Math 615 LECTURE NOTES, WINTER, 2016

by Mel Hochster

Lecture of January 6

This course will deal with several topics in the theory of commutative Noetherian rings, including the following:

(1) The theory of Gröbner bases and applications: a lot more about this momentarily.

(2) The structure theory of complete local rings. One strategy in studying problems over Noetherian rings is to reduce first to the local case, and then to the complete local case. The structure theory of complete local rings can then be applied. There are even deep theorems that permit one to pass from the case of a complete local ring to a finitely generated algebra over a field or complete discrete valuation ring. Other techniques can be used to pass from a problem in such an algebra over a field of characteristic 0 to a corresponding problem over a field, even a finite field, of positive prime characteristic p.

(3) What can one do when a ring is not Cohen-Macaulay?

In particular, we will discuss the theory of Cohen-Macaulay rings, but will focus on techniques that show that all local rings are, in some sense, close to being Cohen-Macaulay.

Although we shall discuss the subject in much greater detail later, we give a brief discussion of Cohen-Macaulay rings here so that we can explain the sort of theorem we want to prove.

Recall the a ring R is *quasilocal* if it has a unique maximal ideal m: in this case we usually denote the residue class field R/m by K, and refer to the quasilocal ring (R, m, K). We reserve the term *local ring* for a Noetherian quasilocal ring.

Let (R, m, K) be a local ring of Krull dimension d. This implies that there exist d elements $x_1, \ldots, x_d \in m$ such that if $I = (x_1, \ldots, x_d)R$, then $\mathrm{Rad}\,(I) = m$. (One cannot use fewer than d elements, by the Krull height theorem.) Such a sequence of elements $x_1, \ldots, x_d$ is called a *system of parameters*. A d-tuple $(r_1, \ldots, r_d)$ is called a *relation* on $x_1, \ldots, x_d$ if

$$\sum_{j=1}^{d} r_j x_j = 0.$$

The relations are easily seen to be an R-submodule of the free R-module R^d. There are some obvious relations: the element

$$(0, \ldots, 0, x_j, 0, \ldots, -x_i, 0, \ldots, 0)$$

where x_j occurs in the i th spot and $-x_i$ occurs in the j th spot, is a relation. The elements in the R-span of these $\binom{d}{2}$ relations are referred to as *trivial relations*.

A local ring is called *Cohen-Macaulay* if there is a system of parameters such that every relation on the parameters is trivial. It then follows by a theorem that this is true for *every* system of parameters. By a theorem, this property passes to localizations. One then defines an arbitrary Noetherian ring to be *Cohen-Macaulay* if all of its localizations at maximal ideals (equivalently, at prime ideals) are Cohen-Macaulay.

In certain graded cases one can give an alternative characterization as follows. Let K be a field and R an $\mathbb{N}$-graded algebra (i.e., R has a direct sum decomposition $R = \bigoplus_{n=0}^{\infty} R_n$ with $1 \in R_0$ satisfying $R_m R_n \subseteq R_{m+n}$ for all $m, n \in \mathbb{N}$) such that R is finitely generated over $R_0 = K$. In this case, it turns out that one can always choose forms $F_1, \ldots, F_d$ of positive degree in R (by raising the F_j to various powers one can even arrange that they all have the same degree) such that $F_1, \ldots, F_d$ are algebraically independent over K and R is module-finite over $A = K[F_1, \ldots, F_d]$. Of course, A is isomorphic with a polynomial ring in d variables over K. In this situation, R is Cohen-Macaulay if and only if R is free as an A-module.

In higher dimension, it is rare for modules over polynomial rings to be free. In fact, relatively few rings are Cohen-Macaulay. In equal characteristic 0, one can start taking module-finite extensions of a polynomial ring: if the dimension is 3 or higher, all sufficiently large such extensions fail to be Cohen-Macaulay.

Examples. Let $S = K[x, y]$ be the polynomial ring in two variables over the field K. Let $R = K[x^2, xy, y^2] \subseteq S$. One may take $A = K[x^2, y^2] \subseteq R$. Then R is free over A on the basis 1, xy, and so is Cohen-Macaulay.

On the other hand, let $R_1 = K[x^2, x^3, xy, y] \subseteq S$ and let $A_1 = K[x^2, y] \subseteq R_1$. Then R_1 is module-finite over A_1 with minimal generators 1, x^3, xy, but is *not* free over A_1. One has that $y(x^3) - x^2(xy) = 0$. This relation on minimal generators shows that R_1 is *not* A_1-free and therefore *not* Cohen-Macaulay. Alternatively, in the local ring of R_1 at its homogeneous maximal ideal, x^2, y is a system of parameters and $(xy, -x^3)$ is a non-trivial relation on x^2, y.

However, many of the rings that arise in natural geometric situations, such as complete intersections and rings defined by the vanishing of minors of a matrix of indeterminates *are* Cohen-Macaulay.

Many problems become easier in Cohen-Macaulay rings. One of the results we are aiming to prove, stated in a very special case, helps to remedy the situation when the ring is not Cohen-Macaulay:

Theorem. *Let R be a complete local domain of prime characteristic $p > 0$. Let $x_1, \ldots, x_d$ be a system of parameters for R, and let $(r_1, \ldots, r_d)$ be a relation on $x_1, \ldots, x_d$. Then there is a complete local module-finite extension domain S of R such that the relation $(r_1, \ldots, r_d)$ becomes trivial over S.*

This result has been known for well over two decades: cf. [M. Hochster and C. Huneke, *Infinite integral extensions and big Cohen-Macaulay algebras*, Annals of Math. **135** (1992), 53–89]. Recently there have been improvements: one can make all relations on all systems of parameters become trivial after just one module-finite extension (but new relations may be introduced). Beyond that, more recently, global versions of this theorem have been proved. We shall discuss the situation in detail later in the course.

The Theorem above turns out to be false when R contains a field of characteristic 0. Nonetheless, one can use the characteristic p results to prove important theorems in equal characteristic 0.

We next want to begin our systematic treatment of the theory of Gröbner bases. Before doing so we shall review some facts about closed algebraic sets in K^n over an algebraically closed field K.

Review of the behavior of closed algebraic sets over an algebraically closed field

This section is meant as an overview of some basic results on closed algebraic sets over an algebraically closed field. We give definitions and statements of some theorems, but most proofs are omitted. For a detailed treatment of this material, the reader may consult [R. Hartshorne, *Algebraic Geometry*, Springer-Verlag Graduate Texts in Mathematics **52**, New York • Berlin • Heidelberg, 1977], Chapter I. There is also a complete discussion in the Lecture Notes from Math 614, Fall 2015: see particularly the Lectures from October 9 through October 26.

Let K be an algebraically closed field, and let $R = K[x_1, \ldots, x_n]$ be a polynomial ring. If $W \subseteq R$ is any set,

$$\mathcal{V}(W) = \{v \in K^n : f(v) = 0 \text{ for all } f \in W\}.$$

It is easy to see that if I is the ideal generated by W, $\mathcal{V}(I) = \mathcal{V}(W)$. Moreover, if $f \in \operatorname{Rad}(I)$, i.e., $f^k \in I$ for some integer $k \geq 1$, then f also must vanish on $\mathcal{V}(I)$, and so $\mathcal{V}\big(\operatorname{Rad}(I)\big) = \mathcal{V}(W)$ as well. If $X = \mathcal{V}(I)$ for some ideal I, we say the X is a *closed algebraic set* in K^n, or a *Zariski* closed set in K^n. In fact, we have

(1) $K^n = \mathcal{V}(0)$ and $\emptyset = \mathcal{V}(R)$ are closed algebraic sets.

(2) $\mathcal{V}(I \cap J) = \mathcal{V}(I) \cup \mathcal{V}(J)$ for any two ideals I and J.

(3) $\mathcal{V}(\sum_{\lambda\in\Lambda} I_\lambda) = \bigcap_{\lambda\in\Lambda} \mathcal{V}(I_\lambda)$ for any family of ideals $\{I_\lambda\}_{\lambda\in\Lambda}$.

The conditions above show that the closed algebraic sets are, in fact, the closed sets of a topology on K^n: this is called the *Zariski topology*.

Suppose that we are given an arbitrary set of points $\mathcal{P} \subseteq K^n$ and we want to understand the Zariski closure $\overline{\mathcal{P}}$ of $\mathcal{P}$. Since this will be the smallest closed set containinng $\mathcal{P}$, we want to find I as *large* as possible such that $V(I) \supseteq \mathcal{P}$. But any element of I must vanish

on $V(I)$, which we want to contain $\mathcal{P}$. Therefore, the largest ideal we can use is the ideal of *all* functions in $K[x_1, \ldots, x_n]$ that vanish on $\mathcal{P}$, and this ideal defines $\overline{\mathcal{P}}$.

Note that if $n = 1$, the closed sets in K are the finite sets and K itself. In K^2 one gets finite unions of points and/or curves defined by one equation, and K^2 itself.

Every closed algebraic set $X \subseteq K^n$ inherits a Zariski topology, whose closed sets are simply the closed algebraic sets in K^n that happen to be contained in X.

The fundamental result in this area is:

Hilbert's Nullstellensatz. *Let K be an algebraically closed field and $R = K[x_1, \ldots, x_n]$ a polynomial ring over K. There is a bijective, order-reversing correspondence between closed algebraic sets in K^n and radical ideals of $K[x_1, \ldots, x_n]$. Under this correspondence, the radical ideal J corresponds to $\mathcal{V}(J)$, and the algebraic set X corresponds to the ideal $\mathcal{I}(X) = \{f \in R : \text{for all } v \in X,\ f(v) = 0\}$. In particular, the maximal ideals of R are in bijective correspondence with the points of K^n: given a point $v = (c_1, \ldots, c_n) \in K^n$, the corresponding maximal ideal consists of all polynomials that vanish at v. (It can also be described in terms of generators as the maximal ideal $(x_1 - c_1, \ldots, x_n - c_n)R$.)*

It follows that polynomials in $K[x_1, \ldots, x_n]$ have a common vanishing point if and only if they do not generate the unit ideal, and that $f \in \mathrm{Rad}\,(f_1, \ldots, f_m)$ if and only if f vanishes on $\mathcal{V}(f_1, \ldots, f_m)$. (Each of these statements is sometimes referred to as "Hilbert's Nullstellensatz.")

When $X = \mathcal{V}(I)$ we shall say that I is a *defining ideal* for X. When, in addition, I is radical we shall sometimes say that I is *the* defining ideal of X: it is now uniquely determined by X.

We want to make the closed algebraic sets over K into a category. When we want to emphasize that K^n is being thought of as an algebraic set, we use the notation $\mathbb{A}^n_K$ for K^n. Given closed algebraic sets $X \subseteq \mathbb{A}^m_K$ and $Y \subseteq \mathbb{A}^n_K$, we define a *$K$-regular* map or *$K$-morphism* from X to Y to be a function $\theta : X \to Y$ that is the restriction of a map $\mathbb{A}^m_K \to \mathbb{A}^n_K$ that is given in terms of coordinates by polynomials. That is, there are n polynomials $f_1, \ldots, f_n \in K[x_1, \ldots, x_m]$ such that for every point $v \in X$, $\theta(v) = \big(f_1(v), \ldots, f_n(v)\big)$.

Note that every K-regular map from X to Y is the restriction (where we restrict both the domain and the target) of a K-regular map $\mathbb{A}^m_K \to \mathbb{A}^n_K$. This may seem at first to be an unreasonably strong requirement, but one should keep in mind that given closed sets $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$, every continuous function from X to Y is the restriction of a continuous function from $\mathbb{R}^m \to \mathbb{R}^n$. To see this, one must show that the composition $X \to Y \subseteq \mathbb{R}^n$ extends to a map on $\mathbb{R}^m$. The composition is given in coordinates by n continuous maps $X \to \mathbb{R}$, and each of these can be extended to $\mathbb{R}^m$ by the Tietze extension theorem.

The identity map $X \to X$ is K-regular, and the composition of two K-regular maps is K-regular, so that the closed algebraic sets and K-regular maps form a category.

Given a closed algebraic set $X \subseteq \mathbb{A}^n_K$, the K-regular maps from X to $\mathbb{A}^1_K = K$ are

simply the maps $X \to K$ arising from the restriction of a polynomial $f \in K[x_1, \ldots, x_n]$ to X. This set of maps forms a K-algebra, denoted $K[X]$, and called the *coordinate ring* of X. We have a surjection $K[x_1, \ldots, x_n] \twoheadrightarrow K[X]$ induced by restriction. The kernel is precisely the set of polynomials that vanish on X, or $\mathcal{I}(X)$, and so

$$K[X] \cong K[x_1, \ldots, x_n]/\mathcal{I}(X)$$

as K-algebras.

Recall that a ring is called *reduced* if every nilpotent element is 0. Then $K[X]$ is a reduced, finitely generated K-algebra. What is more, every reduced, finitely generated K-algebra occurs, up to K-algebra isomorphism, as $K[X]$ for some closed algebraic set X. For given such a K-algebra R, we may map $K[x_1, \ldots, x_n] \twoheadrightarrow R$ by choosing a finite set of, say, n generators for R as a K-algebra and sending the the x_j to these generators. The kernel is a radical ideal J. By Hilbert's Nullstellensatz, $J = \mathcal{I}(X)$ for a unique closed algebraic set X in $\mathbb{A}^n_K$. But then

$$R \cong K[x_1, \ldots, x_n]/J = K[x_1, \ldots, x_n]/\mathcal{I}(X) \cong K[X].$$

The map $X \mapsto K[X]$ is a contravariant functor from closed algebraic sets to reduced finitely generated K-algebras. Given a K-regular map $X \to Y$, one obtains a K-algebra homomorphism $K[Y] \to K[X]$ in an obvious way by composition: an element of $K[Y]$ is precisely a K-regular map $Y \to \mathbb{A}^1_K$, and the composite map $X \to Y \to \mathbb{A}^1_K$ is an element of $K[X]$.

The key result about this is:

Theorem. *Let K be an algebraically closed field. The category of closed algebraic sets over K and K-regular maps is anti-equivalent to the category of reduced, finitely generated K-algebras. The functor $X \mapsto K[X]$ provides the anti-equivalece.*

We shall not give a complete argument here, but we do indicate how one gets a contravariant functor from finitely generated reduced K-algebras to closed algebraic sets over K. The main point is that given a finitely generated reduced K-algebra R, one can give the set X of K-homomorphisms $R \twoheadrightarrow K$, which is in bijective correspondence with the set of maximal ideals of R, the "structure" of a closed algebraic set, i.e., one can put this set in bijective correspondence with the points of a closed algebraic set. Choices are made in setting up this correspondence, but the different closed algebraic sets obtained are canonically isomorphic in the category of closed algebraic sets.

Specifically, one simply maps a polynomial ring $K[x_1, \ldots, x_n] \twoheadrightarrow R$. Suppose that $R \cong K[x_1, \ldots, x_n]/J$, where J will be radical. Each K-algebra homorphism $R \twoheadrightarrow K$ gives a composite homomorphism $K[x_1, \ldots, x_n] \twoheadrightarrow R \twoheadrightarrow K$, and this map corresponds to a maximal ideal of $K[x_1, \ldots, x_n]$ and, hence, to a point of $\mathbb{A}^n_K$. The points obtained are precisely the points of $\mathcal{V}(J)$, which is therefore a closed algebraic set in bijective correspondence with $X = \mathrm{Hom}_{K-\mathrm{alg}}(R, K)$.

Some motivations for introducing Gröbner bases

Gröbner bases are a tool for doing explicit algorithmic calculations in a polynomial ring over a field K (or in a homomorphic image of a polynomial ring over K). Whether Gröbner basis methods actually give an algorithm depends on whether one can perform operations in K algorithmically. We shall not worry about this point. We simply assume that arithmetic operations in K are understood, and seek methods to solve problems in polynomial rings under the presumption that simple manipulations over the field can handled.

In dealing with Gröbner basis questions, unless otherwsie specified, K is always understood to be a field, and a given ring $K[x_1, \ldots, x_n]$ is meant to be assumed to be a polynomial ring in variables $x_1, \ldots, x_n$ over K.

We want to mention right away that while Gröbner bases are tools for calculation, they can also be used to prove substantial theorems, such as the Hilbert basis theorem (ideals in R are finitely generatded) and the Hilbert syzygy theorem (which is discussed further below). There are also many instances in which Gröbner basis techniques have been used to prove that certain infinite classes of rings of a special form have good properties.

Moreover, not surprisingly, the systematic study of Gröbner bases introduces a great many new theoretical problems.

Among the questions we want to consider are the following.

(1) Given generators $f_1, \ldots, f_m$ for an ideal of the polynomial ring $R = K[x_1, \ldots, x_n]$, how do we tell whether a given element $f \in R$ of the polynomial ring is in the ideal?

This is equivalent to determining whether one can solve the equation

$$f = U_1 f_1 + \cdots + U_m f_m$$

where the U_j are unknown elements of R.

If one knows an *a priori* bound D for the degrees of the unknown polynomials U_j in terms of m, n, and the degrees of the f_j, one can think of the U_j as polynomials of degree at most D with unknown coefficients. By working with coefficients, one gets a system of linear equations over K in the unknown coefficients, and the problem becomes pure linear algebra. The trouble with this idea is that while bounds for D are known, they are double exponential, making the implementation of this idea unfeasible. The complexity of the problem is double exponential in theory in worst cases, but the method of Gröbner bases often works in cases that arise in practice.

A similar problem arises in determining whether a given element is in the R-span of finitely many specified elements in the free module R^s. We shall give a Gröbner basis method that can be used for both of these problems.

(2) If we have finitely many generators for an ideal

$$I \subseteq K[x_1, \ldots, x_n] = R,$$

and $1 \leq s \leq n-1$, how can we find finitely many generators for $I \cap K[x_{s+1}, \ldots, x_n]$?

Here, we might intersect with the polynomial subring generated by an arbitrary subset of the variables, but by renumbering the indeterminates we might as well assume that the generators of the subring are $x_{s+1}, \ldots, x_n$. This type of question is part of what is called *elimination theory*: we are eliminating the variables $x_1, \ldots, x_s$ from the equations.

This sort of problem is intimately connected with the problem of solving explicitly the equations obtained by setting the generators of the ideal equal to 0.

To make this connection, we discuss the situation where K is algebraically closed.

We first want to understand the geometric meaning of the intersection of the ideal with the subring.

Proposition. *Let K be algebraically closed and let $I \subseteq K[x_1, \ldots, x_n]$ be any ideal. Suppose that $1 \leq s \leq n-1$ and let $J = I \cap K[x_{s+1}, \ldots, x_n]$. Let $\pi : \mathbb{A}^n_K \to \mathbb{A}^{n-s}_K$ be projection on the last $n-s$ coordinates. Let $X = \mathcal{V}(I) \subseteq \mathbb{A}^n_K$. Then $\mathcal{V}(J)$ is the Zariski closure of the projection $\pi(X)$.*

Proof. Let $f \in K[x_{s+1}, \ldots, x_n]$. By the discussion in the next to last paragraph on p. 3, f is in the defining ideal of Zariski closure of $\pi(X)$ if and only if f vanishes on $\pi(X)$, i.e., $f\big(\pi(X)\big) = 0$. This says that $f \circ \pi$, which is simply f thought of as a function on all of K^n (even though it only involves $x_{s+1}, \ldots, x_n$), vanishes on X, i.e., that $f \in I$. Thus, $I \cap K[x_{s+1}, \ldots, x_n]$ is a defining ideal for the Zariski closure of $\pi(X)$. $\square$

Now suppose that $I = (f_1, \ldots, f_m) \subseteq K[x_1, \ldots, x_n]$. Note that the simultaneous solutions of the system

$$\begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ \qquad\qquad \cdots \\ f_m(x_1, \ldots, x_n) = 0 \end{cases}$$

is the same as the set $V(I)$. Assume that $V(I)$ is a finite set. We next want to show if we have an algorithmic method for doing elimination theory, then we also have an algorithmic method for finding the solutions $V(I)$, *provided that we have a method for solving one polynomial equation in one variable over K*. The assumption that $V(I)$ is finite is not essential: if $V(I)$ is infinite, the method will show that.

The idea is very simple. One calculates $I \cap K[x_n]$. This is a principal ideal, since $K[x_n]$ is a PID. Thus, one gets a single generator $g(x_n) \in K[x_n]$ for the intersection. By the Proposition above, $gR[x_n]$ defines the Zariski closure of the projection of $V(I)$ on $\mathbb{A}^1_K = K$ corresponding to the last coordinate. There are three cases.

First case. The intersection is the (0) ideal. This implies that the Zariski closure of the projection is all of K, which means that the projection is an infinite set.

Second case. The intersection is the unit ideal, i.e., g is a nonzero constant. In this case, the projection is empty, and this means that there are no solutions.

Third case. g is a polynomial of positive degree. We are assuming that in this case we can find the roots of g in K: call them $\lambda_1, \ldots, \lambda_k$. This means that the closure of the projection of $V(I)$ is the set $\{\lambda_1, \ldots, \lambda_k\}$. This implies that the projection is finite, and since finite sets are closed, we must have that $\{\lambda_1, \ldots, \lambda_k\}$ *is* the projection. This means that the last coordinate of each point in $V(I)$ is one of $\lambda_1, \ldots, \lambda_k$, and that every λ_j occurs as the last coordinate of some point of $V(I)$. The problem of solving the original system of equations now breaks up into k separate problems, one for every λ_j. To find the points of $V(I)$ whose last coordinate is λ_j, substitute λ_j for x_n in each of the equations. This produces a new system of equations, but the number of variables is one smaller. Proceeding recursively, we eventually find all solutions of the original system.

(3) Another use of Gröbner bases is in solving the following kind of problem: given elements $f_1, \ldots, f_s \in K[x_1, \ldots, x_n]$, find generators for all the relations on those elements, i.e., for the module of s-tuples of polynomials $(g_1, \ldots, g_s)$ such that $\sum_{j=1}^{m} g_j f_j = 0$. In fact, one can require insteasd that the g_i satisfy several equations like this, i.e., a system

$$\sum_{j=1}^{s} g_j f_{i,j} = 0, \quad 1 \leq i \leq r.$$

This is equivalent to finding the relations on the s columns of the $r \times s$ matrix $\mathcal{M} = \big(f_{i,j}\big)$, i.e., to finding the all the column vectors $\underline{g} = \begin{pmatrix} g_1 \\ \vdots \\ g_s \end{pmatrix}$ such that

$$\mathcal{M}\underline{g} = 0.$$

Consider the R-submodule M of R^s spanned by these columns. The module of relations on the columns is called a *first module of syzygies* of M. More generally, whenever we have a short exact sequence of finitely generated R-modules $0 \to M' \to R^k \to M \to 0$, M' is called a *first module of syzygies* of M. A first module of syzygies of a k th module of syzygies is called a $(k+1)$ *st module of syzygies*: when N is an n th module of syzygies of M there is an exact sequence

$$0 \to N \to R^{b_{n-1}} \to \cdots \to R^{b_0} \to M \to 0$$

of finitely generated R-modules.

Gröbner bases can be used to prove the famous Hilbert syzygy theorem, that every finitely generated module over $K[x_1, \ldots, x_n]$ has an n th module of syzygies that is free. (Equivalently, that every finitely generated R-module has a free resolution of length at most n.) Beyond that, Gröbner bases can be used to compute the resolution.

As a further application, Gröbner basis methods can be used in the graded case to calculate Hilbert functions. We shall discuss this in much greater detail, including a review of what is needed from the theory of Hilbert functions, once we have dome some basic Gröbner basis theory.

Lecture of January 8

Monomial Ideals and Submodules

Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over a field K. When a free R-module F is given, it will typically be assumed to be finitely generated with an ordered free basis $b_1, \ldots, b_n$. The ordered free basis provides an isomorphism with R^n under which $\sum_{i=1}^n r_i b_i$ corresponds to $(r_1, \ldots, r_n)$. Therefore, for the most part, in working with a free module with ordered basis, we might as well assume that it is R^n with the standard basis $e_1, \ldots, e_n$, where e_i has 1 in the ith spot and 0 in the other spots. However, especially when we are working with more than one free module, it may be inconvenient to identify all of the modules with various R^{n_i}.

By a *monomial* μ in R, we mean an element of the form $x_1^{a_1} \cdots x_n^{a_n}$ where the $a_i \in \mathbb{N}$, the nonnegative integers. If $\alpha = (a_1, \ldots, a_n) \in \mathbb{N}^n$, we write x^α for $x_1^{a_1} \cdots x_n^{a_n}$. Thus, there is a bijection betwen monomials of R and elements of $\mathbb{N}^n$. We write $\mathcal{M}$ for the set of monomials of R.

More generally, given a finitely generated free module F with ordered basis, by a *monomial* in F we mean an element of the form μb_i, where $\mu \in M$ and b_i is in the ordered basis. In particular, when $F = R^s$ with the standard basis, we mean an element of the form μe_i with $\mu \in \mathcal{M}$.

The monomials of F form a K-vector space basis for F. Every element $f \in F$ is uniquely expressible as a K-linear combination of mutually distinct monomials (for 0, the set of monomials occurring is the empty set). We refer to the monomials that occur as the *monomials of* f. We shall refer to the product of a nonzero element of K with a monomial as a *term*. Thus, every element of F is uniquely expressible as a sum of terms involving mutually distinct monomials: these terms are referred to as the *terms of* f. In particular, this terminolgy applies in the case where $F = R$.

Proposition. *Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over K. Let $F \cong R^s$ be a free module with ordered basis. The following three conditions on a submodule M of F (respectively, an ideal I of R) are equivalent:*

(1) *M (respectively, I) is generated by monomials.*

(2) *M (respectively, I) is the K-span of monomials.*

(3) *If $f \in M$ (respectively, I), the monomials of f are in M (respectively, I).*

Moreover, if M (respectively, I) is generated by monomials ν_λ (the index set may be infinite), then $f \in M$ if and only if every monomial in f is the product of a monomial $\mu \in \mathcal{M}$ and some ν_λ.

Proof. It suffices to consider the module case. Suppose that $\mathcal{G}$ is a family of monomials in F. The submodule generated by $\mathcal{G}$ must contain all the elements $\{\mu\nu : \mu \in \mathcal{M}, \nu \in \mathcal{G}\}$. The K-span of this set of monomials is closed under multiplication by any element of R, by the distributuve law. It follows that (1) $\Rightarrow$ (2). This implies the final statement. Moreover, (2) $\Rightarrow$ (1) and (2) $\Leftrightarrow$ (3) are obvious. $\square$

Of course, it is *not true* that an arbitrary set of monomials spans a submodule: $\mathcal{G}$ spans a submodule if and only if whenever $\nu \in \mathcal{G}$ and $\mu \in \mathcal{M}$, we have that $\mu\nu \in \mathcal{G}$.

Consider a K-vector space with basis $\mathcal{B}$, and let $\mathcal{S}$ be the set of K-vector subspaces that are spanned by a subset of $\mathcal{B}$. Then there is an order-preserving bijection between $\mathcal{S}$ and the set of subsets of $\mathcal{B}$. This bijection preserves intersection, even infinite intersection, and takes sums, even infinite sums, to unions. Thus, for such a family of vector spaces, intersection distributes over sum (even when the sum is infinite) and union distributes over intersection (even if the intersection is infinite).

Since monomial ideals (respectively, monomial submodules) have K-bases consisting of monomials, it follows that for monomial ideals and submodules, intersection distributes over sums, including infinite sums, and sum distributes over intersections, even infinite intersections.

Let $\alpha = (a_1, \ldots, a_n)$ and let $\beta = (b_1, \ldots, b_n)$. Let $c_i = \min\{a_i, b_i\}$ for each i and let $d_i = \max\{a_i, b_i)$ for each i. Let $\gamma = (c_1, \ldots, c_n)$ and $\delta = (d_1, \ldots, d_n)$. We define $\mathrm{GCD}(x^\alpha, x^\beta) = x^\gamma$, and $\mathrm{LCM}(x^\alpha, x^\beta) = x^\delta$. These definitions agree with the usual UFD notions of greatest common divisor and least common multiple.

In particular, $x^\alpha R \cap x^\beta R = x^\delta R$ where $x^\delta = \mathrm{LCM}(x^\alpha, x^\beta)$. Now suppose that I is generated by monomials x^{α_i} where i varies in some index set, and that J is generated by monomials x^{β_j}, where j varies in some index set. Thus, I is the sum of the ideals $x^{\alpha_i}R$ and J is the sum of the ideals $x^{\beta_j}R$. Since intersection distributes over sum for monomial ideals, it follows that $I \cap J$ is the sum of the ideals $x^{\gamma_{ij}}R$, where $\gamma_{ij} = \mathrm{LCM}(\alpha_i, \beta_j)$, since $x^{\gamma_{ij}}R = x^{\alpha_i}R \cap x^{\beta_j}R$ for all choices of i and j.

Now let F be a finitely generated free module with ordered basis $B_1, \ldots, B_s$. We can extend these definitions to pairs of monomials of F that involve the same basis element, so that $\mathrm{GCD}(\mu_1 b_i, \mu_2 b_i) = \mathrm{GCD}(\mu_1, \mu_2)b_i$ and $\mathrm{LCM}(\mu_1 b_i, \mu_2 b_i) = \mathrm{LCM}(\mu_1, \mu_2)b_i$

Lemma. *If $\{a_n\}_{n\in\mathbb{N}}$ is an infinite sequence of nonnegative integers, then it has an infinite subsequence that is either constant or strictly increasing. In particular, it has an infinite subsequence that is non-decreasing.*

Proof. If the sequence is bounded above, then only finitely many integers occur, and so at least one of them must occur infinitely many times. If the sequence is not bounded,

let $n_1 = 1$ and, recursively, let n_{i+1} be the least integer strictly larger than n_i such that $a_{n_{i+1}} > a_{n_i}$. (If there is no such integer, then the sequence is bounded above.) Clearly, $\{a_{n_i}\}_i$ is strictly inreasing. $\square$

The Lemma above is quite easy, but it has an interesting consequence. Let F be a finitely generated free module with ordered basis over $R = K[x_1, \ldots, x_n]$. The set of monomials of F is partiallly ordered by $\nu_1 \geq \nu_2$ means that $\nu_2 = \mu\nu_1$ for some $\mu \in \mathcal{M}$, i.e., ν_2 is a multiple (necessarily by a monomial) of μ_1. Then:

Proposition. *Let R and F be as above. Then there is no infinite subset of F consisting of mutually incomparable monomials. Equivalently, given any infinite sequence of monomials in F, one of them divides another. In particular, this holds when $F = R$.*

Proof. Suppose that $\nu_1, \nu_2, \nu_3, \ldots$ is an infinite sequence of monomials in F. Then some e_i must occur in infinitely many terms, and so we may pass to an infinite subsequence in which each term has the form $\mu_n e_i$. It therefore suffices to prove the result for an infinite sequence $\mu_1, \mu_2, \mu_3, \ldots$ of monomials in R. Consider the exponents $a_1, a_2, a_3, \ldots$ occurring on the variable x_1 in this sequence. Then we may pass to an infinite subsequence such that these exponents are non-decreasing, by the Lemma above. By the same reasoning we may pass to a still smaller infinite subsequence such that the exponents $b_1, b_2, b_3, \ldots$ on x_2 are *also* non-decreasing. By a straightforward induction, we may repeat this step for each variable, and the nth subsequence obtained will have the property that for all of the variables x_i, where $1 \leq i \leq n$, the sequence of exponents on x_i is non-decreasing. But this means that every monomial in the subsequence divides all monomials that come after it in the subsequence. $\square$

Corollary. *Let R and F be as above. Then every monomial submodule M of F is finitely generated by the set of minimal monomials in M under the partial ordering by divisibility. In particular, this holds for monomial ideals in R.*

Proof. Given any monomial in F, there are only finitely monomials in F that are smaller in the partial ordering, and so given any monomial in M, among the monomials in M that divide it there must be a minimal one. Therefore, M is generated by the minimal monomials in M. Since these are mutually incomparable, the preceding Proposition shows that the set of minimal monomials in M is finite. $\square$

The set of minimal monomials in a monomial submodule (or ideal) is also referred to as the *set of minimal monomial generators.*

Gröbner bases reduce a multitude of problems about ideals of R and about arbitrary submodules of a free module F to the monomial case! In particular we shall use them to give a very simple proof of the Hilbert basis theorem. In order to define Gröbner basis, we need to introduce the idea of a monomial order.

Monomial orders

Let $R = K[x_1, \ldots, x_n]$ and let $\mathcal{M}$ be the set of all monomials in R. A *monomial order* on $\mathcal{M}$ is a *total* ordering $>$ of $\mathcal{M}$ such that

(1) If $\mu, \mu_1, \mu_2 \in \mathcal{M}$ and $\mu_1 > \mu_2$ then $\mu\mu_1 > \mu\mu_2$.

(2) The element 1 is the least element in $\mathcal{M}$.

The second property implies that a monomial order refines the partial ordering by divisibility: since $1 \leq \mu_2$ for all $\mu_2 \in \mathcal{M}$, we have that $\mu_1 \leq \mu_1\mu_2$ for all μ_1, $\mu_2 \in \mathcal{M}$ By renumbering the variables, we may assume that $x_1 > x_2 > \cdots > x_n$, and we shall always assume this about any monomial order that we introduce.

By a *monomial order* on a finitely generated free module F with ordered basis $b_1, \ldots, b_s$ we mean a total ordering of the monomials in F such that

(1) If $\mu \in \mathcal{M}$ and ν_1, ν_2 are monomials in F with $\nu_1 > \nu_2$, then $\mu\nu_1 > \mu\nu_2$.

(2) For every i, the element b_i is least among the elements of the form μb_i for μ in $\mathcal{M}$.

Property (2) implies that if $\nu \in F$ is a monomial and $\mu \neq 1$ is in $\mathcal{M}$, then $\nu < \mu\nu$. Evidently, this agrees with the first definition when $F = R$ with the ordered basis consisting of 1.

Given a monomial order $>$ on $\mathcal{M}$ we can construct a monomial order on F by requiring that $\mu_1 b_i > \mu_2 b_j$ precisely when $\mu_1 > \mu_2$ or $\mu_1 = \mu_2$ and $i < j$ (so that $b_1 > b_2 > \cdots b_s$). Unless otherwise specified, we shall always do this in working with monomial orders on free modules.

To see that monomial orders on R exist, we give the example of *lexicographic* order, frequently referred to simply as *lex* order. If $\alpha = (a_1, \ldots, a_n)$ and $\beta = (b_1, \ldots, b_n)$, the definition is simply that $x^\alpha > x^\beta$ precisely if there exists an integer j, where $1 \leq j \leq n$, such that $a_i = b_i$ for $i < j$ while $a_j > b_j$. It is very easy to see that this satisfies (1) and (2) above. Note that it is true that $x_1 > \cdots > x_n$ as well.

Suppose that x_1, x_2, x_3 $\ldots$, x_{26} are the letters of the Roman alphabet, A, B, C, $\ldots$, Z. Suppose that given a monomial we write it out as a string of letters with letters occurring in alphabetical order, so that $x_1^3 x_2 x_3^2 x_4^5$ would be written out as $AAABCCDDDDD$. The order we have specified is the same order as these "words" would occur in a dictionary or lexicon. This is the reason for the term "lexicographic order."

We shall soon see that if R has two or more variables, there are uncountably many monomial orders! However, we really only need to make use of two or three of them.

Lecture of January 11

The definition of lexicographic order is quite simple, but the totally ordered set that one gets is not — even if there are only two variables one has

$$1 < x_2 < x_2^2 < \cdots < x_2^n < \cdots$$

$$< x_1 < x_1x_2 < x_1x_2^2 < \cdots < x_1x_2^n < \cdots$$

$$< x_1^2 < x_1^2x_2 < x_1^2x_2^2 < \cdots < x_1^2x_2^n < \cdots$$

$$\cdots$$

$$< x_1^m < x_1^mx_2 < x_1^mx_2^2 < \cdots < x_1^mx_2^n < \cdots$$

$$\cdots$$

Thus, there are abundantly many examples of infinite increasing sequences that have an upper bound within the set. This ordered set is not order-isomorphic with $\mathbb{N}$.

We will write $\mu' >_{\text{lex}} \mu$ to indicate that we are using lexicographic order, although the subscript may be omitted if it is clear from context which monomial order we mean.

In a way, it is simpler to consider a variant notion called *homogeneous* lexicographic order. This order will be indicated by the subscript $_{\text{hlex}}$. The definition is simple: we define $\mu' >_{\text{hlex}} \mu$ to mean that either that $\deg(\mu') > \deg(\mu)$ or that $\deg(\mu') = \deg(\mu)$ and $\mu' >_{\text{lex}} \mu$. Thus, monomials of larger degree are always bigger in this order, while we use lexicographic order to decided which is bigger of two monomials of the same degree. It is quite easy to verify that this is also a monomial order. In $K[x_1, x_2]$ note that $x_1 >_{\text{lex}} x_2^2$ but that $x_2^2 >_{\text{hlex}} x_1$. It is still the case that $x_1 > x_2 > \cdots > x_n$ in homogeneous lexicographic order. The totally ordered set one gets is easily seen to be order isomorphic with $\mathbb{N}$ for hlex order. Some authors use the term *graded lexicographic order* instead of homogeneous lexicographic order, and use the subscript $_{\text{grlex}}$ to indicate it.

Another monomial order of great important is reverse lexicographic order, indicated by the subscript $_{\text{revlex}}$. Some authors use the adjectives "graded" or "homogeneous" as well, and one may see the subscript $_{\text{grevlex}}$ as an indicator, but, as we shall explain below, in using this order one must make it homogeneous, so the adjective is redundant. For reverse lexicographic order, given two monomials, the one of larger degree is always bigger. The issue is how to order the monomials of a given degree. Here ones uses the opposite of lexicographic order for the monomials numbered backward. Specifically, if $\alpha = (a_1, \ldots, a_n)$ and $\beta = (b_1, \ldots, b_n)$, then $x^\alpha >_{\text{revlex}} x^\beta$ means that $\deg(x^\alpha) > \deg(x^\beta)$ (i.e., that $\sum_{j=1}^n a_j > \sum_{j=1}^n b_j$) or that $\deg(\alpha) = \deg(\beta)$ and there exists an integer j with $1 \leq j \leq n$ such that $a_i = b_i$ for $i > j$ while $a_j < b_j$.

There is a double reversal of sorts here, since one is using the *opposite* of what lexicographic order gives when the variables are *numbered backwards*. One still has that $x_1 > x_2 > \cdots > x_n$, and in the two variable case hlex and revlex are the same. In the three variable case one has that $x_1x_3 >_{\text{lex}} x_2^2$ while $x_2^2 >_{\text{revlex}} x_1x_3$. For the latter, the variable with the highest index for which the two monomials have different exponents is x_3, and the first monomial has the smaller exponent. The difference between the two conditions might be paraphrased by saying that if two monomials have the same degree, for hlex the greater involves more of the low index variables while for revlex the greater involves

fewer of the high index variables. This statement is quite misleading, however, since it is only the first spot (for hlex) and the last spot (for revlex) where the monomials have differerent exponents that governs which monomial is greater. E.g., with 1000 variables,

$$x_1 x_{999}^{10000000} > x_2^{10000000} x_{1000}$$

for both hlex and revlex.

Note that if we simply reverse the order of the variables and take the opposite of lexicographic order (without putting on the condition that monomials of higher degree are always larger), we do not get a monomial order, even if there is only one variable. We always have

$$1 > x_i > x_i^2 > \cdots,$$

if we reverse lexicographic order, even if we think of the variables numbered backwards, and this is not a monomial order. This is what makes it unnecessary to specify homogeneous or graded when discussing reverse lexicographic order.

We extend lex, hlex, and revlex to free modules by our standard rule. Thus, if $b_1, \ldots, b_s$ is the ordered free basis for F, for μ, $\mu' \in \mathcal{M}$, $\mu b_i > \mu' b_j$ means that $\mu > \mu'$ or that $\mu = \mu'$ and $i < j$, no matter which of the three we are working with.

The ordered set is $\mathbb{N}$ for revlex as well as for hlex.

Experience has shown that revlex tends to shorten calculation times for certain applications. It is of some interest that reverse lexicographic order was first considered by F. S. Macaulay in the early 1900s, long before the computer age.

Recall that a totally ordered set is *well-ordered* if, equivalently, either

(1) Every nonempty subset has a least element.

(2) Every non-increasing infinite sequence of elements is eventually constant.

If (1) fails and we have a nonempty subset with no least element, we can recursively construct an infinite strictly decreasing sequence within the subset: choose any element to be the first element of the subsequence. If we have chosen $\mu_1 > \cdots > \mu_n$ strictly decreasing within the subset, we can choose μ_{n+1} with $\mu_n > \mu_{n+1}$ because otherwise μ_n would be the least element in the subset. On the other hand, if (2) fails, by omitting repeated terms we get an infinite strictly decreasing sequence, and the set of elements in it is a subset with no least element. Note that condition (2) is often referred to as DCC or *Artinian*, especially in reference to partially ordered sets.

The following is a critical property of monomial orders.

Theorem. *Let R be a polynomial ring over K and F be a finitely generated free R-module with ordered basis. Then every monomial ordering on R or F is a well-ordering of the monomials.*

Proof. It suffices to consider the case of F. Let S be any nonempty subset of the monomials in F. Give $\nu \in S$, there are only finitely many monomials ν_1 in F such that ν_1 divides

ν, i.e., such that $\nu = \mu\nu_1$ for some monomial $\mu \in \mathcal{M}$. Among these, at least one must be a minimal element in the partial ordering by divisibility. Thus, every element of S is a multiple of a minimal element of S. The set S_0 of minimal elements of S consists of mutually incomparable monomials: none of them divides any of the others. By the Proposition on the top of p. 3 of the Lecture Notes of January 8, this set is finite. Some element of the finite set S_0 is minimum in the monomial order, since the monomial order is a total order. This element is the least element of S for the monomial order, for given any $\nu \in S$, we can write $\nu = \mu\nu_1$ with ν_1 minimal in S with respect to divisibility, and then $\mu\nu_1 \geq \nu_1 \geq \nu_0$ in the monomial order. $\square$

Initial terms and the division algorithm

In this section, let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over a field K, let F be a finitely generated free R-module with ordered basis, and assume that we have a fixed monomial order $>$ on F. Of course, it may well be that $F = R$.

We are going to make several definitions, such as "initial term" and "initial module." Each of these definitions is relative to a fixed monomial order.

First note that the total ordering of monomials also gives an ordering of terms in a weak sense. Given two terms $c\nu$, $c'\nu'$, where c, $c' \in K - \{0\}$ are nonzero scalars and ν, ν' are monomials in F, we write $c\nu < c'\nu'$ to mean that $\nu < \nu'$ and we write $c\nu \leq c'\nu'$ to mean $\nu \leq \nu'$. There relations are transitive. Given any two terms, they will be comparable. However, if $c\nu \leq c'\nu'$ and $c'\nu' \leq c\nu$, the conclusion that we can draw is that $\nu = \nu'$, and *not* that the terms are equal.

This terminology will be very convenient, especially in discussing the terms occurring in a given element $f \in F - \{0\}$. By definition, these terms involve mutually distinct monomials, and so the relation we have introudced on terms restricts to give a linear ordering of the terms of the element f. In particular, $f \neq 0$ has a unique greatest term under $>$, which is called the *initial term* of f and denoted $\mathrm{in}_>(f)$. However, if it is clear from context which monomial order is being used, we may simply write $\mathrm{in}(f)$ for the initial term of f.

When using lexicographic, homogeneous lexicographic, or reverse lexicographic order, the respective notations $\mathrm{in}_{\mathrm{lex}}(f)$, $\mathrm{in}_{\mathrm{hlex}}(f)$, or $\mathrm{in}_{\mathrm{revlex}}(f)$, are used.

Let $M \subseteq F$ be an arbitrary submodule. The submodule of F spanned by the initial terms of all elements of M is a monomial submodule: instead of using $c\nu$ as a generator, where $c \in K - \{0\}$ and ν is a monomial, we can use ν itself. This submodule of F is denoted $\mathrm{in}_>(M)$ or $\mathrm{in}(M)$ and is called the *initial module* of M. It is typically not contained in M (unless M itself is a monomial module). If $F = R$ and I is an ideal, $\mathrm{in}(I)$ is called the *initial ideal* of I. Just as in the case of individual elements, we may indicate that the monomial order used is lexicographic, homogeneous lexicographic, or reverse lexicographic order with the respective notations $\mathrm{in}_{\mathrm{lex}}(M)$, $\mathrm{in}_{\mathrm{hlex}}(M)$, or $\mathrm{in}_{\mathrm{revlex}}(M)$.

With these notations in place, we want to discuss an analogue of the division algorithm for polynomial rings in one variable over a field. However, in our case, instead of dividing

by one polynomial to get a quotient and remainder, we may be "dividing" by several. Furthermore, instead of working with polynomials, we may be working with elements of F. However, for heuristic reasons, the reader may want to think at first only about the case where $F = R$.

Let $f \in F$ and $g_1, \ldots, g_r \in F$, where the g_i are assumed to be nonzero. By a *standard expression* for f in terms of the g_i we mean an expression of the form

$$f = \sum_{i=1}^{r} q_i g_i + h$$

with every $q_i \in R$ and $h \in F$ (technically, one should work with the $(r+1)$-tuple $(q_1, \ldots, q_r, h)$) such that the following two conditions are satisfied:

(1) No term of h is divisible by any of the terms $\mathrm{in}(g_i)$.

(2) For every i such that $q_i g_i \neq 0$, $\mathrm{in}(q_i g_i) \leq \mathrm{in}(f)$.

The element h in a standard expression as above is called a *remainder* for f with respect to $g_1, \ldots, g_r$. (But, again, all of these definitions depend on fixing a monomial order.)

Note that g_i may occur with coefficient $q_i = 0$, in which case $q_i g_i = 0$ and has no initial term: condition (2) is phrased so that the possibility $q_i g_i = 0$ is allowed. In fact, if f has no term that is divisible by any $\mathrm{in}(g_i)$, we may take all the $q_i = 0$ and $h = f$, and so obtain a standard expression at once. It may well be that $h = 0$ in a standard expression. Condition (2) is then satisfied vacuously because h has *no* terms.

Also note that (2) is equivalent to the following condition that, *a priori*, looks stronger:

(2°) For every i, every term of $q_i g_i$ is $\leq \mathrm{in}(f)$.

When $q_i g_i = 0$, this condition is satisfied vacuously, and so we do not need to make a separate statement about that case. If not, this condition follows at once from (2), because $\mathrm{in}(q_i g_i)$ is the greatest term in $q_i g_i$.

We shall prove that there is always a standard expression for f in terms of the g_i. In fact, we shall prove that the following procedure always yields such an expression:

Deterministic division algorithm. Let $>$, f, and $g_1, \ldots, g_r$ as above be given. Define a finite sequence of elements f_n with $f_0 = f$, expressions

$$(\#_n) \quad f = \sum_{i=1}^{r} q_{i,n} g_i + f_n$$

and monomials ν_n in F (except that ν_n is not defined for the final value of n) as follows. If $n = 0$ the expression is simply given by taking all $q_{i,0} = 0$. If f_n has no term divisible by any of the $\mathrm{in}(g_i)$ the procedure stops, and we have that $(\#_n)$ is a standard expression for f with remainder $h = f_n$. Otherwise, once f_n and the corresponding expression $(\#_n)$ are

known, let $c_n\nu_n$ be the largest term of f_n that is a multiple of one or more of the elements $\mathrm{in}(g_i)$. (The procedure that we are describing will eventually terminate no matter which of the g_i with $\mathrm{in}(g_i)$ dividing ν we choose, but we want to make it deterministic.) Let i_n be the least integer such that $\mathrm{in}(g_{i_n})$ divides $c_n\nu_n$, and choose $c'_n\mu_n$ such that $c\nu_n = c'\mu_n\mathrm{in}(g_{i_n})$. Finally, we let

$$f_{n+1} = f_n - c'_n\mu_n g_{i_n},$$

and then we may take

$$q_{j,n+1} = q_{j,n}$$

for $j \neq i_n$ while

$$q_{i_n,n+1} = q_{i_n,n} + c'_n\mu_n.$$

A straightforward induction then shows the following:

(a) For every j, f_{j+1} and f_j have the same terms for monomials strictly larger than ν_j, and f_j has a ν_j term while f_{j+1} does not. Hence, if $j \geq k$, the terms of f_j and f_k agree for monomials strictly larger than ν_j. Moreover, for every j, the terms of f_j strictly larger than ν_j are not divisible by any of the $\mathrm{in}(g_i)$ (or they would have been subtracted off at an earlier stage).

(b) The sequence

$$\nu_0 > \nu_1 > \nu_2 > \cdots$$

is strictly decreasing. Hence, the procedure *must* stop, because the set of monomials is well-ordered by the Theorem at the top of p. 3.

(c) Every expression $(\#_n)$ satisfies the equivalent conditions (2) and (2°). If this is true for $(\#_n)$, it will continue to be true for $(\#_{n+1})$, because the initial term of

$$c'_n\mu_n g_{i_n}$$

is

$$c'_n\mu_n\mathrm{in}(g_{i_n}) = c_n\nu_n$$

by construction, and $\nu_n \leq \nu_0 \leq \mathrm{in}(f)$.

We have proved:

Theorem. *Given $f, g_1, \ldots, g_r \in F$, the deterministic division algorithm presented above produces a standard expression for f in terms of the $g_1, \ldots, g_r$. Therefore, a standard expression for f in terms of the $g_1, \ldots, g_r$ always exists.* □

In case $F = R = K[x]$, with $r = 1$, so that we are dividing f by $g_1 = g$ in $K[x]$, the standard expression we get must be $f = qg + h$, where $\deg(h) < \deg(g)$ or $h = 0$. Here, if $\deg(g) = d$, $\mathrm{in}(g) = cx^d$ for some $c \neq 0$ in K, and the condition that h has no term divisible by $\mathrm{in}(g)$ is equivalent to the condition that $\deg(h) < \deg(g)$ or $h = 0$. The

individual steps in the algorithm are exactly the steps in the usual division algorithm for polynomials in one variable.

In the general case, we do not have the uniqueness statements that hold for the case of division of a polynomial in one variable by another. Of course, the determinstic algorithm we gave produces a unique result, but it is not the only standard expression. There are important cases where the remainder is unique: we return to this point soon.

Example. Let $f = x_1x_2$, $g_1 = x_1 + x_3$, and $g_2 = x_2 + x_3$. Suppose we use hlex. Then

$$f_1 = f - x_2(x_1 + x_3) = -x_2x_3$$

and

$$f_2 = -x_2x_3 + x_3(x_2 + x_3) = x_3^2,$$

which is the remainder. The standard expression we get is

$$x_1x_2 = x_2(x_1 + x_3) - x_3(x_2 + x_3) + x_3^2,$$

with $q_1 = x_2$ and $q_2 = -x_3$ while $f_2 = h = x_3^2$. However, we also have

$$x_1x_2 = (-x_3)(x_1 + x_3) + x_1(x_2 + x_3) + x_3^2,$$

a different standard expression, although the remainders are the same.

Lecture of January 13

Example. Now consider

$$f = x_1x_2x_3,\ g_1 = x_1x_2 + x_3^2,\ g_2 = x_1x_3 + x_2^2$$

in $F = R = K[x_1, x_2, x_3]$ with hlex as the monomial order. On the one hand,

$$f = x_3g_1 + 0 \cdot g_2 - x_3^3$$

is a standard expression with remainder $-x_3^3$, while

$$f = 0 \cdot g_1 + x_2g_2 - x_2^3$$

is a standard expression with remainder $-x_2^3$. Therefore, even the remainder is not unique in general, although it is in important cases that we shall soon discuss.

Gröbner bases

Before proceeding further, we want to comment on the use of the word "basis." By a *basis* for a module we simply mean a set of generators for the module. There is no implication that these generators are linearly independent. If we are working with a free module, the term *free basis* will mean basis of linearly independent elements. In the phrase "free module with ordered basis" the basis is understood to be a free basis.

Over a field K, every module is free. We shall use the terms "vector space basis" and "K-vector space basis" for a set of linearly independent generators in the field case.

Throughout this section $R = K[x_1, \ldots, x_n]$ is a polynomial ring over a field K, $\mathcal{M}$ denotes the set of monomials in R, F is a finitely generated free R-module with ordered basis, and $>$ is a fixed monomial order on F.

The following very easy result is, nonetheless, extraordinarily useful.

Theorem. *Let $M \subseteq F$ be a submodule. If $N \subseteq M$ is a submodule such that $\mathrm{in}(N) = \mathrm{in}(M)$, then $N = M$.*

Proof. We shall give two proofs. First, suppose $N \neq M$. Consider the set $\mathcal{S}$ of monomials of F that occur in the initial term of an element of $M - N$. If this set is non-empty, it has a least element with respect to $>$, since monomial orders are well-orderings. Suppose that $f \in M - N$ has initial term $c\nu$ where ν is the least element of $\mathcal{S}$. Then $\nu \in \mathrm{in}(M) = \mathrm{in}(N)$ occurs as the initial term of some element $g \in N$, and then $f - cg$ contains only terms strictly smaller than ν. But this element is still in $M - N$, and its initial term must be smaller than ν, contradicting the minimality of ν. $\square$

Here is an alternative argument. We know that $\mathrm{in}(M) = \mathrm{in}(N)$ is finitely generated, since it is a monomial module. We may therefore choose finitely many elements $g_1, \ldots, g_r \in N$ whose initial terms generated $\mathrm{in}(M)$. Let $f \in M$ be given. By the division algorithm, there is a regular expression

$$f = \sum_{j=1}^{r} q_j g_j + h$$

for f in terms of the g_i. Then $h \in M$, but no term of h is divisible by any $\mathrm{in}(g_j)$. This implies that $h = 0$, for otherwise its initial term $\mathrm{in}(h) \in \mathrm{in}(M)$ and so must be divisible by some $\mathrm{in}(g_j)$. But this shows that $f \in N$. $\square$

We are immediately led to make the following definition. Let $M \subseteq F$ be any submodule. Then $g_1, \ldots, g_r$ is called a *Gröbner basis* for M if the elements $\mathrm{in}(g_1), \ldots, \mathrm{in}(g_r)$ are a basis for $\mathrm{in}(M)$. We know that since $\mathrm{in}(M)$ is monomial, it is finitely generated, and so a Gröbner basis for M always exists.

Theorem. *Every submodule M of F has a Gröbner basis.* $\square$

Theorem. *A Gröbner basis for $M \subseteq F$ is a basis for M.*

Proof. This is immediate from the first Theorem on p. 1. $\square$

Corollary (Hilbert basis theorem). *Every submodule of F is finitely generated. In particular, every ideal of $R = K[x_1, \ldots, x_n]$ is finitely generated.*

Proof. The submodule or ideal has a (finite) Gröbner basis, which is then a basis. $\square$

We also have:

Theorem. *Let $M \subseteq F$ be a submodule. The monomials of F* not *in $\mathrm{in}(M)$ give a K-vector space basis for F/M.*

Proof. We first show that the set of monomials $\mathcal{Q}$ in F and not in $\mathrm{in}(M)$ are linearly independent over K. If we have a linear relation on these monomials, we find that a nonzero linear combination of monomials in $\mathcal{Q}$ is an element f of M. But then the initial term of $f \in M$ involves a monomial not in $\mathrm{in}(M)$, a contradiction.

Now let $f \in F$ be given, and let $g_1, \ldots, g_r$ be a Gröbner basis for M. By the division algorithm, we can write $f = \sum_{j=1}^{r} q_j g_j + h$, where h is in the K-span of $\mathcal{Q}$. But then $f \equiv h \bmod M$. $\square$

Corollary. *Let $M \subseteq F$ and let $g_1, \ldots, g_r$ be a Gröbner basis for M. Then for all $f \in F$, the remainder h in any standard expression*

$$f = \sum_{j=1}^{r} q_j g_j + h$$

is unique, i.e., h is the same no matter what standard expression is chosen.

In particular, $f \in F$ is an element of M if and only if the remainder in any standard expression for f in terms of the Gröbner basis $g_1, \ldots, g_r$ is 0.

Proof. The remainder h is a K-linear combination of monomials in $\mathcal{Q}$, the set of monomials of F not in $\mathrm{in}(M)$. Any two remainders represent the same element of F/M, and so the result follows at once from the preceding Theorem.

The final statement is then obvious. $\square$

Notice that if we can find a Gröbner basis $g_1, \ldots, g_r$ for $M \subseteq F$ (or for $I \subseteq R$), the result above gives an effective test for whether an element $f \in F$ (respectively, R) is in M (respectively, I): one simply uses the division algorithm to find a remainder for f in terms of $g_1, \ldots, g_r$, and $f \in M$ (respectively, I) if and only if the remainder is 0.

However, at this point we do not have an effective method for finding a Gröbner basis for M given a set of generators of M. We shall develop such a method, called the *Buchberger algorithm*, at which point we have a solution for the problem of giving an effective test for membership in M or I when we know specific generators for M or I.

Before discussing the Buchberger algorithm, we want to discuss restrictions on a Gröbner basis for M (or I) that make it unique.

A Gröbner basis $g_1, \ldots, g_r$ for $M \subseteq F$ is called *minimal* if the monomials occurring in $\mathrm{in}(g_1), \ldots, \mathrm{in}(g_r)$ are the minimal monomials in $\mathrm{in}(M)$. Evidently, every Gröbner basis for M has a subset that is a minimal Gröbner basis. Notice that every minimal Gröbner basis for M has the same cardinality as the set of minimal monomials $\mathrm{in}(M)$. We shall say that an ordered Gröbner basis $g_1, \ldots, g_r$ for $M \subseteq F$ is *reduced* if it satisfies the following four conditions:

(1) $g_1, \ldots, g_r$ is minimal.

(2) $\mathrm{in}(g_1) > \mathrm{in}(g_2) > \cdots > \mathrm{in}(g_r)$.

(3) Every $\mathrm{in}(g_i)$ is a monomial, i.e., the coefficient in every initial term is 1.

(4) For all $i \neq j$, $\mathrm{in}(g_i)$ does not divide any term in g_j.

There is variation in the literature in the use of the term "reduced Gröbner basis," but conditions (1) and (4) are always assumed. We have chosen the usage that makes a reduced Gröbner basis for M unique, as we shall see below.

As already noted, any Gröbner basis has a subset that is minimal, and the elements can then be ordered uniquely so that the sequence of initial terms is strictly decreasing. Obviously, one can multiply each term by the reciprocal of the coefficient of the initial term, and therefore conditions (1), (2), and (3) are readily achieved. Note that it is obvous that the sequence of initial terms is then the same as the sequence of minimal monomial generators of $\mathrm{in}(M)$, arranged in strictly decreasing order. We can guarantee condition (4) as follows. Replace g_1 by its remainder in a standard expression with respect to division by $g_2, \ldots, g_r$. Then replace g_2 by its remainder in a standard exrpession with respect to division by $g_3, \ldots, g_r$. Continue in this way for $r-1$ steps. At the ith step, replace g_i by its remainder in a standard expression with respect to division by $g_{i+1}, \ldots, g_r$.

It is easy to see that the result satisfies all of the conditions (1) — (4). The first three conditions are not disturbed. Given $i < j$, $\mathrm{in}(g_i)$ is bigger than any term in g_j, and so cannot divide g_j, while no term in g_i is divisible by $\mathrm{in}(g_j)$, because g_i is the remainder in a standard expression for division by elements one of which has the same initial term as g_j. Note that while the g_k change, their initial terms do not change.

Since we can use the deterministic division algorithm at each step, we see that we can pass algorithmically from a Gröbner basis to a reduced Gröbner basis. We have now proved the first statement in the Theorem below.

Theorem. *Let $M \subseteq F$ (or $I \subseteq R$) be given. Then M (respectively, I) has a reduced Gröbner basis, and it is unique.*

Proof. It remains only to prove uniqueness and, as usual, it suffices to consider the case of modules. We need only show that if $g_1, \ldots, g_r$ and $g'_1, \ldots, g'_r$ are two reduced Gröbner bases for M, then $g_i = g'_i$ for all i. We know *a priori* that $\mathrm{in}(g_i) = \mathrm{in}(g'_i)$ for all I. We use reverse induction on i. Apply the division algorithm to find a standard expression for g'_r in terms of $g_1, \ldots, g_r$. We know that the remainder will be 0. Moreover, at every stage, the initial terms of $g_1, \ldots, g_{r-1}$ are too large to be used. At the very first step in the algorithm, we subtract g_r from g'_r to produce an element of M all of whose terms involve only monomials $< \mathrm{in}(g_r) = \mathrm{in}(g'_r)$. Since this is the least monomial in $\mathrm{in}(M)$, it follows that $g_r - g'_r = 0$. This gives the base step of the induction.

Now assume that $i < r$ and that $g_j = g'_j$ for $j > i$. Perform the division algorithm for g'_i with respect to $g_1, \ldots, g_r$. The terms $g_1, \ldots, g_{i-1}$ are all too large ever to be used. At the first step, one gets $g_i - g'_i$: the initial terms cancel, all remaining terms are strictly smaller than $\mathrm{in}(g_i) = \mathrm{in}(g'_i)$, and none of them is divisible by $\mathrm{in}(g_j)$ for $j > i$, since this is true for all terms but the greatest in both g_i and g'_i. Since the remainder must be 0, we must have that $g_i - g'_i = 0$, and so $g_i = g'_i$, as required. $\square$

Revisited example. Consider again the example on p. 1, in which $g_1 = x_1x_2 + x_3^2$ and $g_2 = x_1x_3 + x_2^2$. The elements g_1 and g_2 are certainly minimal generators for an ideal I of $K[x_1, x_2, x_3]$. They are not, however, a Gröbner basis using hlex. The initial terms of these two elements are x_1x_2 and x_1x_3. Note that $x_3g_1 - x_2g_2 = x_3^3 - x_2^3$ has initial term $-x_2^3$, which shows that $\mathrm{in}(g_1)$ and $\mathrm{in}(g_2)$ are not the only minimal elements of $\mathrm{in}(I)$. In fact, we know this *a priori*, since remainders of division with respect to g_1, g_2 are not unique.

Lecture of January 15

The notion of Gröbner basis is non-trivial and of some interest even when there are no indeterminates, i.e., when $R = K$ is a field, and $F = K^s$.

Consider an $r \times s$ matrix $A = \big(a_{i,j}\big)$ over a field K. The leftmost nonzero entry of a nonzero row of A is called the *leading* or *initial* entry. Recall that A is said to be in *reduced row echelon form* if it satisfies the following conditions:

(1) The nonzero rows precede the rows that are 0.

(2) The leading entry of every nonzero row is 1.

(3) If there are ρ nonzero rows, and if the leading entry of the i th row in the j_i th column, then $j_1 < j_2 < \cdots < j_\rho$.

(4) If the leading entry of the i th row occurs in the j_i th column, then all other entries in the j_i th column are 0.

The key result from elementary linear algebra about reduced row echelon form is that every $r \times s$ matrix over K has the same row space as a *unique* matrix in reduced row echelon

form. Moreover, the given matrix can be put into reduced row echelon form by a sequence of elementary row operations (i.e., multiplying a row by a nonzero scalar, permuting the rows, and adding a multiple of one row to another). This gives a canonical basis for the row space of the original matrix (but this canonical basis does depend on having made a choice of basis for K^s).

Suppose that A is in reduced row echelon form and call the nonzero rows $f_1, \ldots, f_\rho$. The initial term of f_i is e_{j_i}. Condition (4) guarantees that the initial term of f_i does not divide any term in any other f_j. If we have any nonzero element $c_1 f_1 + \cdots + c_\rho f_\rho$ of the row space, its initial term will be $c_i e_{j_i}$ for the smallest value of i such that $c_i \neq 0$. Consequently:

Proposition. *Let the monomial order for K^s be such that $e_1 > e_2 > \cdots > e_s$. An $r \times s$ matrix over the field K is in reduced row echelon form if and only if its nonzero rows precede its zero rows and its nonzero rows form a reduced Gröbner basis for its row space.* $\square$

Relations on monomials and terms

Let $M \subseteq F = R^s$ be any monomial submodule. Since M is generated by monomials $\mu_i e_j$, it follows that

$$M = I_1 e_1 \oplus \cdots \oplus I_s e_s$$

where every I_j is a monomial ideal of R. Understanding the relations on generators for M is therefore equivalent to understanding the relations on generators for several separate monomial ideals of R.

We therefore focus first on understanding generators for the module of relations on a sequence $\mu_1, \ldots, \mu_r$ of monomials in the polynomial ring $R = K[x_1, \ldots, x_n]$. For each pair of monomials μ_i and μ_j with $i \neq j$, we get one "obvious" minimal relation: it comes from the trivial relation that corresponds to the equation

$$\mu_j \mu_i - \mu_i \mu_j = 0$$

by dividing both coefficients μ_j and $-\mu_i$ by $\Delta_{ij} = \mathrm{GCD}(\mu_i, \mu_j)$. (Trivial relatons are also called *Koszul* relations, and the relation obtained by dividing by Δ is sometimes called a *divided Koszul relation.*) Thus, if $I = (\mu_1, \ldots, \mu_r)R$ and we map $R^r \twoheadrightarrow R$ by the map that sends $e_i \mapsto \mu_i$, the kernel will contain the elements

$$\theta_{ij} = \frac{\mu_j}{\Delta_{ij}} e_i - \frac{\mu_j}{\Delta_{ij}} e_j \in R^r.$$

In fact, all relations on μ_i and μ_j are multiples of θ_{ij}. This is a consequence of the following:

Lemma. *Let μ_1 and μ_2 be any two nonzero elements of a UFD R, and let $\Delta = GCD(\mu_1, \mu_2)$, so that $\mu_1 = f_1\Delta$ and and $\mu_2 = f_2\Delta$, with $GCD(f_1, f_2) = 1$. Then $(f_2, -f_1)$ generates*

the module of relations on μ_1 and μ_2. In other words, if $g_1\mu_1 + g_2\mu_2 = 0$, then (g_1, g_2) is a multiple of $(f_2, -f_1)$.

Proof. Since $g_1\mu_1 + g_2\mu_2 = 0$, we have that $g_1\Delta f_1 + g_2\Delta f_2 = 0$, and so $g_1f_1 + g_2f_2 = 0$. Since f_2 divides g_1f_1 while $\mathrm{GCD}(f_1, f_2) = 1$, we have that f_2 divides g_1, say $g_1 = qf_2$. Then $qf_2f_1 + g_2f_2 = 0$, and so $g_2 = -qf_1$, i.e., $(g_1, g_2) = q(f_2, -f_1)$, as required. □

Example. If $\mu_1 = x_1^2x_2^3x_3^5x_4$ and $\mu_2 = x_1^3x_3^2x_4^4$, then the trivial or Koszul relation on these two monomials is given by $(x_1^3x_3^2x_4^4, -x_1^2x_2^3x_3^5x_4)$, while $\theta_{1,2}$ is the result of factoring out the GCD, which is $x_1^2x_3^2x_4$, i.e., $\theta_{12} = (x_1x_4^3, -x_2^3x_3^3)$.

We next want to show that the θ_{ij} generate all relations on the μ_j. We first discuss the notion of an $\mathbb{N}^n$-grading, and more general gradings. Let H be a commutative semigroup (which means that the operation is associative) with identity 0, and suppose that the binary operation for H is written additively. An H-graded ring is a ring R with a direct sum decomposition $R = \bigoplus_{h\in H} R_h$ as an abelian group such that $1 \in R_0$ and $R_hR_{h'} \subseteq R_{h+h'}$ for all $h, h' \in H$. An H-graded module M over an H-graded ring R is then an R-module M with a direct sum decomposition $M = \bigoplus_{h\in H} M_h$ as an abelian group such that that $R_hM_{h'} \subseteq M_{h+h'}$ for all $h, h' \in H$. Note that this implies that every M_h is an R_0-module. An element of R or M is called *homogeneous* or a *form* if it is in one of the R_h or M_h.

If f is in an H-graded ring or module, the direct sum decomposition provides a decomposition of f into homogeneous components, one for every element of H, just as in the $\mathbb{N}$-graded case.

An ideal (respectively, a submodule) of an H-graded ring R (respectively, an H-graded module M) is called a *homogeneous* or *graded* ideal (respectively, submodule) if the following two equivalent conditions hold:

(1) It is generated by homogeneous elements.

(2) It contains all of the homogeneous components of all of its elements.

Suppose that we take $H = \mathbb{N}^n$. Then it is easy to see that the polynomial ring $R = K[x_1, \ldots, x_n]$ is $\mathbb{N}^n$-graded, where, if $\alpha \in \mathbb{N}^n$, $R_\alpha = Kx^\alpha$. This is simply a consequence of the fact that $x^\alpha x^\beta = x^{\alpha+\beta}$. In this case, the homogeneous ideals (with respect to the $\mathbb{N}^n$-grading) are precisely the monomial ideals. We can now prove:

Proposition. *The relations θ_{ij} generate all the relations on the monomials $\mu_1, \ldots, \mu_r$.*

Proof. Suppose that we have a relation corresponding to

$$(*) \quad \sum_{k=1}^{r} f_j\mu_j = 0.$$

(Officially, the relation is $\sum_{j=1}^r f_je_j \in R^r$.) Only finitely many degrees occur when we expanded out all the products occurring in the summation: call these degrees $\alpha_1, \ldots, \alpha_t$.

Fix one of these degrees $\alpha_i \in \mathbb{N}^n$. For every i, the the sum of the terms of degree α_i occurring in $(*)$ is 0. If the degree of $\mu_j = \beta_j$, this sum can be represented as

$$(*_i) \quad \sum_{k=1}^{r} [f_j]_{\alpha_i - \beta_j} \mu_j = 0,$$

where $[f_j]_\gamma$ denotes the degree γ component of f_j. The original relation is the sum of the relations corresponding to the equations $(*_i)$. Therefore, it suffices to show that each of the relations corresponding to one of the equations $(*_i)$ is an R-linear combination of the θ_{ij}. Thus, we need only consider relations in which the degree of every product is x^α for some fixed α. These are *homogeneous* relations.

We may drop the terms with coefficient 0. After renumbering the monomials, we may assume without loss of generality that for every j, f_j is a nonzero term $c_j \mu_j'$ where $\mu_j' \mu_j = x^\alpha$, and α is independent of j. The fact that

$$(**) \quad \sum_{j=1}^{r} (c_j \mu_j') \mu_j = 0$$

is then simply the assertion that $\sum_{j=1}^{r} c_j = 0$, and so $c_r = -\sum_{j=1}^{r-1} c_j$.

The given relation is therefore the sum of $r-1$ relations corresponding to equations of the form

$$(***) \quad c_j \mu_j' \mu_j - c_j \mu_r' \mu_r = 0$$

where $1 \leq i \leq r-1$. Since this equation corresponds to a relation on just two monomials, namely, μ_j and μ_r, by the preceding Lemma the corresponding relation must be a multiple of θ_{jr}. $\square$

Example. The θ_{ij} are not necessarily a minimal set of generators for the relations on the μ_j. For example, suppose that $\mu_1 = x_2 x_3$, $\mu_2 = x_1 x_3$ and $\mu_3 = x_1 x_2$. Then we have that $\theta_{12} = (x_1, x_2, 0)$, $\theta_{13} = (x_1, 0, -x_3)$, and $\theta_{23} = (0, x_2, -x_3)$. Since $\theta_{13} = \theta_{12} + \theta_{23}$, we only need two of these three relations in a minimal basis.

We want to extend this type of relation θ_{ij} to terms $\gamma_i = c_i \mu_i e_m$ and $\gamma_j = c_j \mu_j e_m$ in a free module F with ordered basis provided that γ_i and γ_j involve the *same* ordered basis element e_m. Here, c_i and c_j are nonzero scalars in K, while $\mathfrak{m}_i$ and μ_j are monomials in R. In this case we let $\Delta_{ij} = \mathrm{GCD}(\gamma_i, \gamma_j)$, which we define to be $\mathrm{GCD}(\mu_i, \mu_j) e_m$. We also define γ_i / Δ_{ij} to be $c_i \mu_i / \mathrm{GCD}(\mu_i, \mu_j)$, which is a term in R. We still have

$$(\gamma_i / \Delta_{ij}) \Delta_{ij} = \gamma_i.$$

We can now define

$$\theta_{ij} = \frac{\gamma_j}{\Delta_{ij}} e_i - \frac{\gamma_j}{\Delta_{ij}} e_j \in R^r,$$

just as in the case of monomials in R. We have at once:

Lemma. *if $\gamma_1, \ldots, \gamma_r$ are terms in F, the module of relations on the elements $\gamma_1, \ldots, \gamma_r$ is generated by the relations θ_{ij} for those choices of i, j such that γ_i and γ_j involve the same element of the ordered basis for F.* $\square$

We shall later discuss a similar result that gives an entire finite free resolution for monomial ideals and submodules. This resolution was discovered by Diana Taylor in the 1960s. However, it is *not minimal*. In fact, the minimal resolution of a monomial ideal may depend on the characteristic of the field K.

The Buchberger criterion and algorithm

Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over a field K, let F be a finitely generated free module with ordered basis, and let $M \subseteq F$ be a submodule. Let $g_1, \ldots, g_r \in M$ be elements that generate M. The following theorem gives necessary and sufficient conditions for the g_j to be a Gröbner basis for M. Once this result is known, one immediately gets an algorithm for enlarging a given set of generators of M to a Gröbner basis for M.

The idea underlying the criterion is to try to produce new elements of $\text{in}(M)$ from the given g_j in an obvious way: first take an efficient monomial linear combination of g_i and g_j that gets their initial terms to cancel. Divide the result with respect to the $g_1, \ldots, g_r$. If the remainder is nonzero, its initial term cannot be in the R-span of the $\text{in}(g_j)$, and we have taken a further step towards finding a Gröbner basis. If all of the remainders are 0, we hope that we already have a Gröbner basis. This is true.

Here is a precise formulation. Let $g_1, \ldots, g_r$ be generators for $M \subseteq F$ and let $\nu_1, \ldots, \nu_r$ be their respective initial terms. For every pair of indices $i \neq j$ such that ν_i and ν_j involve the same element of the ordered basis for F, let

$$G_{ij} = \frac{\nu_j}{\Delta_{ij}} g_i - \frac{\nu_j}{\Delta_{ij}} g_j,$$

where $\Delta_{ij} = \text{GCD}(\nu_i, \nu_j)$. Let h_{ij} be a remainder for division of G_{ij} with respect to $g_1, \ldots, g_r$ (the remainder in any standard expression can be used: this need not be the result of using the deterministic division algorithm).

Theorem (Buchberger Criterion). *With notation as in the paragraph above, $g_1, \ldots, g_r$ is a Gröbner basis for M if and only if all of the $h_{ij} = 0$.*

We postpone the proof of the sufficiency of the condition momentarily. When we do give the proof, we shall establish a somewhat weaker sufficient condition.

The condition given above is clearly necessary: if the $g_1, \ldots, g_r$ form a Gröbner basis for M, then since G_{ij} is clearly in M, our test for membership in M implies that the remainder in any standard expression when we divide G_{ij} with respect to the Gröbner basis $g_1, \ldots, g_r$ is 0.

We note that this gives an effective algorithm for finding a Gröbner basis for M given generators $g_1, \ldots, g_r$. We calculate values for the h_{ij}. If one of these is nonzero, its initial term cannot be in the span of $\mathrm{in}(g_1), \ldots, \mathrm{in}(g_r)$. (To make the process choice-free, we can use the least value of i for which $h_{ij} \neq 0$, and, for that i, the least value of j.) We then enlarge the original set of generators by including this element h_{ij}. The R-span of the initial terms has increased. Since F is Noetherian, the process must terminate, i.e., eventually we reach a set of generators for which all of the h_{ij} are 0. The Buchberger criterion now implies that we have a Gröbner basis for M. This method is called the *Buchberger algorithm.*

We do not, however, have an *a priori* estimate for how many steps will be needed to find the Gröbner basis. In worst cases, the number of steps is double exponential. However, in practice, the method is useful in many of the examples that come up.

Lecture of January 20

We give one example of how one starts to calculate a Gröbner basis in a specific instance. Let $g_1 = x_1^2x_2x_4 + x_3^4$ and $g_2 = x_1x_3x_4^2 + x_4^4$ be generators for an ideal I in $R = K[x_1, x_2, x_3, x_4]$ and suppose that we are using hlex as the monomial order. Then $\nu_1 = \mathrm{in}(g_1) = x_1^2x_2x_4$ and $\nu_2 = \mathrm{in}(g_2) = x_1x_3x_4^2$ are elements of $\mathrm{in}(I)$. To test whethre this is a Gröbner basis we calculate G_{12} and h_{12}. Here, $\Delta_{12} = \mathrm{GCD}(\nu_1, \nu_2) = x_1x_4$, and so

$$G_{12} = \frac{x_1x_3x_4^2}{x_1x_4}g_1 - \frac{x_1^2x_2x_4}{x_1x_4}g_2 = x_3x_4(x_1^2x_2x_4 + x_3^4) - x_1x_2(x_1x_3x_4^2 + x_4^4).$$

Note that the multiples of the two initial terms cancel. This simplifies to

$$G_{12} = x_3^5x_4 - x_1x_2x_4^4,$$

and no term is a multiple of ν_1 or ν_2, so that we may take $G_{12} = h_{12}$. We see that $x_1x_2x_4^4 \in \mathrm{in}(I)$, and we now consider whether g_1, g_2, h_{12} might be a Gröbner basis.

