Positive Polynomials and Sums of Squares

Murray Marshall
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>vii</td>
</tr>
<tr>
<td>Introduction</td>
<td>ix</td>
</tr>
<tr>
<td>Chapter 0. Preliminaries</td>
<td>1</td>
</tr>
<tr>
<td>0.1 Notations</td>
<td>1</td>
</tr>
<tr>
<td>0.2 Positive Semidefinite Matrices</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 1. Positive Polynomials and Sums of Squares</td>
<td>3</td>
</tr>
<tr>
<td>1.1 Preliminaries on Polynomials</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Positive Polynomials</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Extending Positive Polynomials</td>
<td>8</td>
</tr>
<tr>
<td>1.4 Hilbert's 17th Problem</td>
<td>11</td>
</tr>
<tr>
<td>1.5 Baer-Krull Theorem</td>
<td>14</td>
</tr>
<tr>
<td>1.6 Formal Power Series Rings</td>
<td>17</td>
</tr>
<tr>
<td>Chapter 2. Krivine's Positivstellensatz</td>
<td>21</td>
</tr>
<tr>
<td>2.1 Quadratic Modules and Preorderings</td>
<td>21</td>
</tr>
<tr>
<td>2.2 Positivstellensatz</td>
<td>25</td>
</tr>
<tr>
<td>2.3 The Proof</td>
<td>27</td>
</tr>
<tr>
<td>2.4 The Real Spectrum</td>
<td>29</td>
</tr>
<tr>
<td>2.5 Abstract Positivstellensatz</td>
<td>31</td>
</tr>
<tr>
<td>2.6 Saturation</td>
<td>33</td>
</tr>
<tr>
<td>2.7 Low-Dimensional Examples</td>
<td>35</td>
</tr>
<tr>
<td>Chapter 3. The Moment Problem</td>
<td>41</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>41</td>
</tr>
<tr>
<td>3.2 Proof of Haviland's Theorem</td>
<td>44</td>
</tr>
<tr>
<td>3.3 Uniqueness Question</td>
<td>46</td>
</tr>
<tr>
<td>3.4 The Conditions (SMP) and (MP)</td>
<td>47</td>
</tr>
<tr>
<td>3.5 Schmüdgen's Theorem</td>
<td>48</td>
</tr>
<tr>
<td>3.6 Countable Dimensional Vector Spaces</td>
<td>50</td>
</tr>
<tr>
<td>Chapter 4. Non-Compact Case</td>
<td>55</td>
</tr>
<tr>
<td>4.1 Stability</td>
<td>55</td>
</tr>
<tr>
<td>4.2 Examples where (SMP) and (MP) fail</td>
<td>61</td>
</tr>
<tr>
<td>4.3 Examples where (SMP) and (MP) hold</td>
<td>64</td>
</tr>
<tr>
<td>4.4 Direct Integral Decomposition</td>
<td>65</td>
</tr>
<tr>
<td>Chapter 5. Archimedean T-modules</td>
<td>71</td>
</tr>
<tr>
<td>5.1 Preprimes</td>
<td>71</td>
</tr>
</tbody>
</table>
5.2 $T$-modules 72
5.3 Semiorderings and Valuations 75
5.4 Representation Theorem 78
5.5 Theorems of Pólya and Reznick 80
5.6 Other Applications 83
5.7 Topology on $\mathcal{V}_A = \text{Hom}(A, \mathbb{R})$ 84

Chapter 6. Schmüdgen's Positivstellensatz 87
6.1 Wörmann's Trick 87
6.2 Non-Compact Case 89
6.3 Remarks and Examples 92

Chapter 7. Putinar's Question 97
7.1 Introduction 97
7.2 Stable Compactness 100
7.3 Jacobi-Prestel Counterexample 103
7.4 The Case $\text{dim} \frac{\mathbb{R}[X]}{\mathbb{M} \cap -\mathbb{M}} \leq 1$ 105

Chapter 8. Weak Isotropy of Quadratic Forms 109
8.1 Isotropy and Weak Isotropy 109
8.2 Residue Forms 110
8.3 Local-Global Principle for Weak Isotropy 113
8.4 Pfister Forms 116
8.5 Application to Putinar's Question 117

Chapter 9. Scheiderer's Local-Global Principle 123
9.1 Basic Lemma 123
9.2 Local-Global Principle 125
9.3 The Case $n = 1$ 128
9.4 The Case $n = 2$ 130
9.5 Hessian Conditions 133
9.6 Second Local-Global Principle 134

Chapter 10. Semidefinite Programming and Optimization 137
10.1 The Cone of PSD Matrices 137
10.2 Semidefinite Programming 138
10.3 Max-Cut Problem 142
10.4 Global Optimization 145
10.5 Constrained Optimization 148
10.6 Exploiting the Gradient Ideal 151
10.7 Existence of Feasible Solutions 156

Appendix 1. Tarski-Seidenberg Theorem 161
11.1 Basic Version 161
11.2 Tarski's Transfer Principle 162
11.3 Lang’s Homomorphism Theorem 163
11.4 Geometric Version 165
11.5 General Version 167

Appendix 2. Algebraic Sets 169
12.1 Transcendence Degree and Krull Dimension 169
## CONTENTS

12.2 Non-Singular Zeros  
12.3 Algebraic Sets  
12.4 Dimension  
12.5 Radical Ideals and Real Ideals  
12.6 Simple Point Criterion  
12.7 Sign-Changing Criterion  

Bibliography
Preface

Let me begin with a short history of the book. In 2000 I gave seminar lectures at the University of Saskatchewan and, later, also at Università di Pisa. The notes from these lectures were written up and appeared under the title “Positive polynomials and sums of squares” in the series Dottorato de Ricerca in Matematica, published by Università di Pisa, in 2000 [M1]. In 2004, I gave a graduate course at the University of Saskatchewan on the same topic. At that time, I expanded on the material considerably, including additional explanation, additional material on the moment problem, and material on the newly discovered application to optimization. These revised notes were used later, by Salma Kuhlmann, as a basis for her course at the special trimester in real algebraic and analytic geometry at Institut Henri Poincaré in Paris in 2005. In the spring of 2006 I started working on this again, the goal being to include additional recent developments and also to simplify arguments wherever possible. The reader who is familiar with [M1] will see various parts of [M1] surviving in what is presented here. At the same time, it is important to understand that a lot of new material has been added. The length has increased by a factor of three. Chapters 3, 4, 9 and 10 are completely new and, in the remaining chapters, new sections have been added, and the presentation has been modified in significant ways. Two appendices have also been added.

I want to thank everyone who contributed to the book in one way or another. A complete list, of course, would include everyone mentioned in the bibliography, and all the people who discussed various aspects of the subject with me, over the years. Special thanks are due to those who attended my course and seminar presentations on the subject, and provided feedback, and to Doris Augustin, Igor Klep and Tim Netzer for their valuable commentary in the latter stages of the project.

I should say finally that, in writing such a book, there are bound to be errors and mistakes of various sorts, and situations where the approach taken is wrong, or misleading, or, at the very least, not the best one. Of course, it would be nice to avoid all these things, but this is not going to happen – for a variety of reasons, but partly because the subject itself is still evolving – so I can only apologize to the reader in advance for all errors and mistakes and false impressions that I may convey, and hope there are not too many of them. A list of errata and corrections can be found on the webpage for this book. The URL is printed on the backcover of the book.
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 [Hi1] 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 by 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 semialgebraic geometry. Both consider ideals in the polynomial ring. In complex algebraic geometry, the polynomial ring is $\mathbb{C}[X] := \mathbb{C}[X_1,\ldots,X_n]$, in real algebraic geometry it is $\mathbb{R}[X] := \mathbb{R}[X_1,\ldots,X_n]$. 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 $g_1,\ldots,g_s \in A$ consists of all elements of the form $\sum_{i=1}^s f_i g_i$, $f_i \in A$, $i = 1,\ldots,s$. 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 \(\sum e \sigma eg^e\), where \(e := (e_1, \ldots, e_s)\) runs through the set \(\{0, 1\}^s\), \(g^e := g_1^{e_1} \cdots g_s^{e_s}\), and each \(\sigma e\) is a sum of squares of elements of \(A\).

The classical Nullstellensatz asserts that if \(f \in \mathbb{C}[X]\) vanishes on the subset of \(\mathbb{C}^n\) defined by \(g_1 = 0, \ldots, g_s = 0, (g_1, \ldots, g_s \in \mathbb{C}[X])\) then some power of \(f\) lies in the ideal of \(\mathbb{C}[X]\) 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 \(f \in \mathbb{R}[X]\) is strictly positive on the subset of \(\mathbb{R}^n\) defined by \(g_1 \geq 0, \ldots, g_s \geq 0, (g_1, \ldots, g_s \in \mathbb{R}[X])\) then \(pf = 1 + q\) for some \(p, q\) in the preordering of \(\mathbb{R}[X]\) 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 Schm"udgen’s solution of the moment problem, first in the compact case [Sm2] 1991, and then, later, in a larger number of non-compact cases as well [Sm3] 2003. Schm"udgen’s proofs in [Sm2] and [Sm3] combine the Positivstellensatz with ideas from functional analysis. In [Sm2], Schm"udgen 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 W"ormann’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 \(L : \mathbb{R}[X] \to \mathbb{R}\) corresponds to a (positive) Borel measure \(\mu\) on \(K\) (in the sense that \(L(f) = \int f d\mu \forall f \in \mathbb{R}[X]\)). For example, in the one-variable case, if \(K = [0, \infty)\), this is the case if \(L(f^2) \geq 0\) and \(L(Xf^2) \geq 0\) for all \(f \in \mathbb{R}[X]\) (or, equivalently, if the \(\infty \times \infty\) matrices

\[
S_1 := \begin{pmatrix} s_0 & s_1 & s_2 & \cdots \\ s_1 & s_2 & \cdots \\ s_2 & \cdots \\ \vdots \end{pmatrix} \quad \text{and} \quad S_X := \begin{pmatrix} s_1 & s_2 & s_3 & \cdots \\ s_2 & s_3 & \cdots \\ s_3 & \cdots \\ \vdots \end{pmatrix}
\]

are positive semidefinite, where \(s_i := L(X^i), i = 0, 1, \ldots\) ) 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], Schm"udgen 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 \(g_1 \geq 0, \ldots, g_s \geq 0\) and \(T\) denotes the preordering of \(\mathbb{R}[X]\) generated by \(g_1, \ldots, g_s\), then any linear map \(L : \mathbb{R}[X] \to \mathbb{R}\) satisfying \(L \geq 0\) on \(T\) corresponds to a (positive) Borel measure \(\mu\) on \(K\) (and conversely). In this same paper Schm"udgen states his famous denominator-free Positivstellensatz: If \(K\) is a compact basic closed semialgebraic set in \(\mathbb{R}^n\) defined by the polynomial inequalities \(g_1 \geq 0, \ldots, g_s \geq 0\), then any \(f \in \mathbb{R}[X]\) strictly positive on \(K\) belongs to the preordering of \(\mathbb{R}[X]\) 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 qua-
dratic modules (a sort of denominator-free Positivstellensatz for Archimedean qua-
dratic 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 $g_1, \ldots, g_s \in A$ consists of all elements $\sum_{i=0}^s \sigma_i g_i$, where each $\sigma_i$
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 $\mathbb{R}[X]$ generated by $g_1, \ldots, g_s$
is Archimedean, given that the basic closed semialgebraic set in $\mathbb{R}^n$ defined by
$g_1 \geq 0, \ldots, g_s \geq 0$ is compact.

