High Degree Perturbations of Nonnegative Polynomials

Diploma Thesis

University of Konstanz

Department of Mathematics and Statistics

Thesis advisor: Prof. Dr. A. Prestel

Tim Netzer

June 2005

Contents

Chapter 1

Introduction

1.1 Outline

It is a well known fact that not every real polynomial in n indeterminates nonnegative on $\mathbb{R}^n$ can be written as a sum of squares of polynomials. For example, it is easy to see that the so called Motzkin polynomial

$$X^2Y^2(X^2 + Y^2 - 3) + 1 \in \mathbb{R}[X, Y]$$

is globally nonnegative, but not a sum of squares.

However, Berg, Christensen and Ressel proved that the set of sums of squares of polynomials is dense in the set of polynomials that are nonnegative on the cube $[-1, 1]^n$, with respect to the norm

$$\|p\|_1 := \sum_{\alpha \in \mathbb{N}^n} |p_\alpha|$$

for polynomials $p = \sum_{\alpha \in \mathbb{N}^n} p_\alpha X_1^{\alpha_1} \cdots X_n^{\alpha_n} \in \mathbb{R}[X_1, ..., X_n]$ (see Theorem 4.1.8 in the appendix; it is from [BCR2], Theorem 9.1).

This of course also shows the weaker statement that the set of sums of squares lies dense in the set of polynomials that are globally nonnegative.

The denseness result by Berg, Christensen and Ressel is proven in a completely nonconstructive way. Now Lasserre gave a proof that shows the denseness of the set of sums of squares in the set of globally nonnegative polynomials in a much more explicit way (see [Las1]). By using results from functional analysis and optimization theory, he showed that a certain small perturbation of some coefficients up to a specific degree (which is unfortunately not known in general) makes a nonnegative polynomial a sum of squares.

This proof is the starting point of our work. By using an idea from Marshall (see [Mar]) and by changing Lasserre's proof a little bit, we will show a similar result for polynomials that are nonnegative only on an algebraic set, and so we will get a generalization of Lasserre's result. We will not use any results from optimization theory but show all the claims directly.

Indeed, Lasserre gave an even more general result in [Las2]. But our proof avoids completely any theory about infinite dimensional optimization and is therefore easier to understand.

In the second section of Chapter 2, we give a new proof for the general denseness result by Berg, Christensen and Ressel. In a similar way to Lasserre, we show how a small perturbation of a polynomial nonnegative on a cube makes it a sum of squares. The same solution to the so called moment problem that was used by Berg, Christensen and Ressel in their work is also crucial for our proof. But our result shows the denseness in a more explicit way than their result does.

Unfortunately, the degree up to which we have to perturbe the polynomial is not known in any of the results. In Chapter 3, we deal with the question which factors influence this degree. Roughly spoken, we find that this degree does not depend on the specific polynomial but only on its degree and the size of its coefficients. This is obtained by generalizing the results from Chapter 2 to real closed extension fields of $\mathbb{R}$ and by using the so called $\aleph_1$-saturation of an ultrapower of $\mathbb{R}$ (a similar approach to get degree bounds for other problems can for example be found in [PD], Section 8.2).

For a better understanding of the main ideas, we will give a short sketch of Lasserre's basic proof in the next section. The exact and more general proof is given in Chapter 2.

I would like to thank Dr. Markus Schweighofer very much for his constant effort, support and good ideas which helped me a lot.

1.2 Main idea of Lasserre's proofs

As already mentioned in the Introduction, the starting point of this work is a result about approximation of nonnegative polynomials by sums of squares, given by Lasserre in [Las1].

The result in its simplest form is the following (compare to Corollary 2.1.11 in the second chapter):

Theorem 1.2.1. *Let $f \in \mathbb{R}[X_1, ..., X_n]$ be nonnegative on $\mathbb{R}^n$. Then for every $\varepsilon > 0$ there is an $r \in \mathbb{N}$ such that*

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{k!} X_i^{2k} \in \sum \mathbb{R}[X_1, ..., X_n]^2.$$

Note that Theorem 1.2.1 shows the denseness of the cone of sums of squares of polynomials in the cone of nonnegative polynomials with respect to the $\|.\|_1$-norm on $\mathbb{R}[X_1, ..., X_n]$. For a polynomial

$$g := \sum_{\alpha \in \mathbb{N}^n} g_\alpha X_1^{\alpha_1} \cdots X_n^{\alpha_n},$$

where almost all $g_\alpha \in \mathbb{R}$ equal zero, the $\|.\|_1$-norm is defined as

$$\|g\|_1 := \sum_{\alpha \in \mathbb{N}^n} |g_\alpha|.$$

So one has

$$\left\| \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{k!} X_i^{2k} \right\|_1 \leq n\varepsilon \sum_{k=0}^{\infty} \frac{1}{k!} = n\varepsilon \exp(1) \stackrel{\varepsilon \to 0}{\longrightarrow} 0.$$

This shows that f from Theorem 1.2.1 can be approximated as closely as desired by a sum of squares (although r might perhaps depend badly on ε).

From now on we will write $\mathbb{R}[X]$ instead of $\mathbb{R}[X_1, ..., X_n]$ most of the time, but always mean n indeterminates. By $\mathbb{R}[X]_r$ we will denote the finite dimensional real vector space consisting of all polynomials of degree smaller or equal to r. Of course, we have $\mathbb{R}[X]_r \subset \mathbb{R}[X]$. Now the idea of the proof of Theorem 1.2.1 is to look at two specific optimization problems. For a fixed polynomial $f \in \mathbb{R}[X]$, the first one, $Q_{r,M}$, is defined for $r \in \mathbb{N}$, $2r \geq \deg f$ and $M \in \mathbb{R}_{>0}$ as follows:

$$Q_{r,M}: \quad \begin{cases} & \text{minimize } \ L(f) \\ s.t. & L : \mathbb{R}[X]_{2r} \longrightarrow \mathbb{R} \text{ linear} \\ & L(1) = 1 \\ & L(ne^{M^2} - \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{k!} X_i^{2k}) \geq 0 \\ & L(p^2) \geq 0 \quad \forall p \in \mathbb{R}[X]_r \end{cases}$$

This optimization problem $Q_{r,M}$ can be written as a so called *semidefinite optimization problem.*

Semidefinite optimization problems are generalizations of linear optimization problems, see for example [VB] or [Tod] for a definition. Now there is a so called duality theory for such problems. This means, for a semidefinite (minimization) problem (P), there is a canonical way to construct a semidefinite (maximization) problem (D), which always has an optimal value smaller or equal to the optimal value of (P). Under certain conditions even *strong duality* holds, i.e. these optimal values are equal (see for example [VB], [Tod], or in a more abstract way [NN]).

If we construct the dual problem to $Q_{r,M}$, we will get the following problem, which we will call $Q^*_{r,M}$:

$$Q^*_{r,M}: \quad \begin{cases} & \text{maximize } z \\ s.t. & z \in \mathbb{R} \\ & f - z \in \sum \mathbb{R}[X]_r^2 + \mathbb{R}_{\geq 0}\left(ne^{M^2} - \sum_{i=1}^n \sum_{k=0}^r \frac{1}{k!} X_i^{2k}\right) \end{cases}$$

Here, $\sum \mathbb{R}[X]_r^2$ denotes the set of sums of squares of polynomials from $\mathbb{R}[X]_r$. It is a subset of $\mathbb{R}[X]_{2r}$.

It is not necessary for our purpose to make sure that this is really the canonical dual to $Q_{r,M}$, as we will prove the following result in a direct way (see Lemma 2.1.6), without using any theory of semidefinite programming:

$$\sup Q^*_{r,M} \geq \inf Q_{r,M}. \tag{1.1}$$

In (1.1), $\sup Q^*_{r,M}$ and $\inf Q_{r,M}$ denote the optimal values of $Q^*_{r,M}$ and $Q_{r,M}$, respectively.

Note that (1.1) is the interesting part of the strong duality, as the other inequality ($\sup Q^*_{r,M} \leq \inf Q_{r,M}$, the so called weak duality) is quite easy to obtain, but will not be needed for our proof.

The most important fact for the proof of (1.1) is the closedness of

$$\sum \mathbb{R}[X]_r^2 + \mathbb{R}_{\geq 0}\left(ne^{M^2} - \sum_{i=1}^n \sum_{k=0}^r \frac{1}{k!} X_i^{2k}\right)$$

in $\mathbb{R}[X]_{2r}$ with respect to the topology induced by any norm on the finite-dimensional vector space $\mathbb{R}[X]_{2r}$.

(1.1) is the first important result for the proof of Theorem 1.2.1. The next important step is to show that for every $M \in \mathbb{R}_{>0}$, the increasing sequence $(\inf Q_{r,M})_{r \in \mathbb{N}, 2r \geq \deg f}$ approximates at least the infimum of the polynomial f

on $\mathbb{R}^n$. Indeed, we will show:

$$\forall M > 0\ \exists y_M \in \mathbb{R} : y_M \geq \inf_{x\in\mathbb{R}^n} f(x) \text{ and } \inf Q_{r,M} \overset{r\to\infty}{\longrightarrow} y_M. \tag{1.2}$$

This is done in Proposition 2.1.9. The main idea is the following: We will show first that $Q_{r,M}$ always has an optimal solution, i.e. the infimum (and therefore minimum) is attained in some feasible linear form $L^{(r)}$. Then, for fixed M, we consider a sequence $(L^{(r)})_{2r\geq \deg f}$, where every $L^{(r)}$ is an optimal solution for $Q_{r,M}$. By identifying each linear form with the (finite) tuple of its values on the canonical monomial basis of $\mathbb{R}[X]_{2r}$, and filling up with zeros, we make all of them elements of $\mathbb{R}^{\mathbb{N}^n}$, and therefore get a sequence in $\mathbb{R}^{\mathbb{N}^n}$. Next we will show that this sequence is bounded in every component, and by using Tychonoff's Theorem we can show that there must be a subsequence converging pointwise to some element $(y_\alpha)_{\alpha\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$. Now from the construction of $Q_{r,M}$ and the pointwise convergence, we will find this $(y_\alpha)_{\alpha\in\mathbb{N}^n}$ to be the moment sequence of a probability measure μ on $\mathbb{R}^n$, in other words, the linear form on $\mathbb{R}[X]$, defined on the canonical monomial basis by $(y_\alpha)_{\alpha\in\mathbb{N}^n}$, is integration with respect to this measure μ. This is obtained by using Nussbaums solution to the moment problem under *Carleman's Condition*, which is stated as Theorem 4.1.5 in the appendix. The proof of Nussbaums's result is not easy and uses functional analysis.

From the pointwise convergence we will get that the subsequence of optimal values that corresponds to the converging subsequence of $(L^{(r)})_r$, converges to $\int_{\mathbb{R}^n} f d\mu$, which is greater or equal to $\inf_{x\in\mathbb{R}^n} f(x)$, as μ is a probability measure. Since the whole sequence $(\min Q_{r,M})_{2r\geq \deg f}$ is monotonously increasing, we will get this convergence for the whole sequence, which was to be shown.

But with (1.1) and (1.2) holding, the result of Theorem 1.2.1 is quite easy to obtain.

If $\inf_{x\in\mathbb{R}^n} f(x) > 0$, then (1.1) and (1.2) gives us by definition of $Q^*_{r,M}$ a

$z > 0$ such that

$$f - z = \sum_j p_j^2 + \delta \left(ne^{M^2} - \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{k!} X_i^{2k} \right)$$

for some $p_j \in \mathbb{R}[X]$ and $\delta \geq 0$, if only r is big enough. So considering z and δne^{M^2} as sums of squares, one will get the result

$$f + \delta \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{k!} X_i^{2k} \in \sum \mathbb{R}[X]^2.$$

But a little computation then shows that δ will become as small as desired, if the M in the optimization problems was chosen big enough.

If $\inf_{x \in \mathbb{R}^n} f(x) = 0$, then by adding some small real number to f, we can go back to the first case and easily get the same result.

This is the main idea of the proof. In the first section of the following chapter, we will state it more generally (see Theorem 2.1.10). We will assume that f is nonnegative only on an algebraic set in $\mathbb{R}^n$. We have to change the optimization problems a little bit and have to introduce a third optimization problem, but the procedure of the proof will stay the same. The result is different from Theorem 1.2.1 in the way that the perturbed polynomial f is not a sum of squares but a sum of squares plus a polynomial vanishing on the given algebraic set. Of course, we will get in particular Theorem 1.2.1 from this result.

In the second section of the following chapter, the same proof is modified to be able to apply another solution to the moment problem, instead of Nussbaum's solution. The used result is stated in the appendix as well and uses boundedness by so called *absolute values* (see Corollary 4.1.7). The theorem obtained in this way is different to Theorem 1.2.1 and gives a more explicit version of the known result, that the sums of squares are $\|\cdot\|_1$-dense in the cone of polynomials nonnegative on a big enough centered cube in $\mathbb{R}^n$.

Chapter 2

Sums of Squares Approximation

2.1 Polynomials nonnegative on an algebraic set

To begin, some definitions and terms that will be used during the following sections are given.

First, we assume $\mathbb{N} = \{0, 1, 2, ...\}$ for the rest of this work. Further let $\mathbb{R}[X] = \mathbb{R}[X_1, ..., X_n]$ denote the ring of real polynomials in n indeterminates and let $\mathbb{R}[X]_r$ be the subspace of polynomials of degree smaller or equal to r.

Fix polynomials $f, g_1, ..., g_s \in \mathbb{R}[X]$ and let

$$N = N(g_1, ..., g_s) := \{x \in \mathbb{R}^n \mid g_i(x) = 0 \text{ for } i = 1, ..., s\}$$

denote the common zero set of the g_i in $\mathbb{R}^n$. We will always assume $N \neq \emptyset$.

Let

$$\text{rrad}(g_1, ..., g_s) \subseteq \mathbb{R}[X]$$

denote the real radical of the polynomials $g_1, ..., g_s$. There are several characterizations of this real radical.

In general, the real radical of an ideal I in a ring A can be defined as

$$\mathrm{rrad}(I) := \left\{ a \in A \mid a^{2e} + t \in I \text{ for some } e \in \mathbb{N} \text{ and } t \in \sum A^2 \right\}.$$

Now the real radical of the polynomials $g_1, ..., g_s$ is just the real radical of the ideal generated by those polynomials in $\mathbb{R}[X]$. But in our case there is also a geometric characterization of the real radical: The real radical of $g_1, ..., g_s$ consists of all real polynomials vanishing on the real zero set of the $g_1, ..., g_s$, see for example [PD]. We will use both characterizations in this work.

In the following, $\mathrm{rrad}(g_1, ..., g_s)_r$ will denote $\mathrm{rrad}(g_1, ..., g_s) \cap \mathbb{R}[X]_r$.

Let

$$\overline{\mathbb{R}[X]} := \{p + \mathrm{rrad}(g_1, ..., g_s) \mid p \in \mathbb{R}[X]\}$$

denote the factor ring of $\mathbb{R}[X]$ modulo the ideal $\mathrm{rrad}(g_1, ..., g_s)$, and $\overline{\mathbb{R}[X]}_r$ the image of $\mathbb{R}[X]_r$ under the canonical projection. For all $r \geq 0$, $\overline{\mathbb{R}[X]}_r$ is a finite dimensional real vector space of dimension at least one.

In this chapter, $\overline{p}$ will always denote the coset of a polynomial p mod $\mathrm{rrad}(g_1, ..., g_s)$. As $\mathrm{rrad}(g_1, ..., g_s) \subsetneq \mathbb{R}[X]$ $(N \neq \emptyset)$, we will often write x instead of $\overline{x}$ for elements $x \in \mathbb{R}$.

Now the coefficients $\frac{1}{k!}$ from Theorem (1.2.1) can be chosen more general than it is done in Lasserre's work:

Let $(a_k)_{k \in \mathbb{N}}$ be a sequence of positive real numbers, satisfying that the power series

$$h(x) := \sum_{k=0}^{\infty} a_k x^k$$

is convergent for all $x \in \mathbb{R}$ and in addition

$$\sum_{k=0}^{\infty} a_k^{\frac{1}{2k}} = \infty,$$

as well as $a_0 = 1$. For example, by taking $a_k = \frac{1}{k!}$, one gets $h(x) = e^x$ and all the conditions are satisfied. Indeed, as $k! \le k^k$, we have

$$\sum_{k=0}^{\infty} \frac{1}{k!}^{\frac{1}{2k}} \ge \sum_{k=0}^{\infty} \frac{1}{k^k}^{\frac{1}{2k}} = \sum_{k=0}^{\infty} \frac{1}{\sqrt{k}} = \infty.$$

This function h replaces the exponential function which was used in the introduction.

In addition, we abbreviate

$$\Theta_r := \sum_{i=1}^{n} \sum_{k=0}^{r} a_k X_i^{2k}.$$

Obviously $\Theta_r \in \mathbb{R}[X]_{2r}$.

Further let $M_0 \in \mathbb{R}_{>0}$ be big enough to ensure $N \cap B_{M_0} \neq \emptyset$, where $B_M := [-M, M]^n$ denotes the closed M-ball with respect to the sup-norm on $\mathbb{R}^n$.

Last, let $r_0 \in \mathbb{N}$ be an integral number that satisfies

$$2r_0 \ge \max(\deg f, \deg\ g_1^2, ..., \deg\ g_s^2).$$

From now on, r and M are always supposed to be bigger than r_0 and M_0, respectively (unlike otherwise mentioned).

Now we define three optimization problems that play a crucial role for the following results.