We have yet to prove that the Buchberger criterion stated last time gives a sufficient condition for $g_1, \ldots, g_r$ to be a Gröbner basis. In fact, we shall prove a sharper result. Before stating the new version, we want to observe:

Lemma. *Let $g_1, \ldots, g_r$ be nonzero elements of F, with our usual notation conventions. If g_i and g_j are such that* all *of their terms involve the same element e_t of the ordered basis for F (this condition is automatically satisfied if $F = R$), and if the initial terms ν_i of g_i and ν_j of g_j are relatively prime (i.e., their GCD is e_t), then there is a standard expression for G_{ij} under division with respect to $g_1, \ldots, g_r$ such that the remainder $h_{ij} = 0$.*

The proof is left as an exercise: see Problem Set #1, Problem 6.

We now state our sharpened version of the Buchberger criterion. $R = K[x_1, \ldots, x_n]$ is a polynomial ring over a field K, and $g_1, \ldots, g_r$ are nonzero generators of a module

$M \subseteq F$, where F is a finitely generated free R-module with ordered basis. Let $\nu_j = \text{in}(g_j)$ for $1 \leq j \leq r$. Consider any set of pairs of indices $i_\lambda < j_\lambda$ such that

(1) For every λ, ν_{i_λ} and ν_{j_λ} involve the same basis element of F.

(2) The standard relations $\theta_{i_\lambda j_\lambda}$ generate the module of relations on the terms $\nu_1, \ldots, \nu_r$.

For every λ, let

$$G_{i_\lambda j_\lambda} = \frac{\nu_j}{\text{GCD}(\nu_{i_\lambda}, \nu_{j_\lambda})} g_i - \frac{\nu_i}{\text{GCD}(\nu_{i_\lambda}, \nu_{j_\lambda})} g_j.$$

For every λ, let $h_{i_\lambda j_\lambda}$ be the remainder in *any* standard expression for $G_{i_\lambda j_\lambda}$ divided by $g_1, \ldots, g_r$. (One does not have to use the remainder that arises from the deterministic division algorithm.)

Theorem (sharpened Buchberger criterion). *Let notation be as in the preceding paragraph. A necessary and sufficient condition for $g_1, \ldots, g_r$ to be a Gröbner basis for M is that every $h_{i_\lambda j_\lambda} = 0$. If $F = R$, the condition is still sufficient if one only checks those λ such that $\text{in}(g_{i_\lambda})$ and $\text{in}(g_{j_\lambda})$ are not relatively prime. (More generally, one can omit the check for λ whenever g_{i_λ} and g_{j_λ} have all terms involving the same element of the ordered basis for F, and $\text{in}(g_{i_\lambda})$ and $\text{in}(g_{j_\lambda})$ are relatively prime.)*

The original statement used all pairs ν_i, ν_j involving the same element of the ordered basis in defining the h_{ij}. We have cut down the number of pairs needed in two ways. First, we only need to use enough pairs to get a basis for the relations on $\nu_1, \ldots, \nu_r$. It is often the case that one can use far fewer pairs. Second, when $F = R$, one can omit checking whether the remainder is 0 for any pair such that the monomials in the initial terms are relatively prime.

It is obvious that the condition in the sharpened Buchberger criterion is necessary. Before giving the proof of sufficiency, we make the following observation. Given a monomial order on F, for every element e_i in the ordered basis we get a monomial order on the ring, which we denote $>_t$, defined by the rule $\mu_1 > \mu_2$ precisely when $\mu_1 e_t > \mu_2 e_t$. In many cases all of these monomial orders on R are the same, but this need not be true in general. However, if $f \in R - \{0\}$, $g \in F - \{0\}$, and $\text{in}(g)$ involves e_t, then

$$(\dagger) \quad \text{in}(fg) = \text{in}_{>_t}(f)\text{in}(g).$$

To see why this is true, consider what happens when we calculate fg by applying the distributive law and taking all products of a term of f and a term of g. First consider only those terms that involve e_t. The specified term occurs, and it is clear that all other terms occurring that involve e_t are strictly smaller, so that it cannot be cancelled. Thus, it suffices to show that any product of a term μ_1 of f and a term $\mu_2 e_j$ of g with $j \neq t$ is also $\leq \text{in}_{>_t}(f)\text{in}(g)$ — the inequality must then be strict, because $j \neq t$. But $\mu_2 e_j \leq \text{in}(g)$, and so

$$\mu_1 \mu_2 e_j \leq \mu_1 \text{in}(g).$$

Since $\mu_1 \leq_t \mathrm{in}_{>_t}(f)$ by definition of $\mathrm{in}_{>_t}(f)$, we have that

$$\mu_1 \mathrm{in}(g) \leq \mathrm{in}_{>_t}(f)\mathrm{in}(g),$$

as required. $\square$

We are now ready to give the argument for sufficiency.

Proof of sufficiency for the sharpened Buchberger criterion. First, in the case where $F = R$, note that the omission of checking the remainder when the initial terms are relatively prime is justified by the Lemma above: one can always choose a standard expression for which the remainder is 0, and so checking those pairs is unnecessary.

Now suppose that all the $h_{i_\lambda j_\lambda} = 0$. We must prove that $g_1, \ldots, g_r$ is a Gröbner basis. We assume the contrary, and obtain a contradiction.

If the $g_1, \ldots, g_r$ are not a Gröbner basis, we can choose

$$f = \sum_{j=1}^{r} f_j g_j$$

such that $\mathrm{in}(f)$ is not a multiple of any of the ν_j. We fix one such element f for the remainder of the proof. Let ϕ denote the r-tuple $(f_1, \ldots, f_r)$. Consider those terms on the right such that $f_j \neq 0$ and for these the ones such that the monomial ν_ϕ corresponding to $\mathrm{in}(f_j g_j)$ is largest as j varies: there may be several values of j that give rise to the same largest monomial ν_ϕ. There are typically many ways to write f as a linear combination of $g_1, \ldots, g_r$. Choose such a representation in such a way that ν_ϕ is minimum. This is possible because the monomial ordering on F is a well-ordering. We simply write $\nu = \nu_\phi$.

We shall obtain a contradiction by proving that if $\mathrm{in}(f)$ is not a multiple of any ν_j, then we can find a different representation for f as a linear combination of the g_j such that the value of ν_ϕ is strictly smaller.

After renumbering the g_j, we may assume that a nonzero scalar multiple of ν is the initial term of $f_i g_i$ for $1 \leq i \leq k$, and not for $f_j g_j$ with $j > k$. Each of $f_{k+1}g_{k+1}, \ldots, f_r g_r$ only involves terms that are strictly smaller than ν. To complete the argument, it will suffice to show that $f_1 g_1 + \cdots + f_k g_k$ can be rewritten as a linear combination of $g_1, \ldots, g_r$ so that the initial term of every product occurring in the sum is $< \nu$.

Suppose that ν involves e_t. Observe that by the discussion on p. 2 leading to the displayed formula (†), we know that for $1 \leq i \leq k$, $\mathrm{in}(f_i g_i) = \mu_i \nu_i$, where $\mu_i = \mathrm{in}_{>_t}(f_i)$. Here, each $\mu_i \nu_i$ is a scalar multiple of ν. We consider two cases.

First case. Here, we assume that $\sum_{i=1}^{k} \mu_i \nu_i \neq 0$. In this case, the value of the sum is a nonzero scalar multiple of ν, and so the initial term of f is evidently a nonzero scalar multiple of ν as well. This is an immediate contradiction, because, up to multiplication by a nonzero scalar, ν is the same as $\mu_i \nu_i$ for $1 \leq i \leq k$, and so is a multiple of ν_i for $1 \leq i \leq k$. But this contradicts the assumption that ν is not a multiple of any ν_j.

Second case. We assume that $\sum_{i=1}^{k} \mu_i \nu_i = 0$. We may write $f_i = \mu_i + \widetilde{f_i}$ for $1 \leq i \leq k$, and then we have

$$f = \sum_{i=1}^{k} \mu_i g_i + \sum_{i=1}^{k} \widetilde{f}_i g_i + \sum_{j=k+1}^{r} f_j g_j.$$

All terms occurring in the second and third sums are $< \nu$. Therefore, it will suffice to show that the first term, $\sum_{i=1}^{k} \mu_i g_i$, can be rewritten as a linear combination $\sum_{j=1}^{r} q_j g_j$ in such a way that every $\mathrm{in}(q_j g_j) < \nu$: after combining terms, we will have a new representation for f with a smaller ν_ϕ.

Since

$$\sum_{i=1}^{k} \mu_i \nu_i = 0,$$

we have that

$$\sum_{i=1}^{k} \mu_i e_i$$

is a relation on $\nu_1, \ldots, \nu_r$. This relation has the same degree as ν, in the sense that each of the products has the same degree in $\mathbb{N}^n$ as ν. It follows that it can be written as a linear combination of the $\theta_{i_\lambda j_\lambda}$. Moreover, we may think of $\theta_{i_\lambda j_\lambda}$ as having the same degree as $LCM(\nu_{i_\lambda}, \nu_{j_\lambda})$. We therefore have an equation

$$(\#) \quad \sum_{i=1}^{k} \mu_i e_i = \sum_{\lambda} \zeta_\lambda \theta_{i_\lambda j_\lambda}$$

where the sum is extended over the indices λ that are needed (we do not include summands with coefficient 0), and each ζ_λ is a term such that

$$\deg(\zeta_\lambda) + \deg(\theta_{i_\lambda j_\lambda}) = \deg(\nu).$$

We now apply to (#) the map $R^r \to R$ sending $e_1, \ldots, e_r$ to $g_1, \ldots, g_r$ respectively. This yields

$$(*) \quad \sum_{i=1}^{k} \mu_i g_i = \sum_{\lambda} \zeta_\lambda G_{i_\lambda j_\lambda}.$$

Here,

$$G_{i_\lambda j_\lambda} = \frac{\nu_{j_\lambda}}{\mathrm{GCD}(\nu_{i_\lambda}, \nu_{j_\lambda})} g_i - \frac{\nu_{i_\lambda}}{\mathrm{GCD}(\nu_{i_\lambda}, \nu_{j_\lambda})} g_j.$$

Here, the initial term of each summand on the right is the same

$$\nu_{j_\lambda} \frac{\nu_{i_\lambda}}{\mathrm{GCD}(\nu_{i_\lambda}, \nu_{j_\lambda}))} = \nu_{i_\lambda} \frac{\nu_{j_\lambda}}{\mathrm{GCD}(\nu_{i_\lambda}, \nu_{j_\lambda}))}$$

which is the same up to a nonzero scalar multiple as $\mathrm{LCM}(\nu_{i_\lambda}, \nu_{j_\lambda})$. Since the initial terms cancel, we have that

$$\mathrm{in}(G_{i_\lambda j_\lambda}) < \mathrm{LCM}(\nu_{i_\lambda}, \nu_{j_\lambda}),$$

and it follows that when we multiply by ζ_λ we have that

$$\text{in}(\zeta_\lambda G_{i_\lambda j_\lambda}) < \nu.$$

By hypothesis, every $G_{i_\lambda j_\lambda}$ has a standard expression of the form $\sum_{j=1}^r q_j^\lambda g_j$ in which the initial term of each product in the sum is $\leq \text{in}(G_{i_\lambda j_\lambda})$. We now substitute into $(*)$ above to obtain

$$\sum_{i=1}^k \mu_i g_i = \sum_\lambda \sum_{j=1}^r \zeta_\lambda q_j^\lambda g_j$$

and for all λ and j we have

$$\text{in}(\zeta_\lambda q_j^\lambda g_j) \leq \zeta_\lambda \text{in}(G_{i_\lambda j_\lambda}) = \text{in}(\zeta_\lambda G_{i_\lambda j_\lambda}) < \nu,$$

exactly as required. $\square$

Review of complexes and homology

By a *complex* over a ring A we mean a sequence of A-modules and A-module maps

$$(*) \quad \cdots \to G_{t+1} \xrightarrow{d_{t+1}} G_t \xrightarrow{d_t} G_{t-1} \to \cdots$$

indexed by $\mathbb{Z}$ such that for all t, $d_t \circ d_{t+1} = 0$. However, we shall frequently consider complexes such that $G_t = 0$ for all $t < 0$, and when we talk about the complex

$$\cdots \to G_{t+1} \to G_t \to G_{t-1} \to \cdots \to G_1 \to G_0 \to 0$$

we mean to imply that all negative terms vanish. We refer to a complex in which $G_i = 0$ for $i < 0$ as a *left* complex. Likewise, when we talk about the complex

$$0 \to G_k \to G_{k-1} \to \cdots \to G_t \to \cdots$$

we mean to imply that $G_i = 0$ for $i > k$. We frequently write $(G_\bullet, d_\bullet)$ or simpy $G_\bullet$ to described a complex as in $(*)$. Note that the condition that $d_t \circ d_{t+1} = 0$ is equivalent to the conditon that $\text{Im}\,(d_{t+1}) \subseteq \text{Ker}\,(d_t)$. We define the t th *homology* module of the complex $G_\bullet$, denoted $H_t(G_\bullet)$, by

$$H_t(G_\bullet) = \text{Ker}\,(d_t)/\text{Im}\,(d_{t+1}).$$

The module $\text{Ker}\,(d_t)$ is referred to as the module of *cycles* in G_t (and its elements are called *cycles*, and the module $\text{Im}\,(d_{t+1})$ is referred to as the module of *boundaries* in G_t (and its elements are called *boundaries*). In a complex, every boundary is a cycle.

The complex $G_\bullet$ is called *exact* at G_t or *exact* at the t th spot if, equivalently, $\text{Im}\,(d_{t+1}) = \text{Ker}\,(d_t)$ or $H_t(G_\bullet) = 0$. Thus, when we have exactness at G_t, every cycle in G_t is a boundary (the converse statement always holds in a complex). A complex

is called *exact* if it is exact at every spot. Equivalently, a complex is exact if all of its homology modules vanish. A left complex $G_\bullet$ is called *acyclic* if $H_t(G_\bullet) = 0$ for all $t \geq 1$. This leaves the possibility that $H_0(G_\bullet) \neq 0$. In this $H_0(G) = G_0/\mathrm{Im}\,(G_1)$, and $H_0(G_\bullet)$ is sometimes referred to as the *augmentation* module for $G_\bullet$. The augmented complex

$$\cdots \to G_t \to \cdots \to G_1 \to G_0 \to H_0(G_\bullet) \to 0$$

is exact.

By a map $\phi = \phi_\bullet$ of complexes of A-modules $F_\bullet \to G_\bullet$ we mean a family of A-module maps $\phi_t : F_t \to G_t$ such that the diagram

$$\begin{array}{ccccccccc}
\cdots & \longrightarrow & F_{t+1} & \longrightarrow & F_t & \longrightarrow & F_{t-1} & \longrightarrow & \cdots \\
\cdots & & \downarrow{\scriptstyle \phi_{t+1}} & & \downarrow{\scriptstyle \phi_t} & & \downarrow{\scriptstyle \phi_{t-1}} & & \cdots \\
\cdots & \longrightarrow & G_{t+1} & \longrightarrow & G_t & \longrightarrow & G_{t-1} & \longrightarrow & \cdots
\end{array}$$

commutes. In this case, for each t there is a map of homology $H_t(F_\bullet) \to H_t(G_\bullet)$: if $z \in F_t$ is a cycle representing an element $[z] \in H_t(F_\bullet)$, the value of the induced map on $[z]$ is $[\phi_t(z)]$, which turns out to depend only on $[z]$. If all the ϕ_t are injective, $F_\bullet$ is called a *subcomplex* of $G_\bullet$, and if all the ϕ_t are surjective, $G_\bullet$ is called a *quotient complex* of $F_\bullet$.

One says that

$$0 \to E_\bullet \to F_\bullet \to G_\bullet \to 0$$

is a *short exact sequence of complexes* if for every t the sequence

$$0 \to E_t \to F_t \to G_t \to 0$$

is exact and, in this case, the *Snake Lemma* or *Serpent Lemma* asserts there is a long exact sequence of homology

$$\cdots \to H_{t+1}(G_\bullet) \to H_t(E_\bullet) \to H_t(F_\bullet) \to H_t(G_\bullet) \to H_{t-1}(E_\bullet) \to \cdots .$$

The maps $\partial_t : H_t(G_\bullet) \to H_{t-1}(E_\bullet)$ are referred to as the *connecting homomorphisms*. If $z \in G_t$ is a cycle representing a homology class $[z]$, we can choose an element $\widetilde{z} \in F_t$ that maps to it. The image y of $\widetilde{z}$ in F_{t-1} maps to 0 in G_{t-1}, and so there is an element $\widetilde{y} \in E_{t-1}$ that maps to y. It is easy to see that $\widetilde{y}$ is a cycle in E_{t-1}, and one defines $\partial_t([z]) = [\widetilde{y}]$. The definition turns out to be independent of the choices made.

Whenever $\phi_\bullet : E_\bullet \to F_\bullet$ is a subcomplex, we may form a quotient complex $G_\bullet$ by letting $G_t = \mathrm{Coker}\,(\phi_t) \cong F_t/\mathrm{Im}\,(E_t)$. The differential is induced by the differential on $F_\bullet$. Similarly, whenever $F_\bullet \to G_\bullet$ is a quotient complex, we may let $E_t = \mathrm{Ker}\,(\phi_t) \subseteq F_t$, and $E_\bullet$ is a subcomplex under the restriction of the differential on $F_\bullet$. In both these cases, the sequence $0 \to E_\bullet \to F_\bullet \to G_\bullet \to 0$ is a short exact sequence of complexes.

Some acyclic complexes and Diana Taylor's resolution for monomial ideals

Let B be any commutative ring. Let $k \in \mathbb{N}$ be fixed, and let G_t denote the free module B-module with free basis $u_{i_1,\dots,i_{t+1}}$ where $1 \leq i_1 < i_1 < \cdots < i_{t+1} \leq k$, so that the generators of G_t are in bijective correspondence with the $t+1$ element subsets of $\{1, 2, \dots, k\}$. In fact, if $\sigma = \{i_1, \dots, i_{t+1}\}$ with $1 \leq i_1 < \cdots < i_{t+1} \leq k$, we shall also write u_σ for $u_{i_1,\dots,i_{t+1}}$.

If $t > k-1$ or $t < 0$ we define $G_t = 0$. Then one forms a complex

$$0 \to G_{k-1} \to \cdots \to G_0 \to 0$$

by defining the differential on G_t as follows. Since G_t is free, it suffices to specify the differential d_t on a typical generator, and if σ is the set $\{i_1, \dots, i_{t+1}\}$ with

$$1 \leq i_1 < \cdots < i_{t+1} \leq k$$

then $d_t(u_\sigma) = \sum_{j=1}^{t+1} (-1)^j u_{\sigma - \{i_j\}}$. It is easy to check that $d_{t-1} \circ d_t = 0$. The point is that after applying both maps, one gets a sum of terms $\pm u_{\sigma - \{i_j, i_{j'}\}}$ as $j \neq j'$ run through all pairs of distinct integers in the set $\{1, \dots, t+1\}$. Each term occurs exactly twice, once when i_j is deleted first and then $i_{j'}$, and a second time when $i_{j'}$ is deleted first and then i_j. It easy to verify that the signs one gets on these two occurrences are opposite, so that all terms cancel.

For those familiar with simplicial homology, we remark that this complex is precisely the complex used to calculate the simplicial homology of a $(k-1)$-simplex. It is therefore well-known that:

Proposition. *For all $k \geq 1$, the complex $G_\bullet$ described above is acyclic and $H_0(G_\bullet) \cong B$. Moreover, if we augment $G_\bullet$ by letting $G_{-1} = Bu_\emptyset$, where the new differential maps every u_i to $u_\emptyset$, the complex*

$$0 \to G_{k-1} \to \cdots \to G_0 \to G_{-1} \to 0$$

is exact.

Proof. We shall give two elementary proofs of this. We leave it to the reader to check that the first statement implies the second.

In the first proof proceed by induction on k. If $k = 1$, the complex is simply

$$0 \to Bu_1 \to 0$$

and the result is clear. Suppse $k > 1$. In the general case, note that the complex $F_\bullet$ corresponding to the set $1, 2, \dots k-1$ is a subcomplex. The quotient complex has free generators indexed by subsets of $\{1, 2, \dots, k\}$ such that k is an element of the subset.

These are in bijective correspondence with the subsets of $\{1, \ldots, k-1\}$ (including the empty set), and this gives a complex isomorphic with the augmented complex of F except that degrees are shifted by 1. Thus, the quotient $G_\bullet/F_\bullet$ is not merely acyclic, but exact, because it is augmented, and the result is now immediate from the Snake Lemma. $\square$

We can also prove acyclicty as follows. Let $h_t : G_t \to G_{t+1}$ be the map that sends $u_\sigma \mapsto 0$ if $1 \in \sigma$ and to $U_{\{1\}\cup\sigma}$ otherwise. Then for every σ,

$$d_{t+1}\big(h_t(u_\sigma)\big) + h_{t-1}\big(d_t(u_\sigma)\big) = u_\sigma$$

for $t \geq 1$ (consider the cases where $1 \in \sigma$ and $1 \notin \sigma$ separately). Thus, $d_{t+1}h_t + h_{t-1}d_t$ is the identity map on G_t. Suppose that $z \in G_t$ is a cycle, where $t \geq 1$. Then

$$d_{t+1}\big(h(z)\big) + h_{t-1}\big(d_t(z)\big) = z.$$

Since $d_t(z) = 0$, $d_{t+1}\big(h(z)\big) = z$, so that every cycle z is a boundary for $t \geq 1$. It remains to check that $H_0(G_\bullet) = B$, which we leave as an informal exercise. $\square$

We next want to describe Diana Taylor's resolution of a monomial ideal. We emphasize that these resolutions are rarely minimal.

We can make use of an arbitrary base ring B. Let $A = B[x_1, \ldots, x_n]$ be a polynomial ring and let $\mu_1, \ldots, \mu_k$ be monomials in A. We shall describe the resolution as an $\mathbb{N}^n$-graded complex: the generators of the free modules will typically have degrees in $\mathbb{N}^n$. The free basis of the tth free module will consist of elements $U_{i_1,\ldots,i_{t+1}}$ indexed by sequences $1 \leq i_1 < \cdots < i_{t+1} \leq k$, just as before. We give this generator the same degree as $\mathrm{LCM}(\mu_{i_1}, \ldots, \mu_{i_{t+1}})$. Then F_t is spanned as a free B-module by the elements $\mu U_{i_1,\ldots,i_{t+1}}$, where μ is a monomial in A, and this element will have the same degree as $\mu\mathrm{LCM}(\mu_{i_1}, \ldots, \mu_{i_{t+1}})$. If $\sigma = \{i_1, \ldots, i_{t+1}\}$, it will be convenient to write U_σ for $U_{i_1,\ldots,i_{t+1}}$, and to define

$$\mathrm{LCM}(\mu_\sigma) = \mathrm{LCM}(\mu_{i_1}, \ldots, \mu_{i_{t+1}}).$$

We can now define the differential on $F_\bullet$ by the rule

$$d_t(U_\sigma) = \sum_{j=1}^{t+1} (-1)^j \frac{\mathrm{LCM}(\mu_\sigma)}{\mathrm{LCM}(\mu_{\sigma-\{i_j\}})} U_{\sigma-\{i_j\}}$$

Note that this formula preserves degrees. Let $I = (\mu_1, \ldots, \mu_k)A$, and augment the complex $F_\bullet$ by the map $F_0 \twoheadrightarrow I$ such that $U_i \mapsto \mu_i$ for $1 \leq i \leq k$. Note that the maps d_t preserve degree.

Theorem (Diana Taylor). *Let $A = B[x_1, \ldots, x_n]$, $\mu_1, \ldots, \mu_k$, I, and $F_\bullet$ be as above. Then*

$$0 \to F_{k-1} \to \cdots \to F_0 \to 0$$

is an acyclic complex that gives a free resolution of I, i.e., the augmented complex

$$0 \to F_{k-1} \to \cdots \to F_0 \to I \to 0$$

is exact.

Lecture of January 22

Proof. Because the maps are degree preserving, it suffices to prove that the complex is exact in each degree α. In fact, the full complex

$$0 \to F_{k-1} \to \cdots \to F_0 \to I \to 0$$

is the direct sum of the homogeneous subcomplexes

$$(*_\alpha) \quad 0 \to [F_{k-1}]_\alpha \to \cdots [F_0]_\alpha \to [I]_\alpha \to 0.$$

It will therefore suffice to prove that each of the complexes $(*_\alpha)$ is exact.

Note the following: the contribution to $[F_t]_\alpha$ from AU_σ is 0 unless $\mathrm{LCM}(\mu_\sigma)$ divides x^α. In this case, there is a unique monomial ν_σ such that $\nu_\sigma \mathrm{LCM}(\mu_\sigma) = x^\alpha$, so that $[F_t]_\alpha$ is the free B-module generated by the elements $\nu_\sigma U_\sigma$ such that $\mathrm{LCM}(\mu_\sigma)$ divides x^α. Let $\mu_{j_1}, \ldots, \mu_{j_h}$ with $j_1 < \cdots < j_h$ be the generators of I that divide α. Then $\mathrm{LCM}(\mu_\sigma)$ divides x^α iff μ_i divides x^α for every $i \in \sigma$ iff $\sigma \subseteq \{j_1, \ldots, j_h\}$.

Therefore, if $x^\alpha \notin I$, every $[F_t]_\alpha = 0$ and $[I]_\alpha = 0$, while if $x^\alpha \in I$, and $\mathcal{S}_\alpha = \{\mu_{j_1}, \ldots, \mu_{j_h}\}$ is the set of generators of I that divide x^α, $[F_t]_\alpha$ is the free B-module on the elements $\nu_\sigma U_\sigma$ such that $\sigma \subseteq \mathcal{S}_\alpha$ and σ is a set with $t+1$ elements. The set $\mathcal{S}_\alpha$ is in bijective correspondence with $\{1, \ldots, h\}$, with μ_{j_i} corresponding to i, and for each $t+1$ element subset τ of $\{1, \ldots, h\}$ we may let u_τ denote the element $\nu_\sigma U_\sigma \in [F_t]_\alpha$, where σ is the $t+1$ element subset of $\mathcal{S}_\alpha$ corresponding to τ. The complex $[F_t]_\alpha$ is then isomorphic to an augmented complex $G_\bullet$ over B of the form described at the bottom of p. 6 and on p. 7 of the Lecture Notes of January 20 (but with h replacing k), and so is exact by the Proposition on p. 7 of those notes. $\square$

Finding Hilbert-Poincaré series

Let M be a finitely generated module over $K[x_1, \ldots, x_n]$. When we consider the $\mathbb{N}^n$ grading on R, we shall allow $\mathbb{Z}^n$-gradings on M. When we consider the $\mathbb{N}$-grading on R, we shall allow $\mathbb{Z}$-gradings on M.

Note that, quite generally, when $H \subseteq H'$ is a subsemigroup of the additive semigroup H' and R is an H-graded ring, we can also view R as H'-graded by letting $R_{h'} = 0$ for

$h' \in H' - H$. Therefore, we can consider H'-graded modules M over the H-graded ring R. In effect, the condition becomes that for $h \in H$ and $h' \in H'$, $R_h M_{h'} \subseteq M_{h+h'}$.

In our cases $H = \mathbb{N}^n$ and $H' = \mathbb{Z}^n$ or $H = \mathbb{N}$ and $H' = \mathbb{Z}$.

Because M is finitely generated over $R = K[x_1, \ldots, x_n]$, if $-B$ is the smallest integer such that some generator of M has a degree involving $-B$, then all nonzero homogeneous elements of M have degree $\geq -B$ in every coordinate: when we multiply by monomials in R, degrees can only increase.

When M is $\mathbb{Z}$-graded, this means that there only finitely many nonzero components of M in negative degree.

If $\alpha \in \mathbb{Z}^n$, we define $M(\alpha)$ (sometimes called M *twisted by* α or the α th *twist of* M) to be the $\mathbb{Z}^n$-graded module that is isomorphic to M as an R-module but with grading shifted so that for all $\beta \in \mathbb{Z}^n$,

$$[M(\alpha)]_\beta = M_{\alpha+\beta}.$$

One reason for introducing these shifted gradings is that in considering free resolutions of graded modules one often wants to use maps that preserve degree. In doing this, one may need to shift gradings even when working with free modules.

Consider one of the simplest possible examples, where $R = K[x]$ and $M = K[x]/xK[x]$, which has the free resolution:

$$0 \to K[x] \xrightarrow{x\cdot} K[x] \to M \to 0$$

The element $1 \in K[x]$ in the leftmost module maps to x in the copy of $K[x]$ to the right. If the map is to be degree-preserving, we need $1 \in K[x]$ to have degree 1. If the right hand copy of $K[x]$ has the usual grading, this means that the leftmost copy should be twisted by -1. The resolution is then

$$0 \to R(-1) \xrightarrow{x\cdot} R \to M \to 0.$$

Note that $[R(-1)]_1 = [R]_{1+(-1)} = [R]_0 = K$, so that 1 has degree 1 in $R(-1)$. Typically, 1 has degree t in $R(-t)$ for all $t \in \mathbb{Z}$.

Also note that any finitely generated $\mathbb{Z}^n$-graded module M over $R = K[x_1, \ldots, x_n]$ has a twist $M(\alpha)$ with the property that $[M(\alpha)]_\beta$ is a nonzero component only if $\beta \in \mathbb{N}^n$. If no generator involves a degree smaller than $-B$ in any component, we may take $\alpha = (-B, \ldots, -B)$. If β has any strictly negative entry, $\alpha + \beta$ has entry $< -B$, and $[M(\alpha)]_\beta = 0$.

We next define the Hilbert-Poincaré series $\mathfrak{P}^\mu_M(z_1, \ldots, z_n)$ of an $\mathbb{N}^n$-graded module M over $K[x_1, \ldots, x_n]$ (here, the superscript μ indicates that we are using the $\mathbb{N}^n$-graded version) by the formula

$$\mathfrak{P}^\mu_M(z_1, \ldots, z_n) = \mathfrak{P}^\mu_M(z) = \sum_{\alpha \in \mathbb{Z}^n} \dim_K([M]_\alpha) z^\alpha,$$

which *a priori* is an element of

$$\mathbb{Z}[[z_1, \ldots, z_n]](1/z_1 \cdots z_n).$$

However, we shall soon prove that these series are actually rational functions of $z_1, \ldots, z_n$.

We first consider the case of R itself. Then

$$\mathfrak{P}_M^\mu(z) = \sum_{\alpha \in \mathbb{N}^n} z^\alpha = (1 + z_1 + z_1^2 + \cdots)(1 + z_2 + z_2^2 + \cdots) \cdots (1 + z_n + z_n^2 + \cdots)$$

$$= \prod_{i=1}^n \frac{1}{1 - z_i} = \frac{1}{\prod_{i=1}^n (1 - z_i)}.$$

Note that if we have a short exact sequence of $\mathbb{Z}^n$-graded finitely generated modules and degree-preserving maps, say $0 \to M_2 \to M_1 \to M_0 \to 0$, then we get a short exact sequence of vector spaces

$$0 \to [M_2]_\alpha \to [M_1]_\alpha \to [M_0]_\alpha \to 0$$

for every α. It follows that

$$\mathfrak{P}_{M_1}^\mu(z) = \mathfrak{P}_{M_0}^\mu(z) + \mathfrak{P}_{M_2}^\mu(z).$$

More generally, given a finite exact sequence

$$0 \to M_h \to \cdots \to M_0 \to 0$$

of finitely generated $\mathbb{N}^n$-graded modules and degree preserving maps, we have that

$$\sum_{i=0}^h (-1)^i \mathfrak{P}_{M_i}^\mu(z) = 0.$$

This follows simply because the exact sequence of length h can be broken up into short exact sequences. Diana Taylor's resolution for monomial ideals now yields the following.

Theorem. *Let I be a monomial ideal with generators $\mu_1 = x^{\alpha_1}, \ldots, \mu_k = x^{\alpha_k}$ in $R = K[x_1, \ldots, x_n]$. Then $\mathfrak{P}_{R/I}^\mu(z)$ is a rational function of $z_1, \ldots, z_n$ whose numerator has integer coefficients and whose denominator is at worst $\prod_{i=1}^n (1 - z_i)$. More precisely, let Σ_t denote the sum of the least common multiples of the monomials $z^{\alpha_1}, \ldots, z^{\alpha_k}$ taken t at a time, for $0 \le t \le k$, where $\Sigma_0 = 1$. Then*

$$\mathfrak{P}_{R/I}^\mu(z) = \frac{\Sigma_0 - \Sigma_1 + \Sigma_2 - \cdots + (-1)^k \Sigma_k}{\prod_{i=1}^k (1 - z_i)}.$$

Proof. We can modify Diana Taylor's resolution slightly by putting it together with the short exact sequence $0 \to I \to R \to R/I \to 0$ to give

$$0 \to F_{k-1} \to \cdots \to F_0 \to R \to R/I \to 0.$$

Consequently, we have

$$(*) \quad \mathfrak{P}^{\mu}_{R/I}(z) = \mathfrak{P}^{\mu}_{R}(z) - \sum_{i=0}^{k-1}(-1)\mathfrak{P}^{\mu}_{F_i}(z).$$

F_i is the direct sum of copies of R, one for each $i+1$ element subset σ of $\{1, \ldots, k\}$, with the generator of R in degree $\mathrm{LCM}(\mu_\sigma) = x^{\beta_\sigma}$. The Hilbert-Poincare series of this cyclic free module is $z^{\beta}_{\sigma}\mathfrak{P}^{\mu}_{R}(z)$. It follows that the Hilbert-Poincaré series

$$\mathfrak{P}^{\mu}_{F_i}(z) = \Sigma_{i+1}\mathfrak{P}^{\mu}_{R}(z).$$

The result now follows from substituting this in $(*)$ and noting that

$$\mathfrak{P}^{\mu}_{R}(z) = \frac{1}{\prod_{i=1}^{n}(1-z_i)}. \qquad \square$$

Corollary. *If $F = R^s$ is free, for every monomial submodule M of F, F/M and M have Hilbert-Poincaré series that are rational functions whose numerator is a polynomial with integer coefficients and whose denominator is at worst $\prod_{i=1}^{n}(1-z_i)$.*

Proof. The monomial submodule is a direct sum of monomial ideals, one in each Re_i. $\square$

We want to consider what happens when the generators of F may have degrees shifted by twisting. The key point is that for any finitely generate $\mathbb{N}^n$-graded module M and any α,

$$\mathfrak{P}^{\mu}_{M(\alpha)}(z) = \sum_{\beta\in\mathbb{Z}^n} \dim_K([M]_{\alpha+\beta})z^{\beta} = z^{-\alpha}\sum_{\beta\in\mathbb{Z}^n} \dim_K([M]_{\alpha+\beta})z^{\alpha+\beta} = z^{-\alpha}\mathfrak{P}^{\mu}_{M}(z),$$

since as β runs through all of $\mathbb{Z}^n$, so does $\alpha+\beta$.

We now want use our monomial results to prove theorems about Hilbert-Poincaré series in the $\mathbb{N}$-graded case. As in the $\mathbb{Z}^n$-graded case,

$$\mathfrak{P}_{M(h)}(z) = z^{-h}\mathfrak{P}_{M}(z).$$

Next note:

Proposition. *Let M be a finitely generated $\mathbb{Z}^n$ graded modules over $K[x_1, \ldots, x_n]$. Then $\mathfrak{P}_M(z) = \mathfrak{P}_M^{\mu}(z, z. \ldots, z)$ (i.e., z is substituted for every z_i).*

In particular, $\mathfrak{P}_R(z) = \dfrac{1}{(1-z)^n}$. Hence, if $M \subseteq F = R^s$ (which includes the case $I \subseteq R$) is monomial then both $\mathfrak{P}_M(z)$ and $\mathfrak{P}_{F/M}(z)$ are rational functions in which the numerator is a polynomial in z with integer coefficients and the denominator is at worst $(1-z)^n$.

In general, for any finitely geneated $\mathbb{Z}^n$-graded module M, $\mathfrak{P}_M(z)$ is a rational function of z whose numerator is a polynomial in z with integer coefficients and whose denominator is, at worst, $z^B(1-z)^n$ for some $B \geq 0$.

Proof. If $\alpha = (a_1, \ldots, a_n) \in \mathbb{Z}_n$, we write $|\alpha|$ for $a_1 + \cdots + a_n$. Then for every integer i,

$$[M]_i = \bigoplus_{|\alpha|=i} [M]_\alpha,$$

and so

$$\dim_K([M]_i) = \sum_{|\alpha|=i} \dim_K([M]_\alpha),$$

and the result follows at once from this observation. The remaining statements are immediate. □

We can now obtain a result for arbitrary finitely generated modules in the graded case.

Theorem. *Let N be any finitely generated $\mathbb{Z}$-graded module over $R = K[x_1, \ldots, x_n]$. Suppose that $u_1, \ldots, u_s$ are finitely many homogeneous generators of respective degrees $d_1, \ldots, d_s$. Think of R^s as $\bigoplus_{j=1}^s R(-d_j)$, and map $R^s \twoheadrightarrow N$ so that $1 \in R(-d_j)$, which has degree d_j, maps to u_j. This map preserves degrees, and the kernel M is an $\mathbb{N}$-graded submodule of R^s.*

Refine the $\mathbb{Z}$-grading on R^s to a $\mathbb{Z}^n$-grading, and choose a monomial order. Then N and $F/\mathrm{in}(M)$ have the same Hilbert-Poincaré series! Hence, the Hilbert-Poincaré series of N is a rational function of z with numerator that is a polynomial in z with integer coefficients and denominator at worst $z^B(1-z)^n$ for some $B \in \mathbb{N}$.

Proof. By the Theorem near the bottom of p. 2 of the Lecture of January 13, the monomials of F not in $\mathrm{in}(M)$ are a basis for F/M, and they are clearly a basis of homogeneous elements. Hence, the monomials of a given degree d are a K-vector space basis for $[F/M]_d$, and also for $[F/\mathrm{in}(M)]_d$, and so $\dim_K([F/M]_d) = \dim_K([F/\mathrm{in}(M)]_d)$ for all d. The first conclusion follows at once, and the second then follows as well because we already know the result in the monomial case. □

Lecture of January 25

Hilbert functions

Let M be a finitely generated graded module over $R = K[x_1, \ldots, x_n]$, a polynomial ring over a field. The *Hilbert function* Hilb_M of M is defined by the formula

$$\mathrm{Hilb}_M(d) = \dim_K([M]_d)$$

for all $d \in \mathbb{Z}$. It is always 0 for $d \ll 0$. This means that

$$\mathfrak{P}_M(z) = \sum_{d\in\mathbb{Z}} \mathrm{Hilb}_M(d) z^d,$$

so that the Hilbert function and the Hilbert-Poincaré series carry the same information.

Before going furrther, we consider what happens when $M = R$, in which case we know that

$$\mathfrak{P}(z) = \frac{1}{(1-z)^n} = (1-z)^{-n}.$$

We can evaluate the coefficients using Newton's binomial theorem, which is just a special case of Taylor's formula. Then coefficient of z^d is then

$$\frac{(-n)(-n-1)(-n-2)\cdots\big(-n-(d-1)\big)}{d!}(-1)^d = \frac{n(n+1)\cdots(n+d-1)}{d!}$$

which is

$$\binom{n+d-1}{d} = \binom{d+n-1}{n-1}.$$

We can get the same formula from a purely combinatorial argument. $\mathrm{Hilb}(d)$ is the number of monomials x^α where $\alpha = (a_1, \ldots, a_n)$ where the $a_i \in \mathbb{N}$ and $a_1 + \cdots + a_n = d$. Each such monomial can be represented by a string containing d blanks _ interspersed with $n-1$ slashes /, where there are first a_1 blanks, then a slash as a separator, then a_2 blanks, then a slash as a separator, and so forth. The string will end with a slash, then a_{n-1} blanks, then a slash, and, finally a_n blanks. (For example, if $n = 4$ and $d = 8$, the string corresponding to $x_1^3 x_3 x_4^5$ is

$$\_\ \_\ \_\ /\ /\ \_\ /\ \_\ \_\ \_\ \_\ \_\ .$$

This gives a bijection between monomials of degree d in $x_1, \ldots, x_n$ and strings of length $d+n-1$ consisting of d blanks and $n-1$ slashes. The number of such strings is determined by the choice of which positions are occupied by the slashes among the $d+n-1$ possibilities, and this is $\binom{d+n-1}{n-1}$.

In any case, we see that the Hilbert function of R agrees with $\binom{d+n-1}{n-1}$ for all sufficiently large d, and this is a polynomial in d of degree $n-1$.

We can immediately derive the following result on Hilbert functions from the results we have on Hilbert-Poincaré series.

Theorem. *With hypothesis as the first paragraph, the Hilbert function of a $\mathbb{Z}$-graded finitely generated R-module M agrees with a polynomial of degree at most $n-1$ in d for all $d \gg 0$.*

Proof. By the last statement of the Theorem given at the bottom of p. 4 and the top of p. 5 of the Lecture Notes of January 22, we know that the Hilbert-Poincaré series of $\mathfrak{P}_M(z)$ is a $\mathbb{Z}$-linear combination of functions of the form $\dfrac{z^c}{(1-z)^n}$ for $c \in \mathbb{Z}$. By the discussion above, for such a function the Hilbert function is given by $\binom{d-c+n-1}{n-1}$ for $d \gg 0$, and this is a polynomial in d of degree $n-1$. When we take a $\mathbb{Z}$-linear combination of such polynomials the highest degree terms may cancel, but the degree is still at most $n-1$. $\square$

The polynomial that agrees with $\mathrm{Hilb}_M(d)$ for $d \gg 0$ is called the *Hilbert polynomial* of M. Note that if one has a short exact sequence of finitely generated $\mathbb{Z}$-graded modules and degree preserving maps, say

$$0 \to M_0 \to M_1 \to M_2 \to 0,$$

it follows that

$$\mathrm{Hilb}(M_1) = \mathrm{Hilb}(M_0) + \mathrm{Hilb}(M_2),$$

just as in the case of Hilbert-Poincaré series. Obviously, the same holds for Hilbert polynomials. Likewise, if one has a finite exact sequence of finitely generated $\mathbb{Z}$-graded modules and degree preserving maps, the alternating sum of the Hilbert functions is 0, and the alternating sum of the Hilbert polynomials is likewise 0.

The module of relations on a Gröbner basis: Schreyer's method

Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over a field K and let F be a finitely generated free R-module with ordered basis $b_1, \ldots, b_s$ for which we have fixed a monomial order.

Let $M \subseteq F$ be a submodule of F for which we have a Gröbner basis $g_1, \ldots, g_r$. Consider the module N of relations on $g_1, \ldots, g_r$, i.e.,

$$N = \{(f_1, \ldots, f_r) \in R^r : \sum_{j=1}^{r} f_j g_j = 0\}.$$

It turns out that there is an almost unbelievably simple method for finding a finite set of generators for N: beyond that, for a suitably chosen monomial order on R^r, these generators a Gröbner basis for N. The method, which is due to Schreyer, is *very* closely related to the Buchberger criterion.

This means that once we have a Gröbner basis for M, we immediately get a Gröbner basis for N, which is a first module of syzygies of M. We are then immediately ready to find a module of syzygies of N, and we can continue in this way to get as many iterated modules of syzygies as we wish.

We shall use $e_1, \ldots, e_r$ as the ordered basis for R^r: it will be convenient to have a notation that distinguishes it from the ordered basis for $F \cong R^s$. Let $\nu_j = \mathrm{in}(g_j)$ for $1 \leq j \leq r$. We define a monomial order on R^r as follows: if μ and μ' are monomials in R, then $\mu e_i > \mu' e_j$ if and only if $\mathrm{in}(\mu g_i) > \mathrm{in}(\mu' g_j)$ (which is equivalent to $\mu\nu_i > \mu'\nu_j$) or $\mathrm{in}(\mu g_i) = \mathrm{in}(\mu' g_j)$ and $i < j$. It is quite straightforward to verify that this is a monomial order on R^r.

The Buchberger criterion provides certain relations on $g_1, \ldots, g_r$ which we shall refer to as *the standard relations*. These arise as follows: for each choice of $i < j$, we know that when we take some choice of standard expression for

$$\frac{\nu_j}{\mathrm{GCD}(\nu_i, \nu_j)} g_i - \frac{\nu_i}{\mathrm{GCD}(\nu_i, \nu_j)} g_j$$

with respect to division by $g_1, \ldots, g_r$, we get remainder 0. This means that for each $i < j$ we have

$$(\#_{ij}) \quad \frac{\nu_j}{\mathrm{GCD}(\nu_i, \nu_j)} g_i - \frac{\nu_i}{\mathrm{GCD}(\nu_i, \nu_j)} g_j = \sum_{k=1}^{r} q_{ijk} g_k$$

where every

$$\mathrm{in}(q_{ijk} g_k) \leq \mathrm{in}\left(\frac{\nu_j}{\mathrm{GCD}(\nu_i, \nu_j)} g_i - \frac{\nu_i}{\mathrm{GCD}(\nu_i, \nu_j)} g_j\right).$$

We obtain these relations because the remainders upon division must be 0. Note that, as in the case of Buchberger's criterion, it suffices to choose one standard expression: it need not be the result of the deterministic division algorithm.

The equation displayed in $(\#_{ij})$ corresponds to a relation on the g_{ij}, namely

$$\rho_{ij} = \frac{\nu_j}{\mathrm{GCD}(\nu_i, \nu_j)} e_i - \frac{\nu_i}{\mathrm{GCD}(\nu_i, \nu_j)} e_j - \sum_{k=1}^{r} q_{ijk} e_k.$$

It is the relations ρ_{ij} that we refer to as the "standard" relations on $g_1, \ldots, g_r$. They are not really unique, since the standard expressions for dividing by $g_1, \ldots, g_r$ are not unique, but, as we have already indicated, the result below is correct when one makes just one choice of standard expression for $i < j$. (Recall, however, that when one has a Gröbner basis $g_1, \ldots, g_r$, the *remainder* upon division by $g_1, \ldots, g_r$ is unique, and will always be zero if the element one is dividing is in the R-span of $g_1, \ldots, g_r$.) Here is the punchline:

Theorem (Schreyer). *Let notation be as above. Then the standard relations ρ_{ij} generate the module of relations on the Gröbner basis $g_1, \ldots, g_r$. What is more, the relations ρ_{ij}*

form a Gröbner basis for the module of relations on the $g_1, \ldots, g_r$ with respect to the monomial order on R^r defined above.

Proof. Of course, the second statement implies the first. We begin by studying

$$\mathrm{in}(f_1 e_1 + \cdots + f_r e_r)$$

for an arbitrary relation on $g_1, \ldots, g_r$. All we need to do is show that each such initial term is a multiple of one of the $\mathrm{in}(\rho_{ij})$. Each $\nu_i = \mathrm{in}(g_i)$ involves one element of the free basis $b_1, \ldots, b_s$ for the original free module R^e: call this element $b_{L(i)}$. Then the monomial μ in f_i that gives rise to the largest term of $f_i e_i$ after multiplying out is the same monomial μ that gives the largest term in $f_i g_i$, and this is $\mathrm{in}_{>_{L(i)}}(f_i)\nu_i$ by the displayed formula (†) on p. 2 of the Lecture Notes of January 20. It follows that the largest term in $f_i e_i$ is $\mathrm{in}_{>_{L(i)}} e_i$. Thus, $\mathrm{in}(f_1 e_1 + \cdots f_r e_r)$ may be described as follows. Consider the largest initial term for any $f_i g_i$, call it ν, and choose the smallest i such that ν is $\mathrm{in}(f_i g_i)$, up to a nonzero scalar multiple. Then $\mathrm{in}(f_1 e_1 + \cdots + f_r e_r)$ is $\mathrm{in}(f_i e_i) = \mathrm{in}_{>_{L(i)}}(f_i) e_i$ for this smallest value of i.

This is precisely the same use of ν as in the proof of the Buchberger criterion in the Lecture Notes of January 20.

We next want to understand $\mathrm{in}(\rho_{ij})$. In the equations $(\#_{ij})$ from which the ρ_{ij} are derived, the initial terms of the two products on the left hand side are the same, and cancel, while the initial term of every $q_{ijk} f_k$ is $\leq$ the initial term on the left. Hence, the initial term of every $q_{ijk} f_k$ is strictly smaller than the initial terms of the two products on the left hand side. When we replace the equation by ρ_{ij}, there is no cancellation, because g_i and g_j on the left have been replaced by e_i and e_j. Thus, the initial term of ρ_{ij} is $\dfrac{\nu_j}{\mathrm{GCD}(\nu_i, \nu_j)} e_i$.

Since $f_1 g_1 + \cdots + f_r g_r = 0$, the initial terms of products $f_j g_j$ that are, up to a nonzero scalar multiple, equal to ν must cancel. Suppose the products that have $c\nu$ as initial term for $c \in K - \{0\}$ are indexed by $j_1, \ldots, j_h$ where $j_1 < \cdots < j_h$. Let $\mu_j = \mathrm{in}_{>_{L(j)}}(f_j)$.

Then each $\mu_{j_t}\nu_{j_t}$ has the form $c_t \nu$ for $c_t \in K - \{0\}$, where $1 \leq t \leq h$, and the sum of the c_t is 0. With this notation, we have that

$$\mathrm{in}(f_1 e_1 + \cdots + f_r e_r) = \mu_{j_1} e_{j_1}.$$

We also have the relation $\sum_{t=1}^{h} \mu_t \nu_t = 0$. Exactly as in the proof of the Buchberger criterion, this means that $(\mu_1, \ldots, \mu_h)$ is a homogeneous linear combination, with coefficients that are terms in R, of the relations θ_{ij}: see the displayed line (#) near the top of p. 4 of the Lecture Notes of January 20 and the preceding discussion. However, in fact, we only need those θ_{ij} such that $i = j_a < j_b = j$. This means that μ_{j_1} must be a multiple, by a term in R, of the coefficient of e_{j_1} in some $\theta_{j_1 j_t}$ for $t > 1$. But this also means precisely that $\mu_{j_1} e_1$ is a multiple of $\mathrm{in}(\rho_{j_1 j_t})$ for some $t > 1$. $\square$

Finding the relations on elements that are not a Gröbner basis

We next want to address the problem of finding a basis for the relations on $g_1, \ldots, g_r$ when these elements are not necessarily a Gröbner basis for their span in F. The first step is to enlarge this set of elements to a Gröbner basis using the Buchberger algorithm. Note that if another generator h_{ij} is needed, it arises as a remainder for division of some

$$\frac{\nu_j}{\mathrm{GCD}(\nu_i, \nu_j)} g_i - \frac{\nu_i}{\mathrm{GCD}(\nu_i, \nu_j)} g_j$$

by $g_1, \ldots, g_r$, and so we will have a formula

$$h_{ij} = \frac{\nu_j}{\mathrm{GCD}(\nu_i, \nu_j)} g_i - \frac{\nu_i}{\mathrm{GCD}(\nu_i, \nu_j)} g_j - \sum_{j=1}^{r} q_j g_j,$$

so that we will be able to keep track of h_{ij} as an R-linear combination of the original $g_1, \ldots, g_r$. As we successively find new elements of the Gröbner basis, each can be expressed as an R-linear combination of its predecessors, and then as an R-linear combination of the original $g_1, \ldots, g_r$.

Suppose that the Gröbner basis that we find is $g_1, \ldots, g_{r+k}$, where we might as well assume that $k > 0$, or we already have a method. Moreover, we may assume that for $1 \leq i \leq k$ we have a formula

$$(**_i) \quad g_{r+i} = \sum_{j=1}^{r} f_{ij} g_j$$

We can now construct a surjective R-linear map from the module of relations on the Gröbner basis $g_1, \ldots, g_{r+k}$ onto the module of relations on $g_1, \ldots, g_r$. This is really the obvious thing to do: given the equation of a relation

$$u_1 g_1 + \cdots + u_r g_r + v_1 g_{r+1} + \cdots + v_k g_{r+k} = 0$$

we may substitute using the equations $(**_i)$ to express $g_{r+1}, \ldots, g_{r+k}$ in terms of $g_1, \ldots, g_r$, and then collect terms to get a relation on $g_1, \ldots, g_r$:

$$(u_1 + v_1 f_{11} + \cdots + v_k f_{k1}) g_1 + \cdots + (u_r + v_1 f_{1r} + \cdots + v_k f_{kr}) g_r = 0.$$

Thus, our map sends the vector $(u_1, \ldots, u_r, v_1, \ldots, v_k)$ to the vector whose jth entry is $u_j + v_1 f_{1j} + \cdots v_k f_{kj}$. This map is clearly linear. Moreover, $(u_1, \ldots, u_r, 0, 0, \ldots, 0)$ maps to $(u_1, \ldots, u_r)$, which shows that the map is surjective.

Thus, a basis for the relations on $g_1, \ldots, g_{r+k}$ maps onto a basis for the relations for $g_1, \ldots, g_r$. Since $g_1, \ldots, g_{r+k}$ is a Gröbner basis, we know how to find a basis for the relations, and we can then apply the map to get a basis for the relations on $g_1, \ldots, g_r$.

Finding generators for the intersection of two submodules of a free module

Suppose that we have generators $g_1, \ldots, g_r$ for $M \subseteq F$, and generators $g'_1, \ldots, g'_s$ for $N \subseteq F$. We want to find generators for $M \cap N$. Given any element of $M \cap N$, it can be written as an R-linear combination of the elements $g_1, \ldots, g_r$, and also as an R-linear combination of the elements $g'_1, \ldots, g'_s$. This leads to an equation

$$(\#) \quad f_1 g_1 + \cdots + f_r g_r = f'_1 g'_1 + \cdots + f'_s g'_s,$$

so that $(f_1, \ldots, f_r, -f'_1, \ldots, -f'_s)$ is a relation on $g_1, \ldots, g_r, g'_1, \ldots, g'_s$. (The original element is the common value of the two sides of the equation (#).) Conversely, given a relation, say $(f_1, \ldots, f_{r+s})$, on $g_1, \ldots, g_r, g'_1, \ldots, g'_s$, we have that

$$f_1 g_1 + \cdots + f_r g_r = (-f_{r+1}) g'_1 + \cdots + (-f_{r+s}) g'_s,$$

so that the left hand side represents an element of $M \cap N$. It follows that we have a surjection from the module Q of relations on $g_1, \ldots, g_r, g'_1, \ldots, g'_s$ onto $M \cap N$ that sends $(f_1, \ldots, f_{r+s}) \mapsto f_1 g_1 + \cdots + f_r g_r$. Therefore, we can find a basis for Q, which we already know how to do, and apply the map to obtain a basis for $M \cap N$.

Lecture of January 27

Review of the theory of Krull dimension

We recall that the (*Krull*) *dimension* of a ring R, which need not be Noetherian, is the supremum of lengths k of strictly increasing chains $P_0 \subset P_1 \subset \cdots \subset P_{k-1} \subset P_k$ of chains of prime ideals of R. The *height* of a prime ideal P is, equivalently, either the supremum of lengths of strictly descending chains of primes whose first element is P, or the dimension of the quasilocal ring R_P (a *quasilocal* ring is a ring with a unique maximal ideal).

We have:

Proposition. *If J is an ideal of R consisting of nilpotent elements, then* $\dim(R) = \dim(R/J)$. *Hence, if I and I' are two ideals of R with the same radical,* $\dim(R/I) = \dim(R/I')$.

Proof. There is an order preserving bijection between primes of R and primes of R/J: every prime ideal P of R contains J, and we may let P correspond to P/J. The second statement now follows because if $J = \mathrm{Rad}\,(I) = \mathrm{Rad}\,(I')$, then R/J is obtained from either R/I or R/I' killing an ideal (J/I or J'/I) all of whose elements are nilpotent. $\square$

Theorem. *If $R \subseteq S$ is an integral extension of rings, then* $\dim(R) = \dim(S)$.

Proof. Given any finite strictly ascending chain of primes in R there is a chain of the same length in S by the going up theorem. Hence, $\dim(R) \leq \dim(S)$. On the other hand, given a strictly ascending chain of primes of S, we obtain a strictly ascending chain of primes in R by intersecting its elements with R. The intersections with R of comparable but distinct primes of S are distinct by the lying over theorem. $\square$

If R is Noetherian, every prime has finite height. In fact:

Krull Height Theorem. *If R is Notherian and $I \subseteq R$ is generated by n elements, the height of any minimal prime P of R is at most n. Moreover, every prime ideal of height n is a minimal prime of an ideal generated by n elements.*

By a local ring (R, m, K) we mean a Noetherian ring with a unique maximal ideal m such that $K = R/m$.

Corollary. *If R is a local ring, the dimension of R (which is the same as the height of m) is the least number n of elements $x_1, \ldots, x_n \in m$ such that m is the radical of $(x_1, \ldots, x_n)R$.*

A set of n elements as described above is called a *system of parameters* for the local ring R. When R is zero-dimensional, the system of parameters is empty.

Corollary. *If $f \in m$, where (R, m, K) is local, then* $\dim(R/fR) \geq \dim(R) - 1$.

Proof. Choose a system of parameters for R/fR that are the images of elements $x_2, \ldots, x_s$ in m, where $s - 1 = \dim(R/xR)$. Since m/fR is nilpotent on $(x_2, \ldots, x_s)$, we have that m is nilpotent on $(f, x_2, \ldots, x_s)R$. Therefore, $\dim(R) \leq s = \dim(R/fR) + 1$. $\square$

Theorem. *Let R be a domain finitely generated over a field K. The dimension n of R is the transcendence degree of its fraction field over K. Every maximal ideal of R has height n, and for any two primes $P \subseteq Q$, a maximal ascending chain of primes from P to Q (also called a* saturated chain *from P to Q) has length equal to* $\text{height}(Q) - \text{height}(P)$.

When R is finitely generated over a field K, it is an integral extension of a polynomial subring, by the Noether normalization theorem. This suggests why the statements in this Theorem ought to be true, and a proof can be based on this idea.

Krull dimension for modules

If M is a finitely generated module over a Noetherian ring R, we define the (*Krull*) *dimension* of M to be the Krull dimension of R/I, where $I = \text{Ann}_R M$ is the annihilator of

I. We make the convention that the Krull dimension of the 0 ring is -1, and this means that the Krull dimension of the 0 module is also -1. Recall that the *support* of M, denoted $\mathrm{Supp}\,(M)$ is

$$\{P \in \mathrm{Spec}\,(R) : M_P \neq 0\}.$$

Also recall:

Proposition. *If M is a finitely generated module over a Noetherian ring R, $\mathrm{Supp}\,(M) = V(I)$, the set of prime ideals containing $I = \mathrm{Ann}_R M$.*

Proof. Let $u_1, \ldots, u_k$ generate M. Then the map $R \to M^k$ that sends $r \mapsto (ru_1, \ldots, ru_k)$ has kernel precisely I, which yields an injection $R/I \hookrightarrow M^k$. If $I \subseteq P$, then $(R/I)_P \neq 0$ injects into $(M^k)_P \cong (M_P)^k$, and so $M_P \neq 0$. Conversely, if $f \in I - P$, then M_P is localization of M_f, which is 0 since $fM = 0$. $\square$

Recall that a prime ideal is an *associated* prime of M if there is an injection $f : R/P \hookrightarrow M$. It is equivalent to assert that there is an element $u \in M$ such that $\mathrm{Ann}_R u = P$. The set of associated primes of M is denoted $\mathrm{Ass}\,(M)$. By a theorem, $\mathrm{Ass}\,(M)$ is finite.

Proposition. *Let R be a Noetherian ring and let M be a finitely generated R-module.*

(a) *The dimension of M is* $\sup\{\dim\,(R/P) : P \in \mathrm{Supp}\,(M)\}$.

(b) *The dimension of M is* $\sup\{\dim\,(R/P) : P$ *is a minimal prime of* $M\}$.

(c) *The dimension of M is* $\sup\{\dim\,(R/P) : P \in \mathrm{Ass}\,(M)\}$.

(d) *Let* $0 \to M' \to M \to M'' \to 0$ *be a short exact sequence. Then* $\dim\,(M) = \max\{\dim\,(M'), \dim\,(M'')\}$.

(e) *If* $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_{k-1} \subseteq M_k = M$ *is a finite fliltration of M, then* $\dim\,(M) = \sup\{\dim\,(M_{i+1}/M_i : 0 \leq i \leq k-1\}$.

Proof. (a) and (b). Since $\mathrm{Supp}\,(M)$ is $V(I)$, the assertion comes down to the statement that $\dim\,(R/I) = \sup\{\dim\,(R/P) : I \subseteq P\}$. This is clear, since I has only finitely many minimal primes $P_1, \ldots, P_h$, and so $\dim\,(R/I)$ is the supremum of the integers $\dim\,(R/P_j)$ where $1 \leq j \leq h$.

(c) The minimal primes of M (equivalently, of the support of M) are the same as the minimal primes P of I. As in the proof of the preceding Proposition we have $R/I \hookrightarrow M^k$, and then

$$P \in \mathrm{Ass}\,(R/I) \subseteq \mathrm{Ass}\,(M^k) = \mathrm{Ass}\,(M),$$

so that every minimal prime of I is in $\mathrm{Ass}\,(M)$. On the other hand, if $P \in \mathrm{Ass}\,(M)$ then $R/P \hookrightarrow M$ and so I kills R/P, i.e., $I \subseteq P$. Part (c) follows at once.

(d) Let I', I, and I'' be the annihilators of M', M, and M'' respectively. Then $I \subseteq I'$ and $I \subseteq I''$, so that $I \subseteq I' \cap I''$. If $u \in M$, then $I''u \subseteq M'$ (since I' kills $M/M' = M''$), and so I' kills $I''u$, i.e., $I'I''u = 0$. This implies that $I'I'' \subseteq I$. Now $(I' \cap I'')^2$ is generated

by products fg where $f, g \in I' \cap I''$. Think of f as in I' and g as in I''. It follows that $(I' \cap I'')^2 \subseteq I'I'' \subseteq I' \cap I''$, so that $\operatorname{Rad}(I' \cap I'') = \operatorname{Rad}(I'I'')$, and we have that $\operatorname{Rad}(I) = \operatorname{Rad}(I'I'')$ as well. The result now follows from part (a) and the fact that $V(I'I'') = V(I') \cup V(I'')$.

(e) We use induction on the length of the filtration. The case where $k = 1$ is obvious, and part (d) gives the case where $k = 2$. If $k > 2$, we have that $\dim(M) = \max\{\dim(M_{k-1}, M_k/M_{k-1})$ by part (d), and and

$$\dim(M_{k-1}) = \sup\{\dim(M_{i+1}/M_i : 0 \leq i \leq k-2\}$$

by the induction hypotheis. $\square$

Remark. Let $M \neq 0$ be a finitely generated module over an arbitrary ring R. Then M has a filtration

$$0 = M_0 \subset M_1 \subset \cdots \subset M_{k-1} \subset M_k$$

such that every factor M_{i+1}/M_i, where $0 \leq i \leq k-1$, is a cyclic module. In fact if $u_1, \ldots, u_k$ generate M, we may take $M_i = Ru_1 + \cdots + Ru_i$, $0 \leq i \leq k$. If R is Noetherian, we can find such a filtration such that every M_{i+1}/M_i is a prime cyclic module, i.e., has the form R/P_i for some prime ideal I of R. One first chooses u_1 such that $\operatorname{Ann}_R u_1 = P_1$ is prime in R. Let $M_1 = Ru_1 \subseteq M$. Proceeding recursively, suppose that $u_1, \ldots, u_i$ have been chosen in M such that, with $M_j = Ru_1 + \cdots + Ru_j$ for $1 \leq j \leq i$, we have that $M_j/M_{j-1} \cong R/P_j$ with P_j prime. If $M_i = M$ we are done. If not we can choose $u_{i+1} \in M$ such that the annihilator of its image in M/M_i is a prime ideal P_{i+1} of R. Then $M_{i+1}/M_i \cong R/P_{i+1}$: in particular, the inclusion $M_i \subset M_{i+1}$ is strict. The process must terminate, since M has ACC. This means that evenutally we reach M_k such that $M_k = M$. For this type of filtration, it follows from part (e) of the Proposition above that we have

$$\dim(M) = \sup\{\dim(R/P_i) : 1 \leq i \leq k\}.$$

The graded case

This section contains several results that are useful in studying dimension theory in the graded case.

Proposition. *Let M be an $\mathbb{N}$-graded or $\mathbb{Z}$-graded module over an $\mathbb{N}$-graded or $\mathbb{Z}$-graded Noetherian ring S. Then every associated prime of M is homogeneous. Hence, every minimal prime of the support of M is homogeneous and, in particular the associated (hence, the minimal) primes of S are homogeneous.*

Proof. Any associated prime P of M is the annihilator of some element u of M, and then every nonzero multiple of $u \neq 0$ can be thought of as a nonzero element of $S/P \cong Su \subseteq M$, and so has annihilator P as well. Replace u by a nonzero multiple with as few nonzero

homogeneous components as possible. If u_i is a nonzero homogeneous component of u of degree i, its annihilator J_i is easily seen to be a homogeneous ideal of S. If $J_h \neq J_i$ we can choose a form F in one and not the other, and then Fu is nonzero with fewer homgeneous components then u. Thus, the homogeneous ideals J_i are all equal to, say, J, and clearly $J \subseteq P$. Suppose that $s \in P - J$ and subtract off all components of s that are in J, so that no nonzero component is in J. Let $s_a \notin J$ be the lowest degree component of s and u_b be the lowest degree component in u. Then $s_a\, u_b$ is the only term of degree $a+b$ occurring in $su = 0$, and so must be 0. But then $s_a \in \operatorname{Ann}_S u_b = J_b = J$, a contradiction. $\square$

Corollary. *Let K be a field and let R be a finitely generated $\mathbb{N}$-graded K-algebra with $R_0 = K$. Let $\mathcal{M} = \bigoplus_{d=1}^{\infty} R_j$ be the homogeneous maximal ideal of R. Then* $\dim(R) = \operatorname{height}(\mathcal{M}) = \dim(R_{\mathcal{M}})$.

Proof. The dimension of R will be equal to the dimension of R/P for one of the minimal primes P of R. Since P is minimal, it is an associated prime and therefore is homogenous. Hence, $P \subseteq \mathcal{M}$. The domain R/P is finitely generated over K, and therefore its dimension is equal to the height of every maximal ideal including, in particular, $\mathcal{M}/P$. Thus,

$$\dim(R) = \dim(R/P) = \dim\big((R/P)_{\mathcal{M}}\big) \leq \dim R_{\mathcal{M}} \leq \dim(R),$$

and so equality holds throughout, as required. $\square$

Proposition (homogeneous prime avoidance). *Let R be an $\mathbb{N}$-graded algebra, and let I be a homogeneous ideal of R whose homogeneous elements have positive degree. Let $P_1, \ldots, P_k$ be prime ideals of R. Suppose that every homogeneous element $f \in I$ is in $\bigcup_{i=1}^{k} P_i$. Then $I \subseteq P_j$ for some j, $1 \leq j \leq k$.*

Proof. We have that the set H of homogeneous elements of I is contained in $\bigcup_{i=1}^{k} P_k$. If $k = 1$ we can conclude that $I \subseteq P_1$. We use induction on k. Without loss of generality, we may assume that H is not contained in the union of any $k-1$ if the P_j. Hence, for every i there is a homogeous element $g_i \in I$ that is not in any of the P_j for $j \neq i$, and so it must be in P_i. We shall show that if $k > 1$ we have a contradiction. By raising the g_i to suitable positive powers we may assume that they all have the same degree. Then $g_1^{k-1} + g_2 \cdots g_k \in I$ is a homogeneous element of I that is not in any of the P_j: g_1 is not in P_j for $j > 1$ but is in P_1, and $g_2 \cdots g_k$ is in each of $P_2, \ldots, P_k$ but is not in P_1. $\square$

We can now connect the dimension of a module with the degree of its Hilbert polynomial.

Theorem. *Let R be a polynomial ring $K[x_1, \ldots, x_n]$ over a field K, and let M be a finitely generated $\mathbb{Z}$-graded module over R. If M has dimension 0, the Hilbert polynomial of M is 0. If* $\dim(M) > 0$*, the Hilbert polynomial of M has degree* $\dim(M) - 1$.

Proof. M has dimension 0 if and only if it is killed by a power of $m = (x_1, \ldots, x_n)R$, in which case $[M]_d = 0$ for all $d \gg 0$. We use induction on $\dim(M)$.

If $\dim(M) > 0$, then exactly as in the Remark on p. 3 we may construct a finite filtration of M in which all the factors are prime cyclic modules, but using the fact that associated primes of graded modules are graded, we may assume that every R/P_i occurring is graded, i.e., that every P_i is homogeneous. Then the dimension of M is the same as the largest dimension of any R/P_i, and the degree of the Hilbert polynomial is the same as the largest degree of the Hilbert polynomial of any R/P_i. (The Hilbert polynomial of M is the sum of the Hilbert polynomials of the R/P_i. Note that we cannot have cancellation of leading coefficients in the highest degree because the leading coefficient of a Hilbert polynomial is positive: it cannot be negative, since the vector dimension of the space of forms $[R/P_i]_d$ for $d \gg 0$ cannot be negative.)

We have therefore reduced to the case where M has the form R/P, and has positive dimension. It follows that some x_i is not in P, and so there is a form f of degree 1 that is nonzero in the domain R/P. The dimension of $N = M/fM$ must be exactly $\dim(M) - 1$: the dimension must drop because we are killing a nonzero element in a domain, and it cannot drop by more than one, because the rings R/P and $R/(P + fR)$ have the same dimension when localized at their maximal ideals, and we may apply the Corollary at the top of p. 2.

We then have a short exact sequence of graded modules and degree preserving maps:

$$0 \to M(-1) \xrightarrow{f} M \to M/fM \to 0,$$

so that if H_M denotes the Hilbert polynmial of M we have so that

$$(*) \quad H_M(d) - H_M(d-1) = H_{M/fM}(d)$$

for all d. In general, if $P(d)$ is a polynomial in d of degree $k \geq 1$ and with leading coefficient a, the *first difference* $P(d) - P(d-1)$ is a polynomial of degree $k-1$ with leading coefficient ka. Therefore, the degree of the left hand side is $\deg(H_M) - 1$, while the right hand side, by the induction hypothesis, is a polynomial of degree $\dim(M/fM) - 1$ (if $\dim(M) > 1$) or is 0 (if $\dim(M) = 1$). Since $\dim(M/fM) = \dim(M) - 1$, the result follows. $\square$

We saw in the final Theorem of the Lecture Notes of January 22 that F/M and $F/\mathrm{in}(M)$ have the same Hilbert-Poincaré series when M is a graded submodule of a finitely generated free module over a polynomial ring $R = K[x_1, \ldots, x_n]$. Of course, this also means that F/M and $F/\mathrm{in}(M)$ have the same Hilbert function and, hence, the same Hilbert polynomial. We therefore can reduce the problem of finding the Krull dimension of a module to the monomial case:

Theorem. *Let N be any finitely generated $\mathbb{Z}$-graded module over $R = K[x_1, \ldots, x_n]$. Suppose that $u_1, \ldots, u_s$ are finitely many homogeneous generators of respective degrees $d_1, \ldots, d_s$. Think of R^s as $\bigoplus_{j=1}^s R(-d_j)$, and map $R^s \twoheadrightarrow N$ so that $1 \in R(-d_j)$, which has degree d_j, maps to u_j. This map preserves degrees, and the kernel M is an $\mathbb{N}$-graded submodule of R^s.*

Refine the $\mathbb{Z}$-grading on R^s to a $\mathbb{Z}^n$-grading, and choose a monomial order. Then $\dim(N) = \dim\big(F/\mathrm{in}(M)\big)$. $\square$

Since a monomial submodule M of F is a direct sum $I_1e_1 \oplus \cdots I_se_s$, where every I_j is a monomial ideal, we have that $\dim(F/M) = \sup_j\{\dim(R/I_j)\}$. We have therefore reduced the problem of finding the dimension of a module M to that of finding the dimension of R/I when I is a monomial ideal. We can make one more simplification: since $R/\mathrm{Rad}(I)$ and R/I have the same dimension, it suffices to consider the case where I is a radical ideal generated by monomials. Since $(x_{i_1}\cdots x_{i_h})^k$ is a multiple of $x_{i_1}^{a_1}\cdots x_{i_h}^{a_h}$ (here, the a_i are positive integers) whenever $k \geq \sup_j a_j$, the radical of an ideal generated by monomials is generated by square-free monomials. (It is easy to check that any ideal generated by square-free monomials in $K[x_1, \ldots, x_n]$ is, in fact, radical.)

Rings defined by killing square-free monomials and simplicial complexes

By a finite simplicial complex Σ with vertices $x_1, \ldots, x_n$ we mean a set of subsets of $\{x_1, \ldots, x_n\}$ such that

(1) For $1 \leq i \leq n$, $\{x_i\} \in \Sigma$.

(2) Every subset of a set in Σ is also in Σ.

The sets $\sigma \in \Sigma$ are called the *faces*. The *dimension* of σ is one less than its cardinality: the elements of Σ of dimension i are called *i-simplices* of σ. The *dimension* of Σ is the largest dimension of any face. The maximal faces of Σ are called *facets* and these determine Σ: a set is in Σ if and only if it is a subset of a facet of Σ.

If we think of $x_1, \ldots, x_n$ as the points $e_1, \ldots, e_n$ in $\mathbb{R}^n$, where e_i has 1 in the ith spot and 0 elsewhere, we can define *the geometric realization* $|\Sigma|$ of Σ to be the topological space

$$\bigcup_{\sigma\in\Sigma} \text{convex hull}(\sigma)$$

in $\mathbb{R}^n$. The dimension of Σ then coincides with its dimension as a topological space.

Example. If Σ has three vertices x_1, x_2, x_3 and facets $\{x_1, x_2\}$, $\{x_1, x_3\}$, and $\{x_2, x_3\}$, then $|\Sigma|$ is the union of three line segments: it is a triangle, without the interior. On the other hand, if Σ has one facet, $\{x_1, x_2, x_3\}$, then $|\Sigma|$ is a triangle with interior.

Our reason for discussing simplicial complexes at this point is that there is a bijective correspondence between the square-free monomial ideals in $K[x_1, \ldots, x_n]$ that do not contain any of the variables $x_1, \ldots, x_n$ and the simplicial complexes with vertices $x_1, \ldots, x_n$. One may let the ideal I correspond to the subsets of $\{x_1, \ldots, x_n\}$ whose product is *not in* I. Notice that if a monomial ideal does contain one of the variables x_i, the quotient R/I may be thought of as a quotient of a polynomial ring in fewer variables (omitting x_i) by square-free monomials.

The ring R/I_Σ corresponding to simplicial complex Σ is called the *face* ring or *Stanley-Reisner* ring of Σ over K. Here, I_Σ is simply the ideal generated by all square-free monomials such that the set of variables occurring is not a face of Σ.

We leave it as an exercise to verify the minimal primes of R/I_Σ correspond bijectively to the facets of Σ: each minimal prime Q is generated by the images of the elements in $\{x_1, \ldots, x_n\} - \sigma$ for some facet σ, the quotient by Q is the isomorphic to a polynomial ring in the variables that occur in σ. It then follows that $\dim(R/I_\Sigma) = \dim(\Sigma) + 1$.

Elimination theory

We now return to the problem of finding the intersection of an ideal $I \subseteq K[x_1, \ldots, x_n]$ with $K[x_{k+1}, \ldots, x_n]$, which also gives an algorithm for solving a finite system of polynomial equations over an algebraically closed field when there are only finitely many solutions. The method is incredibly simple!

Theorem. *If $g_1, \ldots, g_r$ is a Gröbner basis for I with respect to lexicographic order, then the elements of this basis that lie in $K[x_{k+1}, \ldots, x_n]$ are a Gröbner basis for the ideal $J = I \cap K[x_{k+1}, \ldots, x_n]$.*

Proof. Let $g_{h+1}, \ldots, g_r$ be the elements of the Gröbner basis that lie in $K[x_{k+1}, \ldots, \boldsymbol{x}_n]$ (if $g_1, \ldots, g_r$ are in order of the sizes of their initial terms, these elements will be consecutive and at the end of the sequence).

Consider any element $f \in J$. Then there is a standard expression for f divided by $g_1, \ldots, g_r$, and the remainder will be zero. Say the expression is $f = \sum_{j=1}^n q_j g_j$. Any g_j that involves one of $x_1, \ldots, x_k$ has initial term involving one of the variables $x_1, \ldots, x_k$, and the initial term of $q_j g_j$ will be too large to use in the standard expression unless $q_j = 0$. Therefore, we actually have $f = \sum_{j=h+1}^n q_j g_j$. The same reasoning shows that any q_j for $j > k$ involves only $x_{k+1}, \ldots, x_n$. The initial term of f must be the same, up to a nonzero scalar multiple, as the intitial term of one of the $q_j g_j$, and so it is in the $K[x_{k+1}, \ldots, x_n]$-span of $g_{h+1}, \ldots, g_r$. $\square$

Lecture of January 29

We next want to discuss the notion of a *regular sequence* in a ring or on a module. We are aiming to discuss criteria, using revlex, for a sequence to be regular on F/M. However, we also want to discuss some theorems that we are aiming to prove eventually about the Cohen-Macaulay property for $\mathbb{N}$-graded algebras finitely generated over a field K.

