

# MUNSHI'S PROOF OF THE NULLSTELLENSATZ

J.PETER MAY

## 1. Introduction.

Following Ritabrata Munshi [5] and using extracts from Kaplansky's book [3], we give a geodesic proof of the Nullstellensatz, aimed at undergraduates. Some historical commentary will be given in Section 7.

We assume familiarity with the definitions of a commutative ring, an integral domain, a field, and the field of fractions of an integral domain. The letter  $R$  will always stand for an integral domain, and  $K$  will stand for its field of fractions. We also assume familiarity with the notion of an ideal. For the present purposes, an  $R$ -algebra  $A$  will mean a commutative ring  $A$  that contains  $R$  as a subring. The polynomial  $R$ -algebra  $R[x_1, \dots, x_n]$  may be viewed as  $R[x_1][x_2 \dots, x_n]$  or as  $R[x_1, \dots, x_{n-1}][x_n]$ . Polynomials in  $R[x_1, \dots, x_n]$  have degrees in each variable, and a total degree. It is convenient to let the zero polynomial have (total) degree  $-\infty$ . In  $R[x]$ , we then have

$$\deg(fg) = \deg(f) + \deg(g)$$

and

$$\deg(f + g) \leq \max(\deg(f), \deg(g))$$

for all polynomials  $f$  and  $g$ . Proofs of results about polynomials often proceed by inductive arguments in which one lowers degrees of polynomials by taking appropriate linear combinations. Munshi's new proof of the Nullstellensatz carries that simple idea to extreme lengths, introducing a simple and precise way to carry out such inductions when there are many variables.

## 2. Preliminaries.

Before explaining Munshi's proof, we summarize all of the relevant classical facts about integral domains that will implicitly or explicitly enter into the argument.

**Proposition 2.1.** *The ring  $R[x_1, \dots, x_n]$  is an integral domain.*

Indeed, by induction on  $n$ , it suffices to show this for  $n = 1$ . Here  $\deg(fg) \geq 0$  if and only if both  $\deg(f) \geq 0$  and  $\deg(g) \geq 0$ , which means that  $fg \neq 0$  if and only if both  $f \neq 0$  and  $g \neq 0$ .

We recall that a (proper) ideal  $P$  in  $R$  is *prime* if  $xy \in P$  implies  $x \in P$  or  $y \in P$ ;  $P$  is *maximal* if it is not properly contained in a larger (proper) ideal. A maximal ideal is a prime ideal. An element  $p$  of an integral domain  $R$  is *irreducible* if it is not zero and not a unit, and if  $p = ab$  implies that either  $a$  or  $b$  is a unit. This is one possible generalization of the notion of a prime number in  $\mathbf{Z}$ . Here is another. An element  $p$  is *prime* if the *principal ideal*  $(p) = \{rp \mid r \in R\}$  is a prime ideal.

**Proposition 2.2.** *Every prime element is irreducible, but not conversely.*

An integral domain  $R$  is a *principal ideal domain* (PID) if every ideal  $I$  in  $R$  is principal. Here the converse does hold in view of the following stronger result.

**Proposition 2.3.** *If  $R$  is a PID, then  $p$  is irreducible if and only if  $(p)$  is maximal.*

Indeed, if  $(p) \subset (q)$ , then  $p = rq$ , and if  $p$  is irreducible, then  $r$  must be a unit and  $(p) = (q)$ .

**Proposition 2.4.** *If  $F$  is a field, then  $F[x]$  is a PID.*

An integral domain  $R$  is a *unique factorization domain* (UFD) if every nonzero element  $a$  that is not a unit can be written as a finite product of irreducible elements, uniquely up to order of factors and multiplication by units. That is, if  $a = p_1 \cdots p_m$  and  $a = q_1 \cdots q_n$ , then  $m = n$  and, after reordering,  $q_i = u_i p_i$  for a unit  $u_i$ .

