ROCKY MOUNTAIN JOURNAL OF MATHEMATICS Volume 14, Number 4, Fall 1984

AN INTRODUCTION TO REAL ALGEBRA

T. Y. LAM

Dedicated to the memory of Gus Efroymson

Introduction. Real Algebra, roughly speaking, is the study of "real" objects such as real rings, real places and real varieties. Becuase of the recent interest in developing algebraic geometry over the real numbers (or, more generally, over a real closed field), the algebraic study of these real objects has attracted considerable attention. In many ways, the role played by real algebra in the development of real algebraic geometry is analogous to the role played by commutative algebra in the development of classical algebraic geometry. Therefore, real algebra, like commutative algebra, is a subject with a great potential for applications to geometric problems.

In the category of fields, the real objects (namely, the formally real fields) have been studied already a long time ago by Artin and Schreier, who recognized that formally real fields are precisely the fields which can be ordered. The idea of exploiting the orderings in a real field, for instance, was central in Artin's solution to Hilbert's 17th Problem. By the 70's, there was already a sizable literature on formally real (or ordered) fields. However, real algebra in the category of rings underwent a much slower development. There are several possible notions of reality for rings and people weren't sure which one to adopt. Similarly, it was not at all clear how one should define for rings the notion of orderings. Fortunately, with the impetus given by real algebraic geometry, these problems have recently been successfully resolved. There is now a consensus (or almost a consensus) about what an ordering on a ring should be, and, with this notion of orderings, there is a remarkably complete analogue of the Artin-Schreier theory valid for rings. Further, Coste and Coste-Roy have introduced the notion of the real spectrum of a ring. On the one hand, this is the correct generalization of the space of orderings of a field, and on the other hand, this offers a "real" analogue of the Zariski prime spectrum of a ring. With the discovery of this notion, one seems to be now fully ready for a

This report was done with support of an NSF grant at the Center for Pure and Applied Mathematics at U.C. Berkeley.

systematic study of real algebra over rings, as well as its applications to real algebraic geometry.

In these notes, we shall give a self-contained introduction to these recent developments in real algebra. We shall not try to go deeply into any single topic, nor shall we try to formulate results in their greatest possible degree of generality. Instead, we shall focus on what we perceive to be the central facts, and try to organize them, with some care, into a coherent picture. By aiming our exposition at the nonexperts, it is hoped that this work will make the techniques of real algebra accessible to a general audience. For the experts, my exposition has probably little to offer, except perhaps by way of terminology. Since the subject of real algebra over rings is relatively new, the terminology used so far in the literature seemed to be in a great disarray. Needless to say, the lack of a consistent terminology in a subject is not in the best interest of its development. In view of this, I take the opportunity here to offer a remedy. In these notes, we shall call a ring A semireal if -1 is not a sum of squares in A; we shall call A real (or formally real) if $\sum a_i^2 = 0 \Rightarrow all a_i = 0$ in A. If (A, M) is a local ring, we shall call A residually real if the residue field A/\mathfrak{M} is real. By an ordering on a ring A, we mean (essentially) an ordering on the quotient field of A/p for some prime ideal $p \subset A$. For us, this system of terminology works very well throughout these notes. For better or for worse, we propose it for possible use by the mathematical community.

Since this work is intended to be an exposition and not a broad survey, we have left out the discussion of many important topics. To get a broader view of the subject of real algebra, we urge our readers to consult also the other available survey/expository articles on the subject, for instance [4], [18] and [19]. Readers desiring more information about the recent literature should consult the two excellent collections of articles [16] and [25] listed in the references.

The work presented here is a revised and somewhat expanded version of my lecture notes circulated at the Sexta Escuela Latino Americana de Matematicas in Oaxtepec, Mexico, July, 1982. I thank Professor José Adem, organizer of ELAM VI, for his kind permission for me to use the same material for the Real Algebraic Geometry Conference. During the preparation of these notes, I have had the great benefit of frequent consultation with J. Merzel and A. Prestel. Both of them have generously given their time in helping me out when I got stuck in the writing. Their valuable suggestions have resulted in many substantial improvements in our exposition. E. Becker and G. Brumfiel have read and made many insightful comments on the earlier version of these notes which greatly helped me understand what I had written, and prompted the revisions incorporated in the present version. M. Coste and M.-F. Coste-Roy made useful comments on the terminology of this work and pointed out to us some pertinent references. To all of them, my sincere thanks.

Finally I would like to remark that the untimely death of Gus Efroymson shortly after the Boulder Conference has saddened us all and has been a great loss to workers in real algebraic geometry.

1. The level of a ring. In this section, we shall collect some elementary results about the level of (commutative) rings. Of course, for "Real Algebra", we shall be primarily interested in rings of infinite level, but for the convenience of this section we shall try to state some of the results in terms of the finiteness of the level. For later sections, these results can simply be applied in their contrapositive forms.

First, we recall the definition of the level (Stufe) of a ring.

DEFINITION 1.1. If -1 is a sum of squares in a ring A, we define the level of A (denoted by s(A)) to be the smallest natural number s such that -1 can be expressed as a sum of s squares in A. If -1 is not a sum of squares in A, we define s(A) to be ∞ .

If F is a field, Pfister has shown that s(F) is either ∞ , or else a power of 2, and that all powers of 2 are possible (cf. [20, Ch. 11]). On the other hand, if A is a commutative ring, then s(A) can be ∞ or any given natural number n; in fact, Dai, Lam and Peng have shown that the **R**-algebra

$$\mathbf{R}[x_1, \ldots, x_n]/(1 + x_1^2 + \cdots + x_n^2)$$

has level = n. The proof is a simple application of the Borsuk-Ulam Theorem; cf. [12].

For any commutative ring A, we shall write $\sum A^2$ for the set of sums of squares in A. If the ring A is understood, we shall simply write \sum (or \sum_A) for $\sum A^2$. This notation will be used freely throughout these notes.

Our first result is the following basic lemma.

LEMMA 1.2. Let A be a commutative ring, and let $\Sigma = \Sigma A^2$. Let $S \subset A$ be a multiplicative set containing 1 but not containing 0, such that $S + \Sigma \subset S$. Let \mathfrak{p} be an ideal of A maximal with respect to the property that \mathfrak{p} is disjoint from S. Then \mathfrak{p} is a prime ideal, and $qf(A/\mathfrak{p})$ (the quotient field of A/\mathfrak{p}) has infinite level.

REMARK 1.3. Assume that $s(A) = \infty$. Then the lemma can be applied to the set $S: = 1 + \Sigma$. We clearly have $1 \in S$ and $0 \notin S$, and it is easy to check that S is a multiplicative set, with $S + \Sigma \subset S$.

PROOF (of 1.2). The fact that p is a prime ideal is well-known from commutative algebra. (For this, we do not need $S + \sum \subset S$.) Thus it only remains to show that $s(qf(A/p)) = \infty$. Assume, instead, that $s(qf(A/p)) < \infty$. Then there exists a relation $b_1^2 + b_2^2 + \cdots + b_n^2 \in p$, where $b_i \in A$ $(1 \le i \le n)$, and, say, $b_1 \notin p$. By the maximality of p, $p + (b_1)$ must meet S, so we have $s \equiv rb_1 \pmod{p}$ for some $s \in S$. But then $s^2 \equiv r^2b_1^2 \pmod{p}$ and so

(1.4) $s^2 + r^2 b_2^2 + \cdots + r^2 b_n^2 \equiv r^2 (b_1^2 + \cdots + b_n^2) \equiv 0 \pmod{\mathfrak{p}}.$

However, $s^2 + r^2 b_2^2 + \cdots + r^2 b_n^2 \in S + \sum \subset S$, so (1.4) contradicts the fact that $S \cap \mathfrak{p} = \emptyset$.

The following result has been observed by many authors, including Coste and Coste-Roy, and Bröcker-Dress-Scharlau (cf. [7]).

THEOREM 1.5. For any commutative ring A, $s(A) < \infty$ if and only if for all prime ideals $\mathfrak{p} \subset A$, $s(qf(A/\mathfrak{p})) < \infty$.

PROOF. The "only if" part is trivial. The "if" part follows immediately from the Lemma above (in view of Remark 1.3).

COROLLARY 1.6. (Local-Global Criterion for Finite Level) For any commutative ring A, $s(A) < \infty$ if and only if for all maximal ideals $\mathfrak{M} \subset A$, $s(A_{\mathfrak{M}}) < \infty$.

PROOF. Again, we need only prove the "if" part. Assume that $s(A) = \infty$. Let $\mathfrak{p} \subset A$ be a prime ideal such that $s(qf(A/\mathfrak{p})) = \infty$. (Such a prime exists by 1.2.) Since $qf(A/\mathfrak{p}) \cong A_\mathfrak{p}/\mathfrak{p}A_\mathfrak{p}$, it follows that $s(A_\mathfrak{p}) = \infty$. Let \mathfrak{M} be any maximal ideal containing \mathscr{Y} . Then we have a homomorphism $A_\mathfrak{M} \to A_\mathfrak{p}$, so $s(A_\mathfrak{p}) = \infty$ implies that $s(A_\mathfrak{M}) = \infty$.

Let A be an integral domain, with quotient field F. Of course, $s(A) < \infty \Rightarrow s(F) < \infty$, but the converse is in general not true. An easy counterexample is given by the **R**-algebra

$$A = \mathbf{R}[x_1, \ldots, x_n]/(x_1^2 + \cdots + x_n^2).$$

In qf(A), -1 is a sum of n-1 squares so $s(F) \leq n-1$, but since A admits a homomorphism into **R**, we have $s(A) = \infty$. If we want a local example, we can replace A by its localization at the maximal ideal generated by $\bar{x}_1, \ldots, \bar{x}_n$.

There are several important classes of integral domains A for which $s(F) < \infty$ does imply $s(A) < \infty$. In the following, we shall collect some of the known facts in this direction.

PROPOSITION 1.7. Let A be a valuation ring with quotient field F. Then s(A) = s(F). (In particular, $s(A) < \infty$ if and only if $s(F) < \infty$.)

PROOF. We claim that, if -1 is a sum of *n* squares in *F*, then -1 is a sum of *n* squares in *A*. This will clearly give the Proposition. Assuming that -1 is a sum of *n* squares in *F*, we can write down an equation $a_0^2 + \cdots + a_n^2 = 0$, where $a_i \in A$ are not all zero. Using the fact_that *A* is a

valuation ring, we may assume, after reindexing, that $a_0 \neq 0$, and $a_i/a_0 \in A$ for all *i*. Then $-1 = (a_1/a_0)^2 + \cdots + (a_n/a_0)^2$ is a sum of *n* squares in *A*.

COROLLARY 1.8. Let A be a Prüfer domain with quotient field F. Then $s(A) < \infty$ if and only if $s(F) < \infty$.

PROOF. One of the characterizations for A to be a Prüfer domain is that, for all maximal (or prime) ideals $\mathfrak{M} \subset A$, $A_{\mathfrak{M}}$ is a valuation ring in F. Assun: that $s(F) < \infty$. Then, by (1.7), $s(A_{\mathfrak{M}}) < \infty$ for all maximal ideals \mathfrak{M} . But then by the Local-Global Criterion for Finite Level, we have $s(A) < \infty$.

The Corollary above, in particular, holds for Dedekind domains A. However, the technique we used to prove 1.8 did not yield any quantitative results relating s(A) and s(F) for Prüfer domains. In the case of Dedekind domains, such a quantitative result is indeed possible. We shall state without proof the following result of R. Baeza [3].

THEOREM 1.9. Let A be a Dedekind domain with quotient field F. Then s(A) is either s(F) or s(F) + 1. In particular, s(A) is either ∞ , or 2^n , or $2^n + 1$ (for some integer $n \ge 0$).

To show the possibility of the values $2^n + 1$ in this theorem, we mention the following example from [9].

EXAMPLE 1.10. Consider the ring $A = \mathbf{Q}[x, y]/(1 + x^2 + 2y^2)$. It can be shown that A is a principal ideal domain. The equation $-1 = \bar{x}^2 + \bar{y}^2 + \bar{y}^2$ shows that $s(F) \leq s(A) \leq 3$. Using elementary considerations, one can show that s(F) = 2 and s(A) = 3. The details can be found in [9].

Even for $n \ge 2$, there probably exist Dedekind domains A with quotient field F such that $s(F) = 2^n$ and $s(A) = 2^n + 1$. However, no such examples seem to have been exhibited in the literature.

We now turn our attention to regular local rings.

LEMMA 1.11. Let (A, \mathfrak{M}) be a regular local ring with quotient field F. Then $s(A/\mathfrak{M}) \leq s(F)$.

PROOF. We shall proceed by induction on $d: = \dim A$. If d = 1, then A is a discrete valuation ring; in this case s(F) = s(A) by 1.7 so the desired inequality is trivial. Now assume d > 1. Fix an element $p \in \mathfrak{M}$ which is part of a regular system of parameters for \mathfrak{M} . Then $A_{(p)}$ is a discrete valuation ring, so by (1.7), $s(A_{(p)}) = s(F)$. Let B: = A/(p). This is a regular local ring of dimension d - 1, with residue field $\cong A/\mathfrak{M}$, so by the inductive hypothesis $s(A/\mathfrak{M}) \leq s(\overline{qf(B)})$. But $qf(B) \cong A_{(p)}/p \cdot A_{(p)}$ so $s(qf(B)) \leq s(A_{(p)}) = s(F)$. Combining the two inequalities, we get $s(A/\mathfrak{M}) \leq s(F)$.

If the local ring (A, \mathfrak{M}) fails to be regular, the inequality $s(A/\mathfrak{M}) \leq s(F)$ may not hold, even if $s(A/\mathfrak{M})$ and s(F) are both finite. Consider, for

example, the ring $B = \mathbb{R}[u, v, x_1, \ldots, x_r]/(u^2 + v^2, 1 + x_1^2 + \cdots + x_r^2)$ and the prime ideal $\mathfrak{p} = (\bar{u}, \bar{v}, 1 + \bar{x}_1^2 + \cdots + \bar{x}_r^2)$ in B. Let A be the localization $B_\mathfrak{p}$ with maximal ideal $\mathfrak{M} = \mathfrak{p} B_\mathfrak{p}$. In qf(A), we have $-1 = (\bar{u}/\bar{v})^2$ so s(qf(A)) = 1. However,

$$A/\mathfrak{M} \cong B_{\mathfrak{g}}/\mathfrak{p}B_{\mathfrak{g}} \cong qf(B/\mathfrak{p}) \cong qf(\mathbb{R}[x_1, \ldots, x_r]/(1 + x_1^2 + \cdots + x_r^2))$$

and, by a theorem of Pfister [20, p. 303], this field has level 2^n where $2^n \leq r < 2^{n+1}$.

Using the Lemma above, we can obtain a result on the level of a nonlocal regular ring. Recall that a commutative ring A is said to be regular if A is Noetherian and the localizations of A at all maximal ideals are regular local rings. It is known that if A is regular, then the localizations of A at all prime ideals are also regular local rings.

THEOREM 1.12. Let A be a regular domain with quotient field F. Then $s(A) < \infty$ if and only if $s(F) < \infty$.

PROOF. To show the "if" part, assume that $s(F) < \infty$. By 1.5, it suffices to show that $s(qf(A/p)) < \infty$ for any prime ideal $p \subset A$. But $qf(A/p) \cong A_p/pA_p$ is the residue field of the regular local ring A_p . By 1.11, we have $s(A_p/pA_p) \leq s(qf(A_p)) = s(F) < \infty$.

Even in the case when A is a regular local ring, the exact relationship between s(A) and s(F) seems to be unknown. It seems reasonable to conjecture that, in this case, s(A) = s(F); this is known to be true for low dimensional regular local rings (in the case when 2 is invertible in A). If A is not local but just a regular domain, the relationship between s(A)and s(F) seems to be even more inaccessible. In view of Baeza's result (1.9), one might perhaps ask whether s(A) is bounded by some function of s(F) and dim A.

2. Two kinds of Reality. For commutative rings A, we can introduce two notions of "reality", a weak one and a strong one, as follows.

÷,

J.

DEFINITION 2.1.

(1) A is said to be semireal if $s(A) = \infty$.

(2) A is said to be formally real (or just real) if $a_1^2 + \cdots + a_n^2 = 0$, $(a_i \in A)$ implies that each $a_i = 0$.

We can also define two similar notions of reality for ideals $\mathfrak{A} \subset A$.

(1') \mathfrak{A} is said to be semireal if A/\mathfrak{A} is semireal

(2') \mathfrak{A} is said to be real if A/\mathfrak{A} is real.

Clearly, (1) and (2) define the same notion if A is a field. Therefore, (1') and (2') define the same notion if A is a maximal ideal.

Taking the definitions literally, we would have to concede that the zero ring is real, but not semireal. This is somewhat embarrassing, but size we seldom consider the zero ring, it should not cause any problem. If $A \neq 0$ and $\mathfrak{A} \neq A$, we clearly have

A is real \Rightarrow A is semireal, and \mathfrak{A} is real \Rightarrow \mathfrak{A} is semireal.

The example $A = \mathbf{R}[x_1, \ldots, x_n]/(x_1^2 + \cdots + x_n^2)$ mentioned already in the last section shows that a ring may just be semireal but not real.

REMARKS 2.2 (1) Suppose we have a ring homomorphism $f: A \rightarrow B$. Then

(a) B is semireal $\Rightarrow A$ is semireal;

(b) If f is injective, then B is real \Rightarrow A is real.

(2) Consider a localization $A \rightarrow S^{-1}A$ of the ring A. Then

(a) $S^{-1}A$ is semireal $\Rightarrow A$ is semireal;

(b) A is real $\Rightarrow S^{-1}A$ is real.

To prove the latter, suppose $(a_1/s)^2 + \cdots + (a_n/s)^2 = 0$ in $S^{-1}A$, where $a_i \in A$ and $s \in S$. Then for some $t \in S$, we have $t(a_1^2 + \cdots + a_n^2) = 0$ in A, and so $(ta_1)^2 + \cdots + (ta_n)^2 = 0$. By the reality of A, we have each $ta_i = 0$ in A, and hence each $a_i/s = 0$ in $S^{-1}A$.

THEOREM 2.3. For any commutative ring $A \neq 0$, the following statements are equivalent.

