Introduction

The main purpose of this course is to explore the fascinating connection existing between positive polynomials and moment problems. The corner stone of this intimate relation is the famous Riesz-Haviland Theorem (proved by Riesz for the one-dimensional case in 1923 and by Haviland for higher dimensions in 1936), which establishes that the problem of characterizing the cone $\mathrm{Psd}(K)$ of all non-negative polynomials on a prescribed subset K of $\mathbb{R}^d$ is the dual facet of the so-called K-moment problem (KMP).

These two problems arose more or less contemporarily at the end of 19th century. In fact, Hilbert's theorem about sum of squares representations of non-negative forms appeared in 1888 and the first formulation of the KMP is due to Stieltjes in 1894, even if moments were already applied by Chebysev, Krein and Markov in the 1880's in studying limit values of integrals. As the characterization of $\mathrm{Psd}(K)$ and the KMP are faces of the same coin, we could start our journey by looking at any of these two problems but, since this course builds up on the contents of the course "Real Algebraic Geometry I" and "Topological Vector Spaces" held during last semester, we are going to start with a quick overview about the main results concerning the fundamental question of characterizing the $\mathrm{Psd}(K)$ cone (e.g. Positivstellensätze, saturation of preorderings and quadratic modules, closure of even power modules, etc.). Then we will rigorously formulate the KMP for $K \subseteq \mathbb{R}^n$ with $n \in \mathbb{N}$, i.e. the problem of establishing whether or not a given sequence of real numbers is the moment sequence of a non-negative Radon measure on K. As Landau brilliantly summarized in [9]: "The moment problem is a classical question in analysis remarkable not only for its own elegance but also for the extraordinary range of subjects theoretical and applied which has illuminated". In this course, we will only discover a small part of the beauty of the moment problem. In particular, after proving the Riesz-Haviland Theorem, we will use it to connect the KMP to the $\mathrm{Psd}(K)$ cone and we will study in detail how the even-power representations/approximations of non-negative polynomials on K influenced the theory of the KMP and at the same time how some of them actually came exactly from the study of the KMP. We will focus only the full finite dimensional KMP for basic closed semi-algebraic sets, i.e. on the case when the starting sequence is infinite and K is a subset of $\mathbb{R}^n$ determined by finitely many polynomial inequalities. Particular attention will be given to the case when K is non-compact which is still open in many of its aspects. Last but not least, we also would like to introduce a very general version of the KMP, namely for linear functionals on any unital commutative real algebra and present some recent results and open problems. Indeed, both the theory of positive polynomials and the moment problem are far to be static and, despite the huge progress of the last 130 years, we can still agree with the statement of Diaconis of 1987 in [9]: "Much is known but still the theory is not up to the demands of the applications" and being motivated to go forward with further research on these topics!

Chapter 1

Positive Polynomials and Sum of Squares

1.1 The ring of multivariate polynomials

Let $n \in \mathbb{N}$. We denote the ring of polynomials in n variables and real coefficients by $\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \ldots, X_n]$. We denote by $0 \in \mathbb{R}[\underline{X}]$ the polynomial with all coefficients equal to zero and by convention we set $\deg(0) = -\infty$. Let us recall some fundamental properties of $\mathbb{R}[\underline{X}]$.

Proposition 1.1.1. *Let $f, g \in \mathbb{R}[\underline{X}]$ s.t. $f \not\equiv 0$ and $g \not\equiv 0$. Then*
(i) $\deg(fg) = \deg(f) + \deg(g)$,
(ii) $\deg(f+g) \leq \max\left\{\deg(f), \deg(g)\right\}$,
(iii) $\deg(f+g) = \max\left\{\deg(f), \deg(g)\right\}$, *if* $\deg(f) \neq \deg(g)$.

