}$ is $\|\cdot\|_K$-continuous $\mathbb{Z}[\frac{1}{2}]$ map, such that $L(a^{2d}) \geq 0$ for all $a, b \in A$, then there exists a unique Borel measure $\mu \in \mathcal{M}_+(K)$ such that $\forall a \in A$ $L(a) = \int_K \hat{a}\, d\mu$.*

Proof. Let $\hat{A} := \{\hat{a} : a \in A\}$ and define $\bar{L} : \hat{A} \longrightarrow \mathbb{R}$ by $\bar{L}(\hat{a}) = L(a)$.

We prove if $\hat{a} \geq 0$, then $L(a) \geq 0$. To see this, let $\epsilon > 0$ be given and find $\delta > 0$ such that $\|a - b\|_K < \delta$ implies $|L(a) - L(b)| < \epsilon$. Take $c_\epsilon \in A$ such that $\|a - c_\epsilon^{2d}\|_K < \delta$. Then

$$L(c_\epsilon^{2d}) - \epsilon < L(a) < L(c_\epsilon^{2d}) + \epsilon,$$

let $\epsilon \to 0$, yields $L(a) \geq 0$.

Note that $\bar{L}$ is well-defined, since $\hat{a} = 0$, implies $\hat{a} \geq 0$ and $-\hat{a} \geq 0$, so $\bar{L}(\hat{a}) \geq 0$ and $\bar{L}(-\hat{a}) \geq 0$, simultaneously and hence $\bar{L}(\hat{a}) = 0$. $\|\cdot\|_K$-continuity of L on A, implies $\|\cdot\|_K$-continuity of $\bar{L}$ on $\hat{A}$. Let $\mathcal{A}$ be the $\mathbb{R}$-subalgebra of $\mathrm{C}(K)$, generated by $\hat{A}$. Elements of $\mathcal{A}$ are of the form $r_1\hat{a}_1 + \cdots + r_n\hat{a}_n$, where $r_i \in \mathbb{R}$ and $a_i \in A$, $i = 1, \cdots, n$. $\bar{L}$ is continuously extensible to $\mathcal{A}$ by $\bar{L}(r\hat{a}) := r\bar{L}(\hat{a})$. By Lemma 6.1.1, $\hat{A}$ and hence $\mathcal{A}$ are dense in the space $(\mathrm{C}(K), \|\cdot\|_K)$. Hahn-Banach Theorem gives a continuous extension of $\bar{L}$ to $\mathrm{C}(K)$. Denoting the extension again by $\bar{L}$, an easy verification shows that $\bar{L}(C^+(K)) \subseteq \mathbb{R}^+$. Applying Riesz Representation Theorem, the result follows. $\square$

6.2 The Non-Compact Case

So far, the compactness of K helped us to define a topology on A such that the equality $\overline{\sum A^2} = \mathrm{Psd}(K)$ holds. If we drop the compactness condition on K, then we are not able to define $\|\cdot\|_K$ anymore. One key step in the described procedure was the continuity of evaluation maps proved in Lemma 6.1.4(1). We now focus on this property.

Definition 6.2.1. To any homomorphism $\alpha \in \mathcal{X}_A$ we associate a seminorm ρ_α on $\mathrm{C}(\mathcal{X}_A)$ by $\rho_\alpha(f) = |f(\alpha)|$ for each $f \in \mathrm{C}(\mathcal{X}_A)$. For any set $K \subset \mathcal{X}_A$ the family of seminorms $T_K = \{\rho_\alpha : \alpha \in K\}$ induces a topology on $\mathrm{C}(\mathcal{X}_A)$. We denote this topology by $\mathcal{F}_K$. The maps $\bar{\rho}_\alpha : a \mapsto \rho_\alpha(\hat{a})$ are seminorms on A, we denote the topology induced on A by $\{\bar{\rho}_\alpha|_A : \alpha \in K\}$ with $\mathcal{T}_K$.

Note that the sets $U_\alpha^\epsilon = \{a : \bar{\rho}_\alpha(a) < \epsilon\}$, form a sub-basis for a fundamental system of neighbourhoods for 0, where $\alpha \in K$ and $\epsilon > 0$. Moreover, the family $\{U^\epsilon_{\alpha_1,\dots,\alpha_m}(x) : m \in \mathbb{N},\ \alpha_i \in K,\ i = 1,\dots,m,\ \epsilon \in \mathbb{R}^{>0},\ x \in A\}$, where

$$U^\epsilon_{\alpha_1,\dots,\alpha_m}(x) = \{a \in A : \bar{\rho}_{\alpha_i}(x - a) < \epsilon, i = 1,\cdots,m\},$$

is a basis for $\mathcal{T}_K$. The following is immediate:

Lemma 6.2.2. *For $\mathcal{X}_A$ and $(A, \mathcal{T}_K)$ we have*

1. *$Psd(K) = Psd(\overline{K})$.*
2. *$Psd(K)$ is closed in $\mathcal{T}_K$.*

Proof. To see (1), note that each $\hat{a}$ is continuous. So $\hat{a} \geq 0$ on K implies $\hat{a} \geq 0$ on $\overline{K}$.
(2) follows from continuity on each $\alpha \in K$ over $(A, \mathcal{T}_K)$ and the fact that $\mathrm{Psd}(K) = \bigcap_{\alpha \in K} \alpha^{-1}([0,\infty))$. $\square$

For the special case $K = \mathbb{R}^n$ and $A = \mathbb{R}[\underline{X}]$, it follows from [57, Proposition 6.2] that the closure of $\sum \mathbb{R}[\underline{X}]^2$ with respect to the finest locally multiplicatively convex topology

η_0 on $\mathbb{R}[\underline{X}]$ is equal to $\text{Psd}(\mathbb{R}^n)$. Since $\mathcal{T}_{\mathbb{R}^n}$ is locally multiplicatively convex and $\text{Psd}(\mathbb{R}^n)$ is closed in $\mathcal{T}_{\mathbb{R}^n}$ we get

$$\text{Psd}(\mathbb{R}^n) = \overline{\sum \mathbb{R}[\underline{X}]^2}^{\eta_0} \subseteq \overline{\sum \mathbb{R}[\underline{X}]^2}^{\mathcal{T}_{\mathbb{R}^n}} \subseteq \overline{\text{Psd}(\mathbb{R}^n)}^{\mathcal{T}_{\mathbb{R}^n}} = \text{Psd}(\mathbb{R}^n).$$

In the next theorem, we show that a similar result holds for arbitrary K and the smaller cone of sums of $2d$-powers $\sum A^{2d} \subset \sum A^2$.

Theorem 6.2.3. *Let $K \subseteq \mathcal{X}_A$ be a closed set and $d \geq 1$, then $\overline{\sum A^{2d}}^{\mathcal{T}_K} = Psd(K)$.*

Proof. Since $\sum A^{2d} \subseteq \text{Psd}(K)$ and $\text{Psd}(K)$ is closed by Lemma 6.2.2, we have $\overline{\sum A^{2d}}^{\mathcal{T}_K} \subseteq \text{Psd}(K)$.

To get the reverse inclusion, let $a \in \text{Psd}(K)$ be given. We show that any neighbourhood of a in $\mathcal{T}_K$ has a non-empty intersection with $\sum A^{2d}$.

Claim. If $\hat{a} > 0$ on K then $a \in \overline{\sum A^{2d}}^{\mathcal{T}_K}$.

To prove this, let U be an open set containing a. There exist $\alpha_1, \ldots, \alpha_n \in K$ and $\epsilon > 0$ such that $a \in U^{\epsilon}_{\alpha_1,\ldots,\alpha_n}(a) \subseteq U$. Chose $m \in \mathbb{N}$ such that $\max_{1\leq i\leq n} \alpha_i(a) < 2^{2dm}$. Now for $b = \frac{a}{2^{2dm}}$ we have $0 < \alpha_i(b) < 1$. By continuity of $f(t) = t^{2d}$, for each $i = 1, \ldots, n$ there exists $\delta_i > 0$ such that for any t, if $|t - \alpha_i(b)^{1/2d}| < \delta_i$, then $|t^{2d} - \alpha_i(b)| < \frac{\epsilon}{2^{2dm}}$. Take $\delta = \min_{1\leq i\leq n} \delta_i$. Let $p(t) = \sum_{j=0}^{N} \lambda_j t^j$ be the real polynomial satisfying $p(\alpha_i(b)) = \sqrt[2d]{\alpha_i(b)}$ for $i = 1, \ldots, n$. Since $\mathbb{Z}[\frac{1}{2}]$ is dense in $\mathbb{R}$ one can choose $\tilde{\lambda}_j \in \mathbb{Z}[\frac{1}{2}]$, such that $|\sum_{j=1}^{N} \lambda_j \alpha_i(b)^j - \sum_{j=1}^{N} \tilde{\lambda}_j \alpha_i(b)^j| < \delta$, for $i = 1, \ldots, n$. Let $c = \sum_{j=1}^{N} \tilde{\lambda}_j b^j \in A$. Then $|\alpha_i(b) - \alpha_i(c)^{2d}| < \frac{\epsilon}{2^{2dm}}$. Multiplying by 2^{2dm}, $|\alpha_i(a) - \alpha_i(2^m c)^{2d}| < \epsilon$ i.e. $\bar{\rho}_{\alpha_i}(a - (2^m c)^{2d}) < \epsilon$ for $i = 1, \ldots, n$. Therefore $U^{\epsilon}_{\alpha_1,\ldots,\alpha_n}(a) \cap \sum A^{2d} \neq \varnothing$ and hence $a \in \overline{\sum A^{2d}}^{\mathcal{T}_K}$ which completes the proof of the claim.

For an arbitrary $a \in \text{Psd}(K)$, and each $k \in \mathbb{N}$, $\widehat{(a + \frac{1}{2^k})} > 0$ on K, so $\forall k \in \mathbb{N}$, $a + \frac{1}{2^k} \in \overline{\sum A^{2d}}^{\mathcal{T}_K}$. Letting $k \to \infty$, $\bar{\rho}_{\alpha}(a + \frac{1}{2^k}) \to \bar{\rho}_{\alpha}(a)$, we get $a \in \overline{\sum A^{2d}}^{\mathcal{T}_K}$ and hence $\text{Psd}(K) \subseteq \overline{\sum A^{2d}}^{\mathcal{T}_K}$ as desired. □

Corollary 6.2.4. *Suppose that K is a closed subset of $\mathcal{X}_A$ such that for some $p \in A$, $K_i = \hat{p}^{-1}([0, i])$ is compact and $L : A \longrightarrow \mathbb{R}$ is $\mathcal{T}_K$-continuous $\mathbb{Z}[\frac{1}{2}]$ map, such that $L(a^{2d}) \geq 0$ for all $a \in A$, then there exists a Borel measure $\mu \in \mathcal{M}_+(K)$ such that*

$$\forall a \in A \;\; L(a) = \int_K \hat{a} \; d\mu.$$

Proof. Following the argument in the proof of Corollary 6.1.7, the map $\bar{L} : \hat{A} \to \mathbb{R}$ is well-defined and has a $\mathcal{F}_K$-continuous extension to the $\mathbb{R}$-subalgebra $\mathcal{A}$ of $\mathrm{C}(K)$, generated by $\hat{A}$. Applying Theorem 2.2.4 to $\bar{L}, \hat{p}$ and $\mathcal{A}$, the result follows. $\square$

According to [8, Proposition 1.8], $\mathcal{T}_K$ is Hausdorff if and only if for every nonzero $a \in A$ there exists $\alpha \in K$ such that $\bar{\rho}_\alpha(a) \neq 0$. But since $\hat{}$ is assumed to be one-to-one, this happens if and only if K is Zariski dense. So we have the following.