(1) A is semireal;

(2) A has a semireal ideal;

(3) A has a real ideal \neq A;

(4) A has a semireal prime ideal;

(5) A has a real prime ideal;

(6) A has a prime ideal p such that A_p is semireal;

(7) A has a maximal ideal \mathfrak{M} such that $A_{\mathfrak{M}}$ is semireal;

(8) Some localization $S^{-1}A$ is semireal.

REMARK 2.4. Statement (5) is clearly equivalent to

(9) A has a homomorphism into a real field,

since a domain is real if and only if its quotient field is real. However, (5) is not equivalent to

(5') A has a maximal ideal which is real.

In fact, the ring Z is real (and semireal), but all its maximal ideals are ponreal.

PROOF (of 2.3). First we prove that (2), (4) and (5) are equivalent. Since clearly $(5) \Rightarrow (4) \Rightarrow (2)$, we need only show that $(2) \Rightarrow (5)$. Let \mathfrak{A} be a semireal ideal. Then \mathfrak{A} is disjoint from $S: = 1 + \sum A^2$. By 1.2, 1.3 and Zorn's Lemma, we can enlarge \mathfrak{A} to a real prime ideal. This proves (5).

Next, note the trivial implications $(5) \Rightarrow (3) \Rightarrow (2)$ and the trivial equivalences $(2) \Leftrightarrow (1) \Leftrightarrow (8)$. These and the above show that all state-

ments other than (6), (7) are equivalent.

It is now easy to account for (6) and (7). Since $(7) \Rightarrow (6) \Rightarrow (1)$ are trivial, we can finish by showing that $(5) \Rightarrow (7)$. Let \mathfrak{p} be a real prime. Let \mathfrak{M} be any maximal ideal containing \mathfrak{p} . We claim that $A_{\mathfrak{M}}$ is semireal. To see this, note that

$$s(qf(A/p)) = \infty \Rightarrow s(A_p/pA_p) = \infty \Rightarrow s(A_p) = \infty$$
.

Since we have a homomorphism $A_{\mathfrak{M}} \to A_{\mathfrak{p}}$, the latter implies that $s(A_{\mathfrak{M}}) = \infty$.

COROLLARY 2.5. Let A be either a Prüfer domain or a regular domain, with quotient field F. Then the statements (1) through (9) are all equivalent to

(10) A is real, and (11) F is real.

PROOF. This follows from our earlier results (1.8) and (1.12).

For local rings, it is technically convenient to introduce a third notion of reality, as in the following

DEFINITION 2.6. A local ring (A, \mathfrak{M}) is said to be residually real if \mathfrak{M} is a real ideal, i.e., if the residue field A/\mathfrak{M} is formally real.

If (A, \mathfrak{M}) is semireal or even real, it does not follow that it is residually real. (An easy counterexample is $\mathbb{Z}_{(p)}$.) On the other hand, if (A, \mathfrak{M}) is residually real, then A must be semireal, but not necessarily real. (An easy counterexample is the localization of $\mathbb{R}[x_1, \ldots, x_n]/(x_1^2 + \cdots + x_n^2)$ at the maximal ideal generated by $\bar{x}_1, \ldots, \bar{x}_n$.) Nevertheless, we do have the following which is a consequence of 2.5 (see also 1.11).

PROPOSITION 2.7. Let A be a valuation ring or a regular local ring. If A is residually real, then A is real.

We shall now conclude this section with the following result which characterizes real rings in terms of real fields.

THEOREM 2.8. A ring A is real if and only if A can be imbedded into a direct product of (formally) real fields.

If A is a domain, this theorem becomes trivial since the quotient field of A will also be real. Thus the work needed for the proof of 2.8 is mainly in the non-domain case. We first prove the following lemma.

LEMMA 2.9. A ring A is real if and only if A is reduced and all minimal primes of A are real.

PROOF. Recall that a ring R is called reduced if its nilradical nil(R) = $\{a \in R: a^n = 0 \text{ for some } n\}$ is zero. To prove the "only if" part, let A be

a real ring. Clearly A is reduced. Let \mathfrak{p} be any minimal prime ideal. By 2.2 (2b), $A_{\mathfrak{p}}$ is real, and hence reduced. Therefore, $\mathfrak{p}A_{\mathfrak{p}} = \operatorname{nil}(A_{\mathfrak{p}}) = 0$. This implies that $\operatorname{qf}(A/\mathfrak{p}) \cong A_{\mathfrak{p}}/\mathfrak{p}A_{\mathfrak{p}} = A_{\mathfrak{p}}$. Since $A_{\mathfrak{p}}$ is real, it follows that the prime ideal \mathfrak{p} is real.

For the converse, assume that A is reduced, and that all minimal primes are real. If $\sum a_i^2 = 0$, then each a_i must lie in each minimal prime, and so each a_i must lie in the intersection of all the minimal primes, which is nil(A). Since A is reduced, it follows that each $a_i = 0$.

PROOF (of 2.8). The "if" part is trivial. For the converse, assume that A is real. Let $\{\mathfrak{p}_i\}$ be the set of minimal primes. By (2.9), the quotient fields $F_i := qf(A/\mathfrak{p}_i)$ are real. The obvious homomorphism from A to $\prod F_i$ has kernel = $\bigcap \mathfrak{p}_i = nil(A) = 0$, so we get an imbedding of A into the direct product of the real fields $\{F_i\}$.

At this point, let us make some remarks about Lemma 2.9. According to this Lemma, if A is real, then A is reduced, and all its minimal primes are real. In view of this, one may ask whether if A is semireal, is A also reduced, and are all its minimal primes semireal? It is easy to see by examples that the answers to both questions are negative. For instance, the ring $\mathbf{R}[x]/(x^2)$ is semireal, but clearly not reduced. On the other hand, consider a direct product $A = F_1 \times F_2$ where F_1 is a real field and F_2 is a nonreal field. Then, since A has a homomorphism onto F_1 , it is semireal. The reduced ring A has two minimal primes, $\mathfrak{p}_1 = F_1 \times \{0\}$ and $\mathfrak{p}_2 = \{0\} \times F_2$. Since $A/\mathfrak{p}_1 \cong F_2$ and $A/\mathfrak{p}_2 \cong F_1$, we see that \mathfrak{p}_2 is real, but \mathfrak{p}_1 is not real or semireal. This example suggests that, to get an analogue of 2.9 for semireal rings, we have to replace the universal quantifier (\forall) by the existential quantifier (\exists). The correct statement is as follows.

LEMMA 2.10. A ring A is semireal if and only if one of its minimal primes is semireal.

This statement is just a slight refinement of the equivalence $(1) \Leftrightarrow (4)$ in Theorem 2.3. We just need to make two additional observations: (a) any prime $p \subset A$ always contains a minimal prime p_0 ; (b) if $p_0 \subset p$, then p is semireal $\Rightarrow p_0$ is semireal.

Finally, we would like to mention another characterization of semireal rings which is important from the viewpoint of studying quadratic forms. For any ring A, let W(A) denote the Witt ring of regular inner product spaces over A. By a signature over A, we shall mean a ring homomorphism from W(A) to the ring of integers which takes $\langle 1 \rangle$ to 1. It can be shown that the ring A is semireal if and only if there is such a signature; for a proof of this fact, we refer the reader to Knebusch's lecture notes in the Kingston Conference on Quadratic Forms (1976).

3. Artin-Schreier theory for commutative rings.

The basic connection between the notion of formal reality and the existence of orderings was first discovered by Artin and Schreier, in the context of fields [2]. They showed that a field F is real if and only if F can be ordered. We shall begin by recalling how this theorem can be proved by using the convenient notion of preorderings.

Let F be a field. A subset $T \subset F$ is called a preordering on F if $T + T \subset T$, $T \cdot T \subset T$, $F^2 \subset T$, and $-1 \notin T$. A preordering T is called an ordering if it satisfies the further condition that $T \cup -T = F$. The following two facts are both easily proved.

(a) A preordering T is an ordering if and only if T is maximal as a preordering;

(b) (By Zorn's Lemma) Any preordering T can be enlarged into an ordering.

If we assume these two facts, the Artin-Schreier Theorem mentioned above can be proved as follows. Assume F is real. Then $-1 \notin \sum F^2$, so $\sum F^2$ is a preordering. Applying (b) above to $T = \sum F^2$, we get an ordering on F. Conversely, if T is an ordering on F, then $T \supset \sum F^2$. Since $-1 \notin T$, we must have $-1 \notin \sum F^2$, so F is real.

Suppose, instead of studying fields, we want to study commutative rings. It is natural to ask whether there is some sort of generalization of the Artin-Schreier Theorem to rings. To find such a generalization, one problem we have to solve is to try to come up with the right definition of an "ordering" for a commutative ring. We proceed as follows.

First, we can define preorderings T in a ring A in the same way as we did for fields, namely, we call $T \subset A$ a preordering if $T + T \subset T$, $T \cdot T \subset T$, $A^2 \subset T$ and $-1 \notin T$. If A were a field, then $T \cap -T$ would be zero, for otherwise there would exist $t_1, t_2 \in T$ such that $t_1 = -t_2 \neq 0$; but then $-1 = t_1/t_2 = t_1t_2 \cdot (t_2^{-1})^2 \in T$, a contradiction. If A is just a ring, then, for a preordering $T, T \cap -T$ need not be zero. It is easy to see that $\mathfrak{A} := T \cap -T$ is an additive subgroup of A, i.e., since $\mathfrak{A} = -\mathfrak{A}$, it suffices to check that \mathfrak{A} is closed under addition. But for $a_i \in \mathfrak{A}$, we have

$$\begin{array}{cc} a_1, & a_2 \in T \Rightarrow a_1 + a_2 \in & T \\ -a_1, & -a_2 \in T \Rightarrow a_1 + a_2 \in -T \end{array} \Rightarrow a_1 + a_2 \in \mathfrak{A}.$$

The group $\mathfrak{A} = T \cap -T$ is clearly the largest additive subgroup of A contained in T. In the following, we shall call it the support of T (abbreviated as supp(T)).

REMARK 3.0. (a) If $1/2 \in A$, then the support \mathfrak{A} of any preordering *T* is an ideal of *A*. For this, we have to show that $a \in \mathfrak{A}$ and $x \in A$ imply that $xa \in \mathfrak{A}$. Write *x* in the form $y^2 - z^2$; this is always possible if $1/2 \in A$, for we can take y = (1 + x)/2 and z = (1 - x)/2. We have

$$\begin{array}{ccc} a \in T \Rightarrow & y^2 a \in T \\ -a \in T \Rightarrow & -z^2 \ a \in T \end{array} \Rightarrow xa = (y^2 a) + (-z^2 a) \in T.$$

Similarly we can get $xa \in -T$. Therefore, $xa \in A$.

(b) If a preordering $T \subset A$ satisfies the condition $T \cup -T = A$, then the support \mathfrak{A} of T will be an ideal of A even if 2 is not invertible. This follows by an easy argument.

We are now ready to define the notion of an ordering in an arbitrary commutative ring.

DEFINITION 3.1. A preordering $T \subset A$ is called an ordering if it satisfies the two further conditions below:

(1) $T \cup -T = A$, and

(2) The support $\mathfrak{p} = T \cap - T$ is a prime ideal of A.

In the literature, such a T is sometimes called a prime ordering. To simplify the terminology, we propose to call it just an ordering.

Note that, given an ordering $T \subset A$ with support p as above, we can go to the domain $\overline{A} = A/p$, on which T induces a preordering \overline{T} satisfying $\overline{T} \cup -\overline{T} = \overline{A}$ and $\overline{T} \cap -\overline{T} = \{\overline{0}\}$. Thus, \overline{T} is an ordering on \overline{A} with support zero, and therefore extends uniquely to an ordering on qf(A/p). Conversely, if we start with a real prime ideal p, and fix an ordering on qf(A/p), then, by restriction to A/p, we get an ordering on A/p with support zero. Taking its preimage in A, we clearly get an ordering on A with support p. Therefore, the prescription of an ordering on A with support p is equivalent to the prescription of an ordering on the quotient field of A/p. The possible supports of orderings on A are precisely all the real prime ideals of A.

The following result offers another characterization of the notion of an ordering.

THEOREM 3.2. A preordering T on a ring A is an ordering if and only if it satisfies the following property.

(*)
$$a b \in -T \Rightarrow a \in T \text{ or } b \in T.$$

PROOF. Suppose T satisfies (*). Taking b = -a in (*), we see that $T \cup -T = A$. Now let p be the support of T. By Remark 3.0 (b), this is an ideal of A. To show that p is prime, we shall check that $x y \in p = T \cap -T$ and $x \notin p$ imply $y \in p$. Without loss of generality, we may assume that $x \notin T$. Then, by (*), we have (for both signs)

$$x \cdot (\pm y) \in -T \Rightarrow \pm y \in T$$

$$\Rightarrow y \in \mathfrak{p}, \text{ as desired.}$$

Conversely, let T be an ordering with support p. Suppose $a \ b \in -T$, but $a \notin T$ and $b \notin T$. Then we have $a, b \in -T$ and so $a \ b \in (-T)(-T) = T$.

Therefore $a b \in T \cap -T = p$. Since p is prime, we have either $a \in p$ or $b \in p$. This is a contradiction since $p \subset T$.

For an element $a \in A$ and an ordering $T \subset A$ as in 3.1, we shall use the following notations.

(3.3)
$$\begin{cases} a \geq 0 : \Leftrightarrow a \in T, \qquad a \leq T \\ a \geq 0 : \Leftrightarrow a \in -T, \\ a \geq 0 : \Leftrightarrow a \notin -T, \qquad a < 0 : \Leftrightarrow a \notin T, \\ a = 0 : \Leftrightarrow (a \leq T) \text{ and } a \geq T \\ a \geq T \cap -T = \operatorname{supp}(T). \end{cases}$$

Thus, for any $a \in A$, we have exactly one of the following possibilities: $a >_T 0$, or $a <_T 0$ or $a =_T 0$. Since $a \ge_T 0$ means simply that $\bar{a} \ge_T 0$ in the quotient field of $A/T \cap -T$, etc., it is easy to see that the usual laws of inequalities remain valid for \ge_T and $>_T$. In terms of this inequality notation, the characterizing property (*) for an ordering T in Theorem 3.2 may be written in the more suggestive form: $ab \le_T 0 \Rightarrow a \ge_T 0$ or $b \ge_T 0$.

Next, we shall try to prove some existence results for orderings. We begin with the following observation of A. Prestel.

LEMMA 3.4. Let T be a preordering on a ring A, and let $x, y \in A$. (1) If $xy \in -T$, then one of $T + x \cdot T$, $T + y \cdot T$ is a preordering. (2) $T + x \cdot T$ is not a preordering if and only if there exists an equation -1 = a + (1 + b)x where $a, b \in T$.

PROOF. For (1), let $T_1 = T + x \cdot T$ and $T_2 = T + y \cdot T$. Both of these are closed under addition, multiplication, and contain A^2 . If neither one is a preordering, we would have the equations

$$-1 = t_1 + xt_2$$
, $-1 = t_3 + yt_4$ where $t_i \in T$.

Multiplying $-xt_2 = 1 + t_1$ with $-yt_4 = 1 + t_3$, we get $xyt_2t_4 = 1 + t_5$ for some $t_5 \in T$. But then $-1 = t_5 - xyt_2t_4 \in T$, a contradiction.

For (2), the "if" part is trivial. For the "only if" part, assume that $T_1 = T + x \cdot T$ is not a preordering; then $-1 \in T_1$. Multiplying this by $1 + x \in T_1$, we get $-(1 + x) \in T_1 \cdot T_1 = T_1 = T + x \cdot T$. Adding x, we get $-1 \in T + (1 + T) \cdot x$, as desired.

Since $-x^2 \in -T$ for all x, 3.4 (1) implies the following.

COROLLARY 3.5. For any $x \in A$, one of $T \pm x \cdot T$ is a preordering.

With the preparation above, it is now an easy matter to deduce the following result of A. Prestel [24].

THEOREM 3.6. Let A be a ring, and T be a maximal preordering on A. Then T is an ordering. **PROOF.** In view of 3.2, we need only check that T has the following property.

(*)
$$xy \in -T \Rightarrow x \in T \text{ or } y \in T.$$

Assume that $xy \in -T$. By 3.4, the maximality of T implies that either $T + x \cdot T = T$ or $T + y \cdot T = T$. Therefore, we have either $x \in T$ or $y \in T$, as desired.

COROLLARY 3.7. Any preordering T on a ring A is contained in an ordering.

PROOF. By Zorn's Lemma, we can enlarge T to a maximal preordering T_1 . By 3.6, T_1 must be an ordering.

Clearly, (3.7) implies

COROLLARY 3.8. Let T be an ordering on a ring A. Then T is maximal as an ordering if and only if it is maximal as a preordering.

In view of this result, there will be no ambiguity in talking about "maximal orderings". Moreover, by 3.6, maximal orderings and maximal preorderings are the same objects, and by Zorn's Lemma, any preordering is contained in a maximal ordering.

We are now ready to state the desired generalization of the Artin-Schreier Theorem for commutative rings.

THEOREM 3.9. A ring A is semireal if and only if A has an ordering.

PROOF. Let T be an ordering on A. If A has finite level, we would have $-1 \in \sum A^2 \subset T$, a contradiction. Conversely, assume that A is semireal. Then $\sum A^2$ does not contain -1, so it is a preordering on A. By what we have said above, this can be enlarged to a (maximal) ordering on A, so A has at least one ordering. Alternatively, we can also argue as follows. Since A is semireal, there exists a real prime ideal p by 2.3. The quotient field of A/p is formally real, so we can fix an ordering on it, by the Artin-Schreier Theorem in the field case. But then, as we have observed before, this ordering on qf(A/p) corresponds to an ordering on A (with support p). Therefore, A has at least one ordering.

By Theorem 3.6, any maximal preordering on a ring A is an ordering. If A happens to be a field, we know that the coverse is also true, but if A is just a commutative ring, the converse is in general false. In other words, there may exist orderings on rings which are not maximal orderings. In the following paragraph, we shall offer such an example on the ring $\mathbf{R}[x]$.