**Theorem 2.5.** *Every principal ideal domain is a unique factorization domain.*

**Theorem 2.6.** *If  $R$  is a unique factorization domain, then so is  $R[x_1, \dots, x_n]$ .*

Again, the proof is by induction on  $n$ . Of course, Theorem 2.6 would be false if UFD were replaced by PID, since  $x_1$  would be an irreducible element such that  $(x_1)$  is not maximal.

### 3. The Nullstellensatz.

Consider the fields  $\mathbf{R}$  and  $\mathbf{C}$  of real and complex numbers. In  $\mathbf{R}[x]$ , the polynomial  $x^2 + 1$  is irreducible. The quotient field  $\mathbf{R}[x]/(x^2 + 1)$  is a copy of  $\mathbf{C}$ : we have adjoined  $i = \sqrt{-1}$ .

**Theorem 3.1 (Fundamental theorem of algebra).** *Every polynomial  $f$  in  $\mathbf{C}[x]$  has a root  $a$  in  $\mathbf{C}$ . Thus, if  $f$  is monic, it splits completely as a product of linear polynomials  $x - a_i$ .*

This means that the only maximal ideals in  $\mathbf{C}[x]$  are the principal ideals  $(x - a)$ . A field  $F$  with the property of the conclusion is said to be *algebraically closed*. The Nullstellensatz says that this property propagates to polynomials in many variables.

**Theorem 3.2 (Nullstellensatz).** *Let  $F$  be an algebraically closed field. Then an ideal  $M$  in  $F[x_1, \dots, x_n]$  is maximal if and only if there are elements  $a_i$  in  $F$  such that  $M$  is the ideal generated by the elements  $x_i - a_i$ ; that is,*

$$M = (x_1 - a_1, \dots, x_n - a_n).$$

The name “Nullstellensatz,” or “zero–place theorem,” comes from the following consequence.

**Corollary 3.3.** *If  $I$  is a proper ideal of  $F[x_1, \dots, x_n]$ , then there is an element  $a = (a_1, \dots, a_n)$  in  $F^n$  such that  $f(a) = 0$  for all  $f$  in  $I$ .*

*Proof.* The ideal  $I$  is contained in some maximal ideal  $(x_1 - a_1, \dots, x_n - a_n)$ .  $\square$

The new proof of the Nullstellensatz is a direct consequence of the following theorem, which a priori has nothing to do with algebraically closed fields.

**Theorem 3.4 (Munshi).** *Assume that the intersection of the nonzero prime ideals of  $R$  is zero. If  $M$  is a maximal ideal in  $R[x_1, \dots, x_n]$ , then  $M \cap R \neq 0$ .*

The following result, which is equivalent to Kaplansky's [3, Thm. 21, p.14], is used in the proof of Munshi's theorem, and a special case of it enters into the application of Munshi's theorem to the proof of the Nullstellensatz.

**Theorem 3.5 (Kaplansky).** *The intersection of the nonzero prime ideals of  $R[x]$  is zero.*

*Proof of the Nullstellensatz.* An ideal  $(x_1 - a_1, \dots, x - a_n)$  is maximal since

$$F[x_1, \dots, x_n]/(x_1 - a_1, \dots, x - a_n)$$

is clearly isomorphic to  $F$  and is thus a field. Conversely, let  $M$  be a maximal ideal in  $F[x_1, \dots, x_n]$ , where  $n \geq 2$ . Regard  $F[x_1, \dots, x_n]$  as  $F[x_1][x_2, \dots, x_n]$ . By Kaplansky's theorem, the integral domain  $F[x_1]$  satisfies the hypothesis on  $R$  in Munshi's theorem. Therefore there is a nonzero element  $f$  in  $M \cap F[x_1]$ . Since  $F$  is algebraically closed,  $f$  splits into a product of linear factors. Because  $f$  is in  $M$  and  $M$  is maximal (and hence prime), at least one of those linear factors, say  $x_1 - a_1$ , is in  $M$ . The same argument gives an element  $x_i - a_i$  in  $M$  for each  $i$ ,  $1 \leq i \leq n$ . Then

$$(x_1 - a_1, \dots, x_n - a_n) \subset M.$$

Since  $(x_1 - a_1, \dots, x_n - a_n)$  is maximal, equality holds and we are done.  $\square$

Theorem 3.2 is actually the "weak form" of the Nullstellensatz. For completeness, and because it is the real starting point of algebraic geometry, we explain how little more is needed to prove the "strong form" of the Nullstellensatz in Section 6.