Proposition 6.2.5. *$\mathcal{T}_K$ is Hausdorff if and only if K is Zariski dense.*

Let $K \subseteq \mathcal{X}_A$ be compact. We can define two topological structures, $\|\cdot\|_K$ and $\mathcal{T}_K$ on A satisfying $\overline{\sum A^{2d}}^{\|\cdot\|_K} = \overline{\sum A^{2d}}^{\mathcal{T}_K}$. Since every evaluation map e_α for $\alpha \in K$ is continuous, ρ_α is continuous as well and hence every open set in $\mathcal{T}_K$ is also open in $\|\cdot\|_K$-topology. But we show that $\|\cdot\|_K$-topology is strictly finer than $\mathcal{T}_K$.

Lemma 6.2.6. *Let K be an infinite, compact subset of $\mathcal{X}_A$ and suppose that $\alpha_1, \dots, \alpha_m \in K$ and $0 < \delta < \epsilon < 1$ are given. There exists $a \in A$ such that $a \in U^\delta_{\alpha_1,\dots,\alpha_m}(0)$ and $\|a\|_K > \epsilon$.*

Proof. Note that $\mathcal{X}_A$ is Hausdorff and so is K. Compactness of K implies that K is a normal space. Take $C = \{\alpha_1, \dots, \alpha_m\}$ and $D = \{\beta\}$, where $\beta \in K \setminus C$. By Urysohn's lemma, there exists a continuous function $f : K \to [0,1]$ such that $f(C) = 0$ and $f(\beta) = 1$. For $\gamma < \min\{\delta, 1-\epsilon\}$, there exists $a \in A$ such that $\|f - \hat{a}\|_K < \gamma$ by Lemma 6.1.1. Clearly $a \in U^\delta_{\alpha_1,\dots,\alpha_m}$ and $\hat{a}(\beta) > \epsilon$ which implies that $\|a\|_K > \epsilon$ as desired. $\square$

Corollary 6.2.7. *If K is an infinite, compact subset of $\mathcal{X}_A$, then the $\|\cdot\|_K$-topology is strictly finer than $\mathcal{T}_K$*

Proof. Let $N_\epsilon = \{a \in A : \|a\|_K < \epsilon\}$ be an open ball around 0 in $\|\cdot\|_K$-topology for $0 < \epsilon < 1$. We show that N_ϵ does not contain any open neighbourhood of 0 in $\mathcal{T}_K$. In contrary, suppose that $0 \in U^\delta_{\alpha_1,\dots,\alpha_m} \subseteq N_\epsilon$. Obviously $\delta \leq \epsilon$ and by Lemma 6.2.6, there exists $a \in U^\delta_{\alpha_1,\dots,\alpha_m}$ such that $\|a\|_K > \epsilon$ and hence $a \notin N_\epsilon$ which is a contradiction. So, N_ϵ is not open in $\mathcal{T}_K$ and hence, $\|\cdot\|_K$-topology is strictly finer that $\mathcal{T}_K$. $\square$

6.3 Application to the Ring of Polynomials

We are mainly interested in the special case of real polynomials. In this case, $\mathbb{R}[\underline{X}]$ is the free finitely generated commutative $\mathbb{R}$-algebra generated by $X_1, \cdots, X_n$ and hence every

$\alpha \in \mathcal{X}_{\mathbb{R}[\underline{X}]}$ is completely determined by $\alpha(X_i)$, $i = 1, \ldots, n$. So, $\mathcal{X}_{\mathbb{R}[\underline{X}]} = \mathbb{R}^n$ with the usual euclidean topology.

Corollary 6.3.1. *Suppose that K is a compact subset of $\mathbb{R}^n$ which is Zariski dense in $\mathbb{R}^n$. There is a norm topology on $\mathbb{R}[\underline{X}]$ such that $\overline{\sum \mathbb{R}[\underline{X}]^{2d}} = Psd(K)$.*

Proof. The existence of such a topology is a consequence of Theorem 6.1.6. The fact that $\|\cdot\|_K$ is actually a norm, follows from Corollary 6.1.3. □

Remark 6.3.2. According to Theorem 2.1.7, $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_K} = \mathrm{Psd}(K)$ is equivalent to

$$(\sum \mathbb{R}[\underline{X}]^{2d})^{\vee}_{\|\cdot\|_K} = \mathrm{Psd}(K)^{\vee}_{\|\cdot\|_K}.$$

One can restate the previous conclusion in terms of moments in the following way. For a linear functional L on $\mathbb{R}[\underline{X}]$ with $L(p^{2d}) \geq 0$ for all $p \in \mathbb{R}[\underline{X}]$, if there exists a positive real number C such that $\forall s \in \mathbb{N}^n$ $|L(\underline{X}^s)| \leq C\|\underline{X}^s\|_K$, then there exists a Borel measure $\mu \in \mathcal{M}_+(K)$, representing L, i.e.

$$\forall f \in \mathbb{R}[\underline{X}]\ L(f) = \int_K f\ d\mu.$$

It worth noting that for a functional L, being positive semi-definite is equivalent to the condition $L(p^2) \geq 0$ for all $p \in \mathbb{R}[\underline{X}]$, and hence, $L(p^{2d}) \geq 0$ for $d \geq 1$. So, it is immediate that every $\|\cdot\|_K$-continuous positive semi-definite functional on $\mathbb{R}[\underline{X}]$, comes from a Borel measure on K.

6.3.1 Comparison with other topologies

In this section we compare the topologies $\|\cdot\|_K$, $\mathcal{T}_K$ of this chapter with other topologies studied in Chapter 4 and the one introduced by Lasserre in [38]. Recall that for an absolute value ϕ on $\mathbb{N}^n$, since $M_{\phi,2d}$ is archimedean (Lemma 4.2.2), the set $\mathcal{K}_\phi$ is compact by Theorem 1.3.10, so $\|\cdot\|_{\mathcal{K}_\phi}$ is defined and we have the following:

Proposition 6.3.3. *The $\|\cdot\|_\phi$-topology is finer than the $\|\cdot\|_{\mathcal{K}_\phi}$-topology and*

$$\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_\phi} = \overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{\mathcal{K}_\phi}} = Psd(\mathcal{K}_\phi),$$

for every $d \geq 1$.

Proof. To prove that $\|\cdot\|_\phi$-topology is finer than $\|\cdot\|_{\mathcal{K}_\phi}$-topology, we show $\|f\|_{\mathcal{K}_\phi} \le \|f\|_\phi$. Note that $\mathcal{K}_\phi$ is compact by Lemma 4.2.2, so

$$\begin{array}{rcl} \|f\|_{\mathcal{K}_\phi} & = & \sup_{x\in\mathcal{K}_\phi} |f(x)| \\ & = & \sup_{x\in\mathcal{K}_\phi} |\sum f_s x^s| \\ & \le & \sup_{x\in\mathcal{K}_\phi} \sum |f_s|\cdot|x^s| \ \cdot \\ & \le & \sum |f_s|\phi(s) \\ & = & \|f\|_\phi. \end{array}$$

Therefore, the identity map $id : (\mathbb{R}[\underline{X}], \|\cdot\|_\phi) \longrightarrow (\mathbb{R}[\underline{X}], \|\cdot\|_{\mathcal{K}_\phi})$ is continuous and so $\|\cdot\|_\phi$-topology is finer than $\|\cdot\|_{\mathcal{K}_\phi}$-topology. Moreover, $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{\mathcal{K}_\phi}} = \mathrm{Psd}(\mathcal{K}_\phi)$ by Corollary 6.3.1. $\square$

Recently, Lasserre [38] proved that there exists a norm $\|\cdot\|_w$ on $\mathbb{R}[\underline{X}]$ such that for any basic semialgebraic set $K \subseteq \mathbb{R}^n$, defined by a finite set of polynomials S, the closure of the quadratic module M_S and the preordering T_S with respect to $\|\cdot\|_w$ are equal to $\mathrm{Psd}(K)$. The $\|\cdot\|_w$ is explicitly defined by

$$\|\sum_{s\in\mathbb{N}^n} f_s \underline{X}^s\|_w = \sum_{s\in\mathbb{N}^n} |f_s| w(s),$$

where $w(s) = (2\lceil |s|/2 \rceil)!$ and $|s| = |(s_1,\ldots,s_n)| = s_1 + \cdots + s_n$.

Proposition 6.3.4. *For any compact basic semi-algebraic set $\mathcal{K}_S \subset \mathbb{R}^n$, the $\|\cdot\|_w$-topology is finer than $\|\cdot\|_{\mathcal{K}_S}$-topology and*

$$\overline{M_S}^{\|\cdot\|_w} = \overline{T_S}^{\|\cdot\|_w} = \overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{\mathcal{K}_S}} = Psd(\mathcal{K}_S),$$

where $d \ge 1$.

Proof. To show that $\|\cdot\|_w$-topology is finer than $\|\cdot\|_{\mathcal{K}_S}$-topology, it suffices to prove that the formal identity map

$$id : (\mathbb{R}[\underline{X}], \|\cdot\|_w) \longrightarrow (\mathbb{R}[\underline{X}], \|\cdot\|_{\mathcal{K}_S})$$

is continuous. Let $p_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ be the projection on ith coordinate and

$$M = \max_{1\le i\le n} \{|p_i(x)| : x \in \mathcal{K}_S\}.$$

So, for each $s \in \mathbb{N}^n$ we have $\|\underline{X}^s\|_{\mathcal{K}_S} \le M^{|s|}$. Also $w(s) \ge |s|!$ for all $s \in \mathbb{N}^n$. By Stirling's formula $|s|! \sim \sqrt{2\pi} e^{(|s|+\frac{1}{2})\ln|s|-|s|}$, we see that

$$\begin{aligned} \frac{\|\underline{X}^s\|_{\mathcal{K}_S}}{\|\underline{X}^s\|_w} &\le \frac{M^{|s|}}{|s|!} \\ &\sim \frac{1}{\sqrt{2\pi}} e^{|s|(\ln M - \ln|s| + 1) - \frac{1}{2}\ln|s|} \xrightarrow{|s|\to\infty} 0. \end{aligned}$$

Therefore for some $N \in \mathbb{N}$, if $|s| > N$ then $\frac{\|\underline{X}^s\|_{\mathcal{K}_S}}{\|\underline{X}^s\|_w} < 1$, which shows that *id* is bounded and hence continuous. The asserted equality follows from Corollary 6.1.5 and [38, Theorem 3.3]. □

Proposition 6.3.5. *For any basic semialgebraic set $\mathcal{K}_S \subset \mathbb{R}^n$, the $\|\cdot\|_w$-topology is finer than $\mathcal{T}_{\mathcal{K}_S}$ and*

$$\overline{M_S}^{\|\cdot\|_w} = \overline{T_S}^{\|\cdot\|_w} = \overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\mathcal{T}_{\mathcal{K}_S}} = Psd(\mathcal{K}_S),$$

where $d \ge 1$.

