Math 711: Lecture of December 6, 2004

We begin our attack on the proof of the Theorem stated at the end of the Lecture Notes from December 3 by making several reductions in the problem. First note that it suffices to consider just one system of equations. We denote the equations

$$F_j(X_1,\, \ldots\, , X_d, Y_1,\, \ldots\, , Y_h) = 0, \quad 1 \le j \le n.$$

Second, we note that if there is a solution $x = x_1,\, \ldots\, , x_d,\ y = y_1,\, \ldots\, , y_h$ in a Noetherian ring R such that $(x_1,\, \ldots\, , x_d)R$ has height d, then we may localize R at a minimal prime of $(x_1,\, \ldots\, , x_d)R$ of height d. Thus, we get a solution of the system satisfying the height constraint if and only if we get a solution in a local ring of dimension d such that $x_1,\, \ldots\, , x_d$ is a system of parameters. We may evidently replace this local ring by its completion. Hence there is a solution satisfying the height constraint if and only if there is a solution in a complete local ring $(R,\, m)$ such that $x_1,\, \ldots\, , x_d$ is a system of parameters.

This solution is preserved if we kill a minimal prime of R such that the quotient still has dimension d. Thus, we may assume that $(R,\, m)$ is a complete local domain. By the structure theory of complete local rings, we then have that R is module-finite over A, where A is a formal power series ring either over a field K or over a complete DVR $(V,\, pV)$ of mixed characteristic p. Specifically, we write $A = K[[u_1,\, \ldots\, , u_d]]$ or $A = V[[u_2,\, \ldots\, ,\, u_d]]$, and in the second case we let $u_1 = p$. In either case, $u_1,\, \ldots\, , u_d$ is a regular system of parameters for A.

Since R is module-finite over A, it has a set of generators $\theta_1 = 1,\, \theta_2,\, \ldots\, ,\, \theta_s$. These are not free generators: let $\mathcal{M} = \bigl(a_{ij}\bigr)$ denote an $s \times t$ matrix over A whose t columns span the module of relations on $\theta_1,\, \ldots\, , \theta_s$ over A.

We next want to describe a finite system of polynomial equations in variables corresponding to the a_{ij} and other scalars from A. The equations have coefficients in $\mathbb{Z}$, and solving them in a ring A_1 enables one to construct a ring S module-finite over A_1 with a set of generators $\theta'_1 = 1,\, \theta'_2,\, \ldots\, , \theta'_s$.

First note that for all $i,\, j$ we can choose $a_{ijk} \in A$ such that

$$(*) \qquad \theta_i\theta_j = \sum_{\kappa=1}^{s} a_{ijk}\theta_k.$$

We need the a_{ij} and the a_{ijk} to satisfy certain equational conditions in order to guarantee that one gets a commutative associative ring with θ_1 as identity. We impose:

$$(1) \qquad a_{1jk} = \delta_{jk}$$

where δ is the Kronecker δ function which is 1 when $j = k$ and 0 otherwise. This guarantees the multiplication by θ_1 is the identity function.

To guarantee that multiplication will be commutative we impose

$$(2) \qquad a_{ijk} = a_{jik}$$

for all i, j, k.

We want to write down equations that will guarantee that multiplication is associative. To do this, for all i, j, k write down the formula for $\theta_i\theta_j$ as a linear combination of $\theta_1, \dots, \theta_s$ using $(*)$, and then multiply on the right by θ_k and simplify each quadratic term by another application of $(*)$. This yields a formula for $(\theta_i\theta_j)\theta_k$ as a linear combination of $\theta_1, \dots, \theta_s$ in which every coefficient is a polynomial in the a_{ijk}. Similarly, use $(*)$ to write down $a_j a_k$ as a linear combination of $\theta_1, \dots, \theta_s$ and then multiply by θ_i and use $(*)$ repeatedly again to give a formula for $\theta_i(\theta_j\theta_k)$ in which every coefficient is a polynomial with integer coefficients in the a_{ijk}. Subtract the two linear combinations to obtain $B_1^{(ijk)}\theta_1 + \cdots + B_s^{(ijk)}\theta_s$ where each $B_\mu^{(ijk)}$ is a polynomial in the a_{ijk} with coefficients in $\mathbb{Z}$. Write the coefficients as an $s \times 1$ column vector. Then this column vector gives a relation on $\theta_1, \dots, \theta_s$, and so is a linear combination of the columns of $\big(a_{ij}\big)$. Thus, if $B^{(ijk)}$ denotes the column vector whose entries are $B_1^{(ijk)}, \dots, B_s^{(ijk)}$ then for all i, j, k we can choose a $t \times 1$ vector $U^{(ijk)}$ of scalars from A such that

$$(3) \qquad B^{(ijk)} = \mathcal{M}U^{(ijk)},$$

where $\mathcal{M}$ is the $s \times t$ matrix $\big(a_{ij}\big)$ whose columns span the relations on $\theta_1, \dots, \theta_s$.

