CHAPTER 9

Hilbert's Nullstellensatz

In something of an algebraic detour, we will now prove Theorem 5.3.1 for affine hypersurfaces. In the general case, we shall also state (but not prove) a reformulation which lays out the correspondence between affine algebraic varieties and ideals in commutative rings.

9.1. Resultants (bis)

We need another result on resultants. As in §8.1 let $\mathbb{D}$ be a UFD with fraction field K; and for $f = a_0Y^n + a_1Y^{n-1} + \cdots + a_n$ and $g = b_0Y^m + b_1Y^{m-1} + \cdots + b_m$ polynomials in $\mathbb{D}[Y]$, define $\mathcal{R}(f,g) := \det M_{(f,g)}$. (In case $\mathbb{D}$ is itself a polynomial ring, we will often write $\mathcal{R}_Y(f,g)$ to make it clear that Y is the variable being eliminated.)

9.1.1. PROPOSITION. *$\mathcal{R}(f,g) = Gf + Fg$ for some $F, G \in \mathbb{D}[Y]$ with* $\deg G < \deg g$, $\deg F < \deg f$.

PROOF. If $\mathcal{R}(f,g) = 0$, then we are done by (8.1.3). Otherwise, write

(9.1.2)
$$\begin{array}{rclllllll}
Y^{m-1}f & = & a_0Y^{n+m-1} & +a_1Y^{n+m-2} & +\cdots & +a_nY^{m-1} & & \\
Y^{m-2}f & = & & a_0Y^{n+m-2} & +\cdots & \cdots & +a_nY^{m-2} & \\
\vdots & & & & & & & \\
f & = & & & a_0Y^n & +\cdots & \cdots & +a_n \\
Y^{n-1}g & = & b_0Y^{n+m-1} & +b_1Y^{n+m-2} & +\cdots & +b_mY^{n-1} & & \\
Y^{n-2}g & = & & b_0Y^{n+m-2} & +\cdots & \cdots & +b_mY^{n-2} & \\
\vdots & & & & & & & \\
g & = & & & b_0Y^m & +\cdots & \cdots & +b_m.
\end{array}$$

Viewing the system (9.1.2) as a vector equation, the RHS is evidently

$$M_{(f,g)} \begin{pmatrix} Y^{n+m-1} \\ Y^{n+m-2} \\ \vdots \\ Y \\ 1 \end{pmatrix}.$$

Moreover, by Cramer's rule we have (in K) $M_{(f,g)}^{-1} = (\det M_{(f,g)})^{-1} A$, where A is the *adjugate matrix* with $(i,j)^{\text{th}}$ entry $(-1)^{i+j}$ times the $(j,i)^{\text{th}}$ minor of $M_{(f,g)}$. In other words, the entries of

$$\mathcal{R}(f,g) M_{(f,g)}^{-1} = A$$

are in $\mathbb{D}$. Applying this to both sides of (9.1.2) thus produces a system of the form

$$\begin{array}{rcllll} ?? & = & \mathcal{R}(f,g)Y^{n+m-1} & & & \\ ?? & = & & \mathcal{R}(f,g)Y^{n+m-2} & & \\ \vdots & & & & \ddots & \\ ?? & = & & & & \mathcal{R}(f,g) \end{array} \tag{9.1.3}$$

where each "??" is a $\mathbb{D}$-linear combination of the entries to the left of "=" in (9.1.2). In particular, the last row of (9.1.3) is

$$G_0 f + F_0 g = \mathcal{R}(f,g),$$

where $G_0, F_0 \in \mathbb{D}[Y]$ satisfy $\deg G_0 \leq m-1$, $\deg F_0 \leq n-1$. $\square$

We should mention the formula for the resultant of two polynomials whose irreducible factors are all linear (or constant) in y, although we will neither use nor prove it:

9.1.4. PROPOSITION. *If f and g decompose into linear factors $f = a_0 \prod_i (Y - x_i)$, $g = b_0 \prod_j (Y - y_j)$ (for $x_i, y_j \in \mathbb{D}$), then $\mathcal{R}(f,g) = a_0^m b_0^n \prod_{i,j}(x_i - y_j)$.*

9.2. Study's lemma

We continue to assume that $\mathbb{D}$ is a UFD with $f \in \mathbb{D}[Y]$ of degree n. Given $\delta \in \mathbb{D}$, we have the ring homomorphism given by "evaluation at δ":

$$\begin{array}{ccc} \mathbb{D}[Y] & \xrightarrow{\theta_\delta} & \mathbb{D} \\ G(Y) & \longmapsto & G(\delta) \end{array}.$$

9.2.1. PROPOSITION. *(i) If $f(\delta)(=\theta_\delta(f)) = 0$, i.e. δ is a root of f, then $(Y-\delta) \mid f(Y)$.*

(ii) f has at most n roots in $\mathbb{D}$.