Proof. It suffices to show that for any $x \in \mathcal{K}_S$, the evaluation map $e_x(f) = f(x)$ is $\|\cdot\|_w$-continuous. Since $\frac{|x^s|}{|s|!} \xrightarrow{|s|\to\infty} 0$, we deduce that $\sup_{f\in\mathbb{R}[\underline{X}]} \frac{|f(x)|}{\|f\|_w}$ is bounded. So, e_x and hence ρ_x is $\|\cdot\|_w$-continuous. Therefore any basic open set in $\mathcal{T}_{\mathcal{K}_S}$ is $\|\cdot\|_w$-open. The asserted equality follows from Theorem 6.2.3 and [38, Theorem 3.3]. □

The following diagram shows the relation of all different topologies we discussed. D is a closed set, K is compact and ϕ is an absolute value. The left side diagram demonstrates topologies and their relations, while the right side consists of the corresponding closures for $\sum \mathbb{R}[\underline{X}]^2$. The arrows show the inclusion in both.

$$\begin{tikzcd}
 & \varphi & \\
 & \|\cdot\|_w \arrow[u] & \\
\mathcal{T}_D \arrow[ur] & & \\
\mathcal{T}_K \arrow[rr] & & \|\cdot\|_K \arrow[uul] \\
\mathcal{T}_{\mathcal{K}_\phi} \arrow[r] & \|\cdot\|_{\mathcal{K}_\phi} \arrow[r] & \|\cdot\|_\phi \arrow[uuul]
\end{tikzcd}
\qquad
\begin{tikzcd}
 & \sum \mathbb{R}[\underline{X}]^2 \arrow[d] & \\
 & \mathrm{Psd}(\mathbb{R}^n) \arrow[dl] \arrow[dd] \arrow[ddr] & \\
\mathrm{Psd}(D) & & \\
 & & \mathrm{Psd}(K) \\
 & \mathrm{Psd}(\mathcal{K}_\phi) &
\end{tikzcd}$$

Summary and Concluding Remarks

This thesis provides some new advancements in two aspects of real algebraic geometry related to the moment problem:

1. SOS representation,
2. SOS approximation.

SOS representation

It is known that identifying whether a polynomial $f \in \mathbb{R}[\underline{X}]$ is PSD is a NP-hard problem. A popular simplification to this problem is considering SOS polynomials instead of PSD ones which is solvable in polynomial time using SDP method. But in practice SDP method is only feasible for small n (number of variables) and small $2d$ (degree of the polynomial). In Chapter 3 Theorem 3.1.3, we applied a result of Hurwitz and Reznick to give a new sufficient condition for a polynomial to be a sum of squares. This sufficient condition is applied to find a lower bound of an even degree polynomial in Theorem 3.2.1. The method described in Section 3.2 associates a geometric program to f which for its optimum value m^*, $f_{\text{gp}} = f_0 - m^*$ is a lower bound for f on $\mathbb{R}^n$.

A computer program implementing this method is designed to measure the performance of computational aspects of f_{gp} compare to f_{sos}. The results are given in tables 3.1, 3.2 and 3.3. Appendix C contains an early version of the program written for Sage and the latest version of the code is available at http://goo.gl/iI3Y0.

The method of SDP is easy to extend to constrained optimization problems over basic closed semialgebraic sets, i.e., the problems of type

$$\begin{cases} \text{Minimize} & f(\underline{X}) \\ \text{Subject to} & g_i(\underline{X}) \geq 0 \quad i = 1, \ldots, s \end{cases} \tag{1}$$

can be solved using semidefinite programming. At the moment it is not known to the us whether there exists a generalization of the method of Chapter 3 to solve (1) optimzation problems or not. However, there exist some progress on approximating a lower for problems of the special type

$$\begin{cases} \text{minimize} & f(\underline{X}) \\ \text{subject to} & M - \sum_{i=1}^n X_i^{2d} \geq 0 \end{cases}$$

for $M > 0$ and $2d \geq \deg f$.

Question. How one can extend the method of geometric program to approximate a lower bound for (1)?

SOS approximation

The rest of the thesis (chapters 4, 5 and 6) is devoted to the K-moment problem for continuous linear functionals. In Section 2.4, it is proved that for a cone $C \subseteq \mathbb{R}[\underline{X}]$, a locally convex topology τ on $\mathbb{R}[\underline{X}]$ and a closed set $K \subseteq \mathbb{R}^n$, the followings are equivalent:

- $\overline{C}^{\tau} = \operatorname{Psd}(K)$,
- For every τ-continuous linear functional L with $L(C) \subseteq \mathbb{R}^+$ there exists a positive Borel measure μ on K such that

$$\forall f \in \mathbb{R}[\underline{X}] \quad L(f) = \int_K f \, d\mu.$$

Then we studied various cases for C, τ and K:

$C = \sum \mathbb{R}[\underline{X}]^{2d}$ and $\tau = \|\cdot\|_{\phi}$-topology

If $\phi : \mathbb{N}^n \longrightarrow \mathbb{R}^+$ satisfies $\phi(1, \ldots, 1) \geq 1$ and $\phi(\alpha + \beta) \leq \phi(\alpha)\phi(\beta)$ then the map $\|\cdot\|_{\phi}$: $\mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}^+$ defined by

$$\|\sum_{\alpha} f_{\alpha}\underline{X}^{\alpha}\|_{\phi} := \sum_{\alpha} |f_{\alpha}|\phi(\alpha),$$

is a submultiplicative seminorm and for

$$\mathcal{K}_{\phi} := \{x \in \mathbb{R}^n \ : \ |x^{\alpha}| \leq \phi(\alpha) \quad \forall \alpha \in \mathbb{N}^n\},$$

the equation $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_\phi} = \mathrm{Psd}(\mathcal{K}_\phi)$ holds. A different but equivalent reading of this result says that for every $\|\cdot\|_\phi$-continuous linear functional L for which $L(f^{2d}) \geq 0$ for all $f \in \mathbb{R}[\underline{X}]$, there exists a positive Borel measure μ on $\mathcal{K}_\phi$ such that $L(f) = \int_{\mathcal{K}_\phi} f \, d\mu$.

Then we show that this result is extendible to the semigroupring $\mathbb{C}[S]$, where S is a commutative semigroup with convolution and neutral element. This result generalizes Berg *et al.*'s 1976 [7].

$C = \sum \mathbb{R}[\underline{X}]^{2d}$ and $\tau = \|\cdot\|_{p,r}$-topology

Taking $r = (r_1, \ldots, r_n) \in (\mathbb{R}^{>0})^n$, the map $\phi(\alpha) = r^\alpha$ is an absolute value on $\mathbb{N}^n$. For $1 \leq p < \infty$ the maps

$$\|\sum_\alpha f_\alpha \underline{X}^\alpha\|_{p,r} := \left(\sum_\alpha |f_\alpha|^p r^\alpha\right)^{\frac{1}{p}},$$

and

$$\|\sum_\alpha f_\alpha \underline{X}^\alpha\|_{\infty,r} := \max_\alpha |f_\alpha| r^\alpha$$

are different norms, inducing different topologies on $\mathbb{R}[\underline{X}]$. The closure result and hence the moment problem corresponding to these norms is the following (Theorem 5.2.2):

- $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{p,r}} = \mathrm{Psd}(\prod_{i=1}^n [-r_i^{\frac{1}{p}}, r_i^{\frac{1}{p}}])$,
- $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_{\infty,r}} = \mathrm{Psd}(\prod_{i=1}^n [-r_i, ri])$.

$C = \sum \mathbb{R}[\underline{X}]^{2d}$ and compact K

For a given compact subset K of $\mathbb{R}^n$,

$$\|f\|_K := \sup_{x \in K} |f(x)|$$

defines a seminorm on $\mathbb{R}[\underline{X}]$. The necessary and sufficient condition on K which guarantees $\|\cdot\|_K$ to be a norm is given (Corollary 6.1.3). It is proved that $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\|\cdot\|_K} = \mathrm{Psd}(K)$ (Theorem 6.1.5).

$C = \sum \mathbb{R}[\underline{X}]^{2d}$ and closed K

Replacing the compactness condition on $K \subseteq \mathbb{R}^n$ with closeness annihilates $\|\cdot\|_K$-topology to be defined. But to each $a \in K$ we can assign a seminorm ρ_a, defined by $\rho_a(f) = |f(a)|$.

The family of seminorms $\{\rho_a : a \in K\}$ induces a locally convex topology $\mathcal{T}_K$ on $\mathbb{R}[\underline{X}]$. It is also proved that $\overline{\sum \mathbb{R}[\underline{X}]^{2d}}^{\mathcal{T}_K} = \mathrm{Psd}(K)$ (Theorem 6.2.3).

All the results on the moment problem are stated for sums of $2d^{th}$ powers which seems to be an improvement. But in contrast to sums of squares, there is no systematic method to study sum of $2d^{th}$ powers decompositions of polynomials for $d > 1$. Note that for $d = 1$, semidefinite programming provides a practical method to work with sums of squares.

Question. Is there an effective decision procedure like semidefinite programming for sums of squares, to determine whether a polynomial is a sum of $2d^{th}$ powers?

We have tried to state the results of chapters 4, 5 and 6 in a more general framework of topological vector spaces or *commutative* topological $\mathbb{R}$-algebras. The reason is that the main focus of this thesis is on the ring of polynomials. The moment problem for *noncommutative* algebras has been studied specially in functional analysis and it is still an active field. A natural question is that

Question. Which parts of the theory that we developed can be carried over noncommutative case?

Appendix A

Some Model Theory of Real Closed Fields

In this appendix, first we state some basic facts about model theory of real closed fields. The definitions and results on the model theory of real closed fields are mainly taken from [40]. The main purpose of this chapter is to provide a concrete version of the Abstract Positivstellensatz for finitely generated $\mathbb{R}$- algebras which is used to prove Theorem 1.4.3 and 1.4.4.

A.1 The Theory of Real Closed Fields

Let $\mathscr{L}_{or}$ be the language of ordered rings $\{+,-,\cdot,<,0,1\}$, where $+$, $-$ and $\cdot$ are binary function symbols, $<$ a binary relation and 0 and 1 are constants. The axioms for ordered fields are the following.

G1 $\forall x\ x+0=x$,

G2 $\forall x\forall y\forall z\ x+(y+z)=(x+y)+z$,

G3 $\forall x\exists y\ x+y=0$,

Gc $\forall x\forall y\ x+y=y+x$,

G- $\forall x\forall y\forall z\ (x-y=z\leftrightarrow x=y+z)$,

R1 $\forall x\ x\cdot 0=0$,

R2 $\forall x\ x\cdot 1=1\cdot x=x$,

R3 $\forall x\forall y\forall z\ (x\cdot(y\cdot z)=(x\cdot y)\cdot z)$,

R4 $\forall x\forall y\forall z\ x\cdot(y+z)=(x\cdot y)+(x\cdot z)$,

R5 $\forall x\forall y\forall z\ (x+y)\cdot z=(x\cdot z)+(y\cdot z)$,

Rc $\forall x\forall y\ x\cdot y=y\cdot x$,

F1 $1\neq 0$,

F2 $\forall x\ (x \neq 0 \to \exists y\ x \cdot y = 1)$,

O1 $\forall x \forall y \forall z (x < y \wedge y < z \to x < z)$,

O2 $\forall x \forall y [x < y \wedge \neg(y \leq x)] \vee [y < x \wedge \neg(x \leq y)] \vee [x = y \wedge \neg(x < y \vee y < x)]$,

O3 $\forall x \forall y \forall z\ (x < y \to x + z < y + z)$,

O4 $\forall x \forall y \forall z\ ((x < y \wedge 0 < z) \to x \cdot z < y \cdot z)$.

The theory derived from the above set of axioms in $\mathscr{L}_{or}$ is called the theory of *ordered fields* and is denoted by OF. This is easy to check that for a model $F \vDash \mathrm{OF}$, the definable set $F^+ = \{x : 0 < x \vee x = 0\}$ is an ordering of the field $(|F|, \cdot, +, 0, 1)$. For convenience, we identify the ordering F^+ by $<$ and the ordered $(|F|, F^+)$ simply by $(F, <)$. OF together with (RC1) and the axiom scheme (RC2$_n$) are the axioms for *the theory of real closed fields (RCF)* in $\mathscr{L}_{or}$.

RC1 $\forall x \exists y\ (y^2 = x \vee y^2 + x = 0)$,

RC2$_n$ $\forall x_0 \ldots \forall x_{2n} \exists y\ y^{2n+1} + \sum_{i=0}^{2n} x_i y^i = 0$.