First of all:

$$Q_{r,M}: \quad \begin{cases} \quad \text{minimize } \; L(f) \\ s.t. \quad L : \mathbb{R}[X]_{2r} \longrightarrow \mathbb{R} \text{ linear} \\ \quad L(1) = 1 \\ \quad L(nh(M^2) - \Theta_r) \geq 0 \\ \quad L(\text{rrad}(g_1, ..., g_s) \cap \mathbb{R}[X]_{2r}) = 0 \\ \quad L(p^2) \geq 0 \quad \forall p \in \mathbb{R}[X]_r \end{cases}$$

Using an idea from [Mar], the next optimization problem is obtained from the first one by passing over from $\mathbb{R}[X]_{2r}$ to $\overline{\mathbb{R}[X]}_{2r}$:

$$\overline{Q}_{r,M}: \quad \begin{cases} \quad \text{minimize } \; \overline{L}(\overline{f}) \\ s.t. \quad \overline{L} : \overline{\mathbb{R}[X]}_{2r} \longrightarrow \mathbb{R} \text{ linear} \\ \quad \overline{L}(1) = 1 \\ \quad \overline{L}(nh(M^2) - \overline{\Theta}_r) \geq 0 \\ \quad \overline{L}(\overline{p}^2) \geq 0 \quad \forall \overline{p} \in \overline{\mathbb{R}[X]}_r \end{cases}$$

From every feasible solution $\overline{L}$ for $\overline{Q}_{r,M}$, one gets a feasible solution $L := \overline{L} \circ \text{pr}$ (pr denotes the canonical projection from $\mathbb{R}[X]_{2r}$ to $\overline{\mathbb{R}[X]}_{2r}$) for $Q_{r,M}$. As

$$\overline{L}(\overline{f}) = \overline{L}(\text{pr}(f)) = L(f),$$

one obtains

$$\inf \; Q_{r,m} \leq \inf \; \overline{Q}_{r,M}.$$

Indeed, from the condition

$$L(\mathrm{rrad}(g_1, ..., g_s) \cap \mathbb{R}[X]_{2r}) = 0$$

in $Q_{r,M}$ we even get equality, but we won't need this.

The last problem is the so called *dual problem* to $\overline{Q}_{r,M}$:

$$\overline{Q}^*_{r,M} : \quad \begin{cases} & \text{maximize } z \\ s.t. & z \in \mathbb{R} \\ & \overline{f} - z \in \sum \overline{\mathbb{R}[X]}^2_r + \mathbb{R}_{\geq 0}(nh(M^2) - \overline{\Theta}_r) \end{cases}$$

We will now prove a result similar to (1.1), namely that $\sup \overline{Q}^*_{r,M}$ is greater or equal to $\inf \overline{Q}_{r,M}$. The main ideas are from [Mar] and [Sch].

First, some auxiliary results are needed. Unlike otherwise mentioned, we always assume finite dimensional real vector spaces to be equipped with the standard topology. All notions like *closed, continuous* etc. refer to this topology.

Lemma 2.1.1. *For $m, m' \in \mathbb{N} \setminus \{0\}$ let $\Phi : \mathbb{R}^m \longrightarrow \mathbb{R}^{m'}$ be a continuous map, homogeneous of degree $d \geq 1$, i.e. $\Phi(\lambda x) = \lambda^d \Phi(x)$ for all $\lambda \in \mathbb{R}$, $x \in \mathbb{R}^m$. Further suppose $\Phi^{-1}(0) = \{0\}$. Then Φ is closed.*

Proof. Let $K \subseteq \mathbb{R}^m$ be closed. Let $(x_l)_{l \in \mathbb{N}}$ be a sequence in K, satisfying $\lim_{l \to \infty} \Phi(x_l) = y$ for some $y \in \mathbb{R}^{m'}$. We have to show that $y \in \Phi(K)$. Assume without loss of generality $y \neq 0$ and therefore again without loss of generality all $x_l \neq 0$. As the unit sphere $S^{m-1} \subset \mathbb{R}^m$ is compact, the sequence defined by $\hat{x}_l := \frac{x_l}{\|x_l\|}$ has an accumulation point und therefore one gets a subsequence $(\hat{x}_{l_k})_{k \in \mathbb{N}}$ that converges to some $x \in S^{m-1}$. As Φ is homogeneous and $\Phi^{-1}(0) = \{0\}$, one obtains

$$\frac{1}{\|x_{l_k}\|^d} \Phi(x_{l_k}) = \Phi\left(\frac{x_{l_k}}{\|x_{l_k}\|}\right) \xrightarrow{k \to \infty} \Phi(x) \neq 0.$$

From this we see, as $(\Phi(x_{l_k}))_k$ is convergent and therefore bounded, that $(\|x_{l_k}\|)_k$ must be bounded as well. But this shows that there is a subsequence $(x_{l_{k_s}})_s$ that converges to some x' in K (as K is closed). So one gets

$$y = \lim_{l\to\infty} \Phi(x_l) = \lim_{s\to\infty} \Phi(x_{l_{k_s}}) = \Phi(x') \in \Phi(K).$$

□

Lemma 2.1.2. *Every element in $\sum \mathbb{R}[X]_r^2$ can be written as a sum of squares of length $s(r) := \dim_{\mathbb{R}} \mathbb{R}[X]_r$.*

Proof. Let $(X^\alpha)_{\alpha\in\mathbb{N}^n,|\alpha|\le r}$ be the canonical monomial basis of $\mathbb{R}[X]_r$. We give those elements an arbitrary but fixed order $X^{\alpha_1}, ..., X^{\alpha_{s(r)}}$. Define $\tilde{X}$ to be the $s(r)\times s(r)$-matrix with elements $\tilde{X}_{ij} := X^{\alpha_i+\alpha_j}$.

Now consider the space $S\mathbb{R}^{s(r)\times s(r)}$ of symmetric matrices with the inner product $\langle A, B\rangle := \mathrm{Tr}(AB) := \sum_{j,i=1}^{s(r)} A_{ij}B_{ji}$. Let $S\mathbb{R}_+^{s(r)\times s(r)} \subseteq S\mathbb{R}^{s(r)\times s(r)}$ denote the convex cone of positive semidefinite symmetric matrices. Then we have

$$\sum \mathbb{R}[X]_r^2 \subseteq \{\langle \tilde{X}, G\rangle \mid G \in S\mathbb{R}_+^{s(r)\times s(r)}\}. \tag{2.1}$$

Indeed, take $q \in \mathbb{R}[X]_r$ and identify it with the $s(r)$-vector of its coefficients. Then $qq^t \in S\mathbb{R}_+^{s(r)\times s(r)}$, and a little computation shows

$$\langle \tilde{X}, qq^t\rangle = q^t\tilde{X}q = q^2. \tag{2.2}$$

As the right hand set in (2.1) is closed under addition, (2.1) holds.

Furthermore we have

$$\{\ \langle \tilde{X}, G\rangle \mid G \in S\mathbb{R}_+^{s(r)\times s(r)}\} \subseteq \{\ q_1^2 + ... + q_{s(r)}^2 \mid q_i \in \mathbb{R}[X]_r\}. \tag{2.3}$$

If $G = uu^t$ for some $u \in \mathbb{R}^{s(r)}$, then (2.3) is clear from (2.2). Otherwise, use the fact that for every $G \in S\mathbb{R}_+^{s(r)\times s(r)}$ there exists an orthogonal $B \in \mathbb{R}^{s(r)\times s(r)}$ such that $G = BDB^t$, where D is a diagonal matrix with only nonnegative entries. So one can define $\sqrt{D}$ and has $G = (B\sqrt{D})(B\sqrt{D})^t$. By taking as u_i the i-th column of $B\sqrt{D}$, one gets $G = \sum_{i=1}^{s(r)} u_i u_i^t$, which together with the first case establishes (2.3).

Since the right hand side of (2.3) is obviously contained in the left hand side of (2.1), the proof is finished. □

In the next Proposition, we will need the notion of a real radical ideal: An ideal I is called a real radical ideal, if it is its own real radical. Obviously, the real radical of an ideal is a real radical ideal.

Proposition 2.1.3. *Let $I \subseteq \mathbb{R}[X]$ be a real radical ideal and let $\overline{\mathbb{R}[X]}_r$ denote the finite-dimensional real vector space of the cosets modulo I of the elements of $\mathbb{R}[X]_r$ (where $r \geq 1$). Then*

$$\sum \overline{\mathbb{R}[X]}_r^2 \subseteq \overline{\mathbb{R}[X]}_{2r}$$

is closed in the standard topology.

Proof. By Lemma 2.1.2, every $x \in \sum \overline{\mathbb{R}[X]}_r^2 = \overline{\sum \mathbb{R}[X]_r^2}$ can in particular be written as a sum of s(r) elements from $\overline{\mathbb{R}[X]}_r^2$. Define

$$\begin{array}{rcl} \Phi : \underbrace{\overline{\mathbb{R}[X]}_r \times ... \times \overline{\mathbb{R}[X]}_r}_{s(r)} & \longrightarrow & \overline{\mathbb{R}[X]}_{2r} \\ (\overline{h}_1, ..., \overline{h}_{s(r)}) & \longmapsto & \sum_{i=1}^{s(r)} \overline{h}_i^2 \end{array}$$

Φ is well defined, as I is an ideal. Obviously Φ is a continuous map of finite-dimensional $\mathbb{R}$-vector spaces and Φ is homogeneous of degree 2. As I is a real radical ideal, one gets $\Phi^{-1}(0) = \{0\}$, which, together with Lemma 2.1.1, proves Φ to be closed. As we have seen, $\sum \overline{\mathbb{R}[X]}_r^2$ is the codomain of Φ and therefore closed. □

Lemma 2.1.4. *For $m \geq 1$ let $K \subseteq \mathbb{R}^m$ be a closed convex cone, i.e. a closed convex set that is closed under multiplication with nonnegative real numbers. Let $v \in \mathbb{R}^m$ and $v \notin -K$. Then $K + \mathbb{R}_{\geq 0} v$ is closed.*

Proof. Let $K^* := \{y \in \mathbb{R}^m \mid \langle y, z\rangle \geq 0 \ \forall z \in K\}$ denote the so called *dual cone* to K. There must be a $y \in K^*$ satisfying $\langle v, y\rangle > 0$. Otherwise $\langle -v, y\rangle \geq 0$ for all $y \in K^*$. So by definition $-v \in (K^*)^*$. As the dual cone of the dual cone is the closure of the original cone, and in our case K is closed, one gets $-v \in K$, a contradiction.

So take such $y \in K^*$ and let H denote the hyperplane orthogonal to y.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in K, let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative real numbers and suppose $x_n + \lambda_n v \overset{n\to\infty}{\longrightarrow} x \in \mathbb{R}^m$. We have to show $x \in K + \mathbb{R}_{\geq 0} v$. As $y \in K^*$, and therefore $\langle x_n, y\rangle \geq 0$ for all $n \in \mathbb{N}$, one gets

$$0 \leq \lambda_n \langle v, y\rangle = \langle \lambda_n v, y\rangle \leq \langle x_n + \lambda_n v, y\rangle \overset{n\to\infty}{\longrightarrow} \langle x, y\rangle.$$

Since $\langle v, y\rangle > 0$, $(\lambda_n)_{n\in\mathbb{N}}$ must be bounded. So one gets a subsequence $(\lambda_{n_k})_{k\in\mathbb{N}}$ and some $v^* \in \mathbb{R}_{\geq 0} v$, such that $\lim_{k\to\infty} \lambda_{n_k} v = v^*$. But then, as the whole subsequence converges, $(x_{n_k})_{k\in\mathbb{N}}$ must converge to some $x^* \in K$ as well. But from this we get $x = x^* + v^* \in K + \mathbb{R}_{\geq 0} v$. $\square$

Remark: The condition $v \notin -K$ cannot generally be omitted in Lemma 2.1.4, which shows the following example that was given to me by Prof. C. Scheiderer. Consider

$$K := \{(x, y, z) \in \mathbb{R}^3 \mid x, y, z \geq 0, xy \geq z^2\} \subseteq \mathbb{R}^3.$$

K is a closed convex cone. For $v := (0, -1, 0) \in \mathbb{R}^3$ we have $v \in -K$, and $K + \mathbb{R}_{\geq 0} v$ is not closed. Indeed, for every $x > 0$ we have $(x, \frac{1}{x}, 1) \in K$ and thus $(x, 0, 1) \in K + \mathbb{R}_{\geq 0} v$. But $(0, 0, 1) \notin K + \mathbb{R}_{\geq 0} v$, as for all $y \in \mathbb{R}$, the elements $(0, y, 1)$ are not in K.

In our case, the closed convex cone K is $\sum \overline{\mathbb{R}[X]}_r^2$ and $v = nh(M^2) - \overline{\Theta}_r \in \overline{\mathbb{R}[X]}_{2r}$, as defined above. Suppose $nh(M^2) - \overline{\Theta}_r \in -\sum \overline{\mathbb{R}[X]}_r^2$. Take some $\hat{x} \in N \cap B_{M_0}$ and evaluate at $\hat{x}$. As $\hat{x} \in N$, this is well-defined, and as

$\hat{x} \in B_{M_0}$, we get

$$nh(M^2) - \Theta_r(\hat{x}) \geq nh(M^2) - \sum_{i=1}^{n}\sum_{k=0}^{r} a_k M_0^{2k} \geq nh(M^2) - nh(M_0^2) > 0,$$

as $M > M_0$. This contradicts $nh(M^2) - \overline{\Theta}_r \in -\sum \overline{\mathbb{R}[X]}_r^2$. So by Lemma 2.1.4, we get the closedness of $\sum \overline{\mathbb{R}[X]}_r^2 + \mathbb{R}_{\geq 0}(nh(M^2) - \overline{\Theta}_r) \subseteq \overline{\mathbb{R}[X]}_{2r}$. This is what we need for the proof of the following Lemma:

Lemma 2.1.5. *For all $r \geq 1$,*

$$\sum \overline{\mathbb{R}[X]}_r^2 + \mathbb{R}_{\geq 0}(nh(M^2) - \overline{\Theta}_r)$$

is the intersection of all halfspaces $\{\, \overline{p} \in \overline{\mathbb{R}[X]}_{2r} \mid \overline{L}(\overline{p}) \geq 0 \,\}$, where $\overline{L}$ ranges over all feasible solutions of $\overline{Q}_{r,M}$.

Proof. Let $U := \sum \overline{\mathbb{R}[X]}_r^2 + \mathbb{R}_{\geq 0}(nh(M^2) - \overline{\Theta}_r)$. It is obvious that U is a subset of all these halfspaces, because this is explicitly demanded in the optimization problem $\overline{Q}_{r,M}$. Therefore we only have to show that for $\overline{p} \notin U$, there exists a feasible solution $\overline{L}$ of $\overline{Q}_{r,m}$ that satisfies $\overline{L}(\overline{p}) < 0$. Choose an inner product on $\overline{\mathbb{R}[X]}_{2r}$. As U is closed with respect to the induced topology (see the above considerations), there exists $\overline{q} \in U$ with minimal distance to $\overline{p}$. Now choose a linear form $\overline{L}_0$ on $\overline{\mathbb{R}[X]}_{2r}$ whose kernel is the halfspace orthogonal to $\overline{p} - \overline{q}$ and that satisfies $\overline{L}_0(\overline{p} - \overline{q}) < 0$. Then we have $\overline{L}_0(\overline{q}) = 0$. This is obvious if $\overline{q} = 0$. Else, $\overline{q}$ must be the orthogonal projection of $\overline{p}$ to the line $\mathbb{R}_{\geq 0}\overline{q}$. Indeed, as U is a cone, there would be smaller distances from elements of U to $\overline{p}$, otherwise.

Now we have

$$\overline{L}_0(\overline{p}) = \overline{L}_0(\overline{p} - \overline{q} + \overline{q}) = \overline{L}_0(\overline{p} - \overline{q}) < 0.$$

And as U is a convex cone, every $u \in U$ with $\overline{L}_0(u) < 0$ would give elements in U, near $\overline{q}$ on the line segment to u, with smaller distances to $\overline{p}$, a contradiction. So we have $\overline{L}_0(u) \geq 0$ for all $u \in U$. If $\overline{L}_0(1) > 0$, we can get $\overline{L}_0(1) = 1$ by scaling with a positive real number, and therefore make $\overline{L}_0$ feasible for $\overline{Q}_{r,M}$. If $\overline{L}_0(1) = 0$, then take $\hat{x} \in N \cap B_{M_0}$, let $\varphi_{\hat{x}}$ be the (well defined) evaluation at $\hat{x}$ and consider $\overline{L} := \varphi_{\hat{x}} + \lambda \overline{L}_0$ for some $\lambda \in \mathbb{R}$ we will choose. Then we have $\overline{L}(u) = u(\hat{x}) + \lambda \overline{L}_0(u)$. As $M > \|\hat{x}\|_\infty$ (see the conventions on page 12), we have $u(\hat{x}) > 0$ for $u \in U$. Therefore $\overline{L}(u) \geq 0$ for all $u \in U$ whenever λ is nonnegative. Further, $\overline{L}(1) = 1 + \lambda \overline{L}_0(1) = 1$, so $\overline{L}$ is feasible for $\overline{Q}_{r,M}$ if λ is nonnegative. And $\overline{L}(\overline{p}) = \overline{p}(\hat{x}) + \lambda \overline{L}_0(\hat{x}) < 0$ if λ is big enough. □

As mentioned before, we finally get the strong duality:

Lemma 2.1.6. *For every $r \geq r_0$ and $M > M_0$, we have*

$$\sup \overline{Q}^*_{r,M} \geq \inf \overline{Q}_{r,M} \geq \inf Q_{r,M}.$$

Proof. The second inequality was proven on page 14. To see that the first inequality holds, take $x \in \mathbb{R}$ satisfying $\overline{f} - x \notin U$ (U as defined in the proof of Lemma 2.1.5). So by Lemma 2.1.5, one gets a linear form $\overline{L}$, feasible for $\overline{Q}_{r,M}$ and $\overline{L}(\overline{f} - x) < 0$. So $\overline{L}(\overline{f}) < x$, and therefore $\inf \overline{Q}_{r,M} < x$. By taking $x = \sup \overline{Q}^*_{r,M} + \varepsilon$ ($\varepsilon > 0$ arbitrary) one obtains $\inf \overline{Q}_{r,M} < \sup \overline{Q}^*_{r,M} + \varepsilon$, which proves the inequality. □

Now we can begin to deal with the result that was mentioned in (1.2). First we need two technical lemmas that are mostly from [Las1]. The first one is needed to prove the second one, which itself is used in Proposition 2.1.9 below.