#### 4. The proof of Kaplansky's theorem.

We need a definition and some lemmas to prove Kaplansky's theorem.

**Definition 4.1.**  $R$  is of *finite type* if  $K$  is finitely generated as an  $R$ -algebra; that is, there are finitely many elements  $k_1, \dots, k_n$  of  $K$  such that every element of  $K$  is a polynomial in the  $k_i$  with coefficients in  $R$ .

**Lemma 4.2.** *If  $R$  is of finite type, then  $K$  is generated over  $R$  by a single element  $k$ , so that every element of  $K$  is a polynomial in  $k$  with coefficients in  $R$ .*

*Proof.* Let  $K$  be generated by elements  $k_1, \dots, k_n$ , where  $k_i = a_i/b_i$  with  $a_i$  and  $b_i$  in  $R$ . Let  $k = 1/b_1 \cdots b_n$ . If  $r_1 = a_1 b_2 \cdots b_n$ , then  $k_1 = r_1 k$ , and similarly for the other  $k_i$ . Therefore, since every element of  $R$  is a polynomial in the  $k_i$ , every element of  $R$  is a polynomial in  $k$ .  $\square$

Setting  $k = 1/c$ , we see that  $c$  is a nonzero element of  $R$  such that every element of  $K$  is a polynomial in  $1/c$  with coefficients in  $R$ . We may write  $K = R[1/c]$ .

**Lemma 4.3.** *The following conditions on a nonzero element  $c$  of  $R$  are equivalent:*

- (i)  $c$  is in the intersection of the nonzero prime ideals of  $R$ ;
- (ii) every nonzero ideal  $I$  of  $R$  contains some power of  $c$ ;
- (iii)  $K = R[1/c]$ .

*Proof.* (i)  $\implies$  (ii). Assume that no power of  $c$  is in  $I$  and let  $P$  be an ideal maximal among those that contain  $I$  but do not contain any power of  $c$ . Such a  $P$  exists by Zorn's lemma (or more directly if  $R$  is Noetherian, when any ascending chain of ideals stabilizes after finitely many stages). Then  $P$  is prime. Indeed, if  $ab$  is in  $P$  and neither  $a$  nor  $b$  is in  $P$ , then both  $(P, a)$  and  $(P, b)$  properly contain  $P$

and therefore each of these ideals contains some power of  $c$ , say  $p + ra = c^m$  and  $q + sb = c^n$  for some elements  $p, q$  in  $P$  and  $r, s$  in  $R$ . The product of these two elements is a power of  $c$  that lies in  $P$ , which is a contradiction. This shows that  $P$  is a prime ideal that does not contain  $c$ , which is contrary to (i). Therefore some power of  $c$  must be in  $I$ .

(ii)  $\implies$  (iii). For any nonzero  $b$  in  $R$ , some power  $c^n$  of  $c$  is in the ideal  $(b)$ , say  $rb = c^n$ . Then, in  $K$ ,  $1/b = r/c^n$ . This implies (iii).

(iii)  $\implies$  (i). Let  $P$  be any nonzero prime ideal of  $R$ , let  $b$  be a nonzero element of  $P$ , and write  $1/b = r/c^n$ . Then  $br = c^n$  is in  $P$ , hence  $c$  is in  $P$ .  $\square$

**Lemma 4.4.** *If  $R$  is a PID, then  $R$  is of finite type if and only if, up to units, it has only finitely many prime elements  $p_i$ .*

*Proof.* If  $c$  is a nonzero element of  $R$ , then  $c$  is a product of finitely many prime elements  $p_i$ , by Proposition 2.3 and Theorem 2.5. The equivalence of (iii) and (i) in Lemma 4.3 implies that  $K = R[1/c]$  if and only if, up to units, the  $p_i$  are the only prime elements in  $R$ .  $\square$