The models of RCF are exactly real closed fields with their canonical orderings. Because the ordering is definable by the $\mathscr{L}_r$-formula[1]

$$\exists z\ (z \neq 0 \wedge x + z^2 = y),$$

any definable set of a model $F \vDash \mathrm{RCF}$ in $\mathscr{L}_{or}$ is also definable in $\mathscr{L}_r$. Also it follows from axioms O1-O4 that $\forall x_1 \ldots \forall x_n\ x_1^2 + \cdots + x_n^2 + 1 \neq 0$, therefore $(|F|, \cdot, +, 0, 1)$ is formally real.

Definition A.1.1. A field F which is formally real as a ring is called a formally real field (Definition 1.1.7). A real closed algebraic extension of a formally real field F is called a *real closure* of F.

By Zorn's Lemma, every formally real field F has a maximal formally real algebraic extension. This maximal extension is a real closure of F (see [40, §3.3]). The real closures of a formally real field may not be unique. For example, for $F = \mathbb{Q}(X)$, $F_0 = F(\sqrt{X})$ and $F_1 = F(\sqrt{-X})$, are both formally real. If R_i denotes the a real closure of F_i, $i = 0, 1$, then R_0 is not isomorphic to R_1, because X is a square in R_0, but not in R_1.

[1] $\mathscr{L}_r$ denotes the language of rings, i.e., $\{+, -, \cdot, 0, 1\}$.

Lemma A.1.2. *If $(F,<)$ is an ordered field, $0 < x \in F$, and x is not a square in F, then we can extend the ordering of F to $F(\sqrt{x})$.*

Proof. Let $P = \{x \in F : x > 0 \vee x = 0\}$. The extension of P to $F(\sqrt{x})$ is $T = \sum F(\sqrt{x})^2 \cdot P$. This is a proper preordering of $F(\sqrt{x})$. Otherwise, if $-1 \in T$ then $-1 = \sum(a_i + b_i\sqrt{x})^2 p_i$, where $a_i, b_i \in F$ and $p_i \in P$. So $-1 = \sum(a_i^2 + b_i^2 x)p_i + 2\sum a_i b_i p_i \sqrt{x}$. Since x is not a square $\sum a_i b_i p_i \sqrt{x}$ must be 0 and hence $-1 = \sum(a_i^2 + b_i^2 x)p_i \in P$ which is a contradiction. Therefore T is proper and by Proposition 1.1.6 there exists an ordering P' of $F(\sqrt{x})$ extending T and $P' \cap F = P$. □

Corollary A.1.3. *If $(F,<)$ is an ordered field, there is a real closure R of F such that the canonical ordering of R extends the ordering on F.*

Proof. By successive application of Lemma A.1.2, we can find an ordered field $(L,<)$ extending $(F,<)$ such that every positive element of F is a square in L. We now apply Zorn's Lemma to find a maximal formally real algebraic extension R of L. Because every positive element of F is a square in R, the canonical ordering of R extends the ordering of F. □

A.2 Quantifier Elimination

Let $\mathscr{L}$ be a language and T a theory in $\mathscr{L}$. We say that T has *quantifier elimination* if for every formula ϕ, there exists a quantifier-free formula ψ such that

$$T \models \phi \leftrightarrow \psi.$$

Theorem A.2.1. *Suppose that $\mathscr{L}$ contains a constant symbol c, T is an $\mathscr{L}$-theory, and $\phi(\underline{v})$ is an $\mathscr{L}$-formula. The following are equivalent:*

1. *There is a quantifier-free $\mathscr{L}$-formula $\psi(\underline{v})$ such that $T \models \forall \underline{v}(\phi(\underline{v}) \leftrightarrow \psi(\underline{v}))$.*

2. *If $\mathcal{M}$ and $\mathcal{N}$ are models of T, $\mathcal{A}$ is an $\mathscr{L}$-structure, $\mathcal{A} \subseteq \mathcal{M}$, and $\mathcal{A} \subseteq \mathcal{N}$, then $\mathcal{M} \models \phi(\underline{a})$ if and only if $\mathcal{N} \models \phi(\underline{a})$ for all $\underline{a} \in \mathcal{A}$.*

Proof. See [40, Theorem 3.1.4]. □

The next lemma shows that we can prove quantifier elimination by getting rid of one existential quantifier at a time.

Lemma A.2.2. *Let T be an $\mathscr{L}$-theory an suppose that for every quantifier-free $\mathscr{L}$-formula $\theta(\underline{v}, w)$ there is a quantifier-free formula $\psi(\underline{v})$ such that $T \vDash \exists w\theta(\underline{v}, w) \leftrightarrow \psi(\underline{v})$. Then, T has quantifier elimination.*

Proof. Let $\phi(\underline{v})$ be an $\mathscr{L}$-formula. We wish to show that $T \vDash \forall\underline{v}(\phi(\underline{v}) \leftrightarrow \psi(\underline{v}))$ for some quantifier-free $\phi(\underline{v})$. We prove this by induction on the complexity of $\phi(\underline{v})$.

If ϕ is quantifier-free, there is nothing to prove. Suppose that for $i = 0, 1$, $T \vDash \forall\underline{v}(\theta_i(\underline{v}) \leftrightarrow \psi_i(\underline{v}))$, where ψ_i is quantifier free.

If $\phi(\underline{v}) = \neg\theta_0(\underline{v})$, then $T \vDash \forall\underline{v}(\phi(\underline{v}) \leftrightarrow \neg\psi_0(\underline{v}))$.

If $\phi(\underline{v}) = \theta_0(\underline{v}) \wedge \theta_1(\underline{v})$, then $T \vDash \forall\underline{v}(\phi(\underline{v}) \leftrightarrow (\psi_0(\underline{v}) \wedge \psi_1(\underline{v})))$.

In either case, ϕ is equivalent to a quantifier-free formula.

Suppose that $T \vDash \forall\underline{v}(\theta(\underline{v}, w) \leftrightarrow \psi_0(\underline{v}, w))$, where ψ_0 is quantifier-free and $\phi(\underline{v}) = \exists w\theta(\underline{v}, w)$. Then $T \vDash \forall\underline{v}(\phi(\underline{v}) \leftrightarrow \exists w\psi_0(\underline{v}, w))$. By our assumptions, there is a quantifier-free $\psi(\underline{v})$ such that $T \vDash \forall\underline{v}(\exists w\psi_0(\underline{v}, w) \leftrightarrow \psi(\underline{v}))$. But then $T \vDash \forall\underline{v}(\phi(\underline{v}) \leftrightarrow \psi(\underline{v}))$. □

Combining A.2.1 and A.2.2 we get the following simple and useful test for quantifier elimination.

Corollary A.2.3. *Let T be an $\mathscr{L}$-theory. Suppose that for all quantifier-free formulas $\phi(\underline{v}, w)$, if $\mathcal{M}, \mathcal{N} \vDash T$, $\mathcal{A}$ is a common substructure of $\mathcal{M}$ an $\mathcal{N}$, $\underline{a} \in \mathcal{A}$, and there is $b \in |\mathcal{M}|$ such that $\mathcal{M} \vDash \phi(\underline{a}, b)$, then there is $c \in |\mathcal{N}|$ such that $\mathcal{N} \vDash \phi(\underline{a}, c)$. Then T has quantifier elimination.*

A *universal sentence* is one of the form $\forall\underline{v}\phi(\underline{v})$, where ϕ is quantifier-free. We say that an $\mathscr{L}$-theory T has a *universal axiomatization* if there is a set of universal $\mathscr{L}$-sentences Γ such that $\Gamma \vDash T$ and $T \vDash \Gamma$.

Lemma A.2.4. *Suppose that $\mathcal{N}$ is a substructure of $\mathcal{M}$, $\underline{a} \in |\mathcal{N}|$, and $\phi(\underline{v})$ is a quantifier-free formula. Then, $\mathcal{N} \vDash \phi(\underline{a})$ if and only if $\mathcal{M} \vDash \phi(\underline{a})$.*

Proof. Use induction on the complexity of ϕ. □

Theorem A.2.5. *An $\mathscr{L}$-theory T has a universal axiomatization if and only if whenever $\mathcal{M} \vDash T$ and $\mathcal{N}$ is a substructure of $\mathcal{M}$, then $\mathcal{N} \vDash T$. In other words, a theory preserved under substructures if and only if it has a universal axiomatization.*

Proof. Suppose that $\mathcal{N} \subseteq \mathcal{M}$. By Lemma A.2.4, if $\phi(\underline{v})$ is quantifier-free and $\underline{a} \in |\mathcal{N}|$, then $\mathcal{N} \vDash \phi(\underline{a})$ if and only if $\mathcal{M} \vDash \phi(\underline{a})$. Thus, if $\mathcal{M} \vDash \forall \underline{v} \phi(\underline{v})$, then so does $\mathcal{N}$. Suppose that T preserved under substructures. Let

$$T_\forall := \{\phi \ : \ \phi \text{ is universal and } T \vDash \phi\}.$$

Clearly, if $\mathcal{N} \vDash T$, then $\mathcal{N} \vDash T_\forall$. For other direction, suppose that $\mathcal{N} \vDash T_\forall$. We claim that $\mathcal{N} \vDash T$. Let $\mathscr{L}_\mathcal{N} = \mathscr{L} \cup |\mathcal{N}|$. Then $\mathcal{N}$ is also a structure of $\mathscr{L}_\mathcal{N}$, interpreting all new symbols $a \in |\mathcal{N}|$ by a itself. We define the *atomic diagram* of $\mathcal{N}$ by $\mathrm{Diag}(\mathcal{N}) := \{\phi(a_1, \dots, a_n) \ : \ \mathcal{N} \vDash \phi(a_1, \dots, a_n), \phi$ an atomic or the negation of an atomic $\mathscr{L}$-fomula$\}$.

Claim. $T \cup \mathrm{Diag}(\mathcal{N})$ is satisfiable.

Suppose not. Then, there is a finite $\Delta = \{\psi_1, dots, \psi_n\} \subseteq \mathrm{Diag}(\mathcal{N})$ such that $T \cup \Delta$ is not satisfiable. Let $\underline{c}$ be the new constant symbol from $|\mathcal{N}|$ used in $\psi_1, \dots, \psi_n$ and say $\psi_i = \phi_i(\underline{c})$, where ϕ_i is a quantifier-free $\mathscr{L}$-formula. Because the constant $\underline{c}$ do not occur in T, if there is a model of $T \cup \{\exists \underline{v} \phi_i(\underline{v})\}$, then by interpreting $\underline{c}$ as witnesses to the existential formula, $T \cup \Delta$ would be satisfiable. Thus $T \vDash \forall \underline{v} \neg \phi(\underline{v})$. As the later formula is universal, $\forall \underline{v} \bigvee \neg \phi_i(\underline{v}) \in T_\forall$, contradicting $\mathcal{N} \vDash T_\forall$. So there is $\mathcal{M}$ satisfying $T \cup \mathrm{Diag}(\mathcal{N})$ which implies $\mathcal{N} \subseteq \mathcal{M}$.