A sequence of elements $f_1, \ldots, f_k \in R$, where R is a ring, is said to a *regular sequence* on the R-module M (when $M = R$, one may refer to a *regular sequence on R* or a *a regular sequence in R* if

(1) $(f_1, \ldots, f_n)M \neq M$,

(2) f_1 is not a zerodivisor on M, i.e., $M \xrightarrow{f_1 \cdot} M$ is injective.

(3) For all i, $1 \leq i \leq k-1$, f_{i+1} is not a zerodivisor on $M/(f_1, \ldots, f_k)M$.

These conditions can be expressed more concisely by allowing $i = 0$ in condition (1), with the interpretation that $(f_1, \ldots, f_i)M = 0$ if $i = 0$.

The empty sequence is regular sequence on every nonzero module M.

Condition (1) is assumed in order to eliminate certain degerate situations. Without it, the sequence 1, 1, , 1, ... , 1 (of any desired length) would be a regular sequence on the 0 module, for example.

Note that $f_1, \ldots, f_h, f_{h+1}, \ldots, f_k$ is a regular sequence on M if and only if $f_1, \ldots, f_h$ is a regular sequence on M and $f_{h+1}, \ldots, f_k$ is a regular sequence on $M/(f_1, \ldots, f_h)M$.

The term *Rees sequence* on M is also used, as well as the term *R-sequence* on M (where "R" may be thought of as standing for "Rees" or "regular"). The term *M-sequence* is also used. We shall always use the term "regular sequence," however.

For example, if $x_1, \ldots, x_n$ are indeterminates, $x_1, \ldots, x_n$ is a regular sequence on $R = K[x_1, \ldots, x_n]$ and on $S = K[[x_1, \ldots, x_n]]$, as well as on any free R-module or free S-module. In fact, we will show that a finitely generated S-module (respectively, a finitely generated $\mathbb{Z}$-graded R-module) M is S-free (respectively, R-free) if and only if $x_1, \ldots, x_n$ is a regular sequence on M.

It is worth noting that, in general, regular sequences are not permutable, even in very well-behaved rings. For example, in the polynomial ring $R = K[x, y, z]$, x, $(1-x)y$, $(1-x)z$ is a regular sequence, but $(1-x)y$, $(1-x)z$, x is not. For the former, modulo xR, the latter two elements become y and z in $K[y, z]$. For the second sequence, $(1-x)z$ is a zerodivisor modulo $(1-x)yR$: the image of y is not 0, but $(1-x)z$ kills the image of y. However, we shall see that regular sequences are permutable in the local case when the module is finitely generated, and in certain graded cases (a precise statement is given below).

Before considering properties of regular sequences further, we want to discuss the local and graded versions of Nakayama's Lemma.

Nakayama's Lemma. *Let R be a ring and let M be an R-module. Suppose that either of the followng two conditions holds:*

(1) *R has a unique maximal ideal m and M is finitely generated.*

(2) *R is $\mathbb{N}$-graded, $m \subseteq R$ consists entirely of elements whose homogeneous components have positive degree, and M is $\mathbb{Z}$-graded, but $[M]_{-d} = 0$ for all $d \gg 0$.*

If $mM = M$ then $M = 0$.

Proof. In case (1) let $u_1, \ldots, u_k$ be a set of generators of M of smallest cardinality. If $k = 0$ then $M = 0$ and we are done. If not, then $u_k \in mM = m(Ru_1 + \cdots + Ru_k) =$

$mu_1 + \cdots + mu_k$, and so $u_k = f_1u_1 + \cdots + f_ku_k$ with every $f_j \in m$. Then $(1 - f_k)u_k = f_1u_1 + \cdots + f_{k-1}u_{k-1}$, and $1 - f_1 \notin m$. It follows that $1 - f_1$ is a unit of R. If $g = (1 - f_k)^{-1}$, then $u_k = gf_1u_1 + \cdots + gf_{k-1}u_{k-1}$, and $u_1, \ldots, u_{k-1}$ generate M, contradicting the minimality of k.

In case (2), let $u \in M$ be a nonzero homogeneous element of smallest possible degree. Then $u \in mM$ implies that u is a sum of elements f_jv_j where the f_j are homogeneous of positive degree and the v_j are homogeneous. Then u is the sum of those nonzero terms f_jv_j such that $\deg(f_j) + \deg(v_j) = \deg(u)$. For those v_j occurring, this implies that $\deg(v_j) = \deg(u) - \deg(f_j) < \deg(u)$, a contradiction. $\square$

Corollary. *Let R be a ring and let M be an R-module. Suppose that either of the followng two conditions holds:*

(1) *R has a unique maximal ideal m and M is finitely generated.*

(2) *R is $\mathbb{N}$-graded, $m \subseteq R$ consists entirely of elements whose homogeneous components have positive degree, and M is $\mathbb{Z}$-graded, but $[M]_{-d} = 0$ for all $d \gg 0$.*

If the images of the elements $\{u_\lambda\}_{\lambda\in\Lambda}$ generate M/mM (and, in case (2), are homogeneous) then the elements $\{u_\lambda\}_{\lambda\in\Lambda}$ generate M.

Proof. Let N be the R-span of $\{u_\lambda\}_{\lambda\in\Lambda}$. In case (2), N and M/N are homogeneous. Since the images of the u_λ span M/mM, we have that $N + mM = M$, and consequently we also have that $(mM + N)/N = M/N$, and this implies that $m(M/N) = M/N$. Thus, by the appropriate case of Nakayama's Lemma, $M/N = 0$, and $M = N$. $\square$

As a consequence of Nakayama's Lemma, we can prove the permutability of regular sequences in local and graded cases.

Proposition (permutability of regular sequences). *Let R be a ring, let M be an R-module, and let $f_1, \ldots, f_k \in R$ be a regular sequence on M. Suppose that either of the following two conditions holds:*

(1) *(R, m, K) is local, $f_1, \ldots, f_k \in m$, and M is finitely generated.*

(2) *R is $\mathbb{N}$-graded, M is $\mathbb{Z}$-graded but $[M]_{-d} = 0$ for all $d \gg 0$, and $f_1, \ldots, f_k$ are homogeneous of positive degree.*

For every permutation π of $1, 2, \ldots, k$, $f_{\pi(1)}, f_{\pi(2)}, \ldots, f_{\pi(k)}$ is a regular sequence on M.

Proof. Because the permutations on $1, 2, \ldots, k$ are generated by transpositions $(i\ i+1)$ of consecutive integers, we need only consider the case where π is such a transposition. We may replace M by $M/(f_1, \ldots, f_{i-1})M$ without affecting any relevant issues. Thus, we may assume without loss of generality that we are simply transposing the first two terms of the regular sequence. But once we have shown that f_2, f_1 is a regular sequence, the

rest is automatic, since $M/(f_1,\, f_2)M = M/(f_2,\, f_1)M$. Therefore, we need only consider the case where $k = 2$ and we are transposing the elements.

We first need to see that f_2 is not a zerodivisor on M. Let $N \subseteq M$ be the annihilator of f_2. (In the graded case, N is graded.) If $u \in N$, then $f_2u = 0$ certainly implies that $f_2u = f_1v$, and so $u = f_1w$ for some $w \in M$. But then $0 = f_2u = f_2f_1w = f_1(f_2w)$, and since f_1 is not a zerodivisor on M, we have that $f_2w = 0$, so that $w \in N$. But we have now shown that if $u \in N$, then $u = f_1w$ with $w \in N$. Thus, $N = f_1N$. By the appropriate form of Nakayama's Lemma, $N = 0$.

Now suppose that $f_1v = f_2u$ where $v,\, u \in M$, so that f_1 kills the image of v in M/f_2M. Then, since f_2 is not a zerodivisor on M/f_1M, we have that $u \in f_1M$, say $u = f_1w$. Then $f_1v = f_2f_1w$ and $f_1(v - f_2w) = 0$. Since f_1 is not a zerodivisor on M, $v = f_2w$. $\square$

Regular local rings

A local ring $(R,\, m,\, K)$ is called *regular* if the Krull dimension of R is equal to the least number of generators of the maximal ideal m. The least number of generators of m is the K-vector space dimension of m/m^2 by Nakayama's Lemma: $\dim_K(m/m^2)$ is called the *embedding dimension* of R. The Krull dimension is the least number of generators of an ideal whose radical is m, and we always have $\dim(R) \leq \dim_K(m/m^2)$.

If $\dim(R) = 0$, R is regular if and only if R is a field.

If $\dim(R) = 1$, then m is generated by one element x, which is not nilpotent. Every nonzero element can be written as a unit times a power of x, since the intersection of the powers of m is 0: simply factor out x as many times as possible. It follows that R is a domain. Thus, the one dimensional regular local rings are precisely the Noetherian discrete valuation rings: we refer to such a ring briefly as a DVR.

Higher dimensional examples include formal power series rings over a field or a DVR.

Note that if R is regular and $x_1, \ldots, x_k$ have images that are linearly indepedent in m/m^2, then $\overline{R} = R/(x_1, \ldots, x_k)R$ is again regular. (Call the maximal ideal in the quotient ring $\overline{m}$. We can extend the sequence to $x_1, \ldots, x_n$, where $n = \dim(R)$, and then the images of the remaining elements $x_{k+1}, \ldots, x_n$ are linearly independent in $\overline{m}/\overline{m}^2$ and are a system of parameters for $\overline{R}$).

We have:

Theorem. *A regular local ring $(R,\, m,\, K)$ is a domain, and a local ring is regular if and only if its maximal ideal m is generated by a regular sequence.*

Proof. We use induction on $\dim(R)$ to prove that R is a domain. Therefore, we may assume that $\dim(R) \geq 2$. Let $x,\, y$ have linearly independent images in m/m^2. It follows that each of the elements $x - y^n$ is prime, for $R/(x - y^n)$ is a regular, and is a domain by

the induction hypothesis. It is easy to see that none of these elements divides any of the others. If $x - y^n$ were a multiple of $x - y^h$ then in $R/(x - y^h)$ the images of $x - y^n$ and $x - y^h$ are both 0, and so $y^n \equiv y^h$. Since $R/(x - y^h)$ is a domain, this forces $y \equiv 0$ (y is in the maximal ideal, and so no power of y can equal 1). But then $(x - y^h)R$ contains y, which is false even modulo m^2. If $uv = 0$ in R, then either u is divisible by infinitely many $x - y^n$ or v is. Suppose u is. But the intersection of ideals generated by prime elements, none of which divides any of the others, is their product. This forces u into arbitrarily high powers of m, and so $u = 0$.

It now follows that if $x_1, \ldots, x_n$ generate m minimally, then $R/(x_1, \ldots, x_k)$ is a domain for every k, and so x_{k+1} is not a zerodivisor modulo $(x_1, \ldots, x_k)R$.

On the other hand, if m is generated by a regular sequence one sees at once that the dimension and embedding dimension of R are the same. □

We can now charactize when a module is free in terms of regular sequences in certain cases. We need Nakayama's Lemma to hold.

Theorem. *Let R be a ring and $M \neq 0$ an R-module. Suppose that one of the following conditions holds:*

(1) *(R, m, K) is regular local, $x_1, \ldots, x_n$ is a regular sequence generating m, and M is finitely generated.*

(2) *$R = K[x_1, \ldots, x_n]$ is a polynomial ring over a field K, and M is $\mathbb{Z}$-graded such that $[M]_{-d} = 0$ for all $d \gg 0$. Then M is free if and only if $x_1, \ldots, x_n$ is a regular sequence on M.*

Proof. In both cases, $x_1, \ldots, x_n$ form a regular sequence on R. If elements form a direct sequence on each module in a family, then they form a regular sequence on the direct sum. Hence, $x_1, \ldots, x_n$ is a regular sequence on any free module.

It remains to show that, under the hypothesis of the Theorem, if $x_1, \ldots, x_n$ form a regular sequence on M then M is free. Choose elements $\{u_\lambda\}_{\lambda \in \Lambda}$ in M whose images are a K-vector space basis for $M/(x_1, \ldots, x_n)M$. Moreover, in case (2) choose these elements to be homogeneous. By the appropriate form of Nakayama's Lemma, they span M. It is therefore sufficient to prove that they are independent over R. We use induction on n. The case $n = 0$ is clear. Assume that $n \geq 1$. It follows that M/x_1M is free on the images of the u_λ over R/x_1R. Consider h elements from this set of generators, say $u_1, \ldots, u_h$, and let $N \subseteq R^h$ be the set of relations on these elements over R. (In the graded case, let u_i have degree s_i and view R_h as $R(-s_1) \oplus \cdots \oplus R(-s_h)$.) In the graded case, N is graded. We can complete the proof by showing that $N = 0$. Now consider any relation $(f_1, \ldots, f_h)$ on $u_1, \ldots, u_h$, so that $f_1u_1 + \cdots + f_hu_h = 0$. Working modulo x_1M (and x_1R), we see that we must have that every f_j is divisible by x_1, say $f_j = x_1g_j$. Then $x_1(g_1u_1 + \cdots g_hu_h) = 0$, and x_1 is not a zerodivisor on M. It follows that $(g_1, \ldots, g_h) \in N$. Thus, $N = x_1N$. By the appropriate form of Nakayama's Lemma, $N = 0$. □

Discussion: homogeneous systems of parameters. Let R be a finitely generated $\mathbb{N}$-graded K-algebra, where $R_0 = K$. Let $m = \bigoplus_{d=1}^{\infty} R_d$ be the homogeneous maximal ideal of R. Since the minimal primes of R are homogeneous, if $\dim(R) > 0$ we can choose a form $F_1 \in m$ such that F_1 is not in any minimal prime of R. Then $\dim(R/F_1R) = \dim(R) - 1$. Now suppose that forms $F_1, \ldots, F_i$ have been chosen such that $\dim\big(R/(F_1, \ldots, F_i)R\big) = \dim(R) - i$. If $i < n = \dim(R)$, we can choose $F_{i+1} \in m$ not in any minimal prime (these are homogeneous) of $(F_1, \ldots, F_i)R$, and it follows that $\dim\big(R/(F_1, \ldots, F_{i+1})\big) = \dim(R) - (i+1)$. Thus, eventually we have a sequence of forms $F_1, \ldots, F_n$ of positive degree such that $\dim\big(R/(F_1, \ldots, F_n)\big) = 0$. Such a sequence of forms is called a *homogeous system of parameters* for R.

Theorem. *Let R be a finitely generated $\mathbb{N}$-graded K-algebra with $R_0 = K$ such that $\dim(R) = n$. A homogeneous system of parameters $F_1, \ldots, F_n$ for R always exists. Moreover, if $F_1, \ldots, F_n$ is a sequence of homogeneous elements of positive degree, then the following statements are equivalent.*

(1) *$F_1, \ldots, F_n$ is a homogeneous system of parameters.*

(2) *m is nilpotent modulo $(F_1, \ldots, F_n)R$.*

(3) *$R/(F_1, \ldots, F_n)R$ is finite-dimensional as a K-vector space.*

(4) *R is module-finite over the subring $K[F_1, \ldots, F_n]$.*

Moreover, when these conditions hold, $F_1, \ldots, F_n$ are algebraically independent over K, so that $K[F_1, \ldots, F_n]$ is a polynomial ring.

Proof. We have already shown existence.

(1) $\Rightarrow$ (2). If $F_1, \ldots, F_n$ is a homogeneous system of parameters, we have that $\dim\big(R/F_1, \ldots, F_n)\big) = 0$. We then know that all prime ideals are maximal. But we also know that the maximal ideals are also minimal primes, and so must be homogeneous. Since there is only one homogenous maximal ideal, it must be $m/(F_1, \ldots, F_n)R$, and so m is nilpotent on $(F_1, \ldots, F_n)R$.

(2) $\Rightarrow$ (3). If m is nilpotent modulo $(F_1, \ldots, F_n)R$, then the homogeneous maximal ideal of $\overline{R} = R/(F_1, \ldots, F_n)R$ is nilpotent, and it follows that $[\overline{R}]_d = 0$ for all $d \gg 0$. Since each $\overline{R}_d$ is a finite dimensional vector space over K, it follows that $\overline{R}$ itself is finite-dimensional as a K-vector space.

(3) $\Rightarrow$ (4). This is immediate from the homogeneous form of Nakayama's Lemma: a finite set of homogeneous elements of R whose images in $\overline{R}$ are a K-vector space basis will span R over $K[F_1, \ldots, F_n]$, since the homogenous maximal ideal of $K[F_1, \ldots, F_n]$ is generated by $F_1, \ldots, F_n$.

(4) $\Rightarrow$ (1). If R is module-finite over $K[F_1, \ldots, F_n]$, this is preserved mod $(F_1, \ldots, F_n)$, so that $R/(F_1, \ldots, F_n)$ is module-finite over K, and therefore zero-dimensional as a ring.

Finally, when R is a module-finite extension of $K[F_1, \ldots, F_n]$, the two rings have the same dimension. Since $K[F_1, \ldots, F_n]$ has dimension n, the elements $F_1, \ldots, F_n$ must be algebraically independent. $\square$

Discussion: making a transition from one system of parameters to another. Let R be a Noetherian ring of Krull dimension n, and assume that either

(1) (R, m, K) is local and $f_1, \ldots, f_n$ and $g_1, \ldots, g_n$ are two systems of parameters.

(2) R is finitely generated $\mathbb{N}$-graded over $R_0 = K$, a field, m is the homogeneous maximal ideal, and $f_1, \ldots, f_n$ and $g_1, \ldots, g_n$ are two homogeneous systems of parameters for R.

We want to observe that in this situation there is a finite sequence of systems of parameters (respectively, homogeneous systems of parameters in case (2)) starting with $f_1, \ldots, f_n$ and ending with $g_1, \ldots, g_n$ such that any two consecutive elements of the sequence agree in all but one element (e.g., after reordering, only the i th terms are possibly different for a single value of i, $1 \leq i \leq n$. We can see this by induction on n. If $n = 1$ there is nothing to prove. If $n > 1$, first note that we can choose h (homogeneous of positive degree in the graded case) so as to avoid all minimal primes of $(f_2, \ldots, f_n)R$ and all minimal primes of $(g_2, \ldots, g_n)R$. Then it suffices to get a sequence from $h, f_2, \ldots, f_n$ to $h, g_2, \ldots, g_n$, since the former differs from $f_1, \ldots, f_n$ in only one term and the latter differs from $g_1, \ldots, g_n$ in only one term. But this problem can be solved by working in R/hR and getting a sequence from the images of $f_2, \ldots, f_n$ to the images of $g_2, \ldots, g_n$, which we can do by the induction hypothesis. We lift all of the systems of parameters back to R by taking, for each one, h and inverse images of the elements in the sequence in R (taking a homogeneous inverse image in the graded case), and always taking the same inverse image for each element of R/hR that occurs. $\square$

Cohen-Macaulay rings were discussed in the first lecture. But we are now in a position to prove several of the assertions made there.

Theorem. *Let R be a finitely generated graded algebra over $R_0 = K$. The following conditions are equivalent.*

(1) *Some homogeneous system of parameters is a regular sequence.*

(2) *Every homogeneous system of parameters is a regular sequence.*

(3) *For some homogeneous system of parameters $F_1, \ldots, F_n$, R is a free-module over $K[F_1, \ldots, F_n]$.*

(4) *For every homogeneous system of parameters $F_1, \ldots, F_n$, R is a free-module over $K[F_1, \ldots, F_n]$.*

Proof. We first show that (1) and (2) are equivalent. We want to show that if one homogeneous system of parameters is a regular sequence, then every homogeneous system of parameters is a regular sequence. By the Discussion above, we may assume that they agree except possibly in one term. Since regular sequences are permutable (and systems of parameters are obviously permutable), we may assume that they agree except possibly for the last term. Call them $F_1, \ldots, F_n$ and $F_1, \ldots, F_{n-1}, G$. The issue is whether the last term is a nonzerodivisor modulo the earlier terms. Therefore, we may pass to

$\overline{R} = R/(F_1, \ldots, F_{n-1})$, which is one-dimensional. It follows that we may assume that R is one-dimensional, and we need only show that if F, G both generate ideals whose radical is m and F is a nonzerodivisor, then G is a nonzerodivisor. But F has a power in GR, say $F^k = GH$. If G is a zerodivisor, it follows that F^k is as well, and then F must be a zerodivisor. This proves the equivalence of (1) and (2). The preceding Theorem yields the equivalence of (1) and (3), as well as the equivalence of (2) and (4), immediately. □

As mentioned earlier, we shall say that R is *Cohen-Macaulay* of these equivalent conditions hold. The same argument as given in the proof just above also shows:

Theorem. *Let (R, m, K) be a local ring. Then one system of parameters is a regular sequence if and only if every system of parameters is a regular sequence.* □

We shall say that the local ring R is *Cohen-Macaulay* if every system of parameters is a regular sequence. Of course, regular rings are Cohen-Macaulay. We shall later show that an $\mathbb{N}$-graded ring over $R_0 = K$ is Cohen-Macaulay if and only if all of its local rings are Cohen-Macaulay.

We shall eventually prove two substantial results about when rings are Cohen-Macaulay. One of them is Reisner's criterion for when the face ring of a finite simplicial complex is Cohen-Macaulay. The other concerns the Cohen-Macaulay property for certain rings of invariants of matrix groups acting on polynomial rings.

To state Reisner's criterion, we need the notion of *link* in a simplicial complex Σ. If x is a vertex of Σ, we define the *link* of x in Σ to be the simplicial complex Λ such that $\tau \in \Lambda$ if and only if $\tau \in \Sigma$, $x \notin \tau$, and $\{x\} \cup \tau \in \Sigma$.

For example, suppose that Σ corresponds to the triangulation of a convex pentagon obtained by connecting an interior point to the vertices, and x is the interior point. If the vertices on the perimeter are $x_1, \ , x_2, x_3, x_4, x_5$, then the facets of Σ are the five 2-simples $\{x, x_i, x_{i+1}\}$, for $1 \leq i \leq 5$, where x_{i+1} is to be interpreted as x_1 when $i = 5$ (i.e., the subscripts are read modulo 5). The link of x is the perimeter of the pentagon: its facets are the five 1-simplices (or *edges*) $\{x_i, x_{i+1}\}$, where $1 \leq i \leq 5$.

If we take Σ to have facets $\{x_1, x_3\}$, $\{x_2, x_3\}$, and $\{x_3, x_4, x_5\}$ (the geometric realization consists of a triangle with interior and two additional line segments jutting out from one vertex), then the link of x_3 has facets $\{x_1\}$, $\{x_2\}$, and $\{x_4, x_5\}$: a line segment with two additional isolated points.

Once one has a link, one can treat it as a new simplicial complex, and take the link of one of its vertices. This may be iterated several times. But these iterated links can be obtained in a single step as follows. If $\sigma_0 \in \Sigma$, define the *link* of $\sigma_0 \in \Sigma$ as the simplical complex $\{\tau \in \Sigma : \tau \cap \sigma_0 = \emptyset \text{ and } \tau \cup \sigma_0 \in \Sigma\}$. One gets the same simplicial complex by iterating the operation of taking links of vertices, using all vetices in σ_0: the iterated link obtained is independent of the order in which one takes links of vertices.

We also recall that the *reduced simplicial homology* of Σ over K is the the same as the simplicial homology over K, except in dimension 0, where it has K-vector space dimension one smaller. (Thus $\widetilde{H}_0(X; K) = 0$ if and only if Σ is connected.)

We can now state:

Theorem (Reisner). *Let K be a field, let Σ be a finite simplical complex with vertices $x_1, \ldots, x_n$, and let I_Σ be the ideal of $R = K[x_1, \ldots, x_n]$ generated by the square free monomials such that the set of variables that occur is not a face of Σ. Then R/I_Σ is Cohen-Macaulay if and only if both of the following conditions hold:*

(1) *The reduced simplicial homology $\widetilde{H}_i(\Sigma; K)$ with coefficients in K vanishes, $0 \leq i \leq \dim(\Sigma) - 1$.*

(2) *For every link Λ, the reduced simplicial homology $\widetilde{H}_i(\Lambda; K) = 0$, $0 \leq i \leq \dim(\Lambda) - 1$.*

We defer the proof. We also note that by a result of Munkres, Reisner's condition is actually a topological property of $|\Sigma|$.

Note that in dimension 0, every finite simplicial complex is Cohen-Macaulay. In dimension 1, Σ is Cohen-Macaulay if and only if it is connected.

In dimension 2, a triangulation of a sphere gives a Cohen-Macaulay ring, a triangulation of a cylinder does not, while what happens with a triangulation of a real projective plane depends on the characteristic. In characteristic 2, the first homology group of the the real projective plane does not vanish, and the ring one gets is not Cohen-Macaulay. In all other characteristics, the ring is Cohen-Macaulay.

Finally, we mention one more Theorem. Let G be a Zariski closed subgroup of $\mathrm{GL}(n, K)$: thus, G is a group of matrices. Suppose that G is linearly reductive, by which we mean that every (algebraic) representation is completely reducible. There are many such groups in characteristic 0: the general and special linear groups, the orthogonal group, and the symplectic group are examples, as well as finite groups, the multiplicative group of the field, and products of the groups already mentioned In characteristic $p > 0$, there are relatively few such groups: products of copies of the multiplicative group of the field and finite groups whose order is not divisibile by p are the main examples .

Then G may be thought of as acting on the space of forms of degree 1 in $K[x_1, \ldots, x_n]$, and the action extends to an action on the polynomial ring R itself. One may form the ring of invariants $R^G = \{f \in R : \gamma(f) = f \text{ for all } \gamma \in G\}$. When G is linearly reductive, this group turns out to be finitely generated. Beyond that:

Theorem. *With hypothesis as in the paragraph above, R^G is a Cohen-Macaulay ring.*

Lecture of February 1

Invariant Theory

We want to present some examples from classical invariant theory to which one can apply the Theorem on the Cohen-Macaulay property for rings of invariants stated at the end of the Lecture Notes of January 29, as well as a strong form, stated below.

For simplicity, in this discussion we assume that we are working over an algebraically closed field K when describing what is meant by a linear algebraic group and an action of such a group: this minimizes prerequisites from algebraic geometry. However, the statements identifying the rings of invariants of various group actions are all valid over any infinite field, and the statements about rings being Cohen-Macaulay are valid over any field. In fact, we note the following result:

Proposition. *If R is a finitely generated graded K-algebra over a field K with $R_0 = K$, then R is Cohen-Macaulay if and only if $L \otimes_K R$ is Cohen-Macaulay.*

The proof is left as an exercise: see problem 4(d). of Problem Set #2.

Next note that if $X \subseteq \mathbb{A}^s_K$ is a closed algebraic set and $f \in K[X]$ is a regular function on X, the open subset $X_f = X - V(f)$ has the structure of a closed algebraic set embedded in $\mathbb{A}^{s+1}_K$: if $X = V(I)$, then X_f is in bijective correspondence with $V(I, fx_{s+1} - 1) \subseteq \mathbb{A}^{s+1}_K$. The coordinate ring of X_f is easily shown to be $K[X]_f$, and the inclusion $X_f \subseteq X$ corresponds to the natural K-algebra homomorphism $K[X] \to K[X]_f$.

Therefore, if we identify $n \times n$ matrices over K with $\mathbb{A}^{n^2}$ and D denotes the determinant function, $\mathrm{GL}(n, K)$ may be identified with $\mathbb{A}^{n^2}_D$, and so has the structure of a closed algebraic set. For any finite-dimensional vector space V over K, by choosing a basis we may identify the group $\mathrm{GL}_K(V)$ of K-linear automorphisms of V with $\mathrm{GL}(r, K)$, where $r = \dim(V)$, and so $\mathrm{GL}(V)$ acquires the structure of an algebraic set. Since conjugation by a fixed invertible $r \times r$ matrix is an automorphism of $\mathrm{GL}(r, K)$ as an algebraic set, the algebraic set structure on $GL(V)$ is independent of the choice of the K-vector space basis for V.

By a *representation* of the linear algebraic group G we mean a group homomorphism $G \to \mathrm{GL}_K(V)$ that is also a K-regular map of closed algebraic sets. The representation evidently gives an action of G on V, and may also be described by giving a K-regular map $G \times V \to V$ satisfying the conditions for a group action. A representation is called *irreducible* if no proper nonzero subspace W of V is stable under the action of V.

As was mentioned in the Lecture of January 29, G is called *linearly reductive* if every representation is *completely reducible*, which means that it is a direct sum of irreducible

representations. As was also mentioned in that lecture, the general linear group and the special linear group are examples in characteristic 0. The multiplciative group of the field is $\mathrm{GL}(1, K)$: finite products of copies of the multiplicative group of the field (such groups are called *algebraic tori*) are examples in all characteristics.

In fact, one has the following more general statement:

Theorem. *Let G be a linearly reductive linear algebraic group over K acting on the vector space of forms of degree one in the polynomial ring $R = K[x_1, \ldots, x_n]$. The action extends uniquely to an action of G on R by degree-preserving K-algebra automorphisms. For this action, the ring of invariants R^G is Cohen-Macaulay.*

The proof is deferred for a while. One of the surprising aspects of this Theorem is that the most interesting examples are in characteristic 0, but the first proof [M. Hochster and J. L. Roberts, *Rings of invariants of reductive groups acting on regular rings are Cohen-Macaulay*, Advances in Math. **13** (1974) 115–175] of the result and, by far, the simplest proof [M. Hochster and C. Huneke, *Tight closure, invariant theory, and the Briançon-Skoda theorem*, Amer. J. Math. **3** (1990) 31–116] use reduction to characteristic $p > 0$.

In classical invariant theory there were two fundamental problems. The first was to determine generators for the ring of invariants of a group action. The second was to give generators for the ideal of relations on these generating invariants. See [Hermann Weyl, *The Classical Groups*, Princeton Univ. Press, Princeton, 1946] for the solution of several important problems of this type. In the light of the Theorem above, the rings of invariants studied classically provide many interesting examples of Cohen-Macaulay rings.

We want to consider some of these examples. We first introduce two notations. If X is a matrix with entries in a K-algebra R, we denote by $I_t(X)$ the ideal of R generated by the $t \times t$ minors (determinants of $t \times t$ submatrices) of X, and by $K[X]$ the K-subalgebra of R generated by the entries of X. More generally, we denote by $K[X/t]$ the K-subalgebra of R generated by the $t \times t$ minors of X.

In the three examples just below, the field is assumed *to be infinite*.

First example. Let $G = K - \{0\} \cong GL(1, K)$ act on the polynomial ring

$$R = K[x_1, \ldots, x_m, y_1, \ldots, y_n]$$

in $m + n$ variables so that if $a \in G$, $x_i \mapsto x_j a^{-1}$ and $y_j \mapsto a y_j$ for all i and j. It is easy to verify that the ring of invariants is

$$K[x_i y_j : 1 \leq i \leq m,\ 1 \leq j \leq n].$$

(It is certainly clear that these elements are invariant: $x_i a^{-1} a y_j = x_i y_j$.) If $U = \big(u_{ij}\big)$ is an $m \times n$ matrix of new indeterminates, we can map $K[U] \twoheadrightarrow R^G = K[x_i y_j : i, j]$ as K-algebras by sending $u_{ij} \mapsto x_i y_j$. Note that

$$(x_i y_j)(x_{i'} y_{j'}) = (x_i y_{j'})(x_{i'} y_j),$$

which shows that $I_2(U)$ is in the kernel. In fact, $R^G \cong K[U]/I_2(U)$. This ring is Cohen-Macaulay in all characteristics.

Second example. We can generalize the preceding example as follows. Let t, m, n be positive integers with $t \leq \min\{m, n\}$, let $X = \left(x_{ij}\right)$ be an $m \times t$ matrix of indeterminates over K, and let $Y = \left(y_{jk}\right)$ be a $t \times n$ matrix of indeterminates over K. Let $G = \mathrm{GL}(t, K)$ act on $K[X, Y]$ as follows: if $A \in G$, A acts by sending the entries of X to the entries of XA^{-1} and the entries of Y to the entries of AY. The preceding example is the case where $t = 1$. It is proved, for example, in Weyl's book that the ring of invariants is generated by the entries of the $m \times n$ product matrix XY. These entries are the scalar products of the various rows of X with the various columns of Y. It is clear that then entries of XY *are* invariant, because $(XA^{-1})(AY) = XY$. Again, one can map $K[U] \twoheadrightarrow K[XY] = R^G$ as K-algebras, where U is an $m \times n$ matrix of new indetermnates, and it is easy to show that the ideal generated by the $(t+1) \times (t+1)$ size minors of U is in the kernel. It turns out that, in fact, $R^G = K[XY] \cong K[U]/I_{t+1}(U)$. The Theorem above then implies that $K[U]/I_{t+1}(U)$ is Cohen-Macaulay in characteristic 0. (This is true in all characteristics: see [M. Hochster and J. A. Eagon, *Cohen-Macaulay rings, invariant theory, and the generic perfection of determinantal loci*, Amer. J. Math. **93** (1971) 1020–1058].)

Third example. Let X be an $n \times s$ matrix of indeterminates over the field K, where $1 \leq n \leq s$, and let $G = \mathrm{SL}(n, K)$ act on on $K[X]$ be sending the entries of X to the corresponding entries of AX. Note that if C denotes any column of X, the entries of C are sent to the corresponding entries of AC. It follows that if Y is any $n \times k$ submatrix of X (so that Y consists of a set of columns of X), then the entries of Y are sent to the corresponding entries of AY. Consequently, if Y is any $n \times n$ submatrix of X, then $\det(AY) = \det(A)\det(Y) = \det(Y)$, since the elements of $\mathrm{SL}(n, K)$ are precisely the $n \times n$ matrics with determinant 1. In this case $R^G = K[X/n]$, the ring generated over K by the $\binom{s}{n}$ $n \times n$ minors of X, the so-called *maximal* minors of X. The relations on the minors are generated by certain standard quadratic relations called the *Plücker relations.*[1]

By the Theorem above, these rings $K[X]^G = K[X/n]$ are Cohen-Macaulay in characteristic 0. (This is also true in characteristic $p > 0$: see for example, [M. Hochster,

[1]These rings are well-known in algebraic geometry: the set of n-dimensional vector subspaces of K^s has the structure of a projective algebraic variety, which can be embedded in a projective space over K of dimension $\binom{s}{n} - 1$. The idea is that given a subspace V, one can choose an $s \times n$ matrix M whose rows are a basis for V: the $\binom{s}{n}$ minors of this matrix do not all vanish, and satisfy the Plücker relations. Therefore they give a point in the algebraic set G defined by the Plücker relations. G turns out to be irreducible. If one changes the matrix, the new matrix can be gotten from M by multiplying on the left by an invertible $n \times n$ matrix A: each of the $n \times n$ minors of AM is the product of $\det(A)$ with the corresponding minor of M, and so one gets the same point in projective space no matter which matrix whose rows are a basis for V is chosen. It can be shown that every point of G can be obtained in this way from a unique subspace of K^s of dimension n, so that this gives a bijective correspondence between the projective variety G and the set of n-dimensional vector subspaces of K^s. The projective variety V is called the *Grassmann variety* or *Grassmannian.*

Grassmannians and their Schubert subvarieties are arithmetically Cohen-Macaulay, J. of Algebra **25** (1973) 40–57]. They are also known to be unique factorization domains.

We shall also deduce from the Theorem stated above that an integrally closed ring that is a subring of $Kx_1, \ldots, x_n]$ generated by monomials is Cohen-Macaulay. In general, normality is far from sufficient for the Cohen-Macaulay property. The proof we give will depend on showing that any such ring is isomorphic with a ring of invariants of an algebraic torus, i.e., a product of copies of $\mathrm{GL}(1, K)$, acting on a polynomial ring. Cf. [M. Hochster, *Rings of invariants of tori, Cohen-Macaulay rings generated by monomials, and polytopes*, Annals of Math. **96** (1972) 318–337].

Monomial submodules and the colon operation

Our next objective is to use revlex to give a criterion for when a sequence of indeterminates is a regular sequence on a module. We need some preliminaries concerning the behavior of monomial submodules and the colon operation.

If $M \subseteq F$ are any two R-modules and J is an ideal of R, we define

$$M :_F J = \{f \in M : Jf \subseteq M\}.$$

When $J = uR$ is the principal, we may write $M :_F u$ instead of $M :_F uR$.

When u is a nonzerodivisor (we shall typically be in this situation, for u will almost always be a nonzero element of a polynomial ring in the sequel), we have the following:

$$(*) \quad u(M :_F u) = M \cap uF \text{ and so } M :_F u = \frac{1}{u}(M \cap uF).$$

In fact, $uf \in M$ means precisely that $f \in M :_F u$, and then $f = \dfrac{1}{u}uf$ is uniquely determined.

We proved early that for monomial submodules and ideals, intersection distributes over sum. Hence $(*)$ yields:

Proposition. *Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over the field K, and F a fintely generated free module. Let $M_1, \ldots, M_k$ be monomial submodules of F, and let $\mu \in R$ be a monomial. Then*

$$(M_1 + \cdots + M_k) :_F \mu = (M_1 :_F \mu) + \cdots + (M_k :_F \mu). \qquad \square$$

This gives a very easy way of calculating $M :_F \mu$ when M is a monomial module. If $\nu_j e_{i_j}$ is a typical generator, M is the sum of the modules $\nu_j e_{i_j} R$. It follows that $M :_F \mu$ is the sum of the modules $\nu_j e_{i_j} R :_F \mu$. Each of these is simply $(\nu_j R :_R \mu)e_{i_j}$. Thus, we

have reduced to calculating $\nu R :_R \mu$ when ν and μ are monomials in R. Each of these is a cyclic module generated by one monomial, namely $\nu/\mathrm{GCD}(\mu, \nu)$.

An alternative description is as follows: if $a, b \in \mathbb{N}$, let $a \dot{-} b = \max\{a - b, 0\}$, and if $\alpha = (a_1, \ldots, a_n)$ and $\beta = (b_1, \ldots, b_n) \in \mathbb{N}^n$, let $\gamma = (a_1 \dot{-} b_1, \ldots, a_n \dot{-} b_n)$. Then $x^\alpha R : x^\beta = x^\gamma R$.

It is quite easy to see that a monomial μ is a nonzero divisor on F/M, where M is monomial, if and only if the variables occurring in μ do not occur in any minimal generator of M. This implies that $\mu_1, \ldots, \mu_h$ is a regular sequence on F/M if and only the variables occurring in μ_i do not occur in any other μ_j nor in any minimal generator of M.

We shall next aim to show that for reverse lexicographic order on F, if $M \subseteq F$ is graded, $x_{k+1}, \ldots, x_n$ is a regular sequence on F/M if and only if it is a regular sequence on $F/\mathrm{in}(M)$.

Lecture of February 3

Regular sequences in the monomial case

We want to analyze what it means for a sequence of monomials $\mu_1, \ldots, \mu_k$ in $R = K[x_1, \ldots, x_n]$ to be a regular sequence on F/M when F is a finitely generated free R-module and M is a monomial submodule of F.

First note that, quite generally, $f \in R$ is not a zerodivisor on Q/N if and only if $N :_Q f = N$. This says precisely that $fu \in N$ if and only if $u \in N$. This yields:

Proposition. *Let $R = K[x_1, \ldots, x_n]$ and let $\mu_1, \ldots, \mu_k$ be a sequence of monomials in R. Let M be a monomial submodule of the finitely generated free module F. Then $\mu_1, \ldots, \mu_k$ is a regular sequence on F/M if and only if no variable that occurs in μ_i occurs in another μ_j, nor in any of the minimal monomial generators of M.*

Proof. Since $M = I_1 e_1 \oplus \cdots I_s e_s$ where the I_j are monomial ideals, we reduce at once to the case where $M = I$ is a monomial ideal: call the minimal monomial generators $\nu_1, \ldots, \nu_h$. We use induction on k. If $k = 1$, note that if μ_1 shares a variable x_t with ν_i then $\nu_i :_R \mu_1$ is generated by a monomial that divides ν_i and has a smaller exponent on x_t then ν_i does. This element is not in I, by the minimality of ν_i, but is in $I : \mu_1$. Hence the condition that μ_1 not involve a variable occurring in any ν_i is necessary. On the other hand, if that is true then $\nu_i :_R \mu = \nu_i R$ for every i, $1 \leq i \leq h$, and since colon distributes over sum we have that

$$I :_R \mu_1 = (\sum_{i=1}^h \nu_i R) :_R \mu_1 = \sum_{i=1}^h (\nu_i R :_R \mu_1) = \sum_{i=1}^h \nu_i R = I,$$

as required. Moreover it is clear that $\nu_1, \ldots, \nu_h, \mu_1$ are minimal generators for $I + \mu_1 R$. The inductive step is then an application of the case where $k = 1$. $\square$

Compatible orders and a sufficient condition for regularity of a sequence

Given a polynomial ring $K[x_1, \ldots, x_n]$ over a field K and a monomial order $>$ on a finitely generated R-free module F with ordered free basis $e_1, \ldots, e_s$, recall that for every t, $1 \leq t \leq s$, there is a monomial order $>_t$ on R defined by the condition $\mu > \mu'$ precisely if $\mu e_t > \mu' e_t$. Moreover, if $g \in R - \{0\}$ and $f \in F - \{0\}$ are such that $\mathrm{in}(f)$ involves e_t, then

$$(\dagger) \quad \mathrm{in}(gf) = \mathrm{in}_{>_t}(g)\mathrm{in}(f).$$

See the second page of the Lecture Notes of January 20. We shall say that a monomial order $>_R$ on R is *compatible* with a given monomial order $>$ on F if all of the orders $>_t$ are the same, and agree with $>_R$. It follows at once that if $>_R$ and $>$ on F are compatible, then for all $g \in R - \{0\}$ and $f \in F - \{0\}$,

$$(\dagger\dagger) \quad \mathrm{in}(fg) = \mathrm{in}_{>_R}(g)\mathrm{in}(f).$$

In fact, condition $(\dagger\dagger)$ is easily seen to be equivalent to compatibility. In working with compatible monomial orders, we typically use the same symbol $>$ for both.

If two of the $>_t$ are distinct, which can happen, there is no compatible order on R. If there is a compatible order on R, it is unique. The standard method of extending a monomial order on R to a monomial order on F (i.e., $\mu e_i > \mu' e_j$ if $\mu > \mu'$ or $\mu = \mu'$ and $i < j$) always produces a monomial order on F with which the original monomial order is compatible. In particular, revlex on F is compatible with revlex on R. In the sequel, when F is graded so that its generators do not necessarily all have degree 0, we give a slightly different way of extending revlex to F — but it is still compatible with revlex on R.

We next observe the following sufficient (but not necessary) condition for elements of R to be a regular sequence on F/M. Notice that we are not assuming that M is graded, nor that $>$ is revlex.

Theorem. *Let $R = K[x_1, \ldots, x_n]$, $f_1, \ldots, f_k \in R$ and let M be any submodule of a finitely generated free R-module F. Suppose that we have compatible monomial orders on R and F. If $\mathrm{in}(f_1), \ldots, \mathrm{in}(f_k)$ form a regular sequence on $M/\mathrm{in}(M)$, then $f_1, \ldots, f_k$ is a regular sequence on F/M and, for $1 \leq i \leq k$, $\mathrm{in}\big(M + (f_1, \ldots, f_i)F\big) = \mathrm{in}(M) + \big(\mathrm{in}(f_1), \ldots, \mathrm{in}(f_i)\big)F$.*

Proof. We use induction on k, and we consequently can reduce at once to the case where $k = 1$. We write f for f_1, and we must show that if $\mathrm{in}(f)$ is a not a zerodivisor on $F/\mathrm{in}(M)$ then (1) f is not a zerodivisor on F/M and (2) $\mathrm{in}(M + fM) = \mathrm{in}(M) + \mathrm{in}(f)F$.

If (1) fails we have $fu \in v \in M$ with $u \notin M$, and we can choose such an example with $\mathrm{in}(u)$ minimum, since the monomial order on F is a well-ordering. By the compatibility of orders, $\mathrm{in}(fu) = \mathrm{in}(f)\mathrm{in}(u) = \mathrm{in}(v) \in \mathrm{in}(M)$, and since $\mathrm{in}(f)$ is not a zerodivisor on $\mathrm{in}(M)$, we have that $\mathrm{in}(u) \in \mathrm{in}(M)$, so that we can choose $u' \in M$ with $\mathrm{in}(u) = \mathrm{in}(u')$. Then fu and fu' are both in M, and so $f(u - u') \in M$. But the initial terms of u and u'

cancel, so that $u = u'$ or $\mathrm{in}(u - u') < \mathrm{in}(u)$. The latter contradicts the minimality of the choice of u, and the former shows that $u \in M$.

To prove (2), note that $\mathrm{in}(M) + \mathrm{in}(f)F \subseteq \mathrm{in}(M + fF)$ is obvious, and so we need only prove the opposite inclusion. If it fails, we can choose $u + fv \in M + fF$ where $u \in M$, $v \in F$, such that $\mathrm{in}(u + fv) \notin \mathrm{in}(M) + \mathrm{in}(f)F$, and, again, we can make this choice so that $\mathrm{in}(v)$ is minimum (note that v cannot be 0). We consider two cases.

First case: $\mathrm{in}(fv) \in \mathrm{in}(M)$. Then $\mathrm{in}(f)\mathrm{in}(v) \in \mathrm{in}(M)$ and, since $\mathrm{in}(f)$ is not a zerodivisor on $\mathrm{in}(M)$, we have that $\mathrm{in}(v) \in \mathrm{in}(M)$ and we can choose $v' \in M$ such that $\mathrm{in}(v) = \mathrm{in}(v')$. Then $u + fv = (u + fv') + f(v - v')$ still has initial form not in $M + fV$, and we have $u + fv' \in M$ while $v - v'$ has smaller initial form than v, a contradiction.

Second case: $\mathrm{in}(fv) \notin \mathrm{in}(M)$. In this case, $\mathrm{in}(fv)$ and $\mathrm{in}(u) \in \mathrm{in}(M)$ cannot cancel, and so one of them must be $\mathrm{in}(u + fv)$. But then either $\mathrm{in}(u + fv) = \mathrm{in}(u) \in \mathrm{in}(M)$ or $\mathrm{in}(u + fv) = \mathrm{in}(fv) = \mathrm{in}(f)\mathrm{in}(v) \in \mathrm{in}(f)F$, as required. $\square$

Special properties of reverse lexicographic order and a converse result

Throughout this section, $R = K[x_1, \ldots, x_n]$ is a polynomial ring over K considered with reverse lexicographic order, F is a finitely generated graded free R-module with ordered free homogeneous basis $e_1, \ldots, e_s$, also with reverse lexicographic order, which we define as follows. In the graded case we still want revlex to define total degree. Therfore, we define $\mu e_i >_{\mathrm{revlex}} \mu' e_j$ to mean either that (1) $\deg(\mu e_i) > \deg(\mu' e_j)$ or (2) $\deg(\mu e_i) = \deg(\mu' e_j)$ and $\mu < \mu'$ in lexicographic order for the variables ordered so that

$$x_n > x_{n-1} > \cdots > x_2 > x_1,$$

or (3) $\deg(\mu e_i) = \deg(\mu' e_j)$, $\mu = \mu'$, and $i < j$.

Let M be a graded submodule of F. We already noted at the end of the Lecture of February 1 that $x_{k+1}, \ldots, x_n$ is a regular sequence on F/M if and only if $x_{k+1}, \ldots, x_n$ is a regular sequence on $F/\mathrm{in}(M)$, which we know is equivalent to the condition that no minimal monomial generator of $\mathrm{in}(M)$ invovles any of the variables $x_{k+1}, \ldots, x_n$. The preceding Theorem already shows that the condition is sufficient. We next want to prove that it is necessary as well. The following very easy result is a key fact about revlex that we shall use repeatedly.

Lemma. *Let notation be as above and let $u \in F - \{0\}$ be a homogeneous element. Then for every positive integer h, x_n^h divides u if and only if x_n^h divides* $\mathrm{in}(u)$.

Proof. "Only if" is obvious. The "if" part is immediate from the definition: since all terms have the same degree, any term not divisible by x_n^h is strictly larger than any term divisible by x_n^h. $\square$

Proposition. *Let notation be as above, with $M \subseteq F$ graded, and let $g_1, \ldots, g_r$ be a Gröbner basis for M consisting of homogeneous elements. Let k be a positive integer.*

(a) $\text{in}(M + x_n^h F) = \text{in}(M) + x_n^h F$, *and* $g_1, \ldots, g_r, x_n^k e_1, \ldots, x_n^k e_s$ *is a Gröbner basis for* $M + x_n^h F$.

(b) $\text{in}(M :_F x_n^h) = \text{in}(M) :_F x_n^h$. *Moreover, if for* $1 \leq j \leq r$, t_j *denotes the greatest integer in the interval* $[0, h]$ *such that* $x^{t_j} | g_j$ *and* $h_j = g_j / x_n^{t_j}$, *then* $h_1, \ldots, h_r$ *is a Gröbner basis for* $M :_F x_n^h$.

Proof. (a) Clearly, $\text{in}(M) + x_n^h F \subseteq \text{in}(M + x_n^h)F$. Now consider $\text{in}(u + x_n^h f)$ where $u \in U$ and $f \in F$. In revlex, the homogeneous component of an element of highest degree has the same initial form as the element, and so we may assume that $u + x_n^h f$ is homogeneous. If the initial term is divisible by x_n^h the result is proved. If not, it must be a term of u, and x_n must occur with a strictly smaller exponent than h. All other terms of u must be smaller: either they are not divisible by x_n^h and persist in $u + x_n^h f$, or they are divisible by x_n^h, which forces them to be smaller than u in revlex, by the definitiion of revlex. The statement about the Gröbner basis is immediate, since the specified elements are in $M + x_n^h F$ and their initial terms span $\text{in}(M) + x_n^h F$.

(b) We have that a monomial $\nu \in \text{in}(M :_F x_n^h)$ iff and $x^h \nu \in \text{in}(M)$ iff $x^h v = \text{in}(w)$ with $w \in M$ homogeneous. But x_n^h divides w if and only x_n^h divides $\text{in}(w)$, by the Lemma above, and the result is immediate. We then have that $\text{in}(M)$ is the span of the $\text{in}(g_j)R : Fx_n^h$, and these are the same as the $\text{in}(g_j / x_j^{t_j})R$. Again, we are using that a power of x_n divides g_j if and only if it divides $\text{in}(g_j)$. $\square$

We can now prove:

Theorem. *Let notation be as above, with $M \subseteq F$ graded, and use revlex order on F and R. Then $x_{k+1}, \ldots, x_n$ is a regular sequence on F/M if and only if it is a regular sequence on $F/\text{in}(M)$.*

Proof. Since regular sequences are permutable in the graded case, we may show instead the same result for $x_n, \ldots, x_{k+1}$. We already know the "if" part. Now suppose that x_n is not a zerodivisor on F/M. Then $M :_F x_n = M$, and so

$$\text{in}(M) = \text{in}(M :_F x_n) = \text{in}(M) : Fx_n = \text{in}(M).$$

The proof is now completed by induction: when we work mod x_n, R is replaced by $R/x_n R = K[x_1, \ldots, x_{n-1}]$, F by $F/x_n F$, and M by $M/x_n M \hookrightarrow F/x_n F$, since x_n is not a zerodivisor on $M/x_n M$. The hypothesis is preserved because of the preceding Proposition. $\square$

Lecture of February 5

Associated primes and primary decompostion for modules

Throughout this section R is a Noetherian ring and M an R-module. Recall that P is an *associated prime* of M if, equivalently

(1) There is an injection $R/P \hookrightarrow M$.

(2) There is an element $u \in M$ such that $\operatorname{Ann}_R u = P$.

The set of associated primes of M is denoted $\operatorname{Ass}(M)$. Although we have made this definition even when M need not be finitely generated, the rest of our study is restricted to the case where M is Noetherian. Note that if $M = 0$, then $\operatorname{Ass}(M) = \emptyset$. The converse is also true, as we shall see below.

Proposition. *Let M be a finitely generated R-module, where R is Noetherian*

(a) *If $u \neq 0$ is any element of M, one can choose $s \in R$ such that $\operatorname{Ann}_R su$ is a prime ideal P of R, and $P \in \operatorname{Ass}(M)$. In particular, if $M \neq 0$, then $\operatorname{Ass}(M)$ is nonempty.*

(b) *If $ru = 0$ where $r \in R$ and $u \in M - \{0\}$, then one can choose $s \in R$ such that $\operatorname{Ann}_R su = P$ is prime. Note that $r \in P$. Consequently, the set of elements of R that are zerodivisors on M is the union of the set of associated primes of M.*

(c) *If $M \neq 0$ it has a finite filtration $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_n = M$ in which all the factors M_i/M_{i-1} for $1 \leq i \leq n$ are prime cyclic modules, i.e., have the form R/P_i for some prime ideal P_i of R.*

(d) *If $N \subseteq M$, then $\operatorname{Ass}(N) \subseteq \operatorname{Ass}(M)$.*

(e) *If $0 \to M' \to M \to M'' \to 0$ is exact, then $\operatorname{Ass}(M) \subseteq \operatorname{Ass}(M') \cup \operatorname{Ass}(M'')$.*

(f) *If $0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_n = M$ is a finite filtration of M, then*

$$\operatorname{Ass}(M) \subseteq \bigcup_{i=1}^{n} \operatorname{Ass}(M_i/M_{i-1}).$$

(g) *If one has a prime cyclic filtration of M as in part (c), $\operatorname{Ass}(M) \subseteq \{P_1, \ldots, P_n\}$. In particular, $\operatorname{Ass}(M)$ is finite.*

(h) *If W is a multiplicative system in R, $\operatorname{Ass}(W^{-1}M)$ over $W^{-1}M$ is the set*

$$\{PW^{-1}R : P \in \operatorname{Ass}(M) \text{ and } P \cap W = \emptyset\}.$$

Proof. (a) The family of ideals $\{\operatorname{Ann}_R tu : t \in R \text{ and } tu \neq 0\}$ is nonempty since we may take $t = 1$. Since R has ACC, it has a maximal element $\operatorname{Ann}_R su = P$. We claim that P is prime. If $ab \in P$, then $abu = 0$. If $a \notin P$, we must have $b \in P$, or else $bu \neq 0$ and has annihilator containing $P + aR$ strictly larger than P.

(b) This is immediate from (a). Note that it is obvious that if $P = \mathrm{Ann}_R u$ with $u \in M$, then $u \neq 0$, and so every element of P is a zerodivisor on M.

(c) Choose a sequence of elements u_1, u_2, $\cdots$ in M recursively as follows. Choose u_1 to be any element of M such that $\mathrm{Ann}_R u_1 = P_1$ is prime. If $u_1, \ldots, u_i$ have been chosen and $Ru_1 + \cdots + Ru_i = M$, the seqeunce stops. If not, choose $u_{i+1} \in M$ such that its image $\overline{u}_{i+1}$ in $M/(Ru_1 + \cdots + Ru_i)$ has annihilator P_{i+1} that is prime. Let $M_i = Ru_1 + \cdots + Ru_i$. The sequence must stop, since the M_i are strictly increasing and M has ACC. By construction, the factors are prime cyclic modules.

(d) This is obvious, since if $R/P \hookrightarrow N$, we have a composite map $R/P \hookrightarrow N \hookrightarrow M$.

(e) Let $u \in M$ be such that $\mathrm{Ann}_R u = P$, which means that $Ru \cong R/P$. If Ru meets $M' - \{0\}$, say $ru = v$ is a nonzero element of M', then $\mathrm{Ann}_R v = P$ since v may be thought of as a nonzero element of R/P, and $P \in \mathrm{Ass}\,(M')$. If $Ru \cap M' = 0$, then the composite map $R/P \cong Ru \subseteq M \twoheadrightarrow M''$ is injective, and so $P \in \mathrm{Ass}\,(M'')$.

(f) We use induction on n. By part (e),

$$\mathrm{Ass}\,(M) \subseteq \mathrm{Ass}\,(M_{n-1}) \cup \mathrm{Ass}\,(M/M_{n-1})$$

and we may apply the induction hypothesis to $0 \subseteq M_1 \subseteq \cdots \subseteq M_{n-1}$.

(g) This is immediate from part (f), since $\mathrm{Ass}\,(R/P_i) = \{P_i\}$.

(h) If $R/P \hookrightarrow M$ and P does not meet W, then $W^{-1}R/PW^{-1}R \hookrightarrow W^{-1}M$. Conversely, suppose that $u/w_0 \in W^{-1}M$ where $u \in M$ and $w_0 \in W$ has annihilator Q in $W^{-1}R$. The same is true for $w_0(u/w_0) = u/1$. We know that $Q = PW^{-1}R$ for some prime P of R such that $P \cap W = \emptyset$. Choose $w \in W$ such that $\mathrm{Ann}_R wu$ is maximal. If $f \in P$, we know that $fwu/1$ is 0 in $W^{-1}M$, and so we can choose $v \in W$ such that $vfwu = 0$. But $\mathrm{Ann}_R(vwu) = \mathrm{Ann}_R(wu)$ by the maximality of $\mathrm{Ann}_R(wu)$, so that we must have $fwu = 0$. On the other hand, if $fwu = 0$ for $f \in R$, then $f/1 \in Q$, and so f is in the contraction of Q to R, which is P. We have shown that $P = \mathrm{Ann}_R(wu)$, and so $P \in \mathrm{Ass}\,(M)$. $\square$

Remark. If M is nonzero module over a Noetherian domain R, then M is torsion-free over R if and only if $\mathrm{Ass}\,(M) = \{(0)\}$, since this says precisely that no nonzero element of R is a zerodivisor on M.

Remark. There does not necessarily exist a filtration of M with prime cyclic factors in which the only primes that occur are associated primes of M. For example, let $R = K[x, y]$ be the polynomial ring in two variables over a field K and let $M = (x, y)R \subseteq R$, which is an ideal of R, but which we are viewing as a torsion-free R-module. Then $\mathrm{Ass}\,(M) = (0)$, but there is no finite filtration of M in which every factor is R, since M needs two generators but is rank one, and so is not free over R.

Recall that if M is finitely generated over a Noetherian ring R and $I = \mathrm{Ann}_R M$, then $P \in \mathrm{Supp}\,(M)$, which means that $M_P \neq 0$, if and only if $I \subseteq P$. The minimal primes of $\mathrm{Supp}\,(M)$ are the same as the minimal primes of I, and are called the *minimal primes* of M.

Proposition. *Let M be a finitely generated module over a Noetherian ring R, and let $I = \mathrm{Ann}_R M$.*

(a) *Every associated prime of R/I is an associated prime of M.*

(b) *Every associated prime of M contains a minimal prime of M. Every minimal prime of M is an associated prime of M, and so the minimal primes of M are the same as the minimal primes of* $\mathrm{Ass}\,(M)$.

(c) *Let m be a maximal ideal of R. Then the following conditions on M are equivalent:*

(1) $\mathrm{Ass}\,(M) = \{m\}$.

(2) $\mathrm{Supp}\,(M) = \{m\}$.

(3) *M is killed by a power of m.*

(4) *M has a finite filtration in which all the factors are $\cong R/m$.*

Proof. (a) Let $u_1, \ldots, u_h$ generate M. The map $R \to M^h$ that sends $r \mapsto (ru_1, \ldots, ru_h)$ has kernel I, yielding an injection $R/I \hookrightarrow M^h$. Since $M \subseteq M^h$, $\mathrm{Ass}\,(M) \subseteq \mathrm{Ass}\,(M^h)$. Since M_h has a finite filtration $0 \subseteq M \subseteq M \oplus M \subseteq \cdots \subseteq M^{h-1} \subseteq M^h$ in which all factors are M, $\mathrm{Ass}\,(M^h) \subseteq \mathrm{Ass}\,(M)$. Thus, $\mathrm{Ass}\,(R/I) \subseteq \mathrm{Ass}\,(M^h) = \mathrm{Ass}\,(M)$.

(b) Since $R/P \hookrightarrow M$, we have that I kills R/P, and so $I \subseteq P$, so that P contains a minimal prime of I. Every minimal prime of M is a minimal prime of R/I and, hence, an associated prime of R/I. Therefore every minimal prime of M is an associated prime of M by part (a). The final statement is now clear.

(c) (1) $\Leftrightarrow$ (2) since in both cases m is the only minimal prime of M. This implies that $\mathrm{Rad}\,(I) = m$, and so $m^h \subseteq I$ for some h and (2) $\Rightarrow$ (3). If $m^h M = 0$, M has a finite filtration $0 = m^h M \subseteq m^{h-1}M \subseteq \cdots \subseteq m^2 M \subseteq mM \subseteq M$ and each factor $m^i M/m^{i-1}M$ is killed by m, and so is a finite-dimensional vector space over $K = R/m$. Hence, this filtration can be refined to one in which every factor is $\cong R/m$, since every $m^i M/m^{i-1}M$ has a finite filtration in which all factors are $\cong R/m$. Thus (3) $\Rightarrow$ (4). Finally, (4) $\Rightarrow$ (1) by part (g) of the earlier Proposition. $\square$

If P is a prime ideal of R, M is called *P-coprimary* if, equivalently,

(1) $\mathrm{Ass}\,(M) = \{P\}$.

(2) $M \neq 0$, for some $h \geq 1$, $P^h M = 0$, and every element of $R - P$ is a nonzerodivisor on M.

(3) $M \hookrightarrow M_P$ is injective, and M_P has finite length over R_P.

We need to check that these three conditions are equivalent. (1) $\Rightarrow$ (3), for if $\mathrm{Ass}\,(M) = P$ all elements of $R - P$ are nonzerodivisors on M and $M \hookrightarrow M_P$. But since $\mathrm{Ass}\,(M_P) = \{PR_P\}$ by part (h) of the Proposition on p. 1, and PR_P is maximal in R_P, this implies that M_P has finite length by the equivalence of (1) and (4) in part (c) of the preceding Proposition.

Assume (3). Then $(PR_P)^h$ kills M_P for some h, and so P^h kills $M \hookrightarrow M_P$. Since the elements of $R-P$ act invertibly on M_P, they are not zerodvisors on $M \subseteq R_P$. This shows that (3) $\Rightarrow$ (2).

Finally, assume (2). Choose k as large as possible such that $P^k M \neq 0$: we allow $k = 0$. By hypothesis, $k \leq h-1$. Choose $u \neq 0$ in $P^k M$. Then $Pu \subseteq P^{k+1}M = 0$, while no element of $R-P$ kills u. It follows that $P \in \operatorname{Ass}(M)$. Moreover $P^h \subseteq \operatorname{Ann}(M)$ shows that every associated prime contains of M contains P. But there cannot be an associated prime strictly larger than P, since it would contain an element of $R-P$, and such an element is a nonzerodivisor on R. Hence, (2) $\Rightarrow$ (1), as required. $\square$

Remark. When $M = R/I$ with R a proper ideal of R, it is easy to see that M is P-coprimary if and only if I is primary to P.

We shall say that a proper submodule N of M is *irreducible* if it is not the intersection of two strictly larger submodules of M. It is easy to see that this is equivalent to the condition that N not be the intersection of finitely many larger submodules of M. Note that in each part of the Lemma below, we can replace M by M/N, N by 0, and each submodule of M containing N by its image modulo N without affecting any relevant issue.

Lemma. *Let R be a Noetherian ring and let $N \subset M$ be finitely generated R-modules, where the inclusion is strict.*

(a) *N is a finite intersection of irreducible submodules of M (this includes the possibility that N itself is irreducible).*

(b) *If N is irreducible, then M/N it is P-coprimary for some prime P.*

(c) *If $N_1, \ldots, N_k$ are submodules such that each M/N_j is P-coprimary to P for the same prime P, then $M/\bigcap_{j=1}^k N_j$ is also P-coprimary.*

Proof. (a) Let $\mathcal{N}$ denote the set of proper submodules of M that are not finite intersections of irreducible submodules. If $\mathcal{N}$ is nonempty, it has a maximal element N. Then N cannot itself be irreducible. Suppose that $N = N_1 \cap N_2$ where N_1 and N_2 are strictly larger. Then each N_i is a finite intersection of strictly larger submodules, and, hence, so is $N_1 \cap N_2 = N$, a contradiction.

(b) We replace M by M/N and so assume that $N = 0$. If $\operatorname{Ass}(M)$ contains two or more relevant primes, then we can choose $u \in M$ such that $\operatorname{Ann}_R u = P$ and $v \in M$ such that $\operatorname{Ann}_R v = Q$, where $P \neq Q$ are distinct primes. Then $Ru \cap Rv$ must be 0: any nonzero element of Rv has annihilator P, while any nonzero element of Ru has annihilator Q. This contradicts the irreducibility of 0.

(c) The map $M \to \prod_{j=1}^k M/N_j$ that sends $u \mapsto (u+N_1, \cdots, u+N_k)$ has kernel $N = \bigcap_{j=1}^k N_j$, and so we have $M/(\bigcap_{j=1}^k N_j) \hookrightarrow \prod_{j=1}^k N_j \cong \bigoplus_{j=1}^k (M/N_j)$ The latter has a filtration by submodules $M/N_1 \oplus \cdots \oplus M/N_j$ with factors M/N_1, M/N_2, $\ldots$, M/N_k. Hence, $\operatorname{Ass}(N) \subseteq \bigcup_{j=1}^k \operatorname{Ass}(M/N_j) = \{P\}$, as required. $\square$

If $N \subset M$ is a strict inclusion of finitely generated modules over a Noetherian ring R, we shall say that $N = N_1 \cap \cdots \cap N_k$ is a *primary decomposition* for N in M if

(1) Each M/N_i is P_i-coprimary for some prime P_i of R.

(2) If $i \neq j$ then $P_i \neq P_j$.

(3) The intersection is *irredundant* in the sense that if any N_j is omitted, the intersection of the others is strictly larger than N.

Theorem (primary decomposition for modules). *Let R be a Noetherian ring and let $N \subset M$, where the inclusion is strict. Then N has a primary decomposition. In any primary decomposition, the primes occurring are precisely the elements of* $\mathrm{Ass}\,(M/N)$*, and the number of terms is the number of primes in* $\mathrm{Ass}\,(M/N)$*. If P_i is a minimal prime of M/N, then the corresponding P_i-coprimary module N_i in the primary decomposition is uniquely determined and is, in fact, $Ker\big(M \to (M/N)_{P_i}\big)$.*

Proof. To prove existence, first write N as a finite intersection of irreducibles N_j, by part (a) of the preceding Lemma. For each prime P such that one of these is coprimary to P, replace those N_j that are P-coprimary by their intersection. Thus, N is an intersection of P-coprimary modules such that the primes that occur are mutually distinct. If the intersection is not irredundant, we may successively omit terms until we reach an intersection that is irredundant.

We now want to prove the uniqueness statement. We pass to M/N and so assume that $N = 0$. Suppose that $0 = N_1 \cap \cdots \cap N_k$ is a primary decomposition for 0 in M, where M/N_j is P_j-coprimary. As in part the proof of part (c) of the preceding Lemma, we have an injection $M \hookrightarrow \bigoplus_{j=1}^{k}(M/N_j)$, and it folows that $\mathrm{Ass}\,(M)$ is contained in the set of primes $\{P_1, \ldots, P_k\}$. To see that $P_i \in \mathrm{Ass}\,(M)$, note that since the intersection is irredundant, we can choose an element $u \in \bigcap_{j\neq i} N_j - N_i$. The image of u under $M \hookrightarrow \bigoplus_{j=1} M/N_j$ is 0 in every M/N_j except M/N_i, and is nonzero in M/N_i. Hence,

$$\mathrm{Ass}\,(Ru) \subseteq \mathrm{Ass}\,(M/N_i) = \{P_i\},$$

and so $\mathrm{Ass}\,(Ru) = \{P_i\}$. But $Ru \subseteq M$, and so $\mathrm{Ass}\,(Ru) \subseteq \mathrm{Ass}\,(M)$, i.e., $P_i \in \mathrm{Ass}\,(M)$. The statement about the number of terms is now obvious.

Finally, suppose that $P = P_i$ is minimal among the associated primes. For every $P_j \neq P_i$, $P_j - P_i$ is nonempty. It follows that $(M/N_j)_{P_i} = 0$, so that $M_{P_i} = (N_j)_{P_i}$. Now,

$$N_{P_i} = (N_1 \cap \cdots \cap N_k)_{P_i} = (N_1)_{P_i} \cap \cdots \cap (N_k)_{P_i} = (N_i)_{P_i},$$

so that $(M/N)_{P_i} \cong (M/N_i)_{P_i}$. Hence, the kernel of $M \to (M/N)_{P_i}$ is the set of $u \in M$ such that for some $w \in R - P_i$, $wu \in N_i$. Since M/N_i is P_i-coprimary, no element of $R - P_i$ is a zerodivisor on M/N_i, it follows that the kernel is N_i. $\square$

Depth

We give a brief introduction to the theory of depth without using homological methods: the homological proofs of certain results, such as the fact that maximal regular sequences in I on a module M all have the same length, are very slick, but in some ways mask the simplicity of what is going on.

We shall assume that $R \to S$ is a homomorphism of Noetherian rings, that I is an ideal of R, and that M is a finitely generated S-module. By far the most important case is the one where $R = S$, and the reader is encouraged to focus on this situation if this is a first encounter with depth. The greater generality is very useful, however, in that one can frequently choose regular sequences that arise from a "smaller" ring.

If $IM = M$, we define the *depth* of M on I to be $+\infty$. If $IM \neq M$, it turns out that all maximal regular sequences on M consisting of elements of I have the same length, and we define this length to be the *depth* of M on I. This fact is proved below. Before giving the proof, we want to characterize the "degenerate" situation in which $IM = M$ in a down-to-earth way.

Proposition. *Let $R \to S$ be a homomorphism of Noetherian rings, let N be a finitely generated R-module, and let M a finitely generated S-module. Let I be the annihilator of N in R, and let J be the annihilator of M in S.*

(a) *The support of $N \otimes_R M$ over S is $\mathcal{V}(IS + J) = \{Q \in \operatorname{Spec}(S) : IS + J \subseteq Q\}$. In particular, $N \otimes_R M = 0$ if and only if $IS + J = S$, the unit ideal.*

(b) *In particular, if $N = R/I$ is cyclic, $IM = M$ if and only if $IS + J = S$.*

Proof. (a) Since I kills N, IS kills $N \otimes_R M$, and since J kills M, J kills $N \otimes_R M$. Thus, any prime in the support of $N \otimes_R M$ must contain $IS + J$. Now suppose that $IS + J \subseteq Q$, a prime of S, and the Q lies over P in R. It suffices to see that $(N \otimes_R M)_Q \neq 0$, and this may be identified with $N_P \otimes_{R_P} M_Q$. Here, $I \subseteq P$, and so $N_P \neq 0$. Let $R_P/PR_P = K$. By Nakayama's Lemma, N_P/PN_P is a nonzero K-vector space, say K^h, $h \geq 1$, and, similarly, M_Q/QM_Q is a nonzero vector space over $L = S_Q/QS_Q$, say L^k. Then $N_P \otimes_{R_P} M_Q$ maps onto $K^h \otimes_K L^k \cong (K \otimes_K L)^{hk} \cong L^{hk} \neq 0$, as required.

(b) This is immediate from part (a), since $N \otimes_R M \cong M/IM$ in this case. □

We will need the following:

Lemma. *Let R be a ring and let $x_1, \ldots, x_n$ be a regular sequence on an R-module M. Suppose that x_2 is not a zerodivisor on M. Then $x_2, x_1, x_3, x_4, \ldots, x_{n-1}, x_n$ is a regular sequence on M.*

Proof. It suffices to show that x_1 is a nonzerodivisor modulo x_2M: since $M/(\boldsymbol{x}_1, x_2)M = M/(x_2, x_1)M$, the remaining conditions are unaffected by the interchange of x_2 and x_1. Suppose that $x_1u \in x_2M$, say $x_1u = x_2v$. Since x_1, x_2 is a regular sequence on M and $x_2v \equiv 0 \bmod x_1M$, we have that $v \in x_1M$, say $v = x_1w$. Then $x_1u = x_2x_1w$, and $x_1(u - x_2w) = 0$. Since x_1 is not a zerodivisor on M, $u \in x_2M$, as required. □

We can now justify the definition we want to give for depth.

Theorem. *Let $R \to S$ be a homomorphism of Noetherian rings, let M be a finitely generated S-module, and let $I \subseteq R$ be an ideal of R. Assume that $IM \neq M$.*

(a) *There is no infinite regular sequence $x_1, x_2, x_3, \ldots$ on M consisting of elements of I.*

(b) *There is no zerodivisor on M in I if and only if I is contained in the contraction of a prime in $\operatorname{Ass}(M)$ to R. Hence, there is a no nonzerodivisor on M in I if and only if there is an element $u \in M - \{0\}$ such that $Iu = 0$.*

(c) *Every regular sequence in I on M (including the empty regular sequence) can be extended to a maximal regular sequence in I on M, and this maximal regular sequence is always finite.*

(d) *All maximal regular sequences in I on M have the same length.*

Proof. (a) Suppose we have such a sequence. Let $I_n = (x_1, \ldots, x_n)R$. Since R is Noetherian, we eventually have $I_n = I_{n+1}$. This means that $x_{n+1} \in I_n$, and so kills M/I_nM. Since x_{n+1} is not a zerodivisor on M/I_nM, we must have $M/I_nM = 0$, i.e., $M = I_nM$. But $I_n \subseteq I$ and $M \neq IM$, a contradiction.

(b) Let θ denote the map $R \to S$. Note that the action of $x \in R$ on M is the same as the action of $\theta(x)$. Hence, $x \in R$ is a zerodivisor on M if and only if $\theta(x)$ is a zerodivisor on M, and this means that $\theta(x)$ is in the union of the associated primes $Q_1, \ldots, Q_k$ of M in S. Let P_i denote the contraction of Q_i to R. We then have that I consists entirely of zerodivisors on M if and only if it is contained in the union of the P_i. But then it is contained in some P_i. Choose $u \in M - \{0\}$ such that $\operatorname{Ann}_S u = Q_i$. Then, since $\theta(I) \subseteq Q_i$, $Iu = 0$, as required.

(c) Suppose that we have a regular sequence $x_1, \ldots, x_k$ and that I_k is the ideal $(x_1, \ldots, x_k)R$. If every element of I is a zerodivisor on M/I_kM, then we have constructed the required maximal regular sequence on M in I. If not, we can enlarge the regular sequence to $x_1, \ldots, x_{k+1}$ by taking x_{k+1} to be an element of I that is not a zerodivisor on M/I_kM. We can continue recursively in this way. The process must terminate by part (a).

(d) Suppose that we have a counterexample. Since M has ACC, among all submodules N of M such that M/N provides a counterexample, there is a maximal one. (The family is nonempty, since it contains 0.) Therefore, we may assume the result holds for every proper homomorphic image of M. If I consists entirely of zerodivisors on M, the empty sequence is the unique maximal regular sequence on M.

Now suppose that $x \in I$ is a maximal regular sequence on M. Then I consists entirely of zerodivisors on M/xM, and by part (b), there exists an element $u \in M - xM$ such that $Iu \subseteq xM$. Now let $y \in I$ be a nonzerodivisor. We want to show that it constitutes a maximal regular sequence. Since $Iu \subseteq xM$, we can write $yu = xv$ for $v \in M$. First note that $v \notin yM$, for if $v = yw$, then $yu = xyw$. Since y is a nonzerodivisor, this implies, $u = xw$, a contradiction. The argument in this case will therefore be complete if we can show that $Iv \subseteq yM$. But if $f \in I$, we have $xfv = f(xv) = f(yu) = y(fu) = y(xw)$ for

some $w \in M$, since $Iu \subseteq xM$. But then $x(fv - yw) = 0$, and since x is not a zerodivisor on M, we have that $fv = yw \in yM$, as required.

Finally, suppose that we have two maximal regular sequences $x_1, \ldots, x_h$ and $y_1, \ldots, y_k$ on M in I where $h \geq 2$ and $k \geq 2$. Then the contractions to R of the associated primes of M/x_1M do not cover I (they miss x_2), and the contractions to R of the associated primes of M/y_1M do not cover I similarly. Likewise, the contractions of the associated primes of M do not cover I (they miss x_1). It follows that the union of all three sets of primes does not cover I: if it did, I would be contained in one of these primes, a contradiction. We can therefore pick $z \in I$ not in any of them. Then x_1, z is a regular sequence on M, and can be extended to a maximal regular sequence on M in I, say x_1, z, x_3', $\ldots$, $x_{h'}'$. Similarly, we can construct a maximal sequence on M in I of the form y_1, z, y_3', $\ldots$, $y_{k'}'$.

But if two maximal regular sequence on I in M have the same first term, say x, then the terms after the first form maximal regular sequences on M/xM, a proper quotient of M. It follows that they have the same length, since we know the result for M/xM, and so the original regular sequences have the same length. Thus, $h' = h$ and $k' = k$. By the Lemma above, z, x_1, x_3', $\ldots$, x_h' is also a regular sequence on M, and so is z, y_1, y_3', $\ldots$, y_k'. Since these two have the same first term, we obtain that $h = k$. $\square$

We are now justified, under the hypotheses of the Theorem above, in defining the *depth* of M on I, which we shall denote $\operatorname{depth}_I M$, to be the length of any maximal regular sequence on M whose terms are in I.