Of course, it is natural to wonder what happens in Schmüdgen’s Positivstel-
len satz 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 [SI] 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]
idea has been exploited to optimize a polynomial using semidefinite programming.

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

$$f_+ := \inf \{L(f) \mid L \in X_d\}$$

is a lower bound for $f$ on $\mathbb{R}^n$. Each $L \in X_d$ is determined by the multisequence $(s_\alpha)_{|\alpha| \leq d}$ defined by $s_\alpha := L(X^\alpha)$. Since $L(1) = 1$, $s_0 = 1$. Since $L$ is $\geq 0$ 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

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

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

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

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

which is the dual problem. Actually, the duality gap here is zero, i.e., $\bar{f}_+ = f_+$. 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 $\bar{f}_+$. $f$ is a sum of squares iff $\bar{f}_+ \geq 0$.)

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.
APPENDIX 1

Tarski-Seidenberg Theorem

Tarski’s Transfer Principle 1.4.2 has played a big role in what we have been doing so far, but, actually, Tarski’s Transfer Principle is just a corollary of an even more basic result called the Tarski-Seidenberg Theorem. Since the Tarski-Seidenberg Theorem has other important applications as well, it is something that everyone needs to learn eventually. There are various versions, see 11.1.1, 11.4.2 and 11.5.1. For the beginner, the best place to start is Chapter 1 of the book [B-C-R]. This contains all the results we need concerning ordered fields, real closed fields and real closures, and it also contains a proof of the basic version of the Tarski-Seidenberg Theorem that we need. This is not an easy proof to read (but, at the same time, there seems to be no way to make it easier). Anyway, we assume the reader is at least somewhat familiar with this material, and take this as our starting point. In Section 11.2 we explain how Tarski’s Transfer Principle can be derived as a consequence of the Tarski-Seidenberg Theorem. In Section 11.3 we explain the relationship to Lang’s Homomorphism Theorem.

11.1 Basic Version

We consider systems of polynomial equations and inequalities of the form

\[ S(X) : \begin{cases} f_1(X) \triangleright_1 0 \\ f_2(X) \triangleright_2 0 \\ \vdots \\ f_k(X) \triangleright_k 0 \end{cases} \]

where \( \triangleright_i \in \{\geq, >, =, \neq\} \) and each \( f_i(X) \) is a polynomial in \( n \) variables \( X_1, \ldots, X_n \) with coefficients in \( \mathbb{Q} \) or in some real closed field \( R \). We consider the following basic version of the Tarski-Seidenberg Theorem, which was announced already by Tarski in 1931, in the case \( R = \mathbb{R} [T] \). The proof of Seidenberg appeared later, in 1954 [Sei].

11.1.1 TARSKI-SEIDENBERG THEOREM (Basic Version). Given a system of polynomial equations and inequalities \( S(T, X) \) in \( m+n \) variables \( T_1, \ldots, T_m, X_1, \ldots, X_n \) with coefficients in \( \mathbb{Q} \), there exist finitely many systems of polynomial equations and inequalities \( S_1(T), \ldots, S_\ell(T) \) with coefficients in \( \mathbb{Q} \) such that, for each real closed field \( R \) and each \( t = (t_1, \ldots, t_m) \in R^m \), the system \( S(t, X) \) has a solution \( x = (x_1, \ldots, x_n) \in R^n \) iff \( t \) is a solution of one of the systems \( S_1(T), \ldots, S_\ell(T) \).

In the applications that we discuss here, one does not need to know how to compute the systems \( S_i(T) \). One only needs to know that they exist.

Exercise: Determine the \( S_i(T) \) in case \( m = 3, n = 1 \) and the system \( S(T, X) \) consists of the single equation \( T_1 X^2 + T_2 X + T_3 = 0 \).
The proof of 11.1.1 is by induction on \( n \). See [B-C-R, Th. 1.4.2] for the proof in the case \( n = 1 \). In contrast to the case \( n = 1 \), the inductive step of the proof is almost trivial:

**Proof.** Assume \( n > 1 \). By the case \( n = 1 \) there exist finitely many systems \( S_i(T, X_1, \ldots, X_{n-1}), i = 1, \ldots, \ell \) such that for each real closed field \( R \) and each choice of \( (t, x_1, \ldots, x_{n-1}) \in R^{n+n-1} \), the system \( S(t, x_1, \ldots, x_{n-1}, X_n) \) has a solution \( x_n \in R \) iff \( (t, x_1, \ldots, x_{n-1}) \) is a solution of some \( S_i(T, X_1, \ldots, X_{n-1}) \). By induction on \( n \), for each \( 1 \leq i \leq \ell \) there exists systems \( S_{ij}(T), j = 1, \ldots, \ell_i \) such that, for each real closed field \( R \) and each \( t \in R^m \), the system \( S_i(t, X_1, \ldots, X_{n-1}) \) has a solution \( (x_1, \ldots, x_{n-1}) \in R^{n-1} \) iff \( t \) is a solution of one of the systems \( S_{ij}(T) \).

Pure logic now shows that for any real closed field \( R \) and any \( t \in R^m \), the system \( S(t, X) \) has a solution \( x \in R^n \) iff \( t \) is a solution of one of the systems \( S_{ij}(T) \). \( \square \)

### 11.2 Tarski’s Transfer Principle

In this section we apply the Tarski-Seidenberg Theorem 11.1.1 to establish various versions of Tarski’s Transfer Principle. As an immediate consequence of 11.1.1, we obtain the following:

**11.2.1 Transfer Principle.** Suppose \((K, \leq)\) is an ordered field, \( R_1 \) and \( R_2 \) are real closed extensions of \((K, \leq)\) and \( t \in K^m \). Then the system \( S(t, X) \) has a solution \( x \in R_1^n \) iff it has a solution \( x \in R_2^n \).

**Proof.** \( S(t, X) \) has a solution \( x \in R_1^n \) iff \( t \) is a solution of some \( S_i(T) \) iff \( S(t, X) \) has a solution \( x \in R_2^n \). \( \square \)

11.2.1 is stated more naturally as follows:

**11.2.2 Transfer Principle.** Suppose \((K, \leq)\) is an ordered field and \( R_1 \) and \( R_2 \) are real closed extensions of \((K, \leq)\). Then a system of polynomial equations and inequalities of the form

\[
S(X) : \begin{cases} f_1(X) \succ_1 0 \\ \vdots \\ f_k(X) \succ_k 0 
\end{cases}
\]

where \( \succ_i \in \{\geq, >, =, \neq\} \) and each \( f_i(X) \) is a polynomial in \( n \) variables with coefficients in \( K \), has a solution \( x \in R_1^n \) iff it has a solution \( x \in R_2^n \).

**Proof.** Let \( t_1, \ldots, t_m \) be the coefficients of the polynomials \( f_1, \ldots, f_k \), listed in some fixed order. Replacing the coefficients \( t_1, \ldots, t_m \) by variables \( T_1, \ldots, T_m \) yields a system \( S'(T, X) \) such that \( S'(t, X) \) is the system \( S(X) \). Now apply 11.2.1. \( \square \)

**11.2.3 Transfer Principle.** Suppose \( R \) and \( R_1 \) are real closed fields, \( R \subset R_1 \). Then a system of polynomial equations and inequalities of the form

\[
S(X) : \begin{cases} f_1(X) \succ_1 0 \\ \vdots \\ f_k(X) \succ_k 0 
\end{cases}
\]

where \( \succ_i \in \{\geq, >, =, \neq\} \) and each \( f_i(X) \) is a polynomial in \( n \) variables with coefficients in \( R \), has a solution \( x \in R_1^n \) iff it has a solution \( x \in R^n \).
Proof. Apply 11.2.2 with $K = R_2 = R$. □

Note: In the version of the Transfer Principle stated above, one implication is completely trivial.

Note: Although we have used 11.2.2 to prove 11.2.3, it is also possible to reverse the process:

Proof. Let $R$ denote the real closure of the ordered field $(K, \leq)$. By the uniqueness of the real closure we have embeddings $\phi_i : R \hookrightarrow R_i$ such that the composite mapping $K \hookrightarrow R \hookrightarrow R_i$ is the inclusion $K \hookrightarrow R_i$, $i = 1, 2$. By 11.2.3, $S(X)$ has a solution in $R_1^n$ iff $S(X)$ has a solution in $R^n$ iff $S(X)$ has a solution in $R_2^n$. □

We now prove the version of the Transfer Principle which is stated without proof in Section 1.4:

11.2.4 Transfer Principle. Suppose $R$ is a real closed field and $(F, \leq)$ is an ordered field extension of $R$. If a system of polynomial equations and inequalities of the form

$$S(X) : \begin{cases} f_1(X) \triangleright_1 0 \\ \vdots \\ f_k(X) \triangleright_k 0 \end{cases}$$

where $\triangleright_i \in \{\geq, >, =, \neq\}$ and each $f_i(X)$ is a polynomial in $n$ variables with coefficients in $R$, has a solution $x \in F^n$ then it has a solution $x \in R^n$.

Proof. Let $R_1$ denote the real closure of the ordered field $(F, \leq)$. Now apply 11.2.3. □

As is explained in Chapters 1 and 2, the solution of Hilbert’s 17th Problem, the Positivstellensatz and the Real Nullstellensatz are applications of 11.2.4.

11.3 Lang’s Homomorphism Theorem

Of course, the Transfer Principle applies, in particular, to systems of polynomial equations:

11.3.1 Homomorphism Theorem. Suppose $R$ and $R_1$ are real closed fields, $R \subseteq R_1$. Then a system of polynomial equations of the form

$$S(X) : \begin{cases} f_1(X) = 0 \\ \vdots \\ f_k(X) = 0 \end{cases}$$

where each $f_i(X)$ is a polynomial in $n$ variables with coefficients in $R$, has a solution $x \in R_1^n$ iff it has a solution $x \in R^n$.