**Lemma 4.5.** *Let  $S$  be an integral domain such that  $R \subset S \subset K$ . If  $R$  is of finite type, then so is  $S$ .*

*Proof.* Observe that  $K$  is also the field of fractions of  $S$ . If  $K = R[1/c]$ , then  $K = S[1/c]$ .  $\square$

**Lemma 4.6.** *The polynomial ring  $K[x]$  has infinitely many prime ideals.*

*Proof.* By Proposition 2.4,  $K[x]$  is a PID, and it has infinitely many monic irreducible polynomials. Indeed, Euclid's proof that there are infinitely many prime numbers applies: if  $p_1, \dots, p_n$  were a complete list of the irreducible monic polynomials in  $K[x]$ , then  $q = 1 + p_1 \cdots p_n$  would be a monic polynomial divisible by none of the  $p_i$ . Since irreducible polynomials are prime elements, by Proposition 2.3, the conclusion follows.  $\square$

*Proof of Kaplansky's theorem.* Suppose that  $c$  is a nonzero element of  $R[x]$  that is in every nonzero prime ideal of  $R[x]$ . Let  $L$  be the field of fractions of  $R[x]$ . By Lemma 4.3,  $L = R[x][1/c]$ . Since  $L$  contains  $K$  and  $x$ , we have  $R[x] \subset K[x] \subset L$ . Since  $R[x]$  is of finite type, so is  $K[x]$ , by Lemma 4.5. Lemma 4.4 implies that  $K[x]$  has only finitely many monic irreducible polynomials, but Lemma 4.6 ensures that  $K[x]$  has infinitely many monic irreducible polynomials. The contradiction proves the result.  $\square$

## 5. The proof of Munshi's theorem.

We first prove the case  $n = 1$ , then the case  $n = 2$ . It will be immediately apparent that the same argument applies to prove the general case, at the price of just a little added notational complexity.

Let  $n = 1$ , write  $x = x_1$ , and assume that  $M \cap R = 0$ , contrary to the conclusion of the theorem. Let  $f(x) = a_0x^k + a_1x^{k-1} + \cdots + a_k$  be a polynomial of minimal degree in  $M$ , where  $a_i$  is in  $R$  and  $a_0 \neq 0$ . Then our assumption is that  $k \geq 1$ . By hypothesis, there is a nonzero prime ideal  $P$  of  $R$  such that  $a_0$  is not in  $P$ . Let  $p$  be a nonzero element of  $P$ . Since  $p$  is in  $R$ ,  $p$  is not in  $M$ . Thus  $(M, p) = R[x]$ . Let  $S = R - P$ . For each  $s$  in  $S$ , we can choose a polynomial  $g_s(x)$  in  $R[x]$  such that  $pg_s(x) + s$  is in  $M$ . Since  $s$  is not in  $P$ ,  $s$  is not in  $(p)$  and  $pg_s(x) + s \neq 0$ . Note

that  $g_s(x)$  and  $g_s(x) + s$  have the same degree. Here  $g_s(x)$  need not be unique, and we agree to choose  $g_s(x)$  to be of minimal degree among all possible choices. Since  $pg_s(x) + s$  is in  $M$ , its degree is at least  $k$ .

Now choose an element  $s_0$  of  $S$  such that  $g_{s_0}(x)$  has minimal degree among all  $g_s(x)$ . Write  $g_{s_0}(x) = b_0x^j + b_1x^{j-1} + \cdots + b_j$  with  $b_0 \neq 0$ . Then  $j \geq k$ . Because  $P$  is prime and both  $a_0$  and  $s_0$  are in  $S$ ,  $t = a_0s_0$  is also in  $S$ . Consider the element  $a_0(pg_{s_0}(x) + s_0) - b_0px^{j-k}f(x)$  of  $M$ . Since the coefficient of  $x^j$  is zero, the degree of this polynomial is at most  $j - 1$ . Clearly, we can rewrite it as an expression of the form  $g_t(x) + t$ . Since  $g_t(x)$  has degree at most  $j - 1$ , this contradicts the choice of  $s_0$ . Thus our original assumption that  $k \geq 1$  is incorrect and  $M \cap P \neq 0$ .