We also note:

Proposition. *Let R be a finitely generated $\mathbb{N}$-graded algebra with $R_0 = K$, a field, let $m = \bigoplus_{d=1}^{\infty} R_d$ be the homogeneous maximal ideal, and let M be a finitely generated $\mathbb{Z}$-graded R-module. Then the depth of M on m is the same as the length of any maximal regular sequence on M consisting of forms of positive degree. Hence, R is Cohen-Macaulay if and only if* $\operatorname{depth}_m R = \dim(R)$.

Proof. First note that if $\operatorname{depth}_m M > 0$, then we can construct a nonzero form F_1 of positive degree that is not a zerodivisor on M, by homogeneous prime avoidance. We can then proceed recursively to construct a maximal regular sequence of such forms on M: we begin by passing to M/F_1M. The final statement is now obvious. $\square$

Lecture of February 8

Remark. Let $R \to S$ be a homomorphism of Noetherian rings, I an ideal of R, and M a finitely generated S-module such that $IM \neq M$. Let $x_1, \ldots, x_k \in I$ be a regular sequence on M. Let $J = (x_1, \ldots, x_k)R$. Then $\operatorname{depth}_I M/JM = \operatorname{depth}_M - k$, and $\operatorname{depth}_{I/J} M/JM = \operatorname{depth}_I M - k$, where in the second equality we have replaced R by R/J and S by S/JS. The point is that if we extend $x_1, \ldots, x_k$ to a maximal regular sequence $x_1, \ldots, x_n$ in I

on M, then $x_{k+1}, \ldots, x_n$ is very easily seen to be a maximal regular sequence in I on M/JM, and its image in R/J is a maximal regular sequence in I/J on M/JM.

Remark. We next want to see that, with the same hypothesis as in the first sentence of the previous remark, we have that $\operatorname{depth}_I M = \operatorname{depth}_{IS} M$. Let $\theta : R \to S$ be the map, and let $x_1, \ldots, x_n$ be a maximal regular sequence in I on M. Clearly, $\theta(x_1), \theta(x_2), \ldots, \theta(x_n)$ is a regular sequence on M in IS because x_i acts on M exactly the way that $\theta(x_i)$ acts on M. We need only see that it is maximal. Again, since the x_i act on M exactly as the $\theta(x_i)$ act on M, we have that

$$M/(x_1, \ldots, x_n)M = M/\big(\theta(x_1), \theta(x_2), \ldots, \theta(x_n)\big)M.$$

Since $x_1, \ldots, x_n$ is a maximal regular sequence on M, there exists an element $u \in M/(x_1, \ldots, x_n)M - \{0\}$ that is killed by I. Since the annihilator of u is an ideal of S, we must have that u is killed by IS as well, which shows that $\theta(x_1), \theta(x_2), \ldots, \theta(x_n)$ is a maximal regular sequence in IS on M, as required. $\square$

Remark. When (R, m, K) is local, $\operatorname{depth}(M)$, with no specification of an ideal, is understood to be $\operatorname{depth}_m M$.

Remark. When I is an ideal of R, $\operatorname{depth}_I R$ is sometimes referred to as the *depth of I as an ideal*. However, the phrase "as as ideal" is frequently omitted. This terminology is flawed, since the two depths may be different. For example, if $R = K[[x, y]]$ and $I = (x, y)R$, the depth of I as an ideal is 2, since x, y is a regular sequence. However, if I is regarded as an R-module, the depth of I on $m = (x, y)R$ is only one: I/xI has depth 0, since the image of x is killed by $m = I$, while $x \notin mI$. However, the situtation is rarely confusing, because when I is an ideal, "the depth of I" is almost always used for $\operatorname{depth}_I R$

Linear systems of parameters for standard graded algebras

We shall refer to a finitely generated $\mathbb{N}$-graded algebra R over $R_0 = K$, a field, such that R_1, the vector space of linear forms, generates R, as a *standard* graded K-algebra. The following fact gives a very strong form of avoidance of ideals, not just prime ideals, and will enable us to prove the existence of regular sequences consisting of linear forms.

Proposition. *Let K be an infinite field, $V \subseteq W$ be vector spaces, and let $V_1, \ldots, V_h$ be vector subspaces of W such that $V \subseteq \bigcup_{i=1}^h V_i$. Then $V \subseteq V_i$ for some i.*

Proof. If not, for each i choose $v_i \in V - V_i$. We may replace V by the span of the v_i and so assume it is finite-dimensional of dimension d. We may replace V_i by $V_i \cap V$, so that we may assume every $V_i \subseteq V$. The result is clear when $d = 1$. When $d = 2$, we may assume that $V = K^2$, and the vectors $(1, c)$, $c \in K - \{0\}$ lie on infinitely many distinct lines. For $d > 2$ we use induction. Since each subspace of $V \cong K^d$ of dimension $d - 1$ is covered by the V_i, each is contained in some V_i, and, hence, equal to some V_i. Therefore it suffices to

see that there are infinitely many subspaces of dimension $d-1$. Write $V = K^2 \oplus W$ where $W \cong K^{d-2}$. The line L in K^2 yields a subspace $L \oplus W$ of dimension $d-1$, and if $L \neq L'$ then $L \oplus W$ and $L' \oplus W$ are distinct subspaces. $\square$

Theorem. *Let R be a standard graded K-algebra, and let M be a $\mathbb{Z}$-graded fnitely generated R-module. Let $m = \bigoplus_{d=1}^{\infty}[R]_d$ denote the homogeneous maximal ideal of R.*

(a) *R has a homogeneous system of parameters consisting of linear forms.*

(b) *The depth of M on m is at least h if and only if there exists a regular sequence on M consisting of linear forms.*

(c) *In particular, the depth of R on m is at least h if and only if there is a regular sequence on R of length h consisting of linear forms.*

Proof. We construct the required sequence of linear forms for (a) and (b) recursively as follows. If the union of minimal primes of R (respectively, the associated primes of M) contains $V = [R]_1$, then by the Lemma above, one of them contains V, and since V generates m, we have that m is a minimal prime of R (respectively, an associated prime of M). In the case of (a), R has dimension 0. In the case of (b), the depth of M on m is 0. In either case, the empty sequence satisfies the condition. If not, we can choose a linear form F_1 that is not in the union of these primes. This gives the first element of a system of parameters in (a), and the first element of a regular sequence in part (b). We can construct the required sequence recursively by passing to R/F_1R for (a) and to M/F_1M for (b). Part (c) is simply the case of (b) where $M = R$. $\square$

Change of field

We want to make several comments about the effect of change of field on various questions.

Discussion: change of field and Gröbner bases. Let $R = K[x_1, \ldots, x_n]$, let F be a finitely generated free R-module with ordered basis $e_1, \ldots, e_s$, let $M \subseteq F$ be a submodule, and fix a monomial order on F. Let $K \subseteq L$ be a field extension. We use L as a subscript to indicate the result of applying $L \otimes_K \_$. Thus, $R_L \cong L[x_1, \ldots, x_n]$, F_L is a finitely generated free R_L-module with ordered basis $1 \otimes e_1, \ldots, 1 \otimes e_s$, $M_L \subseteq F_L$, and the monomials in F_L are the images of the monomials in F under the obvious injection $F \hookrightarrow F_L$ that sends $f \mapsto 1 \otimes f$ for all $f \in F$. We identify F with its image. The monomial order on F then immediately gives a corresponding monomial order on F_L because the two sets of monomials have been identified.

In this situation, let $g_1, \ldots, g_r$ denote a Gröbner basis for M. Then $g_1, \ldots, g_r$ is a Gröbner basis for M_L as well. We can apply the Buchberger criterion to see this: as we apply it, all the divisions can be carried out over K, and so we have standard expressions with remainder 0, as required, independent of whether we think over K or

over L. This implies that $\text{in}(M_L)$ contains the same monomials as $\text{in}(M)$, and we have $\text{in}(M_L) = \text{in}(M)_L$.

In the graded case, the Hilbert functions of M and M_L are the same. We know that R and R_L are both Cohen-Macaulay or not alike: this is problem 4(d). in Problem Set #2.

There are also many properties of rings which, if they hold for R_L, hold for R. If R_L is (1) reduced, or (2) a domain, or (3) normal, so is R. (1) and (2) hold simply because $R \subseteq R_L$. The third may be proved as follows: suppose that a/b is integral over R, where $a \in R$ and b is a nonzerodivisor over R. Because R_L is flat over R, b is a nonzerodivisor on R_L, and a/b is certainly integral over R_L. It follows that $a/b \in R_L$, and so $a \in bR_L \cap R$. When S is faithfully flat over R, for every ideal I of R, $IS \cap R = I$ ($R/I \to S/IS$ is still faithfully flat, which implies injective). Hence, $bR_L \cap R = bR$, and we have that $a \in bR$, which implies that $a/b \in R$.

However, R_L can be a UFD even though R is not. For example, if $R = \mathbb{R}[X, Y](X^2 + Y^2 - 1)$, where X and Y are indeterminates, it turns out that R is not a UFD (the height one prime ideal $(X, Y-1)R$ can be shown not to be principal), but $R_{\mathbb{C}} \cong \mathbb{C}[X, Y]/(X^2+Y^2-1)$ is a UFD: one can use new variables $U = X + iY$, $V = X - iY$, making a linear change of coordinates over $\mathbb{C}$, and then see that $R_{\mathbb{C}} \cong \mathbb{C}[U, V]/(UV - 1) \cong \mathbb{C}[U, 1/U]$.

Generic linear combinations as regular sequences

We want to show that if an ideal contains a regular sequence on a module M, one can use "generic" linear combinations of the generators of the ideal, i.e., linear combinations with indeterminate coefficients, to produce such a regular sequence. We first observe:

Proposition. *Let R be a Noetherian ring, and $S = R[z_1, \ldots, z_h] = R[z]$ a polynomial ring over R. Let $N \subseteq M$ be finitely generated R-modules. We write $M[z]$ for $R[z] \otimes_R M$.*

(a) *If P is prime in R, then PS is prime in S.*

(b) *If M is P-coprimary, then $M[z]$ is PS-coprimary.*

(c) *If $N = N_1 \cap \cdots \cap N_k$ is a primary decomposition of N in M, then we have that $N[z] = N_1[z] \cap \cdots \cap N_k[z]$ is a primary decompostion of $N[z]$ in $M[z]$.*

(d) $\text{Ass}\,(M[z])$ *over S is* $\{PS : P \in \text{Ass}\,(M)\}$.

Proof. (a) There is an obvious surjetion $R[z] \twoheadrightarrow (R/P)[z]$. The result follows because $(R/P)[z]$ is a domain, and the kernel is clearly $PR[z]$.

(b) We may localize at $R - P$, which consists of nonzerodivisors on both M and $M[z]$, without affecting the issue, and so we may assume that (R, P, K) is local. Then M has a finite filtration whose factors are copies of K, and since $R[z]$ is R-flat, $M[z]$ has a finite filtration by copies of $K \otimes R[z] = K[z]$. Since $\text{Ass}\,(K[z])$ over S is clearly $PR[z]$, this is also true for $M[z]$.

(c) We have that $N[z] = N_1[z] \cap \cdots \cap N_k[z]$ since flat base change preserves finite intersection. Suppose that N_i is P_i-coprimary. By part (b), $N_i[z]$ is P_iS-coprimary. It remains only to see that the intersection of the $N_j[z]$, omitting $N_i[z]$, is not $N[z]$. This follows from the fact that the intersetion of the N_j, omitting N_i, is not N, and the fact that S is faithfully flat over R.

(d) This is immediate from the primary decomposition in part (c). $\square$

Corollary. *Let R be a Noetherian ring, let I be an ideal of R with generators $f_1, \ldots, f_h$, and let M be a finitely generated R-module with $IM \neq M$.*

(a) *If $\operatorname{depth}_I M \geq 1$ and $z_1, \ldots, z_h$ are indeterminates over R, then $g = z_1 f_1 + \cdots + z_h f_h$ is a nonzerodivisor on $M[z]$ in $IR[z]$.*

(b) *If $\operatorname{depth}_I M = n$, then for every set of indeterminates z over R, $\operatorname{depth}_I R[z]M[z] = n$. Moreover, if we take z to include indeterminates $z_{i,j}$ where $1 \leq i \leq n$ and $1 \leq j \leq h$ and we let $g_i = z_{i,1} f_1 + \cdots + z_{i,h} f_h$ for $1 \leq j \leq n$, then $g_1, \ldots, g_n$ is a maximal regular sequence in $IR[z]$ on $M[z]$. In particular, we may take $M = R$.*

(c) *Let R be a finitely generated K-algebra. Let notation be as in part (b). Let $L = K(z)$, the fraction field of $K[z]$, where the indeterminates z include $z_{i,j}$ as in part (b). Let the subscript L indicate the result of applying $L \otimes_K \_$. Then $g_1, \ldots, g_n$ is a maximal regular sequence in IR_L on M_L. In particular, if $M = R$, $g_1, \ldots, g_n$ is a maximal regular sequence in R_L.*

Proof. (a) If g is a zerodivisor, it is in $\operatorname{Ass}(M[z])$, and so it is in $PR[z]$ for some associated prime P of M. This implies that all coefficients occurring are in P, and so $I \subseteq P$, which contradicts $\operatorname{depth}_I M \geq 1$.

Part (b) is simply the iterated use of (a). In part (c), it is clear that the $g_1, \ldots, g_n$ will still be a regular sequence after localization, provided that we still have $IM_L \neq M_L$. This follows from the fact that L is free, and, hence, faithfully flat, over K. $\square$

The Zariski topology on K^n over an infinite field K

Let K be an infinite field. We consider the ring $R = K[x_1, \ldots, x_n]$ of polynomials as a ring of functions on K^n. We note that if a polynomial is nonzero as an element of R, then it yields a nonzero function. In fact, this is true if $n = 1$ because a nonzero polynomial of degree at most n has at most n roots, and K is infinite. We may use induction on n. A polynomial $f(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ may be written as a polynomial in x_n with coefficients in $K[x_1, \ldots, x_{n-1}]$. If it is nonzero, we can choose a nonzero coefficient $g(x_1, \ldots, x_{n-1})$, and by the induction hypothesis we can choose a point $(c_1, \ldots, c_{n-1}) \in K^{n-1}$ such that $g(c_1, \ldots, c_{n-1}) \neq 0$. Then $F(c_1, \ldots, c_{n-1}, x_n)$ is a nonzero polynomial in $K[x_n]$, and so by the one variable case we can choose c_n so that it does not vanish for $x_n = c_n$.

If $\mathcal{S}$ as any subset of R, we let

$$\mathcal{V}(\mathcal{S}) = \{(c_1, \ldots, c_n) \in K^n : \text{ for all } f \in \mathcal{S},\ f(c_1, \ldots, c_n) = 0\},$$

and we shall say that these sets are *closed algebraic sets* in K^n. As in the case where K is algebraically closed, $\mathcal{V}(\mathcal{S})$ is the same as $\mathcal{V}(I)$, where I is the ideal generated by $\mathcal{S}$, and it is also the same as $\mathcal{V}\big(\mathrm{Rad}\,(I)\big)$. However, distinct radical ideals may define the same closed algebraic set.

These sets are, likewise, the closed sets of a topology on K^n called the *Zariski topology*. We note that the complement of any proper closed set $\mathcal{V}(I)$ of K^n is Zariski dense in K^n. That is, every nonempty Zariski open set in K^n is dense. To see this, note that we have at least one nonzero polynomial f in I. If the complement of $\mathcal{V}(I)$ were a proper closed set, it would be contained in $\mathcal{V}(g)$ for some nonzero polynomial g. But then the nonzero polynomial fg vanishes everywhere, a contradiction.

We may view $G = \mathrm{GL}(n,\, K)$ as a Zariski open set in K^{n^2}. We may identify an $n \times n$ matrix with a point of K^{n^2}, and then G is the complement of the set where the deteminant function vanishes. We note that, as in the case of an algebraically closed field, the open subset X_f of an algebraic set $X \subseteq K^N$ where a polynomial f does not vanish may be viewed as closed algebraic set in K^{N+1}: it is in bijective correspondence with the set

$$\{(c_1, \ldots, c_{N+1}) \in K^{n+1} : (c_1, \ldots, c_N) \in X \text{ and } c_{N+1} = 1/f(c_1, \ldots, c_n)\},$$

which is the closed set defined by the same polynomials in $x_1, \ldots, x_N$ that define X along with the polynoial $fx_{N+1} - 1$. The inherited Zariski topologies on $X_f \subseteq X$ and on the corresponding set in K^{N+1} are the same.

In particular, we have a Zariski topology on $\mathrm{GL}(n,\, K)$, and every nonempty open subset is dense: such a subset is open in K^{n^2}, and hence dense even in K^{n^2}.

We shall write $\mathcal{B}_n^{\mathrm{U}}$ for the subgroup of upper triangular invertible matrices in $\mathrm{GL}(n,\, K)$ and $\mathcal{B}_n^{\mathrm{L}}$ for the subgroup of lower triangular invertible matrices. The subscript $_n$ will often be ommitted.

Generic initial modules

Let $R = K[x_1, \ldots, x_n]$ where K is an infinite field, let F be a finitely generated free R-module with ordered basis, and fix a monomial order on F. Let M be a submodule of F.

Let $A \in \mathrm{GL}(n,\, K)$. Then $A = \big(a_{i,j}\big)$ acts on the vector space $[R]_1$ of forms of degree 1 by sending the form $c_1x_1 + \cdots + c_nx_n$ to the form $c_1'x_1 + \cdots + c_n'x_n$ where

$$A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} c_1' \\ \vdots \\ c_n' \end{pmatrix}.$$

This means that the coefficients of $A(x_j)$ are given by the entries of the i th column of the matrix, i.e., $Ax_j = \sum_{i=1}^{n} a_{i,j}x_i$. This is a left action of $G = \mathrm{GL}(n, K)$ on the vector space of one-forms.

This action extends to an action of $\mathrm{GL}(n, K)$ on R by K-algebra automorphisms, where $A : f \mapsto f\big(A(x_1), \ldots, A(x_n)\big)$. The action extends also to F in an obvious way by letting $A(f_1e_1 + \cdots + f_se_s) = A(f_1)e_1 + \cdots + A(f_s)e_s$.

We let $A(M)$ denote the image of M under the action of F. We want to prove:

Theorem. *There is a Zariski open subset U of $\mathrm{GL}(n, K)$ such that for all $A \in U$, $\mathrm{in}(AM)$ is the same monomial module. Moreover, if $Z = \big(z_{i,j}\big)$ is an $n \times n$ matrix of indeterminates over K, and $L = K(z_{i,j} : i, j)$, so that $Z \in GL(n, L)$, then $\mathrm{in}\big(ZM_L\big)$ gives a monomial module containing the same monomials.*

Proof. Let $g_1, \ldots, g_r$ be a Gröbner basis for ZM_L containing images for generators of M under Z. We form a finite family of polynomials in $K[Z]$ as follows. We include all denominators of coefficients of the g_r, and all numerators of the coefficients of their initial terms. By the Buchberger criterion, for each i, j there is a standard expression

$$(*) \quad G_{i,j} = \sum_{k=1}^{r} q_{i,j,k}g_k$$

with remainder 0. We include in our family all denominators of coefficients of the $q_{i,j,k}$ and all numerators of the initial terms of the $q_{i,j,k}g_k$. Let f be the product of all the polynomials in this family. The Gröbner basis and all elements in the expressions $(*)$ have coefficients in $K[Z]_f$. For any matrix $A \in \mathrm{GL}(n, K)_f$, there is a K-homomorphism $K[Z]_f \to K$ that maps the entries of Z to the corresponding entries of A. This map carries $g_1, \ldots, g_r$ to a Gröbner basis for AM: we may take the images of the expressions in $(*)$, and these show that we have a Gröbner basis by the Buchberger criterion. The monomial initial terms of $g_1, \ldots, g_r$ therefore generate both $\mathrm{in}(ZM_L)$ and every $\mathrm{in}(AM)$ for $A \in GL(n, K)_f = U$. $\square$

The common initial module that we have proved to exist is denoted $\mathrm{Gin}(M)$, and called the *generic initial module*. Note that even when K is finite, we can still consider the span in F of the monomial terms in $\mathrm{in}(ZM_L)$ as a generic initial module: it becomes one after a base change to any infinite field.

Lecture of February 10

Elementary matrices and unipotent matrices

We shall write $\mathcal{U}_n^{\mathrm{U}} \subseteq \mathcal{B}_n^{\mathrm{U}}$ for the subgroup consisting of upper triangular matrices such that all diagonal entries are equal to 1. This is the group of upper triangular unipotent

matrices. Similarly, $\mathcal{U}_n^{\mathrm{L}} \subseteq \mathcal{B}_n^{\mathrm{L}}$ is the subgroup consisting of lower triangular matrices with all diagonal entries equal to 1, the group of lower triangular unipotent matrices. The subscript $_n$ will often be omitted.

If $i \neq j$ are integers with $1 \leq i, j \leq n$ and $c \in K$, we denote by $E_{ij}(c)$ the matrix obtained by adding c times the j th row of the $n \times n$ identity matrix to the i th row. This matrix has all diagonal entries equal to 1, and precisely one off-diagonal entry that may be nonzero: the entry in the i th row and j th column is c. The field K and the value of n should be clear from context. For any $A \in \mathrm{GL}(n, K)$, $E_{ij}A$ is the matrix obtained from A be adding c times the j th row of A to the i th row of A. If $i < j$, then $E_{ij}(c) \in \mathcal{U}^{\mathrm{U}}$, while if $i > j$, then $E_{ij}(c) \in \mathcal{U}^{\mathrm{L}}$.

Every element A of $\mathcal{B}^{\mathrm{U}}$ is the product (on either side) of the diagonal matrix whose diagonal entries are the same as those of A and an upper triangular unipotent matrix. The upper triangular unipotent matrices are generated by the $E_{ij}(c)$ for $i < j$. Note that $E_{ij}(c)$ and $E_{ij}(-c)$ are inverses. Given any upper triangular unipotent matrix, it can be "brought to" the identity matrix by a finite sequence of elementary row operations corresponding to left multiplication by matrices $E_{ij}(c)$ with $i < j$. One subtracts multiples of the last row from earlier rows to make all entries of the last column except the bottom entry equal to 0. Then one subtracts multiples of the $n-1$ st row from the earlier rows to make all entries in the $n-1$ st column except the $n-1$ st equal to 0. Once the j th column has only one nonzero entry, which is 1, in the j th spot for all $j > i$, one subtracts multiples of the i th row from the earlier rows until all entries of the i th column are 0, except for the i th entry, which is 1. One continues in this way until off-diagonal entries are 0. This means that one can choose upper triangular matrices $E_1, \ldots, E_N$ such that

$$E_N \cdots E_1 A = I.$$

But this in turn implies that

$$A = E_1^{-1} \cdots E_N^{-1},$$

as required. It follows that $\mathcal{B}^{\mathrm{U}}$ is generated by the diagonal matrices and the matrices $E_{ij}(c)$ for $i < j$.

In an exactly similar way (or simply by transposing) we have that every element A of $\mathcal{B}^{\mathrm{L}}$ is the product (on either side) of the diagonal matrix whose diagonal entries are the same as those of A and a lower triangular unipotent matrix. The lower triangular unipotent matrices are generated by the $E_{ij}(c)$ for $i > j$, and $\mathcal{B}^{\mathrm{L}}$ is generated by the diagonal matrices and the matrices $E_{ij}(c)$ for $i > j$.

Lower triangular matrices preserve the initial form

Let $R = K[x_1, \ldots, x_n]$ and let F be a finitely generated free module with ordered basis $e_1, \ldots, e_s$. We assume a monomial order on F such that if $i < j$, then $x_i e_t > x_j e_t$ for $1 \leq t \leq s$. This means that for any term ν,

$$(\#) \quad x_i^h \nu > x_i^{h-d} x_j^d \nu$$

for $1 \leq d \leq h$ and $i < j$. Ignoring the scalar, $\nu = \mu e_t$, and

$$x_i^h \nu e_t = x_i^{h-1} \nu x_i e_t > x_i^{h-1} \nu x_j e_t = x_i^{h-1} x_j \nu e_t$$

which is the case $d = 1$. But then, by induction on d, if $d > 1$ the last term is greater than $x_i^{(h-1)-(d-1)} x_j^{d-1} x_j \nu e_i$, which yields the result.

Theorem. *Let $A \in \mathcal{B}^{\mathrm{L}}$. For every nonzero element $f \in F$, $\mathrm{in}(Af) = \mathrm{in}(f)$. Hence, for every submodule $M \subseteq F$, we have that $\mathrm{in}(AM) = \mathrm{in}(M)$.*

Proof. The second statement is clear from the first. Since A can be written as a product of diagonal matrices and matrices $E_{ij}(c)$ with $i > j$, it suffices to prove the first statement for each of the two types. If A is diagonal, it is clear that the monomials occurring in terms of Af are the same as the monomials occurring in terms of f: the action is such that each term of f is multiplied by a nonzero scalar in the field, and no new terms are introduced.

Therefore we may assume that $A = E_{ji}(c)$ with $j > i$. We consider the effect of the action of A on a typical term of f. Note that A sends $x_i \mapsto x_i + cx_j$ while fixing all the other x_k. The term can be written as $x_i^h \nu$ where ν is a term not divisible by x_i. Then A maps this term to $(x_i + cx_j)^h \nu$. When we expand we get $x_i^h \nu$ and a sum of other terms, which, if nonzero, have the form $c' x_i^{h-d} x_j^d \nu$ where $c' \in K - \{0\}$ and $1 \leq d \leq h$. Thus, the original term occurs, and the other terms are strictly smaller, by (#) displayed above. It follows that if $x_i^h \nu$ is the initial term of f, it still occurs in Af, and all other terms occurring are strictly smaller, so that it remains the initial term. $\square$

Corollary. *If U is a Zariski dense open subset of $\mathrm{GL}(n, K)$ such that $\mathrm{in}(AM)$ is $\mathrm{Gin}(M)$ for all $A \in U$, then $\mathcal{B}^{\mathrm{L}} U$ is a Zariski dense open set with the same property.*

Proof. If $\mathrm{in}(AM) = \mathrm{Gin}(M)$ and $B \in \mathcal{B}^{\mathrm{L}}$, then we have from the preceding Theorem that $\mathrm{in}\big(B(AM)\big) = \mathrm{in}(AM) = \mathrm{Gin}(M)$, from which it follows that every matrix in $\mathcal{B}^{\mathrm{L}} U = \{BA : B \in \mathcal{B}^{\mathrm{L}},\, A \in U\}$ consists entirely of matrices that map M to a module whose initial module is $\mathrm{Gin}(M)$. Multiplication by $B \in GL(n, K)$ is an automorphism of $\mathrm{GL}(n, K)$ as an algebraic set (*not* as a group), so that for all $B \in \mathcal{B}^{\mathrm{L}}$, $BU = \{BA : A \in U\}$ is again a dense open set. Since $\mathcal{B}^{\mathrm{L}} U$ is the union of the family $\{BU : B \in \mathcal{B}^{\mathrm{L}}\}$, it is also a dense open set. $\square$

We next note:

Lemma. *$\mathcal{B}^{\mathrm{L}} \mathcal{U}^{\mathrm{U}} = \{BA : B \in \mathcal{B}^{\mathrm{L}},\ A \in \mathcal{U}^{\mathrm{U}}\}$ is a Zariski dense open set in $\mathrm{GL}(n, K)$.*

Proof. For $1 \leq k \leq n$, let D_k be the polynomial function on $GL(n, K)$ given by the determinant of the $k \times k$ submatrix in the upper left corner. Let D be the product $D_1 D_2 \cdots D_{n-1}$. We claim that $\mathcal{B}^{\mathrm{L}} \mathcal{U}^{\mathrm{U}} = \mathrm{GL}(n, K)_D$, the set of invertible $n \times n$ matrices such that the nested minors in the upper left corner do not vanish. Evidently, $\mathcal{U}^{\mathrm{U}} \subseteq$

$\mathrm{GL}(n, K)_D$. Next note that if $A \in \mathrm{GL}(n, K)_D$ and B is an invertible diagonal matrix, then $BA \in \mathrm{GL}(n, K)_D$ (the relevant minors are each multiplied by a nonzero scalar) and $E_{ij}(c)B \in \mathrm{GL}(n, K)_D$ for all $i > j$ and $c \in K$ (adding a multiple of an earlier row to later row does not change any of the relevant minors). Hence, $\mathcal{B}^{\mathrm{L}}\mathcal{U}^{\mathrm{U}} \subseteq \mathrm{GL}(n, K)_D$.

It remains to prove the opposite inclusion. Now consider any matrix $A \in \mathrm{GL}(n, K)_D$. By the hypothesis on the nonvanishing of D, we have that D_1 does not vanish, i.e., the entry in the upper left hand corner is not 0. Hence, we can subtract multiples of the first row from lower rows to obtain a matrix in which the first column is 0 below the first entry. In the course of this process, at each stage we are multiplying by a lower triangular elementary matrix. We can proceed by, indiuction on j, to multiply by lower triangular elementary matrices unitl we reach a matrix such that all entries below the main diagonal in the first j columns are 0. At every stage, we continue to have a matrix in $\mathrm{GL}(n, K)_D$. Suppose this has been done for all columns preceding the j th column. The hypothesis that D does not vanish implies that D_j does not vanish, and since the $j \times j$ submatrix in the upper left corner is now upper triangular, this implies that the j, j entry on the diagonal is nonzero. We can therefore subtract multiples of the j th row from lower rows until the j th column contains only 0 entries below the main diagonal. In this way, we eventually reach an upper triangular matrix. We have multiplied the original matrix A on the left by a lower triangular unipotent matrix B in the process, thereby obtaining an upper triangular matrix C. Since $BA = C$, we have $A = B^{-1}C$, as required. $\square$

Corollary. *Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over an infinite field K, let F be a finitely generated free R-module with ordered basis $e_1, \ldots, e_s$, and suppose that we have a monomial order on F such that for all t and $i < j$, $x_i e_t > x_j e_t$. Let $M \subseteq F$ be a submodule. Let $U \subseteq \mathrm{GL}(n, K)$ be such that $\mathcal{B}^{\mathrm{L}}U \subseteq U$ and $\mathrm{in}(AM) = \mathrm{Gin}(M)$ for all $A \in U$. Then U has nonempty intersection with $\mathcal{U}^{\mathrm{U}}$.*

Proof. Since U and $\mathcal{B}^{\mathrm{L}}\mathcal{U}^{\mathrm{U}}$ are Zariski dense open sets, their intersection is nonempty. Choose $A \in U$ such that $A = BC$ with $B \in \mathcal{B}^{\mathrm{L}}$ and $C \in \mathcal{U}^{\mathrm{U}}$. Then $C = B^{-1}A \in U$, as required. $\square$

Ideals stable under the action of the group of invertible diagonal matrices

We want to show that when K is infinite, an ideal I of the polynomial ring $R = K[x_1, \ldots, x_n]$ over a field K is stable under the action of $\mathcal{D}_n$, i.e., mapped into itself by every element of $\mathcal{D}_n$, if and only if it is a monomial ideal.

We shall prove some much stronger results. We first want to prove a result on the invertibility of Van der Monde matrices.

Discussion: Van der Monde matrices. Let $u_1, \ldots, u_h$ be elements of a commutative ring. Let Q be the $h \times h$ matrix $\left(u_i^{j-1}\right)$, which is called a *Van der Monde* matrix. We want to show that if the elements $u_i - u_j$ are all invertible, then so is Q. We give two proofs.

(a) We shall show that the determinant of Q is $\prod_{j>i}(u_j - u_i)$. Hence, Q is invertible if $u_j - u_i$ is a unit for $j > i$. It suffices to prove the first statement when the u_i are indeterminates over $\mathbb{Z}$. Call the determinant D. If we set $u_j = u_i$, then D vanishes because two rows become equal. Thus, $u_j - u_i$ divides D in $\mathbb{Z}[u_1, \ldots, u_h]$. Since the polynomial ring is a UFD and these are relatively prime in pairs, the product P of the $u_j - u_i$ divides D. But they both have degree $1 + 2 + \cdots + h - 1$. Hence, $D = cP$ for some integer c. The monomial $u_2 u_3^2 \cdots u_h^{h-1}$ obtained from the main diagonal of matrix in taking the determinant occurs with coefficient 1 in both P and D, so that $c = 1$. $\square$

(b) We can also show the invertibility of Q as follows: if the determinant is not a unit, it is contained in a maximal ideal. We can kill the maximal ideal. We may therefore assume that the ring is a field K, and the u_i are mutually distinct elements of this field. If the matrix is not invertible, there a nontrivial relation on the columns with coefficients $c_0, \ldots, c_{n-1}$ in the field. This implies that the nonzero polynomial

$$c_{h-1}x^{h-1} + \cdots + c_1 x + c_0$$

has h distinct roots, $u_1, \ldots, u_h$, in the field K, a contradiction. $\square$

Next note the following. Suppose that R is an $\mathbb{N}$- or $\mathbb{Z}$-graded algebra and that $u \in R_0$ is a unit. The there is an automorphism $\eta_u : R \to R$ such that if $f \in [R]_d$, then $\eta_u(f) = u^d f$.

Proposition. *Let R be an $\mathbb{N}$- or $\mathbb{Z}$-graded algebra such that R_0 contains an infinite field or, more generally, such that R_0 contains infinitely many elements u_i that are units and such that for all $i \neq j$, the element $u_i - u_j$ is a unit. Let $I \subseteq R$ be any ideal that is stable under all of the automorphism η_{u_i}, with notation as just above. Then I is a homogeneous ideal of R.*

Proof. Let $f_{t+1} + \cdots + f_{t+h} = f$ be an element of I, where the interval $[t+1, \ldots, t+h]$ includes all degrees in which the element has a nonzero homogeneous component, and f_j denotes the homogeneous component in degree j. Choose invertible elements $u_1, \ldots, u_h$ in R_0 such that $u_i - u_j$ is invertible for $i \neq j$. By letting η_{u_i} act we obtain an equation

$$(*_i) \quad u_i^{t+1} f_{t+1} + \cdots + u_i^{t+j} f_{t+j} + \cdots + u_i^{t+h} f_{t+h} = \eta_{u_i}(f) \in I$$

We can multiply this equation by u_i^{-t-1} and let $g_i = u_i^{-t-1}\eta_{u_i}(f) \in I$ to obtain

$$(**_i) \quad f_{t+1} + \cdots + u_i^{j-1} f_{t+j} + \cdots + u_i^{h-1} f_{t+h} = g_i$$

for $1 \leq i \leq h$. In matrix form, these equations can be written as

$$Q \begin{pmatrix} f_{t+1} \\ \vdots \\ f_{t+h} \end{pmatrix} = \begin{pmatrix} g_1 \\ \vdots \\ g_h \end{pmatrix},$$

where Q is the Van der Monde matrix discussed above, and so is invertible over R. We then have

$$\begin{pmatrix} f_{t+1} \\ \vdots \\ f_{t+h} \end{pmatrix} = Q^{-1} \begin{pmatrix} g_1 \\ \vdots \\ g_h \end{pmatrix},$$

and since Q^{-1} has entries in R and the $g_j \in I$, it follows that all of the homogeneous components of f are in I, as required. $\square$

Corollary. *Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over an infinite field K, and let I be an ideal of R that is stable under the action of the diagonal matrices $\mathcal{D}_n \subseteq \mathrm{GL}(n, K)$. Then I is a monomial ideal of R.*

Proof. Let S_k be the polynomial ring in the remaining variables with x_k omitted for $1 \leq k \leq n$, so that $R = S_k[x_k]$. Then R is $\mathbb{N}$-graded thinking of it as a polynomial ring in one variable over S_k, with $[R]_d = S_k x_k^d$ for every $d \in \mathbb{N}$, and $K \subseteq R_0 = S_k$. If $u \in K - \{0\}$, the automorphism η_u conincides with the action of the diagonal matrix with u on the diagonal in the k, k spot and all other entries equal to 1 on R, and so I is stable with respect to this action. Hence, I is homogeneous with respect to each of the x_k gradings. Given an element f of I, it is a sum of x_n-homogeneous componets all of which are in I: these have the form $g_{n-1}x_n^d$ where $g_{n-1} \in K[x_1, \ldots, x_{n-1}]$. Each of these is in turn a sum of x_{n-1}-homogeneous components, all of which are in I. These have the form $g_{n-2}x_{n-1}^{d_{n-1}}x_n^{d_n}$ where $g_{n-2} \in K[x_1, \ldots, x_{n-2}]$. Continuing in this way, we see that every monomial term of f is in I, as required. $\square$

Borel-fixed ideals

We shall show soon that in the graded case, generic initial monomial ideals are stable under the action of $\mathcal{B}_n^{\mathrm{U}}$. In this section we want to charactize the ideals of the polynomial ring $R = K[x_1, \ldots, x_n]$ that are stable under the action of $\mathcal{B}_n^{\mathrm{U}}$.

We first prove an elementary fact about the behavior of binomial coefficients modulo a prime integer p that we shall need to handle the characteristic $p > 0$ case.

In the Lemma below, the binomial coefficients $\binom{k}{h}$, where h, $k \in \mathbb{N}$, are defined to be 0 if $h > k$. Otherwise, they have their usual meaning, $\dfrac{k!}{h!(k-h)!}$. Note that $\binom{k}{h}$ is always nonzero if $0 \leq h \leq k$: in particular, if $h = 0$ its value is 1, even if $k = 0$.

Lemma. *Let h and k be nonnegative integers and let p be a positive prime integer. Let*

$$h = h_d p^d + h_{d-1}p^{d-1} + \cdots + h_0$$

and

$$k = k_d p^d + k_{d-1}p^{d-1} + \cdots + k_0$$

be expansions of h and k respectively in base p, so that $0 \le h_i \le p-1$ and $0 \le k_i \le p-1$ for all i. (The length d of the expansion is permitted to be longer than needed, so that, for example, h_d, or several of the initial h_i, may be 0, and the same holds for k.) Then

$$\binom{k}{h} \equiv \binom{k_d}{h_d}\binom{k_{d-1}}{h_{d-1}} \cdots \binom{k_1}{h_1}\binom{k_0}{h_0} \mod p.$$

Hence, $\binom{k}{h} \not\equiv 0 \mod p$ if and only if $h_i \le k_i$ for all i.

Proof. Let z be an indeterminate over $\mathbb{Z}_p = \mathbb{Z}/pZ$. Then

$$(1+z)^k = \big((1+z)^{p^d}\big)^{k_d}\big((1+z)^{p^{d-1}}\big)^{k_{d-1}} \cdots \big((1+z)^p\big)^{k_1}(1+z)^{k_0},$$

and since we are in prime characteristic $p > 0$ we may rewrite this as

$$(1+z^{p^d})^{k_d}(1+z^{p^{d-1}})^{k_{d-1}} \cdots (1+z^p)^{k_1}(1+z)^{k_0}.$$

If we expand each factor by the binomial theorem and then multiply out, using the generalized distributive law, we obtain the sum of $(k_d+1)(k_{d-1}+1)\cdots(k_0+1)$ terms, one for every choice of integers $h_d, \ldots, h_0$ with $0 \le h_i \le k_i$, namely:

$$\binom{k_d}{h_d}\binom{k_{d-1}}{h_{d-1}} \cdots \binom{k_1}{h_1}\binom{k_0}{h_0}(z^{p^d})^{h_d}(z^{p^{d-1}})^{h_{d-1}} \cdots (z^p)^{h_1} z^{h_0} =$$

$$\binom{k_d}{h_d}\binom{k_{d-1}}{h_{d-1}} \cdots \binom{k_1}{h_1}\binom{k_0}{h_0} z^{h_d p^d + h_{d-1}p^{d-1} + \cdots + h_1 p + h_0}.$$

Because the exponents are distinct expansions of nonnegative integers in base p, they are all distinct, and there are no cancellations of terms. These coefficients are all nonzero, because p does not occur as factor in the formula for the binomial coefficient $\binom{k_i}{h_i}$ when $0 \le h_i \le k_i \le p-1$. There is no nononzero term involving z^h if the expansion of h in base p is such that $h_i > k_i$ for some i, and the formula given remains correct in this case because $\binom{k_i}{h_i} = 0$ when $h_i > k_i$. The final statement is now clear. $\square$

For each integer p in the set $\{0, 2, 3, 5, \ldots\}$ consisting of 0 and the positive prime integers, if $h, k \in \mathbb{N}$ we define $h \le_p k$ to mean $\binom{k}{h}$ does not vanish modulo p. If $p = 0$ this is the usual total order on $\mathbb{N}$, but if $p > 0$ it is a partial ordering because of the characterization in the last statement of the Lemma just above.

Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring in n variables over a field. We define an ideal of R to be *Borel-fixed* if it is stable under the action of $\mathcal{B}_n^{\mathrm{U}}$, the Borel subgroup of $\mathrm{GL}(n, K)$ consisting of upper triangular matrices. Such an ideal is stable under the action of $\mathcal{D}_n$, and so it must be a monomial ideal. We have the following:

Proposition. *Let notation be as in the paragraph above, and let $I \subseteq R$. Then I is Borel-fixed if and only if it is a monomial ideal and has a set of monomial generators μ with the following property:*

(#) if $\mu = x_j^k \nu$ where $x_j \nmid \nu$, then $x_i^h x_j^{k-h} \nu \in I$ for all $i < j$ and all h such $h \leq_p k$.

If I is Borel-fixed, condition (#) is satisfied by every monomial $\mu \in I$.

Proof. I is stable under the action $\mathcal{D}$ if and only if it is monomial, and a monomial ideal I is Borel-fixed if and only if every $E_{ij}(c)$, $c \in K - \{0\}$ and $i < j$, maps I into itself, since the diagonal matrices together with the $E_{ij}(c)$ for $i < j$ generate $\mathcal{B}^{\mathrm{U}}$. It is sufficient that every μ in a set of monomial generators for I map into I, and it is necessary that every $\mu \in I$ map into I. But given μ, its image under the map that sends $x_j \mapsto cx_i + x_j$ while the other variables are fixed is

$$(cx_i + x_j)^k \nu = \sum_{0 \leq h \leq_p k} \binom{k}{h} c^k x_i^h x_j^{k-h} \nu,$$

since the integers h satisfying $0 \leq h \leq_p k$ are precisely the ones that yield a nonzero binomial coefficient. The stated result is now immediate. □

Lecture of February 12

We next want to consider one example where the generic initial ideal depends on the characteristic. The example also illustrates that, even when the given ideal is monomial, the generic initial ideal can be rather different.

Consider $I = (x_1^2, x_2^2)$ in $R = K[x_1, x_2]$ where K is infinite. Suppose that we use either hlex or revlex as the monomial order. If $A = \big(a_{ij}\big)$,

$$\mathrm{Gin}(I) = \mathrm{in}\big(((a_{11}x_1 + a_{21}x_2)^2, (a_{12}x_1 + a_{22}x_2)^2)R\big)$$

for a_{11}, a_{12}, a_{21}, a_{22} in sufficiently general position. In characteristic different from 2, we get x_{11}^2 as the initial term from either generator. The initial term of

$$a_{12}^2(a_{11}x_1 + a_{21}x_2)^2 - a_{11}^2(a_{12}x_1 + a_{22}x_2)^2$$

yields an x_1x_2 term. In degree $d \geq 3$, I contains all monomials of degree d, and, hence, so does AI. It follows that $\mathrm{Gin}(I) = (x_1^2, x_1x_2, x_2^3)$. However, in characteristic two, both squares are linear combinations of x_1^2 and x_2^2, and $\mathrm{Gin}(I) = (x_1^2, x_2^2)$. This is consistent with out characterization of Borel-fixed ideals because it is false that $1 \leq_2 2$: the binomial coefficient $\binom{2}{1}$ vanishes modulo 2.

The following result explains in part why generic initial ideals have great interest.

Theorem. *Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over a field K, and let I be a homogeneous ideal of R. Let m be the homogeneous maximal ideal of R. Then* $\operatorname{depth}_m(R/I) = k$ *if and only if the minimal monomial generators of* $\operatorname{Gin}(R/I)$ *for revlex involve x_{n-k} but not x_j for $j \geq n-k+1$.*

Proof. After the variables are placed in general position, say by a change of coordinates using a matrix of indeterminates, there is a regular sequence of length k on R/I if and only if the last k variables form such a sequence: see the Proposition on p. 3 of the Lecture Notes of February 8. By our results on reverse lexicographic order, this is equivalent to the absence of the last k variables from the initial ideal with respect to revlex: see the final Theorem of the Lecture of February 3. $\square$

Actions on vector spaces and exterior algebra

We are aiming to prove results concerning when the initial ideal is Borel-fixed. The theorems we obtain can actually be viewed as results about actions on on finite-dimensional K-vector subspaces of R.

We assume that $R = K[x_1, \ldots, x_n]$, a polynomial ring over an infinite field K, and we fix a monomial order on R such that $x_1 > x_2 > \cdots > x_n$, as usual.

This gives an ordered basis for every subspace of R spanned by monomials. Recall that when a vector space V has an ordered basis $v_1 > v_2 \cdots > v_h$, the theory of Gröbner bases applies directly to V: the base ring may be thought of as K, the polynomials in 0 variables over the field K. When V is a subspace of R, this gives us, *a priori*, two notions of initial term. We write $\operatorname{in}_{\text{vec}}(f)$ to indicate that we are taking the initial term in a vector space sense. However, in practice, we shall frequently be considering a finite-dimensional K-vector subspace of R spanned by monomials, with the order of the basis elements obtained by restricting the monomial order on R. In this case, $\operatorname{in}(f)$ and $\operatorname{in}_{\text{vec}}(f)$ agree. However, $\operatorname{in}_{\text{vec}}(W)$ is a K-vector subspace of V, not an ideal of R.

Recall that the exterior algebra $\bigwedge^{\bullet}(V) = \bigoplus_{k \in \mathbb{N}} \bigwedge^k(V)$ of a K-vector space V is an $\mathbb{N}$-graded associative algebra generated over $K = \bigwedge^0(V)$ by $V = \bigwedge^1(V)$ with multiplication denoted $\wedge$ satisfying precisely those relations implied by the condition that $v \wedge v = 0$ for every element $v \in V$. Then

$$0 = (v+w) \wedge (v+w) = v \wedge v + v \wedge w + w \wedge v + w \wedge w = v \wedge w + w \wedge v,$$

so that

$$v \wedge w = -w \wedge v$$

for all $v, w \in V$. This implies that if $\{v_j\}_{j \in \mathcal{J}}$ is an ordered basis for V then the elements $v_{j_1} \wedge v_{j_2} \wedge \cdots \wedge v_{j_i}$ such that $v_{j_1} > v_{j_2} > \cdots > v_{j_i}$ form a basis for $\bigwedge^i(V)$. In particular, it follows that if $\dim_K(V) = k$, then

$$\dim_K\left(\bigwedge^i(V)\right) = \binom{k}{i}, \ 0 \leq i \leq k,$$

while $\bigwedge^i(V) = 0$ for $i > \dim(V)$.

Moreover, for any elements $v_1, \ldots, v_k \in V$, we have that for every permutation π of $\{1, \ldots, k\}$,

$$v_{\pi(1)} \wedge \cdots \wedge v_{\pi(k)} = \operatorname{sgn}(\pi)(v_1 \wedge \cdots \wedge v_k),$$

where $\operatorname{sgn}(\pi)$ is the sign of the permutation π. We also have that $v_1 \wedge \cdots \wedge v_k = 0$ if and only if $v_1, \ldots, v_k$ are linearly dependent over K. We know that if $v_1, \ldots, v_k$ is a basis for V, then the single element $v_1 \wedge \cdots \wedge v_k$ is a basis for $\bigwedge^k(V)$, which is a one-dimensional space. If we consider k linear combinations of $v_1, \ldots, v_k$, say

$$w_i = c_{i1}v_1 + \cdots + c_{ik}v_k$$

for $1 \leq i \leq k$, with the elements $c_{ij} \in K$, then

$$w_1 \wedge \cdots \wedge w_k = \det(c_{ij})\, v_1 \wedge \cdots \wedge v_k,$$

which will be another generator of $\bigwedge^k(V)$ precisely when $w_1, \ldots, w_k$ is a basis for V.

Note that forms of even degree in $\bigwedge^\bullet(V)$ are in the center, while if w, w' are forms of odd degree, $w \wedge w' = -w' \wedge w$. Notice also that, by definition, $\bigwedge^0(V) = K$. An $\mathbb{N}$-graded associative algebra such that for any two nonzero forms w, w' of degrees d, d' respectively, $ww' = (-1)^{dd'}w'w$ is called a *skew-commutative graded algebra*. (Some call such graded algebras *commutative*, but we shall not do this.)

If $T : V \to W$ is a K-linear map, it extends uniquely to a degree preserving K-homomorphism of $\mathbb{N}$-graded associative algebras $\bigwedge^\bullet(T) : \bigwedge^\bullet(V) \to \bigwedge^\bullet(W)$. This makes $\bigwedge^\bullet(\_)$ into a covariant functor from K-vector spaces and K-linear maps to skew-commutative graded K-algebras and degree-preserving K-algebra homomorphisms. In particular, we have functorial maps $\bigwedge^i(T) : \bigwedge^i(V) \to \bigwedge^i(W)$ for every $i \in \mathbb{N}$. Observe also that

$$T(v_1 \wedge \cdots \wedge v_k) = T(v_1) \wedge \cdots \wedge T(v_k).$$

If $W \subseteq V$ is a k-dimensional vector space, then $\bigwedge^k(W) \subseteq \bigwedge^k(V)$ is a one-dimensional subspace of $\bigwedge^k(V)$. This one dimensional subspace uniquely determines W, since if $\bigwedge^k(W) = Kw$ then $W = \{v \in V : w \wedge v = 0\}$.

Given an ordered basis for V, we introduce an order on the basis for $\bigwedge^k V$ mentioned above. A typical element of the basis for $\bigwedge^k V$ has the form $v_1 \wedge \cdots \wedge v_k$ where $v_1, \ldots, v_k$ are in the given ordered basis for V and are such that $v_1 > \cdots > v_k$. The ordering is given by the following rule: if $v_1 \wedge \cdots \wedge v_k$ and $w_1 \wedge \cdots \wedge w_k$ are in this basis with $v_1 > \cdots > v_k$ and $w_1 > \cdots w_k$, we define $v_1 \wedge \cdots \wedge v_k > w_1 \wedge \cdots \wedge w_k$ to mean that there exists i, $1 \leq i \leq k$, such that $v_j = w_j$ for $j < i$ and $v_i > w_i$. This ordering resembles lexicographic ordering of monomials.

Remark. Suppose that $v_1, \ldots, v_k$ are distinct elements of the ordered basis, not necessarily in decreasing order, and that $w_1, \ldots, w_k$ are distinct elements of the ordered basis, also

not necessarily in decreasing order. Suppose that for every i, $(*)$ $\quad v_i \geq w_i$. Then this condition also holds when both sequences are arranged in decreasing order. The reason is simply this: let $v_1', \ldots, v_k'$ and $w_1', \ldots, w_k'$ denote the sequences arranged in decreasing order. For every i, each of the elements $w_1', \ldots, w_i'$ is less than some element of $v_1, \ldots, v_k$ coming from the inequalities $(*)$, where these i elements are mutually distinct. Then w_i is less than or equal to each of these i distinct elements of the $v_1, \ldots, v_k$. The smallest of these i elements is evidently at most v_i'. $\square$

We then have:

Proposition. *Let V be a vector space with ordered basis, and let W be a subspace of dimension k. Then a reduced Gröbner basis for W is a basis for W, and given such a basis $w_1, \ldots, w_k$, we have that $\mathrm{in}_{\mathrm{vec}}(w_1) \wedge \cdots \wedge \mathrm{in}_{\mathrm{vec}}(w_k)$ is the initial term of a generator for $\mathrm{in}_{\mathrm{vec}}\bigl(\bigwedge^k(W)\bigr)$.*

Proof. Fix sufficiently many elements $v_1 > \cdots > v_s$ of the ordered basis for V so that W is contained in their span. We have already noted in the Lecture Notes of January 15 that the condition for $w_1, \ldots, w_k$ to be a reduced Gröbner basis is that when each w_i is written in terms of $v_1, \ldots, v_s$, the coefficients used, formed into the rows of a $k \times s$ matrix, produce a reduced row echelon matrix without any rows that are 0. (If any v_i are not actually used, they contribute columns that are entirely zero, and do not affect whether the matrix one obtains is in reduced row echelon form.) The leading entries of the rows correspond to the intial terms of the w_i. It is now clear from the Remark above that when we form $\omega_1 \wedge \cdots \wedge w_k$, the initial term is obtained by forming the product, under $\wedge$, of the initial terms. $\square$

We next observe that we can define a generic vector space of initial forms, $\mathrm{Gin}_{\mathrm{vec}}(W)$ when W is a k-dimensional subspace of V.

Theorem. *Fix a monomial order on $R = K[x_1, \ldots, x_n]$: this yields an ordered basis for $\bigwedge^k(R)$ for all $k \in \mathbb{N}$. Let $W \subseteq R = K[x_1, \ldots, x_n]$ be a given subspace of finite dimension. There is a Zariski open dense subset U of $\mathrm{GL}(n, K)$ such that $\mathrm{in}_{\mathrm{vec}}(AW)$ is the same for all $A \in U$. U may chosen so that $\mathcal{B}_n^{\mathrm{L}} U = U$. If g generates $\bigwedge^k(W)$ as a K-vector space, then $\mathrm{Gin}_{\mathrm{vec}}(g)$ is the greatest term occurring in Ag for any $A \in \mathrm{GL}(n, K)$.*

Proof. Let $Z = \bigl(z_{ij}\bigr)$ be a matrix of new indeterminates. Exactly as in our earlier proof of the existence of generic initial modules, we may consider ZW over $K(Z)$ and construct the reduced Gröbner basis, in the $K(Z)$-vector space sense, there, keeping track of finitely many polynomials in the z_{ij} that are used in denominators and also finitely many polynomials in the z_{ij} that occur as numerators of coefficients of initial terms. We may form the product P of these polynomials in the z_{ij}, and then we may take the set where P does not vanish as a choice of U. Applying a matrix in $\mathcal{B}^{\mathrm{L}}$ does not change the initial term of an element, and hence $\mathcal{B}_n^{\mathrm{L}} U$ is a larger dense open set for which every matrix yields the same initial vector space. Finally, note that a term that occurs in some Ag will occur in Zg. Since

the leading term of Zg gives $\mathrm{Gin}_{\mathrm{vec}}(g)$, and is greater than any other term in Zg, the final statement follows. $\square$

We next want to characterize when a K-vector subspace of R is stable under the action of the diagonal matrices $\mathcal{D}_n$, and when such a subspace is stable under the action of the $\mathcal{B}_n^{\mathrm{U}}$. We first note that over an infinite field K, we have a graded vector space analogue of the Proposition on p. 4 of the Lecture Notes of February 10.

First note that if V is an $\mathbb{N}$ or $\mathbb{Z}$-graded vector space over an infinite field K, then we may define an automorphism η_u of V for each nonzero $u \in K$ such that each element $v \in [V]_d$ maps to $u^d v$.

Proposition. *Let V be an $\mathbb{N}$- or $\mathbb{Z}$-graded vector space over an infinite field K. Then every subspace W of V stable under all the η_u for $u \in K - \{0\}$ is graded.*

Proof. The proof is identical with the proof given for the earlier Proposition: the van Der Monde matrix Q now has entries in K, and W replaces the ideal I. $\square$

Theorem. *Let R be a polynomial ring $K[x_1, \ldots, x_n]$, where K is a field. Let W be a K-vector subspace of R. Then W is stable under the action of $\mathcal{D}_n$ if and only if W is spanned by monomials.*

Suppose that K has characteristic p (which may be 0). Then W is stable under the action of $\mathcal{B}_n^{\mathrm{U}}$ if and only if it is spanned by monomials and for every monomial $\mu \in W$, if $\mu = x_j^k \nu$, where x_j does not divide μ, $h \leq_p k$, and $i \leq j$ then $x_i^h x_j^{k-h} \nu \in W$.

Proof. The proof of the first statement is identical with the proof of the Corollary on p. 5 of the Lecture Notes of February 10, using the Proposition above, and the proof of the second statement is identical with the proof of the Proposition on p. 7 of the Lecture Notes of February 10. $\square$

We refer to the subspaces of R stable under $\mathcal{B}_n^{\mathrm{U}}$ as *Borel-fixed*.

We next note that there is a monomial grading of $\bigwedge^\bullet(R)$. If one has terms $v_1, \ldots, v_k$ involving mutually distinct monomials $\mu_1, \ldots, \mu_k$, we define the *monomial degree* of the element $v_1 \wedge \cdots \wedge v_k$ to be the product $\mu_1 \cdots \mu_k$. If the μ_i are not mutually distinct, then $v_1 \wedge \cdots \wedge v_k = 0$. Let$[\bigwedge^\bullet(R)]_\mu$ denote the K-span of all basis elements whose monomial degree is μ. Then we have a direct sum decomposition

$$\bigwedge^\bullet(R) = \bigoplus_{\mu \in \mathcal{M}} [\bigwedge^\bullet(R)]_\mu,$$

where $\mathcal{M}$ is the set of monomials in R. Note that if $R = K[x_1, x_2]$ and $\mu = x_1^3 x_2^3$, then $[\bigwedge^\bullet(R)]_\mu$ contains $x_1^3 x_2^3$, $x_1^2 \wedge x_1 x_2^3$, and $x_1^2 \wedge x_1 x_2 \wedge x_2^2$, as well as many other elements.

Remark. A critical observation is the following: if W is a k-dimensional K-vector subspace of R and w generates $\bigwedge^k(W)$, then the monomial degree of $\mathrm{in}(w)$ is strictly larger than the monomial degree of any other term of w. Consider a Gröbner basis for W as a vector space:

in each element, the initial monomial is strictly larger than any other monomial occurring. The product of the initial monomials is therefore strictly larger than the product of any other choice of monomials, one from each factor, from which the assertion follows.

The action of $\mathcal{D}_n$ on R induces an action on $\bigwedge^\bullet(R)$. Note that if $\beta = (b_1, \dots, b_n) \in (K - \{0\})^n$, and $\mathrm{diag}(b_1, \dots, b_n)$ is the diagonal matrix with b_i in the i, i position on the main diagonal for $1 \leq i \leq n$, then for any element $v \in [\bigwedge^\bullet(R)]_\mu$, we have that $Bv = \mu(\beta)v$, where $\mu(\beta)$ denotes the result of substituting $x_1 = b_1, \dots, x_n = b_n$ in μ.

Theorem. *Let $W \subseteq R$ be any finite dimensional vector space. Let Then $\mathrm{Gin}_{\mathrm{vec}}(W)$ is Borel-fixed.*

Proof. First replace W by AW such that $\mathrm{in}_{\mathrm{vec}}(AW) = \mathrm{Gin}_{\mathrm{vec}}(W)$. Let w generate the one-dimensional vector space $\bigwedge^k(W)$. It suffices to show that for every upper triangular elementary matrix $E = E_{ij}(c)$, $E\big(\mathrm{in}_{\mathrm{vec}}(w)\big) = \mathrm{in}_{\mathrm{vec}}(w)$. The action of E on a monomial term produces a linear combination of monomial terms one of which is the original term, while the others are strictly larger — if the term has the form $x_j^k\nu$ where ν is a term not divisible by x_j, this this follows from the expansion

$$(cx_i + x_j)^k\nu = \sum_{0 \leq h \leq_p k} \binom{k}{h} c^k x_i^h x_j^{k-h}\nu,$$

which was used in the proof of the Proposition on p. 7 of the Lecture Notes of February 10. It follows from the Remark on p. 3 that if $E\big(\mathrm{in}_{\mathrm{vec}}(w)\big) \neq \mathrm{in}_{\mathrm{vec}}(w)$, all of its nonzero terms other than $\mathrm{in}_{\mathrm{vec}}(w)$ are larger than $\mathrm{in}_{\mathrm{vec}}(w)$. Pick one such term τ. It suffices to show that τ survives in $EB(w)$ for some upper triangular matrix $B = \mathrm{diag}(b_1, \dots, b_n)$ where $\beta = (b_1, \dots, b_n) \in (K - \{0\})^n$. Let μ be the monomial degree of $\mathrm{in}(w)$. By the Remark just above, the monomial degree of every other term in w is strictly smaller than μ, so that

$$w = \mathrm{in}(w) + \sum_{\nu < \mu} w_\nu.$$

Then

$$EB(w) = E(B\mathrm{in}(w)) + \sum_{\nu<\mu} E(B(w_{nu}) = \mu(\beta)E\big(\mathrm{in}(w)\big) + \sum_{\nu<\mu} \nu(\beta)E(w_{nu}).$$

Consider the coefficient of τ in the final expression on the right as a function of β. The first sumand makes a contribution $\mu(\beta)$ to this coefficient. The other contributions to the sum have the form $c_\nu\nu(\beta)$ for $\nu < \mu$. It follows that the coefficient of τ is a nonzero polynomial in β, since the $\mu(\beta)$ term cannot be canceled. Hence τ occurs in $EB(w)$ for some choice of B, which contradicts the last statement in the Theorem stated on the bottom of p. 3 and the top of p. 4. $\square$

Lecture of February 15

We postpone further consideration of Gröbner bases to study some results in invariant theory.

To keep prerequisites from algebraic geometry to a minimum, in our study we will take the ground field K to be an algebraically closed field. For the kinds of results that we will be considering, this is no disadvantage: typically, one can deduce results over any infinite field by passing to the algebraic closure.

Linear algebraic groups and their modules

We have seen that $\mathrm{GL}(n, K)$ has the structure of a closed algebraic set, and that the same is true for the $\mathrm{GL}_n(V)$, the group of K-automorphisms of a finite-dimensional vector space V. See pages 1. and 2. of the Lecture of February 1. One gives $\mathrm{GL}_n(V)$ the structure of a closed algebraic set by choosing a basis for V. If $\dim(V) = n$, this gives an identification of V with $\mathrm{GL}(n, K)$. However, the structure of V as an algebraic set is independent of the choice of basis: if one takes a different basis, the identification of $GL(n, K)$ with V changes, but this is via an automorphism of $\mathrm{GL}(n, K)$ given by conjugating by the change of basis matrix. This map is not only a group automorphism: it is also an automorphism in the category of closed algebraic sets.

A *linear algebraic group* G is a Zariski closed subgroup of some $\mathrm{GL}(n, K)$. Thus, G has the structure of closed algebraic set.

The product of two closed algebraic sets has the structure of a closed algebraic set. If $X = V(I)$ where $I \subseteq K[x_1, \ldots, x_m]$, so that $X \subseteq \mathbb{A}^m_K$, and $Y = V(J)$ where $J = K[y_1, \ldots, y_n]$, so that $Y \subseteq \mathbb{A}^n_K$ (the variables are taken to be $m + n$ algebraically independent elements) then $X \times Y$ may be identified with $V(IT + JT) \subseteq \mathbb{A}^{m+n}_K$, where $T = K[x_1, \ldots, x_m, y_1, \ldots, y_n]$.

It is easy to show that if G is a linear algebraic group, then the map $G \times G \to G$ that corresponds to the group multiplication is regular, as well as the inverse map $G \to G$: this follows from the fact that this is true when $G = \mathrm{GL}(n, K)$.

An *action* of a linear algebraic group G on a finite-dimensional vector space V is then a group action $G \times V \to V$ such that the defining map is a morphism of closed algebraic sets, i.e., a regular map over K. The image of (γ, v) is denoted $\gamma(v)$. Alternatively, it is given by a homomorphism $h : G \to GL_K(V)$: the action is recovered by the rule $\gamma(v) = h(\gamma)(v)$. We then say that V is G-module (over K, but usually we do not mention the field K).

If $W \subseteq V$ is a K-vector subspace such that W is stable under the action of G, the restriction of the map $G \times V \to V$ gives W the structure of a G-module, and we shall say that W is a *G-submodule* of V.

We extend the notion of G-module to infinite-dimensional K-vector spaces as follows: an action of G on an infinite-dimensional vector space V is allowed if V is a directed union of finite-dimensional spaces W such that the restricted action makes W into a G-module.

The direct sum of G-modules becomes a G-module in an obvious way. A G-stable subspace of an infinite-dimensional G-module is again a G-module. If V is a G-module and $W \subseteq V$, then V/W has the structure of G-module such that for all $\gamma \in G$ and $v \in V$, $\gamma(v + W) = \gamma(v) + W$.

A G-module map $f : V \to W$ is a K-linear map such that for all $\gamma \in G$ and $v \in V$, $f\big(\gamma(v)\big) = \gamma\big(f(v)\big)$. The inclusion of a G-submodule $W \subseteq V$ is a G-module map, as is the quotient map $V \twoheadrightarrow V/W$.

A nonzero G-module M is called *irreducible* or *simple* if it has no nonzero proper submodule. If M is irreducible it is necessarily finite-dimensional, as it is a directed union of finite-dimensional G-submodules.

A linear algebraic group is called *linearly reductive* if every finite-dimensional G-module is a direct sum of irreducible G-modules. Over an field, the finite groups G such that the order of G is invertible in the field are linearly reductive, and so is an algebraic torus, i.e., a finite product of copies of $GL(1, K)$. In characteristic $p > 0$, these are the main examples. But over $\mathbb{C}$ the semisimple groups are linearly reductive as well. We shall comment further about this later.

Linearly reductive linear algebraic groups

Theorem. *Let G be a linearly reductive linear algebraic group and let $W \subseteq V$ be G-modules. Then there is a family of irreducible submodules $\{M_\lambda\}_{\lambda\in\Lambda}$ in V such that*

$$V = W + \sum_{\lambda\in\Lambda} M_\lambda$$

and the sum is direct. Hence, if

$$W' = \sum_{\lambda\in\Lambda} M_\lambda,$$

then $V = W \oplus W'$, so that W' is a G-module complement for W in V.

In particular, we may take $W = 0$, and so V itself is a direct sum of irreducible submodules, even if it is infinite-dimensional.

Proof. Consider the set of families of irreducible submodues

$$\{M_\lambda\}_{\lambda\in\Lambda}$$

of V such that the sum

$$W + \sum_{\lambda\in\Lambda} M_\lambda$$

is direct, i.e., such that every module occurring has intersection 0 with the sum of the other modules occurring. The empty set is such a family, and the union of chain of such families

is such a family. Hence, there is a maximal such family, which we denote $\{M_\lambda\}_{\lambda\in\Lambda}$. We claim that $V = V'$, where

$$V' = W + \sum_{\lambda\in\Lambda} M_\lambda.$$

If not, there is a finite-dimensional submodule V_0 of V that is not contained in V'. V_0 is a direct sum of irreducibles: one of these, call it M_0, must also fail to be contained in V'. Then $M_0 \cap V'$ is a proper G-submodule of M_0, and so it is 0. But then the family can be enlarged by including M_0 as a new member, a contradiction. $\square$

If V is G-module, let V^G be the *subspace of invariants*, i.e.,

$$V^G = \{v \in V : \text{for all } \gamma \in G,\ \gamma(v) = v\}.$$

Then V^G is the largest G-submodule of V on which G acts trivially, and it is a direct sum (although not in a unique way) of one-dimensional G-modules on which G acts trivially. Note that if M is an irreducible G-module on which G acts on non-trivially, then $M^G = 0$, for otherwise M^G is a proper nonzero G-submodule of M.

Theorem. *Let V be a G-module, where G is linearly reductive. Then V^G has a* unique *G-module complement V_G, which may also be characterized as the sum of all irreducible submodules M of V on which G acts non-trivially.*

Proof. Let W be any G-module complement for V^G. Let M be any irreducible in G on which G acts non-trivially. If $M \cap W \neq 0$, the $M \cap W = M$, and so $M \subseteq W$ as required. Otherwise M injects into $V/W \cong V^G$, which implies that G acts trivially on M, a contradiction. Thus, every irreducible on which G acts nontrivially is contained in W. But W is a direct sum of irreducibles, and G must act non-trivially on each of these, since there are no invariants in W. Therefore, W is the sum of all irreducible submodules of G on which G acts non-trivially, which proves that W is unique. $\square$

We also have:

Proposition. *If $f : V \to W$ is a map of G-modules, then $f : V^G \to W^G$, i.e., f induces a map of the respective G-invariant subspaces of V and W by restriction. Moreover, $f : V_G \to W_G$. Thus, f preserves the direct sum decompositions $V = V^G \oplus V_G$ and $W = W^G \oplus W_G$.*

Proof. If v is invariant so that $\gamma(v) = v$ for all $\gamma \in G$, then $\gamma\big(f(v)\big) = f\big(\gamma(v)\big) = f(v)$ for all $\gamma \in G$. Thus, $F(V^G) \subseteq W^G$.

Now consider any irreducible M on which G acts non-trivially. The kernel of f intersected with M is a G-submodule of M, and, hence, is 0 or M. If it is 0, then M injects into W, and the image is an isomorphic copy of M, which means that $f(M)$ is an irreducible G-submodule of W on which G acts non-trivially. Hence, $f(M) \subseteq W_G$. On the other hand, if the kernel contains all of M, the image is $0 \subseteq W_G$. $\square$

Dicussion. Let G be a linear algebraic group that is not necessarily lineaarly reductive. Consider a short exact sequence of G-modules

$$0 \to W \to V \to Y \to 0.$$

Clearly, $W^G \subseteq Y^G$, and the kernel of the map $V^G \to Y^G$ is, evidently, $V^G \cap W$, which is obviously W^G. Hence, for any linear algebraic group, we always have that

$$0 \to W^G \to Y^G \to V^G$$

is exact. In general, however, the map $Y^G \to V^G$ need not be onto. However:

Corollary. *If G is linearly reductive and $0 \to W \to V \to Y \to 0$ is an exact sequence of G-modules, then $0 \to W^G \to V^G \to Y^G \to 0$ is exact.*

Proof. The map $V \to Y$ is the direct sum of the maps $V^G \to Y^G$ and $V_G \to Y_G$. Hence, it is surjective if and only if both $V^G \to Y^G$ and $V_G \to Y_G$ are surjective, which, in particular, shows that $V^G \to Y^G$ is surjective. $\square$

When G is linearly reductive, we have a canonical G-module retraction $\rho_V : V \twoheadrightarrow V^G$ that is obtained by killing V_G. This map is called the *Reynolds operator*. Note that if we are given a short exact sequence of G-modules $0 \to W \to Y \to V \to 0$, then we have a commutative diagram:

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & W_G & \longrightarrow & V_G & \longrightarrow & Y_G & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & W & \longrightarrow & V & \longrightarrow & Y & \longrightarrow & 0 \\
 & & {\scriptstyle \rho_W}\downarrow & & {\scriptstyle \rho_V}\downarrow & & {\scriptstyle \rho_Y}\downarrow & & \\
0 & \longrightarrow & W^G & \longrightarrow & V^G & \longrightarrow & Y^G & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

The columns are split exact, and the rows are exact: the middle row is the direct sum of the rows above and below it.

The property that when $V \to W$ is a surjection of finite-dimensional G-modules then $V^G \to W^G$ is surjective actually characterizes linearly reductive groups. To see this, first note that if V and W are finite-dimensional G-modules, we can put a G-module structure on $\mathrm{Hom}_K(V, W)$ (this is simply the vector space of all K-linear maps) as follows: for all $\gamma \in G$ and all $f : V \to W$, $\gamma(f)(v) = \gamma(f(\gamma^{-1}v)$. This is easily verified to give $\mathrm{Hom}_K(V, W)$ the structure of a G-module. Moreover:

Lemma. *Let V, W be finite-dimensional G-modules. Then $\mathrm{Hom}_K(V, W)^G$ is the K-vector space of G-module maps from V to W.*

Proof. Suppose that $f : V \to W$. Then f is fixed by G if and only if for all $\gamma \in G$ and for all $v \in V$, $\gamma\big(f(\gamma^{-1}v)\big) = f(v)$, i.e., $f(\gamma^{-1}v) = \gamma^{-1}f(v)$. Since γ^{-1} takes on every value in G as γ varies, we have that f is fixed by G iff f is a G-module homomorphism. $\square$

Theorem. *Let G be a linear algebraic group. G is linearly reductive if and only if for every surjective G-module map of finite-dimensional G-modules $V \twoheadrightarrow W$, the map $V^G \to W^G$ is also surjective.*

Proof. It suffices to show that every finite-dimensional G-module V is a direct sum of irreducible G-modules: if not, let V be a counter-example of smallest possible vector space dimension. Then V is not irreducible, and we may choose a maximal proper G-submodule $M \neq 0$, so that $W = V/M$ is irreducible. It suffices to show that the exact sequence

$$(*) \quad 0 \to M \to V \xrightarrow{f} W \to 0$$

splits as a sequence of G-modules, since in that case we have that $V \cong M \oplus W$ and $\dim_K(M) < \dim_K(V)$. It is, of course, split as a sequence of K-vector spaces. Apply $\mathrm{Hom}_K(W, \_)$, where this is simply Hom as K-vector spaces. Then

$$0 \to \mathrm{Hom}_K(W, M) \to \mathrm{Hom}_K(W, V) \xrightarrow{f_*} \mathrm{Hom}_K(W, W) \to 0$$

is exact (since the sequence $(*)$ is split as a sequence of K-vector spaces), and the map f_*, which sends $g : W \to V$ to $f \circ g$, is therefore surjective. This is a sequence of G-modules, and so the map

$$\mathrm{Hom}_K(W, V)^G \to \mathrm{Hom}_K(W, W)^G$$

is surjective. That is, the set of G-module maps from $W \to V$ maps onto the set of G-module maps from $W \to W$. Hence, there is a G-module map $g : W \to V$ such that $f_*(g) = f \circ g$ is the identity map on W, and so $(*)$ is split as a sequence of G-modules. $\square$

Remark. The existence of a functorial Reynolds operator that retracts every finite-dimensional G-module onto its invariant submodule and so, for every G-module map $V \to W$, provides a commutative diagram:

$$\begin{array}{ccc} V & \xrightarrow{f} & > W \\ {\scriptstyle \rho_V}\big\downarrow & & \big\downarrow{\scriptstyle \rho_W} \\ V^G & \xrightarrow[f]{} & W^G \end{array}$$

already implies that when the top arrow is surjective, so is the bottom arrow. For if $w \in W^G$ we may choose an arbitrary element $v \in V$ such that $f(v) = w$, and then

$$f\big(\rho_V(v)\big) = \rho_W\big(f(v)\big) = \rho_G(w) = w,$$

as required. Thus, the existence of a functorial retraction onto the modules of invariants is also equivalent to the condition that G be linearly reductive.

Remark. If G is a finite group such that the order $|G|$ of G is invertible in K, the Reynolds operator is given by:

$$\rho(v) = \frac{1}{|G|} \sum_{g \in G} g(v),$$

i.e., averaging over the group G.

It turns out that linear reductive linear algebraic groups over the complex numbers $\mathbb{C}$ are precisely those that have a Zariski dense compact real Lie subgroup H. Then H has Haar measure, a translation-invariant measure μ such that $\mu(H) = 1$, and the Reynolds operator can be obtained by averaging over the group:

$$\rho(v) = \int_{\gamma \in H} \gamma(v)\, d\mu.$$

Early proofs of finite generation for rings of invariants of semisimple groups over $\mathbb{C}$ made use of this idea. Purely algebraic proofs have been available for a long time: these involve the study of modules over the Lie algebra. See, for example, [A. Borel, *Linear Algebraic Groups*, Benjamin, New York, 1969].

Lecture of February 17

The additive group $G = (K, +)$ of the field K may be identified with the group of upper triangular 2×2 unipotent matrices

$$\{\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} : a \in K\},$$

since

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a+b \\ 0 & 1 \end{pmatrix}$$

for all a, $b \in K$. This group is not linearly reductive. Let $V = K^2$, thought of a column vectors, and let G act in the obvious way, by left multiplication on column vectors. Let $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Then $V^G = Ke_1$ is a G-stable subspace of K^2, i.e., e_1 is an eigenvectxor of every matrix in G corresponding to the eigenvalue 1. However, Ke_1 has no G-stable complement in K^2: such a complement would be one-dimensional and that would require matrices such as $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in G$ to have a second eigenvector. $\square$

The Reynolds operator for ring actions and finite generation of R^G

We next want to study the situation where G is a linearly reductive linear algebraic group acting on a K-algebra R by ring automorphisms.

Theorem. *Let G be a linearly reductive algebraic group and let R be a K-algebra that is a G-module such that G acts on R by K-algebra automorphisms. Then the Reynolds operator $R \to R^G$ is R^G-linear.*

Proof. The Reynolds operator arises from the decomposition $R = R^G \oplus R_G$. It suffices to show that R_G is an R^G-module. Let $M \subseteq R_G$ be a typical irreducible G-submodule of R on which G acts non-trivially. Let $a \in R^G$, and consider the map $M \twoheadrightarrow aM$ that sends $r \mapsto ar$ for all $r \in M$. This is a map of G-modules, because for all $\gamma \in G$, $\gamma(ar) = \gamma(a)\gamma(r) = a\gamma(r)$. The kernel is therefore a G-submodule of M. If the kernel is M, then $aM = 0 \subseteq R_G$. If the kernel is 0, then $M \cong aM$ as G-modules. It follows that G acts non-trivially on the irreducible G-module aM, and so $aM \subseteq R_G$, as required. $\square$

We have the following:

Lemma. *Let $A \subseteq R$ be a ring extension such that A is a direct summand of R as an A-module, i.e., there is an A-linear map $R \to A$ that restricts to the identity map on A.*

(a) *For every ideal I of A, $IR \cap A = I$.*

(b) *If R is Noetherian, then A is Noetherian.*

(c) *If R is Noetherian and A is an $\mathbb{N}$-graded algebra over $A_0 = K$, a field, then A is a finitely generated K-algebra.*

Proof. (a) Suppose we have $a = f_1 r_1 + \cdots + f_k r_k$ where $a \in A$, the $f_j \in I \subseteq A$, and the $r_j \in R$. Then $\rho(a) = a$, and by the A-linearity of ρ, we have that $a = \rho(a) = f_1\rho(r_1) + \cdots + f_k\rho(r_k) \in I$, as required, since each $\rho(r_j) \in A$.