Lemma 2.1.7. *Let $L : \mathbb{R}[X_1, X_2]_{2r} \longrightarrow \mathbb{R}$ be a linear form and suppose $L(p^2) \geq 0$ for all $p \in \mathbb{R}[X]_r$, i.e. the bilinear form defined by $(p, q) \mapsto L(pq)$ is positive semidefinite. Then all values $L(X^{2\alpha})$ where $0 \leq |\alpha| \leq r$ are bounded by* $\max_{k=0,\ldots,r} \max\{L(X_1^{2k}), L(X_2^{2k})\}$.

Proof. What we will actually show is that all $L(X^{2\alpha})$, where $|\alpha| = k$, are bounded by $\max\{L(X_1^{2k}), L(X_2^{2k})\}$.

Let $p \in \mathbb{N}$ be such that either $k = 2p$ (if k is even) or $k = 2p + 1$ (if k is odd) and define $\Gamma := \{\ (2a, 2b) \mid a + b = k;\ a, b \neq 0\}$. One has $\Gamma = \Gamma_1 \cup \Gamma_2$ where

$$\begin{aligned} \Gamma_1 &:= \{\ (k, 0) + (k - 2i, 2i) \mid i = 1, ..., p\ \} \\ \Gamma_2 &:= \{\ (0, k) + (2j, k - 2j) \mid j = 1, ..., p\}. \end{aligned}$$

If k is odd, then this union is disjoint, else $\Gamma_1 \cap \Gamma_2 = \{(2p, 2p)\}$. For $s := \max\{L(X^\gamma) \mid \gamma \in \Gamma\ \}$, we get $s = L(X^{\gamma^*})$ for some $\gamma^* \in \Gamma_1$ or $\gamma^* \in \Gamma_2$.

If a linear form L defines a positive semidefinite bilinear form in the above shown way, one alway has

$$L(X^{\alpha+\beta})^2 \leq L(X^{2\alpha})L(X^{2\beta}).$$

So in our case, we obtain

$$L\left(X_1^{2k}\right) \cdot L\left(X_1^{2k-4i} X_2^{4i}\right) \ \geq\ L\left(X_1^{2k-2i} X_2^{2i}\right)^2,\ i = 1, ..., p \qquad (2.4)$$

$$L\left(X_2^{2k}\right) \cdot L\left(X_1^{4j} X_2^{2k-4j}\right) \ \geq\ L\left(X_1^{2j} X_2^{2k-2j}\right)^2,\ j = 1, ..., p. \qquad (2.5)$$

With $s_k := \max\{L\left(X_1^{2k}\right), L\left(X_2^{2k}\right)\}$, by (2.4) and (2.5), one gets either

$$s_k \cdot s \geq L\left(X_1^{2k}\right) \cdot L\left(X^{\gamma^*}\right) \geq L\left(X^{\gamma^*}\right)^2 = s^2$$

or

$$s_k \cdot s \geq L\left(X_2^{2k}\right) \cdot L\left(X^{\gamma^*}\right) \geq L\left(X^{\gamma^*}\right)^2 = s^2.$$

In any case $s_k \geq s$. □

Lemma 2.1.8. *Let $L : \mathbb{R}[X_1, ..., X_n]_{2r} \longrightarrow \mathbb{R}$ be a linear form and suppose $L(p^2) \geq 0$ for all $p \in \mathbb{R}[X]_r$, i.e. the bilinear form defined by $(p, q) \mapsto L(pq)$ is positive semidefinite. Assume that for all i=1,...,n and k=0,...,r, the values $L(X_i^{2k})$ are bounded by some τ. Then all values $L(X^\alpha)$, where $|\alpha| \leq 2r$, satisfy $|L(X^\alpha)| \leq \tau$.*

Proof. We only need to show that all values $L\left(X^{2\alpha}\right)$, where $|\alpha| \leq r$, are bounded by τ. Indeed, if a linear form defines a positive semidefinite bilinear form in the above shown way, one always has $L(X^{\alpha+\beta})^2 \leq L(X^{2\alpha})L(X^{2\beta})$, and therefore, if all the values $L\left(X^{2\gamma}\right)$ are bounded by τ, one gets $|L\left(X^{\alpha}\right)| \leq \tau$ for all $0 \leq |\alpha| \leq 2r$.

The proof is by induction on the number n of variables.

$\underline{n=1}$: Nothing is to be shown in this case, as all the values $L\left(X^{2\alpha}\right)$ are bounded by τ by the assumption.

$\underline{n=2}$: This is an immediate result of Lemma 2.1.7.

$\underline{n-1 \rightsquigarrow n, n>2}$: By the induction hypothesis, the claim is true for all $L\left(X^{2\alpha}\right)$, where $|\alpha| \leq r$ and some $\alpha_i = 0$. Indeed, L restricts to a linear form on the ring of polynomials with $n-1$ indeterminates and satisfies all the assumptions needed. So the induction hypothesis gives the boundedness of all those values $L\left(X^{2\alpha}\right)$.

Now take $L\left(X^{2\alpha}\right)$, where $|\alpha| \leq r$ and all $\alpha_i \geq 1$. With no loss of generality, assume $\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_n$. Consider the two elements

$$\gamma := (2\alpha_1, 0, \alpha_3 + \alpha_2 - \alpha_1, \alpha_4, ..., \alpha_n) \in \mathbb{N}^n \text{ and}$$

$$\gamma^{'} := (0, 2\alpha_2, \alpha_3 + \alpha_1 - \alpha_2, \alpha_4, ..., \alpha_n) \in \mathbb{N}^n.$$

We have $|\gamma|, |\gamma^{'}| \leq r$ and $\gamma_2 = \gamma_1^{'} = 0$. Therefore, by the above result, we get

$$L\left(X^{2\gamma}\right) \leq \tau \text{ and } L\left(X^{2\gamma^{'}}\right) \leq \tau.$$

As L defines a positive semidefinite bilinear form, one gets

$$L\left(X^{2\alpha}\right)^2 = L\left(X^{\gamma+\gamma^{'}}\right)^2 \leq L\left(X^{2\gamma}\right) \cdot L\left(X^{2\gamma^{'}}\right) \leq \tau^2,$$

which yields

$$|L\left(X^{2\alpha}\right)| \leq \tau.$$

□

Finally we can prove the result that was mentioned in (1.2).

Proposition 2.1.9. *For every $M > M_0$ the sequence $(\inf Q_{r,M})_{r\geq r_0}$ is monotonously increasing and converges to some $l_M \geq f^*$, where $f^* := \inf_{x\in N} f(x)$.*

Proof. Fix $M > M_0$. First, for every $r > r_0$, the set of feasible solutions for $Q_{r,M}$ is nonempty. One may take $\hat{x} \in N \cap B_{M_0}$ and consider the evaluation at $\hat{x}$.

Now with r getting bigger, the set of feasible solutions for $Q_{r,M}$ is getting smaller. Indeed, for L feasible for $Q_{r+k,M}$, where $k \in \mathbb{N}$, we have

$$0 \leq L(nh(M^2) - \Theta_{r+k}) = L(nh(M^2) - \Theta_r) - L(\Theta_{r+k} - \Theta_r)$$

As $\Theta_{r+k} - \Theta_r$ is a sum of squares, one obtains $L(\Theta_{r+k} - \Theta_r) \geq 0$ and therefore

$$L(nh(M^2) - \Theta_r) \geq 0.$$

So obviously L is also feasible for $Q_{r,M}$, and that is why $\inf Q_{r,M}$ is monotonously increasing with r.

If a linear form L on $\mathbb{R}[X]_{2r}$ is regarded as the vector $(L(X^\alpha))_{|\alpha|\leq 2r} \in \mathbb{R}^{s(2r)}$, where $s(2r) := \dim_{\mathbb{R}} \mathbb{R}[X]_{2r}$, the set of feasible solutions for $Q_{r,M}$ is obviously closed with respect to the euclidean topology. Further, it is bounded, and therefore compact:
From $L(nh(M^2) - \Theta_r) \geq 0$ and $L(1) = 1$, one gets

$$\sum_{i=1}^{n} \sum_{k=0}^{r} a_k L(X_i^{2k}) \leq nh(M^2).$$

From $L(X_i^{2k}) = L((X_i^k)^2) \geq 0$ (for $i = 1, ..., n$ and $k = 0, ..., r$) therefore follows

$$L(X_i^{2k}) \leq \frac{1}{\hat{a}_r} nh(M^2) =: \tau_r \qquad (i = 1, ..., n;\ k = 0, ..., r), \tag{2.6}$$

where $\hat{a}_r := \min_{k=0,\dots,r} a_k$ (remember the definition from page 11; all a_k are positive). By Lemma 2.1.8, one gets this upper bound τ_r for all elements $L(X^\alpha)$ with $|\alpha| \leq 2r$, since $L(q^2) \geq 0$ for all $q \in \mathbb{R}[X]_r$, as L is feasible for $Q_{r,M}$.

This shows the compactness of the set of feasible solutions. Since the objective function is continuous, $\inf Q_{r,M}$ is attained at some feasible point.

As we have already seen, if L is a feasible solution for $Q_{s,M}$ and $s \geq r$, then L is also feasible for $Q_{r,M}$. Therefore one gets even smaller upper bounds for the elements $L(X^\alpha)$: Take L feasible for $Q_{r,M}$, $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq 2r$ and choose the unique natural number k such that $2k-1 < |\alpha| \leq 2k$. We will abbreviate this k by $\breve{\alpha}$. Then, regarding L as a feasible solution for $Q_{k,M}$, one obtains

$$|L(X^\alpha)| \leq \tau_k = \tau_{\breve{\alpha}},$$

where τ_k is defined as in (2.6) above.

Now let $(L^{(r)})_{r\geq r_0}$ be a sequence of optimal solutions for $(Q_{r,M})_{r\geq r_0}$ and fill up all the sequences $L^{(r)} = (L^{(r)}(X^\alpha))_{|\alpha|\leq 2r}$ with zeros to make them elements of $\mathbb{R}^{\mathbb{N}^n}$. From the above shown boundedness, we get

$$L^{(r)} \in \prod_{\alpha\in\mathbb{N}^n} [-\tau_{\breve{\alpha}}, \tau_{\breve{\alpha}}]$$

for all $r \geq r_0$. By Tychonoff's Theorem, this product is compact in the product topology obtained from the euclidean topology on the intervals. Therefore one obtains a subsequence $L^{(r_k)}$ that converges in this topology to some $(y_\alpha)_{\alpha\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$. In particular, we get pointwise convergence, i.e.

$$L^{(r_k)}(X^\alpha) \overset{k\to\infty}{\longrightarrow} y_\alpha \qquad \forall \alpha \in \mathbb{N}^n.$$

$(y_\alpha)_{\alpha\in\mathbb{N}^n}$ defines a linear form L_y on $\mathbb{R}[X]$ by $L_y(X^\alpha) := y_\alpha$. From the pointwise convergence, one obtains $L_y(q^2) \geq 0$ for all $q \in \mathbb{R}[X]$. Also one gets

$$L_y(nh(M^2) - \Theta_r) \geq 0 \quad \forall r \geq r_0.$$

But this, together with $L_y(1) = 1$, which again is given by the pointwise convergence, implies

$$\sum_{i=1}^{n}\sum_{j=0}^{\infty} a_j L_y(X_i^{2j}) \le nh(M^2) =: \varrho.$$

In particular

$$L_y(X_i^{2k}) \le \frac{1}{a_k}\varrho \qquad \forall k \in \mathbb{N}.$$

But this implies

$$\sum_{j=0}^{\infty}(L_y(X_i^{2j}))^{-\frac{1}{2j}} \ge \sum_{j=0}^{\infty} \varrho^{-\frac{1}{2j}} a_j^{\frac{1}{2j}} \ge c\sum_{j=0}^{\infty} a_j^{\frac{1}{2j}} = \infty \quad (i = 1, ..., n)$$

for some $c > 0$ (remember the definition of $(a_j)_j$ on page 11). Together with $L_y(q^2) \ge 0$ for all $q \in \mathbb{R}[X]$, one obtains that *Carleman's Condition* (see Theorem 4.1.5 in the appendix) is satisfied, and therefore there is a positive measure μ on $\mathbb{R}^n$, satisfying

$$L_y(h) = \int_{\mathbb{R}^n} h d\mu \qquad \text{for all } h \in \mathbb{R}[X].$$

As $L^{(r)}(\mathrm{rrad}(g_1, ..., g_s) \cap \mathbb{R}[x]_{2r}) = 0$ for all r, and again by using the pointwise convergence, in particular

$$\int_{\mathbb{R}^n} g_j^2 \, d\mu = L_y(g_j^2) = 0 \tag{2.7}$$

holds for $j = 1, ..., n$. Thus, for every j, there must be $A_j \subseteq \mathbb{R}^n$ with $\mu(\mathbb{R}^n \backslash A_j) = 0$ and $g_j = 0$ on A_j. Hence $\bigcap_{j=1}^n A_j \subseteq N$ and

$$\mu\left(\mathbb{R}^n \backslash \bigcap_{j=1}^{n} A_j\right) = \mu\left(\bigcup_{j=1}^{n} \mathbb{R}^n \backslash A_j\right) = 0.$$

Together with $L_y(1) = 1$, this shows that μ is a probability measure on N. Finally, once more by the pointwise convergence, one obtains

$$\inf Q_{r_k,M} = L^{(r_k)}(f) \stackrel{k\to\infty}{\longrightarrow} \int_{\mathbb{R}^n} f \, d\mu = \int_N f \, d\mu \ge \int_N f^* \, d\mu = f^*\mu(N) = f^*.$$

From the fact that $(\inf Q_{r,M})_{r\in\mathbb{N}}$ is increasing, one gets this convergence for the whole sequence. □

Now we will get the first interesting result about sum of squares representation of nonnegative polynomials, which even holds without our assumption $N \neq \emptyset$.

Theorem 2.1.10. *Let $s \in \mathbb{N}$, $f, g_1, ..., g_s \in \mathbb{R}[X]$ and let N denote the real zero set of the polynomials g_i. Suppose f is nonnegative on N. Then for every $\varepsilon > 0$ there exists $r \in \mathbb{N}$, such that*

$$f + \varepsilon\Theta_r \in \sum \mathbb{R}[X]_r^2 + \text{rrad}(g_1, ..., g_s)_{2r}.$$

(Remember the notations defined on page 11f.)

Proof. If $N = \emptyset$, then the result is obvious, as $\text{rrad}(g_1, ..., g_s)_{2r} = \mathbb{R}[X]_{2r}$. So let $N \neq \emptyset$ and choose M_0 big enough to ensure $N \cap [-M_0, M_0]^n \neq \emptyset$. We have to distinguish two cases.

Case 1: $f^* := \inf_{x\in N} f(x) > 0$.
Let $M_1 > M_0$ be big enough to ensure $f^* - \frac{1}{M_1} > 0$. By Lemma 2.1.6 and Proposition 2.1.9, for every $M > M_1$ there is an r_M such that

$$\sup \overline{Q}^*_{r_M,M} \geq f^* - \frac{1}{M} > 0.$$

But by definition of $\overline{Q}^*_{r_M,M}$, this implies that there are $z_M \in \mathbb{R}, \lambda_M \in \mathbb{R}_{\geq 0}, q_m \in \sum \mathbb{R}[X]^2_{r_M}$ and $\tilde{h} \in \text{rrad}(g_1, ..., g_s)_{2r_M}$ such that

$$f - z_M = q_m + \lambda_M(nh(M^2) - \Theta_{r_M}) + \tilde{h} \tag{2.8}$$

and

$$z_M \geq f^* - \frac{1}{M} > 0. \tag{2.9}$$

As (2.8) is equivalent to

$$f + \lambda_M\Theta_{r_M} = q_M + z_M + \lambda_M nh(M^2) + \tilde{h} \tag{2.10}$$

and $z_M, \lambda_M nh(M^2) \geq 0$, this shows

$$f + \lambda_M \Theta_{r_M} \in \sum \mathbb{R}[X]^2_{r_M} + \mathrm{rrad}(g_1, ..., g_s)_{2r_M}. \tag{2.11}$$

By evaluating (2.8) at some $\hat{x} \in N \cap B_{M_0}$, one obtains

$$\begin{aligned} f(\hat{x}) - z_M &= q_M(\hat{x}) + \lambda_M(nh(M^2) - \Theta_{r_M}(\hat{x})) \\ &\geq \lambda_M n(h(M^2) - h(M_0^2)). \end{aligned}$$

Together with $\frac{1}{M} \geq f^* - z_M$, which is (2.9), one gets

$$f(\hat{x}) - f^* + \frac{1}{M} \geq f(\hat{x}) - f^* + f^* - z_M \geq \lambda_M n(h(M^2) - h(M_0^2)).$$

Hence

$$\lambda_M \leq \frac{f(\hat{x}) - f^* + \frac{1}{M}}{n(h(M^2) - h(M_0^2))} \overset{M\to\infty}{\longrightarrow} 0.$$

So by starting the proof with M big enough, (2.11) yields the result.