Note that $\mathbb{R}[\underline{X}]$ is a real vector space of countable dimension, since a basis is $\{\underline{X}^\alpha \mid \alpha \in \mathbb{N}_0^n\}$ where $\underline{X}^\alpha := X_1^{\alpha_1} \ldots X_n^{\alpha_n}$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$. In fact, $\mathbb{R}[\underline{X}]$ can be written as a countable union of finite dimensional vector spaces, i.e. $\mathbb{R}[\underline{X}] := \bigcup_{d=0}^{\infty} \mathbb{R}[\underline{X}]_d$ where each $\mathbb{R}[\underline{X}]_d := \{f \in \mathbb{R}[\underline{X}] | \deg(f) \leq d\}$ has dimension $\binom{d+n}{n}$. This structure naturally carries a topology on $\mathbb{R}[\underline{X}]$, which makes it into a Hausdorff topological vector space, namely the *finite topology* τ_f defined by: $U \subseteq \mathbb{R}[\underline{X}]$ is open w.r.t. τ_f iff $\forall d \in \mathbb{N}_0$, $U \cap \mathbb{R}[\underline{X}]_d$ is open in $\mathbb{R}[\underline{X}]_d$ endowed with the euclidean topology (see e.g. [4, Section 4.5] for more details). As we showed in [4, Theorem 4.5.3], τ_f is the finest locally convex topology on $\mathbb{R}[\underline{X}]$ and so every linear functional on $\mathbb{R}[\underline{X}]$ is τ_f-continuous (see [4, Theorem 4.4.3]). This property will be particularly interesting in the study of the n-dimensional moment problem.

Definition 1.1.2. *A polynomial is said to be* homogenous *or* form *if it is the zero polynomial or a linear combination of monomials with same finite degree.*

For $n \in \mathbb{N}$ and $d \in \mathbb{N}_0$, we denote by $\mathcal{F}_{n,d}$ the set of all forms in n variables of degree d, i.e. $\mathcal{F}_{n,d} = \{f \in \mathbb{R}[X_1, \ldots, X_n] \mid f$ is a form and $\deg(f) = d\}$ which is also called set of n-ary d-ics forms.

The set $\mathcal{F}_{n,m}$ is a finite dimensional real vector space of dimension $\binom{d+n-1}{n-1}$.

Definition 1.1.3. *Let $p \in \mathbb{R}[X_1, \ldots, X_n]$ with $\deg(p) = d$. The* homogenization *p_h of p is defined as*

$$p_h(X_0, X_1 \ldots, X_n) := X_0^d \; p\left(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}\right).$$

Note that p_h is a homogeneous polynomial of degree d and in $n+1$ variables i.e. $p_h \in \mathcal{F}_{n+1,d}$.

1.2 When is a psd polynomial a sos?

In this section, we are going to consider the fundamental question of when a non-negative polynomial on $\mathbb{R}^n$ can be written as sum of squares of polynomials in $\mathbb{R}[\underline{X}]$.

Definition 1.2.1. *For $p \in \mathbb{R}[\underline{X}]$ we say that*

- *p is* positive semidefinite (psd) *if $p(x) \geq 0$ for all $x \in \mathbb{R}^n$.*
- *p is a* sum of squares (sos) *if $p = \sum_{i=1}^{s} h_i^2$ for some $s \in \mathbb{N}$ and $h_i \in \mathbb{R}[\underline{X}]$ for $i = 1, \ldots, s$.*

We denote by $\mathrm{Psd}(\mathbb{R}^n)$ the cone of all psd polynomials in $\mathbb{R}[\underline{X}]$ and by $\sum \mathbb{R}[\underline{X}]^2$ the cone of all sos of polynomials in $\mathbb{R}[\underline{X}]$.

Clearly, for any $n \in \mathbb{N}$ we always have that $\sum \mathbb{R}[\underline{X}]^2 \subseteq \mathrm{Psd}(\mathbb{R}^n)$. Hence, it is natural to ask for which $n \in \mathbb{N}$ the converse also holds, i.e. when we have that $\mathrm{Psd}(\mathbb{R}^n) = \sum \mathbb{R}[\underline{X}]^2$.