(b) Suppose that $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ is an infinite non-decreasing chain of ideals in A. Since R is Noetherian, then chain $I_j R$ is eventually stable, and so for some k, $I_k R = I_{k+h} R$ for all $h \geq 0$. Intersecting with A and applying (a), we have that $I_k = I_{k+h}$ for all $h \geq 0$, as required.

(c) By part (b), A is Noetherian, and so its maximal ideal is finitely generated as an ideal. We can take the generators to be forms of positive degree, say $F_1, \ldots, F_h$. Let $B = K[F_1, \ldots, F_h] \subseteq A$. It suffices to show that $B = A$. If not, we can choose a homogeneous element $F \in A - B$ of least degree. Since F is in the maximal ideal of A, we can write $F = \sum_{j=1}^{h} G_j F_j$, and by taking homogenous components we may assume that if $G_j \neq 0$, then $\deg(G_j) = \deg(F) - \deg(F_j) < \deg(F)$, and so every $G_j \in B$ by the fact that F has least degree in $A - B$. But then $F \in B$ as well. $\square$

Corollary. *If G is a linearly reductive linear algebraic group acting by K-automorphisms on a finitely generated K-algebra R, then R^G is finitely generated.*

Proof. If R is graded and the action preserves degree, this follows from part (c) of the Lemma above. In the general case, we can choose a finite-dimensional vector space $V \subseteq R$ that is G-stable and contains generators of R. We may then form the symmetric algebra

S of V over K, which is a polynomial ring over K whose space of forms of degree 1 is is isomorphic with V. We may let G act on V using the G-module structure of G, and this action extends to the polynmial ring S. The map $S \to R$ that sends each element of $V = [S]_1$ to itself, but considered as an element of $V \subseteq R$ extends uniquely to a K-algebra homomorphism $S \to R$. Since V generates R, this map is surjective. It is easy to see that this is also a map of G-modules. Hence, since we have a surjection $S \twoheadrightarrow R$, we also have a surjection $S^G \twoheadrightarrow R^G$. S^G is finitely generated over K by the graded case already considered, and so R^G is finitely generated over K. $\square$

Hilbert's fourteenth problem asks whether every ring of invariants of a linear algebraic group acting on a polynomial ring is finitely generated. This turns out to be false: the first counter-example was given by M. Nagata. It involved the action of the product of a large number of copies of the additive group of the field. Finite generation does hold when the group is linearly reductive and in some other important cases. We mention one here.

Theorem (Emmy Noether). *Let G be a finite group acting on a finitely generated K-algebra R. Then R^G is a finitely generated K-algebra.*

Proof. Let $R = K[r_1, \ldots, r_k]$. Suppose that $|G| = n$, say $G = \{\gamma_1, \ldots, \gamma_n\}$. For each r_i, consider the elements $\gamma_1(r_i), \ldots, \gamma_n(r_i)$. The elementary symmetric functions of these elements are invariant, and give coefficients for an equation of integral dependence of r_i, namely $\prod_{j=1}^n \big(z - \gamma_j(r_i)\big) = 0$. Hence, if R_0 is generated over K by the k sets of elementary symmetric functions of elements $\gamma_1(r_i), \ldots, \gamma_n(r_i)$, $1 \leq i \leq k$, then R_0 is finitely generated over K, and $R_0 \subseteq R^G \subseteq R$. Each r_i is integral over R_0, and so R is integral and finitely generated over R_0. Hence, R is module-finite over R_0, which is Noetherian. It follows that R^G is module-finite over R_0 and, hence, finitely generated over K. $\square$

The Cohen-Macaulay property for certain rings of invariants

Our next main objective is to prove the following:

Theorem. *Let G be a linearly reductive linear algebraic group over a field K, acting by K-automorphisms on a polynomial ring $R = K[x_1, \ldots, x_n]$ by a degree-preserving action, i.e., an action that extends an action of G on $[R]_1$. Then R^G is a Cohen-Macaulay ring.*

The proof will occupy us for a while. One of the subtle points is that a homogeneous system of parameters of R^G, which will generate an ideal of height $d = \dim(R^G)$ in R^G, typically generates an ideal of smaller height in R: in fact, it is hard to say anything special about the expansion to R of the ideal generated by a homogeneous system of parameters of R^G.

The argument we give will depend on reduction to characteristic $p > 0$, which is odd, because there are relatively few linearly reductive groups in positive characteristic. Another proof is known: cf. [J.-F. Boutot, *Singularités rationelles par les groupes réductifs*,

Invent. Math. **88** (1987) 65–68]. However, that argument needs resolution of singularities, Grothendieck duality, and the Grauert-Riemenschneider vanishing theorem. The first proof of this Theorem, which used reduction to prime characteristic $p > 0$, was given in [M. Hochster and J. L. Roberts, *Rings of invariants of reductive groups acting on regular rings are Cohen-Macaulay*, Adv. in Math. **13** (1974) 115–175], but the argument we give here follows a line of thought introduced in [M. Hochster and C. Huneke, *Tight closure, invariant theory, and the Briançon-Skoda Theorem*, J.A.M.S. **3** (1990) 31–116]. The Theorem is actually true whenever A is a graded ring that is a direct summand over itself of a polynomial ring $K[x_1, \ldots, x_n]$. In fact, whenever A is a direct summand of a regular ring R as an A-module, if A contains a field it must be Cohen-Macaulay, but the argument for the general case, which can be achieved along the same lines as the argument given here, is much more technical.

Here is a sharper form of the Theorem:

Theorem. *Let R be a polynomial ring over a field K, let A be a K-subalgebra of R generated by forms, and let $F_1, \ldots, F_d$ be a homogeneous system of parameters of A such that for every i, $1 \leq i \leq d-1$, $(F_1, \ldots, F_i)R \cap A = (F_1, \ldots, F_i)A$. Then A is a Cohen-Macaulay ring.*

If $A = R^G$ for a linearly reductive linear algebraic group G, then every ideal of A is contracted from R, and so we have that the ideals $(F_1, \ldots, F_i)A$ are contracted from R.

We shall first prove the Theorem above in characteristic $p > 0$. The proof depends on the following somewhat technical fact:

Theorem (colon-capturing). *Let A be an $\mathbb{N}$-graded domain finitely generated over a field K of prime characteristic $p > 0$. Let $F_1, \ldots, F_d$ be a homogeneous system of parameters for A. Suppose that one has a relation:*

$$u_{i+1}F_{i+1} = u_1F_1 + \cdots + u_iF_i$$

for some i. Then there exists an element $c \in A - \{0\}$ such that for all nonnegative integers $e \gg 0$,

$$(*) \quad cu_{k+1}^{p^e} \in (F_1^{p^e}, \ldots, F_i^{p^e})A.$$

Before proving this fact, we want to make several comments. When working in prime characteristic $p > 0$, it will be typographically convenient to use the letter q to stand for p^e, where $e \in \mathbb{N}$. Thus, the statement $(*)$ can be expressed instead as

$$(**) \quad cu_{k+1}^{q} \in (F_1^{q}, \ldots, F_i^{q})A.$$

Consider an ideal $J \subseteq A$, where A is any ring of prime characteristic $p > 0$. Then we shall use the notation $J^{[q]}$ for the ideal $(u^q : u \in A)A$, i.e., the ideal generated by all qth powers of elements of J. If J has generators u_i, then $J^{[q]}$ has generators u_i^q, since

$$(r_{i_1}u_{i_1} + \cdots + r_{i_h}u_{i_h})^q = r_{i_1}^q u_{i_1}^q + \cdots + r_{i_h}^q u_{i_h}^q,$$

but $J^{[q]}$ is independent of the choice of generators of J. Note that $J^{[q]} \subseteq J^q$, but, unless J is principal, J^q tends to be considerably larger: it contains all products of q generators of J, while $J^{[q]}$ contains only q th powers of generators of J.

The condition in $(**)$ for all $q \gg 0$ with fixed $c \neq 0$ may be construed, heuristically, as asserting that the element u_{i+1} is "almost" in the ideal generated by $F_1, \ldots, F_i$. We can make this thought somewhat less vague as follows: take q th roots of both sides in a suitable integral extension of A (one must adjoing sufficieintly many q th roots of elements in A). From the equation

$$(\#) \quad cu_{i+1}^q = F_1^q u_1 + \cdots + F_i^q u_i$$

one gets

$$(\#\#) \quad c^{1/q} u_{i+1} = F_1 u_1^{1/q} + \cdots + F_i u_k^{1/q}.$$

As $q \to \infty$, $1/q$ approaches 0, and so one may think of $c^{1/q}$ as approaching 1 in a vague heuristic sense. Thus, elements getting "arbitrarily close to 1" are multiplying u_{i+1} into $(F_1, \ldots, F_i)$, although in a somewhat larger ring than A.

Proof of the Theorem on colon-capturing. A is module-finite over $B = K[F_1, \ldots, F_d]$. Let $u_1, \ldots, u_h$ be a maximal sequence of elements of A that are linearly indpendent over B, so that $G = Bu_1 + \cdots Bu_h$ is a free B-module of rank h. Here, h will be the same as the degree of the extension of fraction fields, $[\mathrm{frac}\,(A) : \mathrm{frac}\,(B)]$. Consequently, A/G is a torsion-module over the domain B: we can see this as follows. If $u \in A - G$, it must have a nonzero multiple in G: otherwise $u_1, \ldots, u_h, u$ are linearly independent over B, contradicting the choice of h. Hence, each generator of A has a nonzero multiple in G. By taking the product of the multipliers, we obtain a nonzero element $c \in B \subseteq A$ such that $cA \subseteq G$. It turns out that c has the property we require.

Suppose that we have a relation

$$F_{i+1} u_{i+1} = F_1 u_1 + \cdots + F_i u_i,$$

where the $u_j \in A$. Taking q th powers where $q = p^e$ we have:

$$F_{i+1}^q u_{i+1}^q = F_1^q u_1^q + \cdots + F_i^q u_i^q,$$

and multiplying by c gives

$$(\#) \quad F_{i+1}^q (cu_{i+1}^q) = F_1^q (cu_1^q) + \cdots + F_i^q (cu_i^q).$$

Since each of the elements $cF_j^q \in cA \subseteq G$, we may think of $(\#)$ as a relation on $F_1^q, \ldots, F_{i+1}^q$ with coefficients in the free B-module G. Since $F_1^q, \ldots, F_{i+1}^q$ is a regular sequence on B, it is a regular sequence on G, and we can conclude that

$$cu_{i+1}^q \in (F_1^q, \ldots, F_i^q)G \subseteq (F_1^q, \ldots, F_i^q)A,$$

for all $q = p^e$, as required. $\square$

We shall prove the following Lemma: we postpone giving the argument for a bit.

Lemma. *Let R be the polynomial ring $K[x_1, \ldots, x_n]$. Let J be any ideal of R. Suppose that there exists $c \in R - \{0\}$ and $f \in R$ such that $cf^q \in J^{[q]}$ for all $q \gg 0$. Then $f \in J$.*

Assuming this result for the moment, we give the proof of the sharper form of the Theorem on the Cohen-Macaulay proprerty for rings of invariants. The argument is amazingly easy now!

Proof of the sharper theorem. We want to show that $F_1, \ldots, F_d$ is a regular sequence in A. Suppose that $uF_{i+1} \in (F_1, \ldots, F_i)A = I$. By the Theorem on colon-capturing above, we have that there exists $c \neq 0$ in A such that $cu^q \in I^{[q]}$ for all $q \gg 0$. Then we may expand I to R to obtain $cu^q \in (IR)^{[q]}$ for all $q \gg 0$. By the Lemma above, we then have $u \in IR$, so that $u \in IR \cap A = I$ by hypothesis. $\square$

It remains to prove the Lemma.

Lecture of February 19

If R is a ring of prime characteristic p we write $F_R : R \to R$ for the *Frobenius endomorphism*: $F_R(r) = r^p$. If $e \in \mathbb{N}$, we write F_R^e for the composition of F_R with itself e times, the iterated Frobenius endomorphism. Thus, $F_R^e(r) = r^{p^e}$. The subscript $_R$ is often omitted.

Quite generally, if R is a regular Noetherian ring, $F^e : R \to R$ is faithfully flat. We shall not prove this fact in general at this point, but we do want to prove that when R is a polynomial ring over a field K, $F^e : R \to R$ makes the right hand copy of R into a free R-module over the left hand copy of R. Note that F^e is an injective homomorphism, since the polynomial ring has no nonzero nilpotents. The image of R under this map is $R^q = \{r^q : r \in R\}$, where $q = p^e$.

We first note the following:

Lemma. *If T is free as S-algebra and S is free as an R-algebra, then T is free as an R-algebra. In fact, if $\{t_j\}_{j \in \mathcal{J}}$ is a free basis for T over S and $\{s_i\}_{i \in \mathcal{I}}$ is a free basis for S over R then the set of products $\{t_j s_i : j \in \mathcal{J}, i \in \mathcal{I}\}$ is a free basis for T over R.*

Proof. If $t \in T$, we can write $t = \sum_{k=1}^n u_k t_{j_k}$, where the $u_k \in S$, and then we may express every u_k as an R-linear combination of finitely many of the elemnts s_i. It follows that the specified products span. If some R-linear combination of the products is 0, we may enlarge the set so that it consists of elements $s_{i_h} t_{j_k}$ for $1 \leq h \leq m$ and $1 \leq k \leq n$. If

$$\sum_{1 \leq h \leq m, 1 \leq k \leq n} r_{hk} s_{i_h} t_{j_k} = 0$$

where the $r_{hk} \in R$. We can write this as

$$\sum_{k=1}^{n}(\sum_{h=1}^{m} r_{hk}s_{i_h})t_{j_k} = 0,$$

from which we first conclude that every $\sum_{h=1}^{m} r_{hk}s_{i_h} = 0$ and then that every $r_{hk} = 0$. $\square$

Proposition. *If B is a free A-algebra, $x_1, \ldots, x_n$ are indeterminates, and $k_1, \ldots, k_n$ are positive integers, then $B[x_1, \ldots, x_n]$ is free over $A[x_1^{k_1}, \ldots, x_n^{k_n}]$.*

Proof. By a straightforward induction, this reduces at once to the case where $n = 1$. We let $x = x_1$ and $k = k_1$. Then $B[x] \cong A[x] \otimes_A B$ is free over $A[x]$. By the preceding Lemma, it suffices to show that $A[x]$ is free over $A[x^k]$. But it is quite easy to verify that the elements x^a for $0 \leq x \leq a-1$ are a free basis. $\square$

Theorem. *Let K be field and let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over K. Then $F_R^e : R \to R$ makes the right hand copy of R into a free module over the left hand copy of R.*

Proof. The image of R under F^e is $R^q = K^q[x_1^q, \ldots, x_n^q]$. It suffices to show that R is free over R^q. Note that since K^q is a field, K is free over K^q. The result is now immediate from the preceding Proposition. $\square$

Lecture of February 22

We need the following:

Lemma. *Let $R \to S$ be flat, and let $I \subseteq R$, $J \subseteq R$ be ideals such that $J = (f_1, \ldots, f_k)R$ is finitely generated. Then $(I :_R J)S = IS :_S JS$.*

Proof. Consider the map $R \to (R/I)^{\oplus k}$ that sends $r \mapsto (\overline{f_1 r}, \ldots, \overline{f_k r})$ where $\overline{u}$ denotes the image of $u \in R$ modulo I. The kernel of this map is precisely $I :_R J$, i.e.,

$$0 \to I :_R J \to R \to (R/I)^{\oplus k}$$

is exact. Thus, this sequence remains exact when we apply $S \otimes_R \_$ to obtain:

$$0 \to (I :_R J) \otimes_R S \to S \to (S/IS)^{\oplus k}.$$

The kernel of $\phi : S \to (S/IS)^{\oplus k}$ is therefore the image of $(I :_R J) \otimes_R S \to S$, which is $(I :_R J)S$. (The map is injective, so that $(I :_R J) \otimes_R S \cong (I :_R J)S$. In general, if $R \to S$ is flat and $\mathfrak{A}$ is an ideal of R, when $S \otimes_R \_$ is applied to the injection $0 \to \mathfrak{A} \to R$ it

yields an isomorphism $\mathfrak{A} \otimes_R S \cong \mathfrak{A}S$.) But the definition of ϕ implies that the kernel is $IS :_S JS$. $\square$

Remark. When $\phi : R \to S$ and I is an ideal of R, IS is generated by the images of the elements of I under ϕ. Suppose that R is a ring of prime characteristic $p > 0$ and let $S = R$, made into an R-algebra by means of the structural homomorphism $F^e : R \to R$. Then for any ideal I of R, $IS = I^{[q]}$.

Then:

Theorem. *Let R be a polynomial ring $K[x_1, \ldots, x_n]$ over a field K of characteristic $p > 0$. For any two ideals I, $J \subseteq R$, $I^{[q]} :_R J^{[q]} = (I :_R J)^{[q]}$.*

Proof. Since $F^e : R \to R$ is flat, this is immediate from the Remark just above and the Lemma. $\square$

The following result now completes, in the case of prime characteristic $p > 0$, the proof of the sharper form of the Theorem on the Cohen-Macaulay property for rings of invariants stated at the top of p. 4 of the Lecture Notes of February 17.

Theorem. *Let R be a polynomial ring $K[x_1, \ldots, x_n]$ over a field K of characteristic $p > 0$. Let I be an ideal of R, let $u \in r$, and let $c \in R - \{0\}$. Suppose that $cu^q \in I^{[q]}$ for all $q = p^e \gg 0$. Then $u \in I$.*

Proof. The fact that $cu^q \in I^{[q]}$ for all $q \gg 0$ may be restated as $c \in I^q :_R (uR)^{[q]}$ for all $q \gg 0$. By the Theorem just above, this means that $c \in (I :_R uR)^{[q]}$ for all $q \gg 0$. If $u \notin I$, then $I :_R uR$ is a proper ideal and is contain in some maximal ideal m of R. Then for some q_0 we have

$$c \in \bigcap_{q \geq q_0} (I :_R Ru)^{[q]} \subseteq \bigcap_{q \geq q_0} m^{[q]} \subseteq \bigcap_{q \geq q_0} (mRm)^{[q]} \subseteq \bigcap_{q \geq q_0} (mR_m)^q = 0,$$

and so $c = 0$, a contradiction. Hence, we must have $u \in I$ after all. $\square$

Our next objective is to prove the Theorem for fields of characteristic 0 as well, by reducing to the characteristic p case.

First step: moving towards characteristic p

We now suppose that we have a counter-example to the Theorem stated at the top of p. 4 over a field K of equal characteristic 0. In the sequel, we want to replace K, insofar as possible, by a finitely generated $\mathbb{Z}$-subalgebra $D \subseteq K$. We then obtain a counterexample by killing a maximal ideal μ of D: it turns out that D/μ must be a finite field.

In order to carry our ideas through, we first need to prove some preliminary results. One is the fact just stated about maximal ideals in finitely generated $\mathbb{Z}$-algebras. However, we also need results of the following kind: suppose that $A_D \subseteq R_D$ are finitely generated D-algebras. Then one can localize at one nonzero element $d \in D - \{0\}$ such that $(R_D/A_D)_d$ is flat over D_d. We shall prove one of the strongest known results of this type. This will enable us to preserve an inclusion $A_D \subseteq R_D$ while killing a maximal ideal of D. We shall need to be able to do this and also preserve various other inclusions like this in order to give the detailed argument.

We first review the Noether Normalization Theorem over a domain. We begin with:

Lemma. *Let D be a domain and let $f \in D[x_1, \ldots, x_n]$. Let $N \geq 1$ be an integer that bounds all the exponents of the variables occurring in the terms of f. Let ϕ be the D-automorphism of $D[x_1, \ldots, x_n]$ such that $x_i \mapsto x_i + x_n^{N^i}$ for $i < n$ and such that x_n maps to itself. Then the image of f under ϕ, when viewed as a polynomial in x_n, has leading term dx_n^m for some integer $m \geq 1$, with $d \in D - \{0\}$. Thus, over D_d, $\phi(f)$ is a scalar in D_d times a polynomial in x_n that is monic.*

Proof. Consider any nonzero term of f, which will have the form $c_\alpha x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$, where $\alpha = (a_1, \ldots, a_n)$ and c_α is a nonzero element in D. The image of this term under ϕ is

$$c_\alpha (x_1 + x_n^N)^{a_1} (x_2 + x_n^{N^2})^{a_2} \cdots (x_{n-1} + x_n^{N^{n-1}})^{a_{n-1}} x_n^{a_n},$$

and this contains a unique highest degree term: it is the product of the highest degree terms coming from all the factors, and it is

$$c_\alpha (x_n^N)^{a_1} (x_n^{N^2})^{a_2} \cdots (x_n^{N^{n-1}})^{a_{n-1}} x_n^{a_n} = c_\alpha x_n^{a_n + a_1 N + a_2 N^2 + \cdots + a_{n-1} N^{n-1}}.$$

The exponents that one gets on x_n in these largest degree terms coming from distinct terms of f are all distinct, because of uniqueness of representation of integers in base N. Thus, no two exponents are the same, and no two of these terms can cancel. Therefore, the degree m of the image of f is the same as the largest of the numbers

$$a_n + a_1 N + a_2 N^2 + \cdots + a_{n-1} N^{n-1}$$

as $\alpha = (a_1, \ldots, a_n)$ runs through n-tuples of exponents occurring in nonzero terms of f, and for the choice α_0 of α that yields m, $c_{\alpha_0} x_n^m$ occurs in $\phi(f)$, is the only term of degree m, and and cannot be canceled. It follows that $\phi(f)$ has the required form. $\square$

Theorem (Noether normalization over a domain). *Let T be a finitely generated extension algebra of a Noetherian domain D. Then there is an element $d \in D - \{0\}$ such that T_d is a module-finite extension of a polynomial ring $D_d[z_1, \ldots, z_h]$ over D_d.*

Proof. We use induction on the number n of generators of T over D. If $n = 0$ then $T = D$. We may take $h = 0$. Now suppose that $n \geq 1$ and that we know the result for algebras

generated by $n-1$ or fewer elements. Suppose that $T = D[\theta_1, \ldots, \theta_n]$ has n generators. If the θ_i are algebraically independent over K then we are done: we may take $h = n$ and $z_i = \theta_i$, $1 \leq i \leq n$. Therefore we may assume that we have a nonzero polynomial $f(x_1, \ldots, x_n) \in D[x_1, \ldots, x_n]$ such that $f(\theta_1, \ldots, \theta_n) = 0$. Instead of using the original θ_j as generators of our K-algebra, note that we may use instead the elements

$$\theta'_1 = \theta_1 - \theta_n^N,\ \theta'_2 = \theta_2 - \theta_n^{N^2},\ \ldots,\ \theta'_{n-1} = \theta_{n-1} - \theta_n^{N^{n-1}},\ \theta'_n = \theta_n$$

where N is chosen for f as in the preceding Lemma. With ϕ as in that Lemma, we have that these new algebra generators satisfy $\phi(f) = f(x_1 + x_n^N, \ldots, x_{n-1} + x_n^{N^{n-1}}, x_n)$ which we shall write as g. We replace D by D_d, where d is the coefficient of x_n^m in g. After multiplying by $1/d$, we have that g is monic in x_n with coefficients in $D_d[x_1, \ldots, x_{n-1}]$. This means that θ'_n is integral over $D_d[\theta'_1, \ldots, \theta'_{n-1}] = T_0$, and so T_d is module-finite over T_0. Since T_0 has $n-1$ generators over D_d, we have by the induction hypothesis that $(T_0)_{d'}$ is module-finite over a polynomial ring $D_{dd'}[z_1, \ldots, z_{d-1}] \subseteq (T_0)_{d'}$ for some nonzero $d' \in D$, and then $T_{dd'}$ is module-finite over $D_{dd'}[z_1, \ldots, z_h]$ as well. $\square$

Theorem. *Let κ be a field that is a finitely generated $\mathbb{Z}$-algebra. Then κ is a finite field. Hence, if μ is any maximal ideal of a finitely generated $\mathbb{Z}$-algebra D, then D/μ is a finite field.*

Proof. If $\mathbb{Z}$ injects into κ (we shall see that this cannot happen) then κ is a module-finite extension of a polynomial ring $\mathbb{Z}[1/d][x_1, \ldots, x_h]$ where $d \in \mathbb{Z} - \{0\}$ (we need not localize κ at d, since d must already be invertible in the field κ). If p is a prime not dividing d, then p is not invertible in $\mathbb{Z}_d$, nor in the polynomial ring, and hence cannot be invertible in a module-finite extension of the polynomial ring, a contradiction.

Hence, $\mathbb{Z}$ does not inject into κ, which implies that κ has characteristic $p > 0$ and is finitely generated over $\mathbb{Z}/pZ$ for some prime $p > 0$. Then κ is module-finite over a polynomial ring $(\mathbb{Z}/p\mathbb{Z})[x_1, \ldots, x_h]$. Since κ has dimension 0, we must have $h = 0$, i.e., that κ is module-finite over $\mathbb{Z}/p\mathbb{Z}$, which implies that κ is a finite field. $\square$

Second step: generic freeness

Before proving a strong form of generic freeness, we need:

Lemma. *Let D be any ring. let*

$$0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_k \subseteq \cdots \subseteq M$$

be a non-decreasing possibly infinite sequence of submodules of the module M over D, and suppose that $\bigcup_{k=1}^{\infty} M_k = M$. If M_{k+1}/M_k is free over D for all $k \geq 0$, then M is free.

Proof. Choose a free basis for every M_{k+1}/M_k and for every $k \geq 0$, let $\mathcal{B}_k$ be a set of elements in M_{k+1} that maps onto the chosen free basis for M_{k+1}/M_k. In particular, $\mathcal{B}_1$ is

a free basis for $M_1 \cong M_1/0$. We first claim that $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ is a free basis for M_{k+1} for every $k \geq 0$. We already have this for $k = 0$, and we use induction. Thus, we may assume that $\mathcal{B}_{k-1}$ is a free basis for $\mathcal{M}_k$, and we must show that $\mathcal{B}_k$ is a free basis for $\mathcal{M}_{k+1}$. This is clear from the fact that the D-linear map $M_{k+1}/M_k \to M_{k+1}$ that sends each element of the chosen free basis of M_{k+1}/M_k to the element of $\mathcal{B}_k$ that lifts it is a splitting of the exact sequence

$$0 \to M_k \to M_{k+1} \to M_{k+1}/M_k \to 0.$$

It then follows at once that $\mathcal{B} = \bigcup_{k=0}^{\infty} \mathcal{B}_k$ is a free basis for M: first, there can be no non-trivial relations, for such a relation involves only finitely many basis elements and so would give a non-trivial relation on the elements of some $\mathcal{B}_k$. Second, since $\mathcal{B}$ evidently contains a set that spans M_k for every k and $\bigcup_{k=1}^{\infty} M_k = M$, $\mathcal{B}$ spans M. $\square$

Theorem (strong form of generic freeness). *Let D be a Noetherian domain, and let $D = T_0 \to T_1 \to T_2 \to \cdots \to T_s$ be a sequence of maps of finitely generated T_0-algebras. Let M be a finitely generated T_s-module, and for every i, where $0 \leq i \leq s$, let N_i be a T_i-submodule of M. Let $Q = M/(N_0 + \cdots + N_s)$. Then there exists a nonzero element d in D such that Q_d is D_d-free.*

Proof. By inserting additional algebras in the chain, we may assume without loss of generality that every T_{i+1} is generated over the image of T_i by one element. We use induction on s. Note also that we can view Q as the quotient of $M' = M/N_s$ by the sum of the images of $N_1, \ldots, N_{s-1}$, so that there is no loss of generality in assuming that $N_s = 0$.

If $s = 0$ we simply have a finitely generated D-module M. In this case, take a maximal sequence of elements $u_1, \ldots, u_h \in M$ that are linearly independent over D, so that $G = Du_1 + \cdots + Du_h$ is free over D. (Such a sequence must be finite, or one would have an infinite strictly ascending chain of submodules of M spanned by the initial segments of the sequence $u_1, u_2, u_3, \ldots$.) It follows that M/G is a torsion-module over D: for every element u of $M - G$ there must be a nonzero element of D that multiplies u into G, or else we may take $u_{h+1} = u$ to get a longer sequence. Thus, there is an element d_j of $D - \{0\}$ that multiplies each element v_j of a finite set of generators for M into G. Let d be a nonzero common multiple of these d_j. Then $M_d = G_d$ is free over D_d.

Now suppose that $s \geq 1$. Take a finite set $\mathcal{S}$ of generators for M that includes a finite set of generators for each of the N_i. Let N be the T_{s-1} submodule of M generated by all of these. By the induction hypothesis, we can choose $d' \in D - \{0\}$ such that $N/(N_0 + \cdots + N_{s-1})$ becomes free when we localize at d'. If we can choose d such that M/N becomes free, then localizing at dd' solves the problem. Let θ be an element of T_s that generates T_s over the image of T_{s-1}. Let $M_0 = 0$ and let $M_i = N + \theta N + \cdots + \theta^{i-1}N$ for $i \geq 1$, so that $M_1 = N$, $M_2 = N + \theta N$, $M_3 = N + \theta N + \theta^2 N$, and so forth. Let $W_i = M_i/M_{i-1}$ for $i \geq 1$. We claim that there are surjections

$$N = W_1 \twoheadrightarrow W_2 \twoheadrightarrow \cdots \twoheadrightarrow W_k \twoheadrightarrow \cdots,$$

where the map $W_i \to W_{i+1}$ is induced by multiplication by θ, which takes $M_i \to M_{i+1}$ for every i. The image of the map on numerators contains $\theta^i N$, which spans the quotient,

so that these are all surjections. The kernels of the maps $N \to W_i$ form an ascending sequence of T_{s-1}-submodules of N, and so the kernels are all eventually the same. This implies that there exists k such that for all $i \geq k$, $W_i \cong W_k$. By the induction hypothesis for each of the modules W_j we can choose $d_j \in D - \{0\}$ such that $(W_j)_{d_j}$ is free over D_{d_j}. Let d be a common multiple of these d_j. By the Lemma above, $(M/N)_d$ is free over D_d. $\square$

Third step: descent to a finitely generated algebra over the integers

The next step in our effort to prove the sharper form of the result on the Cohen-Macaulay property for rings of invariants is to "replace" K by a finitely generated $\mathbb{Z}$-subalgebra D of K. The idea is to make D sufficiently large so that all of the salient features of a counter-example can be discussed in D-algebras instead of K-algebras. We then localize D at one element so as to make certain quotients free, using the Theorem on generic freeness. Finally, we kill a maximal ideal of D and so produce a counter-example to the characteristic $p > 0$ form of the Theorem. Since we have already proved the result in positive characteristic, this is a contradiction, and will complete the proof of the Theorem.

We have a field K of chracteristic 0, a polynomial ring $R = K[x_1, \ldots, x_n]$, a K-subalgebra A of R finitely generated over K by forms $u_1, \ldots, u_s$, and a homogeneous system of parameters $F_1, \ldots, F_d$ for A. We also know that for $1 \leq i \leq d-1$,

$$(F_1, \ldots, F_i)R \cap A = (F_1, \ldots, F_i)A.$$

We want to prove that $F_1, \ldots, F_d$ is a regular sequence. Suppose not, and suppose that

$$(\dagger) \quad GF_{i+1} = G_1F_1 + \cdots + G_iF_i$$

where $G_1, \ldots, G_i$, $G \in A$ and $G \notin (F_1, \ldots, F_i)A$, where $i \leq d-1$. We want to show that we can construct an example with the same properties in prime characteristic $p > 0$.

Since $F_1, \ldots, F_d$ is a homogeneous system of parameters for A, every u_j has a power in the ideal generated by $F_1, \ldots, F_d$. Hence, for every j we can choose $m_j \geq 1$ and an equation

$$u_j^{m_j} = w_{j,1}F_1 + \cdots + w_{j,d}F_d,$$

where the $w_{j,k} \in A$. Moreover, every F_t, G_t, and G, as well as all the $w_{j,k}$, can be expressed as polynomials in $u_1, \ldots, u_s$ with coefficients in K, say $F_k = P_k(u_1, \ldots, u_s)$, $G_k = Q_k(u_1, \ldots, u_s)$ for $1 \leq k \leq d$, $G = Q(u_1, \ldots, u_s)$, and $w_{j,k} = H_{j,k}(u_1, \ldots, u_s)$. As a first attempt at constructing the domain D, we take the $\mathbb{Z}$-subalgebra of K generated by all coefficients of the u_j (as polynomials in $x_1, \ldots, x_n$), the P_k, the Q_k, Q, and the $H_{j,k}$. However, we may (and shall) enlarge D further, specifically, by localizing at one nonzero element.

Let $R_D = D[x_1, \ldots, x_n]$, and let $A_D = D[u_1, \ldots, u_s] \subseteq R_D$. The elements F_j, G_j, G, and $w_{j,k}$ are in A_D, and we still have the relation $(\dagger)$ holding in A_D. Moreover, every u_j is

in the radical of the ideal generated by $(F_1, \ldots, F_d)$ in A_d, and so Rad $\big((F_1, \ldots, F_d)A_D\big)$ is a homogeneous prime ideal of A_D, call it $\mathcal{Q}_D$. It is spanned over D by all forms of positive degree. We have that $A_D/\mathcal{Q}_D = D$.

We are now ready for the dénouement, which involves applying the result on generic freeness to preserve this situation while passing to positive characteristic.

Lecture of February 24

The final step: the application of generic freeness

We have the following:

Lemma. *If $0 \to N \to M \to G \to 0$ is an exact sequence of D-modules and G is D-free, then the sequence is split, so that $M \cong N \oplus G$. In this case, for any D-module or D-algebra Q, the sequence $0 \to Q \otimes_D N \to Q \otimes_D M \to Q \otimes_D G \to 0$ is exact.*

Proof. To construct a splitting $f : G \to M$ choose a free basis $\mathcal{B}$ for G and for every element $b \in \mathcal{B}$, define $f(b)$ to be an element of M that maps to b. Exactness is preserved by $Q \otimes_D \_$ becaue tensor product commutes with direct sum. $\square$

We are now ready to complete the proof.

There are several exact sequences that we are going to want to preserve while passing to characteristic $p > 0$. Since A has Krull dimension d and is module-finite over $K[F_1, \ldots, F_d]$, we know that $F_1, \ldots, F_d$ are algebraically independent over K and, hence, over the smaller ring D. This yields

$$(1) \quad 0 \to D[F_1, \ldots, F_d] \to A_D \to A_D/D[F_1, \ldots, F_d] \to 0$$

where $D[F_1, \ldots, F_d]$ is a polynomial ring over D. After localizing at one element of $D - \{0\}$ we may assume that all these modules are D-free, and, henceforth we assume this. We shall make a number of further localizations like this, but only finitely many. Note that localizing further preserves freeness. So long as there are only finitely many localizations at one element, D remains a finitely generated $\mathbb{Z}$-algebra.

Second, we have

$$(2) \quad 0 \to A_D \to R_D \to R_D/A_D \to 0.$$

We may assume that D has been localized at one more element so that the terms of the exact sequence above are D-free.

For every j, the ideal $(F_1, \ldots, F_j)A$ is contracted from $R = K[x_1, \ldots, x_n]$. This implies that the map $A/(F_1, \ldots, F_j)A \to R/(F_1, \ldots, F_j)R$ is injective. This map arises from the map

$$(*) \quad A_D/(F_1, \ldots, F_j)A_D \to R_D/(F_1, \ldots, F_D)R_D$$

in two steps: we may tensor over D with the fraction field $\mathcal{F}$ of D, and then we may tensor over $\mathcal{F} \subseteq K$ with K. After we tensor with K, we know that the map is injective. Since K is faithfully flat (in fact, free) over its subfield $\mathcal{F}$, $(*)$ is injective once we tensor with $\mathcal{F}$. Therefore the kernel, if any, is torsion over D. Hence, if we localize at one element of $D - \{0\}$ so that $A_D/(F_1, \ldots, F_j)A_D$ becomes D-free, the map $(*)$ is injective. We may also localize at one element of $D - \{0\}$ so that the cokernel is free over D, and therefore we have for every j an exact sequence

$$(3) \quad 0 \to A_D/(F_1, \ldots, F_j)A_D \to R_D/(F_1, \ldots, F_D)R_D \to \frac{R_D/(F_1, \ldots, F_D)R_D}{A_D/(F_1, \ldots, F_j)A_D} \to 0$$

consisting of free D-modules.

Finally, we have that $G\big(A/(F_1, \ldots, F_i)A\big) \neq 0$. It follows that $G\big(A_D/(F_1, \ldots, F_i)A_D\big)$ is not a D-torsion module, since it is nonzero after we apply $K \otimes_D \_$. Hence, after localizing further at one element of $D - \{0\}$, we may assume that

$$(4) \quad 0 \to G\big(A_D/(F_1, \ldots, F_i)A_D\big) \to A_D/(F_1, \ldots, F_i)A_D \to A_D/(F_1, \ldots, F_i, G)A_D \to 0$$

is an exact sequence of free D-modules such that the module $G\big(A_D/(F_1, \ldots, F_i)A_D\big)$ is not zero.

We now choose a maximal ideal μ of D. Then $\kappa = D/\mu$ is a finite field, and has prime characteristic $p > 0$ for some p. We write A_κ and R_κ for $\kappa \otimes_D A_D = A_D/\mu A_D$ and $\kappa \otimes_D R_D = R_D/\mu R_D \cong \kappa[x_1, \ldots, x_n]$, respectively. We use $\overline{w}$ to indicate the image $1 \otimes w$ of w in A_κ or R_κ. By the preceding Lemma, the sequences displayed in (1), (2), (3), and (4) remain exact after applying $\kappa \otimes_D \_$.

From (1) we have an injection of $\kappa[F_1, \ldots, F_d]$, which is a polynomial ring, into A_κ. This shows that the dimension of A_κ is at least d. Since the homogeneous maximal ideal of A_κ is generated by the $\overline{u}_j$ and these are nilpotent on the ideal $(\overline{F}_1, \ldots, \overline{F}_d)A_\kappa$, we have that $\overline{F}_1, \ldots, \overline{F}_d$ is a homogeneous system of parameters for A_κ. From (2) we have an injection $A_\kappa \hookrightarrow R_\kappa$. From (3), we have that $(\overline{F}_1, \ldots, \overline{F}_j)A_\kappa$ is contracted from R_κ for every j. From (4), we have $\overline{G}$ is not in $(\overline{F}_1, \ldots, \overline{F}_i)A_\kappa$, although we still have that

$$\overline{G}\,\overline{F}_{i+1} = \overline{G}_1\overline{F}_1 + \cdots + \overline{G}_i\overline{F}_i$$

in A_κ, so that A_κ is not Cohen-Macaulay. This contradicts the positive characteristic version of the Theorem, which we have already proved. $\square$

Note: we have completed the proof of the sharper form of the result on the Cohen-Macaulay property for rings of invariants stated on p. 4 of the Lecture Notes of February 17 in all characteristics now, and, consequently, we have completed as well the proof of the Theorem stated in the middle of p. 3 of the Lecture Notes of February 17.

Remarks. It might seem more natural to prove the Theorem stated in the mdidle of p. 3 of the Lecture Notes of February 17 by preserving the Reynolds operator, i.e., that the ring

of invariants is a direct summand, while passing to characteristic p. It turns out that this is not possible, as we shall see below. What we actually did was to preserve finitely many specific consequences of the existence of the Reynolds operator, namely the contractedness of the ideals $(F_1, \ldots, F_j)A$ from R, while passing to characteristic p, and this was sufficient to get the proof to work.

Consider the action of $G = \mathrm{SL}(2, K)$ on $\mathbb{C}[X]$, where $X = \left(x_{i,j}\right)$ is a 2×3 matrix of indeterminates that sends the entries of X to the corresponding entries of γX for all $\gamma \in G$. We have already noted that the ring of invariants in this case is $\mathbb{C}[\Delta_1, \Delta_2, \Delta_3]$, where Δ_j is the determinant of the submatrix of X obtained by deleting the j th column of X: see the third Example on p. 3 of the Lecture Notes of February 1. In this case Δ_1, Δ_2, and Δ_3 are algebraically independent: this is true even if we special the entries of the matrix X so as to obtain

$$\begin{pmatrix} 1 & 1 & (y-z)/x \\ 0 & x & y \end{pmatrix},$$

where x, y, and z are indeterminates. It is easy to "descend" the inclusion $A = R^G = \mathbb{C}[\Delta_1, \Delta_2, \Delta_3] \subseteq \mathbb{C}[X]$ to an inclusion of finitely generated $\mathbb{Z}$-algebras: one can take $D = \mathbb{Z}$, and consider the inclusion $\mathbb{Z}[\Delta_1, \Delta_2, \Delta_3] \subseteq \mathbb{Z}[X]$. However, this is *not* split after we localize at one integer of $\mathbb{Z} - \{0\}$, nor even if we localize at all positive prime integers except a single prime $p > 0$. The Reynolds operator needs the presence of *all* prime integers $p \neq 0$ in the denominators. Note that if the map were split after localizing at all integers not divisible by p, we could then apply $\mathbb{Z}/p\mathbb{Z} \otimes_{\mathbb{Z}} \_$ and get a splitting of the map $(\mathbb{Z}/p\mathbb{Z})[\Delta_1, \Delta_2, \Delta_3] \subseteq (\mathbb{Z}/p\mathbb{Z})[X]$. But we shall see below that this map is *not* split.

At the same time, we want to note that in the Theorem on generic freeness, it is important that the algebras T_i are nested, with maps $T_0 \to T_1 \to T_2 \to \cdots \to T_s$. The result is false if one kills a sum of submodules over mutually incomparable subalgebras, or even a sum of such subalgebras.

Both our proof that $(\mathbb{Z}/p\mathbb{Z})[\Delta_1, \Delta_2, \Delta_3] \subseteq (\mathbb{Z}/p\mathbb{Z})[X]$ does not split and our example of the fallure of generic freeness when the T_i are incomparable are based on looking at the same example.

Namely, we consider the module

$$H = \frac{\mathbb{Z}[X]_{\Delta_1\Delta_2\Delta_3}}{\mathbb{Z}[X]_{\Delta_2\Delta_3} + \mathbb{Z}[X]_{\Delta_1\Delta_3} + \mathbb{Z}[X]_{\Delta_1\Delta_2}}$$

where X is the same 2×3 matrix of indeterminates discussed in the action of $\mathrm{SL}(2, \mathbb{C})$ above and $D = T_0 = \mathbb{Z}$. Note that the numerator and the three summands in the denominator are all finitely generated $\mathbb{Z}$-algebras. We shall see that $\mathbb{Q} \otimes_{\mathbb{Z}} H$ is a nonzero vector space over the rational numbers $\mathbb{Q}$, and that H is a divisible abelian group, i.e., that $nH = H$ for every nonzero integer n. It follows that if we localizate at any nonzero integer $n \in \mathbb{Z}$, H_n is nonzero, and is not free over $\mathbb{Z}_n$. If it were free over $\mathbb{Z}_n$, it could not be divisible by p for any integer p that does not divide n, since it is simply a direct sum of copies of $\mathbb{Z}_n$.

It remains to prove the assertions that $\mathbb{Q} \otimes H \neq 0$, that $pH = H$ for every nonzero prime integer $p > 0$, and that the map $(\mathbb{Z}/p\mathbb{Z})[\Delta_1, \Delta_2\, \Delta_3] \to (\mathbb{Z}/p\mathbb{Z})[X]$ is non-split for every prime integer $p > 0$.

We first note that if $Z_1,\, Z_2, Z_3$ are indeterminates and B is any base ring, then

$$H(B,\, Z) = \frac{B[Z_1,\, Z_2,\, Z_3]_{Z_1Z_2Z_3}}{B[Z_1,\, Z_2,\, Z_3]_{Z_2Z_3} + B[Z_1,\, Z_2,\, Z_3]_{Z_1Z_3} + B[Z_1,\, Z_2,\, Z_3]_{Z_1Z_2}}$$

is nonzero: in fact, the numerator is the free B-module spanned by *all* monomnials $Z_1^{a_1}Z_2^{a_2}Z_3^{a_3}$ where $a_1,\, a_2,\, a_3 \in \mathbb{Z}$, and the denominator is the free B-module spanned by all such monomials in which one of the integers $a_1,\, a_2,\, a_3$ is nonnegative. Hence, the quotient may be identified with the free B-module spanned by all monomials $Z_1^{a_1}Z_2^{a_2}Z_3^{a_3}$ such that $a_1, a_2, a_3 < 0$. Since $\Delta_1,\, \Delta_2,\, \Delta_3$ are algebraically independent over $\mathbb{C}$ and, hence, over $\mathbb{Q}$, we have that $H(\mathbb{Q},\, \Delta_1,\, \Delta_2,\, \Delta_3) = H(\mathbb{Q},\, \Delta)$ is a nonzero vector space over $\mathbb{Q}$. We have a comutative diagram:

$$\begin{array}{ccc} H(\mathbb{C},\, \Delta) & \xrightarrow{\quad\iota\quad} & H(\mathbb{C},\, \Delta) \otimes_{\mathbb{C}[\Delta]} \mathbb{C}[X] \\ \uparrow & & \uparrow \\ H(\mathbb{Q},\, \Delta) & \longrightarrow & H(\mathbb{Q},\, \Delta) \otimes_{\mathbb{Q}[\Delta_1,\, \Delta_2,\, \Delta_3]} \mathbb{Q}[X] \end{array} .$$

The top row may be thought of as obtained from the bottom row by applying $\mathbb{C} \otimes_{\mathbb{Q}} \_$.

We next observe that because $\iota : \mathbb{C}[\Delta_1,\, \Delta_2,\, \Delta_3] \subseteq \mathbb{C}[X]$ is split by the Reynolds operator for the action of $\mathrm{SL}(2,\, \mathbb{C})$, and the top row is obtained by tensoring this inclusion over $\mathbb{C}[\Delta_1,\, \Delta_2,\, \Delta_3]$ with $H(\mathbb{C},\, \Delta)$, the top arrow is an injection. Since $\mathbb{C}$ is free and therefore faithfully flat over $\mathbb{Q}$, the arrow in the bottom row is also an injection. Thus, $H(\mathbb{Q},\, \Delta) \otimes_{\mathbb{Q}[\Delta_1,\, \Delta_2,\, \Delta_3]} \mathbb{Q}[X]$ is a nonzero vector space over $\mathbb{Q}$, and this is the same as the result of apply $\mathbb{Q} \otimes_{\mathbb{Z}} \_$ to

$$H(\mathbb{Z},\, \Delta) \otimes_{\mathbb{Z}[\Delta_1,\, \Delta_2,\, \Delta_3]} \mathbb{Z}[X] = \frac{\mathbb{Z}[X]_{\Delta_1\Delta_2\Delta_3}}{\mathbb{Z}[X]_{\Delta_2\Delta_3} + \mathbb{Z}[X]_{\Delta_1\Delta_3} + \mathbb{Z}[X]_{\Delta_1\Delta_2}}$$

which is the module H described earlier.

Finally, we shall show that $H = pH$ for every prime integer $p > 0$, and from this we deduce that $(\mathbb{Z}/p\mathbb{Z})[\Delta_1,\, \Delta_2,\, \Delta_3] \to (\mathbb{Z}/p\mathbb{Z})[X]$ is non-split for every prime integer $p > 0$. Note that $H/pH = (\mathbb{Z}/p\mathbb{Z}) \otimes_{\mathbb{Z}} H$. If $(\mathbb{Z}/p\mathbb{Z})[\Delta_1,\, \Delta_2,\, \Delta_3] \to (\mathbb{Z}/p\mathbb{Z})[X]$ splits over $(\mathbb{Z}/p\mathbb{Z})[\Delta_1,\, \Delta_2,\, \Delta_3]$ then by applying $\_ \otimes_{\mathbb{Z}/p\mathbb{Z}} H(\mathbb{Z}/p\mathbb{Z}, \Delta)$ we obtain in injection

$$H(\mathbb{Z}/p\mathbb{Z}, \Delta) \to H/pH.$$

The lefthand term is not zero, and this will imply that $H/pH \neq 0$. Thus, by showing that $H/pH = 0$, we also show that

$$(\mathbb{Z}/p\mathbb{Z})[\Delta_1,\, \Delta_2,\, \Delta_3] \to (\mathbb{Z}/p\mathbb{Z})[X]$$

does not split.

The final step involves some explicit use of local cohomology theory. We refer to to the Lecture of December 8 from Math 711, Fall 2006, which contains a concise treatment of the

material we need here as well as further references, but we give a brief description, including one definition of the functor Ext. A detailed treatment of Ext is given in the Lecture Notes from Math 615, Winter 2004. There is a discussion of homotopic maps of complexes in the Lectures of February 2 and February 4: it is used to prove the independence of Ext from the choice of projective resolution in the definition below. Ext itself is defined in the Lecture of March 22 from the same set of Lecture Notes.

First recall that if M, N are modules over R, the modules $\mathrm{Ext}^i_R(M,\, N)$ are defined as follows. Choose a free (or projective) resolution of M, i.e., an exact complex

$$\cdots \to P_i \to \cdots \to P_0 \to M \to 0$$

such that the P_i are free (or projective). This complex will frequently be infinite. Let $P_\bullet$ be the complex obtained by replacing M by 0, i.e.,

$$\cdots \to P_i \to \cdots \to P_0 \to 0.$$

Apply the contravariant functor $\mathrm{Hom}_R(\_,\, N)$ to this complex to obtain:

$$0 \to \mathrm{Hom}_R(P_0,\, N) \to \cdots \to \mathrm{Hom}_R(P_i,\, N) \to \cdots .$$

Then $\mathrm{Ext}^i_R(M,\, N)$ is the cohomology of the complex at the $\mathrm{Hom}_R(P_i,\, N)$ spot (this is still the kernel of the outgoing map at that spot modulo the image of the incoming map: it is called *cohomology* because the maps increase the indices).

There are other definitions: one may use an injective resolution of N instead, for example, and there are formulations of the theory where neither projectives nor injectives are used. $\mathrm{Ext}^i(M,\, N)$ is independent of the choice of the projective resolution up to canonical (choice-free) isomorphism. If M is held fixed, $\mathrm{Ext}^i_R(M,\, N)$ is a covariant functor of N. If N is held fixed, it is a contravariant functor of M. The functor $\mathrm{Ext}^0_R(M,\, N)$ may be identified canonically with $\mathrm{Hom}_R(M,\, N)$. The elements of $\mathrm{Ext}^1_R(M,\, N)$ are in bijective correspondence with isomorphism classes of short exact sequence $0 \to N \to W \to M \to 0$: the reason for the name "Ext" is that $\mathrm{Ext}^1_R(M,\, N)$ classifies such extensions.

There are two long exact sequences associated with Ext. If $0 \to N_1 \to N_2 \to N_3 \to 0$ is a short exact sequence of R-modules, then there is a long exact sequence

$$0 \to \mathrm{Hom}_R(M,\, N_1) \to \mathrm{Hom}_R(M,\, N_2) \to \mathrm{Hom}_R(M,\, N_3) \to \mathrm{Ext}^1_R(M,\, N_1) \to \cdots$$

$$\to \mathrm{Ext}^i_R(M,\, N_1) \to \mathrm{Ext}^i_R(M,\, N_2) \to \mathrm{Ext}^i_R(M,\, N_3) \to \mathrm{Ext}^{i+1}_R(M,\, N_1) \to \cdots .$$

Similarly, if $0 \to M_1 \to M_2 \to M_3 \to 0$ is exact there is a long exact sequence

$$0 \to \mathrm{Hom}_R(M_3,\, N) \to \mathrm{Hom}_R(M_2,\, N) \to \mathrm{Hom}_R(M_1,\, N) \to \mathrm{Ext}^1_R(M_3,\, N) \to \cdots$$

$$\to \mathrm{Ext}^i_R(M_3,\, N) \to \mathrm{Ext}^i_R(M_2,\, N) \to \mathrm{Ext}^i_R(M_1,\, N) \to \mathrm{Ext}^{i+1}_R(M_3,\, N) \to \cdots .$$

The module $\operatorname{Ext}^i_R(M,\, N)$ is killed both by $\operatorname{Ann}_R M$ and $\operatorname{Ann}_R N$. When R is Noetherian and M, N are finitely generated, one can calculate the modules $\operatorname{Ext}^i_R(M,\, N)$ using a free resolution of M by finitely generated free R-modules, and it follows that all of the modules $\operatorname{Ext}^i_R(M,\, N)$ are finitely generated R-modules in this case.

If R is Noetherian, $I = (f_1,\, \ldots\, , f_s)$ is an ideal of R, and M is any R-module, the i th *local cohomology module* of M with support in I is defined as

$$\varinjlim_t \operatorname{Ext}^i(R/I_t,\, M)$$

where I_t runs through any sequence of ideals cofinal with the powers of I. In particular, we may take $I_t = I^t$ for all t, but, as we shall see below, other choices of I can be advantageous. It follows that $H^i_I(M)$ depends only on the the radical of I and not on I itself.

The main result that we are going to assume without proof here is that $H^i_I(M)$ is also the cohomology at the i th spot of the complex

$$(*)\quad 0 \to M \to \bigoplus_{1\le j\le s} M_{f_i} \to \cdots \to \bigoplus_{1\le j_1<j_2<\cdots j_i\le s} M_{f_{j_1}f_{j_2}\cdots f_{j_i}} \to \cdots \to M_{f_1 f_2\cdots f_s} \to 0.$$

If we think of the i th term as a direct sum and the $i+1$ st term as a direct product, the maps are determined by specifying maps $M_{f_{j_1}\cdots f_{j_i}} \to M_{f_{k_1}\cdots f_{k_{i+1}}}$, where $j_1 < \cdots < j_i$ and $k_1 < \cdots < k_{i+1}$. The map is 0 unless, $\{j_1,\, \ldots\, , j_i\}$ is obtained from $\{k_1,\, \ldots\, , k_{i+1}\}$ by omitting one term, sayt k_t, and then the map is $(-1)^{t-1}\theta$ where θ is the natural map induced by localizing "further" at f_{k_t}.

By the description of local cohomology in $(*)$ above, the module

$$H/pH = \frac{(\mathbb{Z}/p\mathbb{Z})[X]_{\Delta_1\Delta_2\Delta_3}}{(\mathbb{Z}/p\mathbb{Z})[X]_{\Delta_2\Delta_3} + (\mathbb{Z}/p\mathbb{Z})[X]_{\Delta_1\Delta_3} + (\mathbb{Z}/p\mathbb{Z})[X]_{\Delta_1\Delta_2}}$$

is precisely the local cohomology module $H^3_I\big((\mathbb{Z}/p\mathbb{Z})[X]\big)$ where $I = (\Delta_1,\, \Delta_2,\, \Delta_3)S$, where $S = (\mathbb{Z}/p\mathbb{Z})[X]$. On the other hand, from the definition above this local cohomology module is

$$\varinjlim_t \operatorname{Ext}^3_S(S/I_t,\, S),$$

where I_t is any sequence of ideals cofinal with the powers of I. In our case, we use $I_t = I^{[p^t]}$. The proof is completed by showing that for all t, there is a free resolution of R/I_t over R of length 2. Hence, every $\operatorname{Ext}^3_S(S/I_t,\, S)$ vanishes. For $I = I_1$ itself, we leave it as an exercise to show that

$$0 \to S^2 \xrightarrow{\beta} S^3 \xrightarrow{\alpha} S \to S/I \to 0$$

is such a resolution, where $\alpha = \begin{pmatrix}\Delta_1 & -\Delta_2 & \Delta_3\end{pmatrix}$ and the matrix of β is the transpose of X. The case of I_t follows at once by applying $S \otimes_S \_$, where the map $S \to S$ is the t th iteration F^t of the Frobenius endomorphism, to this complex. Since S is faithfully flat over itself via this map, the new complex is exact, and provides a free resolution of S/I_t of length 2. $\square$

Lecture of February 26

We next want to prove that the algebraic torus $\mathrm{GL}(1, K)^s$, which we shall refer to simply as a *torus*, is linearly reductive, as asserted earlier, over every algebraically closed field K, regardless of characteristic. The notation G_{m} is also used for the multiplicative group of K viewed as a linear algebraic group via its isomorphism with $\mathrm{GL}(1, K)$.

Until further notice, K denotes an algebraically closed field. Let G be any linear algebraic group over K. Let $K[G]$ be its coordinate ring, whose elements may be thought of as the regular maps of the closed algebraic set G to K. (This notation has some danger of ambiguity, since $K[G]$ is also used to denote the group ring of G over K, but we shall only use this notation for the coordinate ring here.) The right action of G on itself by multiplication (i.e., γ acts so that $\eta \mapsto \eta\gamma$) induces a (left) action of G on the K-vector space $K[G]$. Thus, if $f \in K[G]$, $\gamma(f)$ denotes the function whose value on $\eta \in G$ is $f(\eta\gamma)$. Since right multiplication by γ is a regular map of $G \to G$, the composition with $f : G \to K$ is also regular.

Discussion: regularity of the action of G on $K[G]$. We study the map

$$G \times K[G] \to K[G]$$

and prove that it gives an action in our sense. Let $f \in K[G]$. Let μ be the multiplication map $G \times G \to G$. The function $(\eta, \gamma) \mapsto f(\eta\gamma)$ is the composite $f \circ \mu$, and so is a regular function on $G \times G$. Therefore, it is an element of

$$K[G \times G] \cong K[G] \otimes_K K[G],$$

and consequently can be written in the form

$$\sum_{i=1}^{k} g_i \otimes h_i$$

where the g_i, $h_i \in K[G]$. This means that for every fixed γ,

$$(*) \quad \gamma(f) = \sum_{t=1}^{k} h_t(\gamma) g_t.$$

Hence, all of the functions $\gamma(f)$ are in the K-span of the g_i, and this is finite-dimensional. It follows that $K[G]$ is a union of finite-dimensional G-stable subspaces V. Let $f_1, \ldots, f_n$ be a basis for one such V. For every f_i in the basis we have a formula like $(*)$ of the form

$$(*_i) \quad \gamma(f_i) = \sum_{t=1}^{k} h_{it}(\gamma) g_{it}.$$

A priori, k may vary with i but we can work with the largest value of k that occurs. Hence, for $c_1, \ldots, c_n \in K^n$ we have

$$(**) \quad \gamma(\sum_{i=1}^{n} c_i f_i) = \sum_{t=1}^{k} \sum_{i=1}^{n} c_i h_{it}(\gamma) g_{it}.$$

Let Θ be a K-vector space retraction of the K-span of the g_{it} to V. Since Θ fixes the element on the left hand side, which is in V, applying Θ to both sides yields:

$$(\#) \quad \gamma(\sum_{i=1}^{n} c_i f_i) = \sum_{t=1}^{k} \sum_{i=1}^{n} c_i h_{it}(\gamma) \Theta(g_{it}).$$

Here, each $\Theta(g_{it})$ is a fixed linear combination of $f_1, \ldots, f_n$, and although we do not carry this out explicitly, the right hand side can now be rewritten as a linear combination of $f_1, \ldots, f_n$ such that coefficients occurring are polynomials in the regular functions h_{it} on G and the coefficients $c_1, \ldots, c_n$ parametrizing $V \cong K^n$. It follows at once that the action of G on V is regular for every such V. $\square$

We next note:

Theorem. *Let G be a linear algebraic group over a field K, and let N be a fnite dimensional G-module. Then N is isomorphic with a submodule of $K[G]^{\oplus h}$ for some h.*

Proof. Let $\theta : N \to K$ be an arbitary K-linear map. We define a K-linear map

$$\theta^{\vee} : N \to K[G]$$

which will turn out to be a map of G-modules as follows: if $v \in N$, let $\theta^{\vee}(v)$ denote the function on G whose value on $\gamma \in G$ is $\theta\big(\gamma(v)\big)$. Since the map $G \times N \to N$ that gives the action of G on N is a regular map, for fixed $v \in N$ the composite

$$G \cong G \times \{v\} \subseteq G \times N \to N$$

is a regular map from $G \to N$ whose composite with the linear functional $\theta : N \to K$ is evidently regular as well. Hence, $\theta^{\vee}(v) \in K[G]$. This map is clearly linear in v, since θ and the action of γ on N are K-linear. Moreover, for any $\eta \in G$ and $v \in N$, $\theta^{\vee}\big(\eta(v)\big) = \eta\big(\theta^{\vee}(v)\big)$: the value of either one on $\gamma \in G$ is, from the appropriate definition, $\theta\big(\gamma(\eta(v))\big)$.

Choose a basis $\theta_1, \ldots, \theta_h$ for $\mathrm{Hom}_K(N, K)$. Then the map $N \to K[G]^{\oplus h}$ that sends $v \mapsto \theta_1^{\vee}(v) \oplus \cdots \oplus \theta_h^{\vee}(v)$ is a G-module injection of N into $K[G]^{\oplus h}$. To see this, note that if $v \neq 0$, it is part of a basis, and there is a linear functional whose value on v is not 0. It follows that for some i, $\theta_i(v) \neq 0$. But then $\theta_i^{\vee}(v) \neq 0$, since its value on the identity element of G is $\theta_i(v) \neq 0$. $\square$

Lemma. *If M is G-module and is a direct sum of irreducibles $\{N_\lambda\}_{\lambda\in\Lambda}$, then every G-submodule N of M is isomorphic to the direct sum of the irreducibles in a subfamily of $\{N_\lambda\}_{\lambda\in\Lambda}$, and N has a complement that is the (internal) direct sum of a subfamily of the $\{N_\lambda\}_{\lambda\in\Lambda}$.*

Proof. Let N be a given submodule of M. We first construct a complement N' of the specified form. By Zorn's Lemma there is a maximal subfamily of $\{N_\lambda\}_{\lambda\in\Lambda}$ whose (direct) sum N' is disjoint from N. We claim that $M = N \oplus N'$. We need only check that $M = N + N'$. If not, some irreducible N_{λ_0} in the family is not contained in $N + N'$. But then its intersection with $N + N'$ must be 0, and we can enlarge the subfamily by using N_{λ_0} as well.

By the same argument, N' has a complement N'' in M that is a direct sum of a subfamily of $\{N_\lambda\}_{\lambda\in\Lambda}$. Then since $M = N \oplus N'$, $N \cong M/N'$, while since $M = N'' \oplus N'$, $M/N' \cong N''$. Thus, $N \cong N''$, which shows that N is isomorphic with a direct sum of a subfamily of the irreducibles as required. □

Corollary of the Theorem. *If G is a linear algebraic group over K and $K[G]$ is a direct sum of irreducible G-modules $\{N_\lambda\}_{\lambda\in\Lambda}$, then G is linearly reductive, and every G-module is isomorphic to a direct sum of irreducible G-modules in this family. In particular, up to isomorphism, every irreducible G-module is in this family.*

Proof. By the Theorem above, every finite-dimensional G-module N is a submodule of $K[G]^{\oplus h}$ for some h, and this module is evidently a direct sum of irreducibles from the same family. The result now follows from the Lemma just above. □

We next want to apply this Corollary to the case where $G = \mathrm{GL}(1, K)^s$ is a torus. Fix an s-tuple of integers $k_1, \ldots, k_s \in \mathbb{Z}^s$. One example of an action of G on a one-dimensional vector space Kx is the action such that $\gamma = (\gamma_1, \ldots, \gamma_s)$ sends

$$x \mapsto \gamma_1^{k_1} \cdots \gamma_s^{k_s} x$$

for all $\gamma \in G$. Because the vector space is one-dimensional, this G-module is clearly irreducible. We can now prove that for this G, every G-module is a direct sum of irreducibles of this type.

Theorem. *Let K be a field and let $G = \mathrm{GL}(1, K)^s$ be a torus. Then G is linearly reductive, and every G-module is a direct sum of one-dimensional G-modules of the type described just above.*

Proof. $K[G]$ is the tensor product of s copies of the coordinate ring of $GL(1, K)$, and may be identified with $K[x_1, x_1^{-1}, \ldots, x_s, x_s^{-1}]$. The action of G on this ring is such that $\gamma = (\gamma_1, \ldots, \gamma_s)$ sends $x_i \mapsto \gamma_i x_i$, $1 \le i \le s$. It follows at once that $\mu = x_1^{k_1} \cdots x_s^{k_s}$, where $(k_1, \ldots, k_s) \in \mathbb{Z}^s$, is mapped to $\gamma_1^{k_1} \cdots \gamma_s^{k_s} \mu$ for every $\gamma = (\gamma_1, \ldots, \gamma_s) \in G$, and so $K[G]$ is the direct sum of copies of G-modules as described just above, one for every monomial μ. The result is now immediate from the Corollary of the Theorem. □

Discussion: degree-preserving actions of a torus on a polynomial ring. We keep the assumption that K is an algebraically field, although we shall occasionally be able to relax it in the statements of some results: this will always be made explicit. The last statement in the Theorem below is an example.

Let $G = \mathrm{GL}(1, K)^s$ act by degree-preserving K-algebra automorphisms on the polynomial ring R in n variables over K so that R is a G-module. Giving such an action is the same as making the one forms $[R]_1$ of R into a G-module: the action then extends uniquely and automatically to R. Given such an action we may write $[R]_1$ as a direct sum of one-dimensional irreducible G-modules as above. Therefore, we may choose a basis $x_1, \ldots, x_n$ for $[R]_1$ over K so that for every j, Kx_j is a G-stable submodule. It follows that for every j we can choose integers $k_{1,j}, \ldots, k_{s,j} \in \mathbb{Z}$ such that for all $\gamma = (\gamma_1, \ldots, \gamma_s) \in G$, γ sends

$$x_j \mapsto \gamma_1^{k_{1,j}} \cdots \gamma_s^{k_{s,j}} x_j.$$

Thus, the action of G on $R = K[x_1, \ldots, x_n]$ is completely determined by the $s \times n$ matrix $\big(k_{i.j}\big)$ of integers. Every action comes from such a matrix, and for every such matrix there is a corresponding action.

Now consider any monomial $\mu = x_1^{a_1} \cdots x_n^{a_n}$ of R. For all $\gamma = (\gamma_1, \ldots, \gamma_s) \in G$, γ sends

$$\mu \mapsto \big(\prod_{i=1}^{s} (\gamma_i^{k_{i,1}a_1 + \cdots + k_{i,n}a_n})\big)\mu.$$

It is now easy to see that the ring of invariants is spanned over K by all monomials $x_1^{a_1} \cdots x_n^{a_n}$ such that the s homogeneous linear equations

$$\sum_{j=1}^{n} k_{i,j} a_j = 0$$

are satisfied.

We have proved:

Theorem. *A ring generated by monomials arises as the ring of invariants of an action of a torus as above if and only if the ring is spanned over K by the monomials x^α where α runs through the solutions in $\mathbb{N}^n$ of some family of s homogenous linear equations over $\mathbb{Z}$ in n unknowns. Consequently, any such ring is Cohen-Macaulay, whether the field is algebraically closed or not.* $\square$

Of course, the Cohen-Macaulay property follows because of our result on rings of invariants of linearly reductive linear algebraic groups acting on polynomial rings. If the field K is not algebraically closed, we may use the fact that the Cohen-Macaulay property is not affected when we tensor over K with its algebraic closure $\overline{K}$: see problem 4(d) of Problem Set #2 and its solution.

Example: the ring defined by the vanishing of the 2×2 *minors of a generic matrix.* Let $G = GL(1, K)$ acting on $K[x_1, \ldots, x_r, y_1, \ldots, y_s]$, where $x_1, \ldots, x_r, y_1, \ldots, y_s$ are $r+s$ algebraically independent elements, so that if $\gamma \in G$, then $x_i \mapsto \gamma x_i$ for $1 \leq i \leq r$ and $y_i \mapsto \gamma^{-1} y_i$ for $1 \leq i \leq s$. Here, there is only one copy of the multiplicative group, and so there is only one equation in the system:

$$x_1^{a_1} \cdots x_r^{a_r} y_1^{b_1} \cdots y_s^{b_s}$$

is invariant if and only if

$$a_1 + \cdots + a_r - b_1 - \cdots - b_s = 0.$$

That is, the ring of invariants is spanned over K by all monomials μ such that the total degree of μ in the variables $x_1, \ldots, x_r$, which is $a_1 + \cdots a_r$, is equal to the total degre of μ in the variables $y_1, \ldots, y_s$, which is $b_1 + \cdots + b_s$.

Each such monomial can written as product of terms $x_i y_j$, usually not uniquely, by pairing each of the x_i occurring in the monomial with one of the y_j occurring. It follows that

$$R^G = K[x_i y_j : 1 \leq i \leq r, 1 \leq j \leq s].$$

Consider an $r \times s$ matrix of new indeterminates $Z = \left(z_{i,j}\right)$. There is a K-algebra surjection

$$K[Z] \twoheadrightarrow K[x_i y_j : 1 \leq i \leq r, 1 \leq j \leq s] = R^G$$

that sends $z_{i,j} \mapsto x_i y_j$ for all i and j. The ideal $I_2(Z)$ is easily checked to be in the kernel, so that we have a surjection $K[Z]/I_2(Z) \twoheadrightarrow R^G$. It is now easy to check that this map is injective, given the result of problem 6. of Problem Set #3, namely, that $I_2(Z)$ is prime. Assuming the result of problem 6, let $\mathcal{F}$ be the fraction field of the domain $D = K[Z]/I_2(Z)$, and let $\overline{z}_{i,j}$ be the image of $z_{i,j}$. It is clear that $z_{1,1}$ has too small a degree to be in $I_2(Z)$, and so $\overline{z}_{1,1} \neq 0$. Since the 2×2 minors of the image $\overline{Z}$ of Z vanish, the matrix $\overline{Z}$ has rank 1 over $\mathcal{F}$. It follows that the ith row of $\overline{Z}$ is $\overline{z}_{i,1}/\overline{z}_{1,1}$ times the first row. Define a a K-algebra map $K[x_1, \ldots, x_r, y_1, \ldots, y_s] \to \mathcal{F}$ by $x_i \mapsto \overline{z}_{i,1}/\overline{z}_{1,1}$ for $1 \leq i \leq r$ and and $y_j \mapsto \overline{z}_{1,j}$ for $1 \leq j \leq s$. Then the restriction to R^G is a K-algebra map $R^G \to K[Z]/I_2(Z)$ that sends $x_i y_j \mapsto \overline{z}_{i,j}$ for all i, j and so is an inverse for ϕ. $\square$

We can now conclude:

Theorem. *Let Z be an $r \times s$ matrix of indeterminates over any field K. Then $K[Z]/I_2(Z)$ is a Cohen-Macaulay domain.* $\square$

We want to prove a somewhat more general result. Recall that a domain D is called *normal* or *integrally closed* if every element of its fraction field that is integral over D is in D.

Theorem. *Let $x_1, \ldots, x_n$ be indeterminates over the field K and let S be any finitely generated normal subring of $K[x_1, 1/x_1, \ldots, x_n, 1/x_n]$ generated by monomials. Then S is Cohen-Macaulay.*

Recall that if $\mathcal{M}$ is a semigroup under multiplication with identity 1, disjoint from the ring B, the semigroup ring $B\langle\mathcal{M}\rangle$ is the free B-module with basis $\mathcal{M}$ with multiplication defined so that if b, $b' \in B$ and μ, $\mu' \in \mathcal{M}$ then $(b\mu)(b'\mu') = (bb')(\mu\mu')$. The general rule for multiplication is then forced by the distributive law. More precisely,

$$\sum_i b_i\mu_i \sum_j b'_j\mu'_j = \sum_\nu (\sum_{\mu_i\mu'_j=\nu} b_i b'_j)\nu$$

where μ, $\mu' \in \mathcal{M}$. It is understood that there are only finitely many nonzero terms in each summation on the left hand side, and this forces the same to be true in the summation on the right hand side.

We will prove the Theorem by showing that each such ring can be obtained from a monomial ring which has the Cohen-Macaulay property by virtue of our Theorem on rings of invariants of tori by adjoining variables and their inverses.

We shall therefore want to characterize the semigroups of exponent vectors in $\mathbb{N}^n$ corresponding to rings of invariants of tori. We already know that such a semigroup is the set of solutions of a finite system of homogeneous linear equations with integer coefficients (we could also say rational coefficients, since an equation can be replace by a nonzero integer multiple to clear denominators). That is, such a semigroup is the intersection of a vector subspace of $\mathbb{Q}^n$ with $\mathbb{N}^n$. It also follows that H is a such a semigroup if and only if it has the following two properties:

(1) If α, $\alpha' \in H$ and $\beta = \alpha - \alpha' \in \mathbb{N}^n$ then $\beta \in H$.

(2) If $\beta \in \mathbb{N}^n$ and $k\beta \in H$ for some integer $k > 0$, then $\beta \in H$.

If H is the intersection of a $\mathbb{Q}$-subspace of $\mathbb{Q}^n$ with $\mathbb{N}^n$, then it must be the intersection of the subspace it spans with $\mathbb{N}$. The abelian group that H spans is

$$H - H = \{\alpha - \alpha' : \alpha,\, \alpha' \in H\}.$$

Let $\mathbb{Q}^+ = \{u \in \mathbb{Q} : u > 0\}$. The vector space that H spans is then

$$\mathbb{Q}^+(H - H) = \{u\beta : u \in \mathbb{Q}^+,\, \beta \in H - H\}.$$

In fact, this vector space is also

$$\bigcup_{m=1}^{\infty} \frac{1}{m}(H - H)$$

where

$$\frac{1}{m}(H - H) = \{\frac{\beta}{m} : \beta \in H - H\}.$$

The fact that H is the intersection of a $\mathbb{Q}$-vector subspace of $\mathbb{Q}^n$ with $\mathbb{N}^n$ if and only if (1) and (2) hold follows at once.

Lecture of March 7

We next want to consider when a K-subalgebra of $S = K[x_1, 1/x_1, \ldots, x_n, 1/x_n]$ generated by monomials is normal. This is entirely a property of the semigroup of monomials involved, and does not depend on the base field.

We shall typically work with the additive semigroup of exponent vectors, which is a subsemigroup H of $\mathbb{Z}^n$. If $\alpha = (a_1, \ldots, a_n) \in \mathbb{Z}^n$, we write x^α for $x_1^{a_1} \cdots x_n^{a_n}$. Then the K-subalgebras of S generated by monomials correspond bijectively to the subsemigroups H of $\mathbb{Z}^n$: given H, the corresponding subalgebra is the K-span of $\{x^\alpha : \alpha \in H\}$.

If H is an additive (which we intend to imply commutative) semigroup such that cancellation holds, i.e., if $\alpha, \alpha', \beta \in H$ and $\alpha + \beta = \alpha' + \beta$ then $\alpha = \alpha'$, then there is an essentially unique way to enlarge H to group that is generated by H. Define an equivalence relation on $H \times H$ by the rule $(\alpha, \beta) \sim (\alpha', \beta')$ precisely when $\alpha + \beta' = \alpha' + \beta$. The equivalence classes form a semigroup such that

$$[(\alpha_1, \beta_1) + [(\alpha_2, \beta_2)] = [(\alpha_1 + \alpha_2, \beta_1 + \beta_2)].$$

H embeds in this new semigroup by sending $\alpha \mapsto [(\alpha, 0)]$. The 0 element is represented by $(0, 0)$ and also by those elements of the form (α, α). There are now inverses since $[(\alpha, \beta)] + [(\beta, \alpha)] = [(\alpha + \beta, \alpha + \beta)] = [(0, 0)]$. In particular, $[(\beta, 0)]$ has additive inverse $[(0, \beta)]$. Thus, the new semigroup is a group, and if we identify $\alpha \in H$ with its image, then every element of this group has the form $\alpha - \beta$ for choices of $\alpha, \beta \in H$. We denote this group $H - H$. If we have any other injection of H into a semigroup G that is a group, then the subgroup of G generated by H is isomorphic with $H - H$.

In particular, when H is a subsemigroup of $\mathbb{Z}^n$, the group $H - H$ depends only on H, not on its embedding in $\mathbb{Z}^n$.

We define $H \subseteq \mathbb{Z}^n$ to be *normal* if whenever $\alpha, \alpha' \in H$ and there is a positive integer k such that $k(h - h') \in H$, then $h - h' \in H$.

Theorem. *For every field K, $R = K[x^\alpha : \alpha \in H]$ is normal if and only if H is normal.*

Proof. First suppose that the subalgebra R is normal, and that $k(\alpha - \alpha') \in H$, where k is a positive integer. Then $x^\alpha, x^{\alpha'} \in R$, and $f = x^\alpha / x^{\alpha'} = x^{\alpha - \alpha'}$ is an element of the fraction field integral over R, since $f^k \in R$. Hence, $f \in R$, and so $\alpha - \alpha' \in H$.

