Introduction

The purpose of the book is to provide the beginning student with a short introduction to recent work in real algebraic geometry and optimization growing out of Schmüdgen's solution of the moment problem in 1991. It is not intended for experts, and it is not meant to be comprehensive. For a more complete coverage, the reader will need to consult the literature.

For prerequisites, only standard advanced undergraduate or beginning graduate level courses in algebra, analysis and topology are assumed. Most of the book is accessible to a determined reader with this sort of background. For the most part, the objects being dealt with are polynomials in n variables with real coefficients. At certain points a little more commutative algebra is required. Here I refer to the book by Atiyah and MacDonald [A-M]. For results on formal power series, I refer to the book by Zariski and Samuel [Z-S]. The reader may safely omit the most difficult proofs on the first reading with little overall loss of understanding. The two appendices at the end fill in some of the gaps. At some stage, the reader will need to learn some of the material in the book by Bochnak, Coste and Roy [B-C-R]. Chapter 1 of [B-C-R] is a prerequisite for Appendix 1. Chapter 2 of [B-C-R], though less essential perhaps, provides missing background in semialgebraic geometry (e.g., cell decomposition and the geometric description of dimension).

The subject has undergone a rather continuous development over the course of the last 120 years, beginning with the paper of Hilbert [Hil1] in 1888. Major progress came with Artin's solution of Hilbert's 17th Problem [A] in 1927 and then, again, with the results of Tarski [T] in 1931 and Seidenberg [Sei] in 1954. These results led, eventually, to the famous Positivstellensatz, discovered by Krivine [Kr1] in 1964 and rediscovered be Stengle [St1] in 1974, which, in a certain sense, marked the beginning of modern real algebraic geometry.

Classical algebraic geometry deals with subsets of $\mathbb{C}^n$ defined by polynomial equations (algebraic sets). Real algebraic geometry (more precisely, semialgebraic geometry) deals with subsets of $\mathbb{R}^n$ defined by polynomial equations and inequalities (semialgebraic sets). There are certain similarities in the two subjects, but there are also important differences. Ideas and techniques from the former are useful in the latter, but new ideas and techniques are also required to handle the special problems inherent in semialgebaic geometry. Both consider ideals in the polynomial ring. In complex algebraic geometry, the polynomial ring is $\mathbb{C}[\underline{X}] := \mathbb{C}[X_1, \dots, X_n]$, in real algebraic geometry it is $\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \dots, X_n]$. But, in semialgebraic geometry, one also needs to consider preorderings.

A *preordering* of a ring[1] A is a subset T of A satisfying $T + T \subseteq T$, $T \cdot T \subseteq T$, and $f^2 \in T$ for all $f \in A$. The ideal of A generated by $g_1, \dots, g_s \in A$ consists of all elements of the form $\sum_{i=1}^{s} f_i g_i$, $f_i \in A$, $i = 1, \dots, s$. The preordering of A

[1] All rings considered here are commutative with 1.

generated by $g_1, \ldots, g_s \in A$ consists of all elements $\sum_e \sigma_e \underline{g}^e$, where $e := (e_1, \ldots, e_s)$ runs through the set $\{0,1\}^s$, $\underline{g}^e := g_1^{e_1} \cdots g_s^{e_s}$, and each σ_e is a sum of squares of elements of A.

The classical Nullstellensatz asserts that if $f \in \mathbb{C}[\underline{X}]$ vanishes on the subset of $\mathbb{C}^n$ defined by $g_1 = 0, \cdots, g_s = 0$, $(g_1, \ldots, g_s \in \mathbb{C}[\underline{X}])$ then some power of f lies in the ideal of $\mathbb{C}[\underline{X}]$ generated by $g_1, \ldots, g_s$. The analogous result in semialgebraic geometry is the Positivstellensatz. The Positivstellensatz (at least, one version of it) asserts that if $f \in \mathbb{R}[\underline{X}]$ is strictly positive on the subset of $\mathbb{R}^n$ defined by $g_1 \geq 0, \cdots, g_s \geq 0$, $(g_1, \ldots, g_s \in \mathbb{R}[\underline{X}])$ then $pf = 1 + q$ for some p, q in the preordering of $\mathbb{R}[\underline{X}]$ generated by $g_1, \ldots, g_s$.

A big difference between the real case and the complex case is that the basic results in the real case (like the Positivstellensatz) are harder to prove. One needs results from the model theory of real closed fields (the results of Tarski [T] and Seidenberg [Sei] referred to above).