Let $A = \mathbf{R}[x]$, and let \mathfrak{p}_0 be the real prime (x). Pulling back the standard ordering on $A/\mathfrak{p}_0 \cong \mathbf{R}$, we get an ordering T_0 on A with center \mathfrak{p}_0 . Clearly, $T_0 = \{a_0 + a_1x + \cdots + a_nx^n : n \ge 0, a_0 \ge 0\}$. Another ordering S on A

is given by restricting to $\mathbf{R}[x]$ the unique ordering on $\mathbf{R}(x)$ with respect to which x is positive and "infinitesimal". More explicitly, we can define S to be $S = \{0\} \cup \{a_m x^m + a_{m+1} x^{m+1} + \cdots + a_n x^n : 0 \le m \le n, a_m > 0\}$. It is straightforward to check that S is an ordering on A, with support $\{0\}$. Clearly we have $S \subseteq T_0$, but the inclusion is strict, i.e., $\pm x$ both belong to T_0 but, while x belongs to S, -x doesn't. This shows that S is not a maximal ordering.

What are some examples of maximal orderings on A? It is easy to see that T_0 is one. More generally, if we take any $a \in \mathbf{R}$, then the ordering T_a on A obtained by pulling back the standard ordering on $A/(x - a) \cong \mathbf{R}$ is a maximal ordering on A. However, these T_a 's are not all the maximal orderings on A. To construct two more, let T_{∞} (resp. $T_{-\infty}$) be the ordering on A obtained by restricting to A the ordering on $\mathbf{R}(x)$ in which x is larger (resp. smaller) than any element of **R**. We claim that $T_{\pm\infty}$ are both maximal orderings on A. By symmetry, it suffices to show this for T_{∞} . Assume T_{∞} is not maximal, and let T be an ordering properly containing it. The support of T is clearly $\neq 0$. (For any ring A and orderings T, T', we have $T' \subseteq T \Rightarrow \operatorname{supp}(T') \subseteq \operatorname{supp}(T)$. In fact, suppose $T' \subseteq T$ but $\operatorname{supp}(T') =$ supp(T). Take $a \in T \setminus T'$. Then $a \in -T' \subset -T$ and so $a \in supp(T) =$ $supp(T') \subset T'$, a contradiction.) Hence the support of T must be of the form (x - a) for some $a \in \mathbf{R}$. But then T must be T_a . Now consider f =x - (a + 1). Using definitions it is easy to see that $f \in T_{\infty}$ and $f \notin T_a = T$, a contradiction. We leave it as an exercise for the reader to show that the maximal orderings on A are precisely the T_a 's together with $T_{+\infty}$.

From the last paragraph we learn that if an ordering T on a ring A has a support which is maximal as a real ideal, then T is maximal as an ordering; however, the coverse is in general not the case.

At this point, let us also make the following observation about the definition of an ordering given in 3.1. For a preordering T to be an ordering on A, we required in 3.1 that $T \cup -T = A$, and that $T \cap -T$ be a prime ideal. In general, the last condition about the support of T will not follow automatically from the other conditions. While we were discussing the ring $A = \mathbf{R}[x]$, this is a good place to give an example to substantiate this point. Keeping the notations in the last paragraph for A, let us define $T^{(r)} = S + (x^r)$, where r is any integer ≥ 1 . It is easy to verify that $T^{(r)}$ is a preordering on A, with the property that $T^{(r)} \cup -T^{(r)} = A$. Hwoever, $T^{(r)}$ has support $T^{(r)} \cap -T^{(r)} = (x^r)$ which is a prime ideal if and only if r = 1. Thus $T^{(1)} (= T_0!)$ is an ordering, but $T^{(r)} (r \geq 2)$ are not orderings.

4. The real spectrum of a ring. For a formally real field F, it is wellknown that the set X_F of orderings on F can be given a natural topology which is known as the Harrison topology. With this topology, X_F is a Boolean space, i.e., it is compact, Hausdorff and totally disconnected. In the literature, the space X_F has played an important role in the study of ordered fields. For a report on the recent progress in this area, see my survey article [21]; cf. also [22].

Since, in 3.1, we have extended the notion of an ordering from the case of fields to the case of rings, we can form the set X_A of orderings on any semireal commutative ring A. We can introduce the Harrison topology in the same way as we did for fields; the resulting space of orderings X_A is called the real spectrum of A. (In the literature, the notation *r*-spec Ais often used for X_A .) It turns out that X_A is always compact for (semireal) rings A; however, in general, X_A may fail to be Hausdorff, so it won't be a Boolean space. Roughly speaking, the real spectrum of a ring is some kind of "cross-breed" between the space of orderings of a field and the Zariski prime ideal spectrum of a ring.

In this section, we shall give an introduction to the topological structure of the real spectrum. Because of the limited scope of these notes, we shall not go deeply into this subject. Instead, we shall content ourselves with an exposition of the most basic properties of the real spectrum. For a deeper investigation, we refer our readers to the article [11] and the references contained therein. From these works, it will be clear that the study of the real spectrum is an indispensable tool in a systematic development of real algebraic geometry or semialgebraic geometry.

We shall begin by introducing the Harrison topology \mathcal{T} on the real spectrum X_A . (Throughout this section, A shall denote a semireal commutative ring.) By definition, a subbasis of open sets for \mathcal{T} is given by the Harrison sets

$$H(a) \ (= H_A(a)) = \{T \in X_A : a \notin -T\} = \{T \in X_A : a >_T 0\} \qquad (a \in A).$$

(From here on, we shall use freely the inequality notations introduced in (3.3).)

Among these sets, we have, for example, $H(1) = X_A$, and $H(-1) = H(0) = \emptyset$. For $a_1, \ldots, a_n \in A$, we form the finite intersections

$$H(a_1, \ldots, a_n) \coloneqq H(a_1) \cap \cdots \cap H(a_n)$$
$$= \{T \in X_A : a_i >_T 0 \text{ for all } i\}.$$

These sets form a basis of open sets for the topology \mathcal{T} . Recall that any $T \in X_A$ induces an ordering \overline{T} on $qf(A/\mathfrak{p})$, where \mathfrak{p} is the support of T. Thus, we have $T \in H(a_1, \ldots, a_n)$ if and only if the elements $\overline{a}_1, \ldots, \overline{a}_n$ in $qf(A/\mathfrak{p})$ are positive with respect to the ordering \overline{T} .

It is easy to see that the real spectrum formation $A \mapsto X_A$ gives a contravariant functor from the category of semireal commutative rings to the category of topological spaces. In fact, if $f: A \to B$ is a homomorphism of semireal rings, we can define a natural map $f^*: X_B \to X_A$ by taking $f^*(T) = f^{-1}(T)$. It is easy to see that, if T is an ordering on B, then $f^{-1}(T)$ is an ordering on A, and that if T has support p in B, then $f^{-1}(T)$ has support $f^{-1}(p)$ in A. The fact that f^* is a continuous map follows from the easy formula $f^{*-1}(H_A(a)) = H_B(f(a))$, for any $a \in A$.

Note that, in the study of X_A , there is a fundamental difference between the case of a field and the case of a ring. If A is a field and $a \neq 0$, then H(-a) is the complement of H(a); in particular, the sets H(a) are not only open sets, but they are also closed sets. Because of this, the real spectrum of a field is a totally disconnected space. If A is a ring, we will still have $H(a) \cap H(-a) = \emptyset$, but $H(a) \cup H(-a)$ won't be the whole space in general. In fact, for any ordering T, we have

Thus, $X_A \setminus (H(a) \cup H(-a))$ consists precisely of orderings whose supports contain a. Of course, this is in general a non-empty set.

Our first main result about (X_A, \mathcal{T}) is the following

THEOREM 4.1. For any semireal ring A, the real spectrum X_A is compact.

(Throughout this paper, "compact" means quasi-compact, and does not mean quasi-compact and Hausdorff.)

The proof of this is modeled upon the usual argument used to show the compactness of the space of orderings of a field. The key point is to imbed X_A into $Y: = \{0, 1\}^A$, the space of functions from A to $\{0, 1\}$. For any ordering $T \subset A$, we define $f_T: A \to \{0, 1\}$ to be the characteristic function on $T \setminus (-T)$, i.e.,

$$f_T(a) = \begin{cases} 1 & \text{if } a > 0, \\ 0 & \text{if } a \leq T 0. \end{cases}$$

For any ordering $T \subset A$, we shall identify T with the function $f_T \in Y$. Thus, X_A is identified with a certain set of functions $X \subset Y$.

We shall give $\{0, 1\}$ the discrete topology and $Y = \{0, 1\}^A$ the product topology. For $a_1, \ldots, a_n \in A$ and $\varepsilon_1, \ldots, \varepsilon_n \in \{0, 1\}$, we define

$$H_{\varepsilon_1,\ldots,\varepsilon_n}(a_1,\ldots,a_n) \coloneqq \{f \in Y \colon f(a_i) = \varepsilon_i \text{ for all } i\}$$

These sets form a basis of open sets for the product topology of Y. By Tychonoff's Theorem, the product space Y is compact.

Since we identify X_A with the set of functions $X = \{f_T: T \in X_A\}$, we have

$$H_{1,...,1}(a_1,...,a_n) \cap X = \{T: f_T(a_i) = 1 \text{ for all } i\}$$

= $\{T: a_1,...,a_n >_T 0\}$
= $H(a_1,...,a_n).$

Therefore, the topology \mathcal{T} on X_A is coarser than the topology \mathcal{T}_0 on X induced by the product topology on Y. The latter topology on $X_A = X$ will be called the Tychonoff topology on X_A . (In the literature, this topology has also been known as the "constructible topology". The term "Tychonoff topology" used here stems from G. Brumfiel's notes at the University of Hawaii. In the sequel, when we refer to the real spectrum X_A without explicitly mentioning which topology is being used, it will be understood that we are using the Harrison topology, not the Tychonoff topology.) We claim

PROPOSITION 4.2. X is a closed subset of Y.

If we assume this result, then since Y is compact, the closed subspace (X, \mathcal{T}_0) is compact and hence X_A will also be compact relative to the coarser topology \mathcal{T} . This would therefore prove Theorem 4.1.

Our job is now to prove 4.2. Consider a function $f \in Y \setminus X$. We associate to f the subset $S := -f^{-1}(0) = \{a: f(-a) = 0\}$ of A. This subset is not an ordering, so it must violate at least one of the following properties:

(1)
$$S + S \subset S$$
, (2) $S \cdot S \subset S$, (3) $A^2 \subset S$
(4) $-1 \notin S$ (5) $a \cdot b \in -S \Rightarrow a \in S$ or $b \in S$.

Suppose f does not satisfy (1). Then there exist $a, b \in S$ such that $a + b \notin S$. This leads to f(-a) = 0, f(-b) = 0, f(-(a + b)) = 1. So f lies in $H = H_{0,0,1}(-a, -b, -(a + b))$. Since the functions in X "arise" from orderings, the open set H is disjoint from X. If f violates the other properties instead of (1), we can construct a neighborhood of f disjoint from X in a similar way. Therefore, X is closed and we have proved Proposition 4.2 (and Theorem 4.1).

(For a proof of the compactness of (X_A, \mathcal{T}_0) without using Tychonoff's Theorem, see C. Saliba's thesis [27].)

In the case of fields, since $T \cap -T = \{0\}$ for any ordering T, we have

$$H_0(a) \cap X = \begin{cases} H(-a) & \text{if } a \neq 0, \\ X & \text{if } a = 0. \end{cases}$$

Therefore, in this case, the Harrison topology and the Tychonoff topology on X_A are the same, and the basic sets displayed above are clopen (closed and open). This implies

COROLLARY 4.3. If A is a field, the real spectrum X_A is a Boolean space.

Next we prove a result which characterizes the inclusion of one ordering in another in terms of the topology of the real spectrum:

PROPOSITION 4.4. Let T, T' be orderings on A. Then $T \subset T'$ if $T' \in \{\overline{T}\}$. (By $\{\overline{T}\}$, we mean the closure of the singleton set $\{T\}$ in (X_A, \mathcal{T}) .)

PROOF. First suppose $T \subset T'$. We claim that any neighborhood H of T' contains T. It suffices to consider a neighborhood of the type $H \coloneqq H(a_1, \ldots, a_n)$. Since $T' \in H$, we have $a_1, \ldots, a_n \notin -T'$; but then $a_1, \ldots, a_n \notin -T$, and hence $T \in H$.

Conversely, assume that $T \not\subset T'$. Fix an element $a \in T \setminus T'$. Clearly, H(-a) is a neighborhood of T' which does not contain T. Hence $T' \notin \{\overline{T}\}$.

From the viewpoint of algebraic geometry, if $T' \in \{T\}$, it is reasonable to say that T' is a "specialization" of T (and that T is a "generalization" of T'). According to 4.4, the specializations of a given ordering T are simply the various orderings containing T. In particular, we have

COROLLARY 4.5. The maximal orderings on A are precisely the closed points in the real spectrum X_A . (In the following, the subspace of closed points in X_A will be denoted by X_A^m .)

To better understand the topological structure of X_A^m , we need the following lemma.

LEMMA 4.6. For $T_1, T_2 \in X_A$, the following three statements are equivalent: (1) $T_1 \not\subset T_2$ and $T_2 \not\subset T_1$;

- (2) There exists an $a \in A$ such that $T_1 \in H(a)$ and $T_2 \in H(-a)$;
- (3) There exist two disjoint open sets V_1 , V_2 such that $T_1 \in V_1$ and $T_2 \in V_2$.

PROOF. (2) \Rightarrow (3) is trivial since $H(a) \cap H(-a) = \emptyset$. (3) \Rightarrow (1) follows from (4.4). (1) \Rightarrow (2) Let $x \in T_1 \setminus T_2$ and $y \in T_2 \setminus T_1$. Consider the element a = x - y. If $-a \in T_1$, adding $x \in T_1$ yields $y \in T_1$, a contradiction. Therefore, $a \notin -T_1$, i.e., $T_1 \in H(a)$. Similarly, if $a \in T_2$, adding $y \in T_2$ yields $x \in T_2$, a contradiction. Therefore, $a \notin T_2$, i.e., $T_2 \in H(-a)$.

From the Lemma above, we can deduce several important Propositions.

PROPOSITION 4.7. (L. Bröcker [6]) For any semireal ring A, X_A^m is a compact Hausdorff space.

PROOF. Applying the Lemma to two distinct points in X_A^m , we see that X_A^m is Hausdorff. To show that \underline{X}_A^m is compact, consider a covering \mathscr{F} of X_A^m by open sets of X_A . Consider any $T \in X_A$; let T' be a maximal ordering containing T. Then T' is contained in an open set $V \in \mathscr{F}$. Since $T' \in \{T\}$ by 4.4, V must also contain T. Therefore the family \mathscr{F} actually covers the whole space. The compactness of X_A^m now follows from the compactness of X_A .

PROPOSITION 4.8. Let $T \in X_A$, Then the set of orderings containing T forms a chain under inclusion. In particular, there is a unique maximal ordering T' containing T.

PROOF. Suppose $T \subset T_i$ (i = 1, 2). If $T_1 \not\subset T_2$ and $T_2 \not\subset T_1$, then by 4.6 there exist disjoint open sets V_i such that $T_i \in V_i$ (i = 1, 2). But since $T_i \in \overline{\{T\}}$, we must have $T \in V_1 \cap V_2$, a contradiction.

In view of the second conclusion in the Proposition above, there is a retraction map λ from X_A to X_A^m defined by $\lambda(T) = T'$ (in the notation of 4.8). It turns out that this map λ is continuous. To prove its continuity, we shall need the following "regularity" lemma [6].

LEMMA 4.9. Let $T \in X_A^m$ and let C be a closed set of X_A not containing T. Then there exist disjoint open sets V_1 , $V_2 \subset X_A$ such that $T \in V_1$ and $C \subset V_2$. In fact, V_1 and V_2 can be chosen such that $V_1 = H(a_1, \ldots, a_n)$ and $V_2 = H(-a_1) \cup \cdots \cup H(-a_n)$ for suitable $a_i \in A$.

PROOF. We essentially repeat here the usual proof that 'compact Hausdorff' \Rightarrow 'regular'. For any $S \in C$, we have clearly $S \not\subset T$ and $T \not\subset S$. Therefore, by 4.6, there exists an $a_S \in A$ such that $T \in H(a_S)$ and $S \in$ $H(-a_S)$. The $H(-a_S)$'s cover C, and the closed set $C \subset X_A$ is compact. Passing to a finite subcover, we have $C \subset H(-a_1) \cup \cdots \cup H(-a_n)$ and $T \subset H(a_1) \cap \cdots \cap H(a_n)$ for suitable $a_i \in A$.

We now state the following result which was pointed out by N. Schwartz.

PROPOSITION 4.10. [28] The map $\lambda: X_A \rightarrow X_A^m$ is continuous. The topology on X_A^m is just the quotient topology of (X_A, \mathcal{T}) with respect to λ .

PROOF. Let $\lambda(S) = T$. To prove the continuity of λ at S, let U be an open set of X_A containing T, and let $C = X_A \setminus U$. By the Lemma above, there exist disjoint open sets V_1 , V_2 such that $T \in V_1$ and $C \subset V_2$. Since $T \in \{\overline{S}\}$, we have $S \in V_1$. We claim that $\lambda(V_1) \subset U$ (this clearly implies the continuity of λ). In fact, let $S' \in V_1$. If $T' := \lambda(S') \notin U$, then $T' \in C \subset V_2$. But since $T' \in \{\overline{S'}\}$ we have $S' \in V_2$, contradicting the fact that V_1, V_2 are disjoint.

To prove the second conclusion of the Proposition, it suffices to show that a set *B* is closed in X_A^m if and only if $\lambda^{-1}(B)$ is closed in X_A . The 'only if' part is just the continuity of λ . For the 'if' part, assume $\lambda^{-1}(B)$ is closed in X_A . Then it is compact, and so $\lambda(\lambda^{-1}(B)) = B$ is also compact, by continuity. But X_A^m is Hausdorff, so *B* must be closed.

I want to thank J. Merzel and A. Prestel for their valuable help on the writing of the above results on X_A^m .