Proof. Immediate from 11.2.3. □

We have the following, essentially equivalent, version of 11.3.1:
11.3.2 Homomorphism Theorem. Suppose \( R \) and \( R_1 \) are real closed fields, \( R \subseteq R_1 \). Then, for any ideal \( I \) in \( R[X] \), if there exists an \( R \)-algebra homomorphism \( \phi : R[X]/I \to R_1 \), then there exists an \( R \)-algebra homomorphism \( \psi : R[X]/I \to R \).

**Proof.** By the Hilbert Basis Theorem, the ideal \( I \) is finitely generated, say by \( f_1, \ldots, f_k \). Consider the system \( S(X) \) of polynomial equations defined as in the statement of 11.3.1. Observe that solutions of the system \( S(X) \) in \( R_1^n \) correspond bijectively to \( R \)-algebra homomorphisms \( \phi : R[X]/I \to R_1 \): If \( x \in R_1^n \) is a solution of \( S(X) \), then \( \phi(f + I) = f(x) \) is a well-defined \( R \)-algebra homomorphism from \( R[X]/I \) to \( R_1 \). If \( \phi : R[X]/I \to R_1 \) is a \( R \)-algebra homomorphism then, for any

\[
f = \sum_{\epsilon} a_{\epsilon} X^\epsilon = \sum_{\epsilon_1, \ldots, \epsilon_n} a_{\epsilon_1, \ldots, \epsilon_n} X_1^{\epsilon_1} \cdots X_n^{\epsilon_n} \in R[X],
\]

\[
\phi(f + I) = \sum_{\epsilon_1, \ldots, \epsilon_n} a_{\epsilon_1, \ldots, \epsilon_n} \phi(X_1 + I)^{\epsilon_1} \cdots \phi(X_n + I)^{\epsilon_n} = \phi(\phi(X_1 + I), \ldots, \phi(X_n + I)).
\]

Using this, one checks easily that \( (\phi(X_1 + I), \ldots, \phi(X_n + I)) \in R_1^n \) is a solution of the system \( S(X) \). Similarly, solutions of the system \( S(X) \) in \( R^n \) correspond bijectively to \( R \)-algebra homomorphisms \( \psi : R[X]/I \to R \). Now apply 11.3.1. \( \square \)

11.3.2 is often written in the following somewhat more obscure form:

11.3.3 Homomorphism Theorem. Suppose \( R \) and \( R_1 \) are real closed fields, \( R \subseteq R_1 \). Then, for any finitely generated \( R \)-algebra \( A \), if there exists an \( R \)-algebra homomorphism \( \phi : A \to R_1 \), then there exists an \( R \)-algebra homomorphism \( \psi : A \to R \).

**Proof.** Every finitely generated \( R \)-algebra is isomorphic to \( R[X]/I \) for some \( n \geq 1 \) and some ideal \( I \). Now apply 11.3.2. \( \square \)

Lang's original proof of the Homomorphism Theorem in 1953 did not use the Transfer Principle. In fact, it appears that Lang was not aware of the work of Tarski at the time. Rather, the ideas in Lang's proof come from certain 'specialization lemmas' that were developed by his Ph.D. Supervisor, Artin, in his solution of Hilbert's 17th Problem in 1927. See [La1, Th. 7] or [La2, Th. 5, p. 279] for Lang's proof. See T.-Y. Lam's real algebra notes [L3, Th. 5.4 and Cor. 5.5] for another proof of the Homomorphism Theorem.

It would appear that the Homomorphism Theorem is some sort of 'poor cousin' to the Transfer Principle, but this is not quite true: In fact, the Transfer Principle, at least the versions of the Transfer Principle given above, can be recovered easily as a corollary of the Homomorphism Theorem.

**Proof.** We use 11.3.3 to prove 11.2.3. By 1.4.3 any system of polynomial equations and inequalities with coefficients in \( R \) can be written in the form

\[
S(X) : \begin{cases}
f_1(X) \geq 0 \\
\vdots \\
f_k(X) \geq 0 \\
g(X) \neq 0
\end{cases}
\]
where \( f_1(X), \ldots, f_k(X), g(X) \) are polynomials in the variables \( X_1, \ldots, X_n \) with coefficients in \( R \). Consider the finitely generated \( R \)-algebra

\[
B = \frac{R[X_1, \ldots, X_n, Y_1, \ldots, Y_k, Z]}{(Y_1^2 - f_1, \ldots, Y_k^2 - f_k, gZ - 1)}.
\]

If \( S(X) \) has a solution \( x = (x_1, \ldots, x_n) \in R_1 \), where \( R_1 \) is some real closed extension of \( R \), then we have an \( R \)-algebra homomorphism \( \phi : B \to R_1 \) defined by \( X_i \mapsto x_i \), \( Y_j \mapsto \sqrt{f_j(x)} \), \( Z \mapsto 1/g(x) \). By 11.3.3, there exists an \( R \)-algebra homomorphism \( \psi : B \to R \). Now one checks easily that \((\psi(X_1), \ldots, \psi(X_n))\) is the required solution of \( S(X) \) in \( R^n \).

Thus we have a second proof of Tarski’s Transfer Principle (as a Corollary of Lang’s Homomorphism Theorem). But, unfortunately, Tarski’s Transfer Principle, by itself, is not quite enough. For many important applications, one needs the full strength of the Tarski-Seidenberg Theorem.

### 11.4 Geometric Version

Fix a real closed field \( R \) and an integer \( n \geq 1 \). We consider three operations on subsets of \( R^n \):

(1) finite union.
(2) finite intersection.
(3) complementation.

The class of \textit{semialgebraic sets} in \( R^n \) is defined to be the smallest class of subsets of \( R^n \) which contains all sets of the form

\[
\{ x \in R^n : f(x) > 0 \}
\]

where \( > \in \{ \geq, >, =, \neq \} \) and \( f \in R[X] \), and is closed under operations (1), (2) and (3). Given a system of equations and inequalities of the form

\[
S(X) : \begin{cases}
    f_1(X) >_1 0 \\
    \vdots \\
    f_k(X) >_k 0
\end{cases}
\]

where \( >_i \in \{ \geq, >, =, \neq \} \) and each \( f_i(X) \) is a polynomial in \( n \) variables with coefficients in \( R \), the set of solutions of \( S(X) \) in \( R^n \) is precisely the semialgebraic set

\[
\cap_{i=1}^k \{ x \in R^n \mid f_i(x) >_i 0 \}.
\]

Semialgebraic subsets of \( R^n \) of this special form are referred to as a \textit{basic semialgebraic sets} in \( R^n \).

#### 11.4.1 Proposition

A subset of \( R^n \) is semialgebraic iff it is a finite union of basic semialgebraic sets.
PROOF. One implication is clear. For the other, it suffices to show that the class consisting of all subsets of $\mathbb{R}^n$ which are finite unions of basic semialgebraic sets is closed under union, intersection and complementation. Closure under union is clear. Closure under intersection follows from $(\bigcup_i C_i) \cap (\bigcup_j D_j) = \bigcup_{i,j} (C_i \cap D_j)$. Using $\mathbb{R}^n \setminus (\bigcup_i C_i) = \bigcap_i (\mathbb{R}^n \setminus C_i)$ and $\mathbb{R}^n \setminus (\bigcup_j D_j) = \bigcup_j (\mathbb{R}^n \setminus D_j)$, showing closure under complementation reduces to showing it for sets of the form

$$\{ x \in \mathbb{R}^n \mid f(x) \succ 0 \}$$

where $\succ \in \{\geq, >, =, \neq\}$ and $f \in \mathbb{R}[\mathbb{X}]$. But this is clear: The complement of the set defined by $f(x) \geq 0$ is the set defined by $-f(x) > 0$. The complement of the set defined by $f(x) = 0$ is the set defined by $f(x) \neq 0$, and so forth.

A function $f : A \to B$ where $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ is said to be \textit{semialgebraic} if $A$ and $B$ are semialgebraic sets and the graph

$$ \text{gr}(f) = \{(x, y) \in \mathbb{R}^{m+n} \mid x \in A, y \in B, y = f(x)\} $$

is a semialgebraic set in $\mathbb{R}^{m+n}$.

In the language of semialgebraic geometry, the Tarski-Seidenberg Theorem takes on the following simple form (which is the version found in Hörmander’s book [Hör]):

11.4.2 TARSKI-SEIDENBERG THEOREM (Geometric Version). Consider the projection map $\pi : \mathbb{R}^{m+n} \to \mathbb{R}^m$ defined by $\pi(t, x) = t$. Then, for any semialgebraic set $A$ in $\mathbb{R}^{m+n}$, $\pi(A)$ is a semialgebraic set in $\mathbb{R}^m$.

PROOF. Since $\pi(\bigcup_{k=1}^t A_j) = \bigcup_{j=1}^t \pi(A_i)$, we can reduce to the case where $A$ is basic semialgebraic. Denote by $u_1, \ldots, u_k$ the coefficients of the various polynomials appearing in the description of $A$, written in some fixed order. Thus we have some system of polynomial equations and inequalities $S(U, T, X)$ with coefficients in $\mathbb{Q}$ such that $A$ is the set of solutions of the system $S(u, T, X)$ in $\mathbb{R}^{m+n}$. By 11.1.1 we have systems of polynomial equations and inequalities $S_i(U, T), i = 1, \ldots, \ell$, with coefficients in $\mathbb{Q}$ such that, for any $t \in \mathbb{R}^m$, the system $S(u, t, X)$ has a solution $x \in \mathbb{R}^n$ iff $(u, t)$ is a solution of some $S_i(U, T)$. It follows that

$$ \pi(A) = \{ t \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n \ (t, x) \in A \} $$

$$ = \{ t \in \mathbb{R}^m \mid t \text{ is a solution of } S_i(u, T) \text{ for some } i \in \{1, \ldots, \ell\} \} $$

$$ = \bigcup_{i=1}^\ell \{ t \in \mathbb{R}^m \mid t \text{ is a solution of } S_i(u, T) \}, $$

a semialgebraic set in $\mathbb{R}^m$.