Now let  $n = 2$  and assume again that  $M \cap R = 0$ . We must derive a contradiction. Write  $x = x_1$  and  $y = x_2$  to simplify notation. Since Kaplansky's theorem shows that  $R[x]$  and  $R[y]$  satisfy the hypothesis of Munshi's theorem, we conclude from the case  $n = 1$  that  $M \cap R[x]$  and  $M \cap R[y]$  are nonzero. Choose polynomials  $d(x)$  in  $M \cap R[x]$  and  $e(y)$  in  $M \cap R[y]$  of minimal degrees  $m$  and  $n$  among all such polynomials.

Let  $N$  be the nonnegative integers and give  $N \times N$  the reverse lexicographic order:  $(i, j) < (i', j')$  if  $j < j'$  or if  $j = j'$  and  $i < i'$ . Define the *bidegree* of a nonzero polynomial  $h = \sum a_{ij}x^i y^j$  to be the maximal  $(i, j)$  in this ordering such that  $a_{ij} \neq 0$ ; we call  $a_{ij}$  the *leading coefficient* of  $h$ . It is convenient pictorially to think of the points of  $N \times N$  as a lattice in the first quadrant of the plane, with arrows drawn left and downwards to indicate adjacent inequalities.

The polynomials  $y^j d(x)$  and  $x^i e(y)$  in  $M$  have bidegrees  $(m, j)$  and  $(i, n)$ , respectively. Since  $M \cap R = 0$ ,  $m > 0$  and  $n > 0$ , so that  $(0, 0) < (m, 0) < (0, n)$ . Let  $B$  and  $\partial B$  denote the lower left box

$$B = \{(i, j) \mid 0 \leq i \leq m \text{ and } 0 \leq j \leq n\}$$

and its partial boundary

$$\partial B = \{(i, j) \mid i = m \text{ or } j = n\} \subset B.$$

We have an element of  $M$  of bidegree  $(i, j)$  for each  $(i, j)$  in  $\partial B$ .

A *flow*  $F$  from  $(a_q, b_q)$  to  $(0, 0)$  is a finite sequence of adjacent lattice points

$$F : (0, 0) < (a_1, b_1) < \cdots < (a_q, b_q).$$

Here "adjacent" means that, for  $0 \leq i < q$ , either

$$a_i = a_{i+1} - 1 \text{ and } b_i = b_{i+1} \quad \text{or} \quad a_i = a_{i+1} \text{ and } b_i = b_{i+1} - 1.$$

We say that  $(a_i, b_i)$  in  $F$  is a *point on the flow*  $F$ . We have the elementary, but key, observation that a flow from a point outside  $B$  down to  $(0, 0)$  must intersect  $\partial B$  and a flow from a point in  $B$  down to  $(0, 0)$  is part of a flow from  $(m, n)$  down to  $(0, 0)$ . Here, going downstream in a flow corresponds to going down in the reverse lexicographic order. Let  $\mathcal{F}$  denote the set of all flows from  $(m, n)$  to  $(0, 0)$ ; it is nonempty and finite.

Now we mimic the proof in the case  $n = 1$ . For a flow  $F$  in  $\mathcal{F}$ , let  $M_F$  be the set of nonzero polynomials in  $M$  with bidegree on  $F$ . Since there are nonzero polynomials of bidegree  $(m, n)$  in  $M$ ,  $M_F$  is nonempty. Choose a polynomial  $f_F$  in  $M_F$  of minimal bidegree. Since  $M \cap R = 0$ , the bidegree of  $f_F$  is not  $(0, 0)$ . Let  $a_F$  be the leading coefficient of  $f_F$  and let  $a$  be the product over  $F$  in  $\mathcal{F}$  of the  $a_F$ . Since  $a$  is a nonzero element of  $R$ , our hypothesis ensures that there is a nonzero prime ideal  $P$  of  $R$  such that  $a$  is not in  $P$ . Let  $p$  be a nonzero element of  $P$ . Since

