Math 711: Lecture of September 20, 2004

We need one more preliminary result:

Lemma. *If $A \subseteq A[\tau] \subseteq B$ with B integral over $A[\tau]$ and A integrally closed in B, and there is a monic polynomial F with coefficients in A such that $F(\tau)B \subseteq A[\tau]$, then $B = A[\tau]$.*

Proof. Let $b \in B$ be arbitrary. Then $F(\tau)b = G(\tau)$ for some polynomial G with coefficients in A, since $F(\tau)B \subseteq A[\tau]$. By the division algorithm for monic polynomials we can write $G = QF + H$, where H is either 0 or of degree smaller than that of F, and where Q and H are polynomials with coefficients in A. Let $c = b - Q(\tau)$. It will suffice to show that $c \in A$, and for this it will suffice to show that c is integral over A. Now $F(\tau)b = F(\tau)(Q(\tau)+c)$, but $F(\tau)b = G(\tau) = Q(\tau)F(\tau) + H(\tau)$ as well, and so $F(\tau)c = H(\tau)$. Since $\deg(H) < \deg(F)$, this implies that $\tau/1 \in B_c$ is integral over the ring A'_c, where A' denotes the image of A in B_c. Since B is integral over $A[\tau]$, we have that B_c is integral over A'_c, and, in particular, $c/1 \in B_c$ is integral over A'_c. If we write down an equation of integral dependence in which the coefficients have common denominator c^M, we get

$$\left(\frac{c}{1}\right)^d + \frac{a_{d-1}}{c^M}\left(\frac{c}{1}\right)^{d-1} + \cdots + \frac{a_0}{c^M} = 0$$

and, multiplying by c^M, we have that

$$c^{M+d} + a_{d-1}c^{d-1} + \cdots + a_0$$

has image 0 in B_c, and so is killed by a power of c. This shows that c is integral over A, and so is in A, as required. $\square$

Proof of Zariski's Main Theorem: the finale. We shall now show that if J, the conductor of S into $R[t]$, is contained in Q that Q is not isolated in its fiber, which will complete the proof of the theorem. Recall that R is integrally closed in S and that S is module-finite over $R[\theta]$.

If $J \subseteq Q$ we can choose a minimal prime q of J in S such that $q \subseteq Q$. Let $p = q \cap R$. Let t denote the image of θ in the ring S/q. Now, $R/p \subseteq (R/p)[t] \subseteq S/q$, and S/q is integral over $(R/p)[t]$. The fact that Q is isolated in its fiber over P implies that Q/q is isolated in its fiber over $P/p \subseteq R/p$. By the final Lemma in the Lecture Notes from September 17, this means that t cannot be transcendental over R/p. We complete the proof by obtaining a contradiction.

If t is algebraic over R/p, localize at p and consider the image τ of θ in S_p. We then obtain a monic polynomial F with coefficients in R_p such that $F(\tau) \in qS_p$. Since q is a minimal prime of J, there exist a positive integer N and $w \in S - q$ such that $w\big(F(\tau)\big)^N \in JS_p$. Replacing F by F^N, we have F monic over R_p such that $wF(\tau) \in JS_p$. We now apply the Lemma above with $A = R_p$ and

$$B = R_p[\tau,\, wS_p] = R_p[\tau] + wS_p \subseteq S_p.$$

Note that since $F(\tau)w \in JS_p$, we have that

$$F(\tau)B \subseteq A[\tau] + JS_pS_p = A[\tau] + JS_p \subseteq A[\tau],$$

since $JS \subseteq R[\theta]$. Note also that A is integrally closed in B, because R_p is integrally closed in S_p. The Lemma above now applies, and we can conclude that $B = A[\tau]$, and so $wS_p \subseteq A[\tau]$. Since S is module-finite over $R[\theta]$, this implies that for some element $g \in R - p$, $gwS \subseteq R[\theta]$. But $g \notin q$ and $w \notin q$, and so $gw \notin q$, while we have just shown that $gw \in J$. This is a contradiction, since $J \subseteq q$ by our choice of q. $\square$

We next want to define smooth, étale, and unramified algebras. We first need to discuss finitely presented algebras a bit. Let R be any commutative ring. An R-algebra S is called *finitely presented* if it is finitely generated and for some set of R-algebra generators $s_1, \ldots, s_n$ of S, the ideal of polynomial relations on $s_1, \ldots, s_n$ over R is a finitely generated ideal. In more detail, note that a choice of R-algebra generators $s_1, \ldots, s_n$ yields a homomorphism of a polynomial ring $R[X_1, \ldots, X_n] \to S$ that is surjective, and so we have that $S \cong R[X_1, \ldots, X_n]/I$, where I is the kernel. I is the ideal of polynomial relations on $s_1, \ldots, s_n$ over R, and so we are requiring that I have some finite set of generators, say $F_1, \ldots, F_m$. That is, S is finitely presented over R if and only if it has the form $R[X_1, \ldots, X_n]/(F_1, \ldots, F_m)$, where $X_1, \ldots, X_n$ are indeterminates.

