MATH 614 LECTURE NOTES, FALL, 2017

by Mel Hochster

Lecture of September 6

We assume familiarity with the notions of ring, ideal, module, and with the polynomial ring in one or finitely many variables over a commutative ring, as well as with homomorphisms of rings and homomorphisms of R-modules over the ring R.

As a matter of notation, $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ are the non-negative integers, the integers, the rational numbers, the real numbers, and the complex numbers, respectively, throughout this course.

Unless otherwise specified, all rings are commutative, associative, and have a multiplicative identity 1 (when precision is needed we write 1_R for the identity in the ring R). It is possible that $1 = 0$, in which case the ring is $\{0\}$, since for every $r \in R$, $r = r \cdot 1 = r \cdot 0 = 0$. We shall assume that a homomorphism h of rings $R \to S$ preserves the identity, i.e., that $h(1_R) = 1_S$. We shall also assume that all given modules M over a ring R are *unital*, i.e., that $1_R \cdot m = m$ for all $m \in M$.

When R and S are rings we write $S = R[\theta_1, \ldots, \theta_n]$ to mean that S is generated as a ring over its subring R by the elements $\theta_1, \ldots, \theta_n$. This means that S contains R and the elements $\theta_1, \ldots, \theta_n$, and that no strictly smaller subring of S contains R and the $\theta_1, \ldots, \theta_n$. It also means that every element of S can be written (not necessarily uniquely) as an R-linear combination of the monomials $\theta_1^{k_1} \cdots \theta_n^{k_n}$. When one writes $S = R[x_1, \ldots, x_k]$ it often means that the x_i are indeterminates, so that S is the polynomial ring in k variables over R. But one should say this.

The main emphasis in this course will be on Noetherian rings, i.e., rings in which every ideal is finitely generated. Specifically, for all ideals $I \subseteq R$, there exist $f_1, \ldots, f_k \in R$ such that $I = (f_1, \ldots, f_k) = (f_1, \ldots, f_k)R = \sum_{i=1}^k Rf_i$. We shall develop a very useful theory of dimension in such rings. This will be discussed further quite soon. We shall not be focused on esoteric examples of rings. In fact, almost all of the theory we develop is of great interest and usefulness in studying the properties of polynomial rings over a field or the integers, and homomorphic images of such rings.

There is a strong connection between studying systems of equations, studying their solutions sets, which often have some kind of geometry associated with them, and studying commutative rings. Suppose the equations involve variables $X_1, \ldots, X_n$ with coefficients in K. The most important case for us will be when K is an algebraically closed field such as the complex numbers $\mathbb{C}$. Suppose the equations have the form $F_i = 0$ where the F_i are polynomials in the X_j with coefficients in K. Let I be the ideal generated by the F_i in the polynomial ring $K[X_1, \ldots, X_n]$ and let R be the quotient ring $K[X_1, \ldots, X_n]/I$. In R, the images x_j of the variables X_j give a solution of the equations, a sort of "universal" solution. The connection between commutative algebra and algebraic geometry is that algebraic properties of the ring R are reflected in geometric properties of the solution set, and conversely. Solutions of the equations in the field K give maximal ideals of R. This leads to the idea that maximal ideals of R should be thought of as points in a geometric object. Some rings have very few maximal ideals: in that case it is better to consider all of the prime ideals of R as points of a geometric object. We shall soon make this idea more formal.

Before we begin the systematic development of our subject, we shall look at some very simple examples of problems, many unsolved, that are quite natural and easy to state. Suppose that we are given polynomials f and g in $\mathbb{C}[x]$, the polynomial ring in one variable over the complex numbers $\mathbb{C}$. Is there an algorithm that enables us to tell whether f and g generate $\mathbb{C}[x]$ over $\mathbb{C}$? This will be the case if and only if $x \in \mathbb{C}[f, g]$, i.e., if and only if x can be expressed as a polynomial with complex coefficients in f and g. For example, suppose that $f = x^5 + x^3 - x^2 + 1$ and $g = x^{14} - x^7 + x^2 + 5$. Here it is easy to see that f and g do not generate, because neither has a term involving x with nonzero coefficient. But if we change f to $x^5 + x^3 - x^2 + x + 1$ the problem does not seem easy. The following theorem of Abhyankar and Moh, whose original proof was about 150 pages long, gives a method of attacking this sort of problem.