Using 11.4.2 one can deduce various interesting and non-trivial consequences. For example, that the image of a semialgebraic function is semialgebraic and that the closure and interior of a semialgebraic set is semialgebraic. See [B-C-R, Sect. 2.1, 2.2] or [C, Sect. 2.1, 2.2]. If one starts writing out the proofs of these results, one realizes very soon that working with projections is cumbersome, and that it is much more natural to work with formulas.
11.5 General Version

We make precise what is meant by a first-order formula (in the language of real closed fields). A *first-order formula* is a formula obtained by the following constructions:

1. If \( f \in \mathbb{Q}[X_1, \ldots, X_n], n \geq 1 \), then \( f \geq 0, f > 0, f = 0, \) and \( f \neq 0 \) are first-order formulas.

2. If \( \Phi \) and \( \Psi \) are first-order formulas, then “\( \Phi \) and \( \Psi \)”, “\( \Phi \) or \( \Psi \)”, and “not \( \Phi \)” (often denoted by \( \Phi \lor \Psi, \Phi \land \Psi \) and \( \neg \Phi \), respectively) are first-order formulas.

3. If \( \Phi \) is a first-order formula then \( \exists X \Phi \) and \( \forall X \Phi \) are first order formulas.

The formulas obtained using constructions 1 and 2 are called *quantifier-free formulas*.

Two first-order formulas \( \Phi(X_1, \ldots, X_n) \) and \( \Psi(X_1, \ldots, X_n) \) are considered to be *equivalent* if for every real closed field \( R \) and every \( x \in R^n \), \( \Phi(x) \) holds in \( R \) iff \( \Psi(x) \) holds in \( R \).

One can easily, using elementary rules of logic, that every quantifier free-formula is equivalent to a finite disjunction of finite conjunctions of formulas obtained using the constructions in 1. The proof is completely similar to the proof of 11.4.1.

11.5.1 TARSKI-SEIDENBERG THEOREM (General Version). Every first-order formula in the language of real closed fields is equivalent to a quantifier-free formula. Otherwise put, the language of real closed fields admits elimination of quantifiers.

**PROOF.** It suffices to show that the class \( \mathcal{C} \) consisting of those first-order formulas which are equivalent to quantifier-free formulas is closed under the constructions in 2 and 3.

Closure of \( \mathcal{C} \) under \( \lor, \land \) and \( \neg \). This is clear: If \( \Phi \sim \Phi' \) and \( \Psi \sim \Psi' \), then \( \Phi \lor \Psi \sim \Phi' \lor \Psi' \), \( \Phi \land \Psi \sim \Phi' \land \Psi' \) and \( \neg \Phi \sim \neg \Phi' \).

Closure under the constructions in 2: Since \( \forall X \Phi \) is equivalent to \( \forall \exists X \neg \Phi \), it suffices to show that if the formula \( \Phi \) is equivalent to a quantifier-free formula then \( \exists X \Phi \) is equivalent to a quantifier-free formula. Since \( \Phi \) is equivalent to \( \Phi' \) implies \( \exists X \Phi \) is equivalent to \( \exists X \Phi' \), and \( \exists X (\phi_1 \land \cdots \land \phi_k) \) is equivalent to \( (\exists X \Phi_1) \lor \cdots \lor (\exists X \Phi_k) \), we are reduced to the case where \( \Phi \) is a finite conjunctions of polynomial equations and inequalities. Applying Theorem 11.1.1, there exist finitely many finite conjunctions of polynomial equations and inequalities \( \Theta_1, \ldots, \Theta_\ell \) such that \( \exists X \Phi \) is equivalent to \( \Theta_1 \lor \cdots \lor \Theta_\ell \).

In dealing with a particular real closed field \( R \), we often want to consider formulas of the form \( \Phi(t, X) \) where \( \Phi(T, X) \) is a first order-formula in the variables \( T_1, \ldots, T_m, X_1, \ldots, X_n \) and \( t \in R^m \). We refer to such a formula as a *first-order formula with parameters in \( R \)."™, \( t_1, \ldots, t_m \) are called the parameters.

11.5.2 COROLLARY. For any real closed field \( R \), semialgebraic sets in \( R^n \) are precisely the sets having the form \( S = \{ x \in R^n \mid \Phi(t, x) \} \), for some first order formula \( \Phi(t, X) \) with parameter \( t_1, \ldots, t_m \) in \( R \).

**PROOF.** This is immediate from 11.5.1. 

We also have the following general version of the Transfer Principle:
11.5.3 Transfer Principle (General Version). Suppose $R_1$ and $R_2$ are real closed extensions of the ordered field $(K, \leq)$. Let $\Phi(t)$ be any first-order formula with parameters $t_1, \ldots, t_m \in K$, with no free variables. Then $\Phi(t)$ holds in $R_1$ iff $\Phi(t)$ holds in $R_2$.

Proof. If the formula $\Phi(t)$ is quantifier free then the result is clear. But, by 11.5.1, the quantifier-free case is the only case that needs to be considered.

The following special case of 11.5.3 is worth mentioning:

11.5.4 Corollary. Suppose $R_1$ and $R_2$ are real closed fields. Suppose $\Phi$ is a first-order formula with parameters in $\mathbb{Q}$ and no free variables. Then $\Phi$ holds in $R_1$ iff $\Phi$ holds in $R_2$.

Proof. Apply 11.5.3 with $K = \mathbb{Q}$.

11.5.4 asserts that the first-order theory of any two real closed fields is the same. Often this result is used to carry over first-order results, already known for the field $\mathbb{R}$, to an arbitrary real closed field. But often, for these sorts of results, it is already pretty clear, from the proof of the result over $\mathbb{R}$, that the same proof carries over to an arbitrary real closed field.

It is important to realize, in all of this, that first-order formulas form a rather restricted class of formulas. Infinite disjunctions and conjunctions are not allowed. Quantification is always understood to be taken over the real closed field $R$ in question. For example, the statement "$x = 1$" is certainly first-order, but the statements "$x$ is an integer" and "there exists an integer $n$ such that $x < n$" are not first-order and, more to the point, they cannot be rephrased in any first-order way.
APPENDIX 2

Algebraic Sets

This appendix contains background material on algebraic sets. The material presented includes basic properties of dimension, see Section 12.1, beginning results about non-singular zeros, see Section 12.2, basic properties of algebraic sets, see Sections 12.3 and 12.4, radical ideals and real ideals, see Section 12.5, the simple point criterion, see 12.6, and the sign-changing criterion, see Section 12.7. Although we are mainly interested in algebraic sets in $R^n$, where $R$ is a real closed field, many of the results hold for an arbitrary field $K$.

12.1 Transcendence Degree and Krull Dimension

Let $K$ be a field. Denote by $K[X]$ the polynomial ring in $n$ variables $X_1, \ldots, X_n$ with coefficients in $K$. For an integral domain $D$, denote by $\text{ff}(D)$, the field of fractions of $D$. For an extension field $L$ of $K$, denote by $\text{trdeg}(L)$, the transcendence degree of $L$ over $K$.

12.1.1 Proposition. Suppose $I$ and $J$ are prime ideals in $K[X]$ and $I \subseteq J$. Then

(1) $\text{trdeg}(\text{ff} \frac{K[X]}{I}) \geq \text{trdeg}(\text{ff} \frac{K[X]}{J})$.

(2) If $\text{trdeg}(\text{ff} \frac{K[X]}{I}) = \text{trdeg}(\text{ff} \frac{K[X]}{J})$, then $I = J$.

Proof. Consider the natural $K$-algebra homomorphism

$$\varphi : \frac{K[X]}{I} \rightarrow \frac{K[X]}{J}.$$ 

Reindexing, if necessary, we may assume that $X_i + J$, $i = 1, \ldots, s$ is a transcendence basis of $\text{ff} \frac{K[X]}{J}$ over $K$. Then one checks easily that $X_i + I$, $i = 1, \ldots, s$ are algebraically independent over $K$. (If $h(X_1 + I, \ldots, X_s + I) = 0$ then, applying $\varphi$, $h(X_1 + J, \ldots, X_s + J) = 0$.) This proves (1). For (2), suppose $f \in J$, $f \notin I$. Then $X_1 + I, \ldots, X_s + I, f + I$ are algebraically independent over $K$. Otherwise, we would have a non-zero polynomial $h(X_1, \ldots, X_s, Y) = \sum_{i=0}^{k} h_i(X_1, \ldots, X_s)Y^i$ such that

$$h(X_1 + I, \ldots, X_s + I, f + I) = \sum_{i=0}^{k} h_i(X_1 + I, \ldots, X_s + I)(f + I)^i = 0.$$ 

Clearly, $k \geq 1$. If $h_0 = 0$, then, using the fact that $K[X]/I$ is an integral domain, we could divide by $f + I$ and reduce $k$. Thus we can assume $h_0 \neq 0$. Applying $\varphi$, we obtain $h_0(X_1 + J, \ldots, X_s + J) = 0$. This is a contradiction. \qed

169
12.1.2 COROLLARY. For any chain of prime ideals \( I = I_0 \supseteq I_1 \supseteq \cdots \supseteq I_k \) in \( K[X] \),
\[
k \leq \text{trdeg}(\frac{K[X]}{I}).
\]

PROOF. It follows from 12.1.1, by induction on \( k \), that
\[
\text{trdeg}(\frac{K[X]}{I}) \geq k + \text{trdeg}(\frac{K[X]}{I_k}).
\]
Since \( \text{trdeg}(\frac{K[X]}{I_k}) \geq 0 \), the result follows. \( \square \)

12.1.3 PROPOSITION. Suppose \( f \in K[X] \) is irreducible. Then
\[
\text{trdeg}(\frac{K[X]}{(f)}) = n - 1.
\]

PROOF. Since \( f \) is not constant, one of the variables \( X_i \) appears in \( f \). Reindexing, we may assume \( X_n \) appears in \( f \). Consider \( D := K[X_1, \ldots, X_{n-1}] \), \( F := \text{ff}(D) \), and \( D[t] \subseteq F[t] \), where \( t := X_n \). \( D \) is a UFD. Since \( f \) has positive degree (as a polynomial in \( t \)), and is irreducible in \( D[t] (= K[X]) \), \( f \) is also irreducible in \( F[t] \). The natural \( K \)-algebra homomorphism \( D[t]/(f) \to F[t]/(f) \) is injective. (If \( g \in D[t] \) and \( f | g \) in \( F[t] \), then \( f | g \) in \( D[t] \).) It follows that \( \text{ff}\frac{D[t]}{(f)} \) is identified with the field \( \frac{F[t]}{(f)} \). Since \( \frac{F[t]}{(f)} \) is an algebraic extension of \( F \) and \( F \) has transcendence degree \( n - 1 \) over \( K \), the result follows. \( \square \)

12.1.4 THEOREM. Let \( I \) be a prime ideal of \( K[X] \).

(1) If \( I \) is a maximal ideal of \( K[X] \), then \( \frac{K[X]}{I} \) is an algebraic extension of \( K \).