More recently, the subject took a new turn, with Schmüdgen's solution of the moment problem, first in the compact case [Sm2] 1991, and then, later, in a large number of non-compact cases as well [Sm3] 2003. Schmüdgen's proofs in [Sm2] and [Sm3] combine the Positivstellensatz with ideas from functional analysis. In [Sm2], Schmüdgen also proved, as a corollary of his main result, a new unexpected denominator-free version of the Positivstellensatz, in the compact case [Sm2, Cor. 3], although there was a gap in his original proof. This gap was plugged eventually, with Wörmann's proof [W], in 1998.

The moment problem is the question of when, given a closed subset K in $\mathbb{R}^n$, a linear map $L : \mathbb{R}[\underline{X}] \to \mathbb{R}$ corresponds to a (positive) Borel measure μ on K (in the sense that $L(f) = \int f \, d\mu \ \forall \ f \in \mathbb{R}[\underline{X}]$). For example, in the one-variable case, if $K = [0, \infty)$, this is the case iff $L(f^2) \geq 0$ and $L(Xf^2) \geq 0$ for all $f \in \mathbb{R}[X]$ (or, equivalently, if the $\infty \times \infty$ matrices

$$S_1 := \begin{pmatrix} s_0 & s_1 & s_2 & \ldots \\ s_1 & s_2 & \ldots & \\ s_2 & \ldots & & \\ \ldots & & & \end{pmatrix} \text{ and } S_X := \begin{pmatrix} s_1 & s_2 & s_3 & \ldots \\ s_2 & s_3 & \ldots & \\ s_3 & \ldots & & \\ \ldots & & & \end{pmatrix}$$

are positive semidefinite, where $s_i := L(X^i)$, $i = 0, 1, \cdots$.) This was proved already by Stieltjes [Sti], in 1885. Over the years, a variety of results of this sort have been proved, e.g., Hausdorff showed in [Hau] 1923 that, if $K = [0,1]$, the corresponding conditions are that $L(f^2) \geq 0$, $L(Xf^2) \geq 0$ and $L((1-X)f^2) \geq 0$ for all $f \in \mathbb{R}[X]$ (or, equivalently, that the matrices S_1, S_X and S_{1-X} are positive semidefinite, where $S_{1-X} := S_1 - S_X$).

In [Sm2], Schmüdgen shows that a similar result holds when K is any compact basic closed semialgebraic set in $\mathbb{R}^n$: If K is defined by the polynomial inequalities $g_1 \geq 0, \cdots, g_s \geq 0$ and T denotes the preordering of $\mathbb{R}[\underline{X}]$ generated by $g_1, \ldots, g_s$, then any linear map $L : \mathbb{R}[\underline{X}] \to \mathbb{R}$ satisfying $L \geq 0$ on T corresponds to a (positive) Borel measure μ on K (and conversely). In this same paper Schmüdgen states his famous denominator-free Positivstellensatz: If K is a compact basic closed semialgebraic set in $\mathbb{R}^n$ defined by the polynomial inequalities $g_1 \geq 0, \cdots, g_s \geq 0$, then any $f \in \mathbb{R}[\underline{X}]$ strictly positive on K belongs to the preordering of $\mathbb{R}[\underline{X}]$ generated by $g_1, \ldots, g_s$.

Then, bit later, Jacobi [J] 2001, motivated by the results of Schmüdgen [Sm2] and Putinar [Pu] 1993, proves his Representation Theorem for Archimedean quadratic modules (a sort of denominator-free Positivstellensatz for Archimedean quadratic modules) which allows one, under the appropriate conditions, to replace preorderings by quadratic modules in Schmüdgen's results.

A *quadratic module* of a ring A is a subset M of A satisfying $M + M \subseteq M$, $f^2 M \subseteq M$ for all $f \in A$, and $1 \in M$. The quadratic module of A generated by $g_1, \dots, g_s \in A$ consists of all elements $\sum_{i=0}^{s} \sigma_i g_i$, where each σ_i is a sum of squares in A and $g_0 := 1$. A quadratic module M of A is said to be *Archimedean* if, for each $f \in A$, there exists an integer $k \geq 1$ such that $k + f \in M$.

At the same time, Jacobi and Prestel [J-P] 2001 give a valuation-theoretic criterion for deciding when the quadratic module of $\mathbb{R}[\underline{X}]$ generated by $g_1, \dots, g_s$ is Archimedean, given that the basic closed semialgebraic set in $\mathbb{R}^n$ defined by $g_1 \geq 0, \cdots, g_s \geq 0$ is compact.

Of course, it is natural to wonder what happens in Schmüdgen's Positivstellensatz when the condition $f > 0$ on K is replaced by the weaker condition $f \geq 0$ on K. Is it still true that $f \in T$? Scheiderer investigates this rather delicate question in a series of papers, beginning with [S1] 1999, and develops a local-global principle which allows one to reduce the question to a question about formal power series rings (at least, in certain cases). It turns out that it is never true if $\dim(K) \geq 3$, but it is true, in certain cases, if $\dim(K) \leq 2$.