PROOF. (i) By the division algorithm,

$$f = q.(Y-\delta) + r \tag{9.2.2}$$

where $\deg r < \deg(Y-\delta) = 1$, i.e. $r \in \mathbb{D}$. Applying θ_δ to (9.2.2), we have

$$0 = f(\delta) = q(\delta).0 + r$$

and thus $r = 0$, so that $(Y-\delta)$ divides f.

(ii) Follows from (i) (and the fact that $\mathbb{D}[Y]$ is a UFD) since f can have at most $n = \deg(f)$ linear factors. $\square$

Now we will specialize to the case $\mathbb{D} = \mathbb{C}[X]$; more generally, the results of this section will hold with any algebraically closed field replacing $\mathbb{C}$, $\{X_1, \ldots, X_{n-1}\}$ replacing X, and S_n replacing S_2.

Let $F \in \mathbb{D}[Y] = \mathbb{C}[X,Y] = S_2$.

9.2.3. PROPOSITION. *If $V(F) = \mathbb{C}^2$, i.e. F vanishes on all of $\mathbb{C}^2$, then $F = 0$ as an element of S_2.*

PROOF. Suppose $F \neq 0$. By Prop. 9.2.1(ii), viewed as an element of $\mathbb{D}[Y]$, F has a finite number of roots in $\mathbb{D} = \mathbb{C}[X]$. Some of these may be constants in $\mathbb{C}$. Since $\mathbb{C}$ is an infinite field, there exists $\beta \in \mathbb{C}$ such that β is not one of these roots, and then $F(X,\beta)(=\theta_\beta(F)) \neq 0$ in $\mathbb{C}[X]$. Again by Prop. 9.2.1(ii), $F(X,\beta)$ itself has finitely many roots, so there exists $\alpha \in \mathbb{C}$ such that $F(\alpha,\beta) \neq 0$. Hence, F is not identically zero on $\mathbb{C}^2$. $\square$

9.2.4. PROPOSITION. [STUDY'S LEMMA] *Given $f, g \in S_2$, with f irreducible and $V(f) \subseteq V(g)$. Then f divides g.*

9.2.5. REMARK. Suppose we drop the requirement that f be irreducible, so that $f = \prod f_i^{m_i}$ (f_i irreducible in S_2). Then $V(f_i) \subset V(f)$ for each i, and by the Proposition $f_i|g$ for each i. This implies that $f|g^{\sum m_i}$, i.e. f divides a power of g.

PROOF. Since $f|0$ is trivial, we take $g \neq 0$. By Prop. 9.2.3, we have $V(g) \neq \mathbb{C}^2$, which implies $V(f) \neq \mathbb{C}^2$ hence $f \neq 0$. We may assume that $f \notin \mathbb{C}$ (since a constant divides anything), and furthermore that $\deg_Y(f) \neq 0$ (otherwise just swap X and Y). Writing

$$f = a_0(X)Y^n + a_1(X)Y^{n-1} + \cdots + a_n(X) \;\notin\; \mathbb{C}[X]$$

($n > 0$ and $a_0(X) \neq 0$), I make the **claim**:[1] *we can assume that $g \notin \mathbb{C}[X]$.*

Assuming the **claim**, f and g are of degree > 0 in Y, so by Prop. 9.1.1 (with $\mathbb{D} = \mathbb{C}[X]$), $\mathcal{R}_Y(f,g) = Fg + Gf \in \mathbb{C}[X]$ for $\deg_Y F < \deg_Y f$, $\deg_Y G < \deg_Y g$. Given any $\alpha \in \mathbb{C} \backslash V(a_0)$, since $\mathbb{C}$ is algebraically closed there exists a root $\beta \in \mathbb{C}$ of $f(\alpha, Y)$. From $V(f) \subseteq V(g)$ we see that $(\alpha, \beta) \in V(f(\alpha, Y)) \subseteq V(g(\alpha, Y)) \subseteq \mathbb{C}$, so that $f(\alpha, Y)$ and $g(\alpha, Y)$ have a common root for every $\alpha \in \mathbb{C}\backslash V(a_0)$. It follows that $a_0 \mathcal{R}_Y(f,g) \in \mathbb{C}[X]$ evaluates to zero at every $\alpha \in \mathbb{C}$, hence is zero. As $a_0 \neq 0$, we find $\mathcal{R}_Y(f,g) = 0$ in $\mathbb{C}[X]$; and then by Prop. 8.1.2, $\deg_Y(\gcd_{S_2}(f,g)) > 0$. (Alternately, $Fg = (-G)f \implies f, g$ have a divisor of nonzero degree in Y.) But f is irreducible, so divides any nonzero non-unit dividing it; we conclude that $f \,|\, \gcd_{S_2}(f,g) \,|\, g$.