Case 2: $f^* = 0$.

Let $\varepsilon > 0$ be given. Then $f + n\frac{\varepsilon}{2} \geq n\frac{\varepsilon}{2} > 0$ on N. So by the first case, there is some $r \in \mathbb{N}$ such that

$$f + n\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\Theta_r \in \sum \mathbb{R}[X]^2_r + \mathrm{rrad}(g_1, ..., g_s)_{2r}.$$

By adding the sum of squares $\frac{\varepsilon}{2}\Theta_r - n\frac{\varepsilon}{2}$ one gets

$$f + \varepsilon\Theta_r \in \sum \mathbb{R}[X]^2_r + \mathrm{rrad}(g_1, ..., g_s)_{2r}.$$

□

The real number r in Theorem 2.1.10 depends a priori on ε and on the polynomial f. Indeed, it *must* depend on ε in general. For example the Motzkin polynomial defined on page 2, which is not a sum of squares, cannot be approximated in the above way by sums of squares in some fixed vector space $\mathbb{R}[X]_r$, as the set of sums of squares is closed according to Proposition (2.1.3).

Nevertheless, there is no real dependency on the explicit choice of the polynomial f, which will be shown in Chapter 3.

Now we will give some Corollaries of Theorem 2.1.10.

Corollary 2.1.11 (Lasserre). *Let $f \in \mathbb{R}[X]$ be nonnegative on $\mathbb{R}^n$. Then for every $\varepsilon > 0$ there is an $r \in \mathbb{N}$ such that*

$$f + \varepsilon\Theta_r \in \sum \mathbb{R}[X]_r^2.$$

Proof. Theorem 2.1.10 with $s = 0$. □

Remark: The representation given in Theorem 2.1.10 provides an evidence for the non-negativity of f on N. As the power series h is convergent on $\mathbb{R}$, we have for $x \in N$:

$$0 \leq f(x) + \varepsilon\Theta_r(x) \leq f(x) + \varepsilon \underbrace{\sum_{i=1}^{n} h(x_i^2)}_{<\infty}.$$

As $\varepsilon > 0$ was arbitrary, one gets

$$f(x) \geq 0 \qquad \forall x \in N.$$

Remark: Corollary 2.1.11 shows in a more constructive way the already known denseness of the convex cone of sums of squares in the convex cone of nonnegative polynomials. With respect to the norm $\|h\|_1 := \sum_{\alpha \in \mathbb{N}^n} |h_\alpha|$ we have

$$\|f - (\underbrace{f + \varepsilon\Theta_r}_{\in \sum \mathbb{R}[X]^2})\|_1 = \varepsilon\|\Theta_r\|_1 = \varepsilon n \sum_{k=0}^{r} a_k \leq \varepsilon n h(1) \xrightarrow{\varepsilon \to 0} 0.$$

So with respect to the $\| \, \|_1$-norm, f can be approximated as closely as desired by sums of squares.

Now one might want to have a result similar to Theorem 2.1.10 for the more general case of a polynomial nonnegative on a basic closed semialgebraic set

$$W(g_1, ..., g_s) := \{ \; x \in \mathbb{R}^n \mid g_i(x) \geq 0; \; i = 1, ..., s \; \}.$$

First we introduce new variables $Z_1, ..., Z_s$ and define the algebraic set

$$\hat{N} := \hat{N}(g_1 - Z_1^2, ..., g_s - Z_s^2) = \{ \ (x, z) \in \mathbb{R}^{n+s} \mid g_i(x) - z_i^2 = 0 \ \} \subseteq \mathbb{R}^{n+s}.$$

If some $f \in \mathbb{R}[X_1, ..., X_n]$ is nonnegative on W, then f, considered as a polynomial in $\mathbb{R}[X_1, ..., X_n, Z_1, ..., Z_s]$, is nonnegative on $\hat{N}$ as well. So for every $\varepsilon > 0$, one gets from Theorem 2.1.10 an $r \in \mathbb{N}$ such that

$$f + \varepsilon\Theta_r + \varepsilon \sum_{i=1}^{s} \sum_{k=0}^{r} a_k Z_i^{2k} \in \sum \mathbb{R}[X, Z]^2 + \operatorname{rrad}(g_1 - Z_1^2, ..., g_s - Z_s^2). \quad (2.12)$$

Now as done in [Las1], we evaluate (2.12) in $(x, \sqrt{g_1(x)}, ..., \sqrt{g_s(x)})$ for $x \in W$. Then one gets

$$f + \varepsilon\Theta_r + \varepsilon \sum_{i=1}^{s} \sum_{k=0}^{r} a_k g_i^k \geq 0 \text{ on W},$$

which implies

$$f(x) + \varepsilon \sum_{i=1}^{n} h(x_i^2) + \varepsilon \sum_{i=1}^{s} h(g_i(x)) \geq 0 \text{ for } x \in W.$$

Again, as $\varepsilon > 0$ was arbitrary, one obtains

$$f(x) \geq 0 \qquad \forall x \in W.$$

So indeed (2.12) provides a certificate for nonnegativity of f on W.

For very special kinds of semialgebraic sets, namely sets of the form $[-l, l]^n$, it is possible to obtain a representation similar to the one in Theorem 2.1.10, providing a certificate of nonnegativity. The proof is quite similar to the one given in this chapter, but uses a different solution to the moment problem. This is what the next section will deal with.

2.2 Polynomials nonnegative on a cube

The requirements in this section are much the same as in the previous section. Again we have $f, g_1, ..., g_s \in \mathbb{R}[X_1, ..., X_n]$. Likewise N is the real zero set of the g_i. We suppose $N \neq \emptyset$. Now we need a real $l \geq 1$, big enough to ensure $N \cap (-l, l)^n \neq \emptyset$. Take $\hat{x} \in N \cap (-l, l)^n$ and choose $C_0 \in \mathbb{R}_{\geq 0}$ such that

$$C_0 > \frac{n}{1 - \frac{\|\hat{x}\|_\infty^2}{l^2}}.$$

$\overline{\mathbb{R}[X]}$ and $\overline{\mathbb{R}[X]}_r$ are the same as above, r_0 is again big enough to ensure $2r_0 \geq \max(\deg f, \deg g_1^2, ..., \deg g_s^2)$. From now on every C is bigger than C_0 and every r bigger than r_0, unless otherwise mentioned.

Last, our Θ_r is modified a little bit:

$$\Theta_r := \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k}.$$

Once more we define three optimization problems, slightly different from the ones in the previous chapter. First of all:

$$Q_{r,C}: \quad \begin{cases} & \text{minimize } \; L(f) \\ s.t. & L : \mathbb{R}[X]_{2r} \longrightarrow \mathbb{R} \text{ linear} \\ & L(1) = 1 \\ & L(C - \Theta_r) \geq 0 \\ & L(\mathrm{rrad}(g_1, ..., g_s) \cap \mathbb{R}[X]_{2r}) = 0 \\ & L(p^2) \geq 0 \quad \forall p \in \mathbb{R}[X]_r \end{cases}$$

Consider again the "modulo rrad$(g_1, ..., g_s)$"-problem:

$$\overline{Q}_{r,C}: \quad \begin{cases} \text{minimize } \overline{L}(\overline{f}) \\ s.t. \quad \overline{L} : \overline{\mathbb{R}[X]}_{2r} \longrightarrow \mathbb{R} \text{ linear} \\ \overline{L}(1) = 1 \\ \overline{L}(C - \overline{\Theta}_r) \geq 0 \\ \overline{L}(\overline{p}^2) \geq 0 \quad \forall \overline{p} \in \overline{\mathbb{R}[X]}_r \end{cases}$$

Last, as expected, the dual problem:

$$\overline{Q}^*_{r,C}: \quad \begin{cases} \text{maximize } z \\ s.t. \quad z \in \mathbb{R} \\ \overline{f} - z \in \sum \overline{\mathbb{R}[X]}_r^2 + \mathbb{R}_{\geq 0}(C - \overline{\Theta}_r) \end{cases}$$

In the same way as in Lemma 2.1.6, we get

$$\sup \overline{Q}^*_{r,C} \geq \inf \overline{Q}_{r,C} \geq \inf Q_{r,C}. \tag{2.13}$$

The inequality on the right hand side is again obvious, the left one is proven as in Lemma 2.1.6. The only thing we have to ensure is

$$C - \overline{\Theta}_r \notin -\sum \overline{\mathbb{R}[X]}_r^2 \tag{2.14}$$

(compare to Lemma 2.1.4 and the following considerations.) But (2.14) is obtained easily by evaluating (well defined) at $\hat{x}$:

$$\overline{\Theta}_r(\hat{x}) \leq \sum_{i=1}^{n} \sum_{k=0}^{\infty} \frac{1}{l^{2k}} \hat{x}_i^{2k} \leq n \sum_{k=0}^{\infty} \frac{\|\hat{x}\|_\infty^{2k}}{l^{2k}} = \frac{n}{1 - \frac{\|\hat{x}\|_\infty^2}{l^2}} < C,$$

and therefore

$$(C - \overline{\Theta}_r)(\hat{x}) > 0,$$

which yields (2.14).

So now we can get a similar result to Proposition 2.1.9:

Proposition 2.2.1. *For every $C > C_0$, the sequence $(\inf Q_{r,C})_{r \geq r_0}$ is monotonously increasing and converges to some $b_C \geq f^*$, where $f^* := \inf_{x \in N \cap [-l,l]^n} f(x)$.*

Proof. The proof is quite similar to the proof of Proposition 2.1.9. First, $Q_{r,C}$ has a feasible solution, namely the evaluation in $\hat{x}$. Second, and again as in the proof of Proposition 2.1.9, $\inf Q_{r,C}$ is growing as r is growing. Next, the set of feasible solutions, considered as a subset of $\mathbb{R}^{s(2r)}$, is obviously closed.

Now from $L(C - \Theta_r) \geq 0$ and $L(p^2) \geq 0$, one obtains

$$0 \leq L(X_i^{2k}) \leq l^{2k} C \qquad \text{for } k \leq r,\ i = 1, ..., n. \tag{2.15}$$

So take L feasible for $Q_{r,C}$ and $\alpha \in \mathbb{N}^n$, $|\alpha| \leq 2r$. If $|\alpha|$ is even, then consider L as a solution for $Q_{\frac{|\alpha|}{2},C}$ and with Lemma 2.1.8 and (2.15) we get

$$|L(X^\alpha)| \leq C l^{|\alpha|}.$$

If α is odd, then consider L as a solution for $Q_{\frac{|\alpha|+1}{2},C}$ and get

$$|L(X^\alpha)| \leq C l^{|\alpha|+1}.$$

With $\hat{C} := Cl$, one get in any case

$$|L(X^\alpha)| \leq \hat{C} l^{|\alpha|}, \tag{2.16}$$

as $l \geq 1$. In particular, (2.16) shows the boundedness and therefore compactness of the set of feasible solutions for $Q_{r,C}$. So to every $Q_{r,C}$, there exists an optimal solution. Now, for every $r \geq r_0$, take an optimal solution $L^{(r)}$

of $Q_{r,C}$ and consider it as an element in $\mathbb{R}^{\mathbb{N}^n}$ as in the proof of Proposition 2.1.9. (2.16) indeed shows, that there exists an element $(y_\alpha)_{\alpha\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ and a subsequence $L^{(r_k)}$ such that

$$L^{(r_k)}(X^\alpha) \stackrel{k\to\infty}{\longrightarrow} y_\alpha \qquad \forall\alpha \in \mathbb{N}^n. \tag{2.17}$$

This is obtained in quite the same way as in the proof of Proposition 2.1.9, by using Tychonoff's Theorem.

But from (2.17), combined with (2.16), we obtain

$$|y_\alpha| \leq \hat{C} l^{|\alpha|} \qquad \forall\alpha \in \mathbb{N}^n. \tag{2.18}$$

As in Proposition 2.1.9, we can consider $(y_\alpha)_{\alpha\in\mathbb{N}^n}$ as a linear form L_y on $\mathbb{R}[X]$ and get $L_y(p^2) \geq 0$ for all $p \in \mathbb{R}[X]$ because of the pointwise convergence. Thus (2.18) and Corollary 4.1.7 from the appendix state that L is integration with respect to a nonnegative measure μ on $[-l,l]^n$.

By the pointwise convergence and $L^{(r_k)}(1) = 1$ for all $k \in \mathbb{N}$, μ must be a probability measure, and from $L_y(\text{rrad}(g_1,...,g_s)) = 0$, which also follows from the pointwise convergence, we get that μ is supported by N, as in Proposition 2.1.9. Due to the pointwise convergence, we finally get

$$\inf Q_{r_k,C} = L^{(r_k)}(f) \stackrel{k\to\infty}{\longrightarrow} L_y(f) = \int_{N\cap[-l,l]^n} f d\mu \geq f^*\mu(N\cap[-l,l]^n) = f^*.$$

As the sequence $\inf Q_{r,C}$ was monotonously increasing, we get this convergence again for the whole sequence. □

Now we can obtain another sum of squares representation from Proposition 2.2.1.

Theorem 2.2.2. *Let $f, g_1, ..., g_s \in \mathbb{R}[X]$ be polynomials, let $N := N(g_1,...,g_s)$ denote the real zero set of the g_i, and let $l \in \mathbb{R}, l \geq 1$, be big enough to ensure $N \cap (-l,l)^n \neq \emptyset$. (If $N = \emptyset$, then the result of this theorem holds for any*

$l \neq 0$ anyway.) Suppose $f \geq 0$ on $N \cap [-l,l]^n$. Then for every $\varepsilon > 0$ there exists an $r \in \mathbb{N}$ such that

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum \mathbb{R}[X]_r^2 + \mathrm{rrad}(g_1, ..., g_s)_{2r}.$$

Proof. Case 1: $f^* := \inf_{x \in N \cap [-l,l]^n} f(x) > 0$. Choose $\zeta > 0$ such that $f^* - \zeta > 0$. By Proposition 2.2.1 and (2.13), for every C big enough, we get $\gamma \in \mathbb{R}$, $r \geq r_0$, $\delta \in \mathbb{R}_{\geq 0}$, $q \in \sum \mathbb{R}[X]_r^2$ and $h \in \mathrm{rrad}(g_1, ..., g_s)_{2r}$ such that

$$\gamma \geq f^* - \zeta > 0 \qquad \text{and} \tag{2.19}$$

$$f - \gamma = q + \delta(C - \Theta_r) + h. \tag{2.20}$$

This implies

$$f + \delta\Theta_r = q + \gamma + \delta C + h = \hat{q} + h,$$

where $\hat{q}$ is another sum of squares.

Evaluating (2.20) at some $\hat{x} \in N \cap (-l,l)^n$ yields

$$f(\hat{x}) - \gamma \geq \delta(C - \Theta_r(\hat{x})), \tag{2.21}$$

where

$$C - \Theta_r(\hat{x}) \geq C - \frac{n}{1 - \frac{\|\hat{x}\|_\infty^2}{l^2}} > 0$$

if C is big enough (note that $\Theta_r(\hat{x}) \leq \frac{n}{1 - \frac{\|\hat{x}\|_\infty^2}{l^2}}$). Thus, with $D := \frac{n}{1 - \frac{\|\hat{x}\|_\infty^2}{l^2}}$, from (2.19) and (2.21) we get

$$\zeta \geq f^* - \gamma \geq -f(\hat{x}) + f^* + \delta(C - D)$$

and therefore

$$\delta \leq \frac{\zeta + f(\hat{x}) - f^*}{C - D} \overset{C \to \infty}{\longrightarrow} 0.$$

Thus by starting with C big enough, we can make $\delta > 0$ as small as desired.

Case 2: Exactly the same as in the proof of Theorem 2.1.10. □

Corollary 2.2.3. *Let $f \in \mathbb{R}[X]$ be nonnegative on $[-l,l]^n$, where $1 \leq l \in \mathbb{R}$. Then for every $\varepsilon > 0$ there exists an $r \in \mathbb{N}$ such that*

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum \mathbb{R}[X]_r^2.$$

Proof. Theorem 2.2.2 with $s = 0$. □

Remark: The representation from Theorem 2.2.2 is a certificate for nonnegativity of f on $N \cap (-l,l)^n$. This is shown exactly as in the previous chapter.

Unfortunately, we cannot get back the nonnegativity of f on elements $x \in N$ that satisfy $\|x\|_\infty = l$, as the power series

$$\sum_{k=0}^{\infty} \frac{x_i^{2k}}{l^{2k}}$$

is not convergent for $x_i = l$. And as there might be points in $N \cap \partial[-l,l]^n$, which are not densely approximated by points in $N \cap (-l,l)^n$, we can not conclude from the nonnegativity of f on $N \cap (-l,l)^n$ to the nonnegativity on $N \cap [-l,l]^n$.

However, in the case of Corollary 2.2.3, we can conclude from the nonnegativity of f on $(-l,l)^n$ to the nonnegativity on $[-l,l]^n$.