It is reasonable to ask what happens if one chooses a different finite set of algebra generators for S over R. The answer is that the ideal of polynomial relations is still finitely generated. It suffices to compare each set of generators with the union of the two sets, and so we may assume that one set is contained in the other, say $s_1, \ldots, s_n$ and $s_1, \ldots, s_{n+h}$. By induction on h it suffices to consider the case where $h = 1$, i.e., to compare $s_1, \ldots, s_n$ and $s_1, \ldots, s_n, s$ where $s = s_{n+1}$. Since s is in the R-subalgebra generated by $s_1, \ldots, s_n$ we can choose a polynomial $F \in R[X_1, \ldots, X_n]$ such that $s = F(s_1, \ldots, s_n)$. Consider the R-algebra map $T = R[X_1, \ldots, X_n] \twoheadrightarrow S$ such that $X_j \mapsto s_j$, $1 \leq j \leq n$: let I be the kernel. We extend this to $T[X]$ so that $X \mapsto s$. It is easy to verify that the new kernel is $J = IT[X] + (X - F)T[X]$. Clearly, if I is finitely generated, so is J. On the other hand, if J is finitely generated we use the fact that there is an algebra retraction $T[X] \to T$ that fixes T and maps X to F: the image of J under this map is clearly I, and so a finite set of generators for J will map to a finite set of generators for I.

If S is finitely presented over R and T is finitely presented over S then T is finitely presented over R. For suppose that

$$S \cong R[X_1, \ldots, X_n]/(F_1, \ldots, F_m) \text{ and } T \cong S[Y_1, \ldots, Y_k]/(G_1, \ldots, G_h).$$

Each G_j can be lifted to an element H_j of $R[X_1, \ldots, X_n, Y_1, \ldots, Y_k]$ by lifting every coefficient to an element of $R[X_1, \ldots, X_n]$ that maps to it. Then

$$T \cong R[X_1, \ldots, X_n, Y_1, \ldots, Y_k]/(F_1, \ldots, F_m, H_1, \ldots, H_h),$$

as required.

Finally, note that a localization of R at one (and, hence, at finitely many) elements is finitely presented: $R_f \cong R[X]/(fX-1)$. A localization (at any multiplicative system) of a finitely presented R-algebra is called *essentially finitely presented.* If the multiplicative system is finitely generated, the word "essentially" is not needed.

Let S be an R-algebra. Let $(T,\, J)$ be a pair consisting of an R-algebra T and an ideal $J \subseteq T$ such that $J^2 = 0$. Then there is an obvious map

$$\Theta_{T,J} : \mathrm{Hom}_{R-\mathrm{alg}}(S, T) \to \mathrm{Hom}_{R-\mathrm{alg}}(S,\, T/J)$$

that sends $f : S \to T$ to its composition $\gamma \circ f$ with the natural surjection $\gamma : T \twoheadrightarrow T/J$. By a *smooth* R-algebra S we mean a finitely presented R-algebra such that for all R-algebras T and $J \subseteq T$ with $J^2 = 0$, the map $\Theta_{T,J}$ is surjective. By an *étale* R-algebra S we mean a finitely presented R-algebra such that for all R-algebras T and $J \subseteq T$ with $J^2 = 0$, the map $\Theta_{T,J}$ is bijective. By an *unramified* R-algebra S we mean a finitely presented R-algebra such that for all R-algebras T and $J \subseteq T$ with $J^2 = 0$, the map $\Theta_{T,J}$ is injective.

Thus, given an R-algebra map $S \to T/J$ with $J^2 = 0$, if S is smooth over R it has at least one lifting to an R-algebra map $S \to T$. If S is étale, it has a unique such lifting. If S is unramified it has at most one such lifting (but there may not be any lifting).

Note that some authors give these definitions while only requiring S to be essentially finitely presented rather than finitely presented. This creates only small differences in the theory and is convenient in certain ways, while adding a few complications in other ways. Instead, we shall add the word "essentially" for this case and talk about maps that are essentially smooth, essentially étale, or essentially unramified.

Our next objective is to give many other characterizations of these three properties. These characterizations should offer deep insight into the nature of morphisms with these properties.