$p$  is in  $R$ ,  $p$  is not in  $M$ . Thus  $(M, p) = R[x, y]$ . Let  $S = R - P$ . Since  $a$  is in  $S$ ,  $a_F$  is in  $S$  for all  $F$  in  $\mathcal{F}$ . For each  $s$  in  $S$ , we can choose an element  $g_s(x, y)$  of  $R[x, y]$  such that  $pg_s(x, y) + s$  is in  $M$ . Since  $s$  is not in  $P$ ,  $s$  is not in  $(p)$  and  $pg_s(x, y) + s \neq 0$ . Here  $g_s(x, y)$  need not be unique, and we agree to choose  $g_s(x, y)$  to be of minimal bidegree among all possible choices.

Now choose  $s_0$  to be an element of  $S$  such that  $g_{s_0}(x, y)$  has minimal bidegree among all  $g_s(x, y)$ . Let  $b$  be the leading coefficient of  $g_{s_0}(x, y)$ . Consider any flow from the bidegree of  $g_{s_0}(x, y)$  to  $(0, 0)$ . By our key observation, this flow must coincide with a flow  $F$  from  $(m, n)$  to  $(0, 0)$  from some point downwards. Clearly the bidegree of  $f_F$  lies downstream from, or coincides with, the bidegree of  $g_{s_0}(x, y)$ . Let  $(u, v)$  be the difference (in the obvious sense) of the bidegrees of these two polynomials. Then  $x^u y^v f_F$  and  $pg_{s_0}(x, y) + s_0$  are two elements of  $M$  of the same bidegree. Multiplying the former by  $bp$  and the latter by  $a_F$ , we obtain elements of  $M$  with the same leading term. Since  $P$  is prime and both  $a_F$  and  $s_0$  are in  $S$ ,  $t = a_F s_0$  is in  $S$ . The element  $a_F(pg_{s_0}(x, y) + s_0) - bp x^u y^v f_F$  of  $M$  can be rewritten in the form  $g_t(x, y) + t$ , where the bidegree of  $g_t(x, y)$  is less than the bidegree of  $g_{s_0}(x, y)$ . This is a contradiction, hence our original assumption that  $M \cap R = 0$  must be false.

As said at the start, the generalization to  $n$  variables works the same way.

## 6. The strong form of the Nullstellensatz.

Here we need another preliminary, namely the Hilbert basis theorem. A commutative ring  $R$  is said to be *Noetherian* if every ideal of  $R$  is finitely generated. We need only consider integral domains, but the general case of the following result is no more difficult.

**Theorem 6.1 (Hilbert basis theorem).** *If  $R$  is a commutative Noetherian ring, then so is  $R[x]$ . Therefore  $R[x_1, \dots, x_n]$  is Noetherian for all  $n$ .*

An ideal  $I$  of  $R$  is a *radical ideal* if  $x^m \in I$  for some  $m \geq 1$  implies  $x \in I$ . The *radical of an ideal*  $I$ , denoted  $\sqrt{I}$ , is the set of all elements  $x$  some power of which is in  $I$ . It is not hard to see that  $\sqrt{I}$  is in fact an ideal containing  $I$ .

Now focus on  $F[x_1, \dots, x_n]$  for a field  $F$  and a fixed  $n$ . Write  $\mathbf{A}^n = \mathbf{A}^n[F]$  for  $F^n$  regarded as just a set (ignoring its vector space structure), and call it *affine  $n$ -space*. The zeroes  $\mathcal{Z}(I)$  of an ideal  $I$  in  $F[x_1, \dots, x_n]$  are the points  $a$  of  $\mathbf{A}^n$  such that  $f(a) = 0$  for all  $f$  in  $I$ . The *affine algebraic sets* are the subsets  $V$  of  $\mathbf{A}^n$  that are the zeroes of a set of polynomials  $\{f_i\}$ . The ideal  $\mathcal{I}(V)$  is then defined to be the set of all polynomials  $f$  such that  $f(v) = 0$  for all  $v$  in  $V$ . This is an ideal, and it is clearly a radical ideal: if  $(f^m)(v) = f(v)^m = 0$ , then  $f(v) = 0$ .