Theorem (Abhyankar-Moh). *Let f, g in $\mathbb{C}[x]$ have degrees d and e respectively. If $\mathbb{C}[f, g] = \mathbb{C}[x]$, then either $d \,|\, e$ or $e \,|\, d$, i.e., one of the two degrees must divide the other.*

Shorter proofs have since been given. Given this difficult result, it is clear that the specific f and g given above cannot generate $\mathbb{C}[x]$: 5 does not divide 14. Now suppose instead that $f = x^5 + x^3 - x^2 + x + 1$ and $g = x^{15} - x^7 + x^2 + 5$. With this choice, the Abhyankar-Moh result does not preclude the possibility that f and g generate $\mathbb{C}[x]$. To pursue the issue further, note that in $g - f^3$ the degree 15 terms cancel, producing a polynomial of smaller degree. But when we consider f and $g - f^3$, which generate the same ring as f and g, the larger degree has decreased while the smaller has stayed the same. Thus, the sum of the degrees has decreased. In this sense, we have a smaller problem. We can now see whether the Abhyankar-Moh criterion is satisfied for this smaller pair. If it is, and the smaller degree divides the larger, we can subtract off a multiple of a power of the smaller degree polynomial and get a new pair in which the larger degree has decreased and the smaller has stayed the same. Eventually, either the criterion fails, or we get a constant and a single polynomial of degree ≥ 2, or one of the polynomials has degree 1. In the first two cases the original pair of polynomials does not generate. In the last case, they do generate.

This is a perfectly general algorithm. To test whether f of degree d and g of degree $n \geq d$ are generators, check whether d divides n. If so and $n = dk$, one can choose a constant c such that $g - cf^k$ has degree smaller than n. If the leading coefficients of f and g are $a \neq 0$ and $b \neq 0$, take $c = b/a^k$. The sum of the degrees for the pair f, $g - cf^k$ has decreased.

Continue in the same way with the new pair, f, $g - cf^k$. If one eventually reaches a pair in which one of the polynomials is linear, the original pair were generators. Otherwise, one reaches either a pair in which neither degree divides the other, or else a pair in which one polynomial has degree ≥ 2 while the other is constant. In either of these cases, the two polynomials do not generate. The constant does not help, since we have all of $\mathbb{C}$ available anyway, and a polynomial g of degree $d \geq 2$ cannot generate: when g is substituted into a polynomial of degree n, call it F, $F(g)$ has a term of degree dn coming from g^n, and no other term occurring can cancel it. Thus, one cannot have $x = F(g)$.

One can work backwards from a pair in which one of the polynomials is linear to get all pairs of generators. For example, one gets pairs of generators

$$\begin{aligned}
&x,\, 0 \rightarrow \\
&x,\, 1 \rightarrow \\
&x+5,\, 1 \rightarrow \\
&x+5,\, (x+5)^7+1 \rightarrow \\
&\big((x+5)^7+1\big)^{11} + x + 5,\, (x+5)^7+1.
\end{aligned}$$

If one expands the last pair out, it is not very obvious from looking at the polynomials that they generate. Of course, applying the algorithm described above would enable one to see it.

This gives a reasonably appealing method for telling whether two polynomials in one variable generate $\mathbb{C}[x]$.

The step of going from the problem of when two polynomials generate to $\mathbb{C}[x]$ over $\mathbb{C}$ to when three polynomials generate turns out to be a giant one, however! While algorithms are known based on the theory of Gröbner bases, the process is much more complex. There are some elegant conjectures, but there is a lack of elegant theorems in higher dimension.

One might hope that given three polynomials that generate $\mathbb{C}[x]$, say f, g, and h, with degrees d, e, n, respectively, that it might be true that one of the degrees has to be a sum of non-negative integer multiples of the other two, e.g., $n = rd+se$. Then one could reduce to a smaller problem (i.e., one where the sum of the degrees is smaller) by subtracting a constant times $f^r g^s$ from h, while keeping the same f and g. But it is *not* true in the case of three polynomials that one of the degrees must be a sum of non-negative integer multiples of the other two. (See whether $f = x^5$, $g = x^4 + x$, and $h = x^3$ generate $\mathbb{C}[x]$.)

The problem of giving an elegant test for deciding when m polynomials generate the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$ in n variables over $\mathbb{C}$ seems formidable, but when $m = n$ there is at least a tantalizing conjecture.