Remark: Corollary 2.2.3 gives a solution to the question, whether the convex cone of sums of squares is $\| \|_1$-dense in the convex cone of polynomials nonnegative on $[-l,l]^n$, where $l \geq 1$. For $l > 1$, this follows from the fact that

$$\sum_{k=0}^{\infty} \frac{1}{l^{2k}} = \frac{1}{1-l^2} < \infty.$$

For $l = 1$, consider the polynomial $f_\varepsilon := f + \varepsilon$, for some arbitrary $\varepsilon > 0$. As f is nonnegative on $[-1,1]^n$, f_ε is nonnegative on some $[-1-\delta, 1+\delta]^n$, where

$\delta > 0$ depends on ε. By the above considerations, f_ε can be approximated as closely as desired by sums of squares, and thus f is approximated as well. This mere denseness was first proven by Berg, Christensen and Ressel (see appendix, Theorem 4.1.8, or [BCR2]), as mentioned in the introduction.

At the end of this chapter, we will give two more corollaries of Theorem 2.2.2.

Corollary 2.2.4. *Let $f, g_1, ..., g_s \in \mathbb{R}[X]$ be polynomials, let N denote the real zero set of the g_i. Suppose $f \geq 0$ on N. Then for every real $l > 0$ and every $\varepsilon > 0$ there exists r such that*

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum \mathbb{R}[X]_r^2 + \mathrm{rrad}(g_1, ..., g_s)_{2r}.$$

Corollary 2.2.5. *Let $f \in \mathbb{R}[X]$ be nonnegative. Then for every real $l > 0$ and every $\varepsilon > 0$ there exists r such that*

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum \mathbb{R}[X]_r^2.$$

For $l \geq 1$, these are immediate corollaries of Theorem 2.2.2 and Corollary 2.2.3, respectively. In the case $0 < l < 1$ use the result of Theorem 2.2.2 with $l = 1$ and enlarge the coefficients of the X_i^{2k} afterwards, i.e. add a sum of squares.

These two corollaries can also be deduced from Theorem 2.1.10 in the previous section. In fact, for every $l > 0$, there is some $M \in \mathbb{R}_{\geq 0}$ such that

$$(l^2)^k \leq Mk! \tag{2.22}$$

for all $k \in \mathbb{N}$. So for $\varepsilon > 0$ given, use the result of Theorem 2.1.10 with $a_k := \frac{1}{k!}$ and

$$\hat{\varepsilon} := \frac{\varepsilon}{M}$$

and obtain

$$f + \hat{\varepsilon}\Theta_r = f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{Mk!} X_i^{2k} \in \sum \mathbb{R}[X]_r^2 + \operatorname{rrad}(g_1, ..., g_s)_{2r}$$

for some $r \in \mathbb{N}$, and therefore by (2.22) in particular

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum \mathbb{R}[X]_r^2 + \operatorname{rrad}(g_1, ..., g_s)_{2r}.$$

Chapter 3

Degree Bounds

3.1 Polynomials nonnegative on an algebraic set

In this section and the following one we will have a closer look at the question, which factors really influence the degree up to which we have to perturbe a nonnegative polynomial to make it a sum of squares. In other words, on which factors does r in Theorem 2.1.10 depend?

Let us start with the simplest case, Corollary 2.1.11. The r can obviously only depend on the polynomial f and the chosen ε, because there are no other parameters. But as we are about to show, it is not the explicit choice of f, but only the degree of f and the size of its coefficients, that, together with ε, influence the size of r in Corollary 2.1.11. That means, for a given degree d, a positive real number C and a positive ε, we get an $r = r(d, \varepsilon, C)$ such that the result of Corollary 2.1.11 holds for every nonnegative polynomial f of degree smaller or equal to d and coefficients bounded by C, with that special r (and the chosen ε, of course).

In the more general case of Theorem 2.1.10, there is a dependency on the polynomials that define the algebraic set, in addition.

The main idea is the following: We will first try to generalize Theorem 2.1.10 to a real closed extension field of $\mathbb{R}$. This is done in Proposition 3.1.1. Then we will apply the result in an ultrapower $\mathbb{R}^* := \mathbb{R}^{\mathbb{N}}/\mathcal{F}$, where $\mathcal{F}$ is a non-principal ultrafilter on $\mathbb{N}$. This ultrapower has the interesting $\aleph_1$-saturation property, which is, roughly spoken, some kind of special compactness property for subsets defined by first order logic formulas. This allows us to get the desired result in $\mathbb{R}^*$. Then we can transfer the result back to $\mathbb{R}$ by Tarski's Transfer Principle.

This Transfer Principle says that whenever we have two real closed fields over a common subring and a first order logic formula in the language of ordered rings with coefficients from this subring, and this formula holds in one of the two real closed fields, then it holds in the other one as well. This principle is an immediate corollary of the so called Elimination of Quantifiers Theorem, which says that for every such first order formula with integer coefficients there is a quantifier-free first order formula with integer coefficients, such that these two formulas are equivalent in every real closed field.

These and the following results about real closed fields can be found in [PD], for example.

Given a real closed field R, we denote the convex hull of $\mathbb{Z}$ in R by $\mathcal{O}$, i.e

$$\mathcal{O} = \{x \in R \mid \exists m \in \mathbb{N} : \ |x| \leq m\}.$$

$\mathcal{O}$ is a valuation ring of R with maximal ideal

$$\mathfrak{m} = \{x \in \mathcal{O} \mid \forall n \in \mathbb{N} \setminus \{0\} : \ |x| \leq \frac{1}{n}\}.$$

Let $\overline{R} := \mathcal{O}/\mathfrak{m}$ denote the residue field and $\sigma : \mathcal{O} \to \overline{R}$ the residue map.

From now on, R will always denote a real closed extension field of $\mathbb{R}$. Then we will always have $\overline{R} = \mathbb{R}$ and σ is the identity on $\mathbb{R}$. In fact, for every $\beta \in \mathcal{O}$ there is exactly one $b \in \mathbb{R}$ such that $\beta \equiv b \mod \mathfrak{m}$.

The residue map σ defines a map on $\mathcal{O}[X]$, that we will call σ as well, by applying to the coefficients of a polynomial:

$$\begin{array}{rlcl} \sigma: & \mathcal{O}[X] & \longrightarrow & \mathbb{R}[X] \\ & \sum_\alpha f_\alpha X^\alpha & \longmapsto & \sum_\alpha \sigma(f_\alpha) X^\alpha. \end{array}$$

We will also write $\overline{\beta}$ and $\overline{h}$ instead of $\sigma(\beta)$ and $\sigma(h)$ for elements $\beta \in \mathcal{O}$ and polynomials $h \in \mathcal{O}[X]$ from now on and for the rest of this work (not to be confused with the residue classes modulo the real radical as in the chapter before).

Now $\Theta_r \in \mathbb{R}[X]$ is defined as in the first section of the second chapter again, i.e.

$$\Theta_r := \sum_{i=1}^{n} \sum_{k=0}^{r} a_k X_i^{2k}$$

for some suitable real coefficients a_k (see Section 2.1). Last, for any $g_1, ..., g_s \in R[X]$, we denote by $\mathrm{rrad}_R(g_1, ..., g_s)$ the real radical of the polynomials $g_1, ..., g_s$ in $R[X]$ (for a definition see page 11). It can be characterized as the set of all polynomials from $R[X]$ that vanish on the zero set of the g_i in R^n. By $\mathrm{rrad}_R(g_1, ..., g_s)_r$, we denote the set of all polynomials from $\mathrm{rrad}_R(g_1, ..., g_s)$ of degree r or smaller.

We now want to get a result similar to Lemma 2.1.6 for R instead of $\mathbb{R}$. For real polynomials $f, g_1, ..., g_s$, Lemma 2.1.6 said

$$\sup \overline{Q}^*_{r,M} \geq \inf Q_{r,M}$$

(where $Q_{r,M}$ and $\overline{Q}^*_{r,M}$ are the optimization problems associated to the polynomials f and $g_1, ..., g_s$), whenever $M > M_0$ for some M_0 that satisfies $B_{M_0} \cap N_{\mathbb{R}}(g_1, ..., g_s) \neq \emptyset$, and $2r \geq \max(\deg f, \deg g_1^2, ..., \deg g_s^2)$.

So with fixed $n, s, r \in \mathbb{N}$ and a real $M > 0$, this can be formulated as a first order logic statement over $\mathbb{R}$, in which the polynomials f and $g_1, ..., g_s$ are quantified.

If one is used to this kind of formulas, it is quite easy to see that this is possible. It is very technical and difficult to write down these formulas exactly. Therefore we will only say what should be stated in this formula and give some exact examples for parts of it.

The formula should state the following: For all polynomials f of degree smaller or equal to $2r$, all polynomials $g_1, ..., g_s$ of degree smaller or equal to r (the number n of indeterminates fixed as well) and all $\varepsilon' > 0$, if there is some element in $(-M, M)^n \cap N(g_1, ..., g_s)$, then there is a linear form L that satisfies all the conditions in $Q_{r,M}$, and an element z such that $f - z$ is a sum of squares of polynomials of degree smaller or equal to r plus an element of the form $\delta(nh(M^2) - \Theta_r)$ where $\delta \geq 0$, plus a polynomial of degree smaller or equal to $2r$ from the real radical of the $g_1, ..., g_s$, and $L(f) - \varepsilon' \leq z$.

Now we will look at some parts in detail. First, the quantification *for all polynomials f of degree smaller or equal to* $2r$ (and the same for the polynomials $g_1, ..., g_s$) is in fact a quantification over the coefficients of the polynomials $f, g_1, ..., g_s$, what can be done as the degree is bounded by r (and s, n are fixed).

The condition *there is some element in* $(-M, M)^n \cap N(g_1, ..., g_s)$ is obviously easy to write:

$$\exists x_1, \ldots, x_n \left(\bigwedge_{i=1}^{n} (-M < x_i \wedge x_i < M) \wedge \bigwedge_{j=1}^{s} g_j(x_1, \ldots, x_n) = 0 \right).$$

The existence of a linear form L is in fact the existence of a $s(2r)$-tuple of elements (the values of L on the monomial basis of $\mathbb{R}[X]_{2r}$). Now the feasibility of L for $Q_{r,M}$ needs some work. For example, we have to say that $L(p^2) \geq 0$ for all polynomials p of degree smaller or equal to r. This can indeed be written as a first order logic formula, where the coefficients of the polynomial p are quantified.

How to write the conditions $L(1) = 1$ and $L(nh(M^2) - \Theta_r) \geq 0$ is obvious. Note that $nh(M) - \Theta_r$ is a real polynomial and can therefore be used in our formula.

Now we need the condition, that L maps every element from the real radical of the $g_1, ..., g_s$ of degree smaller or equal to $2r$ to $[0, \infty)$. But this can also be done with a first order logic formula, as the membership of a polynomial $\tilde{h}$ of degree smaller or equal to $2r$ in this ideal can be stated as desired (here we use the geometric characterization of the real radical):

$$\forall x : \left[(\bigwedge_{i=1}^{s} g_i(x) = 0) \longrightarrow (\tilde{h}(x) = 0) \right].$$

So we can also state that for all tuples of coefficients of such polynomials $\tilde{h}$, we have $L(\tilde{h}) \geq 0$.

We can now add an additional condition that was not demanded in $Q_{r,M}$, but followed from the conditions in $Q_{r,M}$, namely the boundedness of the set of feasible solutions. We showed that in particular

$$|L(X^\alpha)| \leq \tau_{r,M} \qquad \text{for all } |\alpha| \leq 2r$$

for some appropriate $\tau_{r,M} \in \mathbb{N}$, whenever L is feasible for $Q_{r,M}$ (see the proof of Proposition 2.1.9). As this boundedness is just a boundedness of the $s(2r)$ values L consists of, we can add this to our formula. Then the formula will still hold in $\mathbb{R}$. We will later see why this is useful.

Next, the existence of an element z such that $L(f) - \varepsilon' \leq z$ is easy to write as a formula. But the fact that $f - z$ lies in the cone defined in $\overline{Q}^*_{r,M}$ is in fact the most complicated condition. First we need that every sum of squares of polynomials from $\mathbb{R}[X]_r$ is already a sum of $s(r)$ such squares (see Lemma 2.1.2). Thus we can write the existence of such a sum of squares in a first order logic definiton by existentially quantifying the coefficients of $s(r)$ polynomials.

Next, the existence of an element $\delta(nh(M^2) - \Theta_r)$, where $\delta \geq 0$, is not difficult to write down.

Further, the existence of an element in the real radical with degree bounded by $2r$ can be formulated as shown above.

Finally the belonging of $f - z$ to this above mentioned cone is now written in finitely many equations, namely the equality of all coefficients of $f - z$ with all the respective coefficients of the element obtained by adding up the sum of squares, the element $\delta(nh(M^2) - \Theta_r)$ and the polynomial from the real radical.

It should have become clear by now how such a formula can be constructed. So all in all, we will call this huge formula γ and we know that γ holds in $\mathbb{R}$ for every chosen real $M > 0$ and every $r \in \mathbb{N} \setminus \{0\}$ (s and n of course also fixed).

By Tarski's Transfer Principle, it holds in any real closed extension field of $\mathbb{R}$ with those parameters as well.

We will use this in the proof of the following generalization of Theorem 2.1.10.

The main idea for the proof of this theorem is as follows: Our polynomial f is now from $\mathcal{O}[X]$, the polynomials $g_1, ..., g_s$ are from $\mathbb{R}[X]$. First, we will consider the optimization problems $Q_{r,M}$ associated to the real polynomials $\overline{f}, g_1, ..., g_s$. The optimal values approximate the infimum of $\overline{f}$ on $\mathbb{R}^n$ by Proposition 2.1.9 and are therefore nonnegative if r is big enough and f was positive on R^n.

Second, by using the strong duality obtained in R by our formula γ with the above chosen r big enough, we get a linear form L on $\mathcal{O}[X]_{2r}$, an $z \in R$ which is only infinitesimal different from $L(f)$, and a representation of $f - z$ as in the chapter before. We will find out that $\overline{L}$, which we obtain by applying the residue map to the values of L, is feasible for the above optimization problem and so $\overline{L}(\overline{f})$ (and thus $L(f)$ and also z) must be positive as well. The above mentioned representation of $f - z$ will yield the result, then.

Now we will give the result and proof in detail.

Proposition 3.1.1. *Let R be a real closed proper extension field of $\mathbb{R}$. Let $f \in \mathcal{O}[X]$ and $g_1, ..., g_s \in \mathbb{R}[X]$ and let $N_R := N_R(g_1, ..., g_s)$ denote the zero set of the polynomials g_i in R^n. Suppose $f \geq 0$ on $\mathcal{O}^n \cap N_R$. Then for every $\varepsilon \in \mathcal{O}, \varepsilon > 0$ and $\varepsilon \notin \mathfrak{m}$, there is an $r \in \mathbb{N}$ such that*

$$f + \varepsilon\Theta_r \in \sum R[X]_r^2 + \mathrm{rrad}_R(g_1, ..., g_s)_{2r}.$$

Proof. First, suppose $\mathcal{O}^n \cap N_R = \emptyset$. Then by Tarski's Transfer Principle, one already gets $N_R = \emptyset$. Indeed, as all the g_i have real coefficients, the existence of a zero in R^n of all the g_i could be stated as a first order logic statement over $\mathbb{R}$ that holds in R. Therefore one would also get a zero in $\mathbb{R}^n \subset \mathcal{O}^n$, a contradiction.

But $N_R = \emptyset$ implies $\mathrm{rrad}_R(g_1, ..., g_s) = R[X]$, so the Proposition is immediately seen to hold in this case.

So let $\mathcal{O}^n \cap N_R \neq \emptyset$.

First we assume $f \geq \frac{1}{m}$ on $\mathcal{O}^n \cap N_R$ for some $m \in \mathbb{N}$. Then we have $\overline{f} \geq \frac{1}{m}$ on $\mathbb{R}^n \cap N_\mathbb{R} = N_\mathbb{R}$. This uses the fact that $\mathbb{R}^n \cap N_\mathbb{R} \subset \mathcal{O}^n \cap N_R$ and that σ preserves order and is the identity on $\mathbb{R}$.

But $\overline{f}$ is a real polynomial. So we can use the result of Proposition 2.1.9, if we consider the optimization problem $Q_{r,M}$ associated to the real polynomials $\overline{f}, g_1, ..., g_s$.

First choose $M_0 \in \mathbb{R}_{>1}$ big enough to ensure

$$\inf_{N_\mathbb{R}} \overline{f} - \frac{1}{M_0} > 0 \text{ and } B_{M_0-1} \cap N_R \neq \emptyset. \tag{3.1}$$

Next, by using Proposition 2.1.9, choose $r_0 \in \mathbb{N}$ big enough for $2r_0 \geq \max(\deg \overline{f}, \deg g_1^2, ..., g_s^2)$ and

$$\inf Q_{r_0,M_0} > \inf_{N_\mathbb{R}} \overline{f} - \frac{1}{M_0} \tag{3.2}$$

to hold. This is possible since the second condition in (3.1) in particular gives the existence of some element in $(-M_0, M_0)^n \cap N_\mathbb{R}$ and thus makes sure M_0 is big enough to fulfill the premise of Proposition 2.1.9.

Now consider the statement γ. We fix the number of variables n, the number s, the real M_0, as well as $r_0 \in \mathbb{N}$ and a suitable $\tau_{r_0,M_0} \in \mathbb{N}$ (remember the additional condition introduced on page 41).