(2) If \( J \) is a prime ideal of \( K[X] \), \( I \nsubseteq J \), and there is no prime ideal between \( I \) and \( J \), then \( \text{trdeg}(\text{ff}\frac{K[X]}{I}) - \text{trdeg}(\text{ff}\frac{K[X]}{J}) = 1 \).

PROOF. By Noether’s Normalization Lemma (see [A-M, p. 69] for the proof in case \( K \) is infinite) there is a polynomial ring
\[
K[Y] = K[Y_1, \ldots, Y_m]
\]
which is a subring of \( \frac{K[X]}{I} \) such that \( \frac{K[X]}{I} \) is integral over \( K[Y] \).

(1) Suppose \( I \) is a maximal ideal of \( K[X] \). If \( m \geq 1 \) then there exists an irreducible \( f \in K[Y] \). Then \( (f) \) is a prime ideal of \( K[Y] \). By ‘Going Up’ [A-M, p. 62], there exists a prime ideal of \( \frac{K[X]}{I} \) lying over \( (f) \). Since \( \frac{K[X]}{I} \) is a field, this is impossible. Thus \( m = 0 \), and \( \frac{K[X]}{I} \) is an algebraic extension of \( K \).

(2) Suppose \( I \nsubseteq J \), \( J \) a prime ideal, with no prime ideals between \( I \) and \( J \). Consider the prime ideal \( \frac{I}{J} \cap K[Y] \) of \( K[Y] \). \( \frac{I}{J} \cap K[Y] \neq \{0\} \) by [A-M, Cor. 5.9]. Pick an irreducible polynomial \( f \in \frac{I}{J} \cap K[Y] \). By ‘Going Down’ [A-M, p. 64], there exists a prime ideal of \( \frac{K[X]}{I} \) lying over \( (f) \) and contained in \( \frac{I}{J} \). By hypothesis, this prime ideal is equal to \( \frac{I}{J} \), so \( (f) = \frac{I}{J} \cap K[Y] \). Then \( \frac{K[Y]}{(f)} \hookrightarrow \frac{K[X]}{I} \) and \( \frac{K[X]}{I} \) is integral over \( \frac{K[Y]}{(f)} \), so \( \text{trdeg}(\text{ff}\frac{K[X]}{I}) = \text{trdeg}(\text{ff}\frac{K[Y]}{(f)}) = m - 1 \), using 12.1.3. Since \( \text{trdeg}(\text{ff}\frac{K[X]}{I}) = \text{trdeg}(\text{ff} K[Y]) = m \), the proof is complete. \( \square \)
12.2 Non-Singular Zeros

Assume char(K) = 0. Fix a prime ideal I in K[X]. By the Hilbert Basis Theorem, I is finitely generated, say I = (f_1, \ldots, f_s). Consider the s x n matrix \((\frac{\partial f_i}{\partial X_j} + I)\), viewed as a matrix with entries in the field \(\text{ff} \frac{K[X]}{I}\).

12.2.1 Proposition. Let I = (f_1, \ldots, f_s) be a prime ideal of K[X]. Then

\[ \text{trdeg}(\text{ff} \frac{K[X]}{I}) = n - \text{rank}(\frac{\partial f_i}{\partial X_j} + I). \]

Proof. Let L be a field extension of K. By a derivation on L, we mean a K-linear map \(D: L \rightarrow L\), such that \(D(fg) = fD(g) + gD(f)\), \(\text{Der}(L)\) denotes the set of all derivations of L. \(\text{Der}(L)\) is an L-vector space in a natural way.

We compute the dimension of the vector space \(\text{Der}(L)\), for \(L = \text{ff} \frac{K[X]}{I}\). Since

\[ D(\frac{f}{g}) = \frac{gD(f) - fD(g)}{g^2}, \]

each \(D \in \text{Der}(L)\) is completely determined by its effect on elements of \(\frac{K[X]}{I}\). For \(f \in K[X]\),

\[ D(f + I) = \sum_{j=1}^{n} (\frac{\partial f}{\partial X_j} + I)D(X_j + I), \]

so \(D\) is completely determined by the element \((D(X_1 + I), \ldots, D(X_n + I)) \in L^n\). The only constraints are that \(D(f + I) = 0\) for \(f = \sum_{i=1}^{s} f_ig_i \in I\). Since \(D(f_ig_i + I) = (f_i + I)D(g_i + I) + (g_i + I)D(f_i + I) = (g_i + I)D(f_i + I)\), the constraints are precisely that \(D(f_i + I) = 0\) for \(i = 1, \ldots, s\), i.e., that

\[ \sum_{j=1}^{n} (\frac{\partial f_i}{\partial X_j} + I)D(X_j + I) = 0, \quad i = 1, \ldots, s. \]

Thus, the dimension of \(\text{Der}(L)\) is equal to the dimension of the space of solutions of this linear system, which is \(n - \text{rank}(\frac{\partial f_i}{\partial X_j} + I)\).

We complete the proof by showing that if \(L\) is any finitely generated field extension of \(K\) then the dimension of \(\text{Der}(L)\) is equal to \(\text{trdeg}(L)\). This is clear, by what we have just proved, if \(L\) has a presentation of the form \(\text{ff} \frac{K[Y]}{(f)}\) for some polynomial ring \(K[Y] = K[Y_1, \ldots, Y_m]\) and some irreducible \(f \in K[Y]\). (The 1 x m
matrix \( \frac{\partial f}{\partial y_j} + (f) \) has rank 1 so the dimension of \( \text{Der}(L) \) is equal to \( m-1 \) which, by 12.1.3, is equal to the transcendence degree.) We show that, in fact, this is always the case.

Fix a transcendence basis \( Y_1, \ldots, Y_k \) of \( L \) over \( K \). Let \( R = K[Y_1, \ldots, Y_k] \), \( F = \text{ff}(R) \). \( L \) is an algebraic extension of the field \( F \). Since \( L \) is finitely generated over \( K \), \( L \) is finitely generated over \( F \), so \( [L : F] < \infty \). By the Primitive Element Theorem, there exists \( u \in L \) such that \( L = F(u) \). Let \( f \in F[U] \) be the minimal polynomial of \( u \) over \( F \). Clearing denominators, we can assume \( f \in R[U] \). We can also assume that the GCD of the coefficients of \( f \) is 1. Since \( f \) is irreducible in \( F[U] \), \( f \) is also irreducible in \( R[U] \). Set \( Y_{k+1} = U \) and take \( m = k + 1 \). The field of fractions of \( R[U]/(f) = K[Y_1, \ldots, Y_m]/(f) \) is equal to \( F[U]/(f) = L \).

We say \( p \in K^n \) is a zero of \( I \) if \( f(p) = 0 \) for all \( f \in I \). If this is the case then we have a (well-defined) \( K \)-algebra homomorphism from \( K[X]/I \) to \( K \) defined by \( f + I \mapsto f(p) \), so
\[
\text{rank} \left( \frac{\partial f_i}{\partial x_j} (p) \right) \leq \text{rank} \left( \frac{\partial f_i}{\partial x_j} + I \right).
\]
We say \( p \) is a non-singular zero of \( I \) if \( \text{rank} \left( \frac{\partial f_i}{\partial x_j} (p) \right) = \text{rank} \left( \frac{\partial f_i}{\partial x_j} + I \right) \).

12.2.2. Theorem. If \( p \in K^n \) is a non-singular zero of \( I \), where \( I \) a prime ideal in \( K[X] \), then the completion of \( \frac{K[X]}{I} \) at \( p \) is a formal power series ring \( K[[T_1, \ldots, T_d]] \), where \( d := \text{trdeg}(\text{ff } K[X]/I) \).

By the completion of \( \frac{K[X]}{I} \) at \( p \) we mean the completion of \( \frac{K[X]}{I} \) at the kernel of the homomorphism \( f + I \mapsto f(p) \). See [A-M, Ch. 10] or Section 9.2 for the definition of completion.

Proof. Changing coordinates, i.e., making a change in variables, using the fact that
\[
K[X_1, \ldots, X_n] = K[X_1 - p_1, \ldots, X_n - p_n],
\]
we may assume \( p = (0, \ldots, 0) \). The completion of \( K[X] \) at \( p = (0, \ldots, 0) \) is just the formal power series ring \( K[[X]] := K[[X_1, \ldots, X_n]] \). By the exactness property of completions [A-M, Prop. 10.12], the completion of \( \frac{K[X]}{I} \) at \( (0, \ldots, 0) \) is \( \frac{K[[X]]}{\hat{I}} \), where \( \hat{I} \) denotes the completion of \( I \) at \( (0, \ldots, 0) \). By [A-M, Prop. 10.13], \( \hat{I} \) is the extension of \( I \) to \( K[[X]] \). \( I \) is prime so, by Krull’s Intersection Theorem [A-M, Th. 10.17], the natural map \( \frac{K[X]}{I} \to \frac{K[[X]]}{\hat{I}} \) is injective, i.e., \( \hat{I} \cap K[X] = I \).

Decompose each of the generators \( f_1, \ldots, f_s \) of \( I \) as \( f_i = f_{i0} + f_{i1} + \cdots + f_{id_i} \), where \( f_{ij} \) is homogeneous of degree \( j \). Thus \( f_{i0} = f_i(0, \ldots, 0) = 0 \) and \( f_{i1} = \sum_{j=1}^n \frac{\partial f_i}{\partial X_j}(0, \ldots, 0)X_j \). Using our assumption we see that, after reindexing \( f_1, \ldots, f_s \) suitably, we may assume the linear forms \( f_{i1}, \ldots, f_{n-d,1} \) are linearly independent. Further, reindexing \( X_1, \ldots, X_n \) suitably, we may assume that \( X_1, \ldots, X_d, f_{i1}, \ldots, f_{n-d,1} \) form a basis for the vector space of linear forms. We reparametrize in terms of the new variables \( X_1, \ldots, X_d, f_1, \ldots, f_{n-d} \). By [Z-S, Cor 2., p. 137],
\[
K[[X]] := K[[X_1, \ldots, X_n]] = K[[X_1, \ldots, X_d, f_1, \ldots, f_{n-d}]].
\]
12.3 Algebraic Sets

For a subset $S$ of $K[X]$, define

$$Z(S) = \{ x \in K^n \mid g(x) = 0 \ \forall g \in S \}.$$  

The set $Z(S)$ is called the zero set of $S$.