For any $T \in X_A$, we know that the support of T is a prime ideal. There-

fore, we can define a map $\sigma: X_A \to \text{Spec } A$ (= Zariski spectrum of A) by taking $\sigma(T) = \text{supp}(T)$. The image of σ is precisely the set of all real primes in Spec A. Not surprisingly, we have

PROPOSITION 4.11. The map $\sigma: X_A \to \text{Spec } A$ is continuous (with respect to the Harrison topology on X_A and the Zariski topology on Spec A). For any real prime ideal $\mathfrak{p} \subset A$, the fiber over \mathfrak{p} with respect to σ is homeomorphic to the real spectrum of $qf(A|\mathfrak{p})$.

PROOF. Consider a basic open set $D(a) \coloneqq \{\mathfrak{p} \in \text{Spec } A : a \notin \mathfrak{p}\}$. We need to show that $\sigma^{-1}(D(a)) = \{T \in X_A : a \notin \text{supp}(T)\}$ is open in X_A . But by 4.0, this set is $H(a) \cup H(-a)$, which is of course open. This proves the first statement in the Proposition; the second statement is (essentially) obvious.

Because of 4.11 we would expect that the nature of the Zariski topology on Spec A should have a certain influence on the nature of the topology \mathcal{T} on the real spectrum X_A . One of the main features of Spec A is that any (nonempty) irreducible closed set in Spec A has a generic point. In the following result, we shall prove that the same property holds for the real spectrum.

PROPOSITION 4.12. Let C be a closed set in X_A . Then C is irreducible (i.e., not a union of two proper closed sets of C) if and only if $C = \{\overline{T}\}$ for some $T \in X_A$. (In view of 4.4, such a T is unique if it exists; it is called the generic point of the irreducible closed set C.)

PROOF. First assume that C has the form $\{\overline{T}\}$, for some ordering T. If $C = C_1 \cup C_2$ where C_1 , C_2 are closed sets, then we must have $T \in C_1$ or $T \in C_2$. But then $\{\overline{T}\}$ is contained in either C_1 or C_2 , so $C = C_1$ or C_2 . This shows that C is irreducible.

For the proof of the converse, we proceed along the same lines as in the case of the Zariski prime spectrum. In that case, if C_0 is an irreducible closed set in Spec A, one intersects the prime ideals in C_0 to obtain an ideal p; the irreducibility of C_0 implies that p is prime, from which it follows that p is a generic point of C_0 . Now let C be an irreducible closed set in X_A ; let T be the intersection of all the orderings in C. We make the following three claims:

- (1) T is an ordering on A.
- (2) T lies in C.
- (3) $C = \{T\}.$

Here, (1) is the crucial claim, so let us first assume it. To prove (2), it suffices to show that any neighborhood $H(a_1, \ldots, a_n)$ of T contains a point of C. Since $-a_i \notin T$ for each i, there exist orderings $T_1, \ldots, T_n \in C$ such that $-a_i \notin T_i$ for $1 \leq i \leq n$. By 4.6, the orderings T_1, \ldots, T_n

786

containing T must form a chain under inclusion. Say T_1 is the *smallest* one among them. Then $-a_i \notin T_1$ for all *i*, and we have $T_1 \in H(a_1, \ldots, a_n)$, as desired. This shows that $T \in C$, and hence $\{\overline{T}\} \subset C$. But by (4.4), each ordering in C belongs to $\{\overline{T}\}$, so $\{\overline{T}\} = C$.

It now only remains to prove the claim (1). Since T is clearly a preordering, we need only check that T has the property that $a \cdot b \in -T$ implies $a \in T$ or $b \in T$ (cf. Theorem 3.2). Assume that $a \cdot b \in -T$. It is easy to check that $C = (C \setminus H(-a)) \cup (C \setminus H(-b))$. For, if $P \in C$ belongs to both H(-a)and H(-b), then $a \notin P$, $b \notin P$, but $ab \in -T \subset -P$, contradictory to the fact that P is an ordering. By the irreducibility of C, we must have, say $C = C \setminus H(-a)$, i.e., $a \in P$ for every $P \in C$. But then $a \in T$, as desired.

If we think of X_A as the real analogue of Spec A, then we should think of X_A^m as the real analogue of the maximal ideal spectrum max A. However, X_A^m is always Hausdorff but max A is not; the fact that any point in X_A has a unique specialization in X_A^m is also peculiar to the real spectrum. To get the latter property in the setting of Spec A and max A, one would need to impose a very strong condition on the ring A: by definition, A is called a p-m ring (cf. [13]) if every prime ideal p in A is contained in a unique maximal ideal \mathfrak{M} of A. For such a ring A, de Marco and Orsalli have shown that the retraction Spec $A \rightarrow \max A$ sending p to \mathfrak{M} is continuous (see also [29]). In some sense, this result is a precursor of Proposition 4.10 on the continuity of the retraction $\lambda: X_A \rightarrow X_A^m$.

We shall now close this section with a brief discussion of the "constructible subsets" of the real spectrum. For any semireal ring A, let $\mathscr{C}(A)$ be the family of subsets of X_A which can be obtained from the subbasic Harrison sets $\{H(a)\}$ by using (a finite number of) boolean operations. (By boolean operations we mean the procedures of forming finite unions and intersections, and taking complements in X_A .)

DEFINITION 4.13. Sets in the family $\mathcal{C}(A)$ are called the constructible subsets of X_A .

Recall that, earlier in the section, we have identified X_A with a certain subset X in the function space $\{0, 1\}^A$. The product topology on $\{0, 1\}^A$ induces the Tychonoff topology \mathcal{T}_0 on $X = X_A$ which is in general finer than the Harrison topology \mathcal{T} on the real spectrum. Since X_A is closed in $\{0, 1\}^A$, (X_A, \mathcal{T}_0) is a Boolean space, i.e., it is compact, Hausdorff and totally disconnected. We have the following fact which was observed by van den Dries and Coste and Coste-Roy.

PROPOSITION 4.14. The constructible sets in X_A are precisely the clopen (closed and open) sets in the Tychonoff topology \mathcal{T}_0 .

(Note the following two immediate consequences of the Proposition: (1) the constructible sets are compact with respect to both \mathcal{T} and \mathcal{T}_0 ;

(2) any cover of a constructible set by other constructible sets has a finite subcover.)

PROOF. Recall that a basis for the topology on $\{0, 1\}^A$ is given by the clopen sets $H_{\varepsilon_1,\ldots,\varepsilon_r}(a_1,\ldots,a_r)$ defined earlier in this section. Since $H(a_1,\ldots,a_r) = H_1,\ldots,(a_1,\ldots,a_r) \cap X_A$, the Harrison sets $H(a_1,\ldots,a_r)$ are clopen in \mathcal{T}_0 , and hence the same holds for any set in $\mathcal{C}(A)$. Conversely, consider any clopen set C in (X_A, \mathcal{T}_0) . Let U be an open set of $\{0, 1\}^A$ such that $U \cap X_A = C$. Then U is a union of sets of the form $H_{\varepsilon_1,\ldots,\varepsilon_r}(a_1,\ldots,a_r)$. Since C is closed and hence compact in $\{0, 1\}^A$, it is a finite union of sets of the form

$$H_{\varepsilon_1,\ldots,\varepsilon_r}(a_1,\ldots,a_r)\cap X_A=\bigcap_{i=1}^r(H_{\varepsilon_i}(a_i)\cap X_A),$$

so it suffices to show that each $H_{\varepsilon}(a) \cap X_A$ is constructible. This is clear since $H_1(a) \cap X_A = H(a)$ and $H_0(a) \cap X_A = X_A \setminus H(a)$.

The argument used in the proof above showed a little bit more:

COROLLARY 4.15. Every constructible set in X_A is a finite union of sets of the form $H_1 \cap \cdots \cap H_r$ where each H_i is of the form H(a) or $H'(a) = X_A \setminus H(a)$.

(This can also be proved directly by noting that finite unions of sets of the form $H_1 \cap \cdots \cap H_r$ are closed with respect to boolean operations.)

In the study of affine varieties V over a real closed field, the constructible sets of the real spectrum of the coordinate ring of V are intimately related to the semialgebraic subsets of the variety V. This relationship will be explained in §8 below. For a deeper study of the real spectrum and its applications to real algebraic geometry, we refer the reader to [11] and the literature cited therein. (See also Knebusch's article "An invitation to real spectra" in the Proceedings of the Hamilton Conference on Quadratic Forms.)

5. Artin-Lang theory for affine algebras. (This and the next section can be read essentially without assuming the material in \$3 and \$4. In fact (except in 5.5 (C)), we shall use only the notion of orderings for fields, and not for rings. We have arranged the exposition in this way so as to preserve the classical flavor of the material to be covered here and in \$6.)

By an affine algebra, we mean a finitely generated commutative algebra over a field k. Such an algebra is a homomorphic image of a polynomial ring $k[x_1, \ldots, x_n]$, and conversely. Affine algebras are important in algebraic geometry because they arise as coordinate rings of affine varieties. For a k-affine domain A, the field of quotients F is a function field (i.e., a finitely generated extension) over k, so the transcendence degree of K over k is finite. This transcendence degree is denoted by $\operatorname{tr.} d_k K$, or bytr. $d_k A$. Affine algebras are, in many ways, special objects in the category of commutative rings. For these algebras we can usually get much more information than we can hope to get for arbitrary commutative rings. In the development of real algebra, this is especially the case. In this section, we shall develop the Artin-Lang Theory [1], [23] for affine algebras and their associated function fields which will provide the basis for the applications to real algebraic geometry in the later sections. To simplify the exposition, we shall always work over a real closed ground field k. The case when k is an arbitrary ordered field can usually be treated by passing to the real closure of k, and therefore presents no additional difficulties.

In Lang's paper [23], the principal results about affine algebras and function fields are derived by using Sturm's Theorem on the number of real roots of a real polynomial. In this section, we shall present a slightly different treatment: instead of using Sturm's Theorem, we shall use the following result about the behavior of orderings under a finite algebraic extension.

THEOREM 5.1. Let K/F be a finite (algebraic) extension of formally real fields, and let $\varepsilon_{K/F}: X_K \to X_F$ be defined by the restriction of orderings from K to F. Then the image of $\varepsilon_{K/F}$ is open in X_F .

(Incidentally, since X_K is compact and X_F is Hausdorff, $im(\varepsilon_{K/F})$ is a closed set in X_F . However, the fact that it is also open is deeper and more interesting.)

Actually, a much more general result is true. In [17], it is proved that, for any finitely generated extension K/F of formally real fields, the map $\varepsilon_{K/F}$ is, in fact, an open mapping. To derive Lang's results, however, it is enough to assume only a very special case of this Open Mapping Theorem, as stated in 5.1. In the following, we shall show in detail how to derive Lang's Existence of Rational Place Theorem and Lang's Homomorphism Theorem from 5.1. Then we return at the end of the section to give a proof of 5.1. Our exposition here follows the suggestions of J. Merzel.

Let (K, P) be an ordered field and $k \subset K$ any subfield. The archimedean hull of k in (K, P) is defined to be

$$A(k, P) = \{a \in K \colon |a|_P \underset{P}{<} b \text{ for some } b \in k\}.$$

Here, the absolute value of a with respect to P is defined in the usual way. It is easy to see that A(k, P) is a valuation ring in K, and that the image of $P \cap A(k, P)$ in the residue field \overline{K} of A(k, P) is an ordering on \overline{K} .

Using the notion of the archimedean hull, we shall deduce the following consequence of 5.1.

THEOREM 5.2. Let K be a real function field of transcendence degree 1

T.Y. LAM

over a real closed field k, and let $x \in K$ be a given transcendental element over k. Then there exists an ordering P on K such that $(x - \delta)^{-1} \notin A(k, P)$ for some $\delta \in k$.

PROOF. Fix an ordering P_0 on K and let Q_0 be its restriction to F := k(x). By 5.1, there exists an open neighborhood $H(f_1(x), \ldots, f_n(x))$ of Q_0 which lies in the image of $\varepsilon_{K/F}$. We may assume that $f_i(x) \in k[x] \setminus \{0\}$, and, after dropping from it a factor which is a sum of two squares in k[x], we may assume that each $f_i(x)$ is a product of distinct linear factors. For any ordering Q on k(x), we can define sets

$$A_i(Q) = \{a \in k : f_i(a) = 0, a < x\}$$

$$B_i(Q) = \{b \in k : f_i(b) = 0, x < b\}.$$

Clearly, these sets will completely determine the "sign" of f_i with respect to the ordering Q. In particular, if an ordering Q on k(x) is such that $A_i(Q) = A_i(Q_0)$ for all i, then we have $f_i \in Q_0 \Rightarrow f_i \in Q$, and so $Q \in H(f_1(x), \ldots, f_n(x))$. In particular, Q can be lifted to an ordering on K.

Let $a = \max(A_1(Q_0) \cup \cdots \cup A_n(Q_0))$ and $b = \min(B_1(Q_0) \cup \cdots \cup B_n(Q_0))$, formed with respect to the unique ordering on k. (If the first union is empty, we let $a = -\infty$ and if the second union is empty, we let $b = +\infty$.) Since a < x < b in Q_0 , we have a < b in k. Let $\delta \in k$ be any element between a and b (for instance, $\delta = (a + b)/2$). Let Q be an ordering on k(x) in which x is "infinitely close" to δ (i.e., $|x - \delta|_Q < \varepsilon$ for any positive $\varepsilon \in k$) (In the case $\delta = 0$, we have constructed such an ordering in the examples given after 3.9. If $\delta \neq 0$, we simply view $x - \delta$ as the new transcendental element and use the previous construction on $k(x - \delta)$.) Then clearly $A_i(Q) = A_i(Q_0)$ for all i, and, by what we said in the last paragraph, Q can be lifted to an ordering P on K. Clearly, this is the ordering we want.

SUPPLEMENT TO 5.2. Let P be as above. Then $x \in A(k, P)$ and \overline{K} (the residue field of A(k, P) is equal to k.

PROOF. Since $x - \delta$ lies in the maximal ideal of A(k, P), we clearly have $x \in A(k, P)$. For the second conclusion, note that \overline{K} is ordered by the image of $P \cap A(k, P)$. If we can show that \overline{K}/k is algebraic, then we will have $\overline{K} = k$ since k is real closed. Assume that \overline{K}/k is not algebraic. Let $y \in A(k, P)$ be such that \overline{y} is transcendental over k. Then, since $\overline{x} = \delta$, it is easy to see that x, y must be algebraically independent over k, a contradiction to the fact that tr. $d_k K = 1$.

Before we go on, let us make a simple observation about residually real valuation rings.

LEMMA 5.3. Let A be a residually real valuation ring in a field K, and let $a_1, \ldots, a_n \in K$. If $\sum a_i^2 \in A$, then each $a_i \in A$.

PROOF. Without loss of generality, we may assume that $a_i/a_1 \in A$ for all *i*. Then $a_1^2(1 + (a_2/a_1)^2 + \cdots + (a_n/a_1)^2) \in A$. Here, the expression in parentheses must be a unit in A, for otherwise -1 would be a sum of squares in the residue field. Therefore, $a_1 \in A$, and so all $a_i \in A$.

We now come to the principal result of this section.

THEOREM 5.4. (Lang's Existence of Rational Place Theorem) Let K be a real function field over a real closed field k, and let $a_1, \ldots, a_n \in K$. Then there exists a k-place $\varphi: K \to k \cup \{\infty\}$ such that $\varphi(a_i)$ is finite for all i.

PROOF. Let $d = \text{tr.} d_k K$. First assume d = 0. In this case K is a finite algebraic extension of k. Since K is real, we must have K = k, so the theorem is a tautology.

Next assume d = 1. Let $x = \sum a_i^2 \in K$. First assume that x is transcendental over k. Take the ordering P on K constructed in 5.2, and let φ be the place associated with the valuation ring A(k, P). By the Supplement to 5.2, we have $\varphi: K \to k \cup \{\infty\}$ and $\varphi(x) \neq \infty$. By 5.3, we have $\varphi(a_i) \neq \infty$ for all i so φ is the place we want. If x happens to be algebraic over k, then since k(x) is real we must have $x \in k$. In this case, find a k-place $\varphi: K \to k \cup \{\infty\}$ by using any transcendental x'. Since $\varphi(x)$ is automatically finite, we have again $\varphi(a_i) \neq \infty$ by 5.3.

Finally, we treat the general case $d \ge 1$ by induction on d. Fix a real closure \tilde{K} of K with respect to some ordering on K. Choose a k-function field $L \subset K$ such that $\operatorname{tr.} d_L K = 1$, and let \tilde{L} be the algebraic closure of L in \tilde{K} . Then \tilde{L} is real closed and $K \cdot \tilde{L}$ is a function field over \tilde{L} of transcendence degree 1. By the d = 1 case already verified above, there exists an \tilde{L} -place $\lambda: K \cdot \tilde{L} \to \tilde{L} \cup \{\infty\}$ which is finite on all a_i . Let K' be the residue field of $\lambda | K$. Clearly K' is a function field of transcendence degree d - 1 over k. Also, K' is real since $K' \subset \tilde{L}$. By the inductive hypothesis, K' has a k-place μ into $k \cup \{\infty\}$ which is finite on all $\lambda(a_i)$. Therefore we are done by taking the composition of places $\varphi := \mu \circ (\lambda | K)$.

REMARK. Of course the converse of 5.4 is also true, i.e., if K is any field admitting a place into k, then K is formally real. This follows from 2.7.

COROLLARY 5.5. (Lang's Homomorphism Theorem) Let A be any real affine domain over a real closed field k. Then there exists a k-algebra homomorphism from A to k.

PROOF. Let $A = k[a_1, \ldots, a_n]$ and K = qf(A). Let $\varphi: K \to k \cup \{\infty\}$ be a k-place such that $\varphi(a_i) \neq \infty$ for all *i*. Then $\varphi|A$ gives the desired k-homomorphism from A to k.

For later applications, it is useful to note the following Supplements (A), (B) and (C) to 5.5 which are self-strengthened versions of Lang's Homomorphism Theorem. (In all three Supplements, k denotes a real closed field, as in 5.5.)

COROLLARY 5.5(A). Let A be a real affine k-domain, and $f_i (1 \le i \le n)$ be given nonzero elements in A. Then there exists a k-algebra homomorphism φ from A to k such that $\varphi(f_i) \ne 0$ for all i.

PROOF. This follows by applying 5.5 to the (real) affine domain $A[1/f_1 \cdots f_n]$.