Since T preserved under substructures, $\mathcal{N} \vDash T$ and $T_\forall$ is a universal axiomatization of T. $\square$

Corollary A.2.6. *Let T be an $\mathscr{L}$-theory. Then $\mathcal{A} \vDash T_\forall$ if and only if there is $\mathcal{M} \vDash T$ with $\mathcal{A} \subseteq \mathcal{M}$.*

Proof. If $\mathcal{A} \vDash T_\forall$, follow the same argument as in Theorem A.2.5 to get a model for $T \cup \mathrm{Diag}(\mathcal{A})$ which satisfies $\mathcal{A} \subseteq \mathcal{M}$. For the reverse, note that if $\mathcal{A} \subseteq \mathcal{M} \vDash T$, then $\mathcal{M} \vDash T_\forall$ and by Theorem A.2.5, $\mathcal{A} \vDash T_\forall$ as desired. $\square$

Definition A.2.7. A theory T has *algebraically prime models* if for any $\mathcal{A} \vDash T_\forall$, there is $\mathcal{A} \subseteq \mathcal{M} \vDash T$ such that for all $\mathcal{N} \vDash T$ and embeddings $j : \mathcal{A} \longrightarrow \mathcal{N}$, there is $h : \mathcal{M} \longrightarrow \mathcal{N}$ completing the following diagram:

$$\begin{tikzcd} \mathcal{A} \arrow[r, "i"] \arrow[dr, "j"'] & \mathcal{M} \arrow[d, "h"] \\ & \mathcal{N} \end{tikzcd}$$

If $\mathcal{M}, \mathcal{N} \models T$ and $\mathcal{M} \subseteq \mathcal{N}$, we say $\mathcal{M}$ is *simply closed* in $\mathcal{N}$ and write $\mathcal{M} \prec_s \mathcal{N}$ if for any quantifier-free formula $\phi(\underline{v}, w)$ and any $\underline{a} \in |\mathcal{M}|$, if $\mathcal{N} \models \exists w \phi(\underline{a}, w)$ then so does $\mathcal{M}$.

Corollary A.2.8. *If T has algebraically prime models and $\mathcal{M} \prec_s \mathcal{N}$ whenever $\mathcal{M} \subseteq \mathcal{N}$ are models of T, then T has quantifier elimination.*

Proof. Apply Corollary A.2.3 to $\mathcal{A} \models T_\forall$. □

Recall that for two $\mathscr{L}$-structures, $\mathcal{M}$ and $\mathcal{N}$, we say $\mathcal{M}$ is an *elementary submodel* of $\mathcal{N}$, or $\mathcal{N}$ is an *elementary extension* of $\mathcal{M}$ if $\mathcal{M} \subseteq \mathcal{N}$ and

$$\mathcal{M} \models \phi(a_1, \dots, a_n) \Leftrightarrow \mathcal{N} \models \phi(a_1, \dots, a_n),$$

for all $\mathscr{L}$-formula $\phi(v_1, \dots, v_n)$ and all $a_1, \dots, a_n \in |\mathcal{M}|$.

Definition A.2.9. An $\mathscr{L}$-theory T is said to be *model complete* if for $\mathcal{M}, \mathcal{N} \models T$, $\mathcal{M} \subseteq \mathcal{N}$ implies $\mathcal{M} \prec \mathcal{N}$.

Proposition A.2.10. *If T has quantifier elimination, then T is model complete.*

Proof. Suppose that $\mathcal{M} \subseteq \mathcal{N}$ are models of T. We show that $\mathcal{M}$ is an elementary submodel of $\mathcal{N}$. Let $\phi(\underline{v})$ be an $\mathscr{L}$-formula, and $\underline{a} \in |\mathcal{M}|$. There is a quantifier-free formula $\psi(\underline{v})$ such that $\mathcal{M} \models \forall \underline{v}(\phi(\underline{v}) \leftrightarrow \psi(\underline{v}))$. Since quantifier-free formulas are preserved under substructures and extensions, $\mathcal{M} \models \psi(\underline{a})$ if and only if $\mathcal{N} \models \psi(\underline{a})$. Thus

$$\mathcal{M} \models \phi(\underline{a}) \Leftrightarrow \mathcal{M} \models \psi(\underline{a}) \Leftrightarrow \mathcal{N} \models \psi(\underline{a}) \Leftrightarrow \mathcal{N} \models \phi(\underline{a}).$$

□

A.3 Tarski's Transfer Principle

We now show that RCF has quantifier elimination.

Lemma A.3.1. *RCF$_\forall$ is the theory of ordered integral domains.*

Proof. Clearly, any substructure of a real closed field is an ordered integral domain. If $(D, <)$ is an ordered integral domain and $F = \mathrm{ff}(D)$, then we can order F by

$$\frac{a}{b} > 0 \Leftrightarrow a \cdot b > 0.$$

By Corollary A.1.3, we can find $(R, <) \models \mathrm{RCF}$ such that $(F, <) \subseteq (R, <)$. □

Corollary A.3.2. *If $(F,<)$ is an ordered field and R_1 and R_2 are real closures of F where the canonical ordering extends the ordering of $(F,<)$, then there is a unique field isomorphism between R_1 and R_2 that is the identity on F.*

Corollary A.3.3. *RCF has algebraically prime models.*

Proof. Let $(D,<)$ be an ordered domain and $(R,<)$ be the real closure of $\mathrm{ff}(D)$ compatible with the ordering of D. Let $(F,<)$ be any real closed field extension of $(D,<)$. Let $K = \{\alpha \in F : \alpha \text{ is algebraic over } \mathrm{ff}(D)\}$. K is real closed and since the ordering of K extends $(D,<)$, by Corollary A.3.2, there is an isomorphism between F and K, fixing D. □

Theorem A.3.4. *RCF admits quantifier elimination in $\mathscr{L}_{or}$.*

Proof. Since RCF has algebraically prime models, by Corollary A.2.8, it suffices to show that $F \prec_s K$ when $F, K \models \mathrm{RCF}$ and $F \subseteq K$. Let $\phi(v, \underline{w})$ be a quantifier-free formula and $\underline{a} \in F$, $b \in K$ be such that $K \models \phi(b, \underline{a})$. We must find $b' \in F$ such that $F \models \phi(b', \underline{a})$.

Note that

$$p(\underline{X}) \neq 0 \leftrightarrow (p(\underline{X}) > 0 \vee -p(\underline{X}) > 0)$$

and

$$p(\underline{X}) \not> 0 \leftrightarrow (p(\underline{X}) = 0 \vee -p(\underline{X}) > 0).$$

So, we may assume that ϕ is a disjunction of conjunctions of formulas of the form $p(v, \underline{w}) = 0$ or $p(v, \underline{w}) > 0$. Let us assume that

$$\phi(v, \underline{a}) \leftrightarrow \bigwedge_{i=1}^{n} p_i(v) = 0 \wedge \bigwedge_{i=1}^{m} q_i(v) > 0,$$

where $p_1, \ldots, p_n, q_1, \ldots, q_m \in F[X]$. If any of the polynomials p_i is nonzero, then b is algebraic over F, and since F has no proper formally real algebraic extension, we conclude that $b \in F$. Thus we reduce to case that

$$\phi(v, \underline{a}) \leftrightarrow \bigwedge_{i=1}^{m} q_i(v) > 0.$$

The polynomial $q_i(X)$ can only change signs at zeros of q_i and all zeros of q_i are in F. Therefore, we can find $c_i \in F \cup \{-\infty\}$ and $d_i \in F \cup \{\infty\}$ such that $c_i < b < d_i$ and $q_i(x) > 0$ for all $x \in (c_i, d_i)$. Let $c = \max\{c_1, \ldots, c_m\}$ and $d = \min\{d_1, \ldots, d_m\}$. Then, $c < d$ and $\bigwedge_{i=1}^{m} q_i(x) > 0$ whenever $c < x < d$. So, there is a $b' \in F$ such that $F \models \phi(b', \underline{a})$. □

Corollary A.3.5. *RCF is model complete and* $RCF = Th(\mathbb{R}, +, \cdot, <)$.

The following well-known result about real closed fields has many applications in real algebra. It can be deduced from the Tarski-Seidenberg Theorem [63, 59]. It also can be deduced from Lang's Homomorphism Theorem.

Corollary A.3.6 (Tarski's Transfer Principle)**.** *Suppose* $(F,<)$ *is an ordered field extension of* $(R,<)$ *where* R *is a real closed field. Suppose that* $\exists x_1, \ldots, x_n \in F$ *satisfying some finite system of polynomial equations and inequalities with coefficients in* R*. Then* $\exists x_1, \ldots, x_n \in R$ *satisfying these same equations and inequalities.*

We denote by $\mathrm{Hom}_R(A,R)$ the set of all nonzero R-algebra homomorphisms from A to R. In spite of what described in Remark 1.2.6 for the case $R = \mathbb{R}$, in general a ring homomorphism $\alpha : A \longrightarrow R$ is not necessarily an R-algebra homomorphism. For example let $A = R = \bigcup \mathbb{R}((X^{\frac{1}{n}}))$ be the field of Puiseux series which is real closed [3, Theorem 2.91]. The map $\Theta : R \longrightarrow R$ defined by $\Theta(X) = X^3$ nontrivial ring homomorphism from R to R which is not an R-algebra homomorphism.

Theorem A.3.7. *Let* A *be a finitely generated* R*-algebra, where* R *is a real closed field. Then* $Hom_R(A,R)$ *is dense in* $Sper(A)$ *with patch topology (and hence in spectral topology).*

Note that here we identify orderings of the form $\alpha^{-1}(R^+) \in \mathrm{Sper}(A)$ by $\alpha \in \mathrm{Hom}_R(A,R)$ itself.

Proof of A.3.7. A basic open set of $\mathrm{Sper}(A)$ in patch topology is of the form

$$U = \{P \in \mathrm{Sper}(A) \ : \ a_i \notin P,\ b_j \in P,\ i = 1, \ldots, n, j = 1, \ldots, m\},$$

for $a_i, b_j \in A$. Let $P \in U$ and consider the field extension $(F,<)$ of $(R,<)$ defined by $F = \mathrm{ff}\frac{A}{\mathfrak{p}}$, where $\mathfrak{p} = \mathrm{supp}\ (P)$ and '<' is the ordering on F defined by P using Proposition 1.1.5. Since A is a finitely generated R-algebra, A itself is a homomorphic image of $R[\underline{X}] = R[X_1, \ldots, X_s]$ under an R-algebra homomorphism π and, by Hilbert's Basis Theorem, $\ker \pi$ is a finitely generated, i.e., $\ker \pi = \langle h_1(\underline{X}), \ldots, h_l(\underline{X})\rangle$ for $h_1(\underline{X}), \ldots, h_l(\underline{X}) \in R[\underline{X}]$. For each $i = 1, \ldots, n$ and $j = 1, \ldots, m$, $\exists f_i, g_j \in R[\underline{X}]$ such that $a_i = \pi(f_i(\underline{X}))$ and $b_j = \pi(g_j(\underline{X}))$, where

$$R[\underline{X}] \xrightarrow{\pi} A \xrightarrow{\iota} F = \mathrm{ff}\frac{A}{\mathfrak{p}}.$$

Let $x_i = \pi(X_i) + \mathfrak{p}$, $i = 1, \ldots, s$ and $x = (x_1, \ldots, x_k)$. Then $\iota(a_i) = \iota\pi(f_i) = f_i(x) < 0$, $\iota(b_j) = \iota\pi(g_j) = g_j(x) \geq 0$ and $\pi(h_k) = 0$ for each $i = 1, \ldots, n$, $j = 1 \ldots, m$ and $k = 1 \ldots, l$. By Tarski's Transfer Principle A.3.6, there exists $y = (y_1, \ldots, y_s) \in R^s$ such that $f_i(y) < 0$, $g_j(y) \geq 0$ and $h_k(y) = 0$ for $i = 1, \ldots, n$, $j = 1, \ldots, m$ and $k = 1, \ldots, l$. Now the R-algebra homomorphism $\alpha : A \longrightarrow R$ defined by $\alpha(\pi(X_i)) = y_i$ is well-defined and we have $\alpha^{-1}(R^+) \in U$ as desired. □

Denoting $\mathrm{Hom}_R(A, R) \cap \mathrm{Sper}_{T_S}(A)$ by $\mathcal{K}^R_{T_S}$ we have the concrete version of the Abstract Positivstellensatz 1.2.4.