We also need to write down equations that guarantee that multiplication is well-defined. The relation $\sum_{i=1}^{s} a_{ij}\theta_i = 0$ coming from the j the column of the matrix $\mathcal{M} = \big(a_{ij}\big)$, when multiplied by θ_k, produces a sum $\sum_{i=1}^{s} a_{ij}\theta_i\theta_k$ that can be rewritten, using $(*)$, as a linear combination of the θ_i with coefficients that are polynomials in the a_{ij} and a_{ijk} with integer coefficients. Since this is 0, for each choice of j and k the column vector $C^{(jk)}$ of coefficients of the θ_i can be written in the form $\mathcal{M}W^{(jk)}$, where $W^{(jk)}$ is a $t \times 1$ column vector with entries in A, i.e.,

$$(4) \qquad C^{(jk)} = \mathcal{M}W^{(jk)}.$$

Note that the ring structure of R is completely determined by the choice of the ring A and the scalars $\mathcal{M} = \big(a_{ij}\big)$, a_{ijk}, and the entries of the $U^{(ijk)}$ and $W^{(jk)}$. Moreover, given any ring A_1 and corresponding subscripted elements of A_1 satisfying the equations (1), (2), (3), and (4), say b_{ij}, b_{ijk}, U_1^{ijk}, and $W_1^{(jk)}$ (the latter two are vectors over A_1), one gets a unique commutative, associative, ring S generated as an A_1-module by elements $\theta'_1, \dots, \theta'_s$ such that θ'_1 is the identity, the columns of the matrix $\big(b_{ij}\big)$ span the relations on $\theta'_1, \dots, \theta'_s$, and such that for all i and j

$$(*') \qquad \theta'_i\theta'_j = \sum_{k=1}^{s} b_{ijk}\theta'_k$$

S is defined as the cokernel of the matrix $\big(b_{ij}\big)$ over A_1 with $\theta'_1, \dots, \theta'_s$ as the image of the standard basis. The equations $(*')$ are used to define the ring multiplication. The equations (1) guarantee that multiplication by θ_1 will be the identity map, the equations (2) guarantee that multiplication is commutative, the equations (3) guarantee that multiplication is associative, and the equations (4) guarantee that multiplication is well-defined.

We denote this ring as $\mathcal{R}(A_1; b_{ij},\, b_{ijk},\, U_1^{(ijk)},\, W_1^{(jk)})$.

We use this idea to descend a solution of the equations

$$F_j(X_1,\, \ldots\, ,X_d, Y_1,\, \ldots\, ,Y_h) = 0, \quad 1 \le j \le n$$

over R, which is a module-finite extension of A, to a solution in a ring S with is a module-finite extension of the Henselian ring A_1, where A_1 is the Henselization of the localization at $(u_1,\, \ldots\, ,u_d)$ of the ring $K[u_1,\, \ldots\, ,u_d]$ (respectively, $V[u_2,\, \ldots\, ,\, u_d]$), so that $\widehat{A_1} = A$. However we need three sets of additional equations.

In the original choice of R, we can write each x_μ and each y_ν as a linear combination of the θ_i with coefficients in A. These coefficients γ will all become unknowns, to be solved for in A_1. Each polynomial $F_j(x, y)$ can be expressed, using $(*)$ repeatedly, as a linear combination of the the θ_i each of whose coefficients is a polynomial in the a_{ijk} and the elements γ. This means that the column vector of coefficients of the θ_i can be written as the product of $\mathcal{M}$ with a column vector over A, whose entries will be a new set of unknowns. We refer to the equations obtained in this way as (5).

In R, the elements $x_1,\, \ldots\, ,x_d$ are a system of parameters. This means that there is a fixed integer N such that every u_ν^N is an R-linear combination of $x_1,\, \ldots\, ,x_d$. Each x_ν is already expressed as an A-linear combination of the θ_i. We can also write each coefficient of each x_ν as an A-linear combination of the θ_i: this introduces new elements that initially vary in A. The equations $(*)$ can be used to rewrite $u_\nu^N\theta_1$ minus the R-linear combination of the x_ν as a linear combination of the $\theta_1,\, \ldots\, ,\theta_s$. The coefficients are polynomials in elements that may be thought of a varying in A. As in previous examples we set the column vector of coefficients equal to $\mathcal{M}$ multiplied by a column vector whose entries are in A, but which should be thought of as new variables. We refer to the equations one gets as (6).