COROLLARY 5.5(B). Let A be any semireal k-affine algebra (not necessarily a domain). Then there exists a k-algebra homomorphism from A to k.

PROOF. By 2.3, A admits a real prime \mathfrak{p} . The Corollary then follows by applying 5.5 to the (real) affine domain A/\mathfrak{p} . (Note, however, that we cannot get the same conclusion as in 5.5(A) here. For instance, for the semireal algebra $A = k[x_1, \ldots, x_n]/(\sum x_i^2)$, clearly any homomorphism from A to k must send the \bar{x}_i 's to zero.)

COROLLARY 5.5(C). Let A be any k-affine algebra and f_i , g_j $(1 \le i \le m, 1 \le j \le n)$ be given elements in A. If there exists an ordering T on A such that $f_i >_T 0$ and $g_j \ge_T 0$ for all i, j, then there exists a k-algebra homomorphism φ : $A \to k$ such that $\varphi(f_i) > 0$ and $\varphi(g_j) \ge 0$ (in the unique ordering of k) for all i, j.

PROOF. Let \mathfrak{p} be the support of T. After replacing A by A/\mathfrak{p} we may assume that $\mathfrak{p} = 0$. Then T extends uniquely to an ordering P on K = qf(A) with $f_i \in P \setminus \{0\}$ and $g_j \in P$ for all i, j. The Corollary now follows by applying 5.5 to the (real) affine algebra $A[1/f_1 \dots f_m, \sqrt{f_1}, \dots, \sqrt{f_m}, \sqrt{g_1}, \dots, \sqrt{g_n}]$ in the real closure of (K, P).

In Lang's paper [23] there is another main result concerning the imbedding of real function fields K over k into a prescribed real closed field $R \supset k$ with $\operatorname{tr.} d_k R \ge \operatorname{tr.} d_k K$. This is known as Lang's Imbedding Theorem. For technical convenience, we shall postpone the derivation of this result to §6. This will complete our exposition of the Artin-Lang Theory for real function fields.

To conclude this section, we shall now give the proof of Theorem 5.1. This will be deduced from the following result, which goes back to J. J. Sylvester and was rediscovered by Olga Taussky [31].

THEOREM 5.6. Let (F, P) be an ordered field, and R be its real closure. Let f(t) be a nonconstant separable polynomial over F, and let K = F[t]/(f(t)). Then the number of (distinct) roots of f in R is given by the P-signature of the trace form on K.

792

(The trace form on K is defined by $T_f(x, y) = \text{tr}_{K/F}(xy)$). This is a symmetric bilinear form over F; as usual, we can think of it as an F-quadratic form. Its P-signature means the Sylvester signature with respect to the ordering P.)

PROOF (of 5.6). Let us compute $\operatorname{sgn}_{P} T_{f}$ by looking at the form $(T_{f})_{R}$ which is the trace form of the *R*-algebra $R \otimes_{F} K$. Let $f = g_{1} \cdots g_{r} h_{1} \cdots h_{s}$ be the square-free factorization of f into irreducible factors in R[t], where deg $g_{i} = 1$ and deg $h_{j} = 2$. By the Chinese Remainder Theorem, we have

$$R \otimes_F K \cong \prod R[t]/(g_i) \times \prod R[t]/(h_j)$$
$$\cong R \times \cdots \times R \times \bar{R} \times \cdots \times \bar{R}$$

where $\overline{R} = R(\sqrt{-1})$. Here, different factors have product equal to zero so they are orthogonal under the trace form. On the factor R, the trace form is $\langle 1 \rangle$, but on a factor \overline{R} , the trace form has matrix $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ so it is a hyperbolic plane. With respect to the unique ordering on R, the signature of the trace form on $R \otimes_F K$ is therefore equal to r, the number of roots of f in R.

COROLLARY 5.7. Let (F, P) be an ordered field, with real closure R. Let K = F[t]/(f(t)) be a finite algebraic field extension of F and let T_f be its trace form. Then $P \in im(\varepsilon_{K/F})$ if and only if $sgn_P T_f > 0$.

PROOF. We simply note that

$$P \in \operatorname{im}(\varepsilon_{K/F}) \Leftrightarrow \exists \text{ an } F\text{-imbedding of } K \text{ into } R$$
$$\Leftrightarrow f \text{ has a root in } R$$
$$\Leftrightarrow \operatorname{sgn}_{F} T_{f} > 0.$$

Now the set of orderings P on F with respect to which T_f has signature >0 is easily seen to be an open set in X_F . Therefore, by 5.7, $im(\varepsilon_{K/F})$ is open. This proves 5.1.

6. Hilbert's 17th Problem and the Real Nullstellensatz. (See the initial remark in §5.) In this section, we shall show how the Artin-Lang Theory developed in the last section can be used to lay part of the foundations for the subject of real algebraic geometry (i.e., the study of real algebraic varieties). We shall first apply Lang's Homomorphism Theorem to give, à la Artin, an affirmative answer to Hilbert's 17th Problem (in the case of a real closed ground field). We then go on to prove the Dubois-Risler Real Nullstellensatz for affine algebras defined over a real closed field. In the study of real affine varieties, this fundamental result plays the same role as that played by Hilbert's classical Nullstellensatz in studying affine varieties over algebraically closed fields. As applications of the Real

Nullstellensatz, we shall give the geometric criteria for an affine algebra to be semireal, and respectively, real; see 6.2, 6.10.

Again, because of the modest nature of these notes, our exposition will only barely touch the surface of the fast growing subject of real algebraic geometry. It is hoped, nevertheless, that this exposition will give the reader a picture of the beginning steps in the development of the foundations of this subject. For a more thorough treatment, we refer our readers to [15] and [8].

We begin by recalling the famous 17th Problem of Hilbert. To simplify the presentation, we shall consider this problem only over real closed coefficient fields. (The generalization to arbitrary ordered fields is relatively straightforward.) Therefore, throughout this section, k shall denote a real closed field.

Let f be a polynomial over k in n variables $x = (x_1, \ldots, x_n)$. We say that f is positive semidefinite if $f(k^n) \ge 0$ in the unique ordering of k. It was known to Hilbert that, if f is positive semidefinite, it need not be a sum of squares in $k[x] = k[x_1, \ldots, x_n]$. (It is easy to show that the Motzkin polynomial $f(x_1, x_2) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1$ gives such an example.) In view of this, Hilbert asked, as (part of) his 17th Problem, the following: if $f \in k[x] = k[x_1, \ldots, x_n]$ is positive semidefinite, must it be a sum of squares in the rational function field $k(x) = k(x_1, \ldots, x_n)$? The affirmative solution to this Problem was given by Artin in 1927, i.e.,

THEOREM 6.1. If $f \in k[x]$ is positive semidefinite, then f is a sum of squares in k(x).

The following is a modern version of Artin's proof.

PROOF (of 6.1). Assume that $f \notin \sum F^2$, where F = k(x). By Artin's Criterion for sums of squares in a field, we know that there exists an ordering $P \subset F$ such that $f <_P 0$. Then $-f >_P 0$, so by Lang's Homomorphism Theorem 5.5(C), there exists a k-algebra homomorphism $\varphi: k[x] \to k$ such that $\varphi(-f) > 0$. Letting $a_i = \varphi(x_i) \in k$, we have

$$f(a_1, \ldots, a_n) = f(\varphi(x_1), \ldots, \varphi(x_n)) = \varphi(f(x_1, \ldots, x_n)) < 0,$$

a contradiction.

Our next goal is to try to formulate the Real Nullstellensatz. Consider a polynomial ideal $\mathfrak{A} \subset k[x] = k[x_1, \ldots, x_n]$. We shall write $V(\mathbf{A})$ for the algebraic set defined by \mathfrak{A} over the algebraic closure $k(\sqrt{-1})$ of k, and write $V_k(\mathfrak{A}) = V(\mathfrak{A}) \cap k^n$ for the set of "real points" in $V(\mathfrak{A})$. The affine algebra $A = k[x]/\mathfrak{A}$ is called the k-coordinate ring of $V_k(\mathfrak{A})$. The following result, known as the Weak Real Nullstellensatz, gives the geometric condition for A to be semireal. THEOREM 6.2. A is semireal if and only if $V_k(\mathfrak{A}) \neq \emptyset$, i.e., if and only if $V(\mathfrak{A})$ has at least one k-point.

PROOF. If $a = (a_1, \ldots, a_n)$ is a k-point, then evaluation at a gives a k-homomorphism from A to k. This implies that A is semireal, proving the "if" part. The "only if" part is just a reformulation of Lang's Homomorphism Theorem, i.e., if A is semireal, there exists a k-algebra homomorphism $\varphi: A \to k$, by 5.5(B). Setting $a_i = \varphi(x_i), a = (a_1, \ldots, a_n)$ is clearly a real point for \mathfrak{A} .

Note that the set $V_k(\mathfrak{A})$ of real points is in one-one correspondence with the set of maximal ideals of A which are real. Thus, in essence, the Weak Real Nullstellensatz says that, in Theorem 2.3, we can add the condition (5') to the list of equivalent conditions for A to be semireal, in the case when A is a k-affine algebra.

For any subset S of k^n , let I(S) denote the ideal of polynomials $f \in k[x]$ such that f(S) = 0. With this notation, we can now state the Real Null-stellensatz in the prime case.

THEOREM 6.3. Let \mathfrak{p} be a prime ideal in k[x]. Then $I(V_k(\mathfrak{p})) = \mathfrak{p}$ if and only if \mathfrak{p} is real.

PROOF. If \mathfrak{p} is not real, there will exist $f_1^2 + \cdots + f_m^2 \in \mathfrak{p}$, with each $f_i \notin \mathfrak{p}$. But clearly $f_i \in I(V_k(\mathfrak{p}))$, so $I(V_k(\mathfrak{p})) \supseteq \mathfrak{p}$.

Conversely, assume p is real. We want to prove that $f \notin p f \notin I(V_k(p))$. Consider the affine algebra A = k[x]/p. Since this is a real domain in which $\bar{f} \neq 0$, Lang's Homomorphism Theorem 5.5(A) gives a k-algebra homomorphism $\varphi: A \to k$ such that $\varphi(\bar{f}) \neq 0$. If $a_i = \varphi(x_i)$, we have $(a_1, \ldots, a_n) \in V_k(p)$ and $f(a_1, \ldots, a_n) \neq 0$, so $f \notin I(V_k(p))$, as desired.

In order to formulate the more general version of the Real Nullstellensatz, we need to develop the notion of the real radical of an ideal. For this, we do not have to restrict ourselves to affine algebras. Therefore, in the next few paragraphs, A will stand for any commutative ring.

DEFINITION 6.4. Let \mathfrak{A} be any ideal in A. We define the real radical of \mathfrak{A} to be

r-rad
$$\mathfrak{A} = \{a \in A : \exists m \ge 0 \text{ and } \sigma \in \sum A^2 \text{ such that } a^{2m} + \sigma \in \mathfrak{A}\}.$$

Observe that, if \mathfrak{A} is a real ideal, then *r*-rad $\mathfrak{A} = \mathfrak{A}$. In fact, if $a^{2m} + \sigma \in \mathfrak{A}$ as above, then $a^m \in \mathfrak{A}$, but since A/\mathfrak{A} is reduced, we have $a \in \mathfrak{A}$.

Recall that the usual radical of an ideal in a commutative ring is equal to the intersection of all the prime ideals containing it. Let us now prove that this fact has a nice analogue for the real radical:

THEOREM 6.5. Let \mathfrak{A} be any ideal in A. Then r-rad \mathfrak{A} is equal to the

intersection \mathfrak{A} of the real primes containing \mathfrak{A} , and is the smallest real ideal containing \mathfrak{A} .

PROOF. Let \mathfrak{B} be any real ideal containing \mathfrak{A} . Then *r*-rad $\mathfrak{A} \subset r$ -rad $\mathfrak{B} = \mathfrak{B}$. In particular, *r*-rad $\mathfrak{A} \subset \widetilde{\mathfrak{A}}$. Next, we shall show that

(6.6) $a \notin r$ -rad $\mathfrak{A} \Rightarrow \exists$ a real prime $\mathfrak{p} \supset \mathfrak{A}$ such that $a \notin \mathfrak{p}$.

This will show that r-rad $\mathfrak{A} = \mathfrak{A}$. Since \mathfrak{A} is clearly real, this will also show that r-rad \mathfrak{A} is the smallest real ideal containing \mathfrak{A} .

To prove (6.6), consider the localization $A' = A[a^{-1}]$ and let $(\mathfrak{A}) = A' \cdot \mathfrak{A}$ be the ideal generated by \mathfrak{A} in A'. We claim that this ideal is semireal. In fact, if otherwise, we would have an equation $(b_1/a^n)^2 + \cdots + (b_r/a^n)^2 + 1 = c/a^n \in A'$, where $b_i \in A$, $c \in \mathfrak{A}$, and n is a sufficiently large integer. Pulling back this equation to A, we have, for some integer s, $a^{2s}(b_1^2 + \cdots + b_r^2 + a^{2n}) = a^{2s} \cdot a^n c$ in A. Since $a^{2s+n}c \in \mathfrak{A}$, this implies that $a \in r$ -rad A, a contradiction. Therefore (\mathfrak{A}) is semireal. By 2.3 there exists a real prime \mathfrak{p}' in A' containing (\mathfrak{A}). The preimage \mathfrak{p} of \mathfrak{p}' under the localization map is a real prime in A containing \mathfrak{A} . Since a/1 is a unit in A', we have $a/1 \notin \mathfrak{p}'$. Hence $a \notin \mathfrak{p}$ so \mathfrak{p} is what we looked for in 6.6.

Another, slightly different, argument to prove 6.6 may be given as follows. Given a as in 6.6, let S be the set $\{a^{2m} + \sigma \colon m \ge 0, \sigma \in \Sigma\}$, where $\Sigma = \Sigma A^2$. This is a multiplicative set of A satisfying $S + \Sigma \subset S$. Since $a \notin r$ -rad \mathfrak{A} , \mathfrak{A} is disjoint from S. By 1.2, (and Zorn's Lemma), we can enlarge \mathfrak{A} to a real prime \mathfrak{A} disjoint from S. In particular, $a \notin \mathfrak{p}$, as desired.

We are now ready to prove the following Dubois-Risler Real Nullstellensatz.

THEOREM 6.7. ([14], [26], [15]) Let k be a real closed field, and \mathfrak{A} be an ideal in $A = k[x_1, \ldots, x_n]$. Then $I(V_k(\mathfrak{A}) = r\text{-rad } \mathfrak{A})$.

PROOF. The inclusion " \supset " is easy. For, if $f \in r$ -rad \mathfrak{A} , then $f^{2m} + g \in \mathfrak{A}$ for some $m \ge 0$ and some $g \in \sum A^2$. Then for any point $a \in V_k(\mathfrak{A}), f^{2m}(a) + g(a) = 0 \in k$ implies f(a) = 0, so $f \in I(V_k(\mathfrak{A}))$.

Conversely, let p be any real prime $\supset \mathfrak{A}$. Then

$$V_k(\mathfrak{p}) \subset V_k(\mathfrak{A}) \Rightarrow I(V_k(\mathfrak{p})) \supset I(V_k(\mathfrak{A})).$$

By 6.3, this implies that $\mathfrak{p} \supset I(V_k(\mathfrak{A}))$, and therefore, by 6.5, we get *r*-rad $\mathfrak{A} \supset I(V_k(\mathfrak{A}))$, as desired.

If $\mathfrak{A} \subset k[x_1, \ldots, x_n]$ is a semireal ideal, we know that $V_k(\mathfrak{A})$ is nonempty, by the Weak Real Nullstellensatz. However, $V_k(\mathfrak{A})$ may be too small in general to reflect any of the geometric properties of $V(\mathfrak{A})$. The most extreme example is, of course, $\mathfrak{A} = (x_1^2 + \cdots + x_n^2)$: here, $V_k(\mathfrak{A})$ consists of the origin, which says nothing about the properties of the interesting

hypersurface $V(\mathfrak{A})$ over the algebraic closure of k. However, if \mathfrak{A} is not only semireal but real, then $V_k(\mathfrak{A})$ turns out to be a "significant part" of $V(\mathfrak{A})$. We now deduce the following consequence of the Real Nullstellensatz.

COROLLARY 6.8. Let \mathfrak{A} be a real ideal in $k[x] = k[x_1, \ldots, x_n]$. Then $V_k(\mathfrak{A})$ is Zariski dense in the algebraic set $V(\mathfrak{A})$.

PROOF. Let $\bar{k} = k(i)$ $(i = \sqrt{-1})$ be the algebraic closure of k. Our job is to show that any polynomial $f \in \bar{k}[x]$ vanishing on $V_k(\mathfrak{A})$ must also vanish on $V(\mathfrak{A})$. Write $f = f_1 + if_2$ where $f_j \in k[x]$. Clearly, $f(V_k(\mathfrak{A})) = 0$ implies that the f_j 's also vanish on $V_k(\mathfrak{A})$, so $f_j \in I(V_k(\mathfrak{A})) = r$ -rad \mathfrak{A} . Since \mathfrak{A} is real, r-rad $\mathfrak{A} = \mathfrak{A}$. Therefore $f_j \in \mathfrak{A}$ and so $f = f_1 + if_2$ vanishes on $V(\mathfrak{A})$.

REMARK 6.9. Suppose p is a real prime in k[x]. Then $p\bar{k}[x]$ is also a prime ideal in $\bar{k}[x]$, and hence V(p) is an irreducible affine variety over \bar{k} . To see that $p\bar{k}[x]$ is prime, let $(a + bi)(c + di) \in p\bar{k}[x]$ where $a, b, c, d \in k[x]$. Then $ac - bd \in p$ and $ad + bc \in p$. From these, we get easily $b(d^2 + c^2) \in p$ and $a(c^2 + d^2) \in p$. Suppose $a + bi \notin p\bar{k}[x]$. Then one of a, b is not in p and so $c^2 + d^2 \in p$. Since p is real, we have $c, d \in p$ and so $c + di \in p\bar{k}[x]$. (In the literature, a set of the form $V_k(p)$ discussed above is often called a real algebraic variety; by 6.8, this set is Zariski dense in the irreducible variety V(p) over \bar{k} .)