In order to state it, we first want to point out that derivatives with respect to x can be defined for polynomials in x over any commutative ring R. One way is simply to decree that polynomials are to be differentiated term by term, and that the derivative of rx^n is nrx^{n-1}. A somewhat more conceptual method is to introduce an auxiliary variable h. If one wants to differentiate $F(x) \in R[x]$, one forms $F(x+h) - F(x)$. This is a polynomial in two variables, x and h, and all the terms that do not involve h as a factor cancel. Thus, one can write $F(x+h) - F(x) = hP(x,h)$ for a unique polynomial in two variables P. That is,

$$P(x,h) = \frac{F(x+h) - F(x)}{h}.$$

One then defines the derivative $\dfrac{dF}{dx}$ or $F'(x)$ to be $P(x,0)$, the result of substituting $h=0$ in $P(x,h)$. This is the algebraist's method of taking a limit as $h \to 0$: just substitute $h=0$.

Given a polynomial $F \in R[x_1, \ldots, x_n]$ we may likewise define its partial derivatives in the various x_i. E.g., to get $\dfrac{\partial F}{\partial x_n}$ we identify the polynomial ring with $S[x_n]$ where $S = R[x_1, \ldots, x_{n-1}]$. We can think of F as a polynomial in x_n only with coefficients in S, and $\dfrac{\partial F}{\partial x_n}$ is simply its derivative with respect to x_n when it is thought of this way.

The *Jacobian conjecture* asserts that $F_1, \ldots, F_n \in \mathbb{C}[x_1, \ldots, x_n]$ generate (note that the number of the F_i is equal to the number n of variables) if and only if the Jacobian determinant $\det\big(\partial F_i/\partial x_j\big)$ is identically a nonzero constant. This is true when $n=1$ and is known to be a necessary condition for the F_i to generate the polynomial ring. But even when $n=2$ it is an open question!

If you think you have a proof, have someone check it carefully — there are at least five published incorrect proofs in the literature, and new ones are circulated frequently.

It is known that if there is a counter-example one needs polynomials of degree at least 100. Such polynomials tend to have about 5,000 terms. It does not seem likely that it will be easy to give a counter-example.

Algebraic sets

The problems discussed above are very easy to state, and very hard. However, they are not close to the main theme in this course, which is dimension theory. We are going to assign a dimension, the *Krull dimension*, to every commutative ring. It may be infinite, but will turn out to be finite for rings that are finitely generated over a field or the integers.

In order to give some idea of where we are headed, we shall discuss the notion of a closed algebraic set in K^n, where K is a field. Everyone is welcome to think of the case where $K = \mathbb{C}$, although for the purpose of drawing pictures, it is easier to think about the case where $K = \mathbb{R}$.

Let K be a field. A polynomial in $K[x_1, \ldots, x_n]$ may be thought of as a function from $K^n \to K$. Given a finite set $f_1, \ldots, f_m$ of polynomials in $K[x_1, \ldots, x_n]$, the set of points where they vanish simultaneously is denoted $V(f_1, \ldots, f_m)$. Thus

$$V(f_1, \ldots, f_m) = \{(a_1, \ldots, a_n) \in K^n : f_i(a_1, \ldots, a_n) = 0, 1 \le i \le n\}.$$

If $X = V(f_1, \ldots, f_m)$, one also says that $f_1, \ldots, f_m$ *define* X.

Over $\mathbb{R}[x, y]$, $V(x^2+y^2-1)$ is a circle in the plane, while $V(xy)$ is the union of the coordinate axes. Note that $V(x,y)$ is just the origin.

A set of the form $V(f_1, \ldots, f_m)$ is called a *closed algebraic set* in K^n. We shall only be talking about closed algebraic sets here, and so we usually omit the word "closed."

For the moment let us restrict attention to the case where K is an algebraically closed field such as the complex numbers $\mathbb{C}$. We want to give algebraic sets a dimension in such a way that K^n has dimension n. Thus, the notion of dimension that we develop will generalize the notion of dimension of a vector space.

We shall do this by associating a ring with X, denoted $K[X]$: it is simply the set of functions defined on X that are obtained by restricting a polynomial function on K^n to X. The dimension of X will be the same as the dimension of the ring $K[X]$. Of course, we have not defined dimension for rings yet.