Theorem A.3.8 (Concrete Positivstellensatz). *Let A be a finitely generated R-algebra, R a real closed filed, S a finite subset of A and $a \in A$. Then*

1. $\hat{a} > 0$ *on* $\mathcal{K}^R_{T_S}$ $\Leftrightarrow$ $\exists p, q \in T_S$ $(pa = 1 + q)$.
2. $\hat{a} \geq 0$ *on* $\mathcal{K}^R_{T_S}$ $\Leftrightarrow$ $\exists p, q \in T_S$ $\exists m \in \mathbb{N}$ $(pa = a^{2m} + q)$.
3. $\hat{a} = 0$ *on* $\mathcal{K}^R_{T_S}$ $\Leftrightarrow$ $\exists m \in \mathbb{N}$ $(-a^{2m} \in T_S)$.

Proof. (1) It suffices to show that $\hat{a} > 0$ on $\mathcal{K}_{T_S}$ if and only if $\mathrm{Sgn}_P a > 0$ on $\mathrm{Sper}_{T_S}(A)$, then the conclusion follows from Abstract Positivstellensatz (1). One direction ($\Leftarrow$) is trivial. For the converse, note that $\mathrm{Sgn}_P a = 1$ over $\mathcal{K}^R_{T_S}$ is equivalent to the condition $\alpha(a) > 0$ for all $\alpha \in \mathcal{K}^R_{T_S}$. By assumption

$$\begin{aligned} \mathcal{K}^R_{T_S} &\subseteq \{P \in \mathrm{Sper}_{T_S}(A) \ : \ \mathrm{Sgn}_P a = 1\} \\ &= \{P \in \mathrm{Sper}_{T_S}(A) \ : \ -a \notin P\} \\ &= U(-a). \end{aligned}$$

Now, in contrary, assume that $\exists P \in \mathrm{Sper}_{T_S}(A)$ such that $\mathrm{Sgn}_P a \leq 0$. In other words, $U(-a)^c \cap \mathrm{Sper}_{T_S}(A) \neq \varnothing$. Then

$$\mathrm{Sper}_{T_S}(A) = \bigcap_{s \in S} U(s)^c. \tag{A.3.1}$$

The equality (A.3.1) shows that for finite S, $\mathrm{Sper}_{T_S}(A)$ is open in patch topology[2]. Therefore $U(-a)^c \cap \mathrm{Sper}_{T_S}(A)$ is (a non-empty) open in patch topology. By Theorem A.3.7,

$$\mathcal{A} = \mathrm{Hom}_R(A, R) \cap [U(-a)^c \cap \mathrm{Sper}_{T_S}(A)] \neq \varnothing.$$

[2]Note that in general, (A.3.1) implies the compactness of $\mathrm{Sper}_T(A)$, for any preordering T.

But $\mathcal{A} \subseteq \mathcal{K}_{T_S}^R$ and for every $\alpha \in \mathcal{A}$, we have $\alpha(a) > 0$, a contradiction. Thus $\mathrm{Sgn}_P a = 1$ on $\mathrm{Sper}_{T_S}(A)$ as desired. This completes the proof of (1).

A similar argument can be used to prove (2) and (3). $\square$

Appendix B
Generalized Moment Problem

In [37], Lasserre introduces a general setting for the moment problem which covers a broad range of problems appearing in different areas of mathematics. We show that the two main topics of this thesis can be considered as variants of Lasserre's formulation.

Recall that $\mathcal{M}_+(K)$ denotes the set of all finite Borel measures on K, where K is assumed to be a locally compact topological space. A function $f : K \longrightarrow \mathbb{R}$ is called a Borel function if for every Borel set B of $\mathbb{R}$, $f^{-1}(B)$ is a Borel set in K.

Definition B.9. Let $K \subseteq \mathbb{R}^n$ be a Borel set, Γ a set of indices, $\{\gamma_i : i \in \Gamma\}$ a set of reals, and $f, h_i : K \longrightarrow \mathbb{R}$, $i \in \Gamma$ are Borel functions that are integrable with respect to every measure $\mu \in \mathcal{M}_+(K)$. The *Generalized Moment Problem (GMP)* is the problem of finding

$$\rho_{mom} = \sup\{\int_K f\ d\mu\ :\ \mu \in \mathcal{M}_+(K),\quad \int_K h_i\ d\mu \leq \gamma_i\quad \forall i \in \Gamma\}. \tag{B.1}$$

Example B.10. K-Moment Problem. For the case $f = 0$, $h_\alpha = \pm\underline{X}^\alpha$, $\alpha \in \Gamma \subseteq \mathbb{N}^n$, (B.1) becomes the feasibility problem defined in 2.2.1 in the following settings:

Take f to be a constant function for example $f = 0$ and for each $\alpha \in \Gamma$, let $h_{\alpha_1} = \underline{X}^\alpha$, $h_{\alpha_2} = -\underline{X}^\alpha$, $\gamma_{\alpha_1} = \gamma_\alpha$ and $\gamma_{\alpha_2} = -\gamma_\alpha$. Now, the K-moment problem 2.2.1 has a solution if and only if the following GMP has a feasible solution.

$$\sup\{0\ :\ \mu \in \mathcal{M}_+(K),\ \int_K h_{\alpha_i}\ d\mu \leq \gamma_{\alpha_i}\quad \alpha \in \Gamma,\ i = 1, 2\}.$$

Example B.11. Polynomial Optimization. With $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $K \subseteq \mathbb{R}^n$, consider the constrained optimization problem

$$f^* = \sup_{x \in K} f(x), \tag{B.2}$$

which we rewrite as

$$\rho_{mom} = \sup\{\int_K f\ d\mu\ :\ \mu \in \mathcal{M}_+(K),\quad \int_K\ d\mu = 1\}. \tag{B.3}$$

Clearly, the equality $\int_K\ d\mu = 1$ in (B.3) is obtained by taking $\Gamma = \{1, 2\}$, $\gamma_1 = \gamma_2 = 1$, $h_1 = 1$ and $h_2 = -1$ in (B.1).

We show that (B.2) and (B.3) are equivalent. i.e., $f^* = \rho_{mom}$.
If $f^* = +\infty$, let M be arbitrary large, and let $x \in K$ be such that $f(x) \geq M$. Then, with $\mu = \delta_x \in \mathcal{M}_+(K)$, the Dirac measure at x, we have $\int_K f \; d\mu = f(x) \geq M$ and so $\rho_{mom} = +\infty$. Now, suppose that $f^* < +\infty$. Since $\forall x \in K$, $f(x) \leq f^*$, then $\int_K f \; d\mu \leq f^*$ and hence $\rho_{mom} \leq f^*$. Conversely, with every $x \in K$, we associate the Dirac measure $\delta_x \in \mathcal{M}_+(K)$ which is a feasible solution of (B.3) with value $f(x)$, leading to $\rho_{mom} \geq f^*$. Therefore $\rho_{mom} = f^*$ as desired.

This is a version of the moment problem we discussed in Chapter 3.

Appendix C
Source Code to Find Lower Bounds

This chapter contains source code written in Sage to compute lower bounds f_{sos} and f_{gp}, introduced in Chapter 3.

C.1 Main Package

This section contains the source code of two main classes, `GlOptGeoPrg` to compute f_{gp} and `SosTools` to compute f_{sos}, written for Sage. The solver used for SDP and also GP are `sdp` and `gp` from CvxOpt 1.1.4, developed by "Joachim Dahl" and "Lieven Vandenberghe" for Python. The latest version of this package, which is called CvxAlgGeo is available at http://goo.gl/iI3Y0.

There is also an earlier version written on Mathematica and for a given polynomial, computes r_L, r_{FK}, r_{dmt} and generates a Matlab code to solve the SDP and GP related to the polynomial to find f_{sos} and f_{gp}. It uses SosTools and Sdp3-4.0 to find f_{sos} and GPposy to find f_{gp} in Matlab. A copy of this code is available at http://goo.gl/bpSzg.

C.1.1 GlOptGeoPrg

The following code is the implementation of the geometric program to calculate f_{gp} introduced in Chapter 3.

```
import numpy
RealNumber = float  # Required for CvxOpt
Integer = int  # Required for CvxOpt

from cvxopt.base import matrix as Mtx
from cvxopt import solvers
from array import array
from time import time, clock

class GlOptGeoPrg:
```

```
"""
A class to find lower bounds of an even degree polynomial, using
Geometric Programs.
"""

number_of_variables=1;
total_degree=2;
polynomial=0;
constant_term=0;
fgp=0;
Info={};

def __init__(self,f):
    self.polynomial=f;
    self.number_of_variables=len(f.variables());
    self.total_degree=f.degree();
    self.constant_term=f.constant_coefficient();

def is_square_mono(self,mono,coef):
    """
    This functions gets the coefficient and the exponent of a term and
    returns True if it is an square term and False otherwise.
    """

    flag=True;
    if coef<0:
        return False
    exp=mono.exponents();
    for ex in exp[0]:
        flag=flag & (ex%2==0)
    return flag;
```

```
def Delta(self):
    """
    This function returns a list of pairs (Coefficient, Monomial)
    where the corresponding term is not a square.
    """

    f=self.polynomial;
    d1=[];d2=[];d3=[]
    monos=f.monomials();
    coefs=f.coefficients();
    for i in range(0,len(monos)):
        if not self.is_square_mono(monos[i], coefs[i]):
            tmpexp=monos[i].exponents();
            NAlpha=self.n_alpha(tmpexp[0]);
            if NAlpha!=0:
                d1.append(coefs[i]);
                d2.append(tmpexp[0]);
                d3.append(NAlpha);
    return [d1,d2,d3];

def tuple_to_exp(self,t1,t2):
 """
 This function takes two tuple of real numbers, raise each one
 in the fist one to the power of the corresponding entity in
 the second and then multiply them together.
 """

    mlt=1;
    n=len(t1);
    for i in range(0,n):
```

```
        if (t1[i]==0) and (t2[i]==0):
            continue;
        mlt*=t1[i]^t2[i];
    return mlt;

def n_alpha(self,alpha):
 """
 This function, counts the number of non-zero entities in the given
 exponent.
 """


    num=0;
    for i in alpha:
        if i!=0:
            num+=1;
    return num;

def sum_a_alpha(self,delta):
 """
 Counts number of auxiliary variables for lifting a certain monomial.
 """
    s=0;
    for a in delta:
        s+=self.n_alpha(a);
    return s;

def non_zero_before(self,mon,idx):
    """
    Counts the number of auxiliary lifting variables before a variable
    in a certain monomial.
    """
```

```
        cnt=0;
        for i in range(idx):
            if mon[i]!=0:
                cnt+=1;
        return cnt;

    def init_geometric_program(self,d,Delta):
"""
This function initializes the geometric program associated to
the input a polynomial.
"""

        SumNAlpha=sum(Delta[2]);
        num=len(Delta[0]);
        n=len(Delta[1][0]);
        F=matrix(RR,num+SumNAlpha,SumNAlpha,0);
        G=matrix(RR,num+SumNAlpha,1,0);
        K=[];
        cntr=0;
        for i in range(num):
            absalpha=sum(Delta[1][i]);
            G[i]=log((d-absalpha)*\\
             (self.tuple_to_exp(Delta[1][i],Delta[1][i])*\\
             (1.0*Delta[0][i]/d)^d)^(1.0/(d-absalpha)));
            for j in Delta[1][i]:
                if j!=0:
                    F[i,cntr]=-j*(1.0/(d-absalpha));
                    cntr+=1;
        K.append(num);
        cntr=0;
```

```
    for j in range(n):
        alphaidx=0;
        cnt=0;
        for i in range(num):
            if Delta[1][i][j]!=0:
                F[num+cntr,alphaidx+\\
                self.non_zero_before(Delta[1][i],j)]=1;
                cntr+=1;
                cnt+=1;
            alphaidx+=Delta[2][i];
        if cnt!=0:
            K.append(cnt);
    return [K,F,G];

def Matrix2CVXOPT(self,M):
    """
    Converts a Sage matrix into a matrix acceptable for
    the CvxOpt package.
    """

    n=M.ncols();
    m=M.nrows();
    CM=[];
    for j in range(n):
        tmp=[];
        for i in range(m):
            CM.append(M[i,j]);
    CC=Mtx(array('d', CM), (m,n))
    return CC;

def f_gp(self):
```

```
"""
The main function to compute the lower bound for an
even degree polynomial, using Geometric Program.
"""

n=self.number_of_variables;
d=self.total_degree;
f=self.polynomial
f0=self.constant_term;
delt=self.Delta();
GP=self.init_geometric_program(d,delt);
K=GP[0];
F=self.Matrix2CVXOPT(GP[1]);
g=self.Matrix2CVXOPT(GP[2]);
start = time();
start2 = clock();
sol = solvers.gp(K, F, g);
elapsed = (time() - start);
elapsed2 = (clock()-start2);
self.fgp=f0-e^(sol['primal objective']);
self.Info={"gp":self.fgp, "Wall":elapsed, "CPU":elapsed2};
if (sol['status']=='unknown') and (sol['gap'] > 0.00000001):
    self.Info['status']= 'Singular KKT';
else:
    self.Info['status']= 'Optimal';
#print sol;
return self.fgp;
```