Recall that the Weak Real Nullstellensatz 6.2 is a criterion for an affine algebra to be semireal. As our next goal, we shall try to establish a corresponding criterion for an affine algebra to be real.

THEOREM 6.10. (The Simple Point Criterion) Let \mathfrak{A} be an ideal in k[x], and let $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ be the minimal primes above it. Then the affine algebra $A = k[x]/\mathfrak{A}$ is real if and only if \mathfrak{A} is a radical ideal, and each $V(\mathfrak{p}_i)$ has a real simple point.

PROOF. In view of 2.9, it is enough to treat the case when **A** is a prime, say $\mathfrak{A} = \mathfrak{p}$. First, assume $V(\mathfrak{p})$ has a real simple point *a*. Let \mathfrak{M} be the maximal ideal of *A* corresponding to *a*. Then $A_{\mathfrak{M}}$ is a regular local ring. Also, since $A_{\mathfrak{M}}/\mathfrak{M}A_{\mathfrak{M}} \cong A/\mathfrak{M} \cong k$, *A* is residually real. By 2.7, we see that $A_{\mathfrak{M}}$ is real, and therefore *A* is real.

Conversely, assume A is real, i.e., p is a real prime. Then, by 6.9, V(p) is an irreducible algebraic variety. The set S of simple points in V(p) is the complement of the singular locus, which is a closed set of codimension ≥ 1 in V(p). Therefore, $S \subset V(p)$ is a nonempty Zariski open set. Since $V_k(p)$ is Zariski dense in V(p), we have $S \cap V_k(p) \neq \emptyset$.

In the next result, we shall try to give an interesting criterion for a principal ideal in a polynomial ring to be real. To state this criterion, we

need a definition: a polynomial $f(x) \in k[x] = k[x_1, \ldots, x_n]$ is said to be indefinite if there exist points $a, b \in k^n$ such that f(a) < 0 < f(b) in the unique ordering of k. (In other words, f is neither positive semidefinite nor negative semidefinite.) The following result of Dubois and Efroymson [15] gives a criterion for the coordinate ring of a hypersurface to be real.

THEOREM 6.11. (The Changing Sign Criterion) Let $f \in k[x]$ be a nonzero polynomial. Then f generates a real ideal if and only if f is a square-free product of irreducible indefinite polynomials.

PROOF. It is possible to deduce this theorem as a consequence of the Simple Point Criterion. However, in the following, we shall give a proof of it without appealing to this Criterion. First, we note that, by 2.9, it is enough to prove the theorem in the case when f is an irreducible polynomial. Assume that f is not indefinite. After replacing f by -f if necessary, we may assume that f is positive semidefinite, so by 6.1 there exists a polynomial equation

(6.12)
$$h^2 f = f_1^2 + \dots + f_m^2$$

where $h \neq 0$. Choose this equation such that the total degree of h is as small as possible. We claim that

(6.13) The f_i 's are not all divisible by f.

If we can prove this, then, in k[x]/(f), we have $\bar{f}_1^2 + \cdots + \bar{f}_m^2 = 0$ but \bar{f}_i are not all zero, so k[x]/(f) is not real. To prove our claim 6.13, assume, instead, that $f_i = fg_i$ $(1 \le i \le m)$ in k[x]. Then $h^2 f = f^2 \sum g_i^2$ and so $h = h_0 f$ for some h_0 . But then $h_0^2 f = \sum g_i^2$, contradicting the "minimal" choice of the equation 6.12.

Now assume that f is indefinite (and irreducible). To show that (f) is real, suppose $g_1^2 + \cdots + g_m^2 \in (f)$, say $\sum g_i^2 = fh$. From this we clearly have $V_k(f) \subset V_k(g_i)$ for all *i*. The desired conclusion that each $g_i \in (f)$ can be drawn from the following lemma.

LEMMA 6.14. Let f be an irreducible indefinite polynomial, and g another polynomial such that $V_k(f) \subset V_k(g)$. Then f divides g.

PROOF. After a suitable change of coordinates, we may assume that $f(a, b_1) < 0 < f(a, b_2)$ (in the unique ordering of k), where $a \in k^{n-1}$, and $b_1, b_2 \in k$. Let $R = k[x_1, \ldots, x_{n-1}]$ and F = qf(R). View f and g as polynomials in $t = x_n$ in the ring $R[t] \subset F[t]$. Assume that $f \nmid g$ in R[t] (=k[x]). We know that f remains irreducible in F[t] and f $\nmid g$ also in F[t]. Since F[t] is a PID, there exists an equation $f \cdot \varphi + g \cdot \gamma = 1$ where $\varphi, \gamma \in F[t]$. Write $\varphi = \varphi_0/h$ and $\gamma = \gamma_0/h$, where $\varphi_0, \gamma_0 \in R[t]$, and $0 \neq h \in R$. Then $f \cdot \varphi_0 + g \cdot \gamma_0 = h$. Choose a neighborhood V of a in k^{n-1} such that $f(V, b_1) < 0 < f(V, b_2)$. (We use the topology on k^{n-1} arising from the interval topology

798

of k.) For any $v \in V$, $f(v, b_1) < 0 < f(v, b_2)$ implies that $f(v, b_v) = 0$ for some b_v between b_1 and b_2 , by the Intermediate Value Theorem. By the hypothesis on g, we also have $g(v, b_v) = 0$. Therefore, $f \cdot \varphi_0 + g \cdot \gamma_0 = h$ 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 $h \equiv 0$, a contradiction.

If f is irreducible and indefinite, it is also possible to show that k[x]/(f) is real by constructing an ordering on its quotient field, i.e., the function field of the hypersurface $V_k(f)$. For the details, we refer the reader to [21, p. 82].

As an application of the Changing Sign Criterion, we shall deduce the following result of Lang which will complete our exposition of the Artin-Lang Theory in §5.

THEOREM 6.15. (Lang's Imbedding Theorem) Let K be a real function field of transcedence degree d over a real closed field k. Let R be a real closed field containing k such that $\operatorname{tr.d}_k R \ge d$. Then there exists a k-imbedding of K into R.

We shall give a proof of this result which is due to A. Prestel (see [17]). The proof will be preceded by an easy lemma.

LEMMA 6.16. Let P be an ordering on a rational function field F(t). Then $F(t)\setminus F$ is dense with respect to the interval topology of F(t).

PROOF. Consider an open interval $(f, f')_P$ with respect to P, where $f <_P f'$. We want to show that $(f, f')_P$ contains a nonconstant. We may assume that $(f + f')/2 \in F$ for otherwise we are done. After a translation, we may assume that (f + f')/2 = 0, so it is enough to find a nonconstant h in any $(-g, g)_P$, where $g \in P$. If $g \notin F$, we can take h = g/2. If $g \in F$, we may assume, without loss of generality, that $t \in P$, so $g^{-1} + t >_P g^{-1}$. But then $0 <_P (g^{-1} + t)^{-1} <_P g$ so we can take $h = (g^{-1} + t)^{-1} \notin F$.

PROOF (of 6.15). We shall use the notations in the statement of 6.15. Let x_1, \ldots, x_d be a k-transcendence base of K/k and write $K = k(x_1, \ldots, x_d, \alpha)$. Let f(z) be the minimal polynomial of α over $k(x_1, \ldots, x_d)$. We may assume that $f \in k[x_1, \ldots, x_d, z]$ so $K = qf(k[x_1, \ldots, x_d, z]/(f))$. Since K is real, 6.11 implies that f is indefinite over k. As before, we may assume that $f(a_1, \ldots, a_d, b) < 0 < f(a_1, \ldots, a_d, b')$ for some $a_i, b, b' \in k$. (All inequalities will be with respect to the unique ordering P on the real closed field R.) Let $t_1 \in R$ be a transcendental element over k and consider $k(t_1)$ with the ordering $\overline{P_1} = P \cap k(t_1)$. By the continuity of f with respect to the interval topology of P_1 , and by 6.16, we can find a $y_1 \in k(t_1) \setminus k$ such that $f(y_1, a_2, \ldots, a_d, b) < 0 < f(y_1, a_2, \ldots, a_d, b')$. Now let L_1 be the algebraic closure of $k(y_1)$ in R and let $t_2 \in R$ be a transcendental element

over L_1 . Applying the same argument to $L_1(t_2)$, we can find a $y_2 \in L_1(t_2) \setminus L_1$ such that $f(y_1, y_2, a_3, \ldots, a_d, b) < 0 < f(y_1, y_2, a_3, \ldots, a_d, b')$. Proceeding like this, we can find $y_i \in R$ $(1 \le i \le d)$ such that

(6.17)
$$f(y_1, \ldots, y_d, b) < 0 < f(y_1, \ldots, y_2, b')$$

where y_1, \ldots, y_d are algebraically independent over k. (The fact that $\operatorname{tr.} d_k R \geq d$ is used to guarantee that we can carry out this construction to the very end.) Since R is real closed, 6.17 implies that $f(y_1, \ldots, y_d, \beta) = 0$ for some $\beta \in R$. Therefore we have a k-isomorphism $K \cong k(y_1, \ldots, y_d, \beta) \subset R$ by $x_i \mapsto y_i$, and $\alpha \mapsto \beta$.

Note that the proof above used merely the "only if" part of 6.11, which follows easily from 6.1. On the other hand, 6.1 is a direct consequence of Lang's Homomorphism Theorem. Therefore, the Imbedding Theorem is independent of the main body of this section (the Real Nullstellensatz, etc.), and could very well have been given in §5. The reason we have not done the Changing Sign Criterion earlier is that we want to have it in this section in parallel with the Simple Point Criterion 6.10.

7. Abstract "Stellensätze" for the real spectrum. In the last section we have presented Artin's solution of Hilbert's 17th Problem and the Dubois-Risler Real Nullstellensatz. These are results about polynomial functions on the real euclidean space k^n and their real subvarieties, where k is a realclosed field. In this section, we shall show that it is possible to take a more abstract viewpoint by passing over the real varieties and working directly with certain subsets of the real spectrum of a ring A. (The idea here is similar to the one in algebraic geometry whereby one works with affine schemes instead of affine varieties.) Since it is meaningful to talk about the sign behavior and the vanishing of elements of A on subsets Q in the real spectrum X_A , we can try to give the criteria for a given element $\ell \in A$ to be positive, non-negative, or zero on Ω . Such criteria may then be called, respectively, the (abstract) Positivstellensatz, Nichtnegativstellensatz, and the (real) Nullstellensatz. In the next section, we shall show that these abstract "Stellensätze", when used in conjunction with Lang's Homomorphism Theorem, will give corresponding concrete theorems for polynomial functions which are positive, non-negative or zero on certain kinds of semialgebraic sets. These concrete versions, in particular, will subsume (and generalize) the solution of Hilbert's 17th Problem and the Dubois-Risler Real Nullstellensatz. Our treatment in this section follows largely the suggestions of E. Becker and G. Brumfiel, and is closely modeled upon Brumfiel's lecture notes at the University of Hawaii in Fall, 1982. For a similar treatment in the geometric case, see [10].

We shall begin by setting up some notations. Let A be a ring, L be a subset of A, and Ω be a subset of the real spectrum X_A . We shall write

 $L >_{\Omega} 0$ (or L > 0 on Ω) if, for any $\ell \in L$ and any $P \in \Omega$, we have $\ell >_{P} 0$ (see (3.3)). Similarly, we can define $L \ge_{\Omega} 0$ and $L =_{\Omega} 0$. Given $L \subset A$, we also define the following three subsets of X_A :

(7.1)
$$\begin{cases} \mathscr{U}(L) \coloneqq \{P \in X_A : L > 0\} = \bigcap_{\ell \in L} H(\ell), \\ \mathscr{W}(L) \coloneqq \{P \in X_A : L \ge 0\} = \{P \in X_A : L \subset P\} = X_A \setminus \bigcup_{\ell \in L} H(-\ell), \\ \mathscr{V}(L) \coloneqq \{P \in X_A : L \ge 0\} = \{P \in X_A : L \subset \operatorname{supp}(P)\} \\ = \mathscr{W}(L) \cap \mathscr{W}(-L). \end{cases}$$

Note that in case L is finite, these are all constructible subsets in X_A ; moreover, $\mathcal{U}(L)$ will be open. In any case, $\mathcal{W}(L)$ and $\mathcal{V}(L)$ are both closed in X_A . We also note the following:

(1) $\mathscr{U}(L)$ is unchanged if we replace L by the multiplicative set generated by L.

(2) $\mathscr{W}(L)$ is unchanged if we replace L by the semiring generated by L and A^2 .

(3) $\mathscr{V}(L)$ is unchanged if we replace L by the ideal in A generated by L. Therefore, in dealing with $\mathscr{U}(L)$ (resp. $\mathscr{W}(L)$, $\mathscr{V}(L)$), there is no loss of generality in assuming that such replacements are made. In order to prove the various Stellensätze, it will be important for us to understand the relationship of the sets $\mathscr{U}(L)$, $\mathscr{W}(L)$ and $\mathscr{V}(L)$ to the functorial behavior of the real tpectrum. We shall now describe this relationship below.

(a) (Residue rings) Let *I* be an ideal in *A* and φ be the natural projection $A \to A/I$. It is easy to check that the induced map $\varphi^* \colon X_{A/I} \to X_A$ is injective, and maps $X_{A/I}$ homeomorphically onto $\mathscr{V}(I)$. We shall henceforth identify the former with the latter.

(b) (Localization) Let S be a multiplicative set in A, and φ be the localization map $A \to A_S$. Again, it is easy to check that φ^* is injective, and maps X_{A_S} homeomorphically onto $\{P \in X_A : \forall s \in S, s \neq_P 0\}$. (Any ordering P in the latter set extends to an ordering $P_S := \{p/s^2: p \in P, s \in S\}$ in A_S .) Again, we shall identify X_{A_S} with its image in X_A . This image is compact, and, in case S is finitely generated, it is also open.

(c) (Square root adjunction) Let $L \subset A$ and B be the ring obtained by adjoining formal square roots of all elements $\ell \in L$ to A; in other words, $B = A[x_{\ell}: \ell \in L]/(\{x_{\ell}^2 - \ell: \ell \in L\})$. One can check that the map $\varphi^*: X_B \to X_A$ induced by $\varphi: A \to B$ has image $\mathscr{W}(L) = \mathscr{W}_A(L)$. Here, the inclusion $\operatorname{im}(\varphi^*) \subset \mathscr{W}_A(L)$ is easy since $L \subset B^2$. The reverse inclusion requires some work; since we will not need this result, its proof will be left to the reader.

To formulate the various abstract Stellensätze for the real spectrum, we shall fix the following notations:

(7.2)
$$\begin{cases} \{f_{\alpha}\}, \{g_{\beta}\}, \{h_{\gamma}\}: \text{subsets of } A, \\ S = \text{ multiplicative set in } A \text{ generated by } \{f_{\alpha}\}, \\ T = \text{ subsemiring of } A \text{ generated by } \{f_{\alpha}, g_{\beta}\} \text{ and } A^{2}, \\ I = \text{ ideal in } A \text{ generated by } \{h_{\gamma}\}, \\ \mathcal{Q} = \mathscr{U}\{f_{\alpha}\} \cap \mathscr{W}\{g_{\beta}\} \cap \mathscr{V}\{h_{\gamma}\} = \mathscr{U}(S) \cap \mathscr{W}(T) \cap \mathscr{V}(I). \end{cases}$$

Given an element $\ell \in A$, we shall seek necessary and sufficient conditions for ℓ to be positive (resp. non-negative, zero) on Ω . We shall first write down certain conditions which are easily checked to be sufficient; the goal of the various Stellensätze will then be to prove that these conditions are also necessary.

PROPOSITION 7.3. (1) $\ell \in A$ is positive on Ω if there is a congruence $t \cdot \ell \equiv s + t' \pmod{I}$ where $t, t' \in T$ and $s \in S$.

(2) $\ell \in A$ is non-negative on Ω if there is a congruence $t \cdot \ell \equiv s \cdot \ell^{2e} + t' \pmod{I}$ where $t, t' \in T, s \in S$ and $e \geq 0$ is an integer.

(3) $\ell \in A$ is zero on Ω if there is a congruence $s \cdot \ell^{2e} + t \equiv 0 \pmod{I}$ where $t \in T$, $s \in S$ and $e \ge 0$ is an integer.

PROOF. (1) For any $P \in \Omega$, we have $s + t' >_P 0$ and $I =_P 0$, so $t \cdot t >_P 0$. Since $t \ge_P 0$, we must have $t >_P 0$.

(2) Let $P \in \Omega$. If $\ell <_P 0$, then $s \cdot \ell^{2e} + t' >_P 0$ and so $t \cdot \ell >_P 0$, a contradiction. Therefore we must have $\ell \ge_P 0$.

(3) Let $P \in \Omega$. Then $s \cdot \ell^{2e} + t = P^{0}$. Since $s > P^{0}$ and $t \ge P^{0}$, this clearly implies that $\ell = P^{0}$.

We are now ready to state the three abstract Stellensätze. They provide the converse to (3) and strong converses to (1) and (2) in Proposition 7.3.

THEOREM 7.4. (1) (Positivstellensatz) Let $\ell > 0$ on Ω . Then there is a congruence $(s + t) \cdot \ell \equiv s + t' \pmod{I}$, where $t, t' \in T$ and $s \in S$.

(2) (Nichtnegativstellensatz) Let $\ell \ge 0$ on Ω . Then there is a congruence $(s \cdot \ell^{2e} + t) \cdot \ell \equiv s \cdot \ell^{2e} + t' \pmod{I}$, where $t, t' \in T, s \in S$ and $e \ge 0$ is an integer.

(3) (Nullstellensatz) Let $\ell = 0$ on Ω . Then there is a congruence $s \cdot \ell^{2e} + t \equiv 0 \pmod{I}$, where $t \in T$, $s \in S$ and $e \ge 0$ is an integer.

The proof of (1), (2) and (3) will be based on the following crucial lemma.

LEMMA 7.5. Let $\ell > 0$ on $\mathcal{W}(T)$, where T is a subsemiring of A containing A^2 . Then there exists an equation $(1 + t_1) \cdot \ell = 1 + t_2$ where $t_1, t_2 \in T$.