In order to illustrate the kind of theorem we are going to prove, consider the problem of describing the intersection of two planes in real three-space $\mathbb{R}^3$. The planes might be parallel, i.e., not meet at all. But if they do meet in at least one point, they must meet in a line.

More generally, if one has vector spaces V and W over a field K, both subspaces of some larger vector space, then $\dim(V\cap W) = \dim V + \dim W - \dim(V+W)$. If the ambient vector space has dimension n, this leads to the result that $\dim(V \cap W) \geq \dim V + \dim W - n$. In the case of planes in three-space, we see that that dimension of the intersection must be at least $2 + 2 - 3 = 1$.

Over an algebraically closed field, the same result turns out to be true for algebraic sets! Suppose that V and W are algebraic sets in K^n and that they meet in a point $x \in K^n$. We have to be a little bit careful because, unlike vector spaces, algebraic sets in general may be unions of finitely many smaller algebraic sets, which need not all have the same dimension. Algebraic sets which are not finite unions of strictly smaller algebraic sets are called *irreducible*. Each algebraic set is a finite union of irreducible ones in such a way that none can be omitted: these are called *irreducible components*. We define $\dim_x V$ to be the largest dimension of an irreducible component of V that contains x. One of our long term goals is to prove that for any algebraic sets V and W in K^n meeting in a point x, $\dim_x(V \cap W) \geq \dim_x V + \dim_x W - n$. This is a beautiful and useful result: it can be thought of as guaranteeing the existence of a solution (or many solutions) of a family of equations.

We conclude for now by mentioning one other sort of problem. Given a specific algebraic set $X = V(f_1, \ldots, f_m)$, the set J of all polynomials vanishing on it is closed under addition and multiplication by any polynomial — that is, it is an ideal of $K[x_1, \ldots, x_n]$. J always contains the ideal I generated by $f_1, \ldots, f_m$. But J may be strictly larger than I. How can one tell?

Here is one example of an open question of this sort. Consider the set of pairs of commuting square matrices of size n. Let $M = M_n(K)$ be the set of $n \times n$ matrices over K. Thus,

$$W = \{(A, B) \in M \times M : AB = BA\}.$$

The matrices are given by their $2n^2$ entries, and we may think of this set as a subset of K^{2n^2}. (To make this official, one would have to describe a way to string the entries of the two matrices out on a line.) Then W is an algebraic set defined by n^2 quadratic equations. If $X = (x_{ij})$ is an $n \times n$ matrix of indeterminates and $Y = (y_{ij})$ is another $n \times n$ matrix of indeterminates, then we may think of the algebraic set W as defined by the vanishing of the entries of the matrix $XY - YX$. These are the n^2 quadratic equations.

Is the ideal of all functions that vanish on W generated by the entries of $XY - YX$? This is a long standing open question. It is known if $n \leq 3$. So far as I know, the question remains open over all fields.

Lecture of September 8

The notes for this lecture contain some basic definitions concerning abstract topological spaces that were not given in class. If you are not familiar with this material please read it carefully. I am not planning to do it in lecture.

We mention one more very natural but very difficult question about algebraic sets. Suppose that one has an algebraic set $X = V(f_1, \ldots, f_m)$. What is the least number of elements needed to define X? In other words, what is the least positive integer k such that $X = V(g_1, \ldots, g_k)$?

Here is a completely specific example. Suppose that we work in the polynomial ring in 6 variables $x_1, \ldots, x_3, y_1, \ldots, y_3$ over the complex numbers $\mathbb{C}$ and let X be the algebraic set in $\mathbb{C}^6$ defined by the vanishing of the 2×2 subdeterminants or *minors* of the matrix

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix},$$

that is, $X = V(f, g, h)$ where $f = x_1y_2 - x_2y_1$, $g = x_1y_3 - x_3y_1$, and $h = x_2y_3 - x_3y_2$. We can think of points of X as representing 2×3 matrices whose rank is at most 1: the vanishing of these equations is precisely the condition for the two rows of the matrix to be linearly dependent. Obviously, X can be defined by 3 equations. Can it be defined by 2 equations? No algorithm is known for settling questions of this sort, and many are open, even for relatively small specific examples. In the example considered here, it turns out that 3 equations are needed. I do not know an elementary proof of this fact — perhaps you can find one!

One of the themes of this course is that there is geometry associated with any commutative ring R. The following discussion illustrates this.