To prove the **claim**, suppose $g \in \mathbb{C}[X]\backslash\{0\}$. Then there exists $\alpha \in \mathbb{C}\backslash V(g.a_0)$. Viewed as a function on $\mathbb{C}^2$, g is constant in Y, so $g(\alpha, \beta) \neq 0 \; \forall \beta \in \mathbb{C}$. But since $a_0(\alpha) \neq 0$, $\deg_Y(f(\alpha, Y)) > 0$; and then (as $\mathbb{C}$ is algebraically closed) $\exists \beta \in \mathbb{C}$ such that $f(\alpha, \beta) = 0$. By assumption, $V(f) \subseteq V(g)$ and so $g(\alpha, \beta) = 0$, a contradiction. □

[1] at this point, of course, we can't "just swap X and Y"

9.3. The Nullstellensatz

The proof of Study immediately generalizes to $\mathbb{C}^n$. This yields a version of Hilbert's Nullstellensatz for hypersurfaces:

9.3.1. COROLLARY. *If $V(f) = V(g)$ for $f, g \in S_n$ and ...*
(i) f, g are irreducible, then $f = \lambda g$ ($\lambda \in \mathbb{C}^*$)
(ii) f, g are not irreducible, then $\exists\ M, N \in \mathbb{N}$ such that $f | g^N$, $g | f^M$. Equivalently, f and g have the same irreducible factors.

PROOF. (i) Study $\implies$ $f|g$ and $g|f$; (ii) is by Remark 9.2.5. $\square$

The point of this is that, modulo issues with powers, there is a *bijection* between hypersurfaces and principal ideals (i.e. polynomials up to multiplication by constants) in S_n which reverses inclusion. That is, provided f and g are "reduced" (all irreducible factors occur with multplicity 1), $(f) \supset (g) \iff f|g \iff V(f) \subset V(g)$.

To get a more general perspective on this, we introduce a few new ideas. First, given a subset $\mathfrak{X} \subseteq \mathbb{C}^n$, we define the ideal of $\mathfrak{X}$ by

$$I(\mathfrak{X}) := \{f \in S_n \mid f(\underline{z}) = 0\ \forall \underline{z} \in \mathfrak{X}\}.$$

For example, if f is "reduced", we clearly have $I(V(f)) = (f)$ by Study's Lemma: any g vanishing on $V(f)$ is divisible by f. A subset $\mathfrak{X} \subseteq \mathbb{C}^n$ is *algebraic* if it is of the form $V(J)$ for some ideal $J \subset S_n$. (Indeed, this is just an affine algebraic variety.) The statement $V(I(\mathfrak{X})) = \mathfrak{X}$ is true (almost a tautology) for algebraic subsets. Moreover, $I(\cdot)$ reverses inclusions as $\mathfrak{X}_1 \subset \mathfrak{X}_2 \implies I(\mathfrak{X}_1) \supset I(\mathfrak{X}_2)$.

Given any ideal $J \subset S_n$, we let $\sqrt{J}$ denote the *radical* of J, which is the ideal comprising all elements of S_n some power of which belongs to J. A *radical ideal* is an ideal which equals its own radical. Finally, J is *prime* $\iff$ S_n/J is a domain ($\iff$ J is irreducible in the monoid of ideals in S_n), and *maximal* $\iff$ S_n/J is a field.

9.3.2. THEOREM. *Let $J \subset S_n$ be an ideal.*

(i) J is maximal $\iff$ $J = (Z_1 - \alpha_1, \ldots, Z_n - \alpha_n)$ for some $\alpha_i \in \mathbb{C}$;
(ii) If $J \neq S_n$, then $V(J) \neq \emptyset$;
(iii) $I(V(J)) = \sqrt{J}$.

Theorem 9.3.2(iii) is the standard modern formulation of the Nullstellensatz,[2] and is equivalent to Theorem 5.3.1 (why?). It has the following important consequence, where an algebraic subset is *irreducible* if it is not a union of two proper algebraic subsets:

9.3.3. COROLLARY. *The correspondence*

$$\begin{array}{ccc} \text{ideals} & & \text{subsets} \\ \{J \subset S_n\} & \underset{V}{\overset{I}{\leftrightarrows}} & \{\mathfrak{X} \subset \mathbb{C}^n\} \end{array}$$

induces inclusion-reversing bijections

$$\begin{array}{ccc} \{\text{radical ideals}\} & \longleftrightarrow & \{\text{algebraic subsets}\} \\ \cup & & \cup \\ \{\text{prime ideals}\} & \longleftrightarrow & \{\text{irred. alg. subsets}\} \end{array}.$$

The last correspondence is checked in the exercises, by showing that $V(J_1 J_2) = V(J_1) \cup V(J_2)$; the rest is clear from the Theorem.