PROOF. If $-1 \in T$, we can take $t_1 = t_2 = -1$; therefore, we may assume that $-1 \notin T$, i.e., T is a preordering. If the asserted equation does not exist, then by 3.4 (2), $T + T \cdot (-\ell)$ is also a preordering. Let P be an

802

ordering containing it. Then $P \in \mathscr{W}(T)$ but $-\ell \geq_P 0$, a contradiction. Therefore, the asserted equation must exist.

We now go back to the notation of 7.4 and give its proof in three parts.

PROOF (of 7.4). (Positivstellensatz) Let $B = A_S/I \cdot A_S$. As explained before, we can identify X_B with the subspace $\mathscr{V}(I) \cap \{P \in X_A : \forall s \in S, s \neq_P 0\}$. From this it is easy to see that $Q \cap X_B = \mathscr{W}_B(T_B)$, where T_B is the semiring $\{t/s^2 + I \cdot A_S : t \in T, s \in S\}$ in B. Therefore $\ell + I \cdot A_S > 0$ on $\mathscr{W}_B(T_B)$. By the Lemma above (applied to B and T_B), we have an equation $(1 + t_1) \cdot \ell = 1 + t_2 \in B$ where $t_1, t_2 \in T_B$. Clearing denominators, we get a congruence of the type claimed in the Positivstellensatz.

(Nichtnegativstellensatz) For *B* as above, let $B' = B[\ell^{-1}]$ and $T_{B'} = \{t/\ell^{2n}: t \in T_B, n \ge 0\}$. As before, we can identify $X_{B'}$ with the subspace $X_B \cap \{P \in X_A \colon \ell \neq_P 0\}$ in X_A . We check easily that $\Omega \cap X_{B'} = \mathscr{W}_{B'}(T_{B'})$, and so the hypothesis on ℓ gives that $\ell > 0$ on $\mathscr{W}_{B'}(T_{B'})$. By 7.5 (applied to B' and $T_{B'}$), we have an equation $(1 + t_1') \cdot \ell = 1 + t_2' \in B'$, where t_1' , $t_2' \in T_{B'}$. Clearing denominators, we get a congruence of the type claimed in the Nichtnegativstellensatz.

(Nullstellensatz) By the hypothesis, we have both $\ell \ge 0$ and $-\ell \ge 0$ on Ω . Applying the Nichtnegativstellensatz, we have therefore two congruences.

 $t_1 \cdot \ell \equiv s \cdot \ell^{2n} + t_2 \pmod{I}, \qquad t_3 \cdot (-\ell) \equiv s'(-\ell)^{2m} + t_4 \pmod{I}.$

Multiplying them together, we get a congruence

$$s \ s' \cdot \ell^{2n+2m} + t_5 + t_1 t_3 \cdot \ell^2 \equiv 0 \pmod{I},$$

with $t_5 \in T$, as claimed in the Nullstellensatz.

Let us now examine a few special cases of 7.4. Let $I \subset A$ be an ideal, and $\Omega = \mathscr{V}(I) \subset X_A$. The set $\{\mathscr{I} \in A : \mathscr{I} = 0 \text{ on } \Omega\}$ is precisely the intersection of all the real prime ideals containing *I*. By the Nullstellensatz, this can be characterized as $\{\mathscr{I} \in A : \mathscr{I}^{2e} + t \in I \text{ for some } t \in \Sigma A^2\}$, i.e., the real radical of *I*. Thus, we get back the earlier result Theorem 6.5.

For another illustration of 7.4, let T be a preordering in A. We would like to determine \tilde{T} , the intersection of all the orderings containing T. In the classical case when A is a field, it is well-known that $\tilde{T} = T$, but if A is just a ring, this need not be the case. To determine \tilde{T} , let $\mathcal{Q} = \mathcal{W}(T)$. Then \tilde{T} is just $\{\ell \in A : \ell \geq 0 \text{ on } \mathcal{Q}\}$. According to the Nichtnegativstellensatz, we get

(7.6)
$$\widetilde{T} = \{ \ell \in A : t' \cdot \ell = \ell^{2e} + t \text{ for some } t, t' \in T \}$$
$$= \{ \ell \in A : \ell^{2e+1} + t \cdot \ell \in T \text{ for some } t \in T \}.$$

This preordering is called the radical of T. In general, \tilde{T} may not be equal

to *T*. To give an example of this, take the preordering $T^{(2)}$ in the polynomial ring $A = \mathbf{R}[x]$ constructed at the end of §3. Recall that $T^{(2)}$ consists of polynomials $\alpha_0 + \alpha_1 x + \cdots + \alpha_r x^r$ where either $\alpha_0 = \alpha_1 = 0$, or $\alpha_0 > 0$, or $\alpha_0 = 0$ and $\alpha_1 > 0$. Write $T = T^{(2)}$ and consider $\ell = -x \in A$. We have $\ell^3 \in T$, so $\ell \in \tilde{T}$, sut $\ell \notin T$. Therefore, $T \subsetneq \tilde{T}$. As a matter of fact, there is precisely one ordering containing *T*, namely, $T^{(1)}$, so $\tilde{T} = T^{(1)}$.

COROLLARY 7.7. An element $\ell \in A$ lies in all the orderings of A if and only if ℓ satisfies an equation $\ell^{2e+1} + \sigma \cdot \ell = \sigma'$, where $\sigma, \sigma' \in \sum A^2$.

PROOF. If A is semireal, this follows from the above by taking $T = \sum A^2$. If A is not semireal, we can write down an equation as in 7.7 by noting that $\ell^{2e+1} + (-1)\ell^{2e} \cdot \ell = 0$.

8. Semialgebraic geometry: the beginning steps. The study of semialgebraic sets in an euclidean space k^n over an ordered field (k, P) is called Semialgebraic Geometry. Roughly speaking, a semialgebraic set is a set in k^n which can be defined by finitely many polynomial equalities and inequalities. In this section, we shall give a short introduction to the study of these sets. In particular, we shall try to explain the relationship between semialgebraic sets and constructible sets in the real spectrum, from which we shall deduce various concrete (or geometric) forms of the Positivstellensatz, Nichtnegativstellensatz and Nullstellensatz for polynomial functions over semialgebraic sets. For convenience of exposition, we shall again assume, throughout, that k is a real-closed field, given its unique ordering (k, k^2) .

Let A be a (semireal) ideal in $k[x] = k[x_1, ..., x_n]$, and let $V = V_k(\mathbf{A})$ be the algebraic set of k-points defined by A. For any finite set of polynomials f_i ($1 \le i \le r$), we define the set

$$U(f_1, \ldots, f_r) := \{a \in k^n : f_i(a) > 0 \text{ for all } i\}$$
$$= U(f_1) \cap \cdots \cap U(f_r).$$

Consider the sets $U(f) \cap V$, where f ranges over all polynomials $\in k[x]$. We let $\mathcal{S}(V)$ be the class of subsets of V which can be obtained from the sets $U(f) \cap V$ by using (a finite number of) boolean operations, i.e., forming finite unions and intersections, and taking complements in V. Note, in particular, that sets of the following form are all in $\mathcal{S}(V)$:

$$\{a \in V: g(a) \ge 0\} = V \setminus \{a \in V: -g(a) > 0\}, \\ \{a \in V: g(a) = 0\} = \{b \in V: g(b) \ge 0\} \cap \{c \in V: -g(c) \ge 0\}, \\$$

where $g \in k[x]$.

DEFINITION 8.1. Sets in the family $\mathscr{S}(V)$ are called semialgebraic subsets of V. Sets of the form $U(f_1, \ldots, f_r) \cap V$ are called basic open semialgebraic subsets of V. (Note that the latter sets are, indeed, open in the strong topology of V, i.e., the topology induced by the interval topology of k.)

We make the following easy observation.

PROPOSITION 8.2. $\mathcal{S}(V)$ is just the family \mathcal{S} of finite unions of sets of the form

$$\{a \in V: g_i(a) = 0 \ (1 \le i \le s), f_j(a) > 0 \ (1 \le j \le r) \},$$

where $r, s \ge 0$, and $g_i, f_j \in k[x]$.

PROOF. It suffices to show that

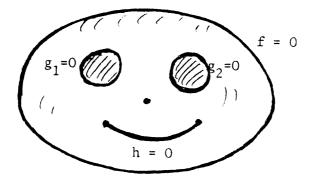
- (i) $A \in \mathcal{S}, B \in \mathcal{S} \Rightarrow A \cup B \in \mathcal{S},$
- (ii) $A \in \mathcal{S}, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S},$
- (iii) $A \in \mathcal{S} \Rightarrow V \setminus A \in \mathcal{S}$.

The first two implications are obvious. For (iii), it suffices (in view of (ii)) to treat the case when A is a set of the form (*). But then $V \setminus A$ is just

$$\bigcup_{i} (\{a \in V : g_{i}(a) > 0\} \cup \{a \in V : -g_{i}(a) > 0\}) \\ \cup \bigcup_{j} (\{a \in V : f_{j}(a) = 0\} \cup \{a \in V : -f_{j}(a) > 0\})$$

which is clearly in \mathcal{S} .

Note that, if we want, we could have expressed a set of the form (*) with s = 1. In fact, the conditions $g_i(a) = 0$ for $1 \le i \le s$ can be expressed by a single equation $g_1(a)^2 + \cdots + g_s(a)^2 = 0$, since k is formally real. The following delightful picture of a semialgebraic set in k^2 first appeared in Brumfiel's book [8, p. 174] and is reproduced here with his kind permission.



The planar semialgebraic set defined by $f \leq 0, g_1 \geq 0, g_2 \geq 0, h \neq 0$, and $(x, y) \neq 0$.

It turns out that the semialgebraic subsets of V are intimately related to the constructible sets in the real spectrum X_A of the affine algebra A = k[x]/A defined in 4.13. In order to explain this, we shall have the opportunity to see how the real spectrum can be brought to bear on the study of the semialgebraic geometry of V.

Let a be any point of V, and let \mathfrak{M}_a be the maximal ideal of A corresponding to a. By pulling back the unique ordering on $A/\mathfrak{M}_a \cong k$, we obtain an ordering on A with support \mathfrak{M}_a . We denote this ordering by $T_a \in X_A$; clearly $T_a = \{f \in A : f(a) \ge 0\}$. It is easy to see that the rule $a \mapsto T_a$ defines an injective mapping of V into X_A . Thus we may identify V with the subset $\{T_a : a \in V\}$ of X_A . Since each T_a is a maximal ordering, we have in fact $V \subset X_A^m$ the subspace of closed points in X_A .

Recall that the family $\mathscr{C}(A)$ of constructible sets in the real spectrum of A is obtained from the subbasic Harrison sets $\{H(f) = \mathscr{U}(f): f \in A\}$ in X_A by using (a finite number of) boolean operations. Since

(8.3)
$$\mathcal{U}(f) \cap V = \{a \in V : f >_{T_a} 0\}$$
$$= \{a \in V : f(a) > 0\}$$
$$= U(f) \cap V,$$

we see that the semialgebraic sets in V are just the contractions of the constructible sets of X_A to V. More precisely, we shall prove the following.

THEOREM 8.4. (1) The contraction map $\alpha \colon \mathscr{C}(A) \to \mathscr{G}(V)$ defined by $\alpha(C) = C \cap V$ (for any $C \in \mathscr{C}(A)$) is a one-one correspondence.

(2) The subset V is dense in X_A with respect to the Tychonoff topology, and therefore also with respect to the Harrison topology.

(3) The topology on V induced from the Harrison topology of X_A is precisely the strong topology on V.

PROOF. For (1), we need only show the injectivity of α . To this end it is enough to show that, for $C \in \mathscr{C}(A)$,

$$(*) C \neq \emptyset \Rightarrow C \cap V \neq \emptyset.$$

Indeed, if $\alpha(C_1) = \alpha(C_2)$, then the symmetric difference $C_1 \triangle C_2$ contracts to the empty set in V, so by (*) we have $C_1 \triangle C_2 = \emptyset$, i.e., $C_1 = C_2$. To show (*), we may assume, in view of 4.15, that

$$C = H(f_1) \cap \cdots \cap H(f_r) \cap H'(g_1) \cap \cdots \cap H'(g_s)$$

where $H'(g_j) = X_A \setminus H(g_j)$. If $C \neq \emptyset$, there is an ordering T on A such that $f_i >_T 0$ and $-g_j \ge_T 0$ for all *i*, *j*. By Lang's Homomorphism Theorem 5.5(C), there exists a k-algebra homomorphism $\varphi: A \to k$ such that $\varphi(f_i) > 0 \ge \varphi(g_j)$ for all *i*, *j*. Let $a_{\ell} = \varphi(\bar{x}_{\ell})$ and $a = (a_1, \ldots, a_n) \in V$. Then we have $f_i >_{T_a} 0 \ge_{T_a} g_j$, so $T_a \in H(f_1) \cap \cdots \cap H(f_r) \cap H'(g_1) \cap$

 $\dots \cap H'(g_s) = C$, showing that $C \cap V$ is nonempty. This proves (*), and shows that V is dense in X_A with respect to the Tychonoff topology; it follows that V is also dense in X_A with respect to the coarser Harrison topology \mathcal{T} . Finally, we have to prove (3). From 8.3, it follows that any open set of \mathcal{T} contracts to an open set in the strong topology of V. Conversely, if we consider a basic open set in the strong topology of V, say $U = \{a \in V: b_i < a_i < c_i \text{ for all } i\}$, then clearly $U = U(x_1 - b_1, \ldots, x_n - b_n, c_1 - x_1, \ldots, c_n - x_n) \cap V$. From 8.3, this set is open in the induced Harrison topology of V.

Note that the same proof also shows the following.

COROLLARY 8.5. Any semialgebraic set $S \in \mathcal{S}(V)$ is dense in the unique constructible set $C \in \mathcal{C}(A)$ corresponding to it under α , with respect to both the Tychonoff and Harrison topologies.

Recall that, by 4.7, X_{A}^{m} is a compact Hausdorff space. However, in general, the variety V with the strong topology may not be compact; for instance, when $k = \mathbf{R}$, V may be unbounded. Therefore, $V \subset X_A^m$ is not an equality in general. In view of Theorem 8.4 (2), (3), one may think of X_{A}^{m} as some sort of natural compactification of V. If V is already compact to begin with, then of course $V = X_A^m$; in this case, $\{T_a: a \in V\}$ are all the maximal orderings on A. For a simple example of $V \subsetneq X_A^m$, consider $\mathbf{A} = (0) \subset \mathbf{R}[x]$ (n = 1); here V is the "line" **R**. We have seen earlier that $A = \mathbf{R}[x]$ has two orderings which are maximal but are different from the orderings $\{T_a: a \in \mathbf{R}\}$. Therefore, X_A^m is a "two-point compactification" of the real line; symbolically, $X_A^m = \{-\infty\} \cup \mathbf{R} \cup$ $\{+\infty\}$. Note that, in many ways, this fits the real picture more nicely than the usual one-point compactification $\mathbf{R} \cup \{\infty\} = S^1$. For a more interesting example, consider A = (0) = k[x] (n = 1), where k is now the (real-closed) field of all real algebraic numbers. What are the maximal orderings on A = k[x]? Besides the obvious ones $T_a(a \in k)$, we must look for maximal orderings on k[x] with support (0). The question is, therefore, which orderings on k(x) will restrict to a maximal ordering on k[x]. As is well-known, the orderings on k(x) correspond to the different Dedekind cuts on k. Orderings $P \in X_{k(x)}$ whose corresponding cuts on k are "realized" by a point $a \in k$ will not restrict to a maximal ordering on k[x] because $P \cap k[x]$ will be properly contained in T_a . On the other hand, if $P \in X_{k(x)}$ defines a cut on k realized by a transcendental number $\alpha \in \mathbf{R} \setminus k$, i.e., $P = \{f(x) \in k(x) : f(\alpha) \ge 0\}$, then $P \cap k[x]$ is a maximal ordering on k[x]. (To see that $P \cap k[x] \in X^m$, assume, instead, that $P \cap$ $k[x] \subseteq T_a$ for some $a \in k$. First assume that $a < \alpha$ in **R**. Take $b \in k$ such that $a < b < \alpha$, and consider $x - b \in A$. Clearly, $x - b \in P \cap k[x]$, but $x - b \notin T_a$, a contradiction. If $\alpha < a$ in **R**, a similar contradiction will

result by considering $\alpha < b < a$ in **R**.) And, of course, there is a cut at $-\infty$, and a cut at $+\infty$. Therefore, in compactifying V = k to X_A^m , we are adding, essentially, the transcendental numbers, and $\pm\infty$. After checking the topology, one can conclude that the compactification X_A^m of k is, again, $\{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. The remarks and examples above are pointed out by E. Becker and help to illustrate the nature of the compactification X_A^m of a real variety V.

Next, we shall try to derive the various "Stellensätze" for polynomial functions on certain semialgebraic subsets of V. The notations associated with V introduced at the beginning of this section will remain in force. Along with the sets $U(f_1, \ldots, f_r) \subset k^n$, let us also define

$$W(g_1, ..., g_m) = \{ a \in k^n : g_i(a) \ge 0 \text{ for all } i \},\$$

$$V(h_1, ..., h_p) = \{ a \in k^n : h_i(a) = 0 \text{ for all } i \}.$$

Comparing these with the notations in 7.1, we have

$$W(g_1,\ldots,g_m) \cap V = \mathscr{W}(g_1,\ldots,g_m) \cap V,$$

$$V(h_1,\ldots,h_p) \cap V = \mathscr{V}(h_1,\ldots,h_p) \cap V.$$

Therefore, the two semialgebraic subsets of V on the left hand side correspond, respectively, to the constructible sets $\mathscr{W}(g_1, \ldots, g_m)$ and $\mathscr{V}(h_1, \ldots, h_p)$ in X_A . It follows that the semialgebraic set $\Omega_0 = U(f_1, \ldots, f_r) \cap W(g_1, \ldots, g_m) \cap V(h_1, \ldots, h_p) \cap V \subset V$ corresponds to the constructible set $\Omega = \mathscr{U}(f_1, \ldots, f_r) \cap \mathscr{W}(g_1, \ldots, g_m) \cap \mathscr{V}(h_1, \ldots, h_p) \subset X_A$.

THEOREM 8.6. Given the notation above, let S be the multiplicative set of A generated by the f_i 's, T be the subsemiring of A generated by A^2 , the f_i 's and g_i 's, and I be the ideal of A generated by the h_i 's. Let $\ell = \ell(x_1, \ldots, x_n) \in A$.