For an algebraic set over an algebraically closed field K, the maximal ideals of the ring $K[X]$ (reminder: functions from X to K that are restrictions of polynomial functions) are in bijective correspondence with the points of X — the point x corresponds to the maximal ideal consisting of functions that vanish at x. This is, essentially, Hilbert's Nullstellensatz, and we shall prove this theorem soon. This maximal ideal may also be described as the kernel of the evaluation homomorphism from $K[X]$ onto K that sends f to $f(x)$.

If R is the ring of continuous real-valued functions on a compact (Hausdorff) topological space X the maximal ideals also correspond to the points of X.

A *filter* $\mathcal{F}$ on a set X is a non-empty family of subsets closed under finite intersection and such that if $Y \in \mathcal{F}$, and $Y \subseteq Y' \subseteq X$, then $Y' \in \mathcal{F}$. Let K be a field. Let S be the ring of all K-valued functions on X. The ideals of S correspond bijectively with the filters on X: given a filter, the corresponding ideal consists of all functions that vanish on some set in the filter. The filter is recovered from the ideal I as the family of sets of the form $F^{-1}(0)$ for some $f \in I$. The key point is that for f and $g_1, \ldots, g_k \in S$, f is in the ideal generated by the g_k if and only if it vanishes whenever all the g_i do. The unit ideal corresponds to the filter which is the set of all subsets of X. The maximal ideals correspond to the maximal filters that do not contain the empty set: these are called *ultrafilters*. Given a point of $x \in X$, there is an ultrafilter consisting of all sets that contain x. Ultrafilters of this type are called *fixed*. If X is infinite, there are always others: the sets with finite complement form a filter, and by Zorn's lemma it is contained in an ultrafilter. For those familiar with the Stone-Cech compactification, the ultrafilters (and, hence, the maximal ideals) correspond bijectively with the points of the Stone-Cech compactification of X when X is given the discrete topology (every set is open).

We shall see that even for a completely arbitrary commutative ring R, the set of all maximal ideals of R, and even the set of all prime ideals of R, has a geometric structure. In fact, these sets have, in a natural way, the structure of topological spaces. We shall give a brief review of the notions needed from topology shortly.

Categories

We do not want to dwell too much on set-theoretic issues but they arise naturally here. We shall allow a class of all sets. Typically, classes are very large and are not allowed to be elements. The objects of a category are allowed to be a class, but morphisms between two objects are required to be a set.

A category $\mathcal{C}$ consists of a class $\mathrm{Ob}\,(\mathcal{C})$ called the *objects* of $\mathcal{C}$ and, for each pair of objects $X, Y \in \mathrm{Ob}\,(\mathcal{C})$ a set $\mathrm{Mor}\,(X, Y)$ called the *morphisms* from X to Y with the following additional structure: for any three given objects X, Y and Z there is a map

$$\mathrm{Mor}\,(X, Y) \times \mathrm{Mor}\,(Y, Z) \to \mathrm{Mor}\,(X, Z)$$

called *composition* such that three axioms given below hold. One writes $f : X \to Y$ or $X \xrightarrow{f} Y$ to mean that $f \in \mathrm{Mor}\,(X, Y)$. If $f : X \to Y$ and $g : Y \to Z$ then the composition is denoted $g \circ f$ or gf. The axioms are as follows:

(0) $\mathrm{Mor}\,(X, Y)$ and $\mathrm{Mor}\,(X', Y')$ are disjoint unless $X = X'$ and $Y = Y'$.

(1) For every object X there is an element denoted 1_X or id_X in $\mathrm{Mor}\,(X, X)$ such that if $g : W \to X$ then $1_X \circ g = g$ while if $h : X \to Y$ then $h \circ 1_X = h$.

(2) If $f : W \to X$, $g : X \to Y$, and $h : Y \to Z$ then $h \circ (g \circ f) = (h \circ g) \circ f$ (*associativity* of composition).

The morphism 1_X is called the *identity* morphism on X and one can show that it is unique. If $f : X \to Y$ then X is called the *domain* of f and Y is called the *codomain*, *target*, or *range* of f, but it is preferable to avoid the term "range" because it is used for the set of values that a function actually takes on. A morphism $f : X \to Y$ is called an *isomorphism* if there is a morphism $g : Y \to X$ such that $gf = 1_X$ and $fg = 1_Y$. If it exists, g is unique and is an isomorphism from $Y \to X$. If there is an isomorphism from $X \to Y$ then X and Y are called *isomorphic*.