12.3.1 Proposition. For a subset $V$ of $K^n$, the following are equivalent:

1. $V = Z(S)$, for some finite set $S$ in $K[X]$.
2. $V = Z(S)$, for some set $S$ in $K[X]$.
3. $V = Z(I)$, for some ideal $I$ in $K[X]$.

Proof. Clearly (1) $\Rightarrow$ (2). For (2) $\Rightarrow$ (3), take $I$ to be the ideal in $K[X]$ generated by $S$. For (3) $\Rightarrow$ (1) use the Hilbert Basis Theorem (every ideal in $K[X]$ is finitely generated). 

By an algebraic set in $K^n$, we mean a subset $V$ of $K^n$ satisfying the equivalent conditions of 12.3.1. Given a set $S$ in $K^n$ we can form

$$\mathcal{I}(S) = \{ f \in K[X] \mid f(x) = 0 \ \forall x \in S \}.$$  

This is an ideal of $K[X]$ (easy to check), called the ideal of polynomials vanishing on $S$.

12.3.2 Proposition. If $V$ is an algebraic set in $K^n$, then $Z(\mathcal{I}(V)) = V$.

Proof. The inclusion $V \subseteq Z(\mathcal{I}(V))$ is clear, and holds for any subset $V$ of $K^n$. For the other inclusion, use the fact that $V = Z(S)$ for some subset $S$ of $K[X]$. Then $S \subseteq \mathcal{I}(Z(S)) = \mathcal{I}(V)$, so $V = Z(S) \supseteq Z(\mathcal{I}(V))$.

For $I$ an ideal of $K[X]$, it is not true in general that $\mathcal{I}(Z(I)) = I$. It follows, as a consequence of 12.3.2, that the map $V \mapsto \mathcal{I}(V)$ is a one-to-one map from the set of algebraic sets in $K^n$ into the set of ideals of $K[X]$. $I = \mathcal{I}(V) \Rightarrow Z(I) = Z(\mathcal{I}(V)) = V$. Thus an ideal $I$ is in the image of the map $V \mapsto \mathcal{I}(V)$ iff $I = \mathcal{I}(Z(I))$. More on this later.

It follows from

$$Z(S) \cup Z(T) = Z(ST) \text{ and } \cap_{i \in I} Z(S_i) = Z(\cup_{i \in I} S_i)$$

(these are easy to check), that the set of all algebraic sets in $K^n$ is closed under finite union and arbitrary intersection. The subsets $K^n$ and $\emptyset$ of $K^n$ are algebraic sets, e.g., $K^n = Z(\emptyset)$, $\emptyset = Z(1)$. Consequently, the algebraic sets in $K^n$ are the closed sets of a certain topology on $K^n$. This topology is called the Zariski topology.
Note: If \( K \) is real closed, then \( K^n \) has another topology, namely the Euclidean topology. This is just the product topology, giving \( K \) the interval topology induced by the ordering on \( K \). Every Zariski open set is an open set in the Euclidean topology (because polynomials are continuous in the Euclidean topology), but the converse is not true.

The Zariski topology has some interesting properties:

12.3.3 Proposition. Every descending chain of algebraic sets in \( K^n \) terminates. Equivalently, every non-empty collection of algebraic sets in \( K^n \) has a minimal element.

Proof. If \( V_1 \supseteq V_2 \supseteq \cdots \) is a descending chain of algebraic sets in \( K^n \), then \( I(V_1) \subseteq I(V_2) \subseteq \cdots \) is an ascending chain of ideals in \( K[X] \). Since \( K[X] \) is Noetherian (by the Hilbert Basis Theorem), the latter sequence terminates. Since the map \( V \mapsto I(V) \) is one-to-one, the former sequence also terminates.

An algebraic set \( V \) in \( K^n \) is said to be irreducible if (1) \( V \) is non-empty, and (2) if \( V_1 \) and \( V_2 \) are algebraic sets in \( V \) satisfying \( V = V_1 \cup V_2 \), then \( V = V_1 \) or \( V = V_2 \) (equivalently, if \( U_1, U_2 \) are non-empty Zariski open sets in \( V \) (in the relative topology), then \( U_1 \cap U_2 \neq \emptyset \).

12.3.4 Proposition. Every algebraic set in \( K^n \) decomposes as a finite union of irreducible algebraic sets.

Proof. This is an easy consequence of 12.3.3.

Given an algebraic set \( V \) in \( K^n \), we can decompose \( V \) as \( V = V_1 \cup \cdots \cup V_k \), with \( V_1, \ldots, V_k \) irreducible algebraic sets in \( K^n \). Refining this decomposition, we can assume the decomposition is irredundant, i.e., \( V_i \not\subseteq V_j \) for \( i \neq j \). \( V_1, \ldots, V_k \) are called the irreducible components of \( V \).

12.3.5 Proposition. The irreducible components of an algebraic set \( V \) in \( K^n \) are well-defined: If \( V = V_1 \cup \cdots \cup V_k \) and \( V = W_1 \cup \cdots \cup W_\ell \) are two irredundant decompositions of \( V \) into irreducible algebraic sets, then \( \ell = k \) and \( W_1, \ldots, W_\ell \) is some permutation of \( V_1, \ldots, V_k \).

Proof. \( V_i = V_i \cap V = V_i \cap (\cup_j W_j) = \cup_j (V_i \cap W_j) \). Since \( V_i \) is irreducible, this implies \( V_i = V_i \cap W_j \), i.e., \( V_i \subseteq W_j \), for some \( j \). Reversing the roles of \( V_i \) and \( W_j \) yields \( W_j \subseteq V_i \) for some \( i' \). Then \( V_i \subseteq V_{i'} \). Since the decomposition \( V = V_1 \cup \cdots \cup V_k \) is irredundant, this implies \( i' = i \) and \( V_i = W_j \). Since the decomposition \( V = W_1 \cup \cdots \cup W_\ell \) is irredundant, the index \( j \) is unique. Similarly, for each index \( j \) there exists a unique index \( i \) such that \( W_j = V_i \).

Given an algebraic set \( V \) in \( K^n \), \( \mathcal{P}(V) \) denotes the the ring (more precisely, the \( K \)-algebra) of all polynomial functions \( f : V \rightarrow K \). \( \mathcal{P}(V) \) is naturally identified with the \( K \)-algebra \( K[X]/I(V) \), by identifying \( f + I(V) \) with the polynomial function \( x \mapsto f(x) \) on \( V \). Since \( X_1, \ldots, X_n \) generate \( K[X] \) as a \( K \)-algebra, and

\[
  f + I(V) = f(X_1 + I(V), \ldots, X_n + I(V)),
\]

the elements \( X_i + I(V) \), \( i = 1, \ldots, n \), generate \( K[X]/I(V) \) as a \( K \)-algebra. Since \( X_i + I(V) \) is identified with the polynomial function \( x \mapsto x_i \) (projection onto the \( i \)-coordinate), the ring \( \mathcal{P}(V) \) is sometimes called the coordinate ring of the algebraic set \( V \).
12.3.6 **Proposition.** For an algebraic set \( V \) in \( K^n \), the following are equivalent:

1. \( V \) is irreducible.
2. \( \mathcal{I}(V) \) is a prime ideal in \( K[X] \).
3. \( \mathcal{P}(V) \) is an integral domain.

**Proof.** Since \( \mathcal{P}(V) = K[X]/\mathcal{I}(V) \), (2) \( \iff \) (3) is clear. (1) \( \Rightarrow \) (2): Suppose \( fg \in \mathcal{I}(V) \). Then \( \mathcal{Z}(f) \cup \mathcal{Z}(g) = \mathcal{Z}(fg) \supseteq V \), so \( V = (V \cap \mathcal{Z}(f)) \cup (V \cap \mathcal{Z}(g)) \).

Since \( V \) is irreducible, this implies \( V = V \cap \mathcal{Z}(f) \) (so \( f \in \mathcal{I}(V) \)) or \( V = V \cap \mathcal{Z}(g) \) (so \( g \in \mathcal{I}(V) \)). (2) \( \Rightarrow \) (1): Suppose \( V = V_1 \cup V_2 \). Then, \( f_1 \in \mathcal{I}(V_1) \) and \( f_2 \in \mathcal{I}(V_2) \) \( \Rightarrow \) \( f_1 f_2 = 0 \) on \( V \), i.e., \( f_1 f_2 \in \mathcal{I}(V) \). This proves \( \mathcal{I}(V_1) \mathcal{I}(V_2) \subseteq \mathcal{I}(V) \). Since \( \mathcal{I}(V) \) is prime, this implies \( \mathcal{I}(V_1) \subseteq \mathcal{I}(V) \), i.e., that \( V_i \supseteq V \), for some \( i \in \{1, 2\} \). \( \square \)

12.3.7 **Proposition.** If \( V \) is an algebraic set with irreducible components \( V_1, \ldots, V_k \) then \( \mathcal{I}(V_1), \ldots, \mathcal{I}(V_k) \) are the minimal prime ideals lying over \( \mathcal{I}(V) \).

**Proof.** Let \( I \) be a prime ideal in \( K[X] \) lying over \( \mathcal{I}(V) \). Then

\[
\prod_{i=1}^{k} \mathcal{I}(V_i) \subseteq \bigcap_{i=1}^{k} \mathcal{I}(V_i) = \mathcal{I}(V) \subseteq I.
\]

Since \( I \) is prime, this implies \( \mathcal{I}(V_i) \subseteq I \) for some \( i \). This proves that every minimal prime ideal lying over \( \mathcal{I}(V) \) belongs to the set \( \{ \mathcal{I}(V_1), \ldots, \mathcal{I}(V_k) \} \). It remains to check that there are no inclusion relations between the \( \mathcal{I}(V_i) \). But this is clear: If \( \mathcal{I}(V_i) \subseteq \mathcal{I}(V_j) \) for some \( i \neq j \) then, by 12.3.2, \( V_i = \mathcal{Z}(\mathcal{I}(V_i)) \supseteq \mathcal{Z}(\mathcal{I}(V_j)) = V_j \), a contradiction. \( \square \)

If \( V \) is an irreducible algebraic set in \( K^n \), the field of fractions of the domain \( \mathcal{P}(V) \) is called the function field of \( V \). Observe: Elements of the function field of \( V \) are not really functions on \( V \). If \( f \) is an element of the function field of \( V \) and \( x \in V \), then \( f(x) \) is only defined if \( f \) has a presentation \( f = \frac{g}{h} \), with \( g, h \in \mathcal{P}(V) \), \( h(x) \neq 0 \). The set of \( x \in V \) where \( f(x) \) is defined is a (non-empty) Zariski open set in \( V \).