C.1.2 SosTools

The following code is the implementation of the semidefinite program to calculate f_{sos} from Chapter 3.

```
import numpy
RealNumber = float  # Required for CvxOpt
Integer = int  # Required for CvxOpt

from cvxopt.base import matrix as Mtx
from cvxopt import solvers
from array import array
from time import time, clock

class SosTools:
    """
    A class to work with sum of square decomposition of polynomials.
    """

    number_of_variables=1;
    total_degree=2;
    polynomial=0;
    Rng=[];
    fullpolynomial=0;
    list=[];
    moment_matrix=[];
    fsos=0;
    Info={};

    def __init__(self,f):
        self.polynomial=f;
        self.number_of_variables=len(f.variables());
        self.total_degree=f.degree();
        self.Rng=PolynomialRing(RR,'x',self.number_of_variables);
        self.fullpolynomial=(1+sum(p for p in self.Rng.gens()))^\\
         self.total_degree;
```

```
        self.list=self.fullpolynomial.monomials();
        self.list.reverse();
        m1=self.MonomialsVector();
        self.moment_matrix=m1.transpose()*m1;

    def NumMonomials(self,n,d):
        """
        Returns the dimension of the space of polynomials
        on n variable with degree at most d.
        """

        return factorial(n+d)/(factorial(n)*factorial(d));

    def MonomialsVector(self):
        """
        Returns a vector consists of all monomials of
        degree at most d.
        """

        mat=matrix(1,len(self.list),self.list);
        return mat;

    def Balpha(self,mono,d,vars):
        """
        Constructs the matrix of multipliers.
        """

        m1=self.MonomialsVector();
        nD=self.NumMonomials(len(vars),d);
        B=matrix(nD,nD);
        M=self.moment_matrix;#m1.transpose()*m1;
```

```
    for i in range(0,nD):
        for j in range(i, nD):
            if M[i,j]==mono:
                B[j,i]=1;
    return B;#.transpose()

def ConstructSDPMat(self,n,d):
    """
    Constructs the main input matrix for the SDP.
    """

    d1=self.NumMonomials(n,d);
    d2=self.NumMonomials(n,2*d);
    G=matrix(d1^2,d2);
    idx=0;
    for p in self.list:
        B=self.VEC(self.Balpha(p,d,self.Rng.gens()));
        for j in range(0,d1^2):
            G[j,idx]=B[j,0];
        idx+=1;
    return G;

def Constraints(self,n,d):
    """
    Generates the constraints for the SDP.
    """

    d1=self.NumMonomials(n,d);
    d2=self.NumMonomials(n,2*d);
    A=matrix(RR,d1,d2);
    A[0,0]=1;
```

```
    b=matrix(RR,d1,1)
    b[0]=1;
    h=matrix(RR,d1,d1,0);
    return [A,b,h]

def VEC(self,M):
    """
    Arranges the columns of the matrix M in a long
    column matrix.
    """

    r=M.nrows();
    c=M.ncols();
    V=matrix(r*c,1);
    for i in range(0,r):
        for j in range(0,c):
            V[i*r+j]=M[j,i];
    return V;

def PolyCoefFullVec(self):
    """
    Returns the coefficients matrix, corresponding
    to the output of Balpha function.
    """

    f=self.polynomial
    c=matrix(RR,len(self.fullpolynomial.coefficients()),1)
    fmono=f.monomials();
    fcoef=f.coefficients()
    idx=0;
    for p in fmono:
```

```
            c[self.list.index(p)]=fcoef[idx];
            idx+=1;
        return c;

    def Matrix2CVXOPT(self,M):
        """
        Converts a Sage matrix into a matrix acceptable for
        the CvxOpt package.
        """

        n=M.ncols();
        m=M.nrows();
        CM=[];
        for j in range(n):
            tmp=[];
            for i in range(m):
                CM.append(M[i,j]);
        CC=Mtx(array('d', CM), (m,n))
        return CC

    def f_sos(self):
        """
        The main function to compute the lower bound for an
        even degree polynomial, using Semidefinite Program.
        """

        f=self.polynomial;
        n=self.number_of_variables;
        d=self.total_degree;
        if d%2==1:
            print 'An odd degree polynomial can not be sos.';
```

```
        return 0;
    d=d/2;
    c=self.Matrix2CVXOPT(self.PolyCoefFullVec());
    G=[self.Matrix2CVXOPT(-self.ConstructSDPMat(n,d))];
    Ab=self.Constraints(n,d);
    H=[self.Matrix2CVXOPT(Ab[2])];
    A=self.Matrix2CVXOPT(Ab[0]);
    B=self.Matrix2CVXOPT(Ab[1]);
    start = time();
    start2=clock();
    sol = solvers.sdp(c, Gl=A, hl=B, Gs = G, hs = H);
    elapsed = (time() - start);
    elapsed2 = (clock()-start2);
    self.fsos = max(sol['dual objective'],sol['primal objective']);
    self.Info={"sos":self.fsos, "Wall":elapsed, "CPU":elapsed2};
    if sol['status']!='optimal':
        self.Info['status']= 'Infeasible';
    else:
        self.Info['status']= 'Optimal';
    return self.fsos;
```

C.2 Sample Usage

The following is a sample code, written for SAGE to use the packages GLOPTGEOPRG and SOSTOOLS. This sample code generates a random polynomial with real coefficients of degree $2d$ on n variables.

```
n=4;        #number of variables
d=3;        #half degree of the polynomial


# Max Number of Monomials:
```

```
numMono=factorial(n+2*d-1)/(factorial(n)*factorial(2*d-1));

R=PolynomialRing(RR,'x',n);
diagPoly=sum(p^(2*d) for p in R.gens());
f=diagPoly+R.random_element(2*d-1,randint(1,numMono/2));

print 'f=',f;

POLY=GlOptGeoPrg(f);
POLY.f_gp();
print POLY.Info;

POLYS=SosTools(f);
POLYS.f_sos();
print POLYS.Info;
```

Sample Output:

```
f= x0^6 + x1^6 + x2^6 + x3^6 + x1^4*x2 - x0*x1*x2^3 - 2*x0^3*x1*x3 -
x1^4*x3 - x0^2*x1*x2*x3 + x0*x1*x2*x3^2 - 17*x0^2*x1^2 + 78*x1^4 +
6*x1*x2^2*x3 + x0*x1*x3^2 - 47*x0*x1*x2 - x0*x1*x3 - 2*x1^2*x3 + x1
     pcost       dcost       gap    pres   dres
 0:  0.0000e+00  1.4433e+01  5e+00  1e+00  9e-01
 1:  6.8148e+00  1.3350e+01  8e-01  7e-01  2e+00
 2: -3.1920e+00  8.0402e+00  7e-01  1e+00  3e-01
 3:  5.1626e+00  8.3013e+00  8e-02  3e-01  1e+00
 4:  6.8078e+00  8.0053e+00  2e-02  1e-01  4e-01
 5:  7.8132e+00  7.9724e+00  1e-03  1e-02  8e-02
 6:  7.9571e+00  7.9692e+00  5e-05  1e-03  1e-02
 7:  7.9689e+00  7.9691e+00  6e-07  2e-05  3e-04
 8:  7.9691e+00  7.9691e+00  6e-09  2e-07  4e-06
 9:  7.9691e+00  7.9691e+00  6e-11  2e-09  4e-08
```

```
Optimal solution found.
{'Wall': 0.035673141479492188, 'status': 'Optimal', 'CPU':
0.030000000000001137, 'gp': -2890.3788408942573}
     pcost       dcost       gap    pres   dres   k/t
 0:  0.0000e+00 -2.7313e+01  2e+03  8e+00  2e+00  1e+00
 1: -1.2955e+01 -1.8073e+01  5e+02  2e+00  5e-01  3e+00
 2: -8.6965e+00 -9.2183e+00  1e+02  7e-01  1e-01  2e+00
 3: -2.0646e+01 -1.9963e+01  1e+02  3e-01  7e-02  2e+00
 4: -2.9934e+01 -2.9485e+01  7e+01  2e-01  4e-02  1e+00
 5: -4.4318e+01 -4.3155e+01  1e+02  1e-01  3e-02  2e+00
 6: -5.5988e+01 -5.5383e+01  8e+01  8e-02  2e-02  9e-01
 7: -7.1915e+01 -7.1160e+01  8e+01  5e-02  1e-02  9e-01
 8: -7.7926e+01 -7.7239e+01  8e+01  4e-02  8e-03  8e-01
 9: -7.7956e+01 -7.7282e+01  8e+01  4e-02  8e-03  8e-01
10: -1.0257e+02 -1.0211e+02  4e+01  1e-02  3e-03  5e-01
11: -1.0477e+02 -1.0421e+02  5e+01  1e-02  3e-03  6e-01
12: -1.1813e+02 -1.1789e+02  2e+01  6e-03  1e-03  3e-01
13: -1.2763e+02 -1.2760e+02  3e+00  6e-04  1e-04  3e-02
14: -1.2877e+02 -1.2876e+02  4e-01  8e-05  2e-05  5e-03
15: -1.2895e+02 -1.2895e+02  6e-03  1e-06  3e-07  7e-05
16: -1.2895e+02 -1.2895e+02  3e-04  7e-08  1e-08  4e-06
17: -1.2895e+02 -1.2895e+02  7e-06  2e-09  3e-10  9e-08
Optimal solution found.
{'Wall': 7.7809391021728516, 'status': 'Optimal', 'sos':
-128.95264644582485, 'CPU': 7.7700000000000102}
```

References

[1] N. I. Akhiezer. *The Classical Moment Problem.* Hafner Publishing Company, New York, NY., 1965.

[2] E. Artin. Über die Zerlegung definiter Funktionen in Quadrate. *Abh. math. Sem. Hamburg*, (5):100–115, 1927.

[3] S. Basu, R. Pollack, and M. F. Roy. *Algorithms in Real Algebraic Geometry*, volume 10 of *Algorithms and Computation in Mathematics.* Springer-Verlag, Berlin Heidelberg, 2003.

[4] E. Beckenstein, L. Narici, and C. Suffel. *Topological Algebras*, volume 24 of *North-Holland Math. Stud.* Elsevier Sci. Publ., Amsterdam, 1977.