Examples. (a) Let the class of objects be the class of all sets, let the morphisms from a set X to a set Y be the functions from X to Y, and let composition be ordinary composition of functions. In this category of sets and functions, two sets are isomorphic if and only if they have the same cardinality.

In the next few examples the objects have underlying sets and composition coincides with composition of functions.

(b) Rings and ring homomorphisms form a category.

(c) Commutative rings with identity and ring homomorphisms that preserve the identity form a category.

(d) For a fixed ring R, R-modules and R-linear homomorphisms form a category.

Examples (c) and (d) give the environments in which we'll be "living" during this course.

(e) Groups and group homomorphisms are another example of a category.

We pause to review some basics about topological spaces before continuing with our examples.

A *topology* on a set X is a family of sets, called the *open sets* of the topology satisfying the following three axioms:

(0) The empty set and X itself are open.
(1) A finite intersection of open sets is open.
(2) An arbitrary union of open sets is open.

A set is called *closed* if its complement is open. A *topological space* is a set X together with a topology. Such a space may be described equally well by specifying what the closed sets are. They must satisfy:

(0) The empty set and X itself are closed.
(1) A finite union of closed sets is closed.
(2) An arbitrary intersection of closed sets is closed.

A subset Y of a topological space X becomes a topological space in its own right: one gets the topology by intersecting the open sets of X with Y. (The closed sets of Y are likewise gotten by intersecting the closed sets of X with Y.) The resulting topology on Y is called the *inherited* topology, and Y with this topology is called a (topological) *subspace* of X.

A topological space is called T_0 if for any two distinct points there is an open set that contains one of them and not the other. It is called T_1 if every point is closed. It is called T_2 or *Hausdorff* if for any two distinct points x and y there are disjoint open sets U and V such that $x \in U$ and $y \in V$.

A family of open subsets of a topological space X (following the usual imprecise practice, we mention the underlying set without mentioning the topology) is called an *open cover* if its union is all of X. A subset of such a family whose union is all of X is called a *subcover*.

A topological space is called *quasicompact* if every open cover has a subcover containing only finitely many open sets, i.e., a finite subcover.

A family of sets is said to have the *finite intersection property* if every finite subfamily has non-empty intersection. Being quasicompact is equivalent to the condition that every family of closed sets with the finite intersection property has non-empty intersection. (This is only interesting when the family is infinite.) A quasicompact Hausdorff space is called *compact*. We assume familiarity with the usual topology on $\mathbb{R}^n$, in which a set is closed if and only if for every convergent sequence of points in the set, the limit point of the sequence is also in the set. Alternatively, a set U is open if and only if for any point x in the set, there exists $a > 0$ in $\mathbb{R}$ such that all points of $\mathbb{R}^n$ within distance of a of x are in U.

The compact subspaces of $\mathbb{R}^n$ are precisely the closed, bounded sets.

A topological space is called *connected* if it is not the union of two non-empty disjoint open subsets (which will then both be closed as well). The connected subsets of the real line are identical with the intervals: these are the subsets with the property that if they contain a and b, they contain all real numbers in between a and b. They include the empty set, individual points, open intervals, half-open intervals, closed intervals, and the whole line.

A function f from a topological space X to a topological space Y is called *continuous* if for every open set V of Y, $f^{-1}V = \{x \in X : f(x) \in V\}$ is open. It is an equivalent condition to require that the inverse image of every closed set be closed.

We are now ready to continue with our discussion of examples of categories.

(f) Topological spaces and continuous maps give a category. In this category, isomorphism is called *homeomorphism*.

We now consider some examples in which composition is not necessarily composition of functions.

(g) A *partially ordered set* (or *poset*) consists of a set P together with a relation $\leq$ such that for all x, y, $z \in P$, (1) if $x \leq y$ and $y \leq x$ then $x = y$ and (2) if $x \leq y$ and $y \leq z$ then $x \leq z$. Given a partially ordered set, we can construct a category in which the objects are the elements of the partially ordered set. We artificially define there to be one morphism from x to y when $x \leq y$, and no morphisms otherwise. In this category, isomorphic objects are equal. Note that there is a unique way to define composition: if we have a morphism f from x to y and one g from y to z, then $x \leq y$ and $y \leq z$. Therefore, $x \leq z$, and there is a unique morphism from x to z, which we define to be the composition gf. Conversely, a category in which (1) the objects form a set, (2) there is at most one morphism between any two objects, and (3) isomorphic objects are equal is essentially the same thing as a partially ordered set. One defines a partial ordering on the objects by $x \leq y$ if and only if there is a morphism from x to y.