We first need to review the notions of *derivation* and of the *module of Kähler differentials.* A *derivation* of a ring R into an R-module M is a map $D : R \to M$ such that for all $f,\, g \in R$, $D(f+g) = D(f) + D(g)$ and $D(fg) = fD(g) + gD(f)$. The kernel of D is a subring of R: note that $D(1 \cdot 1) = 1D(1) + 1D(1)$ and so $D(1) = D(1) + D(1)$ and $D(1) = 0$. A derivation D always kills the image of $\mathbb{Z}$ in R. If R is an A-algebra then a derivation $D : R \to M$ is A-linear if and only if the image of A is killed by D. (If it is A-linear then $D(a \cdot 1) = aD(1) = 0$, while if the image of A is killed then $D(af) = D\big((a \cdot 1)f\big) = (a \cdot 1)D(f) + fD(a \cdot 1) = aD(f) + f \cdot 0 = aD(f)$.) An A-linear derivation is also called an *A-derivation.* The set of A-derivations $\mathrm{Der}_A(R,\, M)$ is an R-submodule of the set of all derivations $\mathrm{Der}\,(R,\, M)$ of R into M. The module structure may be described as follows: the value of $D_1 + D_2$ on $f \in R$ is $D_1(f) + D_2(f)$, and the value of rD on f is $rD(f)$.

There is a "universal" A-derivation from R into a specially constructed R-module $\Omega_{R/A}$. We sketch the construction. Let W be the free R-module with a basis in $\{b_f : f \in R\}$ in bijective correspondence with the elements of R. We want to kill a submodule of W in such a way that the map that takes $f \in R$ to the image of b_f in the quotient of W

by this submodule is a derivation. We therefore kill the R-submodule of W spanned by the elements $b_{f+g} - b_f - b_g$ and $b_{fg} - fb_g - gb_f$ for $f,\, g \in R$, and also $b_{rf} - rb_f$ for all $r \in R$ in the image of A and all $f \in R$. Let $\Omega_{R/A}$ denote the quotient of W by the span V of all these elements. It should be clear that the map $d : R \to \Omega_{R/A}$ that sends f to the image of b_f is an A-derivation. We therefore use *df* to denote the image of b_f. The map $d : R \to \Omega_{R/A}$ has the following universal property: given any A-derivation D of R into an R-module M, there is a unique R-linear map $T : \Omega_{R/A} \to M$ such that $D = T \circ d$. Thus, every A-derivation arises from d, uniquely, by composition with an R-linear map. (It is straightforward to check that the composition of an A-derivation with an R-linear map is an A-derivation.) Otherwise said, for every R-module M we have an isomorphism $\mathrm{Hom}_R(\Omega_{R/A},\, M) \cong \mathrm{Der}_A(R,\, M)$. This is the universal mapping property of $d : R \to \Omega_{R/A}$. (Another way of thinking about this is to note that $d : R \to \Omega_{R/A}$ represents the functor $M \mapsto \mathrm{Der}_A(R,\, M)$ in the category of R-modules.) This mapping property determines $d : R \to \Omega_{R/a}$ uniquely up to unique isomorphism.

The proof that every A-derivation D of $R \to M$ arises uniquely from an R-linear map $\Omega_{R/A} \to M$ is straightforward. Given D, it is clear that if one composes the required linear map with the quotient surjection $W \twoheadrightarrow \Omega_{R/A}$, it must map $b_f \mapsto D(f)$ for every $f \in R$. This shows uniqueness. There is certainly a unique such map from W to M. All that is needed is to show that it kills all the elements whose span V we took in constructing $\Omega_{R/A}$ as W/V. But these are killed precisely because D is an A-derivation. We shall have a lot more to say about derivations and Kähler differentials, but at this point we want to explain how to use them to characterize smooth, étale, and unramified homomorphisms. The proof of the theorem we state next will occupy us for quite a while.

Theorem. *Let S be a finitely presented R-algebra.*
(a) *If R contains the rationals, S is smooth over R if and only if S is flat over R and $\Omega_{S/R}$ is projective as an R-module.*
(b) *S is étale over R if and only if S is flat over R and $\Omega_{S/R}$ is 0.*
(c) *S is unramified over R if and only if $\Omega_{S/R} = 0$.*

In part (a), when R does not necessarily contain the rationals a supplementary condition is needed: e.g., S is smooth over R if and only if for every maximal ideal Q of S lying over P in R, $(\Omega_{S/R})$ is S_Q-free of rank equal to the dimension of S_Q/PS_Q.