Now we will use the fact that the γ we obtain in this way holds in R.

The second condition in (3.1) ensures that the assumption in γ is fulfilled, so for our polynomials f,$g_1, ..., g_s$ and some arbitrary but fixed infinitesimal $\varepsilon' \in \mathfrak{m}_{>0}$, we get the existence of a linear form $L : R[X]_{2r_0} \longrightarrow R$ and an element $z \in R$, which fulfil

$$L(1) = 1 \tag{3.3}$$

$$L(nh(M_0^2) - \Theta_{r_0}) \geq 0 \tag{3.4}$$

$$L(p^2) \geq 0 \text{ for all } p \in R[X]_{r_0} \tag{3.5}$$

$$L(\text{rrad}_R(g_1, ..., g_s)_{2r_0}) = 0 \tag{3.6}$$

$$\|L\|_\infty \leq \tau_{r_0,M_0} \tag{3.7}$$

$$f - z \in \sum R[X]_{r_0}^2 + R_{\geq 0}(nh(M_0^2) - \Theta_{r_0}) + \text{rrad}_R(g_1, ..., g_s)_{2r_0} \tag{3.8}$$

$$L(f) - \varepsilon' \leq z. \tag{3.9}$$

In (3.7), $\|L\|_\infty$ shall denote the maximum size of the values of L on the monomial basis of $R[X]_{2r_0}$.

Condition (3.7) shows that we can consider L as a linear form on $\mathcal{O}[X]_{2r_0}$:

$$L : \mathcal{O}[X]_{2r_0} \longrightarrow \mathcal{O}.$$

Thus we can define $\overline{L}$ in the following way:

$$\begin{array}{rccc} \overline{L}: & \mathbb{R}[X]_{2r_0} & \longrightarrow & \mathbb{R} \\ & X^\alpha & \mapsto & \overline{L(X^\alpha)}. \end{array}$$

As σ is an order preserving ring homomorphism that is the identity on $\mathbb{R}$, we see from the conditions above that $\overline{L}$ is feasible for the optimization problem Q_{r_0,M_0}. For example, (3.6) implies that $\overline{L}$ is zero on $\text{rrad}_\mathbb{R}(g_1, ..., g_s)_{2r_0}$. This is true, because

$$\mathrm{rrad}_{\mathbb{R}}(g_1, ..., g_s)_{2r_0} \subseteq \mathrm{rrad}_R(g_1, ..., g_s)_{2r_0},$$

which is obvious from the algebraic definition of the real radical given on page 11. The other conditions in Q_{r_0,M_0} follow easily from (3.3), (3.4) and (3.5).

So together with (3.2) we get

$$\overline{L}(\overline{f}) \geq \inf Q_{r_0,M_0} > \inf_{N_{\mathbb{R}}} \overline{f} - \frac{1}{M_0}, \tag{3.10}$$

which is an inequality chain in $\mathbb{R}$. But as

$$L(f) \equiv \overline{L(f)} = \overline{L}(\overline{f}) \qquad \mathrm{mod}\ \mathfrak{m}$$

and

$$z \geq L(f) \qquad \mathrm{mod}\ \mathfrak{m}$$

(by (3.9), as $\varepsilon' \in \mathfrak{m}$), this implies

$$z > \inf_{N_{\mathbb{R}}} \overline{f} - \frac{1}{M_0} > 0, \tag{3.11}$$

where the inequality on the right hand side is (3.1). Next, condition (3.8) gives a $\delta \in R_{\geq 0}$ such that

$$f + \delta\Theta_{r_0} \in \sum R[X]^2_{r_0} + \mathrm{rrad}_R(g_1, ..., g_s)_{2r_0}. \tag{3.12}$$

This uses the fact that z as well as $\delta nh(M_0^2)$ are nonnegative and can therefore be considered as a sum of squares.

We only have to show that δ can be chosen small enough.

By evaluating the representation from (3.8) in some $\tilde{x} \in B_{M_0-1} \cap N_R$ ($\neq \emptyset$), we get

$$f(\tilde{x}) - z \geq \delta(nh(M_0^2) - nh((M_0 - 1)^2)) \tag{3.13}$$

and so together with (3.11)

$$\begin{aligned}\frac{1}{M_0} &\geq \inf_{N_\mathbb{R}} \overline{f} - z \\ &= \inf_{N_\mathbb{R}} \overline{f} - f(\tilde{x}) + f(\tilde{x}) - z \\ &\geq \inf_{N_\mathbb{R}} \overline{f} - f(\tilde{x}) + \delta(nh(M_0^2) - nh((M_0-1)^2)),\end{aligned}$$

which shows

$$\delta \leq \frac{\frac{1}{M_0} - \inf \overline{f} + f(\tilde{x})}{nh(M_0^2) - nh((M_0-1)^2)}. \tag{3.14}$$

With growing $M_0 \in \mathbb{R}_{>0}$, the right hand side of (3.14) is getting smaller than any $q \in \mathbb{Q}_{>0}$. This follows from the fact that all the coefficients a_k in our power series h are positive and by using the binomial theorem.

This completes the proof for the first case.

Now consider the general case $f \geq 0$ on $\mathcal{O}^n \cap N_R$. Fix $\varepsilon \in \mathcal{O}, \varepsilon > 0, \varepsilon \notin \mathfrak{m}$. Then we have $f + n\frac{\varepsilon}{2} \geq \frac{1}{l}$ on $\mathcal{O}^n \cap N_R$ for some $l \in \mathbb{N}$. With the above considerations, we have

$$f + n\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\Theta_r \in \sum R[X]_r^2 + \mathrm{rrad}_R(g_1, ..., g_s)_{2r}$$

for some $r \in \mathbb{N}$. But then, as in the proof of Theorem 2.1.10, also

$$f + \varepsilon\Theta_r \in \sum R[X]_r^2 + \mathrm{rrad}_R(g_1, ..., g_s)_{2r}.$$

□

Now we use a specific real closed extension field of $\mathbb{R}$. Let $\mathcal{F}$ be a non-principal ultrafilter on $\mathbb{N}$ and let

$$\mathbb{R}^* := \mathbb{R}^\mathbb{N}/\mathcal{F}$$

be the ultrapower of $\mathbb{R}$ with respect to $\mathcal{F}$. Then $\mathbb{R}^*$ is real closed, extends $\mathbb{R}$ and has the so called $\aleph_1$-saturation property, which says that every countable covering of a set in $(\mathbb{R}^*)^n$, which is defined by a first order logic formula in the language of ordered rings with coefficients from $\mathbb{R}^*$, with sets that are also defined by such formulas over $\mathbb{R}^*$, has a finite subcover. In particular we get this result for formulas with coefficients from $\mathbb{R}$.

Fix the number n of indeterminates, $g_1, ..., g_s \in \mathbb{R}[X], d \in \mathbb{N}$ and $G \in \mathbb{R}_{\geq 0}$. We can now give a formula ϕ over $\mathbb{R}$ that defines, applied in $\mathbb{R}^*$, the set of all coefficients of polynomials f in $\mathbb{R}^*[X]$ of degree smaller or equal to d, that are nonnegative on the set $N_{\mathbb{R}^*}(g_1, ..., g_s)$, have coefficients bounded by G and are therefore in $\mathcal{O}[X]$.

For this purpose let $h_{d,n}(C, X) \in \mathbb{Z}[C, X]$ be the general polynomial of degree d in n indeterminates $X = (X_1, ..., X_n)$ with coefficients $C = (C_1, ..., C_{s(d)})$. Then the formula

$$\phi : \forall x_1, ..., x_n \left[\left(\bigwedge_{i=1}^{s} g_i(x) = 0 \to h_{d,n}(c, x) \geq 0 \right) \wedge \|c\|_\infty \leq G \right]$$

defines exactly the set of all such tuples c of coefficients. Let us denote the defined set by $\phi(\mathbb{R}^*)$.

Next, for $r \in \mathbb{N}$ and some real $\varepsilon > 0$, the statement

$$h_{d,n}(c, X) + \varepsilon\Theta_r \in \sum \mathbb{R}^*[X]_r^2 + \text{rrad}_{\mathbb{R}^*}(g_1, ..., g_s)_{2r} \tag{3.15}$$

can be written as a first order logic formula over $\mathbb{R}$ and defines, applied in $\mathbb{R}^*$, the set of all $c \in (\mathbb{R}^*)^{s(d)}$, such that the polynomial $h_{d,n}(c, X) + \varepsilon\Theta_r$ lies in the set on the right side of (3.15). (Note that every element in $\sum \mathbb{R}^*[X]_r^2$ is a sum of $s(r)$ squares. This is proven exactly like in Lemma 2.1.2). Let us denote this formula by δ_r and the defined set by $\delta_r(\mathbb{R}^*)$ (note that $\varepsilon \in \mathbb{R}_{>0}$ is a fixed parameter in this formula).

Proposition 3.1.1 implies

$$\phi(\mathbb{R}^*) \subseteq \bigcup_{r \in \mathbb{N}} \delta_r(\mathbb{R}^*). \tag{3.16}$$

(Note that $\varepsilon \in \mathbb{R}$ implies $\varepsilon \notin \mathfrak{m}$.) So using the already mentioned $\aleph_1$-saturation and the obvious fact that

$$\dots\ \delta_r(\mathbb{R}^*) \subseteq \delta_{r+1}(\mathbb{R}^*)\ \dots,$$

we obtain

$$\phi(\mathbb{R}^*) \subseteq \delta_{r_0}(\mathbb{R}^*) \tag{3.17}$$

for some $r_0 \in \mathbb{N}$. This r_0 does of course only depend on the specific formulas we used in (3.16), i.e on $n, s, g_1, ..., g_s, d, G$ and ε.

Now the result from (3.17) can be written as a first order logic statement over $\mathbb{R}$, which holds in $\mathbb{R}^*$ (the following is not really a first order logic statement, but can be written as one, again using the upper bound for the length of a sum of squares in $\sum \mathbb{R}^*[X]_r^2$):

$$\begin{aligned}\forall f \in \mathbb{R}^*[X]_d : \quad & (\|f\|_\infty \leq G \wedge f \geq 0 \text{ on } N_{\mathbb{R}^*}(g_1, ..., g_s)) \\ & \longrightarrow \left(f + \varepsilon\Theta_{r_0} \in \sum \mathbb{R}^*[X]_{r_0}^2 + \mathrm{rrad}_{\mathbb{R}^*}(g_1, ..., g_s)_{2r_0}\right)\end{aligned}$$

($\|f\|_\infty$ denotes the maximum size of the finitely many coefficients of f). By applying Tarski's Transfer Principle, we will obtain the same result in $\mathbb{R}$. So all in all one gets the following theorem, which is a generalization of Theorem 2.1.10:

Theorem 3.1.2. *Let $n, s \in \mathbb{N}$ and $g_1, ..., g_s \in \mathbb{R}[X_1, ..., X_n]$. Let $G \in \mathbb{R}_{\geq 0}$, $d \in \mathbb{N}$ and $\varepsilon \in \mathbb{R}_{>0}$. Then there is an*

$$r = r(n, s, g_1, ..., g_s, G, d, \varepsilon) \in \mathbb{N},$$

such that for every $f \in \mathbb{R}[X]_d$ with $\|f\|_\infty \leq G$ and $f \geq 0$ on $N_{\mathbb{R}}(g_1, ..., g_s)$, one has

$$f + \varepsilon\Theta_r \in \sum \mathbb{R}[X]_r^2 + \mathrm{rrad}_{\mathbb{R}}(g_1, ..., g_s).$$

□

Unfortunately, we have a dependency on the explicit polynomials $g_1, ..., g_s$ in Theorem 3.1.2. This emerged as we used the fact that the g_i are real polynomials in Proposition 3.1.1.

By introducing some more constraints, this can be avoided. The proof is quite similar to the one above, but some details have to be changed.

We will again need our first order logic statement γ, introduced above.

Proposition 3.1.3. *Let R be a real closed proper extension field of $\mathbb{R}$. Let $f, g_1, ..., g_s \in \mathcal{O}[X]$. Suppose $N_R(g_1, ..., g_s) \cap \mathcal{O}^n \neq \emptyset$. Further suppose there is an $N \in \mathbb{N}$ such that for all polynomials $g'_1, ..., g'_s \in R[X]$ fulfilling $\deg g'_i \leq \deg g_i$ and $\|g_i - g'_i\|_\infty \leq \frac{1}{N}$ for $i = 1, ..., s$, one has*

$$f \geq 0 \text{ on } N_R(g'_1, ..., g'_s) \cap \mathcal{O}^n.$$

Then for every $\varepsilon \in \mathcal{O}, \varepsilon > 0$ and $\varepsilon \notin \mathfrak{m}$, there is an $r \in \mathbb{N}$ such that

$$f + \varepsilon \Theta_r \in \sum R[X]_r^2 + \mathrm{rrad}_R(g_1, ..., g_s)_{2r}.$$

Proof. As $\deg \overline{g}_i \leq \deg g_i$ and $\|g_i - \overline{g}_i\|_\infty \in \mathfrak{m}$ for all $i = 1, ..., s$, one gets $f \geq 0$ on $N_R(\overline{g}_1, ..., \overline{g}_s) \cap \mathcal{O}^n$ from the assumption (this is where we need the additional assumption on the polynomials $g_1, ..., g_s$ if they are supposed to be from $\mathcal{O}[X]$).

With ε as demanded, one has

$$f + n\frac{\varepsilon}{2} \geq \frac{1}{m} \qquad \text{on} \quad N_R(\overline{g}_1, ..., \overline{g}_s) \cap \mathcal{O}^n$$

for some $m \in \mathbb{N}$ (as $\varepsilon \notin \mathfrak{m}$). Thus

$$\overline{f + n\frac{\varepsilon}{2}} \geq \frac{1}{m} \qquad \text{on} \quad N_{\mathbb{R}}(\overline{g}_1, ..., \overline{g}_s).$$

Now we consider an optimization problem $P_{r,M}$ associated to the real polynomials $\overline{f + n\frac{\varepsilon}{2}}, \overline{g}_1, ..., \overline{g}_s$. $P_{r,M}$ is defined quite similar to $Q_{r,M}$, we only replace the condition $L(\mathrm{rrad}_{\mathbb{R}}(\overline{g}_1, ..., \overline{g}_s)_{2r}) = 0$ by the weaker one

$$L(\overline{g}_j^2) = 0 \text{ for } j = 1, ..., s.$$

As we have seen in the proof of Proposition 2.1.9, we only need this condition to make the result of Proposition 2.1.9 hold for $P_{r,M}$ as well (see (2.7) on page 24). We will see later why it is useful to change from $Q_{r,M}$ to $P_{r,M}$.

Now we can choose $M_0, r_0 \in \mathbb{N}$ big enough to ensure

$$\underbrace{\inf_{N_{\mathbb{R}}(\overline{g}_1,\dots,\overline{g}_s)} \overline{f + n\frac{\varepsilon}{2}}}_{=:(\overline{f+n\frac{\varepsilon}{2}})^*} - \frac{1}{M_0} > 0 \text{ and } B_{M_0-1} \cap N_R(g_1, ..., g_s) \neq \emptyset$$

as well as

$$\inf P_{r_0,M_0} > \left(\overline{f + n\frac{\varepsilon}{2}}\right)^* - \frac{1}{M_0}.$$

By using γ with this r_0, M_0, the polynomials $f + n\frac{\varepsilon}{2}, g_1, ..., g_s$ and again an infinitesimal $\varepsilon' \in \mathfrak{m}_{>0}$, we get a linear form L on $\mathcal{O}[X]_{2r_0}$ and an element $z \in R$ which fulfil all the conditions stated in δ. This ensures $\overline{L}$, defined as in the proof of Proposition 3.1.1, to be feasible for P_{r_0,M_0}. Indeed, we can conclude $\overline{L}(\overline{g}_i^2) = 0$ from $L(\mathrm{rrad}_R(g_1, ..., g_s)_{2r_0}) = 0$ (this is where we need the modified optimization problem $P_{r,M}$; it seems unlikely that we can get $\overline{L}(\mathrm{rrad}_{\mathbb{R}}(\overline{g}_1, ..., \overline{g}_s)_{2r_0}) = 0$ from $L(\mathrm{rrad}_R(g_1, ..., g_s)_{2r_0}) = 0$). The other conditions in P_{r_0,M_0}, which are the same as in Q_{r_0,M_0}, are fulfilled as we showed in the proof of Proposition 3.1.1.