While for $n = 1$, it is easy to show that $\mathrm{Psd}(\mathbb{R}) = \sum \mathbb{R}[X]^2$ (see [7, Proposition 1.2.1]), for $n \geq 2$ it was known already to Hilbert in 1888 that not every psd polynomial is a sos in $\mathbb{R}[\underline{X}]$. Indeed, Hilbert provided a complete characterization of all psd polynomials which are sos. He actually restricted himself only to forms because the property of psd-ness and sos-ness are preserved under homogenization, i.e. for any $p \in \mathbb{R}[\underline{X}]_d$ we have that:

- p is psd iff p_h is psd,
- p is sos iff p_h is sos,

where p_h denotes the homogenization of p (see Definition 1.1.3).

We denote by $\mathcal{P}_{n,d}$ the set of all forms in $\mathcal{F}_{n,d}$ which are psd, and by $\sum_{n,d}$ set of all forms in $\mathcal{F}_{n,d}$ which are sos. It is easy to show that if $p \in \sum_{n,2d}$, then every sos representation of p consists only of homogeneous polynomials of degree d, i.e. $p \in \mathcal{F}_{n,2d},\ p = \sum_{i=1}^{s} p_i^2 \Rightarrow p_i \in \mathcal{F}_{n,d}$.

In [3] Hilbert proved the following result (for a proof see e.g. [5, Lectures 21,22,23], [1, Section 6.3]).

Theorem 1.2.2. *$\mathcal{P}_{n,d} = \sum_{n,d}$ iff*
(i) $n = 2$ [i.e. binary forms] or
(ii) $d = 2$ [i.e. quadratic forms] or
(iii) $(n, d) = (3, 4)$ [i.e. ternary quartics].

The proof of Hilbert was not constructive but in 1927 Motzkin provided the first concrete example of psd form which is not a sos. In addition to Motzkin's form several other examples have been considered. We provide here a short list of the most known ones (for the proofs and references to the original papers see e.g. [5, Lecture 23] and [7, Section 1.2]).

- Motzkin (1927): $z^6 + x^4y^2 + x^2y^4 - 3x^2y^2z^2 \in \mathcal{P}_{3,6} - \sum_{3,6}$
- Robinson (1969): $x^6 + y^6 + z^6 - (x^4y^2 + x^4z^2 + y^4x^2 + y^4z^2 + z^4x^2 + z^4y^2) + 3x^2y^2z^2 \in \mathcal{P}_{3,6} - \sum_{3,6}$.
- Robinson (1969): $w^4 + x^2y^2 + y^2z^2 + x^2z^2 - 4xyzw \in \mathcal{P}_{4,4} - \sum_{4,4}$
- Choi and Lam (1977): $1 + x^2y^2 + y^2z^2 + z^2x^2 - 4xyz \in \mathcal{P}_{3,6} - \sum_{3,6}$.

Hilbert's theorem and these examples which concretely show that in general $\sum \mathbb{R}[\underline{X}]^2 \subsetneq \mathrm{Psd}(\mathbb{R}^n)$ naturally led to relax the original question and investigate when a psd polynomial can be represented (or approximated) by polynomials whose non-negativity is "more evident", e.g. elements of even power modules of $\mathbb{R}[\underline{X}]$. Actually, the need of looking to these further cones in $\mathbb{R}[\underline{X}]$ becomes even more natural when we analyze the more general question of characterizing the cone of all polynomials in $\mathbb{R}[\underline{X}]$ which are non-negative on a prescribed subset K of $\mathbb{R}^n$.

Definition 1.2.3. *Let $K \subseteq \mathbb{R}^n$. A polynomial $p \in \mathbb{R}[\underline{X}]$ is said to be* positive semidefinite *on K if $p(x) \geq 0$ for all $x \in K$. We denote by* $\mathrm{Psd}(K)$ *the cone of all polynomials which are psd on K, i.e.*

$$\mathrm{Psd}(K) := \{p \in \mathbb{R}[\underline{X}] : p(x) \geq 0, \forall x \in K\}.$$

The following results on polynomials in one variable which are psd on intervals were most probably already known in the early 19th century (see [10] for some discussion on the history of such results) as easy consequences of the fundamental theorem of algebra.