We next show that the condition that H be normal is sufficient for R to be normal. Suppose that we can solve the problem when K is an infinite field, e.g., an algebraically closed field. If K is finite, let L be an infinite field containing K. Then

$$R = K[x_1, 1/x_1, \ldots, x_n, 1/x_n] \cap L[x^\alpha : \alpha \in H],$$

and since both the rings being intersected are normal, R is normal as well.

Therefore we may assume that K is infinite. The group of invertible diagonal matrices $\mathcal{D}_n$ acts on S, and R is stable. The Theorem at the top of p. 5 of the Lecture Notes of February 12, which asserts that, when K is infinite, vector subspaces of $K[x_1, \ldots, x_n]$ stable under $\mathcal{D}_n$ are spanned by monomials, extends at once to S: the preceding Proposition from that Lecture is valid for $\mathbb{Z}$-gradings as well as $\mathbb{N}$-gradings. Thus, the integral closure of R will be spanned by monomials as well. Consider the ring obtained by adjoining the inverses of all monomials in R. This ring R_1 corresponds to $H - H$, which is isomorphic with a free abelian group $\mathbb{Z}^h$, and so R_1 is isomorphic with a localized polynomial ring obtained by adjoining h algebraically independent elements and their inverses to K. Thus, R_1 is normal, and so any monomial in the normalization of R is in R_1.

It follows that if R is not normal, then there is a monomial $\mu = x^\alpha / x^{\alpha'}$, where $\alpha, \alpha' \in H$, that is integral over R and not in R. Choose a monic polynomial $F(Z)$ with coefficients in R of degree k satisfied by μ. Assign Z the same monomial degree as μ. Then the sum of the terms whose monomial degree is μ^k must also vanish when we substitute $Z = \mu$, and so we have an equation of integral dependence that is monomially graded. Since R is a domain, there is no loss of generality in assuming that the constant term is nonzero: if necessary, we may factor out a power of Z. We continue to call the degree k. Then μ^k has the same monomial degree ν as the constant term $c\nu$, where $c \in K - \{0\}$, and ν is a monomial in R. This shows that $k(\alpha - \alpha') \in H$, and so $\alpha - \alpha' \in H$ and $\mu \in R$ after all. $\square$

Example. Let K be any field, let $\lambda \geq 0$ be a real number, and let

$$H_\lambda = \{(a, b) \in \mathbb{N}^2 : a/b > \lambda\}.$$

It is easy to see that if $0 \leq \lambda < \lambda'$ then H_λ is strictly larger than $H_{\lambda'}$. Morever, every H_λ is a normal semigroup. Let $R_\lambda = K[x^\alpha : \alpha \in H_\lambda]$. This gives an uncountable chain $\{R_\lambda\}_{\lambda \geq 0}$ of normal subrings of $K[x_1, x_2]$. None of the rings R_λ is Noetherian: if R_λ were Noetherian, the fact that it is $\mathbb{N}$-graded over K would imply that it is finitely generated by elements

$$x_1^{a_1} x_2^{b_1}, \ldots, x_1^{a_n} x_2^{b_n}$$

with every $a_j/b_j > \lambda$. Let $s > \lambda$ be the minimum of the rational numbers $a_1/b_1, \ldots, a_n/b_n$. Then

$$K[x_1^{a_1} x_2^{b_1}, \ldots, x_1^{a_n} x_2^{b_n}]$$

does not contain any monomial $x^a y^b$ with $a/b < s$, and so cannot be equal to R_λ. $\square$

The Example above shows that the condition of being normal is too weak to imply that a semigroup is finitely generated. We next want to consider a much stronger condition on subsemigroups of $\mathbb{N}^n$ which implies *both* normality and finite generation.

We say that a subsemigroup $H \subseteq \mathbb{N}^n$ is *full* if whenever $\alpha, \alpha' \in H$ and $\alpha - \alpha' \in \mathbb{N}^n$ then $\alpha - \alpha' \in \mathbb{N}$. We observed at the end of the previous lecture that the subsemigroups obtained from rings of invariants of torus actions on polynomial rings are full.

It is obvious that full subsemigroups are normal, for if $k(\alpha-\alpha') \in H$, then $k(\alpha-\alpha') \in \mathbb{N}^n$, and since $k > 0$, this implies that $\alpha - \alpha' \in H$. Something much stronger is true.

Theorem. *Let H be a full subsemigroup of $\mathbb{N}^n$. Let $R = K[x^\alpha : \alpha \in H]$, where K is any field. Then $R \hookrightarrow K[x_1, \ldots, x_n]$ is split. Hence:*

(a) *R is a finitely generated K-algebra, and so H is a finitely generated semigroup.*

(b) *R is Cohen-Macaulay.*

Proof. Let W be the K-span of the monomials x^β for $\beta \in \mathbb{N}^n - H$. Evidently, $K[x_1, \ldots, x_n] = R \oplus W$ as K-vector spaces. To complete the proof that we have a splitting, it suffices to show that W is an R-module. This comes down to the assertion that if $\alpha' \in H$, so that $x^{\alpha'} \in R$, and $\beta \in \mathbb{N}^n - H$, so that $x^\beta \in W$, then $x^{\alpha'}x^\beta \in W$. Suppose not. Then $x^{\alpha'+\beta} = x^\alpha$, where $\alpha \in H$. But this means that $\beta = \alpha - \alpha' \in \mathbb{N}^n$. By the definition of full subsemigroup, $\beta \in H$, a contradiction.

The first statement in part (a) follows from the Lemma at the top of p. 2 of the Lecture Notes of February 17, and the second statement in part (a) is an Immediate consequence. Part (b) the follows from the Theorem at the top of p. 4 of the Lecture Notes of February 17. $\square$

We shall complete the proof that finitely generated normal K-subalgebras of S are Cohen-Macaulay by proving the following

Theorem. *Let H be a finitely generated normal subsemigroup of $\mathbb{Z}^n$. Then $H \cong \mathbb{Z}^k \oplus H'$, where H' is isomorphic to a full subsemigroup of $\mathbb{N}^n$.*

It will then follow that $K[x^\alpha : \alpha \in H]$ is the polynomial ring in k variables with the inverses of the variables adjoined over $K[x^\alpha : \alpha \in H']$. Thus, the remaining work is in the proof of the Theorem just above, most of which we postpone for a bit. However, we can immediately give the part of the argument in which we split off $\mathbb{Z}^k$.

First part of the proof of the Theorem. First, replace $\mathbb{Z}^n$ by $H - H \subseteq \mathbb{Z}^n$. Since a subgroup of $\mathbb{Z}^n$ will also be a finitely generated free abelian group, we may assume that $H - H = \mathbb{Z}^n$ (the property of being a normal semigroup is not affected). Let G be the set of all elements of H with additive inverses in H. Then G contains 0 and is closed under addition. It follows that G is a subgroup of $\mathbb{Z}^n$, and so $G \cong \mathbb{Z}^k$ for some $k \in \mathbb{N}$. We next claim that $\mathbb{Z}^n/G$ is torsion-free. Suppose $\beta \in \mathbb{Z}^n = H - H$ and $k\beta \in G$. Then $k(-\beta) \in G$ as well, and both β and $-\beta$ are in $\mathbb{Z}^n = H - H$. It follows that β and $-\beta$ are both in H, and so $\beta \in G$, as required.

Thus, $\mathbb{Z}^n/G$ is a finitely generated torsion-free abelian group, and it follows that it is free. Thus,

$$0 \to G \to \mathbb{Z}^n \to \mathbb{Z}^n/G \to 0$$

splits. Let $G' \cong \mathbb{Z}^h \cong \mathbb{Z}^n/G$ be a free complement for G in H. Every element $\beta \in H$ can be expressed uniquely as $\alpha + \alpha'$ where $\alpha \in G$ and $\alpha' \in G'$. But $-\alpha \in H$, and so

$\alpha' \in H$. Thus, $H = G \oplus H'$, where $H' = H \cap G'$, and may also be viewed as the image of H under the projection $\mathbb{Z}^n = G \oplus G' \cong G \times G' \twoheadrightarrow G'$. It follows that H' is a finitely generated subsemigroup of G'. Evidently, H' does not contain the additive inverse of any of its nonzero elements, since $G \cap H' = 0$. Moreover, H' is normal: if $\beta \in H - H'$ and $\kappa\beta \in H'$, then $\beta \in H$, and may be written uniquely as $\alpha + \alpha'$ with $\alpha \in G$ and $\alpha' \in H'$. Then $k\alpha + k\alpha' \in H'$, and so $k\alpha = 0$. It follows that $\alpha = 0$, and $\beta = \alpha' \in H'$, as required. The proof of the Theorem above therefore reduces to establishing the following

Lemma. *Let H be a finitely generated normal subsemigroup of $\mathbb{Z}^n$ such that there is no nonzero element with an additive inverse in H. Then H is isomorphic with a full subsemigroup of $\mathbb{N}^s$ for some nonnegative integer s.*

The proof of this Lemma will be carried through by studying a class of semigroups in $\mathbb{Q}^n$ that are closed under multiplication by elements of $\mathbb{Q}^+$, the positive rational numbers. What we need is an understanding of convex geometry over $\mathbb{Q}$.

Geometry in vector spaces over the rational numbers

The results in this section are proved over $\mathbb{Q}$: the statements and proofs are valid with no changes whatsoever if $\mathbb{Q}$ is replaced by any field between $\mathbb{Q}$ and $\mathbb{R}$, including $\mathbb{R}$, or any ordered field. The results are, in fact, more "standard" over $\mathbb{R}$.

Let V be a vector space over $\mathbb{Q}$. By a $\mathbb{Q}^+$-*subsemigroup* C of V we mean a subsemigroup that is closed under multiplication by elements of $\mathbb{Q}^+$. (It would also be natural to refer to C as a *convex cone*: it will be closed under taking all linear combinations with nonnegative coefficients, and will be a union of "rays" emanating from the origin.) Henceforth, V will be assumed finite-dimensional. We say that C is *finitely generated over* $\mathbb{Q}^+$ if it has finitely many elements $\alpha_1, \ldots, \alpha_h$ such that every element of C is a $\mathbb{Q}^+$-linear combination of the elements $\alpha_1, \ldots, \alpha_h$. We write V^* for the $\mathbb{Q}$-vector space $\mathrm{Hom}_{\mathbb{Q}}(V, \mathbb{Q})$, which is finite-dimensional of the same dimension as V. Its elements will be called *linear functionals* on V.

If L is a nonzero linear functional on V, the set $\{\alpha \in V : L(\alpha) \geq 0\}$ is called a *half-space*. The set $\{\alpha \in V : L(\alpha) \leq 0\}$ is also a half-space, since we may replace L by $-L$. We can always choose a basis for V consisting of $n-1$ vectors $e_1, \ldots, e_{n-1}$ in the kernel of V and a vector e_n on which L has the value 1. If we identify V with $\mathbb{Q}^n$ using this basis, the half-space determined by L is is identified with $\{(q_1, \ldots, q_n) \in \mathbb{Q}^n : q_n \geq 0\}$: we refer to this as the *standard example* of a half-space. A half-space is a $\mathbb{Q}^+$-subsemigroup that is finitely generated: it suffices to see this for the standard example. Then generators are the vectors $e_1, \ldots, e_{n-1}$, $-e_1, \ldots, -e_{n-1}$, and e_n.

We shall say that a $\mathbb{Q}^+$-subsemigroup C *has no line* or is a $\mathbb{Q}^+$-subsemigroup *with no line* if there is no nonzero vector in C whose additive inverse is in C: it is equivalent that C does not contain a one-dimensional vector subspace of the ambient space.

If C is a finitely generated $\mathbb{Q}^+$-subsemigroup we may take any set of generators, and choose a minimal subset with the property of generating C over $\mathbb{Q}^+$. We shall call these elements *a minimal set of generators of* C.

Lemma. *Let V be a finite-dimensional $\mathbb{Q}$-vector space.*

(a) *Every finite intersection of half-spaces in V is a finitely generated $\mathbb{Q}^+$-subsemigroup.*

(b) *Let C be be any $\mathbb{Q}^+$-subsemigroup in V. Let W be the subset of C consisting of elements with an additive inverse in C. Then W is a vector subspace of V, and if W' is a vector space complement for W in V, then $C = W \oplus C'$, where $C' = C \cap W'$ is also the projection of C on W'. C' is a finitely generated $\mathbb{Q}^+$-subsemigroup with no line.*

(c) *If C is a $\mathbb{Q}^+$-subsemigroup with no line, $\alpha_1, \ldots, \alpha_h \in C$, $c_1, \ldots, c_h \in \mathbb{Q}^+$, and $c_1\alpha_1 + \cdots + c_h\alpha_h = 0$, then $\alpha_1 = \cdots = \alpha_h = 0$.*

(d) *Let C be a finitely generated $\mathbb{Q}^+$-subsemigroup with no line and let α, β be part of a minimal set of generators for C. Then $C_1 = C + \mathbb{Q}\alpha$, which is the $\mathbb{Q}^+$-subsemigroup generated by C and $-\alpha$, does not contain $-\beta$.*

Proof. For part (a) we use induction on the number of half-spaces. We have already proved the result in the discussion above if there is just one half-space. Thus, we may assume that the intersection of all but one of the half-spaces is a finitely generated $\mathbb{Q}^+$-subsemigroup C, and it suffices to show that the intersection of C with remaining half-space is finitely generated. After a change of basis, we may assume that the last half-space D is the standard example. Let $\alpha_1, \ldots, \alpha_h$ generate C, and let c_j be the last coordinate of α_j, $1 \leq j \leq h$. We may multiply each α_j by $1/|c_j|$ if $c_j \neq 0$ and so assume that every nonzero c_j is 1 or -1. Then

$$C \cap D = \{q_1\alpha_1 + \cdots + q_h\alpha_h : q_j \in \mathbb{Q}_j^+ \text{ for all } j \text{ and } \sum_{j=1}^{h} q_j c_j \geq 0\}.$$

It therefore suffices to show that

$$E = \{(q_1, \ldots, q_h) \in (\mathbb{Q}^+)^h : \sum_{j=1}^{h} q_j c_j \geq 0\}$$

is finitely generated as a $\mathbb{Q}^+$-subsemigroup, because we have a surjective map $E \twoheadrightarrow C \cap D$ sending

$$(q_1, \ldots, q_h) \mapsto q_1\alpha_1 + \cdots + q_h\alpha_h.$$

This map will carry a finite set of generators for E to a finite set of generators for $C \cap D$. We may assume that coordinates have been permuted so that we have $c_1 = \cdots = c_a = 1$, $c_{a+1} = \cdots = c_{a+b} = -1$, and the remaining c_j are 0. It is easy to verify that the e_i for $1 \leq i \leq a$, the $e_i + e_j$ for $1 \leq i \leq a$ and $a+1 \leq j \leq b$, and the e_k for $a+b+1 \leq k \leq h$ generate E over $\mathbb{Q}^+$.

Part (b) is entirely similar to the construction of the splitting $H = G \oplus H'$ except that it is much simpler in the present context, and the proof is left as an exercise.

For part (c), if some c_j is not 0, say c_h, then

$$-\alpha_h = \frac{c_1}{c_h}\alpha_1 + \cdots + \frac{c_{h-1}}{c_h}\alpha_{h-1},$$

contradicting the assumption that C has no line.

Finally, for part (d), suppose

$$-\beta = \eta - c\alpha,$$

where we may assume $c > 0$ or else $-\beta \in H$. The element η can be written as a nonnegative linear combination of α, β, and the other minimal generators, say

$$\eta = q\alpha + r\beta + \eta',$$

where η' does not involve α or β. Then

$$-\beta = q\alpha + r\beta + \eta' - c\alpha,$$

and so

$$(q - c)\alpha + (r + 1)\beta + \eta' = 0.$$

If $q \geq c$ this contradicts part (c). If $q < c$, then

$$\alpha = \frac{r+1}{c-q}\beta + \frac{1}{c-q}\eta',$$

which means that α is not needed as a generator, a contradiction. $\square$

Proposition. *Let V be a finite-dimensional vector space over $\mathbb{Q}$ and let $C \subseteq V$ be a finitely generated $\mathbb{Q}^+$-subsemigroup. If C is proper, then C is contained in a half-space, i.e., there is a nonzero linear functional that is nonnegative on C. If $\alpha \in C$ and $-\alpha \notin C$ then one can choose L nonnegative on C so that $L(\alpha) > 0$. If C contains no line, one can choose L so that it is positive on all nonzero elements of C.*

Proof. We use induction on $\dim_{\mathbb{Q}}(V)$, and assume that all of the statements are true for vector spaces of smaller dimension. We may replace V by $C - C$, and so assume that C spans V. If $\dim(V) = 1$ then C is either $\{0\}$, a half-line, or all of V, and the result is trivial.

In general, we have a decomposition $C = W + C'$ where W is a vector space as in part (b) of the Lemma, and $C' \subseteq W'$, a complement for W. If $W \neq 0$ then all of the statements can now be deduced from the induction hypothesis applied to $C' \subseteq W'$: one extends the

functional on W' by letting it be 0 on W. Note that if $\alpha \in C$ and $-\alpha \notin C$ then $\alpha = \beta + \alpha'$ where $\beta \in W$ and $\alpha' \in C' - \{0\}$, and has no additive inverse in C'.

This means that we can assume without loss of generality that C has no line, and we may choose minimal generators $\alpha_1, \ldots, \alpha_h$. We must have $h \geq 2$, or else $\dim_{\mathbb{Q}}(V) \leq 1$, since C spans V. It will suffice to construct a linear functional L_i that is positive on α_i and nonnegative on C for every i. The sum of these linear functionals will be positive on all of $C - \{0\}$, since every element is nonnegative linear combination of the α_i. Thus, it suffices to construct such a functional that is nonnegative on, say, α_1. Let $\alpha = \alpha_2$ and $\beta = \alpha_1$. We apply part (d) of the Lemma above, and replace C by $C_1 = C + \mathbb{Q}\alpha$. Then β does not have an inverse, but C_1 contains a line, and so we can construct a linear functional nonnegative on C_1 and positive on $\beta = \alpha_1$ by reducing to a lower-dimensional case, as in the preceding paragraph. $\square$

Theorem. *Let V be a finite-dimensional vector space over $\mathbb{Q}$. Then $C \subseteq V$ is a finitely generated $\mathbb{Q}^+$-subsemigroup if and only if C is a finite-intersection of half-spaces.*

Proof. The "if" part is part (a) of the Lemma. It remains to see that every $\mathbb{Q}^+$-subsemigroup is a finite intersection of half-spaces. Let $\alpha_1, \ldots, \alpha_h$ be a finite set of generators. The set of linear functionals nonnegative on α_i is a half-space H_i in the dual vector space V^*, and so the intersection of the H_i is a finitely generated $\mathbb{Q}^+$-subsemigroup in V^*. Let $L_1, \ldots, L_s$ be generators. It suffices to show that C is the intersection of the half-spaces determined by the L_j. Let β be any vector not in C. It will suffice to show that there exists a linear functional that is nonnegative on C and negative on β, for this functional is a nonnegative linear combination of the L_j, and so at least one of the L_j will have the same property. Consider

$$C_1 = C + \mathbb{Q}^+(-\beta),$$

the $\mathbb{Q}^+$-subsemigroup generated by C and $-\beta$. If $\beta \in C_1$ we have

$$\beta = \alpha - c\beta$$

with $\alpha \in C$ and $c > 0$ and then

$$\beta = \frac{1}{1+c}\alpha \in C,$$

a contradiction. Since $\beta \notin C_1$, by the Proposition above there is a linear functional that is positive on $-\beta$ and nonnegative on C_1, and this has the required property. $\square$

Lecture of March 9

We now have established the results that we need about convex geometry over the rational numbers, and we are ready to prove the Lemma from the top of p. 4 of the Lecture Notes of March 7, which will also complete the proof that normal subrings of $K[x_1, 1/x_1, \ldots, x_n, 1/x_n]$ generated by finitely many monomials are Cohen-Macaulay.

Proof of the Lemma on embedding normal subsemigroups as full subsemigroups of $\mathbb{N}^s$. Let $H \subseteq \mathbb{Z}^n$ be a finitely generated normal subsemigroup that does not contain the additive inverse of any of its nonzero elements. We want to show that H can be embedded as a full subsemigroup in $\mathbb{N}^s$ for some s. First note that $H - H$ is a free abelian group, and so we may replace $\mathbb{Z}^n$ by $H - H$. Henceforth, we assume that $H - H = \mathbb{Z}^n$. This does not affect the condition that H be normal. Second, let $C = \mathbb{Q}^+H$ be the $\mathbb{Q}^+$-subsemigroup generated by H. It is generated over $\mathbb{Q}^+$ by the generators of H, and so is finitely generated as a $\mathbb{Q}^+$-subsemigroup of $\mathbb{Q}^n$. It contains no line, for if we had β and $-\beta$ both in $\mathbb{Q}^+H$, we could choose a positive integer N such that $N\alpha, -N\alpha \in H$, a contradiction.

Let $\alpha_1, \ldots, \alpha_h$ be nonzero generators of H, and, hence, of C. Let $V = \mathbb{Q}^n$ and $V^* = \mathrm{Hom}_{\mathbb{Q}}(V, \mathbb{Q})$. Let $C' \subseteq V^*$ be the set of all linear functionals in V^* that are nonnegative on C. Since all elements of C are nonnegative rational linear combinations of $\alpha_1, \ldots, \alpha_h$,

$$C' = G_1 \cap \cdots \cap G_h,$$

where

$$G_j = \{L \in V^* : L(\alpha_j) \geq 0\}$$

for $1 \leq j \leq h$. We may think of α_j as an element of $(V^*)^* \cong V$. Then every G_j is a half-space in V^*, and so C' is a fintely generated $\mathbb{Q}^+$-subsemigroup in V^*. Choose $L_1, \ldots, L_s \in V^*$ that generate C' over $\mathbb{Q}^+$. Each $L_i(\alpha_j)$ is nonnegative rational number. We may therefore replace L_i by a multiple by a suitable positive integer, and so assume that for all i, j, the value of $L_i(\alpha_j)$ is in $\mathbb{N}$. Since every element of H is a linear combination of the α_j with coefficients in $\mathbb{N}$, it follows that all values of every L_i on H are in $\mathbb{N}$. We therefore have a map

$$\Phi = (L_1, \ldots, L_s) : H \to \mathbb{N}^s$$

where

$$\alpha \mapsto \big(L_1(\alpha), \ldots, L_s(\alpha)\big).$$

To complete the proof, we shall show that this map is one-to-one and that its image in $\mathbb{N}^s$ is a full subsemigroup of $\mathbb{N}^s$. First, suppose that α, $\beta \in H$ are distinct. Then $\alpha - \beta$ is nonzero, and so either $\alpha - \beta \notin H$ or $\beta - \alpha \notin H$. Suppose, say, that $\alpha - \beta \notin H$. The $\alpha - \beta \notin C$ as well: otherwise, $k(\alpha - \beta) \in H$ for some integer $k > 0$, and, since H is normal, we then have $\alpha - \beta \in H$, a contradiction. Hence, there is a linear functional nonnegative on C and negaqtive on $\alpha - \beta$. This linear functional is in C' and so is a nonnegative rational linear combination of the L_i. It follows that some L_i is negative on $\alpha - \beta$. But then $L_i(\alpha) \neq L_i(\beta)$. Thus, Φ is injective.

Finally, we need to show that the image of H under Φ is a full subsemigroup of $\mathbb{N}^s$. Suppose that $\Phi(\alpha) - \Phi(\alpha') \in \mathbb{N}^s$. We want to show that $\alpha - \alpha' \in H$. But $\Phi(\alpha - \alpha') \in \mathbb{N}^s$, and so $L_i(\alpha - \alpha') \geq 0$ for al i. If $\alpha - \alpha' \notin C$, we know that there is a linear functional L that is nonnegative on C and negative on $\alpha - \alpha'$. But then $L \in C'$, and this is impossible because every L_i is nonnegative on $\alpha - \alpha'$. Thus, $\alpha - \alpha' \in C$. But then for some positive integer k, we have that $k(\alpha - \alpha') \in H$, and so $\alpha - \alpha' \in H$, since H is normal. $\square$

Tight closure

We have shown in a graded instance that a direct summand of a polynomial ring is Cohen-Macaulay, and we have applied that result to show that finitely generated integrally closed rings generated by monomials are also Cohen-Macaulay.

The idea of the proof can be used to establish the result in much greater generality. In fact, it is known that if R is a Noetherian regular ring contain a field and $A \subseteq R$ is a direct summand of R as A-modules, then A is Cohen-Macaulay. This is an open question if R does not contain a field (e.g., R might be a finitely generated extension of $\mathbb{Z}$).

The tool that one needs to establish this result in characteristic $p > 0$ is called *tight closure theory*. A similar theory, defined by reduction to positive characteristic, exists for Noetherian rings containing the rationals. Whether there exists a comparable theory for rings that need not contain a field is a very important open question.

We are going to develop part of the theory in positive characteristic, and explain how the theory is extended to rings that contain $\mathbb{Q}$ without giving full details. We shall also explain why having such a theory would solve many open problems in mixed characteristic.

We begin by defining tight closure for ideals in Noetherian rings of positive prime characteristic p, and discussing some of its good properties. The notion was introduced implicitly in the Theorem on colon-capturing, which is the second Theorem on p. 4 of the Lecture Notes of February 17, but the explicit definition was not made at that point.

Definition: tight closure. Let R be a Noetherian ring of prime characteristic $p > 0$, let I be an ideal of R, and let $f \in R$. We say that f is in the *tight closure* of I if there exists an element $c \in R$, not in any minimal prime of R, such that for all $e \gg 0$, $cf^{p^e} \in I^{[p^e]}$. The set of elements in the tight closure of I is called the *tight closure* of I, and is denoted I^*.

In the earlier Theorem on colon-capturing, R was a domain. Notice that when R is a domain, the condition that c not be in any minimal prime of R is simply the condition that c not be 0. We note some elementary properties of the tight closure operation. Until further notice, R is a Noetherian ring of prime characteristic $p > 0$.

(1) *I^* is an ideal of R, and $I \subseteq I^*$. If $I \subseteq J \subseteq R$ are ideals, then $I^* \subseteq J^*$.*

As we did earlier in this context, we use q to stand for p^e. If $cf^q \in I^{[q]}$ for all $q \gg 0$, then $c(rf)^q \in I^{[q]}$ for all $q \gg 0$. If also $c'g^q \in I^{[q]}$ for all $q \gg 0$, then $(cc')(f+g)^q = c'cf^q + cc'g^q \in I^{[q]}$ for all $q \gg 0$. If $f \in I$ then $1 \cdot f^q \in I^{[q]}$ for all q, which shows that $I \subseteq I^*$. The fact that $I \subseteq J \Rightarrow I^* \subseteq J^*$ is obvious from the definition. $\square$

We shall use the notation R° for the set of elements of R not in any minimal prime of R. The element c used in checking whether a given element of $u \in R$ is in I^* is allowed to depend on u. However, there is a single element $c \in R^\circ$ that can be used for all elements of I^*: that is, if $u \in I^*$, then $cu^q \in I^{[q]}$ for all $q \gg 0$. The point is that I^* is finitely generated: suppose that $u_1, \ldots, u_h$ are generators. Let $c_j \in R^\circ$ be such that $c_j u_j^q \in I^{[q]}$ for all $q \gg 0$, $1 \leq j \leq h$. Let $c = c_1 \cdots c_h$. Then since every $u \in I^*$ is an R-linear combination

of $u_1, \ldots, u_h$, we have that $cu^q \in I^{[q]}$ for all $q \gg 0$. This implies that $c(I^*)^{[q]} \subseteq I^{[q]}$ for all $q \gg 0$.

One can use this to see that $(I^*)^* = I^*$. For suppose that u is such that $c'u^q \in (I^*)^{[q]}$ for all $q \gg 0$. Then $(cc')u^q = c(c'u^q) \in c(I^*)^{[q]} \subseteq I^{[q]}$ for all $q \gg 0$, and so $u \in I^*$. We state this formally:

(2) *If I is any ideal of R, $(I^*)^* = I^*$.*

We note that if R is a domain or if I is not contained in any minimal prime of R, then $u \in I^*$ iff there exists $c \in R^\circ$ such that $cu^q \in I^{[q]}$ for all q. In the second case we can choose $c' \in I - R^\circ$. If $cu^q \in I^{[q]}$ for $q \geq q_0$, we can replace c by $c(c')^{q_0}$. In the domain case we can use this idea unless $I = (0)$. But then $I^* = (0)$, and we automatically have that $cu^q \in I^{[q]}$ for all q when $u \in I^*$, since $u = 0$.

We also note:

(3) *If $R \subseteq S$ are domains, and $I \subseteq R$ is an ideal, $I^* \subseteq (IS)^*$, where I^* is taken in R and $(IS)^*$ in S.*

This is immediate from the definition of tight closure, since nonzero elements of R map to nonzero elements of S and $I^{[q]} \subseteq (IS)^{[q]} = I^{[q]}S$. More generally, this holds when $R \to S$ is a homomorphism such that R° maps into S°. In fact, under mild conditions on the rings, for any map $R \to S$ (it need not be injective) the tight closure of every ideal $I \subseteq R$ maps into the tight closure of IS in S, but the proofs are difficult.

Note that Theorem on colon-capturing from p. 4 of the Lecture Notes of February 17 can now be re-stated as follows:

Theorem (colon-capturing). *Let A be an $\mathbb{N}$-graded domain finitely generated over a field K of prime characteristic $p > 0$. Let $F_1, \ldots, F_d$ be a homogeneous system of parameters for A. Then for $0 \leq i \leq d-1$, $(F_1, \ldots, F_i)A :_A F_{i+1} \subseteq (F_1, \ldots, F_i)^*$.* $\square$

We shall see that there is a local version of this result. Mild conditions on the local ring are needed: for the reader is familiar with the notion of "excellent" local ring, we note that being excellent suffices. It is also sufficient if the ring is a homomorphic image of a regular local ring or even of a Cohen-Macaulay local ring. Since we shall show that every complete local ring is a homomorphic image of a regular local ring, the result is valid in the complete case.

(4) *If A is a local domain of characteristic $p > 0$ that is a homomorphic image of a Cohen-Macaulay ring and $f_1, \ldots, f_d$ is a system of parameters for A, then for $1 \leq i \leq d-1$, $(f_1, \ldots, f_i)A :_A f_{i+1} \subseteq \big((f_1, \ldots, f_i)A\big)^*$.*

The proof is postponed.

We next note that the Lemma on p. 5 of the Lecture Notes of February 17 may now be stated as follows:

Lemma. *Every ideal of the polynomial ring $K[x_1, \ldots, x_n]$ over a field K of prime characteristic $p > 0$ is tightly closed.* $\square$

We shall eventually show the following:

(5) *If R is a regular Noetherian ring of characteristic $p > 0$, then every ideal of R is tightly closed.*

The key point in the proof is that the Frobenius endomorphism is flat for all regular rings of characteristic $p > 0$. We shall prove this making use of the structure theory of complete local rings.

We note that given a theory of tight closure satisfying conditions (1) — (5), one immediately gets the following:

Theorem. *Let R be a regular ring of characteristic $p > 0$ and let $A \subseteq R$ be a subring such that A is a direct summand of R as A-modules. Then A is Cohen-Macaulay.*

Sketch of proof, assuming (1) — (5). The issue is local on A. Assume that (A, m) is local. One may replace A by its completion and R by its completion at mR. Thus, we may assume that the Theorem on colon-capturing holds for A, i.e., that (4) holds. Let $f_1, \ldots, f_d$ be a system of parameters for A. Suppose $uf_{i+1} \in (f_1, \ldots, f_i)A$. Then $u \in \big((f_1, \ldots, f_i)A\big)^*$ by (4). By (3), we have that $u \in \big((f_1, \ldots, f_i)R\big)^*$. By (5), we have that $u \in (f_1, \ldots, f_i)R \cap A$. Since A is a direct summand of R, it follows that $u \in (f_1, \ldots, f_i)A$. Thus, $f_1, \ldots, f_d$ is a regular sequence in A, and A is Cohen-Macaulay. $\square$

Thus, the development of a sufficiently good tight closure theory in characteristic $p > 0$ yields a proof that direct summands of regular rings are Cohen-Macaulay.

There is also a theory of tight closure for Noetherian rings containing $\mathbb{Q}$ that has properties (1) — (5). It is defined in a convoluted way using reduction to positive characteristic p. In consequence, it is known that direct summands of regular rings are Cohen-Macaulay in equal characteristic 0. It remains an open question if the ring does not contain a field.

We shall also see that the existence of a good tight closure theory has many other applications.

Lecture of March 11

Tight closure for modules

We want to extend tight closure theory to modules. Suppose we are given $N \subseteq M$, finitely generated modules over a Noetherian ring R of prime characteristic $p > 0$. We can define v^{p^e} for $v \in R^h$ as follows: if $v = (f_1, \ldots, f_h)$, then $v^{p^e} = (f_1^{p^e}, \ldots, f_h^{p^e})$. If $G \subseteq R^h$ we define G^{p^e} as the R-span of all the elements $\{v^{p^e} : v \in G\}$. One gets the same module if one takes only the R-span of the p^e th powers of generators of G. This agrees with our definition of $I^{[p^e]}$ when $I \subseteq R$ is an ideal. If $G \subseteq R^h$, we define $G^*_{R^h}$, the *tight closure* of

G in R^h as the set of elements $v \in R^h$ such that for some $c \in R^\circ$, $cv^q \in G^{[q]}$ for all $q \gg 0$, where q is p^e.

Given $N \subseteq M$ where M is finitely generated over R, we define the *tight closure* N^*_M of N in M as follows. Map a free module $R^h \twoheadrightarrow M$, and let G be the inverse image of N in R^h, so that we also have a surjection $G \twoheadrightarrow N$. Let v be any element of R^h that maps to u. Then $u \in N^*$ precisely if $v \in G^*_{R^h}$ as defined above. This is independent of the choice of v mapping to u. It is also independent of the choice of surjection $R^h \twoheadrightarrow M$.

It is understood that the tight closure of an ideal is taken in R unless otherwise specified.

Note that:

(0) *$u \in N^*_M$ if and only if the image $\overline{u}$ of u in M/N is in $0^*_{M/N}$.*

As in the ideal case:

(1) *N^*_M is a submodule of M and $N \subseteq N^*_M$. If $N \subseteq Q \subseteq M$ then $N^*_M \subseteq Q^*_M$.*

(2) If $N \subseteq M$, then $(N^*_M)^*_M = N^*_M$.

An example of tight closure

Let K be any field of characteristic $p > 0$ with $p \neq 3$. Let

$$R = K[X, Y, Z]/(X^3 + Y^3 + Z^3) = K[x, y, z].$$

This is a normal ring with an isolated singularity. It is Cohen-Macaulay. It is also a standard graded K-algebra. (This ring is sometimes called a *cubical cone*. It is also the homogeneous coordinate ring of an elliptic curve.)

We claim that $z^2 \in (x, y)^* - (x, y)$ in R. In fact, if we kill $I = (x, y)R$, we have $R/I = K[Z]/(Z^3)$, and the image of Z^2 is not 0. Take $c = z$ (the choices $c = x$ and $c = y$ also work). We need to check that

$$z(z^{2q}) \in (x^q, y^q)$$

for all $q \gg 0$. Let ρ be the remainder when $2q + 1$ is divided by 3, so that $\rho = 0$ or $\rho = 2$. We can write $2q + 1 = 3k + \rho$. Then

$$c(z^2)^q = z^{2q+1} = z^{3k+\rho} = (z^3)^k z^\rho = (-1)^k (x^3 + y^3)^k z^\rho.$$

To conclude the proof that $z^2 \in (x, y)^*$, it suffices to show that $(x^3 + y^3)^k \in (x^q, y^q)$. But otherwise we have $i + j = k$ with $i \geq 0$ and $j \geq 0$, and this implies that $3i \leq q - 1$ and that $3j \leq q - 1$. Adding these inequalities gives $3k = 3i + 3j \leq (q - 1) + (q - 1) = 2q - 2$, so that $2q + 1 - \rho \leq 2q - 2$ which implies that $\rho \geq 3$, a contradiction. $\square$

This gives a non-trivial example where the tight closure of an ideal is larger than the ideal.

Defining tight closure for Noetherian rings containing the rational numbers

We want to discuss very briefly how one extends the theory to all Noetherian rings containing $\mathbb{Q}$. For a detailed account see, [M. Hochster and C. Huneke, *Tight closure in equal characteristic zero*, preprint] available at

`http://www.math.lsa.umich.edu/~hochster/msr.html`

— the notion discussed here corresponds to $^{*\mathrm{eq}}$. There is also an exposition in [M. Hochster, *Tight closure in equal characteristic, big Cohen-Macaulay algebras, and solid closure*, in *Commutative Algebra: Syzygies, Multiplicities and Birational Algebra*, Contemp. Math. **159**, Amer. Math. Soc., Providence, R. I., 1994, 173–196].

We first define a notion of tight closure in finitely generated $\mathbb{Q}$-algebras. In fact, any finitely generated $\mathbb{Q}$-algebra can be obtained as the tensor product over $\mathbb{Z}$ of $\mathbb{Q}$ with a finitely generated $\mathbb{Z}$-algebra. If our original $\mathbb{Q}$ algebra is $R = \mathbb{Q}[X_1, \ldots, X_n]/(F_1, \ldots, F_m)$, note that one can choose a single integer d divisible by all denominators in the polynomials $F_1, \ldots, F_m$, and then

$$R = \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}[1/d][X_1, \ldots, X_n]/(F_1, \ldots, F_m).$$

We want to keep track of the behavior of this finitely generated $\mathbb{Z}$-algebra as we localize at finitely many nonzero integers: of course, this has the same effect as localizing $\mathbb{Z}$ at a single nonzero integer. Therefore we shall think of our finitely generated $\mathbb{Q}$-algebra R as $\mathbb{Q} \otimes_D R_D$, where $D = \mathbb{Z}[1/d]$ is the localization of $\mathbb{Z}$ at a single nonzero integer. But we shall allow that integer d to change so that it has more factors: in effect, as we localize further, we exclude finitely many more prime integers from consideration. By localizing at one element of $\mathbb{Z} - \{0\} \in D$ we may assume that R_D is D-free, by the Theorem on generic freeness. If B is D-algebra, which typically will be either $\mathbb{Q}$ or $\kappa = D/pD$ for some prime integer $p > 0$ not invertible in D, we write R_B for $B \otimes_D R_B$. Thus, $R = R_{\mathbb{Q}}$. Moreover, if M_D is an R_D-module, we write M_B for $B \otimes_D M_D$.

Given a finitely generated R-module M, we may think of it as the cokernel of a finite matrix with entries in D. This matrix will have entries in R_D if we localize D sufficiently, so that we have an R_D-module M_D such that $\mathbb{Q} \otimes_D M_D \cong M$. If D is large enough, we can assume that a given element of M is in D. If N is a finitely generated submodule of M, we may assume that D is large enough to contain a given finite set of generators of N over R, and we consider the R_D-submodule N_D of M_D generated by these elements. By localizing D at one more nonzero integer, we may assume that all of the terms of

$$0 \to N_D \to M_D \to M_D/N_D \to 0$$

are D-free. It follows that

$$0 \to N_B \to M_B \to M_B/N_B \to 0$$

is exact for every D-algebra B. We then have that $N \subseteq M$ arises from the inclusion $N_D \subseteq M_D$ by applying $\mathbb{Q} \otimes_D \_$. Note that when $M = R$ and $N = I$ is an ideal of R, we localize so that R_D/I_D is D-free.

Now suppose whether we want to test whether $u \in M$ is in the tight closure of N in M in *the affine $\mathbb{Q}$-algebra* sense. We choose R_D and $N_D \subseteq M_D$ as above, and take D sufficiently large that $u \in M_D$. We then define $u \in N^*_M$ if the image of $1 \otimes u$ of u is in $N^*_\kappa \subseteq M_\kappa$, where $\kappa = D/pD = \mathbb{Z}/p\mathbb{Z}$, for all but finitely many prime integers $p > 0$ that are prime in D. This condition can be shown to be independent of the choice of D, R_D, and $N_D \subseteq M_D$. This turns out to give a very good notion of tight closure when the base ring is a finitely generated $\mathbb{Q}$-algebra.

Example. Consider $R = \mathbb{Q}[X, Y, Z]/(X^3 + Y^3 + Z^3) = \mathbb{Q}[x, y, z]$. Then in this ring we have $z^2 \in (x, y)^*$, just as we did in positive characterisitic $p \neq 3$. In fact, we can take $D = \mathbb{Z}$ and $R_D = \mathbb{Z}[X, Y, Z]/(X^3 + Y^3 + Z^3)$. We can let $I_D = (x, y)R_D$. For every $p \neq 3$, with $\kappa = \mathbb{Z}/p\mathbb{Z}$, the image of z^2 in $R_\kappa = \kappa[X, Y, Z]/(X^3 + Y^3 + Z^3)$ is in the tight closure, in the characteristic $p > 0$ sense, of $I_\kappa = (x, y)_\kappa$.

This notion can be extended to arbitrary Noetherian rings containing $\mathbb{Q}$ as follows. Let S be any such ring, let M be a finitely generated S-module and $N \subseteq M$ a submodule. Let $u \in M$. Then we define $u \in N^*_M$ if for every map $S \to C$, where C is a complete local domain, there exists and affine $\mathbb{Q}$-algebra R_0, a finitely generated R_0-module M_0, a submodule $N_0 \subseteq M_0$, an element $u_0 \in M_0$, and a map $R_0 \to C$ such that:

(1) $C \otimes_{R_0} M_0 \cong C \otimes_S M$.

(2) The image of $C \otimes_{R_0} N_0$ in $C \otimes_{R_0} M_0 \cong C \otimes_S M$ is the same as the image of $C \otimes_S N$ in $C \otimes_S M$.

(3) The image $1 \otimes u_0$ of u_0 in $C \otimes_{R_0} M_0 \cong C \otimes_S M$ is the same as the $1 \otimes u$ of u in $C \otimes M$.

(4) The element u_0 is in the tight closure of N_0 in M_0 in the affine $\mathbb{Q}$-algebra sense.

That, is roughly speaking, u is in the tight closure of $N \subseteq M$ if for every base change to a complete local domain, the new u, N, M also arise by base change from an instance of tight closure over an affine $\mathbb{Q}$-algebra.

This is a highly technical, convoluted definition, and working with it presents substantial technical difficulties. Nonetheless, with the help of some very deep results about the behavior of complete local rings, including a form of the Artin Approximation Theorem, one can show that this notion satisfies the conditions (1) — (5) discussed in the Lecture Notes for March 9 for a "good" tight closure theory. For the colon-capturing property (4) it suffices if the local ring is an *excellent* domain: we shall not define the property of being excellent here, but all rings that are localizations of finitely generated algebras over either a complete local ring (fields are included) or over $\mathbb{Z}$ are excellent.

We shall not pursue these ideas further in this course, but this should give the reader some feeling for how one extends the theory to all Noetherian rings containing $\mathbb{Q}$ in a manner that ultimately rests on reduction to characteristic $p > 0$.

Another use of tight closure: contracted expansions from module-finite extension rings

Let R be a domain. Suppose that $R \subseteq S$ is a module-finite extension. In general, $I \subseteq IS \subseteq R$, but $IS \cap R$ may be larger than I. The main case is where S is also a domain. For S has a minimal prime $\mathfrak{p}$ disjoint from the multiplicative system $R - \{0\}$, and R injects into $\overline{S} = S/\mathfrak{p}$, which is a domain module-finite over R. Moreover, if $r \in R$ is in IS, then the image of r in $S/\mathfrak{p}$ is in $I\overline{S}$.

Suppose that $f \in R$, $g \in R - \{0\}$, and f/g is integral over R but not in R, which means that $f \notin gR$. We may take $S = R[f/g]$. Then $f \in gS \cap R - gR$, so that when R is not normal even principal ideals fail to be contracted from module-finite extensions. But if R is normal and contains $\mathbb{Q}$, then every ideal is contracted from every module-finite extension S. To see this, first note that it suffices to consider the case where S is a domain, by the argument above. Let $\mathcal{K}$ and $\mathcal{L}$ be the respective fraction fields of R and S. Multiplication by an element of $\mathcal{L}$ gives a map $\mathcal{L} \to \mathcal{L}$ which is $\mathcal{K}$-linear. If we simply think of this map as an endomorphism of the finite-dimensional $\mathcal{K}$-vector space $\mathcal{L}$, we may take its trace: i.e., pick a basis for $\mathcal{L}$ over $\mathcal{K}$, and take the sum of the diagonal entries of the matrix of the multiplication map with respect to this basis. This is independent of the choice of basis.

This trace map $\mathrm{Tr}_{\mathcal{L}/\mathcal{K}} : \mathcal{L} \to \mathcal{K}$ is $\mathcal{K}$-linear (hence, R-linear) and has value h on 1, where $h = [\mathcal{L} : \mathcal{K}]$. When R is a normal Noetherian ring, it turns out that the values of this map on S are in R. (One can see this as follows. First, R is the intersection of its localizations R_P at height one primes P. For if $f, g \in R$, $g \neq 0$, and f/g is in the fraction field of R but not in R, then $f \notin gR$. The associated primes of gR have height one, because R is normal. Using the primary decomposition of gR, we see that $f \notin \mathfrak{A}$ for some ideal $\mathfrak{A}$ primary to an associated P of gR of height one, and since elements of $R - P$ are not zerodivisors on $\mathfrak{A}$, $f \notin \mathfrak{A}R_P$ and so $f \notin gR_P$, i.e., $f/g \notin R_P$. If $Tr_{\mathcal{L}/\mathcal{K}}$ has a value on S not in R, we may preserve this while localizing at a height one prime P of R. But then we may replace R, S by R_P, S_P and assume that $R = R_P$ is a Noetherian discrete valuation ring. Since S is a torsion-free module over R, it is free, and has a free basis over R, say $s_1, \ldots, s_j$, consisting of elements of S. This is also a basis for $\mathcal{L}$ over $\mathcal{K}$, and can be used to calculate the trace of s. But now the matrix for multiplication by s has entries in R: for every s_i we have

$$ss_i = \sum_{j=1}^{h} r_{ij} s_j$$

with the $r_{ij} \in R$. But then the trace is $\sum_{i=1}^{h} r_{ii}$ and is in R after all. The condition that R be Noetherian is not really needed: for example, in the general case, an integrally closed domain can be shown to be a directed union of Noetherian integrally closed domains, from which the general case can be deduced. There are several other lines of argument.)

Finally, $\dfrac{1}{h}\mathrm{Tr}_{\mathcal{L}/\mathcal{K}} : S \to R$ splits $R \hookrightarrow S$ as a map of R-modules: by R-linearity, the fact that 1 maps to itself implies that the same holds for every element of R. Since we have a splitting, it follows that every ideal of R is contracted from S.

Although ideals are contracted from module finite-extensions of normal Noetherian domains that contain $\mathbb{Q}$, this is false in positive characteristic p.

Example. Let $R = K[X, Y, Z]/(X^3+Y^3+Z^3)$ where K is a field of characteristic 2. Then $z^2 \notin (x, y)R$, as noted earlier. But if we make a module-finite domain extension S of R that contains $x^{1/2}$, $y^{1/2}$, and $z^{1/2}$, then since $z^3 = x^3 + y^3$ (we are in characteristic 2, so that minus signs are not needed) we have $z^{3/2} = x^{3/2} + y^{3/2}$ (since squaring commtutes with addition and elements have at most one square root in domains of characteristic 2, taking square roots also commutes with addition in domains of characteristic 2). But then

$$z^2 = z^{1/2}z^{3/2} = z^{1/2}(xx^{1/2} + yy^{1/2}) = x^{1/2}z^{1/2}x + y^{1/2}z^{1/2}y \in (x, y)S \cap R - R.$$

However, tight closure "captures" the contracted expansion to a module-finite extension, which gives another proof that $z^2 \in (x, y)^*$ in the Example just above.

Theorem. *Let R be a Noetherian domain, and let S be any integral extension of R. Then for every ideal I of R, $IS \cap R \subseteq I^*$.*

Proof. Suppose that $f \in R$ and

$$(*) \quad f = \sum_{i=1}^{h} f_j s_j$$

where the $f_j \in I$ and the $s_j \in S$. We may replace S by $R[s_1, \ldots, s_h] \subseteq S$, and so assume that S is module-finite over R. Second, we may kill a minimal prime of S disjoint from $R-\{0\}$ and so assume that S is a module-finite domain extension of R. Choose a maximal set of R-linearly independent elements of S, say $u_1, \ldots, u_k$, so that $Ru_1 + \cdots + Ru_k$ is R-torsion. It follows that some nonzero element $r \in R$, we have that

$$S \cong rS \subseteq Ru_1 + \cdots + Ru_k.$$

Thus, we have an embedding $S \hookrightarrow R^k$. Suppose that $1 \in S$ has as its image in R^k an element whose ith coordinate is nonzero, so that the composite map $S \hookrightarrow R^k \xrightarrow{\pi_i} R$ is nonzero on the element $1 \in S$, where π_i is the ith coordinate projection of $R^k \twoheadrightarrow R$. This gives an R-linear map $\theta : S \to R$ such that $\theta(1) = c \in R$ is nonzero. Now take qth powers of both sides of $(*)$, yielding

$$(**) \quad f^q \cdot 1 = \sum_{i=1}^{h} f_j^q s_j^q.$$

Since θ is R-linear and $f, f_1, \ldots, f_h \in R$, this yields

$$f^q\theta(1) = \sum_{i=1}^{h} f_j^q \theta(s_j^q),$$

and so $cf^q \in I^{[q]}$ for all q. This implies that $f \in I^*$. $\square$

Lecture of March 14

Open questions: tight closure, plus closure, and localization

We want to consider some open questions in tight closure theory, and some related problems about when rings split from their module-finite extension algebras. After we do this, we shall prove some specific results in the characteristic p theory. It will turn out that to proceed further, we will need the structure theory of complete local rings, which we will develop next.

One of the longest standing and most important questions about tight closure is when tight closure commutes with localization. E.g., if R is Noetherian of prime characteristic $p > 0$, I is an ideal of R, and W is a multiplicative system of R, when is $W^{-1}(I^*_R)$ the same as $(W^{-1})^*_{W^{-1}R}$? It is easy to prove that $W^{-1}(I^*_R) \subseteq (W^{-1})^*_{W^{-1}R}$. This was an open question for more than twenty years. It is known to be true in many cases, but false in general, by a result of [H. Brenner and P. Monsky, See, for example, [I. Aberbach, M. Hochster, and C. Huneke, *Localization of tight closure and and modules of finite phantom projective dimension*, J. Reine Angew. Math. (Crelle's Journal) **434** (1993), 67–114], and [M. Hochster and C. Huneke, *Test exponents and localization of tight closure*, Michigan Math. J. **48** (2000), 305–329] for a discussion of the problem.

We saw in the last Theorem of the Lecture Notes of March 11 that tight closure "captures" contracted extension from module-finite and even integral extensions. We shall add this as (6) to our list of desirable properties for a tight closure theory, which becomes the following:

(0) *$u \in N^*_M$ if and only if the image $\overline{u}$ of u in M/N is in $0^*_{M/N}$.*

(1) *N^*_M is a submodule of M and $N \subseteq N^*_M$. If $N \subseteq Q \subseteq M$ then $N^*_M \subseteq Q^*_M$.*

(2) If $N \subseteq M$, then $(N^*_M)^*_M = N^*_M$.

(3) *If $R \subseteq S$ are domains, and $I \subseteq R$ is an ideal, $I^* \subseteq (IS)^*$, where I^* is taken in R and $(IS)^*$ in S.*

(4) *If A is a local domain then, under mild conditions on A (the class of rings allowed should include local rings of a finitely generated algebra over a complete local ring or over $\mathbb{Z}$), and $f_1, \ldots, f_d$ is a system of parameters for A, then for $1 \leq i \leq d-1$, $(f_1, \ldots, f_i)A :_A f_{i+1} \subseteq \big((f_1, \ldots, f_i)A\big)^*$.*

(5) If R is regular, then $I^* = I$ for every ideal I of R.

(6) For every module-finite extension ring R of S and every ideal I of R, $IS \cap R \subseteq I^*$.

These are all properties of tight closure in prime characteristic $p > 0$, and also of the theory of tight closure for Noetherian rings containing $\mathbb{Q}$ that we described in the Lecture

of March 11. In characteristic $p > 0$, (4) holds for homomorphic images of Cohen-Macaulay rings, and for excellent local rings. If $R \supseteq \mathbb{Q}$, (4) holds if R is excellent. We will prove that (4) holds in prime characteristic for homomorphic images of Cohen-Macaulay rings quite soon. We have proved (5) in prime characteristic $p > 0$ for polynomial rings over a field, but not yet for all regular rings. To give the proof for all regular rings we need to prove that the Frobenius endomorphism is flat for all such rings, and we shall eventually use the structure theory of complete local rings to do this.

An extremely important open question is whether there exists a closure theory satisfying (1) — (6) for Noetherian rings that need not contain a field.

The final Theorem of the Lecture of March 11 makes it natural to consider the following variant notion of closure. Let R be any integral domain. Let R^+ denote integral closure of R in an algebraic closure $\overline{\mathcal{K}}$ of its fraction field $\mathcal{K}$. We refer to this ring as the *absolute integral closure* of R. R^+ is unique up to non-unique isomorphism, just as the algebraic closure of a field is. Any module-finite (or integral) extension domain S of R has fraction field algebraic over $\mathcal{K}$, and so S embeds in $\overline{\mathcal{K}}$. It follows that S embeds in R^+, since the elements of S are integral over R. Thus, R^+ contains an R-subalgebra isomorphic to any other integral extension domain of R: it is a maximal extension domain with respect to the property of being integral over R. R^+ is the directed union of its finitely generated subrings, which are module-finite over R. R^+ is also charactized as follows: it is a domain that is an integral extension of R, and every monic polynomial with coefficients in R^+ factors into monic linear polynomials over R^+.

Given an ideal $I \subseteq R$, the following two conditions on $f \in R$ are equivalent:

(1) $f \in IR^+ \cap R$.

(2) For some module-finite extension S of R, $f \in IS \cap R$.

The set of such elements, which is $IR^+ \cap R$, is denoted I^+, and is called the *plus closure* of I. (The definition can be extended to modules $N \subseteq M$ by defining N^+_M to be the kernel of the map $M \to R^+ \otimes_R (M/N)$.)

By the last Theorem of the Lecture Notes of March 11, which is property (6) above in characteristic $p > 0$, we have that

$$I \subseteq I^+ \subseteq I^*$$

in prime characteristic $p > 0$. Whether $I^+ = I^*$ in general under mild conditions for Noetherian rings of prime characterisitc $p > 0$ is another very important open question. It is not known to be true even in finitely generated algebras of Krull dimension 2 over a field.

However, there are some substantial positive results. It is known that under the mild conditions on the local domain R (e.g., when R is excellent), if I is generated by part of a system of parameters for R, then $I^+ = I^*$. See [K. E. Smith, *Tight closure of parameter ideals*, Inventiones Math. **115** (1994) 41–60]. Moreover, H. Brenner [H. Brenner, *Tight closure and plus closure in dimension two*, Amer. J. Math. **128** (2006) 531–539] proved that

if R is the homogeneous coordinate ring of a smooth projective curve over the algebraic closure of $\mathbb{Z}/p\mathbb{Z}$ for some prime integer $p > 0$, then $I^* = I^+$ for homogeneous ideals primary to the homogeneous maximal ideal. In [G. Dietz, *Closure operations in positive characteristic and big Cohen-Macaulay algebras*, Thesis, Univ. of Michigan, 2005] the condition that the ideal be homogeneous is removed: in fact, there is a corresponding result for modules $N \subseteq M$ when M/N has finite length. Brenner's methods involve the theory of semi-stable vector bundles over a smooth curve (in fact, one needs the notion of a *strongly* semi-stable vector bundle, where "strongly" means that the bundle remains semi-stable after pullback by the Frobenius map).

One reason for the great interest in whether plus closure commutes with tight closure is that it is known that plus closure commutes with localization. Hence, if $I^* = I^+$ in general (under mild conditions on the ring) one gets the result that tight closure commutes with localization.

The notion of plus closure is of almost no help in understanding tight closure when the ring contains the rationals. The reason for this is the result established on pp. 4–5 of the Lecture Notes of March 11, which we restate formally here.

Theorem. *Let R be a normal Noetherian domain with fraction field $\mathcal{K}$ and let S be a module-finite extension domain with fraction field $\mathcal{L}$. Let $h = [\mathcal{L} : \mathcal{K}]$. If $\mathbb{Q} \subseteq R$, or, more generally, if h has an inverse in R, then $\frac{1}{h}\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}$ gives an R-module retraction $S \to R$.* $\square$

It follows that if $\mathbb{Q} \subseteq R$ and R is a normal domain, then $I^+ = I$ for every ideal I of R. Many normal rings (in some sense most normal rings) that are essentially of finite type over $\mathbb{Q}$ are not Cohen-Macaulay, and so contain parameter ideals that are not tightly closed. This shows that plus closure is not a greatly useful notion in Noetherian domains that contain $\mathbb{Q}$.

Weakly F-regular rings and F-regular rings

We define a Noetherian ring R of prime characteristic $p > 0$ to be *weakly F-regular* if every ideal is equal to its tight closure, i.e., every ideal is tightly closed. We define R to be *F-regular* if all of its localizations are weakly F-regular. It is not known whether weakly F-regular implies F-regular, even for domains finitely generated over a field. This would follow if tight closure were known to commute with localization.

We have already proved that polynomial rings over a field of positive characteristic are weakly F-regular, and we shall prove that every regular ring of positive characteristic is F-regular. This is one reason for the terminology. The "F" suggest the involvement of the Frobenius endomorphism.

We shall soon show that a weakly F-regular ring is normal, and, if it is a homomorphic image of a Cohen-Macaulay ring, is itself Cohen-Macaulay.

Theorem. *A direct summand A of a weakly F-regular domain is weakly F-regular, and a direct summand of an F-regular domain is F-regular.*

Proof. Assume that R is weakly F-regular. If $f \in I_A^*$, then $f \in (IR)^* \cap A = IR \cap A = I$. Since the direct summand condition is preserved by localization on A, it follows that a direct summand of an F-regular domain is F-regular. $\square$

Examples of F-regular rings. Fix a field K of characteristic $p > 0$. Normal rings finitely generated over K by monomials are direct summand of regular rings, and so are F-regular. If X is an $r \times s$ matrix of indeterminates over K with $1 \leq t \leq r \leq s$, then it is known that $K[X]/I_t(X)$ is F-regular, and that the ring generated by the $r \times r$ minors of X over K is F-regular (this is the homogeneous coordinate ring of the Grassmann variety). See [M. Hochster and C. Huneke, *Tight closure of parameter ideals and splitting in module-finite extensions*, J. of Algeraic Geometry **3** (1994) 599–670], Theorem (7.14). We have already observed that these rings are direct summands of polynomial rings when K has characteristic 0, but this is not true in any obvious way when the characteristic is positive.

Splitting from module-finite extension rings

It is natural to attempt to characterize the Noetherian domains R such that R is a direct summand, as an R-module, of every module-finite extension ring S. We define a Noetherian ring R with this property to be a *splinter*. We then have the following result, which was actually proved in the preceding lecture, although it was not made explicit there.

Theorem. *Let R be a Noetherian domain.*

(a) *If R is a splinter, then every ideal of R is contracted from every integral extension.*

(b) *If R is a splinter, then R is normal.*

(c) *R is a splinter if and only if it is a direct summand of every module-finite domain extension.*

(d) *If $\mathbb{Q} \subseteq R$, then R is a splinter if and only if R is normal.*

Proof. For part (a), suppose $f, f_1, \ldots, f_h \in R$ and $f = \sum_{i=1}^h f_i s_i$ with the s_i in S. Then we have the same situation when S is replaced by $R[s_1, \ldots, s_h]$. Hence, it suffices to show that every ideal of R is contracted from every module-finite extension S. But then we have an R-linear retraction $\phi : S \to R$, and the result is part (a) of the Lemma at the top of p. 2 of the Lecture of February 17.

Part (b) has already been established in the fourth paragraph on p. 4 of the Lecture of March 11.

For part (c), we have already observed that S has a minimal prime $\mathfrak{p}$ disjoint from $R - \{0\}$, and it suffices to split the injection $R \hookrightarrow S/\mathfrak{p}$.

Finally, for part (d), the existence of the required splitting when S is a domain is proved at the bottom of p. 4 and top of p. 5 of the Lecture Notes of March 11, using field trace, and restated on p. 3 here. $\square$

The example on p. 5 of the Lecture Notes of March 11 shows that in positive characteristic p, a normal domain need not be a splinter. The property of being a splinter in characteristic p is closely related to the property of being weakly F-regular.

We first note the following fact: we shall not give the proof in these lectures, but refer the reader to [M. Hochster, *Contracted ideals from integral extensions of regular rings*, Nagoya Math. J. **51** (1973) 25–43] and [M. Hochster, *Cyclic purity versus purity in excellent Noetherian rings*, Trans. Amer. Math. Soc. **231** (1977) 463–488].

Theorem. *Let R be a normal Noetherian domain. Then R is a direct summand of a module-finite extension of S if and only if every ideal of R is contracted from S.*

Of course, we know the "only if" part.

Corollary. *Let R be a normal Noetherian domain of positive characteristic p. Then R is a splinter if and only if for every ideal $I \subseteq R$, $I = I^+$.*

Corollary. *If R is a normal Noetherian domain and R is weakly F-regular, then R is a splinter.*

Proof. This is immediate from the preceding result, since $I^+ \subseteq I^*$. $\square$

We shall see quite soon that if R is weakly F-regular it is automatic that R is normal. If plus closure is the same as tight closure, then it would follow that R is weakly F-regular if and only if R is a splinter. This is an open question.

We have already observed that in characteristic $p > 0$, regular rings are weakly F-regular, although we have not prove this. Assuming this for the moment we have:

Corollary. *A regular ring that contains a field is a direct summand of every module-finite extension ring.*

This was conjectured by the author in 1969, and has been open question for regular rings that do not contain a field, such as polynomial rings over the integers, for 37 years. The case of dimension 3 was recently settled affirmatively in [R. C. Heitmann, *The direct summand conjecture in dimension three*, Annals of Math. (2) **156** (2002) 695–712]. The case of dimension 4 remains open for regular rings that do not contain a field.

It is also a major open question whether there exists a tight closure theory satisfying conditions (0) — (6) of p. 1 for Noetherian rings that need not contain a field. The existence of such a theory would imply that direct summands of regular rings are Cohen-Macaulay in general, and that regular rings are direct summands of all of their module-finite extensions in general. Such a theory would also settle many other open questions.

Lecture of March 16

We next want to study weakly F-rings, i.e., Noetherian rings of prime characteristic $p > 0$ such that every ideal is tightly closed. Until further notice, all given rings R are assumed to be Noetherian, of prime characteristic $p > 0$.

Proposition. *The tight closure of the* (0) *ideal in* R *is the ideal of all nilpotent elements. Hence, if* $(0) = (0)^*$*, the* R *is reduced. In particular, every weakly F-regular ring is reduced.*

Proof. If u is nilpotent then $1 \cdot u^q = 0$ for all $q \gg 0$. Conversely, if $c \in R^\circ$ and $cu^q = 0$ for all $q \gg 0$, then for every minimal prime $\mathfrak{p}$ we have that $cu^q \in \mathfrak{p}$ for some q. Since $c \notin \mathfrak{p}$, we have that $u^q \in \mathfrak{p}$ and so $u \in \mathfrak{p}$. But the intersection of the minimal primes is the set of nilpotent elements of R, and so u is nilpotent. The remaining statements are now obvious. $\square$

Proposition. *Suppose that* $R = S \times T$ *is a product ring, with* $S, T \neq 0$*. Then for every ideal* $I \times J$ *of* $S \times T$*, where* $I \subseteq S$ *and* $J \subseteq T$ *are ideals,* $(I \times J)^*_R = I^*_S \times J^*_T$*.*

Proof. The first point is that $(S \times T)^\circ = (S^\circ) \times (T^\circ)$. Hence if $cs^q \in I^{[q]}$ for all $q \gg 0$ and $dt^q \in J^{[q]}$ for all $q \gg 0$, we have that

$$(c, d)(s, t)^q \in I^{[q]} \times J^{[q]} = (I \times J)^{[q]}$$

for all $q \gg 0$. The converse is also immediate. $\square$

Corollary. *A finite product* $R_1 \times \cdots \times R_h$ *is weakly F-regular if and only if every factor is weakly F-regular.* $\square$

Theorem. *If every principal ideal of* R *is tightly closed, then* R *is a product of normal domains.*

Proof. The fact that $(0) = (0)^*$ implies that R is reduced. We first show that R is a product of domains. If there are two or more minimal primes, the minimal primes can be partitioned into two nonempty sets. Call the intersection of one set I and the intersection of the other set J. Then $I \cap J = 0$, and $I + J$ is not contained in any minimal prime $\mathfrak{p}$, for otherwise, $\mathfrak{p}$ would have to contain both a minimal prime of I and a minimal prime of J, and would be equal to both of these. Hence we can choose $f \in I$ and $g \in J$ such that $f + g$ is not in any minimal prime of R, and so is a nonzerodivisor. Note that $fg \in I \cap J$, and so $fg = 0$. Now

$$(f + g)f^q = f^{q+1} = f(f + g)^q$$

for all q, so that $f \in (f + g)^* = (f + g)R$. Thus, we can choose $r \in R$ such that $f = r(f + g) = rf + rg$, and the $f - rf = rg$. Since $f \in I$ and $g \in J$, both sides must vanish, and so $f = rf$ and $rg = 0$. Now $r(f + g) = rf = f$, and

$$r^2(f + g) = r(rf + rg) = r(f + 0) = rf = f,$$

so that

$$(f+g)(r^2-r)=0.$$

Since $f+g$ is not a zerodivisor, we have that $r^2-r=0$. Since $rf=f$ is not 0 (or $f+g$ would be in the minimal primes containing g) $r\neq 0$. Since $rg=0$, $r\neq 1$. Therefore, R contains a non-trivial idempotent, and is a product of two rings. Both have the property that principal ideals are tightly closed, because a principal ideal of $S\times T$ is the product of a principal ideal of S and a principal ideal of T, and we may apply the Proposition above.

We may apply this argument repeatedly and so write R as a finite product of rings with the property that every principal ideal is tightly closed, and such that none of the factors is a product. Each of the factors must have just one minimal prime, and so is a domain. It remains to see that if principal ideals are tightly closed in a domain R, then R is normal. Suppose that f, $g\in R$, $g\neq 0$, and f/g is integral over R. Let $S=R[f/g]$, which is module-finite over R. Then $f=g(f/g)\in gS$, and so $f\in (gR)^*$. But $(gR)^*=gR$, and so $f\in gR$, i.e., $f/g\in R$, as required. □

We next want to show that, under mild conditions on R, if R is weakly F-regular then R is Cohen-Macaulay. Before giving the proof, we make some comments about Cohen-Macaulay rings in general.

Cohen-Macaulay rings

In this section, we assume that given rings are Noetherian, but make no assumption about the characteristic. In particular, given rings need not contain a field.

We have defined the notion of a Cohen-Macaulay ring in the case of a finitely generated $\mathbb{N}$-graded K-algebra R with $R_0=K$. We have also defined the notion of a Cohen-Macaulay local ring, and define a Noetherian ring to be Cohen-Macaulay if all of its local rings are Cohen-Macaulay. We first note:

Lemma. *Let (R, m, K) be a local ring and let I be an ideal of height h in R. Then there is a sequence of elements $x_1, \ldots, x_h$ in I that is part of a system of parameters for R.*

Proof. If $h=0$ we may take the empty sequence. If $h\geq 1$, then I is not contained in the union of the minimal primes of R, or else we would have that I is contained in one of them and has height 0. Choose $x_1\in I$ not in any minimal prime of R. Then x_1 is part of a system of parameters. We use induction. Suppose that $x_1, \ldots, x_i\in I$ have been chosen so that they are part of a system of paramters with $i<h$. The minimal primes of $(x_1, \ldots, x_i)R$ all have height $\leq i<h$, and so I is not contained in any of them and also not contained in their union. Choose $x_{i+1}\in I$ not in any minimal prime of $(x_1, \ldots, x_i)R$. Then $x_1, \ldots, x_{i+1}$ is also part of a system of parameters. □

Corollary. *If (R, m) is Cohen-Macaulay and P is a prime ideal of R, the R_P is Cohen-Macaulay.*

Proof. Suppose that $h = \text{height}\,(P) = \dim\,(R_P)$. Choose $x_1, \ldots, x_h \in P$ part of a system of parameters for R. Then $x_1, \ldots, x_h$ is a regular sequence in R, and, hence, also in R_P, by flatness. $\square$

Theorem. *Let R be a Noetherian ring. The following conditions are equivalent:*

(1) *R is Cohen-Macaulay, i.e., R_P is Cohen-Macaulay for every prime ideal P.*

(2) *R_m is Cohen-Macaulay for every maximal ideal m.*

(3) *For every proper ideal I of R,* $\text{depth}_I R = \text{height}\,(I)$.

Proof. (1) $\Rightarrow$ (2) is obvious, while (2) $\Rightarrow$ (1) because each R_P is a localization of R_m for some maximal ideal containing P. Now assume (2) and suppose that I has height h. Choose a maximal regular sequence $x_1, \ldots, x_d$ in I on R. Then $R/(x_1, \ldots, x_d)R$ has depth 0 on $I/(x_1, \ldots, x_d)R$, and this remains true after we localize at an associated prime P of $R/(x_1, \ldots, x_d)R$ that contains $I(x_1, \ldots, x_d)R$. Hence, $x_1, \ldots, x_d$ is also a maximal regular sequence in R_P, which shows that $d = h$, since R_P is Cohen-Macaulay of dimension h. Thus, (2) $\Rightarrow$ (3).

Finally, assume (3). Let P be any prime ideal of R of height h. Then P contains a regular sequence of length $\text{height}\,(P) = \dim\,(R_P)$, and this sequence remains a regular sequence when we localize at P. Hence, (3) $\Rightarrow$ (1). $\square$

Theorem. *If R is Cohen-Macaulay, so is the polynomial ring in n variables over R.*

Proof. By induction, we may assume that $n = 1$. Let $\mathcal{M}$ be a maximal ideal of $R[X]$ lying over m in R. We may replace R by R_m and so we may assume that (R, m, K) is local. Then $\mathcal{M}$, which is a maximal ideal of $R[x]$ lying over m, corresponds to a maximal ideal ideal of $K[x]$: each of these is generated by a monic irreducible polynomial f, which lifts to a monic polynomial F in $R[x]$. Thus, we may assume that $\mathcal{M} = mR[x] + FR[X]$. Let $x_1, \ldots, x_d$ be a system of parameters in R, which is also a regular sequence. We may kill the ideal generated by these elements, which also form a regular sequence in $R[X]_{\mathcal{M}}$. We are now in the case where R is an Artin local ring. It is clear that the height of $\mathcal{M}$ is one. Because F is monic, it is not a zerodivisor: a monic polynomial over any ring is not a zerodivisor. This shows that the depth of $\mathcal{M}$ is one, as needed. $\square$

Theorem. *If R is a finitely generated graded K-algebra with $[R]_0 = K$, then R is Cohen-Macaulay in the graded sense if and only if R is Cohen-Macaulay.*

Proof. Let m be the homogeneous maximal ideal. If R_m is Cohen-Macaulay, choose a maximal regular sequence in m consisting of homogeneous elements (necessarily of positive degree), say $F_1, \ldots, F_h$. When we kill these elements, we know that in $R/(F_1, \ldots, F_h)R$, the homogeneous elements of the ideal $m/(F_1, \ldots, F_h)R$ are all contained in the union of the associated primes of $R/(F_1, \ldots, F_h)R$. By the Proposition on homogeneous prime avoidance from the bottom of p. 4 of the Lecture of January 27, $m/(f_1, \ldots, f_h)R$ itself is contained in one of these associated primes, and so $m/(f_1, \ldots, f_h R)$ is an associated

prime. This is preserved when we localize at $m/(f_1, \ldots, f_h)R$, and so R_m has depth 0 once we kill $(f_1, \ldots, f_h)R_m$. Therefore, $f_1, \ldots, f_h$ is a maximal regular sequence in R_m, and this implies that $h = \dim(R_m) = \dim(R)$. Thus, R is Cohen-Macaulay in the graded sense.

Now suppose that R is Cohen-Macaulay in the graded sense. Then R is a module-finite extension of a polynomial ring $A = K[X_1, \ldots, X_n]$, and the polynomial ring is Cohen-Macaulay. Any maximal ideal $\mathfrak{q}$ of R lies over a maximal ideal $\mathfrak{n}$ of A. These have the same height, since we have both the going up and going down theorems in this situation: A is normal, and R is A-free and, hence, torsion-free over A. Since R is A-free, a regular sequence in $A_{\mathfrak{n}}$ is regular on $R_{\mathfrak{n}}$, which is free and, hence, faithfully flat over A, and will remain regular on $R_{\mathfrak{q}}$, which is a localization of $R_{\mathfrak{n}}$. $\square$

We next observe:

Theorem. *Let (R, m, K) be a local ring and $M \neq 0$ a finitely generated R-module of depth s on m. Then every nonzero submodule N of M has dimension at least s.*

Proof. We use induction on s. If $s = 0$ there is nothing to prove. Assume $s > 0$ and that the result holds for smaller s. If M has a submodule $N \neq 0$ of dimension $\leq s-1$, we may choose N maximal with respect to this property. If N' is any nonzero submodule of M of dimension $< s$, then $N' \subseteq N$. To see this, note that $N \oplus N'$ has dimension $< s$, and maps onto $N + N' \subseteq M$, which therefore also has dimension $< s$. By the maximality of N, we must have $N + N' = N$. Since $\operatorname{depth}_m M \geq 1$, we can choose $x \in m$ not a zerodivisor on M, and, hence, also not a zerodivisor on N. We claim that x is not a zerodivisor on $\overline{M} = M/N$, for if $u \in M - N$ and $xu \in N$, then $Rxu \subseteq N$ has dimension $< s$. But this module is isomorphic with $Ru \subseteq M$, since x is not a zerodivisor, and so $\dim(Ru) < s$. But then $Ru \subseteq N$. Consequently, multiplication by x induces an isomorphism of the exact sequence $0 \to N \to M \to \overline{M} \to 0$ with the sequence $0 \to xN \to xM \to x\overline{M} \to 0$, and so this sequence is also exact. But we have a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N & \longrightarrow & M & \longrightarrow & \overline{M} & \longrightarrow & 0 \\ & & \uparrow & & \uparrow & & \uparrow & & \\ 0 & \longrightarrow & xN & \longrightarrow & xM & \longrightarrow & x\overline{M} & \longrightarrow & 0 \end{array}$$

where the vertical arrows are inclusions. By the nine lemma, or by an elementary diagram chase, the sequence of cokernels $0 \to N/xN \to M/xM \to \overline{M}/x\overline{M} \to 0$ is exact. Since x is a nonzerodivisor on N and M, $\dim(N/xN) = \dim(N) - 1 < s-1$, while $\operatorname{depth}_m M/xM = s - 1$. This contradicts the induction hypothesis. $\square$

Corollary. *If (R, m, K) is a Cohen-Macaulay local ring, then for every minimal prime $\mathfrak{p}$ of R, $\dim(R/\mathfrak{p}) = \dim(R)$.*

Proof. If $\mathfrak{p}$ is minimal, then $\mathfrak{p} \in \operatorname{Ass}(R)$ and so $R/\mathfrak{p} \hookrightarrow R$. By the preceding Theorem, $\dim(R/\mathfrak{p}) \geq \operatorname{depth}_m R = \dim(R)$, while the other inclusion is obvious. $\square$

Thus, a Cohen-Macaulay local ring cannot exhibit the kind of behavior one observes in $R = K[[x, y, z]]/\big((x, y) \cap (z)\big)$: this ring has two minimal primes. One of them, $\mathfrak{p}_1$, generated by the images of x and y, is such that $R/\mathfrak{p}_1$ has dimension 1. The other, $\mathfrak{p}_2$, generated by the image of z, is such that $R/\mathfrak{p}_2$ has dimension 2.

A Noetherian ring is called *catenary* if for any two prime ideals $P \subseteq Q$, any two saturated chains of primes joining P to Q have the same length. If R is catenary, then so is R/I for every ideal I, since primes containing I are in bijective correspondence with primes of R containing I, and saturated chains of primes in R/I joining P/I to Q/I, where $I \subseteq P \subseteq Q$ and P, Q are primes of R, correspond to saturated chains of primes of R joinig P to Q. Similarly, any localization of a catenary ring is catenary. M. Nagata gave the first examples of Noetherian rings that are not catenary: there is a local domain (R, m, K) of dimension 3, for example, containing saturated chains $0 \subset Q \subset m$ and $0 \subset P_1 \subset P_2 \subset m$, where all inclusions are strict. See [M. Nagata, *Local rings*, Interscience, New York, 1962], Appendx A1, pp. 204–205. Although Q has height one and $\dim(R) = 3$, the dimension of R/Q is 1. Nagata also showed that even when a Noetherian ring is catenary, the polynomial ring in one variable over it need not be.

A Noetherian ring R is called *universally catenary* if every finitely generated R-algebra is catenary. Cohen-Macaulay rings are universally catenary, as we show in the two results below.