But now we can show just as in the proof of Proposition 3.1.1 that the obtained z is positive and get

$$f + n\frac{\varepsilon}{2} + \delta\Theta_{r_0} \in \sum R[X]^2_{r_0} + \mathrm{rrad}_R(g_1, ..., g_s)_{2r_0},$$

for some $\delta \in R_{\geq 0}$ and $r \in \mathbb{N}$. Finally, and exactly like above, δ can be chosen smaller than $\frac{\varepsilon}{2}$ (resulting in growing r_0), by starting with some M_0 big enough. This shows the claim. □

This result is now used in the case $R = \mathbb{R}^*$ again. Fix $n, s, d, N \in \mathbb{N}$ and $M, G \in \mathbb{R}_{\geq 0}$. Let ϕ be a first order logic formula over $\mathbb{R}$ that defines the set of all coefficients of polynomials $f, g_1, ..., g_s$ in n indeterminates and of degree smaller or equal to d, that fulfil $\|f\|_\infty, \|g_1\|_\infty, ..., \|g_s\|_\infty \leq G$, $B_M \cap N(g_1, ..., g_s) \neq \emptyset$ and the property that for all $g_1', ..., g_s'$ with $\deg g_i' \leq \deg g_i$ and $\|g_i - g_i'\|_\infty \leq \frac{1}{N}$ for $i = 1, ..., s$ one has $f \geq 0$ on $N(g_1', ..., g_s')$. This can obviously be done. For example, the condition $B_M \cap N(g_1, ..., g_s) \neq \emptyset$ is written as

$$\exists x_1, ..., x_n (\bigwedge_{i=1}^{n} (-M \leq x_i \wedge x_i \leq M) \wedge \bigwedge_{j=1}^{s} g_j(x_1, ..., x_n) = 0).$$

The condition that for all $g_1', ..., g_s'$ with $\deg g_i' \leq \deg g_i$ and $\|g_i - g_i'\|_\infty \leq \frac{1}{N}$ for $i = 1, ..., s$ one has $f \geq 0$ on $N(g_1', ..., g_s')$ can also be written in a first order logic form, using that the upper boundedness of the $\|.\|_\infty$-norm of a polynomial (of degree smaller or equal to d) can be written as finitely many inequalities:

$$\forall c_1', ..., c_s' \forall x :$$

$$\bigwedge_{i=1}^{s} \left(\|h_{d,n}(c_i', X) - g_i\|_\infty \leq \tfrac{1}{N} \wedge \deg h_{d,n}(c_i', X) \leq \deg g_i\right)$$

$$\rightarrow [(\bigwedge_{i=1}^{s} h_{d,n}(c_i', x) = 0) \rightarrow f(x) \geq 0] .$$

Here, the c_i' are meant to be the coefficients of the polynomials g_i', i.e every c_i' consists in fact of $s(d)$ elements, whereas x is an n-tupel. $h_{d,n}(C, X)$ denotes the general polynomial in n indeterminates of degree d again. A condition of the form $\deg h' \leq \deg h$, where h and h' are polynomials of degree smaller or equal to some d can indeed be written as finitely many semialgebraic conditions, for example: If all the coefficients of the polynomial h, that belong to monomials of highest possible degree (here: d) are zero, then all those coefficients of the polynomial h' have to be zero as well. If in addition all the coefficients of h, belonging to monomials of degree only one smaller than d are zero, then the same coefficients of h' must also be zero, and so on.

So all in all, any tuple of polynomials $f, g_1, ..., g_s \in \mathbb{R}^*[X]$, whose coefficients make ϕ hold in $\mathbb{R}^*$, fulfil the assumptions of Proposition 3.1.3. Indeed,

$$\|f\|_\infty, \|g_1\|_\infty, ..., \|g_s\|_\infty \leq G$$

ensures the polynomials to have coefficients in $\mathcal{O}$, from $B_M \cap N_{\mathbb{R}^*}(g_1, ..., g_s) \neq \emptyset$ we get

$$N_{\mathbb{R}^*}(g_1, ..., g_s) \cap \mathcal{O}^n \neq \emptyset$$

and finally we get that for all $g_1', ..., g_s' \in \mathbb{R}^*[X]$ with $\deg g_i' \leq \deg g_i$ and $\|g_i - g_i'\|_\infty \leq \frac{1}{N}$ for $i = 1, ..., s$ one has

$$f \geq 0 \qquad \text{on} \quad N_{\mathbb{R}^*}(g_1', ..., g_s').$$

The set of all coefficients of polynomials $f, g_1, ..., g_s$ of degree smaller or equal to some d, that fulfil the resulting statement of Proposition 3.1.3, can now be defined by suitable formulas again. Fix a real $\varepsilon > 0$ and let δ_r be a formula over $\mathbb{R}$ that defines the set of all coefficients of polynomials $f, g_1, ..., g_s$ of degree smaller or equal to d, which fulfill that $f + \varepsilon\Theta_r$ is a sum of squares of length $s(r)$ of polynomials of degree smaller or equal to r, plus a polynomial of degree smaller or equal to $2r$ from the real radical of the g_i. This can be done as shown above.

With Proposition 3.1.3, we get

$$\phi(\mathbb{R}^*) \subseteq \bigcup_{r \in \mathbb{N}} \delta_r(\mathbb{R}^*),$$

which implies

$$\phi(\mathbb{R}^*) \subseteq \delta_{r_0}(\mathbb{R}^*) \tag{3.18}$$

for some $r_0 \in \mathbb{N}$, again using the $\aleph_1$-saturation. This r_0 depends on the used formulas ϕ and δ_r, i.e. on n, d, s, G, M, N and ε.

But the statement of (3.18) can be written as a formula over $\mathbb{R}$ which holds in $\mathbb{R}^*$. Indeed, (3.18) says that for all polynomials $f, g_1, ..., g_s$ in n indeterminates and of degree smaller or equal to d, whose coefficients are bounded by G, which fulfil that $B_M \cap N_{\mathbb{R}^*}(g_1, ..., g_s) \neq \emptyset$ and the property that for all $g_1', ..., g_s'$ with $\deg g_i' \leq \deg g_i$ and $\|g_i - g_i'\|_\infty \leq \frac{1}{N}$ for $i = 1, ..., s$ one has $f \geq 0$ on $N(g_1', ..., g_s')$, we get $f + \varepsilon\Theta_{r_0} \in \sum \mathbb{R}^*[X]_{r_0}^2 + \mathrm{rrad}_{\mathbb{R}^*}(g_1, ..., g_s)_{2r_0}$. By Tarski's Transfer Principle we get the same result in $\mathbb{R}$:

Theorem 3.1.4. *Let $n, s, d, N \in \mathbb{N}$, $\varepsilon \in \mathbb{R}_{>0}$ and $G, M \in \mathbb{R}_{\geq 0}$. Then there is an*

$$r = r(n, s, d, N, G, M, \varepsilon) \in \mathbb{N}$$

such that for all polynomials $f, g_1, ..., g_s \in \mathbb{R}[X_1, ..., X_n]_d$, which have the properties

- *(i)* $B_M \cap N_{\mathbb{R}}(g_1, ..., g_s) \neq \emptyset$
- *(ii)* $\|f\|_\infty, \|g_1\|_\infty, ..., \|g_s\|_\infty \leq G$
- *(iii) for all $g_1', ..., g_s' \in \mathbb{R}[X]$ with $\deg g_i' \leq \deg g_i$ and $\|g_i - g_i'\|_\infty \leq \frac{1}{N}$ for $i = 1, ..., s$, one has $f \geq 0$ on $N_{\mathbb{R}}(g_1', ..., g_s')$,*

one gets

$$f + \varepsilon\Theta_r \in \sum \mathbb{R}[X]_r^2 + \mathrm{rrad}_{\mathbb{R}}(g_1, ..., g_s)_{2r}.$$

□

In the next section, we will show similar results for the representation we obtained in the second section of the second chapter.

3.2 Polynomials nonnegative on a cube

For $l \in \mathbb{R}$, $l \geq 1$, we define Θ_r as in the second section of the second chapter, i.e.

$$\Theta_r := \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k}.$$

We obtained a result about the strong duality of the optimization problems defined in that section, which was stated in (2.13), namely

$$\sup \overline{Q}^*_{r,C} \geq \inf Q_{r,C}.$$

That result was proven for all such problems associated to polynomials $f, g_1, ..., g_s$, if only $2r \geq \deg f, \deg g_1^2, ..., g_s^2$, and as long as there is an $\tilde{x} \in N_{\mathbb{R}}(g_1, ..., g_s) \cap (-l, l)^n$ and the occuring C is big enough to ensure

$$C > \frac{n}{1 - \frac{\|\tilde{x}\|_\infty^2}{l^2}}.$$

We want to write this result as a first order logic statement over $\mathbb{R}$ that holds in $\mathbb{R}$. The same considerations as in the previous section show that this is possible. Therefore we will only explain what this statement should say. Fix a number n of indeterminates, a number s, a degree $r \in \mathbb{N} \setminus \{0\}$ and some $C \in \mathbb{R}_{\geq 0}$ as well as $l \in \mathbb{R}_{\geq 1}$.

We now state the following: For all polynomials f of degree smaller or equal to $2r$, all polynomials $g_1, ..., g_s$ of degree smaller or equal to r and every $\varepsilon' > 0$, if there exists some element $\tilde{x}$ in $(-l, l)^n \cap N(g_1, ..., g_s)$ that fulfils

$$\frac{n}{1 - \frac{\|\tilde{x}\|_\infty^2}{l^2}} < C,$$

then there is a linear form L feasible for $Q_{r,C}$, in addition the values of L on the monomial basis are bounded by some $\tau_{r,C} \in \mathbb{N}$, and there is an element z, such that $f - z$ is a sum of squares of polynomials of degree smaller or equal to r, plus a polynomial of the form $\delta(C - \Theta_r)$, where $\delta \geq 0$, plus a polynomial of degree smaller or equal to $2r$ from the real radical of the g_i, and $L(f) - \varepsilon' \leq z$.

So with all those parameters fixed, write this huge statement as a first order logic statement γ and note that γ holds in every real closed extension field of $\mathbb{R}$ by Tarski's Transfer Principle.

Similar as in the previous chapter, we will proof a proposition which gives a generalization of Theorem 2.2.2:

Proposition 3.2.1. *Let R be a real closed proper extension field of $\mathbb{R}$. Let $f \in \mathcal{O}[X]$, $g_1, ..., g_s \in \mathbb{R}[X]$ and let $N_R(g_1, ..., g_s)$ denote the zero set of the polynomials g_i in R^n. Let $l \in \mathbb{R}$, $l \geq 1$ and $\beta \in \mathbb{R}_{>0}$ be such that*

$$[-(l-\beta), l-\beta]^n \cap N_R(g_1, ..., g_s) \neq \emptyset.$$

Suppose $f \geq 0$ on $[-l, l]^n \cap N_R(g_1, ..., g_s)$. Then for every $\varepsilon \in \mathcal{O}$, $\varepsilon > 0$ and $\varepsilon \notin \mathfrak{m}$, there is an $r \in \mathbb{N}$ such that

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum R[X]_r^2 + \mathrm{rrad}_R(g_1, ..., g_s)_{2r}.$$

Proof. Let $\varepsilon \in \mathcal{O}$, $\varepsilon > 0$ and $\varepsilon \notin \mathfrak{m}$. Then we have

$$f + n\frac{\varepsilon}{2} \geq \frac{1}{m} \text{ on } N_R(g_1, ..., g_s) \cap [-l, l]^n$$

for some $m \in \mathbb{N}$, as $\varepsilon \notin \mathfrak{m}$. Thus

$$\overline{f + n\frac{\varepsilon}{2}} \geq \frac{1}{m} \text{ on } N_{\mathbb{R}}(g_1, ..., g_s) \cap [-l, l]^n,$$

as in the proof of Proposition 3.1.1.

Now choose $C_0 \in \mathbb{N}$ big enough to ensure

$$\underbrace{\inf_{N_{\mathbb{R}} \cap [-l,l]^n} \overline{f + n\frac{\varepsilon}{2}}}_{=:(\overline{f + n\frac{\varepsilon}{2}})^*} - \frac{1}{C_0} > 0 \tag{3.19}$$

as well as

$$C_0 > \frac{n}{1 - \frac{\|\tilde{x}\|_\infty^2}{l^2}} \tag{3.20}$$

for some $\tilde{x} \in [-(l-\beta), l-\beta]^n \cap N_R(g_1, ..., g_s)$. Note that $\frac{n}{1-\frac{\|\tilde{x}\|_\infty^2}{l^2}} \in \mathcal{O}$, as $\tilde{x} \in [-(l-\beta), (l-\beta)]^n$ and $\beta \in \mathbb{R}_{>0}$. So C_0 can be chosen from $\mathbb{N}$. This is the reason why we need β in this proposition.

Now by using the result of Proposition 2.2.1 for the optimization problem $Q_{r,C}$ associated to the real polynomials $\overline{f + n\frac{\varepsilon}{2}}, g_1, ..., g_s$, we can choose $r_0 \in \mathbb{N}$ big enough to ensure

$$\inf Q_{r_0, C_0} > \left(\overline{f + n\frac{\varepsilon}{2}}\right)^* - \frac{1}{C_0}. \tag{3.21}$$

Now consider the statement γ with l, C_0 and r_0. As γ holds in R, we use it for the polynomials $f + n\frac{\varepsilon}{2}, g_1, ..., g_s$ and some infinitesimal $\varepsilon' \in \mathfrak{m}_{>0}$. As the assumption in γ holds by (3.20), we get a linear form $L : R[X]_{2r_0} \to R$ and an element $z \in R$ such that

$$L(1) = 1 \tag{3.22}$$

$$L(C_0 - \Theta_{r_0}) \geq 0 \tag{3.23}$$

$$L(p^2) \geq 0 \text{ for all } p \in R[X]_{r_0} \tag{3.24}$$

$$L(\mathrm{rrad}_R(g_1, ..., g_s)_{2r_0}) = 0 \tag{3.25}$$

$$\|L\|_\infty \leq \tau_{r_0, C_0} \tag{3.26}$$

$$f + n\frac{\varepsilon}{2} - z \in \sum R[X]_{r_0}^2 + R_{\geq 0}(C_0 - \Theta_{r_0}) + \mathrm{rrad}_R(g_1, ..., g_s) \tag{3.27}$$

$$L(f + n\frac{\varepsilon}{2}) - \varepsilon' \leq z. \tag{3.28}$$

The condition $\|L\|_\infty \leq \tau_{r_0, C_0} \in \mathbb{N}$ again ensures L to be a linear form on $\mathcal{O}$, so we can define $\overline{L}(X^\alpha) := \overline{L(X^\alpha)}$ and find $\overline{L}$ feasible for Q_{r_0, C_0}, by the above conditions. For example, we get $\overline{L}(\mathrm{rrad}_{\mathbb{R}}(g_1, ..., g_s)_{2r_0}) = 0$ from (3.25), just as in the proof of Proposition 3.1.1.

So we have

$$\underbrace{\overline{L}(\overline{f+n\frac{\varepsilon}{2}})}_{\in\mathbb{R}} > \underbrace{\left(\overline{f+n\frac{\varepsilon}{2}}\right)^* - \frac{1}{C_0}}_{\in\mathbb{R}}$$

by (3.21). And as

$$L(f+n\frac{\varepsilon}{2}) \equiv \overline{L}(\overline{f+n\frac{\varepsilon}{2}}) \mod \mathfrak{m},$$

as well as

$$z \geq L(f+n\frac{\varepsilon}{2}) \mod \mathfrak{m}$$

(by (3.28), since $\varepsilon' \in \mathfrak{m}$), this implies

$$z > \left(\overline{f+n\frac{\varepsilon}{2}}\right)^* - \frac{1}{C_0} > 0, \tag{3.29}$$

(where the right part of the equation is from (3.19)). Now we can go on as usual. Using the fact that z is a square in R, from (3.27) we get the following representation:

$$f+n\frac{\varepsilon}{2}+\delta\Theta_{r_0} \in \sum R[X]^2_{r_0} + \mathrm{rrad}_R(g_1,...,g_s)_{2r_0} \tag{3.30}$$

for some $\delta \in R_{\geq 0}$. Evaluating the representation from (3.27) in some $\tilde{x} \in [-(l-\beta), l-\beta]^n \cap N_R(g_1,...,g_s)$ yields

$$f(\tilde{x})+n\frac{\varepsilon}{2}-z \geq \delta(C_0-\Theta_{r_0}(\tilde{x})). \tag{3.31}$$

Together with (3.29), the same computation as in the proofs before shows

$$\delta \leq \frac{\frac{1}{C_0}+f(\tilde{x})+n\frac{\varepsilon}{2}-\left(\overline{f+n\frac{\varepsilon}{2}}\right)^*}{C_0-\frac{n}{1-\frac{\|\tilde{x}\|^2_\infty}{l^2}}},$$

where the right hand side becomes smaller than any $q \in \mathbb{Q}_{>0}$ for $C_0 \in \mathbb{N}$ big enough. In particular, δ can be chosen smaller than $\frac{\varepsilon}{2}$, which completes the proof. □

We use this result in $R = \mathbb{R}^*$ again.

Fix a number n of indeterminates and some $g_1, ..., g_s \in \mathbb{R}[X]$, as well as some real $l \geq 1$ and $\beta \in \mathbb{R}_{>0}$. Finally choose a natural number d (the upper bound for the degree) and a $G \in \mathbb{R}_{\geq 0}$ (the upper bound for the coefficients).

Let ϕ be a first order logic formula over $\mathbb{R}$ that defines the set of all polynomials f of degree smaller or equal to d (or better: that defines the set of all tuples of coefficients of such polynomials), which fulfil $\|f\|_\infty \leq G$ as well as $f \geq 0$ on $[-l, l]^n \cap N(g_1, ..., g_s)$. In addition, we need the constraint $[-(l - \beta), l - \beta]^n \cap N(g_1, ..., g_s) \neq \emptyset$ in ϕ.

As shown in the previous chapters, this can be done. Next, with fixed $\varepsilon \in \mathbb{R}_{>0}$, let δ_r be a formula over $\mathbb{R}$ that defines the set of all polynomials f of degree smaller or equal to d, such that $f + \varepsilon\Theta_r$ is a sum of $s(r)$ squares of polynomials of degree smaller or equal to r plus a polynomial of degree smaller or equal to $2r$ from the real radical of the g_i. This can be done as well.