Proposition 1.2.4.
a) $\mathrm{Psd}(\mathbb{R}^+) = \{\sigma_1 + X\sigma_2 : \sigma_1, \sigma_2 \in \sum \mathbb{R}[X]^2\}$.
b) For $a, b \in \mathbb{R}$ *with* $a < b$*, we have*
$\mathrm{Psd}([a,b]) = \{\sigma_1 + (b-x)\sigma_2 + (x-a)\sigma_3 : \sigma_1, \sigma_2, \sigma_3 \in \sum \mathbb{R}[X]^2\}$.

Proof. Set $Q := \sum \mathbb{R}[X]^2 + X\sum \mathbb{R}[X]^2$. Clearly, $Q \subseteq \mathrm{Psd}(\mathbb{R}^+)$ and $Q \cdot Q \subseteq Q$. We want to show that $\mathrm{Psd}(\mathbb{R}^+) \subseteq Q$.

Let $0 \not\equiv p \in \mathrm{Psd}(\mathbb{R}^+)$. By the fundamental theorem of algebra, p has the following factorization into irreducibles:

$$p = a\prod_{j=1}^{r}(X-\alpha_j)^{n_j}\prod_{k=1}^{s}\left[(X-u_k)^2 + v_k^2\right]^{\ell_j}, \tag{1.1}$$

where $r, s, n_1, \ldots, n_r, \ell_1, \ldots, \ell_s \in \mathbb{N}$, $a, \alpha_1, \ldots, \alpha_r, u_1, \ldots, u_s, v_1, \ldots, v_s \in \mathbb{R}$ with $\alpha_j \neq \alpha_i$ whenever $j \neq i$ and $\lambda_k := u_k + iv_k$ s.t. $\lambda_k \neq \lambda_i$ and $\lambda_k \neq \overline{\lambda_i}$ whenever $k \neq i$. Since Q is closed under multiplication, it is enough to show that all factors in (1.1) belong to Q. Clearly, $(X-u_k)^2 + v_k^2 \in Q$ and for n_j even also $(X-\alpha_j)^{n_j} \in Q$, so we just need to show that

$$a \prod_{\substack{j \in \{1,\ldots,r\} \\ \text{s.t. } n_j \text{ odd}}}^{r} (X-\alpha_j)^{n_j} \in Q.$$

As $p(x) \geq 0$ for all $x \in \mathbb{R}^+$, we obtain that $a > 0$ by letting $x \to +\infty$ and so $a \in Q$. Also, if α_j is a real root of p with odd multiplicity n_j then p must change sign in a neighbourhood of α_j and so $\alpha_j \leq 0$, which gives in turn that $X - \alpha_j = (-\alpha_j) + X \cdot 1^2 \in Q$. Hence, $p \in Q$.

For a proof of b) see e.g. [12, Proposition 3.3]. □

Proposition 1.2.4 shows that for $K = \mathbb{R}^+$ or $K = [a,b]$ the cone $\mathrm{Psd}(K)$ actually coincides with a certain quadratic module. Let us define such an object for any unital commutative ring.

Definition 1.2.5. *Let* A *be a commutative ring with* 1*,* $d \in \mathbb{N}$ *and denote by* $\sum A^{2d}$ *the set of all finite sums* $\sum a_i^{2d}$*,* $a_i \in A$*.*
a) A $2d$-power module M *in* A *is a subset* $M \subseteq A$ *such that* $M + M \subseteq M$*,* $a^{2d}M \subseteq M \;\forall\, a \in A, 1 \in M$*.*
b) A $2d$-power preordering T *in* A *is a* $2d$*-power module such that* $T \cdot T \subseteq T$*.*
In the case $d = 1$*,* $2d$*-power modules (resp.,* $2d$*-power preorderings) are referred to as* quadratic modules *(resp.,* quadratic preorderings*)*

Clearly, $\sum A^{2d}$ is a $2d$-power module in A and it is actually the unique smallest one. As $\sum A^{2d}$ is closed under multiplication, we have that $\sum A^{2d}$ is also the unique smallest $2d$-power preordering of A.