Theorem. *A Cohen-Macaulay ring R is catenary, and for any two prime ideals $P \subseteq Q$ in R, every saturated chain of prime ideals joining P to Q has length* $\operatorname{height}(Q) - \operatorname{height}(P)$. *Hence, every finitely generated algebra over a Cohen-Macaulay ring is catenary.*

Proof. The issues are unaffected by localizing at Q. Thus, we may assume that R is local and that Q is the maximal ideal. There is part of a system of parameters of length $h = \operatorname{height}(P)$ contained in P, call it $x_1, \ldots, x_h$, by the Lemma at the beginning of this section. This sequence is a regular sequence on R and in so on R_P, which implies that its image in R_P is system of parameters. We now replace R by $R/(x_1, \ldots, x_h)$. Both the dimension and depth of R have decreased by h, so that R is still Cohen-Macaulay. Q and P are replaced by their images, which have heights $\dim(R) - h$ and 0, and $\dim(R) - h = \dim(R/(x_1, \ldots, x_h)$. We have therefore reduced to the case where R is local and P is a minimal prime. We know that $\dim(R) = \dim(R/P)$, and so at least one saturated chain from P to Q has length $\operatorname{height}(Q) - \operatorname{height}(P) = \operatorname{height}(Q) - 0 = \dim(R)$. To complete the proof, it will suffice to show that all saturated chains from P to Q have the same length, and we may use induction on $\dim(R)$. Consider two such chains, and let their smallest elements other than P be P_1 and P_1'. Choose an element x in P_1 not in any minimal prime, and an element y of P_1' not in any minimal prime. Then xy is a nonzerodivisor in R, and P_1, P_1' are both minimal primes of xy. The ring $R/(xy)$ is Cohen-Macaulay of dimension $\dim(R) - 1$. The result now follows from the induction hypothesis applied to $R/(xy)$: the images of the two saturated chains (omitting P from each) give saturated chains joining $P_1/(xy)$ (respectively, $P_1'/(xy)$) to $Q/(xy)$ in $R/(xy)$. These have the same length, and, hence, so did the original two chains. $\square$

Corollary. *Cohen-Macaulay rings are universally catenary, i.e., a finitely generated algebra over a Cohen-Macaulay ring is catenary.*

Proof. Such an algebra is a homomorphic image of a polynomial ring in finitely many variables over a Cohen-Macaulay ring, which is again Cohen-Macaulay, and homomorphic images of catenary rings are catenary. □

Lecture of March 18

Colon-capturing in homomorphic images of Cohen-Macaulay rings

We will need the following two preliminary results:

Lemma (prime avoidance for cosets). *Let S be any commutative ring, $x \in S$, $I \subseteq S$ an ideal and $P_1, \ldots, P_k$ prime ideals of S. Suppose that the coset $x + I$ is contained in $\bigcup_{i=1}^{k} P_i$. Then there exists j such that $Sx + I \subseteq P_j$.*

Proof. If $k = 1$ the result is clear. Choose $k \geq 2$ minimum giving a counterexample. Then no two P_i are comparable, and $x + I$ is not contained in the union of any $k - 1$ of the P_i. Now $x = x + 0 \in x + I$, and so x is in at least one of the P_j: say $x \in P_k$. If $I \subseteq P_k$, then $Sx + I \subseteq P_k$ and we are done. If not, choose $i_0 \in I - P_k$. We can also choose $i \in I$ such that $x + i \notin \bigcup_{j=1}^{k-1} P_i$. Choose $u_j \in P_j - P_k$ for $j < k$, and let u be the product of the u_j. Then $ui_0 \in I - P_k$, but is in P_j for $j < k$. It follows that $x + (i + ui_0) \in x + I$, but is not in any P_j, $1 \leq j \leq k$, a contradiction. □

Lemma. *Let S be a Cohen-Macaulay local ring, let P be a prime ideal of S of height h, and let $x_1, \ldots, x_{i+1}$ be part of a system of parameters of $R = S/P$. Let $y_1, \ldots, y_h \in P$ be part of a system of parameters for S (we have such a sequence by the first Lemma of the preceding section on Cohen-Macaulay rings). Then there exist elements $\widetilde{x}_1, \ldots, \widetilde{x}_{i+1}$ of S such that $\widetilde{x}_j$ maps to x_j modulo P, $1 \leq j \leq i + 1$, and $y_1, \ldots, y_h, \widetilde{x}_1, \ldots, \widetilde{x}_{i+1}$ is part of a system of parameters for S.*

Proof. We construct the $\widetilde{x}_j$ recursively. Suppose that the $\widetilde{x}_j$ for $j < k+1 \leq i+1$ have been chosen so that $y_1, \ldots, y_h, \widetilde{x}_1, \ldots, \widetilde{x}_k$ is part of a system of parameters for S. Here, k is allowed to be 0 (i.e., we may be choosing $\widetilde{x}_1$). We want to choose an element of $x_{k+1} + P$ that is not in any minimal prime of $y_1, \ldots, y_h, \widetilde{x}_1, \ldots, \widetilde{x}_k$, and these all have height at most $h + k$. By the Lemma on prime avoidance for cosets, if $\widetilde{x}_{k+1} + P$ is contained in the union, then $Sx_{k+1} + P$ is contained in one of them, say Q. Working modulo P we have that Q/P is a minimal prime $x_1, \ldots, x_{k+1}$ of height at most $h + k - h = k$. This is a contradiction, since $x_1, \ldots, x_{k+1}$ is part of a system of parameters in S/P, and so any minimal prime must have height at least $k + 1$. □

Theorem (colon-capturing). *Let (R, m, K) be a local domain of prime characteristic $p > 0$, and suppose that R is a homomorphic image of a Cohen-Macaulay ring of characteristic p. Let $x_1, \ldots, x_{i+1}$ by part of a system of parameters in R. Then*

$$(x_1, \ldots, x_i) :_R x_{i+1} \subseteq (x_1, \ldots, x_i)^*.$$

Proof. Suppose that $R = S/P$, where S is Cohen-Macaulay of characteristic p, and let Q be the inverse image of m in S. Then R is also a homomorphic image of S_Q, since $S_Q/PS_Q \cong (S/P)_Q = R_Q = R_m = R$. Hence, we may assume that S is local. Choose $y_1, \ldots, y_h$ and $\widetilde{x}_1, \ldots, \widetilde{x}_{i+1}$ as in the preceding Lemma. Since P is a minimal prime of $(y_1, \ldots, y_h)$ in S, we can choose $\widetilde{c} \in S - P$ and an integer $N > 0$ such that $\widetilde{c}P^N \in (y_1, \ldots, y_h)S$. Let $c \neq 0$ be the image of $\widetilde{c}$ in R. Suppose that $fx_{i+1} = f_1x_1 + \cdots + f_ix_i$ in R. Then we can choose elements $\widetilde{f}$ and $\widetilde{f}_1, \ldots, \widetilde{f}_i$ in S that lift f and $f_1, \ldots, f_i$ respectively to S. This yields an equation

$$\widetilde{f}\widetilde{x}_{i+1} = \widetilde{f}_1\widetilde{x}_1 + \cdots + \widetilde{f}_i\widetilde{x}_i + \Delta$$

in S, where $\Delta \in P$. Then for all $p^e = q \geq N$ we have

$$\widetilde{f}^q\widetilde{x}_{i+1}^q = \widetilde{f}_1{}^q\widetilde{x}_1^q + \cdots + \widetilde{f}_i^q\widetilde{x}_i^q + \Delta^q$$

We may multiply both sides by $\widetilde{c}$, and use the fact that $\widetilde{c}\Delta^q \in cP^N \subseteq (y_1, \ldots, y_h)$ to conclude that

$$(*) \quad \widetilde{c}\widetilde{f}^q\widetilde{x}_{i+1}^q \in (\widetilde{x}_1^q, \ldots, \widetilde{x}_i^q, y_1, \ldots, y_h)S$$

But $y_1, \ldots, y_h, \widetilde{x}_1^q, \ldots, \widetilde{x}_{i+1}^q$ is a permutable regular sequence in S, and so $(*)$ implies that

$$\widetilde{c}\widetilde{f}^q \in (\widetilde{x}_1^q, \ldots, \widetilde{x}_i^q, y_1, \ldots, y_h)S.$$

When we consider this modulo P, We have that $(y_1, \ldots, y_h)$ is killed, and so

$$cf^q \in (x_1^q, \ldots, x_i^q)$$

for all $q \geq N$, and this gives the desired conclusion. $\square$

Weak F-regularity: localization at maximal ideals and the Cohen-Macaulay property

We next want to prove that the property of being weakly F-regular is local on the maximal ideals of R. From this we will deduce that a weakly F-regular ring that is a homomorphic image of a Cohen-Macaulay ring is Cohen-Macaulay. We need two preliminary results.

Lemma. *Let R be any Noetherian ring, let M be a finitely generated R-module and $N \subseteq M$ a submodule. Then N is the intersection of a (usually infinite) family of submodules Q of M such that every M/Q is killed by a power of a maximal ideal of R.*

In particular, every ideal I of R is an intersection of ideals that are primary to a maximal ideal of R.

Proof. Let $u \in M - N$. Consider the family of submodules $M_1 \subseteq M$ such that $N \subseteq M$ and $u \notin M_1$. This family is nonempty, since it contains N. Therefore it has a maximal element Q. It will suffice to show that M/Q is killed by a power of a maximal ideal of R. Note that every nonzero submodule of M/Q contains the image of u, or else its inverse image in M will strictly contain Q but will not contain u.

We may replace M by M/Q and u by its image in M/Q. It therefore suffices to show that if $u \neq 0$ is in every nonzero submodule of M, then M is killed by a power of a maximal ideal, which is equivalent to the assertion that $\operatorname{Ass}(M)$ consists of a single maximal ideal. Let $P \in \operatorname{Ass}(M)$ and suppose that $P = \operatorname{Ann}_R v$, where $v \neq 0$ is in M. Then $Rv \cong R/P$, and every nonzero element has annihilator P. But $u \in Rv$, and so $P = \operatorname{Ann}_R u$. It follows that every associated prime of M is the same as $\operatorname{Ann}_R u$, and so there is only one associated prime. It remains to show that P is maximal. Suppose not, and consider $R/P \hookrightarrow M$. It will suffice to show that there is no element in all the nonzero ideals of R/P. Thus, it suffices to show that if $S = R/P$ is a Noetherian domain of dimension at least one, there is no nonzero element in all the nonzero ideals. This is true, in fact, even if we localize at a nonzero prime ideal m of S, for in S_m, there is no element in all of the ideals $m^n S_m$. $\square$

Proposition. *Let R be a Noetherian ring of prime characteristic $p > 0$, and let $\mathfrak{A}$ be an ideal of R.*

(a) *If $\theta : R \to S$ is such that S is flat Noetherian R-algebra and, in particular, if S is a localization of R, then $\theta(\mathfrak{A}^*_R) \subseteq (\mathfrak{A}S)^*_S$.*

(b) *Let m be a maximal ideal of R and suppose that $\mathfrak{A}$ is an m-primary ideal. Let $f \in R$. Then $f \in \mathfrak{A}^*_R$ if and only if $f/1 \in (\mathfrak{A}R_m)^*_{R_m}$.*

(c) *Under the hypotheses of part (b), $\mathfrak{A}$ is tightly closed in R if and only if $\mathfrak{A}R_m$ is tightly closed in R_m.*

Proof. (a) Let $f \in \mathfrak{A}^*_R$. The equation $cf^q \in \mathfrak{A}^{[q]}$ implies $\theta(c)\theta(f)^q \in (\mathfrak{A}S)^{[q]}$, and so we need only see that if $c \in R^\circ$ then $c \in S^\circ$. Suppose, to the contrary, that c is in a minimal prime $\mathfrak{q}$ of S. It suffices to see that the contraction $\mathfrak{p}$ of $\mathfrak{q}$ to R is minimal. But $R_\mathfrak{p} \to S_\mathfrak{q}$ is still faithfully flat, and the maximal ideal of $S_\mathfrak{q}$ is nilpotent, which implies that $\mathfrak{p}R_\mathfrak{p}$ is nilpotent, and so $\mathfrak{p}$ is minimal.

For part (b), we see from (a) that if $f \in \mathfrak{A}^*$ then $f \in (\mathfrak{A}R_m)^*$. We need to prove the converse. Suppose that $c_1 \in R^\circ_m$ has the property that $cf_1^q \in \mathfrak{A}^{[q]}R_m = (\mathfrak{A}R_m)^{[q]}$ for all $q \gg 0$. Then c_1 has the form c/w where $c \in R$ and $w \in R - m$. We may replace c_1 by wc_1, since w is a unit, and therefore assume that $c_1 = c/1$ is the image of $c \in R$. We next

want to replace c by an element with the same image in R_m that is not in any minimal prime of R. Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_k$ be the minimal primes of R that are contained in m, so that the ideals $\mathfrak{p}_j R_m$ for $1 \leq j \leq k$ are *all* of the minimal primes of R_m. It follows that the image of $\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_k$ is nilpotent in R_m, and so we can choose an integer $N > 0$ such that $I = (\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_k)^N$ has image 0 in R_m. If $c + I$ is contained in the union of the minimal primes of R, then by the coset form of prime avoidance, it follows that $cR + I \subseteq \mathfrak{p}$ for some minimal prime $\mathfrak{p}$ of R. Since $I \subseteq \mathfrak{p}$, we have that $\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_k \subseteq \mathfrak{p}$, and it follows that $\mathfrak{p}_j = \mathfrak{p}$ for some j, where $1 \leq j \leq k$. But then $c \in \mathfrak{p}_j$, a contradiction, since $c/1$ is not in any minimal prime of R°. Hence, we can choose $f \in I$ such that $c + f \in R^\circ$, and $c + f$ also maps to $c/1$ in R. We change notation and assume $c \in R^\circ$. Then $cf^q/1 \in \mathfrak{A}^{[q]} R_m$ for all $q \gg 0$. Since $\mathfrak{A}^{[q]}$ is primary to m, the ring $R/\mathfrak{A}^{[q]}$ has only one maximal ideal, $m/\mathfrak{A}^{[q]}$, and is already local. Hence,

$$R/\mathfrak{A}^{[q]} \cong (R/fA^{[q]})_m = R_m/\mathfrak{A}^{[q]} R_m.$$

It follows that $cf^q \in \mathfrak{A}^{[q]}$ for all $q \gg 0$, and so $f \in \mathfrak{A}^*_R$, as required.

Part (c) is immediate from part (b) and the observation above that $R_m/\mathfrak{A}R_m = R/\mathfrak{A}$, so that any element of $R_m/\mathfrak{A}R_m$ is represented by an element of R. $\square$

Remark. Part (a) holds for any map $R \to S$ of Noetherian rings of prime characteristic $p > 0$ such that R° maps into S°. We have already seen another example, namely when $R \hookrightarrow S$ are domains.

Theorem. *The following conditions on R are equivalent.*

(1) *R is weakly F-regular.*

(2) *Every ideal of R primary to a maximal ideal of R is tightly closed.*

(3) *For every maximal ideal m of R, R_m is weakly F-regular.*

Proof. Statements (2) and (3) are equivalent by part (c) of the preceding Proposition, and (1) $\Rightarrow$ (2) is clear. Assume (2), and let I be any ideal of R. We need only show that I is tightly closed. If not, let $f \in I^* - I$. Since I is the intersection of the ideals containing I that are primary to maximal ideals, there is an ideal $\mathfrak{A}$ of R primary to a maximal ideal m such that $I \subseteq \mathfrak{A}$ and $f \notin \mathfrak{A}$. Since $\mathfrak{A}$ is tightly closed and $I \subseteq \mathfrak{A}$, we have $I^* \subseteq \mathfrak{A}$, and so $f \in \mathfrak{A}$, a contradiction. $\square$

Lecture of March 21

We shall no longer be assuming that all rings have prime characteristic $p > 0$. Our objective is to prove some basic results about the structure of complete local rings. We shall begin by studying complete local rings that contain a field. Here are three major results that we are aiming to prove:

Theorem. *Let (R, m, K) be a complete local ring that contains a field.*

(a) *If R is regular, then $R \cong K[[x_1, \ldots, x_d]]$, a formal power series ring in n variables over K, where $d = \dim(R)$.*

(b) *R is a homomorphic image of a formal power series ring $K[[x_1, \ldots, x_n]]$ over a field K.*

(c) *R is a module-finite extension ring of a formal power series ring $K[[x_1, \ldots, x_d]]$, where $d = \dim(R)$.*

Note that part (c) is an analogue, for complete local rings, of the Noether normalization theorem.

We shall later analyze the situation where R does not contain a field in detail. But this is more difficult, and we begin with the field case.

By a *coefficient field* for a local ring (R, m) we mean a subring $K \subseteq R$ such that the composite map

$$K \hookrightarrow R \twoheadrightarrow R/m$$

is an isomorphism. This implies that K is a field, since it is isomorphic with R/m. One may think of K as an isomorphic "copy" of the residue class field that is contained in R. The most difficult part in proving the structure theorems stated above is establishing:

Theorem. *A complete local ring that contains a field contains a coefficient field.*

Proving the preceding two Theorems will take a while. Note that if a local ring R has characteristic 0, which means that it contains $\mathbb{Z}$, the hypothesis that it contains a field is equvalent to the statement that it contains $\mathbb{Q}$. But $\mathbb{Q}$ will typically be much smaller than the residue field of R. The hypothesis that R has prime characteristic $p > 0$ already implies that R contains a field: R will contain the field $\mathbb{Z}/p\mathbb{Z}$.

Example. Let $p > 0$ be a prime integer, let P denote the prime ideal $p\mathbb{Z}$ in $\mathbb{Z}$, and let $\mathcal{Z}_p$ denote the completion of the Noetherian discrete valuation ring $\mathbb{Z}_P$ at its maximal ideal. The ring $\mathcal{Z}_p$ is called the *ring of p-adic integers*. Both Z_P and the $\mathcal{Z}_p$ are examples of local rings that do not contain a field. The ring $\mathcal{Z}_p$ may also be obtained by completing $\mathbb{Z}$ with respect to $p\mathbb{Z}$ without localizing first. The maximal ideal of $\mathcal{Z}_p$ is generated by p: every nonzero element is a power of p times a unit. Every elenent of $\mathcal{Z}_p$ can be represented uniquely as a formal series

$$a_0 + a_1 p + a_2 p^2 + a_3 p^3 + \cdots + a_n p^n + \cdots$$

such that every a_i is an integer between 0 and $p - 1$ inclusive. If the coefficients are eventually all zero, we have the base p representation of an element of $\mathbb{N}$. Note, for example, that in $\mathcal{Z}_2$, we have

$$-1 = 1 + 2 + 4 + 8 + \cdots + 2^n + \cdots$$

Example. Local rings that contain a field but do not have a coefficient field are abundant. Here is a simple example of a local ring that contains a field but does not have a coefficient field. Let V be the localization of the polynomial ring $\mathbb{R}[t]$ in one variable over the real numbers $\mathbb{R}$ at the prime ideal $P = (t^2+1)$, and let $m = PV$. Note that V is a Noetherian discrete valuation ring. Then V/PV is the field of $\mathbb{R}[t]/(t^2+1) \cong \mathbb{C}$, which is $\mathbb{C}$. But $S \subseteq \mathbb{R}(t)$ does not contain any element whose square is -1: the square of a non-constant rational function is non-constant, and the square of a real scalar cannot be -1.

The completion of $\widehat{V}$ of V is also a DVR with residue class field $\mathbb{C}$, and so it must contain a square root of -1. The reader may want to attempt to find an explicit power series in t^2+1 that represents a square root of -1. Note that the structure theorems imply that there is an isomorphism $\mathbb{C}[[z]] \cong \widehat{V}$, and one can show that there is such an isomorphism sending $z \mapsto t^2+1$.

In characteristic 0 we shall show that any subring of the complete local ring R that is maximal with respect to the property of being a field is a coefficient field. The proof will depend on Hensel's Lemma. In characteristic $p > 0$, there may be maximal fields within the complete local ring R that are not coefficient fields. The proof we give will be quite different, and will not make any use of Hensel's Lemma at all.

We begin our analysis of the structure of complete local rings by proving Hensel's lemma.

Theorem (Hensel's Lemma). *Let (R, m, K) be a complete local ring (or a completed and m-adically separated quasilocal ring) and let f be a monic polynomial of degree d in $R[x]$. Suppose that $^-$ indicates images in $K[x]$ under the the ring homomorphism $R[x] \twoheadrightarrow K[x]$ induced by $R \twoheadrightarrow K$. If $\overline{f} = GH$ where $G, H \in K[x]$ are monic of degrees s and t, respectively, and G, H* are relatively prime *in $K[x]$, then there are unique monic polynomials $g, h \in R[x]$ such that $f = gh$ and $\overline{g} = g$ while $\overline{h} = h$.*

Before giving the proof, we want to provide some examples that illustrate how powerful Hensel's Lemma is, as well as an instance where it cannot be applied.

Example 1. Let $R = \mathbb{Q}[[z_1, z_2, z_3]]$. Suppose that we want find a power series which is a square root of $1 + z_1z_2^{11}z_3 + z_1^7 + z_2^5z_3^3$. That is, we want to solve the equation

$$(*) \quad x^2 - (1 + z_1z_2^{11}z_3 + z_1^7 + z_2^5z_3^3) = 0$$

in the formal power series ring $\mathbb{Q}[[z_1, z_2, z_3]]$. This is equivalent to factoring the left hand side of $(*)$ in the form $(x-g)(x-h)$ for elements $g, h \in \mathbb{Q}[[z_1, z_2, z_3]]$. Hensel's Lemma enables us to solve this problem by solving it modulo (z_1, z_2, z_3). Modulo the maximal ideal, the equation becomes $x^2 - 1 = 0$, and the left hand side factors $(x-1)(x+1)$. Moreover, $x-1$ and $x+1$ are relatively prime over $\mathbb{Q}[x]$. We can therefore lift this factorization. This provides two square roots of $1 + z_1z_2^{11}z_3 + z_1^7 + z_2^5z_3^3$. These can also be found using Newton's binomial theorem: let $u = z_1z_2^{11}z_3 + z_1^7 + z_2^5z_3^3$. Then

$$(1+u)^{1/2} = 1 + \frac{1}{2}u + \frac{\frac{1}{2}(\frac{1}{2}-1)}{2!}u^2 + \frac{\frac{1}{2}(\frac{1}{2}-1)(\frac{1}{2}-2)}{3!}u^3 + \cdots$$

and one may substitute the expression $z_1 z_2^{11} z_3 + z_1^7 + z_2^5 z_3^3$ for u. Both methods may be used to show that if n is invertible in $K = R/m$ and $u \in m$, then $1 + u$ has an n th root in the complete local ring R. But Hensel's Lemma is much more general, as the next example shows.

Example 2. Let $R = K[[z_1, z_2, z_3]]$. We shall consider the cases where $K = \mathbb{Q}$ and $K = \mathbb{C}$. Suppose that we want to solve the eqation

$$(\#) \quad x^3 + (z_1^{17} - z_2 z_3^5)x^2 + (z_1 z_2 z_3^8)x - 1 + z_2^7 + z_3^9 = 0$$

over R. When the equation is considered modulo the maximal ideal of R, it becomes $x^3 - 1 = 0$ and has the three roots 1, ω, $\overline{\omega}$ where $\omega = \dfrac{-1 + \sqrt{-3}}{2}$ is a primitive cube root of unity, and $\overline{\omega}$ is the conjugate root $\dfrac{-1 + \sqrt{-3}}{2}$ (we also have $\overline{\omega} = 1/\omega = \omega^2$). Hensel's Lemma applied over $\mathbb{C}$ yields unique roots of the equation (#) with constant terms 1, ω, and $\overline{\omega}$, respectively. If we apply Hensel's Lemma over $\mathbb{Q}$, we still have the factorization

$$x^3 - 1 = (x - 1)(x^2 + x + 1)$$

and the factors are relatively prime over $\mathbb{Q}[x]$. This factorization can therefore be lifted, and this shows that there is a unique root of the equation with constant term 1. This is, of course, the same root with constant term 1 that we found over $\mathbb{C}[[z_1, z_2, z_3]]$, but we have gained the information that the coefficients are rational numbers.

Example 3. Consider the equation $x^2 + 1 = 0$ in $\mathcal{Z}_{13}$. Modulo the maximal ideal, we find that there are two roots in $\mathbb{Z}/13\mathbb{Z}$, represented by 5 and $-5 = 8$. It follows that -1 has two square roots in $\mathcal{Z}_{13}$. Similarly, the reader may verify that 3 has a cube root in $\mathcal{Z}_{61}$ that is congruent to 5 modulo the maximal ideal of $\mathcal{Z}_{61}$.

Example 4. Let $R = \mathbb{C}[[z_1, z_2]]$ and consider the equation $x^2 - z_1^2 - z_2^3 = 0$. Modulo the maximal ideal, this becomes $x^2 = 0$. Of course, x^2 factors as $x \cdot x$, but *the factors are not relatively prime.* Therefore, Hensel's Lemma does not apply. In fact, $z_1^2 + z_2^3$ has no square root in the formal power series ring. Similarly, Hensel's Lemma does not give information about solving $x^2 - z_1 = 0$, which also has no solution.

Proof of Hensel's Lemma. Let F_n denote the image of f in $(R/m^n)[x]$. We recursively construct monic polynomials $G_n \in (R/m^n)[x]$, $H_n \in (R/m^n)[x]$ such that $F_n = G_n H_n$ for all $n \geq 1$, where G_n and H_n reduce to G and H, respectively, mod m, and show that F_n and G_n are unique. Note that it will follow that for all n, G_n has the same degree as G, namely s, and H_n has the same degree as H, namely t, where $s + t = d$. The uniqueness implies that mod m^{n-1}, G_n, H_n become G_{n-1}, H_{n-1}, respectively. This yields that the sequence of coefficients of x^i in the G_n is an element of $\varprojlim_n (R/m^n) = R$, since R is complete. Using the coefficients determined in this way, we get a polynomial g in $R[x]$, monic of degree s. Similarly, we get a polynomial $h \in R[x]$, monic of degree t. It is clear

that $\overline{g} = G$ and $\overline{h} = H$, and that $f = gh$, since this holds mod m^n for all n: thus, every coefficient of $f - gh$ is in $\bigcap_n m^n = (0)$.

It remains to carry through the recursion, and we have $G_1 = G$ and $H_1 = H$ from the hypothesis of the theorem. Now assume that G_n and H_n have been constructed and shown unique for a certain $n \geq 1$. We must construct G_{n+1} and H_{n+1} and show that they are unique as well. It will be convenient to work mod m^{n+1} in the rest of the argument: replace R by R/m^{n+1}. Construct G^*, H^* in $R[x]$ by lifting each coefficient of G_n and H_n respectively, but such that the two leading coefficients occur in degrees s and t respectively and are both 1. Then, mod m^n, $F \equiv G^*H^*$, i.e., $\Delta = F - G^*H^* \in m^n R[x]$. We want to show that there are unique choices of $\delta \in m^n R[x]$ of degree at most $s-1$ and $\epsilon \in m^n R[x]$ of degree at most $t-1$ such that $F - (G^* + \delta)(H^* + \epsilon) = 0$, i.e., such that $\Delta = \epsilon G^* + \delta H^* + \delta\epsilon$. Since $\delta, \epsilon \in m^n R[x]$ and $n \geq 1$, their product is in $m^{2n}R[x] = 0$, because $2n \geq n+1$. Thus, our problem is to find such ϵ and δ with $\Delta = \epsilon G^* + \delta H^*$. Now, G and H generate the unit ideal in $K[x]$, and $R[x]_{\text{red}} = K[x]$. It follows that G^* and H^* generate the unit ideal in $R[x]$, and so we can write $1 = \alpha G^* + \beta H^*$. Multiplying by Δ, we get $\Delta = \Delta\alpha G^* + \Delta\beta H^*$. Then $\Delta\alpha$ and $\Delta\beta$ are in $m^n R[x]$, since Δ is, but do not yet satisfy our degree requirements. Since H^* is monic, we can divide $\Delta\alpha$ by H^* to get a quotient γ and remainder ϵ, i.e., $\Delta\alpha = \gamma H^* + \epsilon$, where the degree of ϵ is $\leq t-1$. If we consider this mod m^n, we have $0 \equiv \gamma H_n + \epsilon$, from which it follows that γ, $\epsilon \in m^n R[x]$. Then $\Delta = \epsilon G^* + \delta H^*$ where $\delta = \gamma G^* + \Delta\beta$. Since Δ and ϵG^* both have degree $< n$, so does δH^*, which implies that the degree of δ is $\leq s-1$.

Finally, suppose that we also have $\Delta = \epsilon' G^* + \delta' H^*$ where ϵ' has degree $\leq t-1$ and δ' has degree $\leq s-1$. Subtracting, we get an equation $0 = \mu G^* + \nu H^*$ where the degree of $\mu = \epsilon - \epsilon'$ is $\leq t-1$ and the degree of $\nu = \delta - \delta'$ is $\leq s-1$. Since G^* is a unit considered mod H^*, it follows that $\mu \in (H^*)$, i.e., that H^* divides μ. But H^* is monic, and so this cannot happen unless $\mu = 0$: the degree of μ is too small. Similarly, $\nu = 0$. $\square$

We can now deduce:

Theorem. *Let (R, m, K) be a complete local ring that contains a field of characteristic 0. Then R has a coefficient field. In fact, R will contain a maximal subfield, and any such subfield is a coefficient field.*

Proof. Let $\mathcal{S}$ be the set of all subrings of R that happen to be fields. By hypothesis, this set is nonempty. Given a chain of elements of $\mathcal{S}$, the union is again a subring of R that is a field. By Zorn's lemma, $\mathcal{S}$ will have a maximal element K_0. To complete the proof of the theorem, we shall show that K_0 maps isomorphically onto K. Obviously, we have a map $K_0 \subseteq R \twoheadrightarrow R/m = K$, and so we have a map $K_0 \to K$. This map is automatically injective: call the image K_0'. To complete the proof, it suffices to show that it is surjective.

If not, let θ be an element of K not in the image of K_0. We consider two cases: the first is that θ is transcendental over K_0'. Let t denote an element of R that maps to θ. Then $K_0[t]$ is a polynomial subring of R, and every nonzero element is a unit: if some element

were in m, then working mod m we would get an equation of algebraic dependence for θ over K_0' in K. By the universal mapping property of localization, the inclusion $K_0[t] \subseteq R$ extends to a map $K_0(t) \subseteq R$, which is necessarily an inclusion. This yields a subfield of R larger than K_0, a contradiction.

We now consider the case where θ is algebraic over the image of K_0. Consider the minimal polynomial of θ over K_0', and let f be the corresponding polynomial with coefficients in $K_0[x] \subseteq R[x]$. Modulo m, this polynomial factors as $(x-\theta)H(x)$, where these are relatively prime because θ is separable over K_0': this is the only place in the argument where we use that the field has characteristic 0. The factorization lifts uniquely: we have $f = (x-t)h(x)$ where $t \in R$ is such that $t \equiv \theta \bmod m$. That is, $f(t) = 0$. We claim that the map $K_0[t] \subseteq R \twoheadrightarrow R/m$, whose image is $K_0'[\theta]$, gives an isomorphism of $K_0[t]$ with $K_0'[\theta]$: we only need to show injectivity. But if $P(x) \in K_0[x]$ is a polynomial such that $P(t)$ maps to 0, then f divides $P(x)$, which implies that $P(t) = 0$. Since $K_0[t] \cong K_0'[\theta]$ (both are $\cong K_0[t]/\big(f(t)\big)$), $K_0[t]$ is a field contained in R that is strictly larger than K_0, a contradiction. $\square$

Remark. If R is a complete local domain of positive prime characteristic $p > 0$, the same argument shows that R contains a maximal subfield K_0, and that K is algebraic and purely inseparable over the image of K_0.

Lecture of March 23

Remark. It is worth noting that Cauchy sequences in an I-adic topology are much easier to study, in some ways, than Cauchy sequences of, say, real numbers. In an I-adic topology, for $\{r_n\}_n$ to be a Cauchy sequence it suffices that $r_n - r_{n+1} \to 0$ as $n \to \infty$, i.e., that for any specified $N \in \mathbb{N}$, the differences $r_n - r_{n+1}$ are eventually in I^N. The reason is that if this is true for all $n \geq n_0$, we also have that

$$r_{n'} - r_n = r_{n'} - r_{n'-1} + \cdots + r_{n+1} - r_n \in I^n$$

for all $n' \geq n \geq n_0$. In consequence, a necessary and *sufficient* condition for an infinite series $\sum_{n=0}^{\infty} r_n$ to converge in the I-adic topology is that $r_n \to 0$ as $n \to \infty$, which, of course, is false over $\mathbb{R}$: the series $\sum_{n=1}^{\infty} 1/n$ does not converge, and the corresponding sequence of partial sums $\{r_n\}_n$ does not converge, even though $r_{n+1} - r_n = 1/(n+1) \to 0$ as $n \to \infty$.

Our next result on coefficient fields uses a completely different argument:

Theorem. *Let (R, m, K) be a complete local ring of positive prime characteristic p. Suppose that K is perfect. Let $R^{p^n} = \{r^{p^n} : r \in R\}$ for every $n \in \mathbb{N}$. Then $K_0 = \bigcap_{n=0}^{\infty} R^{p^n}$ is a coefficient field for R, and it is the only coefficient field for R.*

Proof. Consider any coefficient field L for R, assuming for the moment that one exists. Then $L \cong K$, and so L is perfect. Then

$$L = L^p = \cdots = L^{p^n} = \cdots,$$

and so for all n,

$$L \subseteq L^{p^n} \subseteq R^{p^n}.$$

Therefore, $L \subseteq K_0$. If we know that K_0 is a field, it follows that $L = K_0$, proving uniqueness.

It therefore suffices to show that K_0 is a coefficient field for K. We first observe that K_0 meets m only in 0. For if $u \in K_0 \cap m$, then u is a p^n th power for all n. But if $u = v^{p^n}$ then $v \in m$, so $u \in \bigcap_n m^{p^n} = (0)$.

Thus, every element of $K_0 - \{0\}$ is a unit of R. Now if $u = v^{p^n}$ and u is a unit of R, then $1/u = (1/v)^{p^n}$. Therefore, the inverse of every nonzero element of K_0 is in K_0. Since K_0 is clearly a ring, it is a subfield of R.

Finally, we want to show that given $\theta \in K$ some element of K_0 maps to θ. Let r_n denote an element of R that maps to $\theta^{1/p^n} \in K$. Then $r_n^{p^n}$ maps to θ. We claim that $\{r_n^{p^n}\}_n$ is a Cauchy sequence in R, and so has a limit r. To see this, note that r_n and r_{n+1}^p both map to θ^{1/p^n} in K, and so $r_n - r_{n+1}^p$ is in m. Taking p^n powers, we find that

$$r_n^{p^n} - r_{n+1}^{p^{n+1}} \in m^{p^n}.$$

Therefore, the sequence is Cauchy, and has a limit $r \in R$. It is clear that r maps to θ. Therefore, it suffices to show that $r \in R^{p^k}$ for every k. But

$$r_k, r_{k+1}^p, \ldots, r_{k+h}^{p^h} \cdots$$

is a sequence of the same sort for the element θ^{1/p^k}, and so is Cauchy and has a limit s_k in R. But $s_k^{p^k} = r$ and so $r \in R^{p^k}$ for all k. $\square$

Before pursuing the issue of the existence of coefficient fields further, we show that the existence of a coefficient field implies that the complete local ring is a homomorphic image of a power series ring in finitely many variables over a field, and is also a module-finite extension of such a ring.

We first prove the following result, which bears some resemblance to Nakayama's Lemma, but is rather different, since M is not assumed to be finitely generated.

Proposition. *Let R be separated and complete in the I-adic topology, where I is a finitely generated ideal of R, and let M be an I-adically separated R-module. Let $u_1, \ldots, u_h \in M$ have images that span M/IM over R/I. Then $u_1, \ldots, u_h$ span M over R.*

Proof. Since $M = Ru_1 + \cdots + Ru_h + IM$, we find that for all n,

$$(*) \qquad I^n M = I^n u_1 + \cdots + I^n u_h + I^{n+1} M.$$

Let $u \in M$ be given. Then u can be written in the form $r_{01}u_1 + \cdots + r_{0h}u_h + \Delta_1$ where $\Delta_1 \in IM$. Therefore $\Delta_1 = r_{11}u_1 + \cdots r_{1h}u_h + \Delta_2$ where the $r_{1j} \in IM$ and $\Delta_2 \in I^2M$. Then

$$u = (r_{01} + r_{11})u_1 + \cdots + (r_{0n} + r_{1h})u_h + \Delta_2,$$

where $\Delta_2 \in I^2M$. By a straightforward induction on n we obtain, for every n, that

$$u = (r_{01} + r_{11} + \cdots + r_{n1})u_1 + \cdots + (r_{0h} + r_{1h} + \cdots + r_{nh})u_n + \Delta_{n+1}$$

where every $r_{jk} \in I^j$ for $1 \leq k \leq h$ and all $j \geq 0$ and $\Delta_{n+1} \in I^{n+1}M$. In the recursive step, the formula $(*)$ is applied to the element $\Delta_{n+1} \in I^{n+1}M$.

For every k, $\sum_{j=0}^{\infty} r_{jk}$ represents an element s_k of the complete ring R. We claim that

$$u = s_1u_1 + \cdots + s_hu_h.$$

The point is that if we subtract

$$\sigma_n = (r_{01} + r_{11} + \cdots + r_{n1})u_1 + \cdots + (r_{0h} + r_{1h} + \cdots + r_{nh})u_h$$

from u we get $\Delta_{n+1} \in I^{n+1}M$, and if we subtract σ_n from

$$s_1u_1 + \cdots + s_hu_h$$

we also get an element of $I^{n+1}M$, which we shall justify in greater detail below. Therefore,

$$u - (s_1u_1 + \cdots + s_hu_h) \in \bigcap_n I^{n+1}M = 0,$$

since M is I-adically separated.

It remains to see why $s_1u_1 + \cdots + s_hu_h - \sigma_n$ is in $I^{n+1}M$. This difference can be rewritten as $s_1'u_1 + \cdots + s_h'u_h$ where $s_k' = r_{n+1,k} + r_{n+2,k} + \cdots$. Hence, we simply need to justify the assertion that assertion that if $r_{jk} \in I^j$ for $j \geq n+1$ then

$$r_{n+1,k} + r_{n+2,k} + \cdots + r_{n+t,k} + \cdots \in I^{n+1},$$

which needs a short argument. Since I is finitely generated, we know that I^{n+1} is finitely generated by the monomials of degree $n+1$ in the generators of I, say, $g_1, \ldots, g_d$. Then $r_{n+1+t,k} = \sum_{\nu=1}^{d} q_{t\nu}g_\nu$ with every $q_{t\nu} \in I^t$ and $\sum_{t=0}^{\infty} r_{n+1+t,k} = \sum_{\nu=1}^{d}(\sum_{t=0}^{\infty} q_{t\nu})g_\nu$. $\square$

We also note:

Proposition. *Let $f : R \to S$ be a ring homomorphism. Suppose that S is J-adically complete and separated for an ideal $J \subseteq S$ and that $I \subseteq R$ maps into J. Then there is a unique induced homomorphism $\widehat{R}^I \to S$ that is continuous (i.e., preserves limits of Cauchy sequences in the appropriate ideal-adic topology).*

Proof. $\widehat{R}^I$ is the ring of I-adic Cauchy sequences mod the ideal of sequences that converge to 0. The continuity condition forces the element represented by $\{r_n\}_n$ to map to

$$\lim_{n\to\infty} f(r_n)$$

(Cauchy sequences map to Cauchy sequences: if $r_m - r_n \in I^N$, then $f(r_m) - f(r_n) \in J^N$, since $f(I) \subseteq J$.) It is trivial to check that this is a ring homomorphism that kills the ideal of Cauchy sequences that converge to 0, which gives the required map $\widehat{R}^I \to S$. $\square$

A homomorphism of quasilocal rings $h : (A, \mu, \kappa) \to (R, m, K)$ is called a *local homomorphism* if $h(\mu) \subseteq m$. If A is a local domain, not a field, the inclusion of A in its fraction field is not local. If A is a local domain, any quotient map arising from killing a proper ideal is local. A local homomorphism induces a homomorphism of residue class fields $\kappa = A/\mu \to R/m = K$.

Proposition. *Let A be a Noetherian ring that is complete and separated with respect to an ideal μ, which may be 0, let (R, m, K) be a complete local ring, and let $h : A \to R$ be a homomorphism, so that R is an A-algebra and μ maps into m. Thus, if (A, μ) is local, we are requiring that $A \to R$ be local. Suppose that $f_1, \ldots, f_n \in m$ together with μR generate an m-primary ideal. Then:*

(a) *There is a unique continuous homomorphism $h : A[[X_1, \ldots, X_n]] \to R$ extending the A-algebra map $A[X_1, \ldots, X_n]$ taking X_i to f_i for all i.*

(b) *If K is module-finite over the image of A, then R is module-finite over the image of $A[[X_1, \ldots, X_n]]$ under the map discussed in part (a).*

(c) *If the composite map $A \to R \twoheadrightarrow K$ is surjective, and $\mu R + (f_1, \ldots, f_n)R = m$, then the map h described in (a) is surjective.*

Proof. (a) This is immediate from the preceding Proposition, since $(X_1, \ldots, X_n)$ maps into m.

(b) $A[[X_1, \ldots, X_n]]$ is complete and separated with respect to the the $\mathfrak{A}$-adic topology, where $\mathfrak{A} = (\mu, X_1, \ldots, X_n)A[[X_1, \ldots, X_n]]$. Given a Cauchy sequence of power series $\{f_k\}_k$, it is easy to see that the sequence of coefficients of a fixed monomial $X_1^{\nu_1} \cdots X_n^{\nu_n} = X^\nu$ is a Cauchy sequence in A in the μ-adic topology, and so has a limit $a_\nu \in A$. The only possible limit for the Cauchy sequence $\{f_k\}_k$ is the power series

$$\sum_{\nu \in \mathbb{N}^n} a_\nu X^\nu,$$

and it is easy to verify that this is the limit.

The expansion of the ideal $\mathfrak{A}$ of $A[[X_1, \ldots, X_n]]$ to R is $\mu R + (f_1, \ldots, f_n)R$, which contains a power of m, say m^N. Thus, $R/\mathcal{M}R$ is a quotient of R/m^N and has finite length: the latter has a filtration whose factors are the finite-dimensional K-vector spaces m^i/m^{i+1}, $0 \leq i \leq N-1$. Since K is module-finite over the image of A, it follows that $R/\mathfrak{A}R$ is module finite over over $A[[X_1, \ldots, X_n]]/\mathfrak{A} = A/\mu$. Choose elements of R whose images in $R/\mathfrak{A}R$ span it over A/μ. By the Proposition stated on p. 2, these elements span R as an $A[[X_1, \ldots, X_n]]$-module. We are using that R is $\mathfrak{A}$-adically separated, but this follows because $\mathfrak{A}R \subseteq m$, and R is m-adically separated.

(c) We repeat the argument of the proof of part (b), noting that now $R/\mathfrak{A}R \cong K \cong A/\mu$, so that $1 \in R$ generates R as an $A[[X_1, \ldots, X_n]]$ module. But this says that R is a cyclic $A[[X_1, \ldots, X_n]]$-module spanned by 1, which is equivalent to the assertion that $A[[X_1, \ldots, X_n]] \to R$ is surjective. $\square$

We have now done all the real work needed to prove the following:

Theorem. *Let (R, m, K) be a complete local ring with coefficient field $K_0 \subseteq K$, so that $K_0 \subseteq R \twoheadrightarrow R/m = K$ is an isomorphism. Let $f_1, \ldots, f_n$ be elements of m generating an ideal primary to m. Let $K_0[[X_1, \ldots, X_n]] \to R$ be constructed as in the preceding Proposition, with X_i mapping to f_i and with $A = K_0$. Then:*

(a) *R is module-finite over $K_0[[X_1, \ldots, X_n]]$.*

(b) *Suppose that $f_1, \ldots, f_n$ generate m. Then the homomorphism $K_0[[x_1, \ldots, x_n]] \to R$ is surjective. (By Nakayama's lemma, the least value of n that may be used is the dimension as a K-vector space of m/m^2.)*

(c) *If $d = \dim(R)$ and $f_1, \ldots, f_d$ is a system of parameters for R, the homomorphism*

$$K_0[[x_1, \ldots, x_d]] \to R$$

is injective, and so R is a module-finite extension of a formal power series subring.

Proof. (a) and (b) are immediate from the preceding Proposition. For part (c), let $\mathfrak{A}$ denote the kernel of the map $K_0[[x_1, \ldots, x_d]] \to R$. Since R is a module-finite extension of the ring $K_0[[x_1, \ldots, x_d]]/\mathfrak{A}$, $d = \dim(R) = \dim(K_0[[x_1, \ldots, x_d]]/\mathfrak{A})$. But we know that $\dim(K_0[[x_1, \ldots, x_d]]) = d$. Killing a nonzero prime in a local domain must lower the dimension. Therefore, we must have that $\mathfrak{A} = (0)$. $\square$

Thus, when R has a coefficient field K_0 and $f_1, \ldots, f_d$ are a system of parameters, we may consider a formal power series

$$\sum_{\nu \in \mathbb{N}^d} c_\nu f^\nu,$$

where $\nu = (\nu_1, \ldots, \nu_d)$ is a multi-index, the $c_\nu \in K_0$, and f^ν denotes $f_1^{\nu_1} \cdots f_d^{\nu_d}$. Because R is complete, this expression represents an element of R. Part (c) of the preceding Theorem implies that this element is not 0 unless all of the coefficients c_ν vanish. This fact is sometimes referred to as the *analytic independence of a system of parameters*. The elements of a system of parameters behave like formal indeterminates over a coefficient field. Formal indeterminates are also referred to as *analytic indeterminates*.

Lecture of March 25

The results of the preceding Lecture imply that a complete local ring (R, m) that has a coefficient field K is a homomorphic image of a formal power series ring in n variables over K, where n is the least number of elements needed to generate m. Of course, by Nakayama's Lemma, $n = \dim_K(m/m^2)$. This integer is called the *embedding dimension* of R.

To understand why, consider the analogous situation with finitely generated reduced algebras S over an algebraically closed field K. The ring S corresponds to an affine algebraic set X, whose points are in bijective correspondence with the maximal ideals of S. Giving a surjection $K[X_1, \ldots, X_n] \twoheadrightarrow S$ as K-algebras is equivalent to giving an embedding $X \hookrightarrow \mathbb{A}^n_K$ as a closed algebraic set. The least n for which such an embedding is possible is the smallest dimension of an affine space in which X can be embedded, and it is natural to think of n as the embedding dimension of X, and hence, of S, in this context. The terminology "embedding dimension" for $\dim_K(m/m^2)$ is used even when the local ring (R, m, K) does not contain a field.

The general construction of coefficient fields in positive characteristic

We now discuss the construction of coefficient fields in local rings (R, m, K) of prime characteristic $p > 0$ (these automatically contain the field $\mathbb{Z}/p\mathbb{Z}$) when K need not be perfect. If $q = p^n$ we write

$$K^q = \{c^q : c \in K\},$$

the subfield of K consisting of all elements that are q th powers.

It will be convenient to call a polynomial in several variables *n-special*, where $n \geq 1$ is an integer, if every variable occurs with exponent at most $p^n - 1$ in every term. This terminology is not standard.

Let K be a field of characteristic $p > 0$. Finitely many elements $\theta_1, \ldots, \theta_n$ in K (they will turn out to be, necessarily, in $K - K^p$) are called *p-independent* if the following three equivalent conditions are satisfied:

(1) $[K^p[\theta_1, \ldots, \theta_n] : K^p] = p^n$.

(2) $K^p \subseteq K[\theta_1] \subseteq K^p[\theta_1, \theta_2] \subseteq \cdots \subseteq K^p[\theta_1, \theta_2, \ldots, \theta_n]$ is a strictly increasing tower of fields.

(3) The p^n monomials $\theta_1^{a_1} \cdots \theta_n^{a_n}$ such that $0 \leq a_j \leq p-1$ for all j with $1 \leq j \leq n$ are a K^p-vecctor space basis for K over K^p.

Note that since every θ_j satisfies $\theta_j^p \in K^p$, in the tower considered in part (2) at each stage there are only two possibilities: the degree of θ_{j+1} over $K^p[\theta_1, \ldots, \theta_j]$ is either 1, which means that

$$\theta_{j+1} \in K^p[\theta_1, \ldots, \theta_j],$$

or p. Thus, $K[\theta_1, \ldots, \theta_n] = p^n$ occurs only when the degree is p at every stage, and this is equivalent to the statement that the tower of fields is strictly increasing. Condition (3) clearly implies condition (1). The fact that (2) $\Rightarrow$ (3) follows by mathematical induction from the observation that

$$1, \theta_{j+1}, \theta_{j+1}^2, \ldots, \theta_{j+1}^{p-1}$$

is a basis for $L_{j+1} = K^p[\theta_1, \ldots, \theta_{j+1}]$ over $L_j = K[\theta_1, \ldots, \theta_j]$ for every j, and the fact that if one has a basis $\mathcal{C}$ for L_{j+1} over L_j and a basis $\mathcal{B}$ for L_j over K^p then all products of an element from $\mathcal{C}$ with an element from $\mathcal{B}$ form a basis for L_{j+1} over K^p.

Every subset of a p-independent set is p-independent. An infinite subset of K is called *p-independent* if every finite subset is p-independent.

A maximal p-independent subset of K, which will necessarily be a subset of $K - K^p$, is called a *p-base* for K. Zorn's Lemma guarantees the existence of a p-base, since the union of a chain of p-independent sets is p-independent. If Θ is a p-base, then $K = K^p[\Theta]$, for if there were an element θ' of $K - K^p[\Theta]$, it could be used to enlarge the p-base. The empty set is a p-base for K if and only if K is perfect. if K is not perfect, a p-base for K is *never* unique: one can change an element of it by adding an element of K^p.

It is easy to see that Θ is a p-base for K if and only if every element of K is uniquely expressible as a polynomial in the elements of Θ with coefficients in K^p such that the exponent on every $\theta \in \Theta$ is at most $p-1$, i.e., the monomials in the elements of Θ of degree at most $p-1$ in each element are a basis for K over K^p. An equivalent statement is that every element of K is uniquely expressible as as 1-special polynomial in the elements of Θ with coefficients in K^p.

If $q = p^n$, then the elements of $\Theta^q = \{\theta^q : \theta \in \Theta\}$ are a p-base for K^q over K^{pq}: in fact we have a commutative diagram:

$$\begin{array}{ccc} K & \xrightarrow{F^q} & K^q \\ \uparrow & & \uparrow \\ K^p & \xrightarrow[F^{pq}]{} & K^{pq} \end{array}$$

where the vertical arrows are inclusions and the horizontal arrows are isomorphisms: here, $F^q(c) = c^q$. In particular, $\Theta^p = \{\theta^p : \theta \in \Theta\}$ is a p-base for K^p, and it follows by multiplying the two bases together that the monomials in the elements of Θ of degree at most $p^2 - 1$ are a basis for K over K^{p^2}. By a straightforward induction, the monomials in the elements of Θ of degree at most $p^n - 1$ in each element are a basis for K over K^{p^n} for every $n \in \mathbb{N}$. An equivalent statement is that every element of K can be written uniquely as an n-special polynomial in the elements of Θ with coefficients in K^{p^n}.

Theorem. *Let (R, m, K) be a complete local ring of positive prime characteristic p, and let Θ be a p-base for K. Let T be a subset of R that maps bijectively onto Θ, i.e., a lifting of the p-base to R. Then there is a unique coefficient field for R that contains T, namely, $K_0 = \bigcap_n R_n$, where $R_n = R^{p^n}[T]$. Thus, there is a bijection between liftings of the p-base Θ and the coefficient fields of R.*

Proof. Note that any coefficient field must contain some lifting of Θ. Observe also that K_0 is clearly a subring of R that contains T. It will suffice to show that K_0 is a coefficient field and that any coefficient field L containing T is contained in K_0. The latter is easy: the isomorphism $L \to K$ takes T to Θ, and so T is a p-base for L. Every element of L is therefore in $L^{p^n}[T] \subseteq R^{p^n}[T]$. Notice also that every element of $R^{p^n}[T]$ can be written as a polynomial in the elements of T of degree at most $p^n - 1$ in each element, i.e., as an n-special polynomial, with coefficients in R^{p^n}. The reason is that any $N \in \mathbb{N}$ can be written as $ap^n + b$ with $a, b \in \mathbb{N}$ and $b \leq p^n - 1$. So t^N can be rewritten as $(t^a)^{p^n} t^b$, and, consequently, if t^N occurs in a term we can rewrite that term so that it only involves t^b by absorbing $(t^a)^{p^n}$ into the coefficient from R^{p^n}. Thus, every element of $R^{p^n}[T]$ is represented by an n-special polynomial. Note that n-special polynomials in elements of T with coefficients in R^{p^n} map mod m onto the n-special polynomials in elements of Θ with coefficients in K^{p^n}, which we know give all of K.

We next observe that

$$R^{p^n}[T] \cap m \subseteq m^{p^n}.$$

Write the element of $u \in R^{p^n}[T] \cap m$ as an n-special polynomial in elements of T with coefficients in R^{p^n}. Then its image in K, which is 0, is an n-special polynomial in the elements of Θ with coefficients in K^{p^n}, and so cannot vanish unless every coefficient is 0. This means that each coefficient of the n-special polynomial representing u must have been in $m \cap R^{p^n} \subseteq m^{p^n}$. Thus,

$$K_0 \cap m = \bigcap_n (R^{p^n}[T] \cap m) \subseteq \bigcap_n m^{p^n} = (0).$$

We can therefore conclude that K_0 injects into K. It will suffice to show that $K_0 \to K$ is surjective to complete the proof.

Let $\lambda \in K$ be given. Since $K = K^{p^n}[\Theta]$, for every n we can choose an element of $R^{p^n}[T]$ that maps to λ: call it r_n. Then $r_{n+1} \in R^{p^{n+1}}[T] \subseteq R^{p^n}[T]$, and so $r_n - r_{n+1} \in$

$R^{p^n}[T] \cap m \subseteq m^{p^n}$ (the difference $r_n - r_{n+1}$ is in m because both r_n and r_{n+1} map to λ in K). This shows that $\{r_n\}_n$ is Cauchy, and has a limit r_λ. It is clear that $r_\lambda \equiv \lambda$ mod m, since that is true for every r_n. Moreover, r_λ is independent of the choices of the r_n: given another sequence r'_n with the same property, $r_n - r'_n \in R^{p^n}[T] \cap m \subseteq m^{p^n}$, and so $\{r_n\}_n$ and $\{r'_n\}_n$ have the same limit. This implies that the map $K \to R$ such that $\lambda \mapsto R_\lambda$ is a ring homomorphism: if we have two Cauchy sequences whose terms map to λ and λ' respectively mod K, and whose nth terms are both in $R^{p^n}[T]$ for all n, when we add (respectively, multiply) the Cauchy sequences term by term, we get a Cauchy sequence whose limit is $r_{\lambda+\lambda'}$ (respectively, $r_{\lambda\lambda'}$). Moreover, if $t \in T$ maps to $\theta \in \Theta$ then the Cauchy sequence with constant term t can be used to find r_θ, and so $r_\theta = t$.

It remains only to show that for every n, $r_\lambda \in R^{p^n}[T]$. To see this, write λ as an n-special polynomial in the elements of Θ with coefficients in K^{p^n}. Explicitly,

$$\lambda = \sum_{\mu \in \mathcal{F}} c_\mu^{p^n} \mu$$

where $\mathcal{F}$ is some finite set of n-special monomials in the elements of Θ, and every $c_\mu \in K$. If $\mu = \theta_1^{k_1} \cdots \theta_s^{k_s}$, let $\mu' = t_1^{k_1} \cdots t_s^{k_s}$, where t_j is the element of T that maps to θ_j. Then $r_\mu = \mu'$ and

$$r_\lambda = \sum_{\mu \in \mathcal{F}} r_{c_\mu}^{p^n} \mu' \in R^{p^n}[T]. \qquad \square$$

Remark. The proof is valid for every complete and m-adically separated quasilocal ring (R, m, K) such that R has prime characteristic $p > 0$. We made no use of the fact that R is Noetherian.

Remark. This result shows that if (R, m, K) is a complete local ring that is not a field and K is not perfect, then the choice of a coefficient field is *never* unique. Given a lifting of a p-base T, where $T \neq \emptyset$ because K is not perfect, we can always change it by adding nonzero elements of m to one or more of the elements in the p-base.

Lecture of March 28

Consider a complete local ring (R, m, K). If K has characteristic 0, then $\mathbb{Z} \to R \to K$ is injective, and $\mathbb{Z} \subseteq R$. Moreover, no element of $W = \mathbb{Z} - \{0\}$ is in m, since no element of W maps to 0 in $R/m = K$, and so every element of $\mathbb{Z} - \{0\}$ has an inverse in R. By the universal mapping property of localization, we have a unique map of $W^{-1}\mathbb{Z} = \mathbb{Q}$ into R, and so R is an equicharacteristic 0 ring. We already know that R has a coefficient field. We also know this when R has prime characteristic $p > 0$, i.e., when $\mathbb{Z}/p\mathbb{Z} \subseteq R$.

We now want to develop the structure theory of complete local rings when R need not contain a field. From the remarks above, we only need to consider the case where K

has prime characteristic $p > 0$, and we shall assume this in the further development of the theory. The coefficient rings that we are about to describe also exist in the complete separated quasi-local case, but, for simplicity, we only treat the Noetherian case.

We shall say that V is a *coefficient ring* if it is a field or if it is complete local of the form (V, pV, K), where K has characteristic $p > 0$. If R is complete local we shall say that V is a *coefficient ring for* R if V is a coefficient ring, $V \subseteq R$ is local, and the induced map of residue fields is an isomorphism. We shall prove that coefficient rings always exist.

In the case where the characteristic of K is $p > 0$, there are three possibilities. It may be that $p = 0$ in R (and V), in which case V is a field: we have already handled this case. It may be that p is not nilpotent in V: in this case it turns out that V is a Noetherian discrete valuation domain (DVR), like the p-adic integers. Finally, it may turn out that p is not zero, but is nilpotent.

We are aiming to prove the following two results. Like the other theorems we have been proving about the structure of complete local rings, they are due to I. S. Cohen.

Theorem. *Let (R, m, K) be a complete local ring of mixed characteristic. Then R has a coefficient ring.*

Theorem. *Let (W, pW, K) be a coefficient ring of mixed charateristic such that p is nilpotent. Then W has the form V/p^hV, where (V, pV, K) is a coefficient ring that is a complete Noetherian discrete valuation ring.*

Before proving these two results, which will take a considerable effort, we want to give several consequences.

Theorem. *Let R be a complete local ring of mixed characteristic.*

(a) *R is a homomorphic image of a power series ring $V[[X_1, \ldots, X_n]]$ over a complete Noetherian discrete valuation ring (V, pV, K), where n is the embedding dimension of R/pR.*

(b) *If R is a domain, or more generally, if p is part of a system of parameters for R, then R is module-finite over a formal power series ring $V[[x_2, \ldots, x_{d-1}]]$, where $d = \dim(R)$ and V is a complete Noetherian discrete valuation ring that is a coefficient ring for R.*

(c) *Suppose that R is regular of Krull dimension d and that V is a complete Noetherian discrete valuation ring that is a coefficient ring for R. If $p \notin m^2$, then $R \cong V[[x_2, \ldots, x_d]]$, a formal power series ring. If R is regular and $p \in m^2$, then $R \cong V[[x_1, \ldots, x_d]]/(f)$, where the numerator is a formal power series ring and $f = p - g$ with g is in the square of the maximal ideal of $V[[x_1, \ldots, x_d]]$.*

Proof. (a) Let W be a coefficient ring for R and let V be a coefficient ring that is a discrete valuation ring that maps onto W. Choose $f_1, \ldots, f_n \in R$ that map onto a minimal set

of generators of the maximal ideal of R/pR. Then p together with the $f_1, \ldots, f_n$ map onto generators of m. By part (a) of the Proposition stated at the top of p. 4 of the Lecture Notes of March 23, there is a map $W[[x_1, \ldots, x_n]] \to R$ that takes $x_1, \ldots, x_n$ to $f_1, \ldots, f_n$ respectively, and this map is a surjection by part (c) of that same Proposition, with $A = W$ and $\mu = pW$. Hence, we have surjections

$$V[[x_1, \ldots, x_n]] \twoheadrightarrow W[[x_1, \ldots, x_n]] \twoheadrightarrow R,$$

as required.

(b) Since p is part of a system of parameters, it is not nilpotent, and a coefficient ring (V, p, K) for R must be a Noetherian discrete valuation ring. Let $f_2, \ldots, f_d \in m$ be elements that extend p to a system of parameters for R. By parts (a) and (b) of the Proposition cited above, we have a map $V[[x_2, \ldots, x_d]] \to R$ such that R is module finite over the image. Since the $\dim(R) = d$, the image has dimension d, and since $V[[x_2, \ldots, x_d]]$ is a domain of dimension d, the map cannot have a kernel.

(c) If R is regular and $p \notin m^2$, then we can extend p to a minimal set of generators $p, f_2, \ldots, f_d$ of m, and we have a map $V[[x_2, \ldots, x_d]] \to R$ that is injective and such that R is module-finite over the image by part (b). But we are also in the situation of part (a), so that this map is surjective, and this gives the required isomorphism of R with a formal power series ring.

Now suppose that $p \in m^2$. We proceed as in part (a), but choose $f_1, \ldots, f_d$ so that they are a minimal set of generators of m. Let $T = V[[x_1, \ldots, x_d]]$, the formal power series ring, and let m_T be its maximal ideal. Then we have a surjection $T \twoheadrightarrow R$. Since $p \in m^2$, the kernel of this map must contain an element of the form $p - g$, where $g \in m_T^2$. But $f = p - g \in m_T - m_T^2$, and so $T/(f)$ is a regular local ring of dimension d that maps onto R. Since $T/(f)$ is regular, it is a domain, and it follows that the map $T/(f) \twoheadrightarrow R$ cannot have a non-trivial kernel. Thus, $T/(f) \cong R$, as required. $\square$

A regular local ring of mixed characteristic $p > 0$ is called *unramified* if $p \notin m^2$ and *ramified* if $p \in m^2$.

Example. Let $R = V[[x]]/(px)$, where (V, pV, K) is a coefficient ring, and x is a power series indeterminate over V. The image of V in R is isomorphic with V and is a coefficient ring. R is one-dimensional, and is not module-finite over a regular ring: cf. problem 5. of Problem Set #5.

It remains to prove the results of I. S. Cohen about coefficient rings for complete local rings of mixed characteristic, including the statement that they exist. The following elementary fact is critical in carrying this through.

Lemma. *Let (R, m, K) be local with K of prime characteristic $p > 0$. If $r, s \in R$ are such that $r \equiv s \bmod m$, and $n \geq 1$ is an integer, then for all $N \geq n - 1$, with $q = p^N$ we have that $r^q \cong s^q \bmod m^n$.*

Proof. This is clear if $n = 1$. We use induction. If $n > 1$, we know from the induction hypothesis that $r^q \equiv y^q \bmod m^N$ if $N \geq n-2$, and it suffices to show that $r^{pq} \equiv y^{pq} \bmod m^{N+1}$. Since $r^q = s^q + u$ with $u \in m^N$, we have that $r^{pq} = (s^q + u)^p = s^{pq} + puw + u^p$, where puw is a sum of terms from the binomial expansion each of which has the form $\binom{pq}{j} s^j u^{p-j}$ for some j, $1 \leq j \leq p-1$, and in each of these terms the binomial coefficient is divisible by p. Since $u \in m^N$ and $p \cdot 1_R \in m$, $puw \in m^{N+1}$, while $u^p \in m^{Np} \subseteq m^{N+1}$ as well. $\square$

Lecture of March 30

The following Theorem, which constructs coefficient rings when the maximal ideal of the ring is nilpotent, is the heart of the proof of the existence of coefficient rings in complete mixed characteristic local rings. Before giving the proof, we introduce the following notation, which we will use in another argument later. Let x, y be indeterminates over $\mathbb{Z}$. Let q be a power of p, a prime. Then $(x+y)^q - x^q - y^q$ is divisible by p in $\mathbb{Z}[x, y]$, since the binomial coefficients that occur are all divisible by p, and we write $G_q(x, y) \in \mathbb{Z}[x, y]$ for the quotient, so that $(x+y)^q = x^q + y^q + pG_q(x, y)$.

Theorem. *Suppose that (R, m, K) is local where K has characteristic $p > 0$, and that $m^n = 0$. Choose a p-base Θ for K, and a lifting of the p-base to R: that is, for every $\theta \in \Theta$ choose an element $t_\theta \in R$ with residue θ modulo m. Let $T = \{t_\theta : \theta \in \Theta\}$. Then R has a unique coefficient ring V that contains T. In fact, suppose that we fix any sufficiently large power $q = p^N$ of p (in particular, $N \geq n-1$ suffices) and let S_N be the set of all expressions of the form $\sum_{\mu \in \mathcal{M}} r_\mu^q \mu$, where the $\mathcal{M}$ is a finite set of mutually distinct N-special monomials in the elements of T and every $r_\mu^q \in R^q = \{r^q : r \in R\}$. Then we may take*

$$V = S_N + pS_N + p^2 S_N + \cdots + p^{n-1} S_N,$$

which will be the same as the smallest subring of R containing R^q and T.

Before giving the proof, we note that it is not true in general that R^q is closed under addition, and neither is S_N, but we will show that for large N, V is closed under addition and multiplication, and this will imply at once that it is the smallest subring of R containing R^q and T. Of course, R^q *is* closed under multiplication.

Proof of the Proposition. We first note if $r \equiv s \bmod m$ then $r^q \equiv s^q \bmod m^n$ if $N \geq n-1$, by the Lemma at then end of the Lecture Notes of March 25. Therefore R^q maps *bijectively* onto $K^q = \{\lambda^q : \lambda \in K\}$ when we take residue classes mod m. It follows from our analysis of the properties of p-bases that the residue class map $R \to K$ sends S_N *bijectively* onto K.

Suppose that W is a coefficient ring containing T. For each $r \in R$, if $w \equiv r \bmod m$, then $w^q = r^q$. Thus, $R^q \subseteq W$. Then $S_N \subseteq W$, and so $V \subseteq W$. Now consider any

element $w \in W$. Since S_N contains a complete set of representatives of elements of K, every element of W has the form $\sigma_0 + u$ where $\sigma_0 \in S_N$ and $u \in m \cap W = pW$, and so $w = \sigma_0 + pw_1$. But we may also write w_1 in this way and substitute, to get an expression

$$w = \sigma_0 + p\sigma_1 + p^2 w_2,$$

where σ_0, $\sigma_1 \in S_N$ and $w_2 \in W$. Continuing in this way, we find, by a straightforward induction, that

$$W = S_N + pS_N + \cdots + p^j S_N + p^{j+1} W$$

for every $j \geq 0$. We may apply this with $j = n-1$ and note that $p^n = 0$ to conclude that $W = V$. Thus, if there is a coefficient ring, it must be V. However, at this point we do not even know that V is closed under addition.

We next claim that V is a ring. Let $\widetilde{V}$ be the closure of V under addition. Then we can see that $\widetilde{V}$ is a ring, since, by the distributive law, it suffices to show that the product of two elements $p^i r^q \mu$ and $p^j r'^q \mu'$ has the same form. The point is that $\mu\mu'$ can be rewritten in the form $\nu^q \mu''$ where μ'' has all exponents $\leq q-1$, and $p^{i+j}(rr'\nu)^q \mu''$ has the correct form. Thus, $\widetilde{V}$ is the smallest ring that contains R^q and T.

We next prove that V itself is closed under addition. We shall achieve this by proving by reverse induction on j that $p^j V = p^j \widetilde{V}$ for all j, $0 \leq j \leq n$. The case that we are really aiming for is, of course, where $j = 0$. The statement is obvious when $j = n$, since $p^n = 0$ and $p^n V = p^n \widetilde{V} = 0$. Now suppose that $p^{j+1} V = p^{j+1} \widetilde{V}$ for some fixed j. We shall show that $p^j V = p^j \widetilde{V}$, thereby completing the inductive step. Since $p^j \widetilde{V}$ is spanned over $p^{j+1}\widetilde{V} = p^{j+1} V$ by $p^j S_N$, it will suffice to show that given any two elements of $p^j S_N$, their sum differs from an element of $p^j S_N$ by an element of $p^{j+1}\widetilde{V} = p^{j+1} V$. Call the two elements

$$v = p^j \sum_{\mu \in \mathcal{M}} r_\mu^q \mu$$

and

$$v' = p^j \sum_{\mu \in \mathcal{M}} r'^q_\mu \mu,$$

where $r_\mu, r'_\mu \in R$ and $\mathcal{M}$ is a finite set of n-special monomials in elements of T large enough to contain all those monomials that occur with nonzero coefficient in the expressions for v and v'. Since S_N gives a complete set of representatives of K and r^q only depends on what r is modulo m, we may assume that all of the r_μ and r'_μ are elements of S_N. Let

$$v'' = p^j \sum_{\mu \in \mathcal{M}} (r_\mu + r'_\mu)^q \mu.$$

Then

$$v'' - v - v' = p^j \sum_{\mu \in \mathcal{M}} pG_q(r_\mu, r'_\mu)\mu = p^{j+1} \sum_{\mu \in \mathcal{M}} G_q(r_\mu, r'_\mu)\mu \in p^{j+1} V',$$

as required, since all the r_μ, $r'_\mu \in S_N$ and $\widetilde{V}$ is a ring. This completes the proof that $\widetilde{V} = V$, and so V is a subring of R.

We have now shown that V is a subring of R, and that it is the only possible coefficient ring. It is clear that $pV \subseteq m$, while an element of $V - pV$ has nonzero image in K: its constant term in S_N is nonzero, and S_N maps bijectively to K. Thus, $m \cap V = pV$, and we know that $V/pV \cong K$, since S_N maps onto K. It follows that pV is a maximal ideal of V generated by a nilpotent, and so pV is the only prime ideal of V. Any nonzero element of the maximal ideal can be written as $p^t u$ with t as large as possible (we must have that $t < n$), and then u must be a unit. Thus, every nonzero element of V is either a unit, or a unit times a power of p. It follows that every nonzero proper ideal is generated by p^k for some positive integer k, where k is as small as possible such that p^k is in the ideal. It follows that V is a principal ideal ring. Thus, V is a Noetherian local ring, and, in fact, an Artin local ring. $\square$

We want to extend this result to complete local rings in which m is not nilpotent. We first need:

Lemma. *Let K be a field of characteristic $p > 0$ and let (V, pV, K), (W, pW, K) and (V_n, pV_n, K), $n \in \mathbb{N}$, be coefficient rings.*

(a) *If $p^t = 0$ while $p^{t-1} \neq 0$ in V, which is equivalent to the statement that p^t is the characteristic of V, then $\mathrm{Ann}_V p^j V = p^{t-j}V$, $0 \leq j \leq t$. Moreover, if $p^s = 0$ while $p^{s-1} \neq 0$ in W, and $W \twoheadrightarrow V$ is a surjection, then $V = W/p^t W$.*

(b) *Suppose that*

$$V_0 \twoheadleftarrow V_1 \twoheadleftarrow \cdots \twoheadleftarrow V_n \twoheadleftarrow \cdots$$

is an inverse limit system of coefficient rings and surjective maps, and that the characteristic of V_n is $p^{t(n)}$ where $t(n) \geq 1$. Then either $t(n)$ is eventually constant, in which case the maps $h_n : V_{n+1} \twoheadrightarrow V_n$ are eventually all isomorphisms, and the inverse limit is isomorphic with V_n for any sufficiently large n, or $t(n) \to \infty$ as $n \to \infty$, in which case the inverse limit is a complete local principal ideal domain V with maximal ideal pV and residue class field K. In particular, the inverse limit V is a coefficient ring.

Proof. (a) Every ideal of V (respectively, W) has the form $p^k V$ (respectively, p^W)for a unique integer k, $0 \leq k \leq t$ (respectively, $0 \leq k \leq s$). The first statement follows because $k + j \geq n$ iff $k \geq n - j$. The second statement follows because V must have the form $S/p^k S$ for some k, $0 \leq k \leq S$, and the characteristic of $S/p^k S$ is p^k, which must be equal to p^t.

(b) If $t(n)$ is eventually constant it is clear that all the maps are eventually isomorphisms. Therefore, we may assume that $t(n) \to \infty$ as $n \to \infty$. By passing to an infinite subsequence of the V_n we may assume without loss of generality that $t(n)$ is strictly increasing with n. We may think of an element of the inverse limit as a sequence of elements $v_n \in V_n$ such that v_n is the image of v_{n+1} for every n. It is easy to see that one of the v_n

is a unit if and only if all of them are. Suppose on the other hand that none of the v_n is a unit. Then each v_n can be written as pw_n for $w_n \in V_n$. The problem is that while pw_{n+1} maps to pw_n, for all n, it is not necessarily true that w_{n+1} maps to w_n.

Let h_n be the map $V_{n+1} \to V_n$. For all n, let $w'_n = h_n(w_{n+1})$. We will show that for all n, $v_n = pw'_n$ and that $h_n(w'_{n+1}) = w'_n$ for all n. Note first that $h_n(pw_{n+1}) = pw_n = v_n$, and it is also pw'_n. This establishes the first statement. Since $p(w_{n+1} - w'_{n+1}) = 0$, it follows that $w_{n+1} - w'_{n+1} = p^{t(n+1)-1}\delta$, by part (a). Then

$$w'_n = h_n(w_{n+1}) = h_n(w'_{n+1}) + p^{t(n+1)-1}h_n(\delta) = h_n(w'_{n+1}),$$

as required, since $p^{t(n+1)-1}$ is divisible by $p^{t(n)}$, the characteristic of V_n.

It follows that the inverse limit has a unique maximal ideal generated by p. No nonzero element is divisible by arbitrarily high powers of p, since the element will have nonzero image in V_n for some n, and its image in this ring is not divisible by arbitrarily high powers of p. It follows that every nonzero element can be written as a power of p times a unit, and no power of p is 0, because the ring maps onto V_t for arbitrarily large values of t. It is forced to be a principal ideal domain in which every nonzero ideal is generated by a power of p. The fact that the ring arises as an inverse limit implies that it is complete. □

We can now prove:

Theorem (I. S. Cohen). *Every complete local ring (R, m, K) has a coefficient ring. If the residue class field has characteristic $p > 0$, there is a unique coefficient ring containing a given lifting T to R of a p-base Θ for K.*

Proof. We may assume that K has characteristic $p > 0$: we already know that there is a coefficient field if the characteristic of K is 0.

Any coefficient ring for R containing T must map onto a coefficient ring for $R_n = R/m^n$ containing the image of T. Here, there is a unique coefficient ring V_n, which may be described, for any sufficiently large $q = p^N$, as the smallest subring containing all qth powers and the image of T . We may take q large enough that it may be used in the description of coefficient rings V_{n+1} for R_{n+1} and V_n for R_n, and it is then clear that $R_{n+1} \twoheadrightarrow R_n$ induces $V_{n+1} \twoheadrightarrow V_n$. If we construct $V = \varprojlim_n V_n$ and $\varprojlim_n R_n = R$ as sequences of elements $\{r_n\}_n$ such that r_{n+1} maps to r_n for all n, it is clear that

$$V = \varprojlim_n V_n \subseteq \varprojlim_n R_n = R.$$

By part (b) of the preceding Lemma, V is a coefficient ring, and it follows that V is a coefficient ring for R. □

Lecture of April 1

We next prove that, up to non-unique isomorphism, a coefficient ring of mixed characteristic p in which p is nilpotent is determined by its residue class field and and characteristic (the latter is a power of p). However, there is a uniqueness statement about the isomorphism once liftings of a p-base for K are chosen.

Theorem. *Let K, K' be isomorphic fields of characteristic $p > 0$ and let $g : K \to K'$ be the isomorphism. Let (V, pV, K) and (V', pV', K') be two coefficient rings of the same characteristic, $p^n > 0$. We shall also write λ' for the image of $\lambda \in K$ under g. Let Θ be a p-base for K and let $\Theta' = g(\Theta)$ be the corresponding p-base for K'. Let T be a lifting of Θ to V and let T' be a lifting of Θ' to T'. We have an obvious bijection $\widetilde{g} : T \to T'$ such that if $t \in T$ lifts $\theta \in \Theta$ then $\widetilde{g}(t) \in T'$ lifts $\theta' = g(\theta)$. Then $\widetilde{g}$ extends uniquely to an isomorphism of V with V' that lifts $g : K \to K'$.*

Proof. As in the proof of the Theorem on existence of coefficient rings stated on the first page of the Lecture Notes of March 30, we choose $N \geq n-1$ and let $q = p^N$. For every element $\lambda \in K$ there is a unique element $\rho_\lambda \in V^q$ that maps to $\lambda^q \in K^q$. Similarly, there is a unique element $\rho'_{\lambda'} \in V'^q$ that maps to λ'^q for every $\lambda' \in K'$. If there is an isomorphism $V \cong V'$ as stated, it must map $\rho_\lambda \to \rho'_{\lambda'}$ for every $\lambda \in K$. Said otherwise, we have an obvious bijection $V^q \to V'^q$, and $\widetilde{g}$ must extend it. Just as in the proof of the Theorem on existence of coefficient rings, we can define $S_N = S$ to consist of linear combinations of distinct N-special monomials in T such that every coefficient is in V^q. Then S will map bijectively onto K. We define $S'_N = S' \subseteq V'$ analogously. Since S' maps bijectively onto K', we have an obvious bijection $\widetilde{g} : S \to S'$. We use σ' for the element of S' corresponding to $\sigma \in S$.