12.3.8 **Proposition.** Assume \( \text{char}(K) = 0 \). Let \( V \subseteq K^n \) be an irreducible algebraic set with associated prime ideal \( I := \mathcal{I}(V) \). Then the set of non-singular zeros of \( I \) is non-empty and Zariski open in \( V \).

**Proof.** The set of non-singular zeros of \( I \) is the set of all \( p \in V \) satisfying \( \Delta(p) \neq 0 \) for some \( k \times k \) minor \( \Delta \) of the matrix \( \left( \frac{\partial f}{\partial X_j} \right) \), where \( k := \text{rank}(\frac{\partial f}{\partial X_j} + I) \).

Since \( \Delta \in K[X] \), the set of non-singular zeros of \( I \) is Zariski open in \( V \). Since \( \Delta \notin I \) for some such minor, the existence of a non-singular zero is clear. \( \square \)

### 12.4 Dimension

The dimension of an algebraic set \( V \) is defined to be the Krull dimension of the coordinate ring \( \mathcal{P}(V) \). If \( V \) has irreducible components \( V_1, \ldots, V_k \) then

\[
\dim(V) = \max \{ \dim(V_i) \mid i = 1, \ldots, k \},
\]

by 12.3.7. If \( V \) is irreducible then

\[
\dim(V) = \text{trdeg}(\mathcal{P}(V)),
\]
12.4.1 Proposition. Suppose $V,W$ are irreducible algebraic sets with $V \subseteq W$. Then

1. $\dim(V) \leq \dim(W)$.
2. If $\dim(V) = \dim(W)$, then $V = W$.

Proof. Apply 12.1.1 to the prime ideals $\mathcal{I}(W) \subseteq \mathcal{I}(V)$ in $K[X]$. $\square$

12.4.2 Proposition. Let $I$ be a prime ideal of $K[X]$. Then

1. $\dim(Z(I)) \leq \text{trdeg}(K[X]/I)$.
2. $\dim(Z(I)) = \text{trdeg}(K[X]/I)$ iff $I = \mathcal{I}(Z(I))$.

Proof. Let $V$ be an irreducible component of $Z(I)$ chosen so that $\dim(V) = \dim(Z(I))$. Then $\mathcal{I}(V) \supseteq \mathcal{I}(Z(I)) \supseteq I$, so (1) holds, using 12.1.1(1). (2) ($\Rightarrow$): By 12.1.1 (2), $\mathcal{I}(V) = I$, so $I = \mathcal{I}(Z(I))$. (2) ($\Leftarrow$): This is clear, by the definition of dimension. $\square$

12.4.3 Proposition. Let $V \subseteq K^n$, $W \subseteq K^m$ be irreducible algebraic sets. Then

1. $V \times W$ is an irreducible algebraic set in $K^n \times K^m = K^{n+m}$.
2. $\dim(V \times W) = \dim(V) + \dim(W)$.

Proof. (1) Denote by $K[X,Y]$ the polynomial ring in $n+m$ variables $X_1, \ldots, X_n, Y_1, \ldots, Y_m$. Writing $V = Z(S)$, $W = Z(T)$, $S \in K[X]$, $T \subseteq K[Y]$, it is clear that $V \times W = Z(S \cup T)$, viewing $S \cup T$ as a subset of $K[X,Y]$. Suppose $f_1, f_2 \in K[X,Y]$, $f_1f_2$ vanishes on $V \times W$. For each $x \in V$, $f_1(x,Y)f_2(x,Y)$ vanishes on $W$, so $f_1(x,Y)$ vanishes on $W$ or $f_2(x,Y)$ vanishes on $W$. Set $V_i = \{x \in V \mid f_i(x,Y) \text{ vanishes on } W\} = V \cap Z(f_i(x,y) \mid y \in W)$. Then $V_1 \cup V_2 = V$, so $V = V_i$, i.e., $f_i$ vanishes of $V \times W$, for some $i \in \{1,2\}$.

(2) It suffices to prove that $X_i + I(V \times W), \ldots, X_i + I(V \times W), Y_j + I(V \times W), \ldots, Y_j + I(V \times W)$ are algebraically independent elements of $\mathcal{P}(V \times W)$ iff $X_i + I(V), \ldots, X_i + I(V)$ are algebraically independent elements of $\mathcal{P}(V)$ and $Y_j + I(W), \ldots, Y_j + I(W)$ are algebraically independent elements of $\mathcal{P}(W)$. The implication ($\Rightarrow$) is clear. For the proof of ($\Leftarrow$), suppose $f$ is some polynomial in $X_{i_1}, \ldots, X_{i_k}, Y_{j_1}, \ldots, Y_{j_l}$ vanishing on $V \times W$. Write $f = f_\alpha$ where $f_\alpha$ is a polynomial in $X_{i_1}, \ldots, X_{i_k}$, and $Y^\alpha$ is a monomial in $Y_{j_1}, \ldots, Y_{j_l}$. For each $x \in V$, the polynomial $\sum_\alpha f_\alpha(x)Y^\alpha$ vanishes on $W$. Since $V \times W = I(V \times W), \ldots, I(W)$ are algebraically independent, this implies each $f_\alpha(x) = 0$. Thus each $f_\alpha$ vanishes on $V$. Since $X_{i_1} + I(V), \ldots, X_{i_k} + I(V)$ are algebraically independent, this implies each $f_\alpha = 0$, so $f = 0$. $\square$

The argument in the proof of 12.4.3 (2) shows that if $\sum_\alpha f_\alpha Y^\alpha$ vanishes on $V \times W$, where the $f_\alpha$ are polynomials in $X_1, \ldots, X_n$, and the $Y^\alpha$ are monomials in $Y_1, \ldots, Y_m$ such that the corresponding elements $Y^\alpha + I(W) \in \mathcal{P}(W)$ are linearly independent, then each $f_\alpha$ vanishes on $V$. This implies $\mathcal{P}(V \times W) = \mathcal{P}(V) \otimes_K \mathcal{P}(W)$.

Note: If the field $K$ is real closed, then there is also a geometric description of dimension. (For example, $Y^2 = X^3 - X^2$ and $XY = 0$ both have dimension 1—just look at the graphs.) This is explained in [B-C-R, Sect. 2.8]. The material presented here provides the prerequisite material on algebraic sets needed to read [B-C-R, Sect. 2.8].
12.5 Radical Ideals and Real Ideals

It is important to understand the image of the map \( V \mapsto \mathcal{I}(V) \), from the set of algebraic sets in \( K^n \), to the set of ideals of \( K[X] \). This is a non-trivial problem. It is equivalent to the problem of determining the ideals \( I \) in \( K[X] \) which satisfy \( I = \mathcal{I}(\mathbb{Z}(I)) \). When \( K \) is algebraically closed (resp., real closed), the answer is known. Namely, the answer is the Nullstellensatz, due to Hilbert (resp., the Real Nullstellensatz, due to Krivine, Dubois and Risler).

We need some terminology. Let \( A \) be any commutative ring with 1. For an ideal \( I \) of \( A \) define

\[
\sqrt{I} = \{ a \in A \mid \exists \text{ an integer } m \geq 1 \text{ such that } a^m \in I \}.
\]

\[
\sqrt{I} = \{ a \in A \mid \exists \text{ an integer } m \geq 1 \text{ and } b_1, \ldots, b_k \in A \text{ such that } a^{2m} + b_1^2 + \cdots + b_k^2 \in I \}.
\]

\( \sqrt{I} \) is called the radical of \( I \); \( \sqrt{I} \) is called the real radical of \( I \).

12.5.1 Proposition. For any ideal \( I \) of \( A \),

(1) \( I = \sqrt{I} \) iff \( \forall a \in A, a^2 \in I \Rightarrow a \in I \).

(2) \( I = \sqrt{I} \) iff \( \forall a_1, \ldots, a_k \in A, a_1^2 + \cdots + a_k^2 \in I \Rightarrow a_1 \in I \).

Proof. (1) \((\Rightarrow)\) is clear. To prove \((\Leftarrow)\), suppose \( a^m \in I \) for some integer \( m \geq 1 \). Pick \( k \) so that \( 2^k \geq m \). Then \( a^{2^k} = a^m a^{2^k-m} \in I \). By induction on \( k \), this implies, in turn, that \( a \in I \).

(2) \((\Rightarrow)\) is clear. To prove \((\Leftarrow)\), suppose \( a^2 + b_1^2 + \cdots + b_k^2 \in I \). Then \( a^m \in I \). As above, this implies \( a^{2k} \in I \), if \( 2^k \geq m \). By induction, this implies, in turn, that \( a \in I \).

If the equivalent conditions of 12.5.1(1) hold, we say the ideal \( I \) is radical. If the equivalent conditions of 12.5.1(2) hold, we say the ideal \( I \) is real.

If an ideal \( I \) is real, then \( I \) is radical. A prime ideal \( I \) is automatically radical, but it need not be real.

12.5.2 Proposition. For a prime ideal \( I \) of \( A \), the following are equivalent:

(1) \( I \) is real

(2) The field of fractions of the integral domain \( A/I \) is formally real.

Proof. Denote the coset \( a + I \) by \( \bar{a} \). Assume \( I \) is not real, so there exist \( a, a_1, \ldots, a_k \in A \) with \( a^2 + a_1^2 + \cdots + a_k^2 \in I \), \( a \notin I \). Then \( \bar{a}^2 + \bar{a}_1^2 + \cdots + \bar{a}_k^2 = 0 \) and \( \bar{a} \neq 0 \), i.e., \(-1 = (\frac{a}{\bar{a}}) \bar{a})^2 + \cdots + (\frac{a}{\bar{a}}) \bar{a}_k)^2 \), so \( \text{ff}(A/I) \) is not formally real. Conversely, by reversing this argument, if \( \text{ff}(A/I) \) is not formally real, then \( I \) is not real.

12.5.3 Theorem. Let \( I \) be an ideal of \( K[X] \). Then:

(1) (Nullstellensatz) If \( K \) is algebraically closed, then \( \mathcal{I}(\mathbb{Z}(I)) = \sqrt{I} \).

(2) (Real Nullstellensatz) If \( K \) is real closed, then \( \mathcal{I}(\mathbb{Z}(I)) = \sqrt{I} \).

Proof. (1) can be deduced from 12.1.4 (1) by the argument outlined in [A-M, p. 85]. See Chapter 2, Section 2 for the proof of (2).
12.5.4 COROLLARY. Let \( I \) be an ideal in \( K[X] \). Then:

(1) For \( K \) algebraically closed, \( I \) is in the image of the map \( V \mapsto \mathcal{I}(V) \) iff \( I \) is radical.