(1) (Positivstellensatz) If $\ell > 0$ on Ω_0 , then there is a congrucence $(s + t) \cdot \ell \equiv s + t' \pmod{I}$, where $t, t' \in T$ and $s \in S$.

(2) (Nichtnegativstellensatz) If $\ell \ge 0$ on Ω_0 , then there is a congruence $(s \cdot \ell^{2e} + t) \cdot \ell \equiv s \cdot \ell^{2e} + t' \pmod{I}$, where $t, t' \in T, s \in S$ and $e \ge 0$ is an integer.

(3) (Nullstellensatz) If $\ell = 0$ on Ω_0 , then there is a congruence $s \cdot \ell^{2e} + t \equiv 0 \pmod{I}$, where $t \in T$, $s \in S$ and $e \ge 0$ is an integer.

PROOF. These will follow from the corresponding abstract Stellensätze 7.4 if we can show that

(8.7)
$$\begin{cases} \ell > 0 \text{ on } \Omega_0 \Rightarrow \ell > 0 \text{ on } \Omega, \\ \ell \ge 0 \text{ on } \Omega_0 \Rightarrow \ell \ge 0 \text{ on } \Omega, \\ \ell = 0 \text{ on } \Omega_0 \Rightarrow \ell = 0 \text{ on } \Omega. \end{cases}$$

To see these, we exploit the fact that the one-one correspondence α in

8.4 (1) preserves inclusion relations. Suppose $\ell > 0$ on Ω_0 . Then we have $\Omega_0 \subset U(\ell) \cap V$. Since $\alpha(\Omega) = \Omega_0$ and $\alpha(\mathcal{U}(\ell)) = U(\ell) \cap V$, we must have $\Omega \subset \mathcal{U}(\ell)$, i.e., $\ell > 0$ on Ω . The other two implications in 8.7 are proved similarly.

Let us note several special cases of 8.6 and make some historical remarks to put the results in perspective. When there are no f_i 's and g_i 's, (3) gives back the Dubois-Risler Real Nullstellensatz (6.7). When $\mathfrak{A} = 0$ and there are no f_i 's, g_i 's or h_i 's, (2) gives a strengthened solution to Hilbert's 17th Problem, i.e., any positive semidefinite polynomial $\ell \in k[x_1, \dots, x_n]$..., x_n can be expressed in the form $(\ell^{2e} + t')/(\ell^{2e} + t)$ where t, $t' \in \sum k[x]^2$. Also, in the same case, (1) gives an expression for positive polynomials ℓ , namely $\ell = (1 + t')/(1 + t)$ where $t, t' \in \sum k[x]^2$. These refinements are due to A. Prestel and G. Stengle. When there are no g_i 's, h_i 's but we allow the f_i 's, (2) in coarser form gives A. Robinson's Nichtnegativstellensatz for basic open semialgebraic sets in a variety. When there are no f_i 's but we allow the g_i 's and h_i 's, (3) is the Semialgebraic Real Nullstellensatz found by G. Stengle [30]; when $\mathfrak{A} = 0$ and there are no f_i 's and h_i 's, (2) is the Nichtnegativstellensatz for basic closed semialgebraic sets of Stengle. Thus, 8.6 subsumes a large number of earlier results of its kind obtained by different authors over the years.

In the balance of this section, we shall describe a few further results in the recent literature on semialgebraic sets. We shall not give the detailed proofs of these results as the techniques of these proofs lie beyond the scope of our notes. In fact, for the several theorems to be discussed below, the shortest proofs found so far are all based on techniques of logic. To present these techniques in full will considerably lengthen this paper, so we shall content ourselves with a few brief remarks. This will give us an opportunity to at least say something about the important role played by logic in the recent investigations on semialgebraic geometry.

The basic connection between logic and semialgebraic geometry is provided by Tarski's Theorem on quantifier elimination over real closed fields. Stated roughly, this theorem says that, if φ is an elementary formula in the first order language of ordered fields, with quantifications over a real closed field k, then it is possible to eliminate one quantifier at a time, thus leading, eventually, to a logically equivalent quantifier-free formula. In particular, if we can prove a certain elementary statement over a real closed extension K containing k, then the same statement is guaranteed to hold over any real closed extension $K' \supset k$ (including k itself). This remarkable principle is, of course, very powerful when brought to bear upon questions concerning polynomial equalities and inequalities over a real closed field. In similar ways, other results in model theory can be applied profitably to the study of semialgebraic geometry. To give an example, consider a semialgebraic set Ω in the euclidean space k^n over a real closed field k. Essentially by definition, such a set is defined by a quantifier-free formula φ with n free variables. Consider Ω' , the projection of Ω to k^{n-1} . This set consists of points $(a_1, \ldots, a_{n-1}) \in k^{n-1}$ satisfying the formula " $\exists x_n \varphi(x_1, \ldots, x_n)$ ". By Tarski's Theorem, we can eliminate the quantifier " $\exists x_n$ " from this formula, to get another, equivalent, formula which is quantifier-free. Therefore, we have the following pleasant conclusion.

THEOREM 8.8. The projection of a semialgebraic set in k^n to k^{n-1} is again a semialgebraic set.

Similarly, one can use the elimination of quantifiers to show

THEOREM 8.9. If Ω is a semialgebraic set in k^n , then the closure, interior and boundary of Ω (with respect to the strong topology) are also semialgebraic subsets of k^n .

These results are certainly not obvious if we do not assume Tarski's Theorem on quantifier elimination!

As a second example, we shall consider the following result which was observed by L. Bröcker.

THEOREM 8.10. Let (k_0, P) be an ordered field, and k be its real closure. Let Ω be any semialgebraic set in k^n . Then Ω can be defined in k^n by (a finite number of) polynomial equalities and inequalities with coefficients from k_0 ,

At first sight, this may appear to be a somewhat unlikely result, but it is actually a true theorem! To give an illustrating example, let $k_0 = \mathbf{Q}$ with *P* its usual ordering, and let *k* be the field of real algebraic numbers. Consider the line in the *k*-plane given by $\Omega = \{(x, y) \in k^2 : x + y = \sqrt{2}\}$. To begin with, the coefficients used here are only in *k*, and not all in k_0 . However, it is easy enough to rewrite the definition of Ω using only coefficients from k_0 .

$$Q = \{(x, y) \in k^2 \colon (x + y)^2 = 2 \text{ and } x + y > 0\}.$$

As a second illustrating example, consider the case when n = 1 and Ω is a singleton set $\{\alpha\}$ ($\alpha \in k$). How is it possible to "define" Ω over k_0 ? Let $f(x) \in k_0[x]$ be the minimal polynomial of α and write down its roots in kin increasing order, i.e., $\alpha_1 < \cdots < \alpha_m$. We can then specify α by saying that it is, say, the r^{th} root in this sequence. Now, let us return to the general case. The defining polynomials for Ω may have coefficients $\{\beta_i\}$ in k, but we can "define" the β_i 's over k_0 by the method explained above. Therefore, Ω can be defined in k^n by a formula with parameters from k_0 —with a finite number of quantifiers. Applying Tarski's Theorem over k, we can replace this with a quantifier-free formula, still with parameters from k_0 . Therefore, Ω is now defined over k_0 ! For instance, to define the line $\Omega_1 := \{(x, y) \in k^2 : x + \sqrt{2} \ y = 1\}$ over k_0 , we eliminate the coefficient $\sqrt{2}$ by introducing an existential quantifier, i.e., $\Omega_1 = \{(x, y) \in k^2 : \exists w(w^2 = 2) \land (w > 0) \land (x + wy = 1)\}$. Then we can eliminate " $\exists w$ " by applying Tarski's Theorem over k, to get a description of Ω_1 over k_0 . We shall leave it as an exercise for the reader to find such an explicit description!

The statement of 8.10 was pointed out to me by A. Prestel, who also explained to me the idea of the proof sketched above. Bröcker's original proof, not using logic, is apparently much more complicated. For us, the important thing to note here is that Tarski's Theorem not only lends itself to an easy proof of 8.10 but also it shows in a very conceptual way why such a result should be true.

As a final example, we shall mention an important result concerning the structure of open (resp. closed) semialgebraic sets. Take a variety $V = V_k(\mathfrak{A})$ defined over a real-closed field k; let $A = k[x_1, \ldots, x_n]/\mathfrak{A}$ and use the notations introduced at the beginning of this section. Recall from 8.1 that sets of the form $U(f_1, \ldots, f_r) \cap V$ are called basic open semialgebraic subsets of V; they form a basis for the strong topology of V. Similarly, sets of the form $W(g_1, \ldots, g_m) \cap V$ are called basic closed semialgebraic subsets of V. We have the following remarkable

FINITENESS THEOREM 8.11. (1) Any open semialgebraic subset $U \subset V$ is a finite union of basic open semialgebraic subsets.

(2) Any closed semialgebraic subset $W \subset V$ is a finite union of basic closed semialgebraic subsets.

By taking complements in V and using the two distributive laws for unions and intersections, it is easy to see that (1) and (2) are equivalent results. However, the truth of either one is far from obvious. Let U be as in (1). We know by 8.2 that U is a finite union of sets of the form $U(f_1, \ldots, f_r) \cap V(f) \cap V$; however, these sets are not necessarily open, so although we are down to a finite union, the expression ignores the fact that U is open. On the other hand, expressing U as a finite union of basic open semialgebraic subsets reflects the openness of U as well as the fact that U is semialgebraic. Similar remarks can be made about the closed semialgebraic set $W \subset V$ in 8.11 (2).

It is of interest to note that there is another equivalent formulation of the Finiteness Theorem above in terms of the constructible subsets of the real spectrum. This can be stated as follows.

THEOREM 8.12. Under the one-one correspondence $\alpha \colon \mathscr{C}(A) \to \mathscr{S}(V)$ in 8.4 (1), open (closed) constructible sets correspond to open (closed) semi-

algebraic sets, i.e., a set $C \in \mathscr{C}(A)$ is open (closed) in X_A iff its contraction $\alpha(C) = C \cap V$ is open (closed) in V.

Of course, the substance of this theorem is in the "if" part. To see that this is equivalent to the Finiteness Theorem, we proceed as follows.

 $(8.11) \Rightarrow (8.12)$ Assume $C \cap V$ is open in V. By 8.11, this is a finite union of sets of the form $U(f_1, \ldots, f_r) \cap V$. Since the one-one correspondence α preserves inclusion, it follows that C is the (finite) union of the corresponding $\mathcal{U}(f_1, \ldots, f_r)$'s. Therefore, C is open in X_A .

 $(8.12) \Rightarrow (8.11)$ Given U as in 8.11, let $C \in \mathscr{C}(A)$ be the corresponding constructible set. By 8.12, C is open in X_A , so it is a union of basic Harrison open sets $H(f_1, \ldots, f_r) = \mathscr{U}(f_1, \ldots, f_r)$. But, being a constructible set, C is closed, and therefore compact, in the Tychonoff topology. It follows that C can be covered by a finite number of the $\mathscr{U}(f_1, \ldots, f_r)$'s. Contracting to V, we see that U is a finite union of the $V \cap U(f_1, \ldots, f_r)$'s.

REMARK 8.13. In a recent written communication to us, M.A. Dickmann has pointed out that it is possible to prove a refinement of the statement of the Finiteness Theorem 8.11 in the spirit of 8.10. Suppose k is the real closure of an ordered field (k_0, P) , and V is defined over k as before. Then any open (resp. closed) semialgebraic subset of V is a finite union of $U(f_1, \ldots, f_r)$ (resp. $W(g_1, \ldots, g_s)$) where each f_i (resp. g_i) is a polynomial over k_0 .

The Finiteness Theorem has proved to be of considerable importance in the recent development of semialgebraic geometry. It is perhaps worthwhile to say something about its short, but interesting history. This theorem was stated as an "Unproved Proposition" in Brumfiel's book [8]; Brumfiel did not give a proof, but managed to develop an extensive theory of semialgebraic sets without it. At about the time when Brumfiel's book appeared, several others have discovered and proved the same result. In the recent literature, there are five different proofs of the Finiteness Theorem, due respectively to T. Récio, C. Delzell, M. Coste and M.-F. Coste-Roy, L. van den Dries, J. Bochnak and G. Efroymson. (Delzell has pointed out that a sixth proof can be obtained from a paper of McEnerney on semianalytic sets by adapting his arguments to semialgebraic sets.) Several of these proofs used ideas from logic, and involved, in one form or another, Tarski's elimination of quantifiers. The use of Tarski's Principle is, of course not surprising, since writing an open semialgebraic set as a finite union of the $V \cap U(f_1, \ldots, f_r)$'s means expressing the definition of the given set without quantifiers or negations, in terms of strict inequalities. Tarski's Principle gets rid of the quantifiers, but may leave behind non-strict inequalities. The point of the Finiteness Theorem is, therefore, that the openness of the set assures at least one way of eliminating the quantifiers which leaves no non-strict inequalities behind. In this sense, the Finiteness Theorem may be thought of as a refined version of Tarski's elimination of quantifiers.

It is hoped that the several examples given above served to show the natural and fruitful ways in which ideas of logic can be used in semialgebraic geometry. In recent years, the subject of semialgebraic geometry has flourished on a successful combination of algebraic and logical techniques. This is by all means a healthy trend, and is expected to continue into the future. It is based on this expectation that we have chosen to include a brief discussion of the interaction between logic and semialgebraic geometry in our elementary exposition.

References

(These are the papers actually referred to in the text, and are only a very small sample of the recent work related to real algebra. Readers desiring more information about the current literature should consult the two excellent collections of articles in [16] and [25].)

1. E. Artin, Über die Zerlegung definiter Funktionen in Quadrate, Abh. Math. Sem. Univ. Hamburg **5** (1927), 100–115.

2. — and O. Schreier, *Algebraische Konstruktion reeller Körper*, Abh. Math. Sem. Univ. Hamburg **5** (1927), 85–99.

3. R. Baeza, Über die Stufe von Dedekind-Ringen, Arch. Math. 33 (1979), 226-231.

4. E. Becker, Valuations and real places in the theory of formally real fields, in the collection [25] below, 1–40.

5. L. Bröcker, *Positivbereiche in kommutativen Ringen*, Abh. Math. Sem. Univ. Hamb. 52 (1982), 170–178.

6. — , Unveröffentlichtes Manuskript über topologische Basisräume.

7. _____, A. Dress and R. Scharlau, An (almost) trivial local-global principle for the representation of -1 as a sum of squares in an arbitrary commutative ring, in the collection [16] below, 99–106.

8. G. Brumfiel, *Partially ordered rings and semialgebraic geometry*, L.M.S. Lecture Notes Series 37, Cambridge University Press, 1979.

9. M. D. Choi, T. Y. Lam, B. Reznick and A. Rosenberg, Sums of squares in some integral domains, J. Algebra 65 (1980), 234–256.

10. J.-L. Colliot-Thélène, Variantes du Nullstellensatz réel et anneaux formellement réels, in the collection [25] below, 98–108.

11. M. Coste and M.-F. Coste-Roy, *La topologie du spectre réel*, in the collection **[16]** below, 27–59.

12. Z. D. Dai, T. Y. Lam and C. K. Peng, *Levels in algebra and topology*, Bull. Amer. Math. Soc. (new series) **3** (1980), 845–848.

13. G. de Marco and A. Orsalli, *Commutative rings in which every prime ideal is con*tained in a unique maximal ideal, Proc. Amer. Math. Soc. 30 (1971), 459–466.

14. D. Dubois, A Nullstellensatz for ordered fields, Arkiv för Mat. 8 (1969), 111–114.

15. — and G. Efroymson, *Algebraic theory of real varieties*, in Studies and Essays presented to Yu-Why Chen on his 60th birthday, 107–135, Taiwan University, 1970.

16. — and T. Recio, *Ordered fields and real algebraic geometry*, Proceedings of the Special Session in 87th Annual Meeting of AMS in San Francisco, January, 1981, Contemporary Mathematics 8, Amer. Math. Soc., 1982.

17. R. Elman, T. Y. Lam and A. Wadsworth, Orderings under field extensilns, J. reine angew. Math. 306 (1979), 7-27.

18. M. Knebusch, Signaturen, reelle Stellen und reduzierte quardatische Formen, Jber. d. Dt. Math.-Verein. 82 (1980), 109–127.

19.——, On the local theory of signatures and reduced quadratic forms, Abh. Math. Sem. Univ. Hamburg **51** (1981), 149–195.

20. T. Y. Lam, *Algebraic theory of quadratic forms*, Benjamin, 1973 (revised printing: 1981), Reading, Mass.

21.——, *The theory of ordered fields*, in Ring Theory and Algebra III (ed. B. McDonald), Proceedings of Algebra Conference at University of Oklahoma, 1–152, M. Dekker, 1980.

22. ——, Orderings, valuations and quadratic forms, CBMS Lecture Notes Series 52, Amer. Math. Soc., 1983.

23. S. Lang, The theory of real places, Ann. Math. 57 (1953), 378-391.

24. A. Prestel, *Lectures on formally real fields*, IMPA Lecture Notes, 22, Rio de Janeiro, 1975.

25. J.-L. Colliot-Thélène, M. Coste, L. Mahé and M.-F. Roy, *Géométrie Algébrique Réelle et Formes Quadratiques*, Proceedings, Rennes 1981, Springer Lecture Notes in Math., **959**, Berlin-Heidelberg-New York, 1982.

26. J.-J. Risler, Une caractérization des idéaux des variétés algébriques réelles, C. R. Acad. Sci. Paris 271 (1970), 1171–1173.

27. C. Saliba, Sur la quasi-compacité de la topologie constructible du spectre réel, Appendice, Thèse de 3°-cycle, Rennes, France, 1983.

28. N. Schwartz, *Der Raum der Zusammenhangskomponenten einer reellen Varietät*, Geom. Dedicata **13** (1983), 361–397.

29. H. Simmons, Reticulated rings, J. Algebra 66 (1980), 169-192.

30. G. Stengle, A Nullstellensatz and a Positivstellensatz in semialgebraic geometry, Math. Ann. **207** (1974), 87–97.

31. O. Taussky, *The discriminant matrices of an algebraic number field*, J. London Math. Soc. **43** (1968), 152–154.

UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720