One can push the relation between affine algebraic geometry and commutative algebra much further. For example, the *ring of regular functions* on an irreducible affine variety $V = V(\mathfrak{P})$ ($\mathfrak{P}$ a prime ideal) is defined by

$$\mathbb{C}[V] := S_n/\mathfrak{P},$$

and it is easy to see that this embeds (say, for V smooth) in $\mathcal{O}(V)$. (The idea is that $\mathfrak{P}$ is the kernel of the map from S_n to $\mathcal{O}(V)$ given by restricting polynomial "functions" to V, and so $S_n/\mathfrak{P}$ is its image.) $\mathbb{C}[V]$ is sometimes also called the *coordinate ring* of V. Furthermore, if V is the affine part of a smooth projective variety $\bar{V}$, the field of meromorphic functions $\mathcal{K}(\bar{V})$ is isomorphic to the fraction field $\mathbb{C}(V)$ of $\mathbb{C}[V]$. Usually $\mathbb{C}(V)$ is called the *function field* of $\bar{V}$ (or V).

There is even a way to recover varieties from their coordinate rings; this is the "Spec" operation. Very roughly speaking, the affine story is this: any commutative ring A which is finitely generated

[2] Again, we can replace $\mathbb{C}$ here with any algebraically closed field. (A proof of Theorem 9.3.2 is in my Algebra II notes, but would take us too far afield here.)

over $\mathbb{C}$ may be presented as $\mathbb{C}[z_1, \ldots, z_N]/I$ (where $I \subseteq \mathbb{C}[z_1, \ldots, z_N]$ is an ideal), and then you take $V(I) \subseteq \mathbb{C}^N$. This gives one realization of Spec(A); of course, there are many ways of writing A in this form (different N, different I, etc.). From the standpoint of scheme theory, Spec(A) is something intrinsic, an *affine scheme* which exists in the absence of any particular embedding in an affine space $\mathbb{C}^N$. The best resources on this are the book by E. Kunz and the classic text by R. Hartshorne.

The exercises that follow explore some consequences of the Nullstellensatz.

Exercises

(1) Prove: (i) that for any algebraic subset $\mathfrak{X} \subseteq \mathbb{C}^n$, $V(I(\mathfrak{X})) = \mathfrak{X}$; (ii) that for any two ideals $J_1, J_2 \subseteq \mathbb{C}[Z_1, \ldots, Z_n]$, $V(J_1 J_2) = V(J_1) \cup V(J_2)$.

(2) For any finite collection of ideals $\{J_i\}_{i=1}^m$, show that (i) $V(\sum_i J_i) = \cap_i V(J_i)$ and (ii) $V(\cap_i J_i) = V(J_1 \cdots J_m) = \cup_i V(J_i)$. [Hint for (ii): $V(J) = V(\sqrt{J})$ (why?); so start by checking $\sqrt{\cap_i J_i} = \sqrt{J_1 \cdots J_m}$.]

(3) Show that an affine variety V is irreducible if and only if $I(V)$ is a prime ideal. [Hint for one direction: if $J := I(V)$ is not prime, then $\exists f_1, f_2 \in S_n \setminus J$ with $f_1 f_2 \in J$. Take $J_i := (f_i) + J$, show $V(J_i) \subsetneq V$, and consider $J_1 J_2$.]

(4) Prove that any decreasing chain $V_1 \supset V_2 \supset \cdots$ of affine varieties (in $\mathbb{C}^n$) "stabilizes" at some m: i.e., $V_m = V_{m+1} = \cdots$. [Hint: you may assume that every ideal in S_n is finitely generated (Hilbert basis theorem). Why does this imply that any ascending chain of ideals must stabilize?]

(5) (i) Show that every nonempty affine variety $V = V(J) \subset \mathbb{C}^n$ may be written uniquely as a finite union $V_1 \cup \cdots \cup V_r$, where each V_i is irreducible and $V_j \not\subset V_i$ for $i \neq j$. [Hint: suppose otherwise, and use Exercise (4).] (ii) Work this out for $V(J)$, where $J = (z_1 z_2 - z_3,\, z_1 z_3 - z_2^2)$.