(2) For \( K \) real closed, \( I \) is in the image of the map \( V \mapsto \mathcal{I}(V) \) iff \( I \) is real.

PROOF. Recall that an ideal \( I \) in \( K[X] \) is in the image of the map \( V \mapsto \mathcal{I}(V) \) iff \( I = \mathcal{I}(\mathbb{Z}(I)) \). In view of this, the result is immediate from 12.5.3.

\[ \square \]

12.6 Simple Point Criterion

In view of 12.5.4 (2), it is important to be able to recognize real primes.

12.6.1 THEOREM. Suppose \( I \) is a prime ideal of \( K[X] \), \( K \) real closed. The following are equivalent:

(1) \( I \) is real.

(2) \( I = \mathcal{I}(\mathbb{Z}(I)) \).

(3) \( \dim(\mathbb{Z}(I)) = \text{trdeg } K[X] \).

(4) \( I \) has a non-singular zero \( p \) in \( K^n \).

PROOF. (1) \( \leftrightarrow \) (2) \( \leftrightarrow \) (3) by 12.4.2(2) and 12.5.4(2). (2) \( \Rightarrow \) (4) by 12.3.8. (4) \( \Rightarrow \) (1): Picking a non-singular zero \( p \) of \( I \) we have an embedding \( K[X] \hookrightarrow K[[T_1, \ldots, T_d]] \) by 12.2.2. To show \( I \) is real, it suffices to show that \( K((T_1, \ldots, T_d)) \) (the field of fractions of \( K[[T_1, \ldots, T_d]] \)) has at least one ordering. But this is more or less clear. Recall that \( K[[T_1, \ldots, T_d]] \) is a UFD. Localizing \( K[[T_1, \ldots, T_d]] \) at the prime ideal generated by the irreducible \( T_d \) yields a discrete valuation ring of \( K((T_1, \ldots, Y_d)) \) with residue field \( K((T_1, \ldots, T_{d-1})) \). By induction on \( d \), the latter has an ordering. Pulling this back, via Baer- Krull, it follows that \( K((T_1, \ldots, T_d)) \) also has an ordering.

Non-singular zeros are called simple points. Condition (4) of 12.6.1 is called the Simple Point Criterion. For \( I = \mathcal{I}(V) \), where \( V \subseteq K^n \) is an irreducible algebraic set, a non-singular zero of \( I \) is called a non-singular point of \( V \).

Caution: For \( K \) algebraically closed, it not necessary to distinguish between non-singular points and non-singular zeros. For \( K \) real closed, the connection is more subtle — a prime ideal need not be real — e.g., \( (0,0) \) is a non-singular point of the irreducible algebraic set \( Z(X^2 + Y^2) = \{(0,0)\} \), but \( (0,0) \) is not a non-singular zero of the prime ideal \( (X^2 + Y^2) \).

12.7 Sign-Changing Criterion

An irreducible algebraic set \( V \) in \( K^n \) is called a hypersurface in \( K^n \) if \( \mathcal{I}(V) = (f) \) (the principal ideal generated by \( f \)) for some (necessarily irreducible) \( f \in K[X] \).

12.7.1 THEOREM. Let \( K \) be a real closed field, and suppose \( f \in K[X] \) is irreducible. The following are equivalent:

(1) \( (f) \) is real.

(2) \( (f) = \mathcal{I}(Z(f)) \).

(3) \( \dim(Z(f)) = n - 1 \).

(4) The polynomial \( f \) has a non-singular zero in \( K^n \) (i.e., there is an \( x \in K^n \) such that \( f(x) = 0 \) and \( \frac{\partial f}{\partial x_i}(x) \neq 0 \) for some \( i \in \{1, \ldots, n\} \)).
(5) The polynomial \( f \) changes sign on \( K^n \) (i.e., there exist \( x, y \in K^n \) such that \( f(x)f(y) < 0 \)).

**Proof.** \((1) \iff (2) \iff (3) \iff (4)\) by 12.6.1.

(4) \( \implies (5) \). Let \( x = (x_1, \ldots, x_n) \). By hypothesis, the polynomial function \( \varphi : K \to K \) defined by \( \varphi(t) = f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n) \) satisfies \( \varphi(x_i) = 0 \) and \( \varphi'(x_i) \neq 0 \). Thus there exist \( t_1 < x_i < t_2 \) such that \( \varphi(t_1)\rho(t_2) < 0 \), i.e.,

\[
f(x_1, \ldots, x_{i-1}, t_1, x_{i+1}, \ldots, x_n)f(x_1, \ldots, x_{i-1}, t_2, x_{i+1}, \ldots, x_n) < 0.
\]

(5) \( \implies (2) \): Assume \( g \in \mathcal{I}(\mathcal{Z}(f)) \) (i.e., \( \mathcal{Z}(f) \subseteq \mathcal{Z}(g) \)). We want to show \( g \in (f) \).

Making an affine change of coordinates, we can assume

\[
f((a, b_1) < 0 < f(a, b_2)),
\]

where \( a \in K^{n-1} \), and \( b_1, b_2 \in K \). Let \( R = K[X_1, \ldots, X_{n-1}] \) and \( F = ff(R) \). View \( f \) and \( g \) as polynomials in \( t = X_n \) in the ring \( R[t] \subseteq F[t] \). Assume that \( f \nmid g \) in \( R[t] \) (= \( K[X] \)). From the theory of UFDs, we know that \( f \) remains irreducible in \( F[t] \) and \( f \nmid g \) also holds in \( F[t] \). Since \( F[t] \) is a PID, there exists an equation \( pf + qg = 1 \) with \( p, q \in F[t] \). Write \( p = \frac{r}{h} \) and \( q = \frac{s}{h} \), where \( r, s \in R[t], 0 \neq h \in R \). Then \( fr + gs = h \). Choose a Euclidean neighbourhood \( V \) of \( a \) in \( K^{n-1} \) such that \( f(V, b_1) < 0 < f(V, b_2) \). For any \( v \in V, f(v, b_1) < 0 < f(v, b_2) \) implies that \( f(v, b_0) = 0 \) for some \( b_0 \) between \( b_1 \) and \( b_2 \), by the Intermediate Value Theorem. By the hypothesis on \( g \), we also have \( g(v, b_0) = 0 \). Therefore, \( fr + gs = h \) implies that \( h(v) = 0 \) and so \( h(X_1, \ldots, X_{n-1}) \) vanishes on a nonempty open set in \( K^{n-1} \).

This forces \( h = 0 \), by 1.1.2, a contradiction. \( \Box \)

Condition (5) of 12.7.1 is called the **Sign-Changing Criterion**.

### 12.7.2 Examples.

(1) \( X^2 + Y^2 = 0 \) does not define a curve in \( \mathbb{R}^2 \): \( X^2 + Y^2 \) is irreducible in \( \mathbb{R}[X, Y] \), but \( \mathcal{I}(\mathcal{Z}(X^2 + Y^2)) = (X, Y) \), and \( (X^2 + Y^2) \not\subset (X, Y) \). Alternatively, just note that \( X^2 + Y^2 \) does not change sign on \( \mathbb{R}^2 \).

(2) \( Y = X^2 \) defines a curve in \( \mathbb{R}^2 \): \( Y - X^2 \) is irreducible in \( \mathbb{R}[X, Y] \) and obviously changes sign on \( \mathbb{R}^2 \). All points on \( Y = X^2 \) are non-singular.

(3) \( Y^2 = X^3 \) defines a curve in \( \mathbb{R}^2 \): \( Y^2 - X^3 \) is irreducible in \( \mathbb{R}[X, Y] \) and obviously changes sign on \( \mathbb{R}^2 \). The point \((0, 0)\) is a singular point on this curve. (In fact, it is the only singular point.) One can ‘see’ this singular point on the graph. (It is a cusp.)

(4) The point \((0, 0)\) is a singular point on the curve \( Y^3 + 2X^2Y - X^4 = 0 \). On the other hand, it is hard to ‘see’ this singular point on the graph. See Figure 10 or [B-C-R, Ex. 3.3.11].

(5) \( X^2Z = Y^2 \) defines a surface in \( \mathbb{R}^3 \). \( X^2Z - Y^2 \) is irreducible in \( \mathbb{R}[X, Y, Z] \) and one checks easily that it does indeed change sign on \( \mathbb{R}^3 \). See Figure 11 for the graph of this surface. The set of singular points is the \( Z \)-axis. This is Whitney’s Umbrella.

(6) \( X^2(1 - X^2 - Z^2) = Y^2 \) defines a surface in \( \mathbb{R}^3 \): \( Y^2 - X^2(1 - X^2 - Z^2) \) is irreducible in \( \mathbb{R}[X, Y, Z] \) and one checks easily that it does indeed change sign on \( \mathbb{R}^3 \). See Figure 12 for the graph of this surface. The set of singular points is the \( Z \)-axis. The set of non-singular points is bounded. This example appears in [Sw2, Ex. 3.12]. See [B-C-R, Ex. 3.1.2(f)] for a similar example.
**Figure 10.** \( Y^3 + 2X^2Y - X^4 = 0 \)

**Figure 11.** \( X^2Z = Y^2 \)
Figure 12. $X^2(1 - X^2 - Z^2) = Y^2$
Bibliography


183


[S3] ——, Sums of squares on real algebraic curves, Math. Z. 245 (2003), 725–760.


[Sw5] ——, Ideas concerning the maximum cut problem, unpublished notes.


The study of positive polynomials brings together algebra, geometry and analysis. The subject is of fundamental importance in real algebraic geometry when studying the properties of objects defined by polynomial inequalities. Hilbert’s 17th problem and its solution in the first half of the 20th century were landmarks in the early days of the subject. More recently, new connections to the moment problem and to polynomial optimization have been discovered. The moment problem relates linear maps on the multidimensional polynomial ring to positive Borel measures.

This book provides an elementary introduction to positive polynomials and sums of squares, the relationship to the moment problem, and the application to polynomial optimization. The focus is on the exciting new developments that have taken place in the last 15 years, arising out of Schmüdgen’s solution to the moment problem in the compact case in 1991. The book is accessible to a well-motivated student at the beginning graduate level. The objects being dealt with are concrete and down-to-earth, namely polynomials in \( n \) variables with real coefficients, and many examples are included. Proofs are presented as clearly and as simply as possible. Various new, simpler proofs appear in the book for the first time. Abstraction is employed only when it serves a useful purpose, but, at the same time, enough abstraction is included to allow the reader easy access to the literature. The book should be essential reading for any beginning student in the area.