Thus an algebraic set  $V$  gives rise to a radical ideal  $\mathcal{I}(V)$ , and an ideal  $I$  gives rise to an algebraic set  $\mathcal{Z}(I)$ . Because we start with sets  $V$  that are the zeroes of a set of polynomials, it is immediate that  $V = \mathcal{Z}(\mathcal{I}(V))$ . On the other hand, for an arbitrary ideal  $I$ , it is immediate that  $I$  is contained in  $\mathcal{I}(\mathcal{Z}(I))$ . Since  $\mathcal{I}(\mathcal{Z}(I))$  must be a radical ideal, equality cannot be expected in general. However, even if we start with a radical ideal, equality need not hold. The point is that not all radical ideals are of the form  $\mathcal{I}(V)$  for some  $V$ . For example, any prime ideal is a radical ideal, and the prime ideal  $(x^2 + 1)$  of  $\mathbf{R}[x]$  has no zeroes in  $\mathbf{R}$ . The strong form of the Nullstellensatz says that these conclusions do hold when the field we start with is algebraically closed.

**Theorem 6.2 (Strong form of the Nullstellensatz).** *Let  $F$  be an algebraically closed field. Then, for any ideal  $I$  of  $F[x_1, \dots, x_n]$ ,  $\mathcal{S}(\mathcal{Z}(I)) = \sqrt{I}$ . Therefore the correspondences  $\mathcal{Z}$  and  $\mathcal{S}$  between algebraic sets and radical ideals are inverse bijections.*

*Proof.* We must prove that  $\mathcal{S}(\mathcal{Z}(I))$  is contained in  $\sqrt{I}$ . By the Hilbert basis theorem,  $I$  is generated by a finite set  $\{f_1, \dots, f_q\}$  of polynomials. Consider an element  $g$  of  $\mathcal{S}(\mathcal{Z}(I))$ . We must prove that some power of  $g$  is in  $I$ . Introduce a new variable  $y$ , and let  $J$  be the ideal of  $F[x_1, \dots, x_n, y]$  generated by the  $f_i$  and  $yg - 1$ . Clearly  $g$  and the  $f_i$  depend only on the  $x_i$ , and  $g$  vanishes at any point  $a$  of  $\mathbf{A}^n$  at which each  $f_i$  vanishes. Therefore, if  $a$  is a point of  $\mathbf{A}^{n+1}$  such that  $f_i(a) = 0$  for all  $i$ , then  $a_{n+1}g(a) - 1 = -1$ . Thus  $\mathcal{Z}(J)$  is empty and  $J$  cannot be a proper ideal. Therefore  $J = F[x_1, \dots, x_n, y]$  and we can write

$$1 = h_1 f_1 + \dots + h_q f_q + h_{q+1}(yg - 1)$$

for some  $h_i$  in  $F[x_1, \dots, x_n, y]$ . Working in the field of fractions, say, we may set  $z = y^{-1}$  and think of the  $h_i$  as polynomials in the  $x_i$  and  $z^{-1}$ , and we may think of the last summand as  $z^{-1}h_{q+1}(g - z)$ . Multiplying by  $z^N$  for  $N$  large enough to clear denominators, we obtain

$$z^N = j_1 f_1 + \dots + j_q f_q + j_{q+1}(g - z)$$

for some  $j_i$  in  $F[x_1, \dots, x_n, z]$ . Setting  $z = g$  in this polynomial equation, we conclude that  $g^N$  is in  $I$ .  $\square$

It is convenient to let radical ideals  $I$  correspond to their quotient  $F$ -algebras  $F[x_1, \dots, x_n]/I$ . If  $I = \mathcal{S}(V)$ , this quotient ring is called the *coordinate ring* of  $V$  and denoted  $F[V]$ . It is to be thought of as the ring of polynomial functions on  $\mathbf{A}^n$  that vanish on  $V$ . The passage back and forth between algebraic sets and their coordinate rings is an algebraization of the geometry of solutions to polynomial equations.