Definition 1.2.6. *Let A be a commutative ring with 1 and $d \in \mathbb{N}$. For an arbitrary family $S := \{g_j\}_{j\in J}$ of elements in A (note that J is an arbitrary index set possibly uncountable), the $2d$-power module of A generated by S is defined as*

$$M_S = \left\{\sigma_0 + \sigma_1 g_{j_1} + \ldots + \sigma_s g_{j_s} : s \in \mathbb{N}, j_1, \ldots, j_s \in J, \sigma_0, \ldots, \sigma_s \in \sum A^{2d}\right\}$$

while the $2d$-power preordering of A generated by S as

$$T_S := \left\{ \sum_{e=(e_1,\ldots,e_s)\in\{0,1\}^s} \sigma_e \; g_1^{e_1} \ldots g_s^{e_s} : s \in \mathbb{N}, j_1, \ldots, j_s \in J, \sigma_e \in \sum A^{2d}, \forall e \in \{0,1\}^s \right\}.$$

Note that for a fixed $d \in \mathbb{N}$ and $S \subseteq A$ we have $M_S \subseteq T_S \subseteq \mathrm{Psd}(K_S)$. We say that a module (resp. a preordering) $M \subseteq A$ is *finitely generated* if there exist a finite subset $S \subseteq A$ such that $M = M_S$. For example: ΣA^2 is finitely generated with $S = \emptyset$.

Let us come back now to the ring of polynomials in n variables $\mathbb{R}[\underline{X}]$ and to the question of relating the cone $\mathrm{Psd}(K)$ to even power modules in $\mathbb{R}[\underline{X}]$. In the light of the definitions above, we can restate Proposition 1.2.4 by saying that $\mathrm{Psd}(\mathbb{R}^+)$ coincide with the quadratic preordering generated by $\{X\}$, and that for any $a, b \in \mathbb{R}$ with $a < b$ the cone $\mathrm{Psd}([a, b])$ is the quadratic module generated by $\{b - X, X - a\}$. One can also easily see that $\mathbb{R}^+ = \{x \in \mathbb{R} : p(x) \geq 0\}$ and $[a, b] = \{x \in \mathbb{R} : r(x) \geq 0, q(x) \geq 0\}$ with $p := X, r := b - X, q := X - a$.

This leads us to focus our attention on the special class of closed subsets of $\mathbb{R}^n$ having this same structure.

Definition 1.2.7. *Given $S := \{g_1, \ldots, g_s\} \subset \mathbb{R}[\underline{X}]$, we call the following subset of $\mathbb{R}^n$ the* basic closed semialgebraic set (bcsas) *generated by S:*

$$K_S := \{x \in \mathbb{R}^n : g_i(x) \geq 0,\ i = 1, \ldots, s\}.$$

Hence, we are naturally brought to ask whether for any bcsas K_S of $\mathbb{R}^n$ we have $\mathrm{Psd}(K_S) = T_S$ (resp. $\mathrm{Psd}(K_S) = M_S$), where T_S (resp. M_S) is the quadratic preordering (resp. module) associated to S.

We already know that this is not always true, because for $S = \emptyset$ we have $K_S = \mathbb{R}^n$, $T_S = M_S = \sum \mathbb{R}[\underline{X}]^2$ and we have already seen that for $n \geq 2$ it does not always hold $\mathrm{Psd}(\mathbb{R}^n) = \sum \mathbb{R}[\underline{X}]^2$. However, the results in Proposition 1.2.4 give already a motivation for investigating more deeply the connection between $\mathrm{Psd}(K_S)$, T_S and M_S for finite $S \subset \mathbb{R}[\underline{X}]$.