[5] M. Bellare and P. Rogaway. The complexity of approximating a nonlinear program. *Mathematical Programming*, 69:429–441, 1993.

[6] C. Berg, J. P. R. Christensen, and C. U. Jensen. A remark on the multidimensional moment problem. *Math. Ann.*, 243:163–169, 1979.

[7] C. Berg, J. P. R. Christensen, and P. Ressel. Positive definite functions on abelian semigroups. *Math. Ann.*, 223:253–272, 1976.

[8] C. Berg, J. P. R. Christensen, and P. Ressel. *Harmonic Analysis on Semigroups, Theory of Positive Definite and Related Functions*, volume 100 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.

[9] C. Berg and P. H. Maserick. Exponentially bounded positive definite functions. *Illinois J. Math.*, 28:162–179, 1984.

[10] J. Bochnak, M. Coste, and M. F. Roy. *Géométrie algébrique réelle*, volume 12 of *Ergeb. Math.* Springer, 1987. English translation, *Real algebraic geometry*, vol. 36, Ergeb. Math., Springer, 1998.

[11] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[12] C. N. Delzel. A continuous, constructive solution to Hilbert's 17th problem. *Inventiones Mathematicae*, 76(3):365–384, 1984.

[13] E. Deutsch. Bounds for the zeros of polynomials. *Amer. Math. Monthly*, 88:205–206, 1981.

[14] W. D. Dubois. A note on David Harrison's theory of preprimes. *Pac. J. Math.*, 21:15–19, 1967.

[15] C. Fidalgo and A. Kovacec. Positive semidefinite diagonal minus tail forms are sums of squares. *Math. Z.*, 269(3-4):629–645, 2011.

[16] M. Ghasemi, S. Kuhlmann, and E. Samei. The moment problem for continuous positive semidefinite linear functionals. *preprint*.

[17] M. Ghasemi and M. Marshall. Lower bounds for a polynomial in terms of its coefficients. *Arch. Math. (Basel)*, 95:343–353, 2010.

[18] M. Ghasemi and M. Marshall. Lower bounds for polynomials using geometric programming. *SIAM J. Optim.*, 22(2):460–473, 2012.

[19] M. Ghasemi, M. Marshall, and S. Wagner. Closure results for the cone of sums of $2d$-powers. *preprint*.

[20] H. Hamburger. Uber eine Erweiterung des Stieltjesschen Momentproblems Parts I, II, III. *Math. Ann.*, 81-82:235–319, 20–164, 168–187, 1920-1921.

[21] G. H. Hardy and E. W. Wright. *An introduction to the theory of numbers.* Oxford Science Publications, 1960.

[22] F. Hausdorff. Summationsmethoden und Momentfolgen I. *Math. Z.*, 9:74–109, 1921.

[23] E. K. Haviland. On the momentum problem for distribution functions in more than one dimension. *Amer. J. Math.*, 57:562–572, 1935.

[24] E. K. Haviland. On the momentum problem for distribution functions in more than one dimension II. *Amer. J. Math.*, 58:164–168, 1936.

[25] D. Hilbert. Über die Darstellung definiter Formen als Summe von Formenquadraten. *Math. Ann.*, 32:342–350, 1888.

[26] A. Hurwitz. Uber den Vergleich des arithmetischen und des geometrischen Mittels. *J. Reine Angew. Math.*, 108:266–268, 1891. See also: Math. Werke, Basel (505-507), 1933.

[27] T. Jacobi. A representation theorem for certain partially ordered commutative rings. *Math. Z.*, 237:259–273, 2001.

[28] T. Jacobi and A. Prestel. Distinguished representations of strictly positive polynomials. *J. rein angew. Math.*, 532:223–235, 2001.

[29] H. Jarchow. *Locally convex spaces.* B.G. Teubner Stuttgart, 1981.

[30] R. V. Kadison. A representation theory for commutative topological algebra. *Memories of the Amer. Math. Soc.*, 7, 1951.

[31] D. Knuth. *The Art of Computer Programming*, volume 2. Addison-Wesley, New York, 1969.

[32] J. L. Krivine. Anneaux préordonnés. *J. Analyse Math.*, 12:307–326, 1964.

[33] J. L. Krivine. Quelques propriétés des préordres dans les anneaux commutatifs unitaires. *C.R. Acad. Sc. Paris*, 258:3417–3418, 1964.

[34] S. Kuhlmann and M. Marshall. Positivity, sums of squares and the multidimensional moment problem. *Trans. Amer. Math. Soc.*, 354:4285–4301, 2002.

[35] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2001.

[36] J. B. Lasserre. Sufficient conditions for a real polynomial to be a sum of squares. *Arch. Math. (Basel)*, 89:390–398, 2007.

[37] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications*, volume 1 of *Optimization Series.* Imperial College Press, 2010.

[38] J. B. Lasserre. The K-moment problem for continuous linear functionals. *to appear*, 2011.

[39] J. B. Lasserre and T. Netzer. SOS approximation of nonnegative polynomials via simple high degree perturbations. *Math. Z.*, 256:99–112, 2006.

[40] D. Marker. *Model Theory: An Introduction*, volume 217 of *Graduate Texts in Mathematics.* Springer, 2002.

[41] M. Marshall. A general representation theorem for partially ordered commutative rings. *Math Z.*, 242:217–225, 2002.

[42] M. Marshall. Approximating positive polynomials using sum of squares. *Canad. Math. Bull.*, 46(3):400–418, 2003.

[43] M. Marshall. Optimization of polynomial functions. *Canad. Math. Bull.*, 46(4):575–587, 2003.

[44] M. Marshall. *Positive Polynomials and Sum of Squares*, volume 146 of *Mathematical Surveys and Monographs.* AMS., 2008.

[45] M. Marshall. Representation of non-negative polynomials, degree bounds and applications to optimization. *Canad. J. Math.*, 61:205–221, 2009.

[46] T. S. Motzkin. The arithmetic-geometric inequalities. *Proc. Symp. Wright-Patterson AFB*, pages 19–27, August 1965. In: Inequalities (O. Shisha, ed.).

[47] P. Parrilo and B. Sturmfels. Minimizing polynomial functions. *Algorithmic and quantitative real algebraic geometry (Piscataway, NJ)*, pages 88–99, 2001. DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 60, AMS., 2003.

[48] A. L. Peressini, F. E. Sullivan, and Jr. J. J. Uhl. *The Mathematics of Nonlinear Programming.* UTM series. Springer, 1987.

[49] V. V. Prasolov. *Polynomials*, volume 11. Algorithms and Computation in Mathematics, 2004.

[50] A. Prestel. *Lectures on formally real fields*, volume 22. IMPA Lecture Notes, Rio de Janeiro, 1975. also Lecture Notes in Math. 1093, Springer, 1984.

[51] M. Putinar. Positive polynomials on compact semialgebraic sets. *Indiana Univ. Math. J.*, 3(43):969–984, 1993.

[52] B. Reznick. A quantitative version of Hurwitz's theorem on the arithmetic-geometric inequality. *J. reine angew. Math.*, 377:108–112, 1987.

[53] B. Reznick. Forms derived from the arithmetic geometric inequality. *Math. Ann.*, 283:431–464, 1989.

[54] W. Rudin. *Functional Analysis, 2nd ed.* International series in pure and applied mathematics, 1991.

[55] C. Scheiderer. Sums of squares of regular functions on real algebraic varieties. *Trans. AMS.*, 352:1030–1069, 1999.

[56] C. Scheiderer. Non-existence of degree bounds for weighted sum of squares representation. *J. of Complexity*, 21:823–844, 2005.

[57] K. Schmüdgen. Positive cones in enveloping algebras. *Reports on Mathematical Physics*, 14(3):385–404, 1978.

[58] K. Schmüdgen. The K-moment problem for compact semialgebraic sets. *Math. Ann.*, 289:203–206, 1991.

[59] A. Seidenberg. A new decision method for elementary algebra. *Ann. of Math.*, 60:365–374, 1954.

[60] W. A. Stein et al. *Sage Mathematics Software (Version 4.6.2).* The Sage Development Team, 2012. http://www.sagemath.org.

[61] G. Stengle. A Nullstellensatz and a Positivstellensatz in semialgebraic geometry. *Math. Ann.*, 207:87–97, 1974.

[62] T. J. Stieljes. Recherches sur les fractions continues. *Ann. Fac. Sci. Toulouse*, pages A5–A47, 1885. Reprinted in Ann. Fac. Sci. Toulouse Math. 4, no.4, A5-A47, (1995).

[63] A. Tarski. A decision method for elementary algebra and geometry. *Fund. Math.*, 17:210–239, 1931. Rand Corp., 1948; UC Press, Berkeley, 1951; announced in Ann. Soc. Pol. Math. 9 (1930, published 1931) 207-207;.

[64] K. I. Švekov. On Hamburger's moment problem with supplementary requirement that masses are absent on a given interval. *Commun. Sco. Math. Kharkov*, 16:121–128, 1939. (in russian).

[65] S. Warner. *Topological Rings*, volume 178 of *North-Holland Math. Stud.* Elsevier Sci. Publ., Amsterdam, 1993.

[66] S. Willard. *General Topology.* Addison-Wesley, 1970.

[67] T. Wörmann. *Strikt positive Polynome in der semialgebraischen Geometrie.* Dissertation, Univ. Dortmund, 1998.

Glossary

$\overline{C}^{\tau}$	The closure of the set C with respect to the topology τ.
$\mathrm{C}_c(X)$	The algebra of all continuous real valued functions on X with compact support.
$\mathrm{C}(X)$	The algebra of all continuous real valued functions on X.
V_{τ}^{*}	The set of all τ-continuous linear functionals $L : V \longrightarrow \mathbb{R}$.
$C_{\tau}^{\vee}$	$\{L \in V_{\tau}^{*} : L \geq 0 \text{ on } C\}$.
$C_{\tau}^{\vee\vee}$	$\{a \in V : \forall L \in C_{\tau}^{\vee},\ L(a) \geq 0\}$.
f_{gp}	The lower bound for f obtained from its associated geometric program.
f_{sos}	$\sup\{r \in \mathbb{R} \ : \ f - r \in \sum \mathbb{R}[\underline{X}]^2\}$.
GMP	Generalized Moment Problem.
GP	geometric program.
$\mathrm{Hom}_R(A, R)$	The set of all nonzero R-algebra homomorphisms from A to R.
$\mathcal{M}_{+}(K)$	The set of all finite positive Borel measures on K.
$\mathcal{M}(K)$	The set of all finite signed Borel measures on K.
$\mathbb{N}$	Set of natural numbers: $0, 1, 2, 3, \ldots$.
$\mathbb{Z}$	Set of integer numbers: $\ldots, -2, -1, 0, 1, 2, \ldots$.
$\mathbb{Q}$	Set of rational numbers.
$\mathbb{R}$	Set of real numbers.

$\mathbb{C}$	Set of complex numbers, $\mathbb{R}[i]$.
PD	positive definite.
Positivstellensatz	A result which relates nonnegativity of a polynomial to its representations.
PSD	positive semidefinite.
$\mathrm{Psd}(K)$	$\{a \in A \ : \ \forall \alpha \in K \ \hat{a}(\alpha) \geq 0\}$.
RCF	the theory of real closed fields.
SDP	semidefinite programming.
$\sum A^2$	set of all finite sums of squares of elements of A.
SMP	strong moment property.
SOBS	sum of binomial squares.
SOS	sum of squares.
$\mathrm{Sper}(A)$	The real spectrum of A, i.e., the set of all orderings of A.
supp M	The ideal $M \cap -M$ where M is a quadratic module or a preordering.
$\mathcal{X}_A$	The set of all real valued ring homomorphisms on A, i.e., $\mathrm{Hom}(A, \mathbb{R})$.

Index