Pedagogically, this development of the Nullstellensatz gives the starting point of a rigorous introduction to algebraic geometry that requires the absolute minimum of ring and field theoretic prerequisites.

## 7. A little background.

This is not the place to give full historical background, so I will restrict myself to a short discussion of previous elementary proofs of the Nullstellensatz. Oscar Zariski [6] first advertised the desirability of a proof that avoids advanced techniques and, in particular, avoids use of the Noether normalization theorem. He himself gave two such proofs. A very brief modern account of his first proof is posted on Dan Grayson's website [2]. While both of Zariski's proofs use some language beyond the level to which I have restricted myself, in particular the language of integrality and algebraic extensions, they could be rewritten in terms comparably accessible to undergraduates. However, while Zariski's proofs are as elementary as the proof given here and use some of the same details, I find the conceptual structure of the present proof especially appealing. Incidentally, Zariski ascribes the ingenious standard deduction of the strong form of the Nullstellensatz from the weak form to "A. Rabinowitsch," who is not to be found on MathSciNet.

A few years after Zariski's note, Oscar Goldman [1] and Wolfgang Krull [4], independently, gave two quite similar elementary proofs of the Nullstellensatz, both

of which start out with some of the ideas used here. Kaplansky's book [3] gave a reworking of Goldman's argument that was the starting point for Munshi's proof. Incidentally, while Kaplansky proved our Theorem 3.5, he didn't state it in the same form. He called an integral domain of finite type a " $G$ -domain," in honor of Goldman, and the material in [3, pp. 12–15] includes our Lemmas 4.2–4.6 and the theorem that  $R[x]$  is never a  $G$ -domain, which is equivalent to our Theorem 3.5.

Munshi, who is currently a graduate student at Princeton University, found his proof while visiting Bombay as an undergraduate at Calcutta. He published it in [5] (in a journal with a very small circulation, restricted to a few institutes in India) after receiving rejections from several other journals, including this Monthly. He included his proof with his applications to graduate school. I was impressed by the argument, and I reworked it for use in the University of Chicago's 2000 summer REU program, posting it on my web page at that time. Several people persuaded me that it should be published. I offered Munshi three options: a paper by Munshi rewritten in the light of my reworking, a joint paper, or publication in the present form. Munshi chose to have me publish this version. I make no claim to originality.

#### REFERENCES

- [1] O. Goldman, Hilbert rings and the Hilbert Nullstellensatz, *Math. Z.* **54** (1951), 136–140.
- [2] D. Grayson, The Hilbert Nullstellensatz, <http://www.math.uiuc.edu/~dan/ShortProofs/>.
- [3] I. Kaplansky, *Commutative Rings*, rev. ed., University of Chicago Press, Chicago, 1974.
- [4] W. Krull, Jacobson'sches Radikal und Hilbert'scher Nullstellensatz, in *Proceedings of the International Congress of Mathematicians*, Cambridge, MA, 1950, vol. 2, American Mathematical Society, Providence, 1952 pp. 56–64.
- [5] R. Munshi, Hilbert's Nullstellensatz, *Bulletin of the Bombay Mathematical Colloquium*, vol 15, number 1-2, January, 1999, 20-24.
- [6] O. Zariski, A new proof of Hilbert's Nullstellensatz, *Bull. Amer. Math. Soc.* **53** (1947) 362–368.

**J.Peter May** is an algebraic topologist with a longstanding fondness for commutative algebra, a fondness that was much abetted by many discussions with Oscar Zariski at Cambridge University, where both were visitors in 1971–72. May is a professor at the University of Chicago, where he is a past chair of the Department of Mathematics and of the University's Council on Teaching. He is a recipient of the University's Faculty Award for Excellence in Graduate Teaching. He is currently director of Chicago's NSF VIGRE program, in which role he organizes and teaches in Chicago's summer REU program.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, CHICAGO, IL 60637  
*E-mail address:* [may@uchicago.edu](mailto:may@uchicago.edu)