Finally, we need equations which keep track of the condition that the map $A \to R$ is injective. If we tensor with the fraction field $\mathcal{F}$ of A, we get an injection $\mathcal{F} \to \mathcal{F} \otimes_A R$. Hence we have a $\mathcal{F}$-linear map $\mathcal{F} \otimes_A R \to \mathcal{F}$ that is nonzero. This gives a composite A-linear map $R \to \mathcal{F} \otimes_A R \to \mathcal{F}$ that is not zero. We can choose a common denominator in $A - \{0\}$ for the values of this nonzero A-linear map $R \to \mathcal{F}$ on the θ_i and multiply by it to obtain an A-linear map $\phi : R \to A$ that is not zero. Let $a_1,\, \ldots\, ,a_s$ be the values on the θ_i. The condition that there exist an A-linear map with these values is that the matrix product

$$(7) \qquad (a_1 \;\ldots\; a_s)\mathcal{M} = 0.$$

Suppose that $a_i \neq 0$. Then we can choose Q such that $a_i \notin m^Q$.

We now think of every subscripted element from A as a variable. However, in the equations (6) the elements u_ν^N are treated as fixed elements of A_1. The resulting system of polynomial equations over A_1 has a solution in A, and it therefore has a solution in A_1 congruent to the original solution mod m^Q, by the Artin Approximation Theorem.

The solution in A_1 gives rise to a ring S that is module-finite over A_1, generated as an A_1-module by $\theta'_1 = 1, \ldots, \theta'_d$. The equations (1), (2), (3), and (4) guarantee that one has a well-defined commutative associative multiplication on A_1 such that multiplication by θ'_1 is the identity map. The equations (5) guarantee that we have a solution for the equations

$$F_j(X_1, \ldots, X_d, Y_1, \ldots, Y_h) = 0, \quad 1 \le j \le n,$$

and the equations (6) guarantee that $(x_1, \ldots, x_d)S$ is primary to the maximal ideal of S. From the equations (7) we get a nonzero A_1-linear map $S \to A_1$ whose values on the θ'_i are the values corresponding to the variables introduced to replace the a_i. This map is not zero because the value on θ_i is congruent to the original value mod m^Q.

Finally, we may descend further, from A_1 to a suitable étale extension of $K[u_1, \ldots, u_d]$ or $V[u_2, \ldots, u_d]$. After localizing at one element, if necessary, we will still have that the radical of $(x_1, \ldots, x_d)S$ is maximal of height d. This completes the proof of the theorem in case we are working over a DVR.

In the case of a field K, we want to show that we can replace the field K by a finite field, even if K has characteristic 0 initially.

Henceforth, we assume that we have a solution in a finitely generated K-algebra, where K is a field that may be of positive characteristic or characteristic 0, and where the values $x_1, \ldots, x_d$ for $X_1, \ldots, X_d$ generate an ideal whose radical is a maximal ideal of height d. We may write this K-algebra in the form

$$R = K[X_1, \ldots, X_d, Y_1, \ldots, Y_h, Z_1, \ldots, Z_s]/\big(F_j(X,Y), G_k(X,Y,Z)\big)$$

where the X_d and Y_h map to the solution in the quotient ring. We may choose a finitely generated $\mathbb{Z}$-subalgebra B of K that contains all the coefficients of the polynomials generating the ideal we are killing, and then we may define

$$R_B = B[X_1, \ldots, X_d, Y_1, \ldots, Y_h, Z_1, \ldots, Z_s]/\big(F_j(X,Y), G_k(X,Y,Z)\big).$$

If K has characteristic $p > 0$, then B is a finitely generated $(\mathbb{Z}/p\mathbb{Z})$-algebra.

We make the convention that if C is a B-algebra then

$$R_C = C \otimes_B R_B \cong C[X_1, \ldots, X_d, Y_1, \ldots, Y_h, Z_1, \ldots, Z_s]/\big(F_j(X,Y), G_k(X,Y,Z)\big).$$

To complete the proof of the theorem, we shall prove two facts: the first is that for every maximal ideal μ of B, B/μ is a finite field. The second is that for the maximal ideals μ in some nonempty open subset of MaxSpec (B), $(x_1, \ldots, x_d)R_{B/\mu}$ has height d.