Every element $v \in V$ must have the form $\sigma_0 + pv_1$ where σ_0 is the unique element of S that has the same residue as v modulo pV. Continuing this way, as in the proof of the Theorem on existence of coefficient rings, we get a representation

$$v = \sigma_0 + p\sigma_1 + p^2\sigma_2 + \cdots + p^{n-1}\sigma_{n-1}$$

for the element $v \in V$, where the $\sigma_j \in S$. We claim this is unique. Suppose we have another such representation

$$v = \sigma_0^* + p\sigma_1^* + \cdots + p^{n-1}\sigma_{n-1}^*.$$

Suppose that $\sigma_i = \sigma_i^*$ for $i < j$. We want to show that $\sigma_j = \sigma_j^*$ as well. Working in $V/p^{j+1}V$ we have that $\sigma_j p^j = \sigma_{j+1} p^j$, i.e., that $(\sigma_j - \sigma_j^*)$ kills p^j working mod p^{j+1}. By part (a) of the Lemma from p. 3 of the Lecture Notes of March 30 we have that $\sigma_j - \sigma_j^* \in pV$, and so σ_j and σ_j^* represent the same element of $K = V/pV$, and therefore are equal.

Evidently, any isomorphism $V \cong V'$ satisfying the specified conditions must take

$$\sigma_0 + p\sigma_1 + \cdots + p^{n-1}\sigma_{n-1}$$

to

$$\sigma'_0 + p\sigma'_1 + \cdots + p^{n-1}\sigma'_{n-1}.$$

To show that this map really does give an isomorphism of V with V' one shows simultaneously, by induction on j, that addition is preserved in p^jV, and that multiplication is preserved when one multiplies elements in p^hV and p^iV such that $h + i \geq j$. For every element $\lambda \in K$, let σ_λ denote the unique element of S that maps to λ. Note that we may write ρ_λ as σ_λ^q, since σ_λ has residue λ mod pV.

Now,

$$p^j\rho_\lambda\mu + p^j\rho_\eta\mu = p^j(\sigma_\lambda^q + \sigma_\eta^q)\mu = p^j\big((\sigma_\lambda + \sigma_\eta)^q - pG_q(\sigma_\lambda, \sigma_\eta)\big),$$

where $G_q(x, y) \in \mathbb{Z}[x, y]$ is such that $(x + y)^q = x^q + y^q + pG_q(x, y)$. Since $\sigma_\lambda + \sigma_\eta$ has residue $\lambda + \eta$ mod pV, we have that $(\sigma_\lambda + \sigma_\eta)^q = \rho_{\lambda+\eta}$, and it follows that

$$p^j\rho_\lambda\mu + p^j\rho_\eta\mu = p^j\rho_{\lambda+\eta}\mu - p^{j+1}G_q(\sigma_\lambda, \sigma_\eta)\mu.$$

We have similarly that

$$p^j\rho'_{\lambda'}\mu' + p^j\rho'_{\eta'}\mu' = p^j\rho'_{\lambda'+\eta'}\mu' - p^{j+1}G_q(\sigma'_{\lambda'}, \sigma'_{\eta'})\mu',$$

and it follows easily that addition is preserved by our map $p^jV \to p^jV'$: note that $p^{j+1}G_q(\sigma_\lambda, \sigma_\eta)\mu$ maps to $p^{j+1}G_q(\sigma'_{\lambda'}, \sigma'_{\eta'})\mu'$ because all terms are multiples of p^{j+1} (the argument here needs that certain multiplications are preserved as well addition).

Once we have that our map preserves addition on terms in p^jV, the fact that it preserves products of pairs of terms from $p^hV \times p^iV$ for $h + i \geq j$ follows from the distributive law, the fact that addition in p^jV is preserved, and the fact that there is a unique way of writing $\mu_1\mu_2$, where μ_1 and μ_2 are monomials in the elements of T with all exponents $\leq q - 1$, in the form $\nu^q\mu_3$ where all exponents in μ_3 are $\leq q - 1$, and

$$(p^h\rho_\lambda\mu_1)(p^i\rho_\eta\mu_2) = p^{h+i}(\sigma_\lambda\sigma_\eta\nu)^q\mu_3$$

in V, while

$$(p^h\rho'_{\lambda'}\mu'_1)(p^i\rho'_{\eta'}\mu'_2) = p^{h+i}(\sigma'_{\lambda'}\sigma'_{\eta'}\nu')^q\mu'_3$$

in V'. $\square$

Theorem. *Let K be a field of characteristic $p > 0$. Then there exists a complete Noetherian valuation domain (V, pV, K) with residue class field K.*

Proof. It suffices to prove that there exists a Noetherian valuation domain (V, pV, K): its completion will then be complete with the required properties. Choose a well-ordering of K in which 0 is the first element. We construct, by transfinite induction, a direct limit system of Noetherian valuation domains $\{V_\lambda, pV_\lambda, K_\lambda\}$ indexed by the well-ordered set K and injections $K_a \hookrightarrow K$ such that

(1) $K_0 \cong \mathbb{Z}/p\mathbb{Z}$

(2) The image of K_λ in K contains a.

(3) The diagrams

$$\begin{array}{ccccc} V_{\lambda'} & \twoheadrightarrow & K_{\lambda'} & \hookrightarrow & K \\ \uparrow & & \uparrow & & \| \\ V_\lambda & \twoheadrightarrow & K_\lambda & \hookrightarrow & K \end{array}$$

commute for all $\lambda \leq \lambda' \in K$.

Note the given a direct limit system of Noetherian valuation domains and injective local maps such that the same element, say, t (in our case $t = p$) generates all of their maximal ideals, the direct limit, which may be thought of as a directed union, of all of them is a Noetherian discrete valuation domain such that t generates the maximal ideal, and such that the residue class field is the directed union of the residue class fields. Every element of any of these rings not divisible by t is a unit (even in that ring): thus, if W is the directed union, pW is the unique maximal ideal. Every nonzero element of the union is a power of t times a unit, since that is true in any of the valuation domains that contain it, and it follows that every nonzero ideal is generated by the smallest power of p that it contains. The statement about residue class fields is then quite straightforward.

Once we have a direct limit system as described, the direct limit will be a discrete Noetherian valuation domain in which p generates the maximal ideal and the residue class field is isomorphic with K.

It will therefore suffice to construct the direct limit system.

We may take $V_0 = \mathbb{Z}_P$ where $P = p\mathbb{Z}$. We next consider an element $\lambda' \in K$ which is the immediate successor of $\lambda \in K$. We have a Noetherian discrete valuation domain $(V_\lambda, pV_\lambda, K_\lambda)$ and an embedding $K_\lambda \hookrightarrow K$. We want to enlarge V_λ suitably to form $V_{\lambda'}$. If λ' is transcendental over K_λ we simply let $V_{\lambda'}$ be the localization of the polynomial ring $V_\lambda[x]$ in one variable over V_λ at the expansion of pV_λ: the residue class field may be identified with $K_\lambda(x)$, and the embedding of $K_\lambda \hookrightarrow K$ may be extended to the simple transcendental extension $K_\lambda(x)$ so that x maps to $\lambda' \in K$.

If λ' is already in the image of K_λ we may take $V_{\lambda'} = V_\lambda$. If instead λ' is algebraic over the image of K_λ, but not in the image, then it satisfies a minimal monic polynomial

$g = g(x)$ of degree at least 2 with coefficients in the image of K_λ. Lift the coefficients to V_λ so as to obtain a monic polynomial $G = G(x)$ of the same degree over V_λ. We shall show that $V_{\lambda'} = V_\lambda[x]/\big(G(x)\big)$ has the required properties. If G were reducible over the fraction field of V_λ, by Gauss' Lemma it would be reducible over V_λ, and then g would be reducible over the image of K_λ in K. If follows that $\big(G(x)\big)$ is prime in $V_\lambda[x]$ and so $V_{\lambda'}$ is a domain that is a module-finite extension of V_λ. Consider a maximal ideal m of $V_{\lambda'}$. Then the chain $m \supset (0)$ in V_b lies over a chain of distinct primes in V_λ: since V_λ has only two distinct primes, we see that m lies over pV_λ and so $p \in m$. But

$$V_{\lambda'}/pV_{\lambda'} \cong \mathrm{Im}\,(K_\lambda)[x]/g(x) \cong \mathrm{Im}\,(K_\lambda)[\lambda'],$$

and so p must generate a unique maximal ideal in $V_{\lambda'}$, and the residue class field behaves as we require as well.

Finally, if λ' is a limit ordinal, we first take the direct limit of the system of Noetherian discrete valuation domains indexed by the predecessors of λ', and then enlarge this ring as in the preceding paragraph so that the image of its residue class field contains λ'. □

Corollary. *If p is a positive prime integer and K is field of characteristic p, there is, up to isomorphism, a unique coefficient ring of characteristic $p > 0$ with residue class field K and characteristic p^t, and it has the form V/p^tV, where (V, pV, K) is a Noetherian discrete valuation domain.*

Proof. By the preceding Theorem, we can construct V so that it has residue field K. Then V/p^tV is a coefficient ring with residue class field K of characteristic p, and we already know that such all rings are isomorphic, which establishes the uniqueness statement. □

Corollary. *Let p be a positive prime integer, K a field of characteristic p, and suppose that (V, pV, K) and (W, pW, K) are complete Noetherian discrete valuation domains with residue class field K. Fix a p-base Θ for K. Let T be a lifting of Θ to V and T' a lifting to W. Then there is a unique isomorphism of V with W that maps each element of T to the element with the same residue in Θ in T'.*

Proof. By our results for the case where the maximal ideal is nilpotent, we get a unique such isomorphism $V/p^nV \cong W/p^nW$ for every n, and this gives an isomorphism of the inverse limit systems

$$V/pV \twoheadleftarrow V/p^2V \twoheadleftarrow \cdots \twoheadleftarrow V/p^nV \twoheadleftarrow \cdots$$

and

$$W/pW \twoheadleftarrow W/p^2W \twoheadleftarrow \cdots \twoheadleftarrow W/p^nW \twoheadleftarrow \cdots$$

that takes the image of T in each V/p^nV to the image of T' in the corresponding W/p^nW. This induces an isomorphism of the inverse limits, which are V and W, respectively. □

Lecture of April 4

Let $p > 0$ be a prime integer. We now know that a coefficient ring of mixed characteristic p and characteristic p^n, where $n \geq 2$, is determined up to isomorphism by its residue class field. Let K be a given field of characteristic p. We also know that there is a complete mixed characteristic Noetherian discrete valuation ring (V, pV, K) with residue class field K. This implies that V/p^nV is a coefficient ring of characteristic p^n. Hence, as asserted earlier:

Theorem. *A mixed characteristic coefficient ring of characteristc p^n, where $p > 0$ is prime, has the form V/p^nV, where V is a complete Noetherian discrete valuation ring that is a coefficient field.* □

This completes our proof of all of the structure theorems for complete local rings. We restate the following:

Theorem. *Every complete local ring is either a homomorphic image of $K[[x_1, \ldots, x_n]]$, a power series ring over a field K, or of $V[[x_1, \ldots, x_n]]$, a power series ring over a mixed characteristic coefficient ring (V, pV, K) that is a Noetherian discrete valuation ring.*

Both $K[[x_1, \ldots, x_n]]$ and $V[[x_1, \ldots, x_n]]$ are Cohern-Macaulay, as is every regular local ring, since a minimal system of generators for the maximal ideal is a regular sequence. But Cohen-Macaulay rings are universally catemary. Hence:

Corollary. *Every complete local ring is universally catenary.* □

Complete regular local rings are formal power series rings in equal chacteristic, and also in mixed characteristic if unramified. The following is an important tool in working with formal power series rings.

Theorem (Weierstrass preparation theorem). *Let (A, m, K) be a complete local ring and let x be a formal indeterminate over A. Let $f = \sum_{n=0}^{\infty} a_n x^n \in A[[x]]$, where $a_h \in A - m$ is a unit and $a_n \in m$ for $n < h$. (Such an element f is said to be* regular *in x of order h.) Then the images of $1, x, \ldots, x^{h-1}$ are a free basis over A for the ring $A[[x]]/fA[[x]]$, and every element $g \in A[[x]]$ can be written uniquely in the form $qf + r$ where $q \in A[[x]]$, and $r \in A[x]$ is a polynomial of degree $\leq h - 1$.*

Proof. Let $M = A[[x]]/(f)$, which is a finitely generated $A[[x]]$-module, and so will be separated in the $\mathcal{M}$-adic topology, where $\mathcal{M} = (m, x)A[[x]]$. Hence, it is certainly separated in the m-adic topology. Then $M/mM \cong K[[x]]/(\overline{f})$, where $\overline{f}$ is the image of f under the map $A[[x]] \twoheadrightarrow K[[x]]$ induced by $A \twoheadrightarrow K$: it is the result of reducing coefficients of f mod m. It follows that the lowest nonzero term of $\overline{f}$ has the form cx^h, where $c \in K$, and so $\overline{f} = x^h\gamma$ where γ is a unit in $K[[x]]$. Thus,

$$M/mM \cong K[[x]]/(\overline{f}) = K[[x]]/(x^h),$$

which is a K-vector space for which the images of 1, x, ... , x^{h-1} form a K-basis. By the Proposition on p. 2 of the Lecture Notes of March 23, the elements 1, x, ... , x^{h-1} span $A[[x]]/(f)$ as an A-module. This means precisely that every $g \in A[[x]]$ can be written $g = qf + r$ where $r \in A[x]$ has degree at most $h-1$.

Suppose that $g'f + r'$ is another such representation. Then $r' - r = (q - q')f$. Thus, it will suffice to show if $r = qf$ is a polynomial in x of degree at most $h-1$, then $q = 0$ (and $r = 0$ follows). Suppose otherwise. Since some coefficient of q is not 0, we can choose t such that q is not 0 when considered mod $m^t A[[x]]$. Choose such a t as small as possible, and let d be the least degree such that the coefficient of x^d is not in m^t. Pass to R/m^t. Then q has lowest degree term ax^d, and both a and all higher coefficients are in m^{t-1}, or we could have chosen a smaller value of t. When we multiply by f (still thinking mod m^t), note that all terms of f of degree smaller than h kill q, because their coefficients are in m. There is at most one nonzero term of degree $h + d$, and its coefficient is not zero, because the coefficient of x^h in f is a unit. Thus, qf has a nonzero term of degree $\geq h + d > h - 1$, a contradiction. This completes the proof of the existence and uniqueness of q and r. □

Corollary. *Let $A[[x]]$ and f be as in the statement of the Weierstrass Preparation Theorem, with f regular of order h in x. Then f has a unique multiple fq which is a monic polynomial in $A[x]$ of degree h. The multiplier q is a unit, and qf has all non-leading coefficients in m. The polynomial qf called the* unique monic associate *of f.*

Proof. Apply the Weierstrass Preparation Theorem to $g = x^h$. Then $x^h = qf + r$, which says that $x^h - r = qf$. By the uniqueness part of the theorem, these are the only choices of q, r that satisfy the equation, and so the uniqueness statement follows. It remains only to see that q is a unit, and that r has coefficients in m. To this end, we may work mod $mA[[x]]$. We use $\overline{u}$ for the class of $u \in A[[x]]$ mod $mA[[x]]$, and think of $\overline{u}$ as an element of $K[[x]]$.

Then $x^h - \overline{r} = \overline{q}\,\overline{f}$. Since $\overline{f}$ is a unit γ times x^h, we must have $\overline{r} = 0$. It follows that $x^h = x^h \overline{q}\gamma$. We may cancel x^h, and so $\overline{q}$ is a unit of of $K[[x]]$. It follows that q is a unit of $A[[x]]$, as asserted. □

Discussion. This result is often applied to the formal power series ring in n variables, $K[[x_1, \ldots, x_n]]$: one may take $A = K[[x_1, \ldots, x_{n-1}]]$ and $x = x_n$, for example, though, obviously, one might make any of the variable play the role of x. In this case, a power series f is regular in x_n if it involves a term of the form cx_n^h with $c \in K - \{0\}$, and if one takes h as small as possible, f is regular of order h in x_n. The regularity of f of order h in x_n is equivalent to the assertion that under the unique continuous $K[[x_n]]$-algebra map $K[[x_1, \ldots, x_n]] \to K[[x_n]]$ that kills $x_1, \ldots, x_{n-1}$, the image of f is a unit times x_n^h. A logical notation for the image of f is $f(0, \ldots, 0, x_n)$. The Weierstrass preparation theorem asserts that for any g, we can write $f = qg + r$ uniquely, where $q \in K[[x_1, \ldots, x_n]]$, and $r \in K[[x_1, \ldots, x_{n-1}]][x_n]$. In this context, the unique monic associate of f is sometimes call the *distinguished pseudo-polynomial* associated with f. If $K = \mathbb{R}$ or $\mathbb{C}$ one can consider instead the ring of convergent (on a neighborhood of 0) power series. One can carry through the proof of the Weierstrass preparation theorem completely constructively, and show that

when g and f are convergent, so are q and r. See, for example, [O. Zariski and P. Samuel, *Commutative Algebra*, Vol. II, D. Van Nostrand Co., Inc., Princeton, 1960], pp. 139–146.

Any nonzero element of the power series ring (convergent or formal) can be made regular in x_n by a change of variables. The same applies to finitely many elements $f_1, \ldots, f_s$, since it suffices to make the product $f_1 \cdots f_s$ regular in x_n, (if the image of $f_1 \cdots f_s$ in $K[[x_n]]$ is nonzero, so is the image of every factor). If the field is infinite one may make use of a K-automorphism that maps $x_1, \ldots, x_n$ to a different basis for $Kx_1 + \cdots + Kx_n$. One can think of f as $f_0 + f_1 + f_2 + \cdots$ where every f_j is a homogeneous polynomial of degree j in $x_1, \ldots, x_n$. Any given form occurring in $f_j \neq 0$ can be made into a monic polynomial by a suitable linear change of variables, by problem **3.** of Problem Set #3 for Math 614, Fall 2003 and its solution.

If K is finite one can still get the image of f under an automorphism to be regular in x_n by mapping $x_1, \ldots, x_n$ to $x_1 + x_n^{N_1}, \ldots, x_{n-1} + x_n^{N_{n-1}}, x_n$, respectively, as in the proof of the Noether normalization theorem, although the details are somewhat more difficult. Consider the monomials that occur in f (there is at least one, since f is not 0), and totally order the monomials so that $x_1^{j_1} \cdots x_n^{j_n} < x_1^{k_1} \cdots x_n^{k_n}$ means that for some i, $1 \leq i \leq n$, $j_1 = k_1$, $j_2 = k_2$, …, $j_{i-1} = k_{i-1}$, while $j_i < k_i$. Let $x_1^{d_1} \cdots x_n^{d_n}$ be the smallest monomial that occurs with nonzero coefficient in f with respect to this ordering, and let $d = \max\{d_1, \ldots, d_n\}$. Let $N_i = (nd)^{n-i}$, and let θ denote the continuous K-automorphism of $K[[x_1, \ldots, x_n]]$ that sends $x_i \mapsto x_i + {x_n}^{N_i}$ for $1 \leq i \leq n-1$, and $x_n \mapsto x_n$. We claim that $\theta(f)$ is regular in x_n. The point is that the value of $\theta(f)$ after killing $x_1, \ldots, x_{n-1}$ is

$$f(x_n^{N_1}, x_n^{N_2}, \ldots, x_n^{N_{n-1}}, x_n),$$

and the term $c' x_1^{e_1} \cdots x_n^{e_n}$ where $c' \in K - \{0\}$ maps to

$$c' x_n^{e_1 N_1 + e_2 N_2 + \cdots + e_{n-1} N_{n-1} + e_n}.$$

In particular, there is a term in the image of $\theta(f)$ coming from the $x_1^{d_1} \cdots x_n^{d_n}$ term in f, and that term is a nonzero scalar multiple of

$$x_n^{d_1 N_1 + d_2 N_2 + \cdots + d_{n-1} N_{n-1} + d_n}.$$

It suffices to show that no other term cancels it, and so it suffices to show that if for some i with $1 \leq i \leq n$, we have that $e_j = d_j$ for $j < i$ and $e_i > d_i$, then

$$e_1 N_1 + e_2 N_2 + \cdots + e_{n-1} N_{n-1} + e_n > d_1 N_1 + d_2 N_2 + \cdots + d_{n-1} N_{n-1} + d_n.$$

Subtracting the right hand side of the inequality above from the left hand side yields

$$(e_i - d_i) N_i + \sum_{j>i} (e_j - d_j) N_j,$$

since $d_j = e_j$ for $j < i$. It will be enough to show that this difference is positive. Since $e_i > d_i$, the leftmost term is at least N_i. Some of the remaining terms are nonnegative,

and we omit these. The terms for those j such $e_j < d_j$ are negative, but what is being subtracted is bounded by $d_j N_j \leq dN_j$. Since at most $n-1$ terms are being subtracted, the sum of the quantities being subtracted is strictly bounded by $nd \max_{j>i}\{dN_j\}$. The largest of the N_j is N_{i+1}, which is $(dn)^{n-(i+1)}$. Thus, the total quantity being subtracted is strictly bounded by $(dn)(dn)^{n-i-1} = (dn)^{n-i} = N_i$. This completes the proof that

$$e_1N_1 + e_2N_2 + \cdots + e_{n-1}N_{n-1} + e_n > d_1N_1 + d_2N_2 + \cdots + d_{n-1}N_{n-1} + d_n,$$

and we see that $\theta(f)$ is regular in x_n, as required. $\square$

If the Weierstrass Preparation Theorem is proved directly for a formal or convergent power series ring R over a field K (the constructive proofs do not use *a priori* knowledge that the power series ring is Noetherian), the theorem can be used to prove that the ring R is Noetherian by induction on n. The cases where $n = 0$ or $n = 1$ are obvious: the ring is a field or a discrete valuation ring. Suppose the result is known for the power series ring A in $n-1$ variables, and let R be the power series ring in one variable x_n over A. Let I be an ideal of R. We must show that I is finitely generated over R. If $I = (0)$ this is clear. If $I \neq 0$ choose $f \in I$ with $f \neq 0$. Make a change of variables such that f is regular in x_n over A. Then $I/fR \subseteq R/fR$, which is a finitely generated module over A. By the induction hypothesis, A is Noetherian, and so R/fR is Noetherian over A, and hence I/fR is a Noetherian A-module, and is finitely generated as an A-module. Lift these generators to I. The resulting elements, together with f, give a finite set of generators for I.

Although we shall later give a quite different proof valid for all regular local rings, we want to show how the Weierstrass preparation theorem can be used to prove unique factorization in a formal power series ring.

Theorem. *Let K be a field and let (V, π, K) be a Noetherian discrete valuation ring. $R = K[[x_1, \ldots, x_n]]$ or $V[[x_1, \ldots, x_n]]$ be the formal power series ring in n variables over K or V. Then R is a unique factorization domain.*

Proof. We use induction on n. If $n = 0$ then R is a field or a discrete valuation ring. In the latter case, R is a principal ideal domain and, hence, a unique factorization domain.

Suppose that $n \geq 1$. It suffices to prove that if $f \in m$ is irreducible then f is prime. If π divides f, the f is a multiple of π by a unit, since f is irreducible. We know that π is prime, since $R/(\pi) \cong K[[x_1, \ldots, x_n]]$, a domain. Hence, we may assume that π does not divide f. Suppose that f divides gh, where it may be assumed without loss of generality that $g, h \in m$. Then we have an equation $fw = gh$, and since f is irreducible, we must have that $w \in m$ as well. If some power of π divides w, then π divides g or h. We may factor out π and obtain a similar equation in which a lower power of π divides w. Eventually, we obtain an equation in which π does not divide w: otherwise, w would be in every power of the maximal ideal. Then π does not divide g nor h as well. Hence, π does not divide fgh.

Therefore, by the Discussion on pp. 3 and 4, we can make a change of variables in the formal power series ring such that fgh is regular in x_n modulo π. Since an element of the

ring that is a unit modulo π is a unit, we have that fgh is regular in x_n in R as well. Then f, g, and h are all regular in x_n, and we may multiply each by a unit so as to replace it by its unique monic associate: here we view R as $A[[x_n]]$ where $A = K[[x_1, \ldots, x_{n-1}]]$ or $V[[x_1, \ldots, x_{n-1}]]$. Thus, we may assume without loss of generality that f, g, and h are monic polynomials in $A[x_n]$ whose non-leading coefficients are in $Q = (x_1, \ldots, x_{n-1})A$. In the process of replacing f, g, h by their products units, w is replaced by its product with a certain unit as well, so that we still have $fw = gh$. However, *a priori*, w may be a power series in x_n rather than a polynomial.

It is easy, however, to see that $w \in A[x_n]$ as well. We can divide $gh \in A[x_n]$ by f, which is monic in x_n, to get a unique quotient and remainder, say $gh = qf + r$, where the degree of r is less than the degree d of f. The Weierstrass preparation theorem guarantees a unique such representation in $A[[x_n]]$, and in the larger ring we know that $r = 0$. Therefore, the equation $gh = qf$ holds in $A[x_n]$, and this means that $q = w$ is a monic polynomial in x_n as well.

By the induction hypothesis, A is a UFD, and so $A[x_n]$ is a UFD. We first note that f is still irreducible in $A[x_n]$ (this is an issue because it might factor as a polynomial with an invertible constant term in one factor: such a factorization does not contradict irreducibility in $A[[x_n]]$). But if f factors non-trivially $f = f_1 f_2$ in $A[x_n]$, the factors f_1, f_2 must be polynomials in x_n of lower degree which can be taken to be monic. Mod Q, f_1, f_2 give a factorization of x_n^d, and this must be into two powers of x_n of lower degree. Therefore, f_1 and f_2 both have all non-leading coefficients in Q, and, in particular their constant terms are in Q. This implies that neither f_1 nor f_2 is a unit of R, and this contradicts the irreducibility of f in R. Thus, f must be irreducible in $A[x_n]$ as well. But then, in $A[x_n]$ we have that $f \,|\, g$ or $f \,|\, h$, and the same obviously holds in the larger ring R, as required. $\square$

Lecture of April 6

We next want to prove unique factorization in all regular local rings, and we shall use an entirely different method. We first discuss the basic facts about the divisor class group $\mathcal{C}\ell\,(R)$ of a normal Noetherian domain R.

Primary decomposition of principal ideals in a normal Noetherian domain has a particularly simple form: there are no embedded primes, and so if $0 \neq a \in P$ the P-primary component is unique, and corresponds to the contraction of an ideal primary to the maximal ideal in R_P, a discrete valuation ring. But the only ideals primary to PR_P in R_P are the powers of PR_P, and so every P-primary ideal has the form $P^{(n)}$ for a unique positive integer n, where $P^{(n)}$ denotes the nth sybolic power of P, the contraction of $P^n R_P$ to R. Thus, if $a \neq 0$ is not a unit, then aR is uniquely an intersection

$$P_1^{(k_1)} \cap \cdots \cap P_n^{(k_n)}.$$

Form the free abelian group G on generators that are taken either to be the height one primes of R (as we shall do) or elements in bijective correspondence with the height one primes of R. The elements of G are called *divisors*. If the ideal aR has the primary decomposition indicated, the element $\sum_{i=1}^{n} k_i P_i$ is called the *divisor* of a, and denoted $\operatorname{div}(a)$. The coefficient of P is the same as the order of a in the discrete valuation ring R_P. By convention, the divisor of a unit of R is 0. The quotient of G by the span of all the divisors is called the *divisor class group* of R, and denoted $\mathcal{C}\ell(R)$. It turns out to vanish if and only if R is a UFD. In fact, P maps to 0 in $\mathcal{C}\ell(R)$ iff P is principal. One can say something even more general. An ideal I of a Noetherian ring R is said to have *pure height* h if all associated primes of I as an ideal have height h. The unit ideal, which has no associated primes, satisfies this condition by default. If I is an ideal of a Noetherian normal domain of pure height one, then I has a primary decomposition $P_1^{(k_1)} \cap \cdots \cap P_n^{(k_n)}$, and so there is a divisor $\operatorname{div}(I)$ associated with I, namely $\sum_{i=1}^{n} k_i P_i$. If $I = R$ is the unit ideal, we define $\operatorname{div}(I) = 0$.

Theorem. *Let R be a Noetherian normal domain. If I has pure height one, then so does fI for every nonzero element f of R, and $\operatorname{div}(fI) = \operatorname{div}(f) + \operatorname{div}(I)$. For any two ideals I and J of pure height one, $\operatorname{div}(I) = \operatorname{div}(J)$ iff $I = J$, while the images of $\operatorname{div}(I)$ and $\operatorname{div}(J)$ in $\mathcal{C}\ell(R)$ are the same iff there are nonzero elements f, g of R such that $fI = gJ$. This holds iff I and J are isomorphic as R-modules. In particular, I is principal if and only if $\operatorname{div}(I)$ is 0 in the divisor class group. Hence, R is a UFD if and only if $\mathcal{C}\ell(R) = 0$.*

The elements of $\mathcal{C}\ell(R)$ are in bijective correspondence with isomorphism classes of pure height one ideals considered as R-modules, and the inverse of the element represented by $\operatorname{div}(I)$ is given by $\operatorname{div}(J)$, for a pure height one ideal $J \cong \operatorname{Hom}_R(I, R)$. In fact, if $g \in I - \{0\}$, we may take $J = gR :_R I$.

Proof. $I = J$ iff $\operatorname{div}(I) = \operatorname{div}(J)$ because, for pure height one ideals, the associated divisor completely determines the primary decomposition of the ideal. Observe that we have $0 \subseteq fR/fI \subseteq R/fI$ and that the cokernel is isomorphic with R/fR while $fR/fI \cong R/I$. Since $\operatorname{Ass}(R/I)$ contains only height one primes and $\operatorname{Ass}(R/fR)$ contains only height one primes (since R is normal), it follows that $\operatorname{Ass}(R/aI)$ contains only height one primes. The statement that $\operatorname{div}(fI) = \operatorname{div}(f) + \operatorname{div}(I)$ may be checked locally after localizing at each height one prime ideal Q, and is obvious in the case of a discrete valuation ring. In particular, $\operatorname{div}(fg) = \operatorname{div}(f) + \operatorname{div}(g)$ when $f, g \in R - \{0\}$. It follows easily that

$$\operatorname{Span}\{\operatorname{div}(f) : f \in R - \{0\}\} = \{\operatorname{div}(g) - \operatorname{div}(f) : f, g \in R - \{0\}\}.$$

Thus, if $\operatorname{div}(I) = \operatorname{div}(J)$ in $\mathcal{C}\ell(R)$, then $\operatorname{div}(I) - \operatorname{div}(J) = \operatorname{div}(g) - \operatorname{div}(f)$ and so $\operatorname{div}(fI) = \operatorname{div}(gJ)$ and $fI = gJ$. Then $I \cong fI = gJ \cong J$ as modules. Now suppose $\theta : I \cong J$ as modules (it does not matter whether I, J have pure height one) and let $g \in I - \{0\}$ have image f in J. For all $a \in I$, $g\theta(a) = \theta(ga) = a\theta(g) = af$, and so $\theta(a) = (f/g)a$, and θ is precisely multiplication by f/g. This yields that $(f/g)I = J$ and, hence, $fI = gJ$.

Now fix $I \neq (0)$ and $g \in I - \{0\}$. Any map $I \to R$ is multiplication by a fraction f/g, where f is the image of g in R: thus, $\mathrm{Hom}_R(I, R) \cong \{f \in R : (f/g)I \subseteq R\}$, where the homomorphism corresponding to multiplication by f/g is mapped to f. But $(f/g)I \subseteq R$ iff $fI \subseteq gR$, i.e., iff $f \in gR :_R I$. Thus, $\mathrm{Hom}_R(I, R) \cong gR :_R I = J$. We claim that J has pure height one (even if I does not) and that if I has pure height one then $\mathrm{div}\,(J) + \mathrm{div}\,(I) = \mathrm{div}\,(g)$, which shows that $\mathrm{div}\,(J) = -\mathrm{div}\,(I)$ in $\mathcal{C}\ell\,(R)$. Let $f_1, \ldots, f_k$ generate I. Then we have an exact sequence

$$0 \to gR :_R I \to R \to (R/gR)^{\oplus k}$$

where the map from R sends $r \mapsto (\overline{rf_1}, \ldots, \overline{rf_k})$ with the overlines indicating residues modulo gR. It follows that $R/(gR :_R I)$ embeds in $(R/gR)^{\oplus k}$, and so

$$\mathrm{Ass}\left(R/(gR :_R I)\right) \subseteq \mathrm{Ass}\left((R/gR)^{\oplus k}\right) = \mathrm{Ass}\,(R/gR),$$

which shows that all associated primes of $gR :_R I$ have pure height one. Now localize at any height one prime P to check that $\mathrm{div}\,(J) + \mathrm{div}\,(I) = \mathrm{div}\,(g)$. After localization, if x generates the maximal ideal we have that $I = x^m R$, $g = x^{m+n}R$, where m, $n \in \mathbb{N}$, and, since localization commutes with formation of colon ideals, that $J = x^{m+n}R : x^n R$, which is $x^m R$. This is just what we need to show that the coefficients of P in $\mathrm{div}\,(I)$ and $\mathrm{div}\,(J)$ sum to the coefficient of P in $\mathrm{div}\,(g)$.

It remains only to show that every element of $\mathcal{C}\ell\,(R)$ is represented by $\mathrm{div}\,(I)$ for some ideal I. But this is clear, since the paragraph above shows that inverses of elements like $[P]$ are represented by divisors of ideals. $\square$

Remarks. A further related result is that a finitely generated torsion-free module M of torsion-free rank one over a Noetherian normal domain R is isomorphic with a pure height one ideal if and only if it is a *reflexive* R-module, i.e, if and only if the natural map $M \to M^{**}$ is an isomorphism, where $\_^*$ indicates $\mathrm{Hom}(\_, R)$, and the natural map sends $u \in M$ to the map $M^* \to R$ whose value on $f \in M^*$ is $f(u)$. In fact, a finitely generated torsion-free module of rank one over a Noetherian domain is always isomorphic to an ideal $I \neq 0$ of R, and if R is normal, I^{**} may be identified with the intersection of the primary components of I corresponding to height one minimal primes of I. (If there are no such minimal primes then I^{**} may be identified with R.) One can define the divisor class group of the Noetherian normal domain R to be the isomorphism classes of rank one reflexive R-modules with multiplication given by $[I][J] = [(I \otimes_R J)^{**}]$. See the Lecture Notes for March 29 and p. 1 for March 31 from Math 615, Winter 2004 for an analysis of the behavior of reflexive modules over a normal Noetherian domain and a proof that the rank one reflexive modules coincide, up to isomorphism, with the ideals of pure height one.

Our next objective is to construct the divisor class group in a different way, using Grothendieck groups. The second point of view gives information that is not readily available directly.

Let R be a Noetherian ring. Let $\mathcal{M}$ denote the set of modules

$$\{R^n/M : n \in \mathbb{N},\, M \subseteq R^n\}.$$

Every finitely generated R-module is isomorphic to one in $\mathcal{M}$, which is all that we really need about $\mathcal{M}$: we can also start with some other set of modules with this property without affecting the Grothendieck group, but we use this one for definiteness.

Consider the free abelian group with basis $\mathcal{M}$, and kill the subgroup generated by all elements of the form $M - M' - M''$ where

$$0 \to M' \to M \to M'' \to 0$$

is a short exact sequence of elements of $\mathcal{M}$. The quotient group is called the *Grothendieck group* $G_0(R)$ of R. It is an abelian group generated by the elements $[M]$, where $[M]$ denotes the image of $M \in \mathcal{M}$ in $G_0(R)$. Note that if $M' \cong M$ we have a short exact sequence

$$0 \to M' \to M \to 0 \to 0,$$

so that $[M] = [M'] + [0] = [M']$, i.e., isomorphic modules represent the same class in $G_0(R)$.

A map L from $\mathcal{M}$ to an abelian group $(A, +)$ is called *additive* if whenever

$$0 \to M' \to M \to M'' \to 0$$

is exact, then $L(M) = L(M') + L(M'')$. The map θ sending M to $[M] \in G_0(R)$ is additive, and is a universal additive map in the following sense: given any additive map $L : \mathcal{M} \to A$, there is a unique homomorphism $h : G_0(M) \to A$ such that $L = h \circ \theta$. Since we need $L(M) = h([M])$, if there is such a map it must be induced by the map from the free abelian group with basis $\mathcal{M}$ to A that sends M to $h(M)$. Since h is additive, the elements $M - M' - M''$ coming from short exact sequences

$$0 \to M' \to M \to M'' \to 0$$

are killed, and so there is an induced map $h : G_0(R) \to A$. This is obviously the only possible choice for h.

Over a field K, every finitely generated module is isomorphic with $K^{\oplus n}$ for some $n \in \mathbb{N}$. It follows that $G_0(K)$ is generated by $\gamma = [K]$, and in fact it is $\mathbb{Z}\gamma$, the free abelian group on one generator. The additive map associated with the Grothendieck group sends M to $\dim_K(M)\gamma$. If we identify $\mathbb{Z}\gamma$ with $\mathbb{Z}$ by sending $\gamma \mapsto 1$, this is the (K-vector space) dimension map.

If R is a domain with fraction field $\mathcal{F}$, we have an additive map to $\mathbb{Z}$ that sends M to $\dim_{\mathcal{F}} \mathcal{F} \otimes_R M$, which is called the *torsion-free rank* of M. This induces a surjective map

$G_0(R) \to \mathbb{Z}$. If R is a domain and *if* $\gamma = [R]$ generates $G_0(R)$, then $G_0(R) \cong \mathbb{Z}\gamma \cong \mathbb{Z}$, with the isomorphism given by the torsion-free rank map.

Notice that if L is additive and

$$0 \to M_n \to \cdots \to M_1 \to M_0 \to 0$$

is exact, then

$$L(M_0) - L(M_1) + \cdots + (-1)^n L(M_n) = 0.$$

If $n \leq 2$, this follows from the definition. We use induction. In the general case note that we have a short exact sequence

$$0 \to N \to M_1 \to M_0 \to 0$$

and an exact sequence

$$0 \to M_n \to \cdots \to M_3 \to M_2 \to N \to 0,$$

since

$$\operatorname{Coker}(M_3 \to M_2) \cong \operatorname{Ker}(M_1 \to M_0) = N.$$

Then

$$(*) \quad L(M_0) - L(M_1) + L(N) = 0,$$

and

$$(**) \quad L(N) - L(M_2) + \cdots + (-1)^{n-1} L(M_n) = 0$$

by the induction hypothesis. Subtracting $(**)$ from $(*)$ yields the result. $\square$

Our proof of unique factorization in arbitrary regular local rings is based on the following two theorems, whose proofs we postpone momentarily.

To state the first of these theorems, observe that we can define a filtration of $G_0(R)$ by letting $\langle G_0(R) \rangle_i$ denote the subgroup spanned by classes of primes P such that height $(P) \geq i$. This filtration decreases as i increases. From it, we obtain an associated graded group: we write

$$[G_0(R)]_i = \langle G_0(R) \rangle_i / \langle G_0(R) \rangle_{i+1}.$$

Theorem (M. P. Murthy). *If R is a normal domain, then $\mathcal{C}\ell(R) \cong [G_0(R)]_1$ in such a way that the generator of $\mathcal{C}\ell(R)$ corresponding to a height one prime P is mapped to the image of R/P in $[G_0(R)]_1$.*

The second of these theorems is the following, which is a local version of the Hilbert syzygy theorem.

Theorem (Hilbert syzygy theorem for regular local rings). *Let (R, m, K) be a regular local ring of Krull dimension n, and let M be a finitely generated R-module. Then M is free if and only if* $\operatorname{depth}(M) = n$. *If M is not free and M_1 is any first module of syzygies of M,* $\operatorname{depth}(M_1) = \operatorname{depth}(M) + 1$. *Hence, M has a finite free resolution by finitely generated free modules, and any shortest such free resolution of M has length* $n - \operatorname{depth}(M)$.

Once we have proved this, we have:

Corollary. *If R is a regular local ring, $G_0(R) = \mathbb{Z}\gamma$, where $\gamma = [R]$, and so for every finitely generated module M, $[M] \in G_0(R)$ is* rank $(M)\gamma$, *where rank indicates torsion-free rank. In particular, if M is a torsion-module, $[M] = 0$, and so $[R/P] = 0$ for every prime ideal P with height $P \geq 1$.*

Proof. R is a domain, and we have the additive map given by torsion-free rank. It will suffice to show that $[R]$ generates $G_0(R)$. But if M is any finitely generated R-module, we know that M has a finite free resolution

$$0 \to R^{b_k} \to \cdots \to R^{b_1} \to R^{b_0} \to M \to 0,$$

and so the element $[M]$ may be expressed as

$$[R^{b_0}] - [R^{b_1}] + \cdots + (-1)^k [R^{b_k}] = b_0\gamma - b_1\gamma + \cdots + (-1)^k b_k\gamma = (b_0 - b_1 + \cdots + (-1)^k b_k)\gamma$$

□

Hence:

Corollary (Auslander-Buchsbaum). *Every regular local ring is a UFD.*

Proof. (M. P. Murthy) The universal additive map is the same as torsion-free rank, so that if $P \neq (0)$, we have that $[R/P] = 0$ in $G_0(R)$. It follows that $\langle G_0(R)\rangle_i = 0$ for all $i \geq 1$, and, hence, $\mathcal{C}\ell\,(R) = [G_0(R)]_1 = 0$. □

It remains to prove the local version of the Hilbert syzygy theorem and Murthy's characterization of the divisor class group.

Lecture of April 8

Note that given a finite filtration

$$0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_{n-1} \subseteq M_n = M$$

of a finitely generated R-module M and an additive map L we have that

$$L(M) = L(M_n/M_{n-1}) + L(M_{n-1}),$$

and, by induction on n, that

$$L(M) = \sum_{j=1}^{n} L(M_j/M_{j-1}).$$

In particular, $[M] \in G_0(R)$ is

$$\sum_{j=1}^{n} [M_j/M_{j-1}].$$

The following result gives a presentation of the Grothendieck group.

Theorem. *Let R be a Noetherian ring. $G_0(R)$ is generated by the elements $[R/P]$, as P runs through all prime ideals of R. If P is prime and $x \in R - P$, then $[R/(P + xR)] = 0$, and so if $R/Q_1, \ldots, R/Q_k$ are all the factors in a prime filtration of $[R/(P + xR)]$, we have that $[R/Q_1] + \cdots + [R/Q_k] = 0$. The relations of this type are sufficient to generate all relations on the classes of the prime cyclic modules.*

Proof. The first statement follows from the fact that every finitely generated module over a Noetherian ring R has a finite filtration in which the factors are prime cyclic modules. The fact that $[R/(P + xR)] = 0$ follows from the short exact sequence

$$0 \to R/P \xrightarrow{x} R/P \to R/(P + xR) \to 0,$$

which implies $[R/P] = [R/P] + [R/(P + xR)]$ and so $[R/(P + xR)] = 0$ follows.

Now, for every $M \in \mathcal{M}$, fix a prime cyclic filtration of M. We need to see that if we have a short exact sequence

$$0 \to M' \to M \to M'' \to 0$$

that the relation $[M] = [M'] + [M'']$ is deducible from ones of the specified type. We know that M' will be equal to the sum of the classes of the prime cyclic modules occurring in its chosen prime filtration, and so will M''. These two prime cyclic filtrations together induce a prime cyclic filtration $\mathcal{F}$ of M, so that the information $[M] = [M'] + [M'']$ is conveyed by setting $[M]$ equal to the sum of the classes of the prime cyclic modules in these specified filtrations of $[M]$ and $[M']$. But $\mathcal{F}$ will not typically be the specified filtration of $[M]$, and so we need to set the sum of the prime cyclic modules in the specified filtration of M equal to the sum of all those occurring in the specified filtrations of M' and M''.

Thus, we get sufficiently many relations to span all relations if for all finitely generated modules M and for all pairs of possibly distinct prime cyclic filtrations of M, we set the

sum of the classes of the prime cyclic modules coming from one filtration equal to the corresponding sum for the other. But any two filtrations have a common refinement. Take a common refinement, and refine it further until it is a prime cyclic filtration again. Thus, we get sufficiently many relations to span if for every finitely generated module M and for every pair consisting of a prime cyclic filtration of M and a refinement of it, we set the sum of the classes coming from one filtration to the sum of those in the other. Any two prime cyclic filtrations may then be compared by comparing each to a prime cyclic filtration that refines them both.

In refining a given prime cyclic filtration, each factor R/P is refined. Therefore, we get sufficiently many relations to span if for every R/P and every prime cyclic filtration of R/P, we set $[R/P]$ equal to the sum of the classes in the prime cyclic filtration of R/P. Since $\mathrm{Ass}\,(R/P) = P$, the first submodule of a prime cyclic filtration of R/P will be isomorphic with R/P, and will therefore have the form $x(R/P)$, where $x \in R - P$. If the other factors are $R/Q_1, \ldots, R/Q_k$, then these are the factors of a filtration of $(R/P)/x(R/P) = R/(P + xR)$. Since $[x(R/P)] = [R/P]$, the relation we get is

$$[R/P] = [R/P] + [R/Q_1] + \cdots + [R/Q_k],$$

which is equivalent to

$$[R/Q_1] + \cdots + [R/Q_k] = 0,$$

and so the specified relations suffice to span all relations. □

We can immediately deduce as a consequence the theorem of Murthy stated in the Lecture Notes of April 6.

Theorem (M. P. Murthy). *If R is a normal domain, then $\mathcal{C}\ell\,(R) \cong [G_0(R)]_1$ in such a way that the generator of $\mathcal{C}\ell\,(R)$ corresponding to a height one prime P is mapped to the image of R/P in $[G_0(R)]_1$.*

Proof. We know that $G_0(R)$ is the free group on the classes of the R/P, P prime, modulo relations obtained from prime cyclic filtrations of $R/(P + xR)$, $x \notin P$. We shall show that if we kill all the $[R/Q]$ for Q of height 2 or more, all relations are also killed except those coming from $P = (0)$, and the image of any relation corresponding to a prime cyclic filtration of R/xR corresponds precisely to $\mathrm{div}\,(x)$. Clearly, if $P \neq 0$ and $x \notin P$, any prime containing $P + xR$ strictly contains P and so has height two or more. Thus, we need only consider relations on the R/P for P of height one coming from prime cyclic filtrations of R/xR, $x \neq 0$. Clearly, R does not occur, since R/xR is a torsion module, and occurrences of R/Q for Q of height ≥ 2 do not matter. We need only show that for every prime P of height one, the number of occurrences of R/P in any prime cyclic filtration of R/xR is exactly k, where $P^{(k)}$ is the P-primary component of xR. But we can do this calculation after localizing at P: note that all factors corresponding to other primes become 0, since some element in the other prime not in P is inverted. Then $xR_P = P^k R_p$, and we need to show that any prime cyclic filtration of R_P/xR_P has k copies of R_P/PR_P, where we know

that $xR_P = P^k R_P$. Notice that $(R_P,\, PR_P)$ is a DVR, say (V, tV), and $xR_P = t^k V$. The number of nonzero factors in any prime cyclic filtration of V/t^kV is the length of V/t^kV over V, which is k, as required: the only prime cyclic filtration without repetitions is

$$0 \subset t^{k-1}V \subset t^{k-2}V \subset \cdots \subset t^2V \subset tV \subset V. \qquad \square$$

We next restate and then prove the local form of the Hilbert syzygy theorem stated in the Lecture Notes of April 6. The result is entirely analogous to the corresponding result in the graded case treated in the second problem of Problem Set #3.

Theorem (Hilbert syzygy theorem for regular local rings). *Let $(R,\, m,\, K)$ be a regular local ring of Krull dimension n, and let M be a finitely generated nonzero R-module. Then M is free if and only if* $\operatorname{depth}(M) = n$. *If M is not free and M_1 is any first module of syzygies of M,* $\operatorname{depth}(M_1) = \operatorname{depth}(M) + 1$. *Hence, M has a finite free resolution by finitely generated free modules, and any shortest such free resolution of M has length* $n - \operatorname{depth}(M)$.

Proof. For the first statement we use induction on $\dim(R)$. If $n = 0$ then R is a field, every module is free, and there is nothing to prove. Assume that $n > 0$. It is clear that if M is a nonzero free module then its depth is n. Suppose that M has depth n. In particular, $\operatorname{depth}(M) \geq 1$ and we can choose $x \in m$ not in m^2 nor in any minimal prime of M. Then M/xM has depth $n-1$ over R/xR, which is again regular. Thus, M/xM is free by the induction hypothesis: let $u_1, \ldots, u_h$ be elements of M whose images are a free basis for M/xM. These elements span M by Nakayama's Lemma. To complete the proof of this part, it suffices to show that they have no nonzero relation over R. Let N denote the module of all relations on $u_1, \ldots, u_h$ over R. If $(f_1, \ldots, f_h) \in N$ is a relation, so that $f_1u_1 + \cdots f_hu_h = 0$, then we may consider this relation modulo xR. Since the images of the u_j are a free basis for M/xM, it follows that every f_j is in xR, and can be written xg_j for some $g_j \in R$. Then $x(g_1u_1 + \cdots g_hu_h) = 0$, and since x is not a zerodivisor on M, we have that $g_1u_1 + \cdots g_hu_h = 0$. Thus $(f_1, \ldots, f_h) = x(g_1, \ldots, g_h)$ with $(g_1, \ldots, g_h) \in N$, and we consequently have that $N = xN$. By Nakayama's Lemma, $N = 0$, and it follows that M is free on the basis $u_1, \ldots, u_h$.

The remaining statements now follow from part (a) of the second problem of Problem Set #2 exactly as in the graded case. $\square$

We have now completed the proof of unique factorization in regular local rings, following M. P. Murthy.

We want to note another proof of a variant of the Hilbert syzygy theorem for finitely generated modules over polynomial rings, based on Gröbner basis ideas. The argument is based on Schreyer's method for computing modules of relations or syzygies, which is described beginning near the bottom of p. 2 of the Lecture of January 25, and continuing on pp. 3, 4, and 5. We review the method, which is very simple.

Let $R = K[x_1, \ldots, x_n]$ denote the polynomial ring in n variables over a field K, and let $M \subseteq F$ be a submodule of F, where F is free with ordered basis $b_1, \ldots, b_s$. Let $g_1, \ldots, g_r$ be a Gröbner basis for M. (We shall momentarily impose a mild condition on the ordering of the g_i.) We may view the relations on $g_1, \ldots, g_r$ as a submodule R^r, for which we denote the standard basis as $e_1, \ldots, e_r$. Schreyer's method asserts that the module of all relations on $g_1, \ldots, g_r$ is generated by certain standard relations as follows. Suppose that $i < j$ and that $\mathrm{in}(g_i) = \mu_i b_k$, $\mathrm{in}(g_j) = \mu_j b_k$ involve the same element b_k of $b_1, \ldots, b_s$. Then we can write

$$(*_{ij}) \quad \frac{\mu_j}{\mathrm{GCD}(\mu_i, \mu_j)} g_i - \frac{\mu_i}{\mathrm{GCD}(\mu_i, \mu_j)} g_j = \sum_{t=1}^{r} q_{ijt} g_t$$

where the left hand side is a standard expression for division of the left hand side by $g_1, \ldots, g_r$. The remainder is 0 in each case by the Buchberger criterion. The displayed equation implies that

$$(\#_{ij}) \quad \frac{\mu_j}{\mathrm{GCD}(\mu_i, \mu_j)} e_i - \frac{\mu_i}{\mathrm{GCD}(\mu_i, \mu_j)} e_j - \sum_{t=1}^{r} q_{ijt} e_t$$

is a relation on $g_1, \ldots, g_r$. This is a typical standard relation, and we saw that these not only generate the module of all relations, but are, in fact, a Gröbner basis for it with respect to a suitable monomial order on R^r. Moreover, the initial term of $(\#_{ij})$ is

$$(\dagger_{ij}) \quad \frac{\mu_j}{\mathrm{GCD}(\mu_i, \mu_j)} e_i.$$

We now make an almost trivial observation:

Lemma. *Let hypotheses and notations be as above and suppose that $g_1, \ldots, g_r$ have been ordered so that if $i > g$ and $\mathrm{in}(g_i) = \mu_i b_k$ and $\mathrm{in}(g_j) = \mu_j b_k$ involve the same element b_k of the ordered basis $b_1, \ldots, b_s$ for F then $\mu_i > \mu_j$ in lexicographic order on the monomials of R. (This does not depend on what the monomial order on F is: one can always order the g_i so that this condition is satsified.) Suppose that the initial terms of the g_i involve only the x_i for $i \geq h$. Then the initial terms of the standard relations on $g_1, \ldots, g_r$ involve only the variables x_i for $i \geq h + 1$.*

Proof. Since only the variables $x_h, \ldots, x_n$ occur and $\mu_i > \mu_j$ in lexicographic order, we must have that the highest power of x_h occurring in μ_i is at least that occurring in μ_j: call the latter x_h^a. It follows that x_h^a is also the highest power of x_h occurring in $\mathrm{GCD}(\mu_i, \mu_j)$, and so x_h does not occur in the initial term shown in $(\dagger_{ij})$ of the standard relation $(\#_{ij})$. □

Given any finitely generated module M over R its first module of syzygies M_1 is a submodule of a free module. Even if all the variables occur in generators of the initial module for M_1, after at most n repetitions of Schreyer's method, each time with the

Gröbner basis obtained ordered as indicated in the Lemma above, one obtains a Gröbner basis for the module of syzygies such that every initial term is simply one of the e_j. We can now complete our variant proof of the HIlbert syzygy theorem by showing that a module with a Gröbner basis of this form is free.

Lecture of April 11

We next note the following fact:

Proposition. *Let R be any ring and $F = R^n$ a free module. If $f_1, \ldots, f_n \in F$ generate F, then $f_1, \ldots, f_n$ is a free basis for F.*

Proof. We have a surjection $R^n \twoheadrightarrow F$ that maps $e_i \in R^n$ to f_i. Call the kernel N. Since F is free, the map splits, and we have $R^n \cong F \oplus N$. Then N is a homomorphic image of R^n, and so is finitely generated. If $N \neq 0$, we may preserve this while localizing at a suitable maximal ideal m of R. We may therefore assume that (R, m, K) is quasilocal. Now apply $K \otimes_R \_$. We find that $K^n \cong K^n \oplus N/mN$. Thus, $N = mN$, and so $N = 0$. $\square$

The final step in our variant proof of the Hilbert syzygy theorem is the following:

Lemma. *Let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over a field K, let F be a free R-module with ordered free basis $e_1, \ldots, e_s$, and fix any monomial order on F. Let $M \subseteq F$ be such that $\mathrm{in}(M)$ is generated by a subset of $e_1, \ldots, e_s$, i.e., such that M has a Gröbner basis whose initial terms are a subset of $e_1, \ldots, e_s$. Then M and F/M are R-free.*

Proof. Let S be the subset of $e_1, \ldots, e_s$ generating $\mathrm{in}(M)$, and suppose that S has r elements. Let $T = \{e_1, \ldots, e_s\} - S$, which has $s - r$ elements. Let $G \cong R^{s-r}$ be the free submodule of F spanned by T. By the Theorem on the bottom of p. 2 of the Lecture Notes of January 13, the images of the monomials not in $\mathrm{in}(M)$ are a K-vector space basis for F/M. These monomials, which are simply those involving an element of T, are obviously also a K-vector space basis for G. It follows that the composite R-linear map $G \subseteq F \twoheadrightarrow F/M$ is an isomorphism of K-vector spaces. Since it is R-linear, it is also an isomorphism of R-modules. It is clear that $M + G = F$, since $\mathrm{in}(M) \cup \mathrm{in}(G) = S \cup T = \mathrm{in}(F)$. Since no element of $G - \{0\}$ is killed in F/M, the sum is direct, i.e., $F = M \oplus G$. Let $g_1, \ldots, g_r$ be elements of a Gröbner basis for M whose initial terms are the elements of S. Then $g_1, \ldots, g_r$ together with T are s elements that generate $F \cong R^s$, and so they form a free basis for R^s by the preceding Proposition. It follows that $g_1, \ldots, g_r$ is a free basis for M. $\square$

We have now proved a "global" result on modules of syzygies over a polynomial ring: every finitely generated module has an nth module of syzygies that is free. It follows that every finitely generated module has a finite resolution by finitely generated free modules. This means in turn that if $R = K[x_1, \ldots, x_n]$, a polynomial ring over a field,

then $G_0(R) = \mathbb{Z}\gamma$, where $\gamma = [R]$, and the universal additive map is given by torsion-free rank. It follows just as in the local case that $[G_0(R)]_1 = 0$, i.e., that $\mathcal{C}\ell\,(R) = 0$, which gives a new proof that a polynomial ring over a field is a UFD, quite different from the usual one.

We next want to discuss projective modules over a Noetherian ring. A module P over R is *projective* if for every surjective map $M \twoheadrightarrow N$ the map $\mathrm{Hom}_R(P, M) \to \mathrm{Hom}_R(P, N)$ is surjective. It follows at once that $\mathrm{Hom}_R(P, \_)$ is a (covariant) exact functor from R-modules to R-modules. It is easy to see that free modules are projective: to lift a map $f : F \to N$ when F has free basis $\mathcal{B}$, for every $b \in \mathcal{B}$ one chooses $u_b \in M$ such that $u_b \mapsto f(b) \in N$, and one may then define $g : F \to M$ such that $g(b) = u_b$ for all $b \in \mathcal{B}$. The direct sum of two modules is projective if and only if both are projective: this follows from the fact that $\mathrm{Hom}_R(P \oplus Q, M)$ may be identified, functorially in M, with $\mathrm{Hom}_R(P, M) \oplus \mathrm{Hom}_R(Q, M)$. Hence, a direct summand of a free module is projective. Note that since free modules are flat, and since a direct sum of two modules is flat if and only if both are (because of the identification, functorial in M, of

$$(P \oplus Q) \otimes_R M \cong (P \otimes_R M) \oplus (Q \otimes_R M), \ )$$

it follows that projective modules are flat.

We have the following in great generality:

Proposition. *Let R be any ring and P an R-module. The following conditions are equivalent:*

(1) *P is projective.*

(2) *Every surjective map $f : M \twoheadrightarrow P$ splits, where M is an aribtrary R-module.*

(3) *P is a direct summand of a free module.*

Moreover, if P is finitely generated, then P is projective if and only if it is a direct summand of a finitely generated free module.

Proof. We have seen in the paragraph above that (3) $\Rightarrow$ (1). If P is projective and we have $f : M \twoheadrightarrow P$, the identity map $\mathbf{1}_P : P \to P$ lifts to a map $g : P \to M$: this means that $f \circ g = \mathbf{1}_P$, so that g is the required splitting. Finally, (2) $\Rightarrow$ (3) because if P satisfies (2) and we map a free module $F \twoheadrightarrow P$, the map splits, and so P is a direct summand of F. If P is finitely generated, we may take F to be finitely generated. $\square$

If P is a finitely generated projective module, we know that there exists a projective module Q such that $P \oplus Q$ is free. Q is called a *complement* for P. Q itself need not be free. If there exists a free module G such that $P \oplus G$ is a finitely generated free module, G is called a *free complement* for P.

Proposition. *Let R be any ring and let P be a projective module that has a finite resolution by finitely generated free modules. Then P has a free complement.*

Proof. We use induction on the length of the free resolution. If the resolution is

$$0 \to F_1 \to F_0 \to P \to 0$$

then the map $F_0 \twoheadrightarrow P$ splits, and $F_0 \cong P \oplus F_1$. Now suppose that the resolution has length $k > 1$. Let $Q = \mathrm{Ker}\,(F_0 \twoheadrightarrow P)$. Then $F_0 \cong P \oplus Q$, so that Q is projective, and Q has a free resolution of length at most $k-1$. By the induction hypothesis, we can choose a finitely generated free module G such $Q \oplus G = H$ is a finitely generated free module. Then $P \oplus (Q \oplus G) = F_0 \oplus G$ is free, and since $Q \oplus G = H$, we have that H is a free complement for P. $\square$

Hence, given our results for polynomial rings, we have an easy proof of the following:

Theorem. *Let R be a polynomial ring $K[x_1, \ldots, x_n]$ over a field K. Then every finitely generated projective R-module has a free complement.* $\square$

In the mid 1950s Serre asked whether every finitely generated projective module over a polynomial ring is free. This was not answered until 1976, when D. Quillen and A. Suslin gave proofs indepedently. Another, simpler, proof was found soon thereafter by Vaserstein. One way of attacking the problem is as follows.

One wants to prove that when R is a polynomial ring over a field, if $P \oplus R^k$ is free, then P is free. It suffices to show this when $k = 1$. For then, since $P' \oplus R$ is free, with $P' = P \oplus R^{k-1}$, one can conclude that P' is free, and the result follows by induction on k. Thus, once one knows that every finitely generated projective module has a free complement, showing that every finitely generated projective module is free is equivalent to showing that if $P \oplus R = R^n$ then P is free.

Over any ring R, giving a projective module P such that $P \oplus R = R^n$ is equivalent to giving a surjective map $R^n \twoheadrightarrow R$. The kernel of this map, call it P, then has the property that $P \oplus R = R^n$, for the map $R^n \twoheadrightarrow R$ is split, and so the short exact sequence

$$0 \to P \to R^n \to R \to 0$$

is split. Giving a surjective map $R^n \to R$ is the same as giving a $1 \times n$ matrix $\begin{pmatrix} g_1 & \ldots, & g_n \end{pmatrix}$ whose entries generate the unit ideal of R. This determines P. Note that we have elements $f_1, \ldots, f_n \in R$ such that $f_1 g_1 + \cdots + f_n g_n = 1$, and the map of $R \to R^n$ sending $1 \mapsto (f_1, \ldots, f_n)$ gives the splitting.

The projective module P, if free, must have rank $n-1$. In fact, P is free if and only if it is generated by $n-1$ elements $\rho_2, \ldots, \rho_n$. For in this case, these elements together with $\rho_1 = (f_1, \ldots, f_n)$ generate R^n, and so we have a surjetion $R^n \to R^n$ which is, necessarily, a bijection. If we make the ρ_i into the rows of a matrix, the condition that the rows generate R^n (equivalently, that the rows be a free basis for R^n) is that the matrix be invertible.

A row whose entries generate the unit ideal is called a *unimodular* row. The *unimodular row problem* asks the following: given a unimodular row over a ring R, can it be completed

to a square matrix whose determinant is 1? This question has an affirmative answer over R for all size rows if and only if for every projective module P over R such that $P \oplus R$ is free of finite rank, P is free. If every finitely generated projective R-module has a free complement, an affirmative answer to the unimodular row problem implies that every finitely generated projective R-module is free.

As mentioned above, this is the case for polynomial rings in finitely many variables over a field. We want to give one example where there is a projective module with free complement of rank one that is not free. Let $R = \mathbb{R}[X, Y, Z]/(X^2+Y^2+Z^2) = \mathbb{R}[x, y, z]$, which may be thought of as the coordinate ring of the real 2-sphere S^2. Elements of R may be thought of as real-valued continuous functions on S^2. Note that $\begin{pmatrix} x & y & z \end{pmatrix}$ is a unimodular row, since $x^2 + y^2 + z^2 = 1$ in R. This row cannot be completed to a 3×3 matrix whose determinant is 1 if the entries of the matrix are in R, nor even if the entries are allowed to be arbitrary continuous real-valued functions on S^2. It follows that

$$P = \operatorname{Ker}(R^3 \twoheadrightarrow R),$$

where the map has matrix $\begin{pmatrix} x & y & z \end{pmatrix}$, is a projective module over R that is not free but such that $P \oplus R = R^3$. To show that we cannot complete the matrix, suppose that we can, and let the second row be $\begin{pmatrix} f & g & h \end{pmatrix}$ where f, g, h are continuous real-valued functions on S_2. Since the determinant of the matrix is constantly equal to 1, for every point $(a, b, c) \in S^2$, if we substitute a, b, c for the variables the first two rows of the matrix are linearly independent. Thus, (a, b, c) and

$$w(a, b, c) = \big(f(a, b, c), g(a, b, c), h(a, b, c)\big)$$

are linearly independent, and since the vector (a, b, c) is normal to the tangent plane to the sphere at the point (a, b, c), the vector $w(a, b, c)$ has a nonzero projection $v(a, b, c)$ on the tangent plane to S^2 at (a, b, c) that varies continuously with (a, b, c). This constructs a continuous nonzero vector field on S^2, which contradicts a well-known theorem in toplogy ("you can't comb the hair on a billiard ball").

After tensoring with the complex numbers, one can complete the row. Working now over $\mathbb{C}[X, Y, Z]/(X^2 + Y^2 + Z^2 - 1)$, we see that the matrix

$$\begin{pmatrix} x & xi + y & z \\ 0 & -z & -xi + y \\ 1 & 0 & 0 \end{pmatrix}$$

has determinant 1: expand with respect to the third row. If we subtract i times the first column from the second, we get the matrix we want:

$$\begin{pmatrix} x & y & z \\ 0 & -z & -xi + y \\ 1 & -i & 0 \end{pmatrix}.$$

The flatness of the Frobenius endomorphism for regular rings

We shall return to the subject of projective modules, but we first want to establish the assertion made earlier that the Frobenius endomorphism is flat for every regular Noetherian ring of prime characteristic $p > 0$. To do so, we want to reduce to the case where the ring is complete local. We first observe the following:

Proposition. *Let $\theta : (R, m, K) \to (S, \mathfrak{n}, L)$ be a homomorphism of local rings that is* local*, i.e., $\theta(m) \subseteq \mathfrak{n}$. Then S is flat over R if and only if for every injective map $N \hookrightarrow M$ of finite length R-modules, $S \otimes_R N \hookrightarrow S \otimes_R M$ is injective.*

Proof. The condition is obviously necessary. We shall show that it is sufficient. Since tensor commutes with direct limits and every injection $N \hookrightarrow M$ is a direct limit of injections of finitely generated R-modules, it suffices to consider the case where $N \subseteq M$ are finitely generated. Suppose that some $u \in S \otimes_R N$ is such that $u \mapsto 0$ in $S \otimes_R M$. It will suffice to show that there is also such an example in which M and N have finite length. Fix any integer $t > 0$. Then we have an injection

$$N/(m^t M \cap N) \hookrightarrow M/m^t M$$

and there is a commutative diagram

$$\begin{array}{ccc} S \otimes_R N & \xrightarrow{\iota} & S \otimes_R M \\ {\scriptstyle f}\big\downarrow & & {\scriptstyle g}\big\downarrow \\ S \otimes_R \big(N/(m^t M \cap N)\big) & \xrightarrow{\iota'} & S \otimes_R \big(M/m^t M\big) \end{array} .$$

The image $f(u)$ of u in $S \otimes_R \big(\big(N/(m^t M \cap N)\big)$ maps to 0 under ι', by the commutativity of the diagram. Therefore, we have the required example provided that $f(u) \neq 0$. However, for all $h > 0$, we have from the Artin-Rees Lemma that for every sufficiently large integer t, $m^t M \cap N \subseteq m^h N$. Hence, the proof will be complete provided that we can show that the image of u is nonzero in

$$S \otimes_R (N/m^h N) \cong S \otimes_R \big((R/m^h) \otimes_R N\big) \cong (R/m^h) \otimes_R (S \otimes_R N) \cong (S \otimes_R N)/m^h (S \otimes_R N).$$

But

$$m^h (S \otimes_R N) \subseteq \mathfrak{n}^h (S \otimes_R N),$$

and the result follows from the fact that the finitely generated S-module $S \otimes_R N$ is $\mathfrak{n}$-adically separated. $\square$

Lemma. *Let $(R, m, K) \to (S, \mathfrak{n}, L)$ be a local homomorphism of local rings. Then S is flat over R if and only if $\widehat{S}$ is flat over $\widehat{R}$, and this hold iff $\widehat{S}$ is flat over R.*

Proof. If S is flat over R then, since $\widehat{S}$ is flat over S, we have that $\widehat{S}$ is flat over R. Conversely, if $\widehat{S}$ is flat over R, then S is flat over R because $\widehat{S}$ is faithfully flat over S: if $N \subseteq M$ is flat but $S \otimes_R N \to S \otimes_R M$ has a nonzero kernel, the kernel remains nonzero when we apply $\widehat{S} \otimes_S \_$, and this has the same effect as applying $\widehat{S} \otimes_R \_$ to $N \subseteq M$, a contradiction.

We have shown that $R \to S$ is flat if and only $R \to \widehat{S}$ is flat. If $\widehat{R} \to \widehat{S}$ is flat then since $R \to \widehat{R}$ is flat, we have that $R \to \widehat{S}$ is flat, and we are done. It remains only to show that if $R \to S$ is flat, then $\widehat{R} \to \widehat{S}$ is flat. By the Proposition, it suffices to show that if $N \subseteq M$ have finite length, then $\widehat{S} \otimes N \to \widehat{S} \otimes M$ is injective. Suppose that both modules are killed by m^t. Since S/m^tS is flat over R/m^t, if Q is either M or N we have that

$$\widehat{S} \otimes_{\widehat{R}} Q \cong \widehat{S}/m^t\widehat{S} \otimes_{\widehat{R}/m^t\widehat{R}} Q \cong \widehat{S}/m^t\widehat{S} \otimes_{R/m^t} Q \cong \widehat{S} \otimes_R Q,$$

and the result now follow because $\widehat{S}$ is flat over R. $\square$

We are now ready to prove:

Theorem. *Let R be a regular Noetherian ring of prime characteristic $p > 0$. Then the Frobenius endomorphism $F : R \to R$ is flat.*

Proof. To distinguish the two copies of R, we let S denote the right hand copy, so that $F : R \to S$. The issue of flatness is local on R, and if P is prime, then $(R - P)^{-1}S$ is the localization of S at the unique prime Q lying over P (if we remember that S is R, then Q is P), since the pth power of every element of $S - Q$ is in the image of $R - P$. Hence, there is no loss of generality in replacing R by R_P, and we henceforth assume that (R, m, K) is local. By the preceding Lemma, $F : R \to R$ is flat if and only if the induced map $\widehat{R} \to \widehat{R}$ is flat, and this map is easily checked to be the Frobenius endomorphism on $\widehat{R}$. We have now reduced to the case where R is a complete regular local ring. By the structure theory for complete local rings, we may assume without loss of generality that $R = K[[x_1, \ldots, x_n]]$ where K is a field of characteristic p. By the Theorem on p. 2 of the Lecture Notes of February 19, the Frobenius endomorphism $F : K[x_1, \ldots, x_n] \to K[x_1, \ldots, x_n]$ makes $K[x_1, \ldots, x_n]$ into a free algebra over itself. It follows that it is flat over itself, and this remains true when we localize at $(x_1, \ldots, x_n)$. By the preceding Lemma, we still have flatness after we complete both rings. Completing yields

$$F : K[[x_1, \ldots, x_n]] \to K[[x_1, \ldots, x_n]],$$

which proves the flatness result we need. $\square$

We can now give the application of this result that we have been intending for some time.

Theorem. *Let R be a regular Noetherian ring of prime characteristic $p > 0$. Then every ideal I of R is tightly closed.*

Proof. Suppose $u \in I^* - I$ and $c \in R$ is not in any minimal prime and satisfies $cu^q \in I^{[q]}$ for all $q \gg 0$. We may replace R by its localization at a maximal ideal in the support of $(I + Ru)/I$, I by its expansion to the local ring, and u by its image in the local ring. The image of c in this local ring is still not in any minimal prime, i.e., it is not 0. We still have that $u \in I^* - I$. Thus, we may assume without loss of generality that R is local. Then for some q_0,

$$c \in \bigcap_{q \geq q_0} I^{[q]} :_R u^q = \bigcap_{q \geq q_0} (I :_R u)^{[q]} \subseteq \bigcap_{q \geq q_0} m^{[q]} \subseteq \bigcap_{q \geq q_0} m^q = (0),$$

a contradiction. Note that the fact that $I^{[q]} :_R u^q = (I :_R u)^{[q]}$ used in this argument is a consequence of the flatness of the Frobenius endomorphism. $\square$

Projective modules over Noetherian rings

We now return to the subject of projective modules. For finitely generated projective modules over Noetherian rings there are some interesting characterizations.

Theorem. *Let P be a finitely presented module over a quasilocal ring (R, m, K) (in particular, it suffices if R is local and P is finitely generated). Then the following conditions are equivalent:*

(1) *P is free.*

(2) *P is projective.*

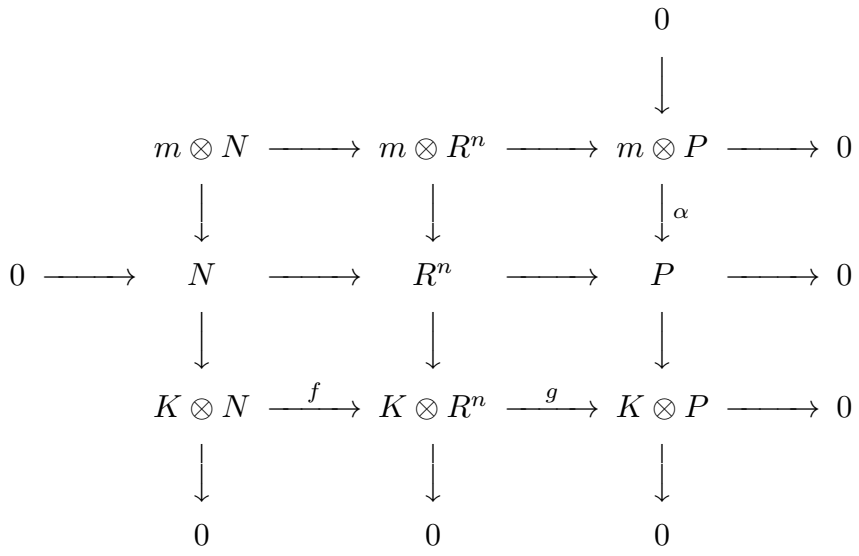
(3) *P is flat.*

(4) *The map $m \otimes_R P \to P$ sending $u \otimes v \mapsto uv$ is injective (and so gives an isomorphism $m \otimes_R P \cong mP$).*

Proof. The implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3), while (3) $\Rightarrow$ (4) follows by applying $\_ \otimes_R P$ to the injection $m \hookrightarrow R$. It remains to prove the difficult implication (4) $\Rightarrow$ (1).

Choose a minimal set of generators $u_1, \ldots, u_n$ for M and map R^n onto P such that $(r_1, \ldots, r_n)$ is sent to $r_1u_1 + \cdots + r_nu_n$. Let N be the kernel of the surjection $R^n \twoheadrightarrow P$, so that we have a short exact sequence $0 \to N \to R^n \to P \to 0$. We also have a short exact sequence $0 \to m \to R \to K \to 0$: think of this as written vertically with m at the top and K at the bottom. Then we may tensor the two sequences together to get the following

array (all tensor products are taken over R):



The rows are obtained by applying $m \otimes \_$, $R \otimes \_$, and $K \otimes \_$, respectively to the short exact sequence $0 \to N \to R^n \to P \to 0$, and the columns are obtained by applying $\_ \otimes N$, $\_ \otimes R^n$, and $\_ \otimes P$, respectively, to the short exact sequence $0 \to m \to R \to K \to 0$. The exactness of the rows and columns shown follows from the right exactness of tensor, with two exceptions: the injective arrow on the left in the middle row comes from the fact that tensoring with R simply yields the short exact sequence with which we started, while the injectivity of α is the hypothesis in (4). (We also have an injection at the top of the middle column because R^n is free, but we don't need this.)

It is easy to see that the four squares in the diagram commute.

The minimality of the set of generators $u_1, \ldots, u_n$ implies that g is an isomorphism of K^n with K^n, and the fact that M is finitely presented implies that N is finitely generated. To complete the proof it suffices to show that $K \otimes N = 0$, for then, by Nakayama's lemma, we have that $N = 0$. But if $N = 0$ then $R^n \to M$ is an isomorphism. To show that $K \otimes N$ is 0, it suffices to prove that the map f is injective.

Suppose that u is an element in the kernel of f. Choose $v \in N$ that maps to u. The image of v in R^n (we still call it v) maps to 0 in $K \otimes R^n$: we can go around the square on the lower left the other way, and u is killed by f. It follows that v is the image of an element w in $m \otimes R^n$. Suppose that w maps to x in $m \otimes M$. Then $\alpha(x) = 0$, because we can go around the square on the upper right the other way, and the image of v in M must be 0 because $v \in N$. But α is injective! Therefore, $x = 0$, which shows that w is the image of an element y in $m \otimes N$. Since w maps to v, y maps to v in N (the map $N \to R^n$ is injective), and this implies that v maps to 0 in $K \otimes N$. But v maps to u, and so $u = 0$. We are done: we have shown that f is injective! $\square$