Now from Proposition 3.2.1, used in the case $R = \mathbb{R}^*$, we get

$$\phi(\mathbb{R}^*) \subseteq \bigcup_{r \in \mathbb{N}} \delta_r(\mathbb{R}^*)$$

and with the $\aleph_1$-saturation indeed

$$\phi(\mathbb{R}^*) \subseteq \delta_{r_0}(\mathbb{R}^*) \tag{3.32}$$

for some $r_0 \in \mathbb{N}$. This r_0 depends on the parameters that we used in ϕ and δ_r, i.e. on $n, s, g_1, ..., g_s, d, G, l, \beta$ and ε. Now the result from (3.32) can be written as a first order logic statement over $\mathbb{R}$ that holds in $\mathbb{R}^*$, and by Tarski's Transfer Principle we get the same result in $\mathbb{R}$. This is stated in the following theorem, which generalizes Theorem 2.2.2:

Theorem 3.2.2. *Let $n, s \in \mathbb{N}$ and $g_1, ..., g_s \in \mathbb{R}[X_1, ..., X_n]$. Let $G \in \mathbb{R}_{\geq 0}$, $d \in \mathbb{N}$ and $\varepsilon \in \mathbb{R}_{>0}$. Further let $l \in \mathbb{R}_{\geq 1}$ and $\beta \in \mathbb{R}_{>0}$ be such that*

$$[-(l - \beta), l - \beta]^n \cap N_{\mathbb{R}}(g_1, ..., g_s) \neq \emptyset.$$

Then there is an

$$r = r(n, s, g_1, ..., g_s, G, d, l, \beta, \varepsilon) \in \mathbb{N}$$

such that for all polynomials $f \in \mathbb{R}[X]_d$ *which satisfy* $\|f\|_\infty \leq G$ *as well as* $f \geq 0$ *on* $[-l, l]^n \cap N_{\mathbb{R}}(g_1, ..., g_s)$*, one has*

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum \mathbb{R}[X]_r^2 + \mathrm{rrad}_{\mathbb{R}}(g_1, ..., g_s)_{2r}.$$

□

Unfortunately, we have a dependency on the polynomials $g_1, ..., g_s$ in this result, just like in Theorem 3.1.2. Again it is possible to avoid this by introducing some more constraints:

Proposition 3.2.3. *Let* R *be a real closed proper extension field of* $\mathbb{R}$*. Let* f*,* $g_1, ..., g_s \in \mathcal{O}[X]$*,* $l \in \mathbb{R}_{\geq 1}$ *and* $\beta \in \mathbb{R}_{>0}$*. Suppose*

$$[-(l-\beta), l-\beta]^n \cap N_R(g_1, ..., g_s) \neq \emptyset.$$

Further suppose that there is an $N \in \mathbb{N}$ *such that for all polynomials* $g'_1, ..., g'_s \in R[X]$ *fulfilling* $\deg g'_i \leq \deg g_i$ *and* $\|g_i - g'_i\|_\infty \leq \frac{1}{N}$ *for* $i = 1, ..., s$*, one has*

$$f \geq 0 \text{ on } [-l, l]^n \cap N_R(g'_1, ..., g'_s).$$

Then for every $\varepsilon \in \mathcal{O}$*,* $\varepsilon > 0$ *and* $\varepsilon \notin \mathfrak{m}$*, there is an* $r \in \mathbb{N}$ *such that*

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum R[X]_r^2 + \mathrm{rrad}_R(g_1, ..., g_s)_{2r}.$$

Proof. As $\deg \overline{g}_i \leq \deg g_i$ and $\|g_i - \overline{g}_i\|_\infty \in \mathfrak{m}$ for all $i = 1, ..., s$, one gets

$$f \geq 0 \qquad \text{on} \quad [-l, l]^n \cap N_R(\overline{g}_1, ..., \overline{g}_s)$$

from the assumption.

With ε as demanded, one has

$$f + n\frac{\varepsilon}{2} \geq \frac{1}{m} \qquad \text{on} \quad [-l,l]^n \cap N_R(\overline{g}_1,...,\overline{g}_s)$$

for some $m \in \mathbb{N}$ (as $\varepsilon \notin \mathfrak{m}$). Thus

$$\overline{f + n\frac{\varepsilon}{2}} \geq \frac{1}{m} \qquad \text{on} \quad [-l,l]^n \cap N_{\mathbb{R}}(\overline{g}_1,...,\overline{g}_s).$$

We again consider the optimization problem $P_{r,C}$ associated to the real polynomials $\overline{f + n\frac{\varepsilon}{2}}, \overline{g}_1,...,\overline{g}_s$, which is obtained by replacing the condition $L(\text{rrad}_{\mathbb{R}}(\overline{g}_1,...,\overline{g}_s)_{2r}) = 0$ in $Q_{r,C}$ by the weaker one

$$L(\overline{g}_j^2) = 0 \text{ for } j = 1,...,s.$$

Again this condition suffices to make Proposition 2.2.1 hold with $P_{r,C}$ instead of $Q_{r,C}$ (see (2.7) on page 24).

Choose $C_0 \in \mathbb{N}$ big enough to ensure

$$\underbrace{\inf_{[-l,l]^n \cap N_{\mathbb{R}}(\overline{g}_1,...,\overline{g}_s)} \overline{f + n\frac{\varepsilon}{2}}}_{:=(\overline{f+n\frac{\varepsilon}{2}})^*} - \frac{1}{C_0} > 0$$

and

$$\frac{n}{1 - \frac{\|\tilde{x}\|_\infty^2}{l^2}} < C_0$$

for some $\tilde{x} \in [-(l-\beta), l-\beta]^n \cap N_R(g_1,...,g_s)$. Then choose $r_0 \in \mathbb{N}$ big enough for

$$\inf P_{r_0,C_0} > \left(\overline{f + n\frac{\varepsilon}{2}}\right)^* - \frac{1}{C_0}$$

to hold.

By using our statement γ with this r_0, C_0, some infinitesimal $\varepsilon' \in \mathfrak{m}_{>0}$ and the polynomials $f + n\frac{\varepsilon}{2}, g_1,...,g_s$, we get a linear form L on $\mathcal{O}[X]_{2r_0}$ and an element $z \in R$, which fulfil all the conditions stated in γ. This ensures $\overline{L}$ to be feasible for P_{r_0,C_0}, which uses the fact that we get $\overline{L}(\overline{g}_j^2) = 0$ for $j = 1,...,s$ from $L(\text{rrad}_R(g_1,...,g_s)_{2r_0}) = 0$.

Now we show as in Proposition 3.2.1 that z is positive and get

$$f + n\frac{\varepsilon}{2} + \delta\Theta_{r_0} \in \sum R[X]^2_{r_0} + \mathrm{rrad}_R(g_1, ..., g_s)_{2r_0}$$

for some $\delta \in R_{\geq 0}$ and $r \in \mathbb{N}$. Finally, and again as in the proof of Proposition 3.2.1, δ can be chosen smaller than $\frac{\varepsilon}{2}$, resulting in growing r_0. This proves the claim. □

As before, we use this result in $\mathbb{R}^*$. Therefore fix $n, s, N, d \in \mathbb{N}$, $l \in \mathbb{R}_{\geq 1}, \beta \in \mathbb{R}_{>0}$ and some $G \in \mathbb{R}_{\geq 0}$. Let ϕ be a first order logic formula over $\mathbb{R}$ that defines the set of all coefficients of polynomials $f, g_1, ..., g_s$ in n indeterminates of degree smaller or equal to d, which fulfil

$$\|f\|_\infty, \|g_1\|_\infty, ..., \|g_s\|_\infty \leq G,$$

$$[-(l-\beta), l-\beta]^n \cap N(g_1, ..., g_s) \neq \emptyset$$

and the property, that for all polynomials $g_1', ..., g_s'$ with $\deg g_i' \leq \deg g_i$ and $\|g_i - g_i'\|_\infty \leq \frac{1}{N}$ one has

$$f \geq 0 \qquad \text{on} \quad [-l, l]^n \cap N(g_1', ..., g_s').$$

As we have shown before, this can be done.

Now for fixed $\varepsilon \in \mathbb{R}_{>0}$, let δ_r be a formula over $\mathbb{R}$ that defines the set of all coefficients of polynomials $f, g_1, ..., g_s$ in n indeterminates of degree smaller or equal to d, which fulfil

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum R[X]^2_r + \mathrm{rrad}_R(g_1, ..., g_s)_{2r}.$$

This can also be done, as we have shown in similar cases several times before.

Proposition 3.2.3 now gives us

$$\phi(\mathbb{R}^*) \subseteq \bigcup_{r \in \mathbb{N}} \delta_r(\mathbb{R}^*),$$

whereas the $\aleph_1$-saturation ensures

$$\phi(\mathbb{R}^*) \subseteq \delta_{r_0}(\mathbb{R}^*) \tag{3.33}$$

for some r_0 depending in this case on n, s, N, d, l, β, G and ε.

Now the result from (3.33) can be written as a first order logic statement over $\mathbb{R}$, so by Tarski's Transfer Principle, we get the same result in $\mathbb{R}$, which is stated in the following theorem:

Theorem 3.2.4. *Let $n, s, d, N \in \mathbb{N}$, $\varepsilon, \beta \in \mathbb{R}_{>0}$, $l \in \mathbb{R}_{\geq 1}$ as well as $G \in \mathbb{R}_{\geq 0}$. Then there is an*

$$r = r(n, s, d, N, G, l, \beta, \varepsilon)$$

such that for all polynomials $f, g_1, ..., g_s \in \mathbb{R}[X_1, ..., X_n]_d$, which have the properties

(i) $[-(l-\beta), l-\beta]^n \cap N_{\mathbb{R}}(g_1, ..., g_s) \neq \emptyset$

(ii) $\|f\|_\infty, \|g_1\|_\infty, ..., \|g_s\|_\infty \leq G$

(iii) for all $g'_1, ..., g'_s \in \mathbb{R}[X]$ with $\deg g'_i \leq \deg g_i$ and $\|g_i - g'_i\|_\infty \leq \frac{1}{N}$ for $i = 1, ..., s$, one has $f \geq 0$ on $[-l, l]^n \cap N_{\mathbb{R}}(g'_1, ..., g'_s)$,

one gets

$$f + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{r} \frac{1}{l^{2k}} X_i^{2k} \in \sum \mathbb{R}[X]_r^2 + \mathrm{rrad}_{\mathbb{R}}(g_1, ..., g_s)_{2r}.$$

□

Chapter 4

Appendix

We used some non-trivial results in this work. The first one provides a sufficient condition, known as *Carleman's Condition*, for a linear form on $\mathbb{R}[X]$ to be given by a measure. A proof can be found in [Nus], Theorem 10.

Theorem 4.1.5 (Nussbaum). *Let $L : \mathbb{R}[X] \longrightarrow \mathbb{R}$ be a linear form and suppose $L(p^2) \geq 0$ for all $p \in \mathbb{R}[X]$. If in addition*

$$\sum_{k=1}^{\infty} L(X_i^{2k})^{-1/2k} = \infty \qquad \textit{for } i = 1, ..., n,$$

then there is a unique positive measure μ on $\mathbb{R}^n$ such that

$$L(h) = \int h d\mu \qquad \textit{for all } h \in \mathbb{R}[X].$$

Before we can give another solution to the moment problem, we need some more definitions.

Definition 3.2: A function $\varphi : \mathbb{N}^n \longrightarrow \mathbb{R}_{\geq 0}$ is called an *absolute value* if

(i) $\varphi((0, ..., 0)) = 1;$

(ii) $\varphi(\alpha + \beta) \leq \varphi(\alpha)\varphi(\beta)$ for $\alpha, \beta \in \mathbb{N}^n$.

Now we can state the following theorem, which is from [Ber], Theorem 8, or [BCR1], Chapter 4, Theorem 2.5:

Theorem 4.1.6. *Let $L : \mathbb{R}[X] \longrightarrow \mathbb{R}$ be a linear form and suppose $L(p^2) \geq 0$ for all $p \in \mathbb{R}[X]$. If there is an absolute value φ and a constant $C > 0$ such that*

$$|L(X^\alpha)| \leq C\varphi(\alpha) \qquad \text{for all } \alpha \in \mathbb{N}^n,$$

then there is a positive measure μ on $\mathbb{R}^n$ such that

$$L(h) = \int h d\mu \qquad \text{for all } h \in \mathbb{R}[X].$$

Furthermore, the measure is supported by the set

$$\{x \in \mathbb{R}^n \mid \forall \alpha \in \mathbb{N}^n \quad |x^\alpha| \leq \varphi(\alpha)\}.$$

In Proposition 2.2.1, we used the following corollary, which is also stated in [Ber], Theorem 9:

Corollary 4.1.7. *Let $L : \mathbb{R}[X] \longrightarrow \mathbb{R}$ be a linear form and suppose $L(p^2) \geq 0$ for all $p \in \mathbb{R}[X]$. If there are constants $C, l > 0$ such that*

$$|L(X^\alpha)| \leq Cl^{|\alpha|} \qquad \text{for all } \alpha \in \mathbb{N}^n,$$

then there is a positive measure on $[-l, l]^n$ such that

$$L(h) = \int h d\mu \qquad \text{for all } h \in \mathbb{R}[X].$$

Proof. $\varphi : \alpha \mapsto l^{|\alpha|}$ is obviously an absolute value. And we have

$$\{x \in \mathbb{R}^n \mid \forall \alpha \in \mathbb{N}^n \quad |x^\alpha| \leq \varphi(\alpha)\} = [-l, l]^n.$$

□

Last, we give the proof for the denseness of the cone of sums of squares in the cone of polynomials nonnegative on $[-1,1]^n$ from [BCR2], Theorem 9.1. Remember that the $\|.\|_1$-norm of a polynomial $f = \sum_{\alpha\in\mathbb{N}^n} f_\alpha X^\alpha$ is defined as

$$\|f\|_1 := \sum_{\alpha\in\mathbb{N}^n} |f_\alpha|.$$

Theorem 4.1.8 (Berg, Christensen, Ressel). *Let $f \in \mathbb{R}[X]$ be such that $f(x) \geq 0$ for all $x \in [-1,1]^n$. Then f can be approximated in the $\|.\|_1$-norm by sums of squares of polynomials from $\mathbb{R}[X]$.*

Proof. Let Δ be the closure of the convex cone of sums of squares with respect to the $\|.\|_1$-norm and suppose $f \notin \Delta$. Then we get a $\|.\|_1$-continuous linear form $\psi : \mathbb{R}[X] \longrightarrow \mathbb{R}$ that satisfies

$$\psi(\Delta) \subseteq [0,\infty) \text{ and } \psi(f) < 0.$$

This is an immediate corollary of Hahn-Banach's Theorem as for example stated in [Wer], Theorem III.2.5.

As ψ is continuous and $\|X^\alpha\|_1 = 1$ for all $\alpha \in \mathbb{N}^n$, we find $|\psi(X^\alpha)| \leq C$ for some $C \in \mathbb{R}_{>0}$ and all $\alpha \in \mathbb{N}^n$. By Corollary 4.1.7, ψ is given by a measure on $[-1,1]^n$, which contradicts $\psi(f) < 0$ (as f is nonnegative on $[-1,1]^n$). $\square$

Bibliography

[Ber] C. Berg: *The multidimensional moment problem and semigroups*, Proc. Symp. Appl. Math. **37**, 110-124 (1987).

[BCR1] C. Berg, J. P. R. Christensen, P. Ressel: *Harmonic analysis on semigroups*, Springer, New York (1984).

[BCR2] C. Berg, J. P. R. Christensen, P. Ressel: *Positive definite functions on abelian semigroups*, Math. Ann. **223**, 253-274 (1976).

[Las1] J. B. Lasserre: *A sum of squares approximation of nonnegative polynomials*, `http://arxiv.org/pdf/math.AG/0412398`.

[Las2] J. B. Lasserre: *S.O.S. approximation of polynomials nonnegative on a real algebraic set*, `http://arxiv.org/pdf/math.AG/0412400`.

[Mar] M. Marshall: *Optimization of polynomial functions*, Canad. Math. Bull. **46**, 575-587 (2003).

[NN] Y. E. Nesterov, A. S. Nemirovski: *Interior-point polynomial algorithms in convex programming*, SIAM, Philadelphia (1994).

[Nus] A. E. Nussbaum: *Quasi-analytic vectors*, Arkiv för matematik **6** (10), 179-191 (1966).

[PD] A. Prestel, C. N. Delzell: *Positive polynomials*, Springer, Berlin (2001).

[Sch] M. Schweighofer: *Optimization of polynomials on compact semialgebraic sets*, SIAM Journal on Optimization **15**, No. 3, 805-825 (2005).

[Tod] J. M. Todd: *Semidefinite optimization*, Acta Numerica **10**, 515-560 (2001).

[VB] L. Vandenberghe, S. Boyd: *Semidefinite programming*, SIAM review **38** (1), 49-95 (1996).

[Wer] D. Werner: *Funktionalanalysis. 4., überarbeitete Auflage*, Springer, Berlin (2002).

Erklärung

Ich versichere hiermit, dass ich die vorliegende Diplomarbeit mit dem Thema

High Degree Perturbations of Nonnegative Polynomials

selbständig verfasst und keine anderen Hilfsmittel als die angegebenen benutzt habe. Die Stellen, die aus anderen Werken dem Wortlaut oder dem Sinne nach entnommen sind, habe ich in jedem einzelnen Fall durch Angabe der Quelle, auch der benutzten Sekundärliteratur, als Entlehnung kenntlich gemacht.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Konstanz, den 20. Juni 2005

Tim Netzer