It has been understood for some time that the problem of deciding when a polynomial is a sum of squares is 'easier' than deciding when it is non-negative. Recently, beginning with the papers of Shor [Sho] 1987, Shor and Stetsyuk [S-S] 1997, Lasserre [Las1] 2000, [Las2] 2001 and Parrilo and Sturmfels [Pa-S] 2003, this idea has been exploited to optimize a polynomial using semidefinite programming.

The basic algorithm goes as follows: Suppose $f \in \mathbb{R}[\underline{X}]$ has degree $\leq d$. Thus f has a presentation $f = \sum_{|\alpha| \leq d} c_\alpha \underline{X}^\alpha$, $c_\alpha \in \mathbb{R}$, where $\underline{X}^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n}$, $\alpha = (\alpha_1, \dots, \alpha_n)$, each α_i is an integer ≥ 0, and $|\alpha| := \alpha_1 + \cdots + \alpha_n$. Let $\mathbb{R}[\underline{X}]_d$ denote the vector space consisting of all polynomials of degree $\leq d$, and let $\mathcal{X}_d$ denote the set of all linear maps $L : \mathbb{R}[\underline{X}]_d \to \mathbb{R}$ such that $L(1) = 1$ and $L \geq 0$ at each element of $\mathbb{R}[\underline{X}]_d$ which is a square.

$$f_+ := \inf\{L(f) \mid L \in \mathcal{X}_d\}$$

is a lower bound for f on $\mathbb{R}^n$. Each $L \in \mathcal{X}_d$ is determined by the multisequence $(s_\alpha)_{|\alpha| \leq d}$ defined by $s_\alpha := L(\underline{X}^\alpha)$. Since $L(1) = 1$, $s_0 = 1$. Since L is ≥ 0 on squares, the $|\Lambda| \times |\Lambda|$ matrix $(s_{\alpha+\beta})_{\alpha,\beta \in \Lambda}$ is positive semidefinite, where $\Lambda := \{\alpha \mid |\alpha| \leq \frac{d}{2}\}$. Thus, to compute f_+ one needs to

$$\begin{cases} \text{minimize } \sum_{|\alpha| \leq d} c_\alpha s_\alpha \\ \text{subject to } s_0 = 1 \text{ and } (s_{\alpha+\beta})_{\alpha,\beta \in \Lambda} \text{ is positive semidefinite,} \end{cases}$$

which is a semidefinite programming problem. Every semidefinite programming problem has an associated dual problem, which is also a semidefinite programming problem [Lo] [V-B]. In this case, the dual program computes

$$\overline{f}_+ := \sup\{\lambda \in \mathbb{R} \mid f - \lambda \text{ is a sum of squares in } \mathbb{R}[\underline{X}]\}.$$

This comes from the fact that $g \in \mathbb{R}[\underline{X}]_d$ is a sum of squares iff g has a presentation $g = \sum_{\alpha,\beta\in\Lambda} a_{\alpha\beta}\underline{X}^\alpha \underline{X}^\beta$ where the $|\Lambda| \times |\Lambda|$ matrix $(a_{\alpha\beta})_{\alpha,\beta\in\Lambda}$ is positive semidefinite. It follows that, to compute $\overline{f}_+$ one must

$$\begin{cases} \text{maximize } \lambda \\ \text{subject to } c_0 - \lambda = a_{00},\ c_\gamma = \sum_{\alpha+\beta=\gamma} a_{\alpha\beta} \text{ for } \gamma \neq 0 \\ \text{and } (a_{\alpha\beta})_{\alpha,\beta\in\Lambda} \text{ is positive semidefinite,} \end{cases}$$

which is the dual problem. Actually, the duality gap here is zero, i.e., $\overline{f}_+ = f_+$. Of course, since the computation is based on semidefinite programming, it can be carried out in polynomial time. This explains why the problem of deciding if f is a sum of squares is 'easy'. (Just compute $\overline{f}_+$. f is a sum of squares iff $\overline{f}_+ \geq 0$.)

A refinement of this basic algorithm, due to Lasserre [Las1] [Las2], allows one to compute lower bounds for f on any basic closed semialgebraic set K in $\mathbb{R}^n$. The various results of Schmüdgen, Putinar, Jacobi, Jacobi and Prestel, and Scheiderer, referred to above, provide theoretical justification for the method. Another refinement of the algorithm, this one exploiting the gradient ideal of f, is given by Nie, Demmel and Sturmfels, in [N-D-S] 2006. The theoretical justification in this case comes from a consideration of the connected components of the set of complex zeros of the gradient ideal.