(h) A category with just one object in which every morphism is an isomorphism is essentially the same thing as a group. The morphisms of the object to itself are the elements of the group.

Lecture of September 11

Given any category $\mathcal{C}$ we can construct an *opposite* category $\mathcal{C}^{\mathrm{op}}$. It has the same objects as $\mathcal{C}$, but for any two objects X and Y in $\mathrm{Ob}\,(\mathcal{C})$, $\mathrm{Mor}_{\mathcal{C}^{\mathrm{op}}}(X, Y) = \mathrm{Mor}_{\mathcal{C}}(Y, X)$. There turns out to be an obvious way of defining composition using the composition in $\mathcal{C}$: if $f \in \mathrm{Mor}_{\mathcal{C}^{\mathrm{op}}}(X, Y)$ and $g \in \mathrm{Mor}_{\mathcal{C}^{\mathrm{op}}}(Y, Z)$ we have that $f : Y \to X$ in $\mathcal{C}$ and $g : Z \to Y$, in $\mathcal{C}$, so that $f \circ g$ in $\mathcal{C}$ is a morphism $Z \to X$ in $\mathcal{C}$, i.e., a morphism $X \to Z$ in $\mathcal{C}^{\mathrm{op}}$, and thus $g \circ_{\mathcal{C}^{\mathrm{op}}} f$ is $f \circ_{\mathcal{C}} g$.

By a (covariant) *functor* from a category $\mathcal{C}$ to a category $\mathcal{D}$ we mean a function F that assigns to every object X in $\mathcal{C}$ an object $F(X)$ in $\mathcal{D}$ and to every morphism $f : X \to Y$ in $\mathcal{C}$ a morphism $F(f) : F(X) \to F(Y)$ in $\mathcal{D}$ such that

(1) For all $X \in \mathrm{Ob}\,(\mathcal{C})$, $F(1_X) = 1_{F(X)}$ and
(2) For all $f : X \to Y$ and $g : Y \to Z$ in $\mathcal{C}$, $F(g \circ f) = F(g) \circ F(f)$.

A *contravariant functor* from $\mathcal{C}$ to $\mathcal{D}$ is a covariant functor to $\mathcal{C}$ to $\mathcal{D}^{\mathrm{op}}$. This means that when $f : X \to Y$ in $\mathcal{C}$, $F(f) : F(Y) \to F(X)$ in $\mathcal{D}$, and $F(g \circ f) = F(f) \circ F(g)$ whenever $g \circ f$ is defined in $\mathcal{C}$.

Here are some examples.

(a) Given any category $\mathcal{C}$, there is an identity functor $1_{\mathcal{C}}$ on $\mathcal{C}$: it sends the object X to the object X and the morphism f to the morphism f. This is a covariant functor.

(b) There is a functor from the category of groups and group homomorphisms to the category of abelian groups and homomorphisms that sends the group G to G/G', where G' is the commutator subgroup of G: G' is generated by the set of all commutators $\{ghg^{-1}h^{-1} : g, h \in G\}$: it is a normal subgroup of G. The group G/G' is abelian. Note also that any homomorphism from G to an abelian group must kill all commutators, and factors through G/G', which is called the abelianization of G.

Given $\phi : G \to H$, ϕ automatically takes commutators to commutators. Therefore, it maps G' into H' and so induces a homomorphism $G/G' \to H/H'$. This explains how this functor behaves on homomorphisms. It is covariant.

(c) Note that the composition of two functors is a functor. If both are covariant or both are contravariant the composition is covariant. If one is covariant and the other is contravariant, the composition is contravariant.

(d) There is a contravariant functor F from the category of topological spaces to the category of rings that maps X to the ring of continuous $\mathbb{R}$-valued functions on X. Given a continuous map $f : X \to Y$, the ring homomorphism $F(Y) \to F(X)$ is induced by composition: if $h : Y \to \mathbb{R}$ is any continuous function on Y, then $h \circ f$ is a continuous function on X.

(e) Given a category such as groups and group homomorphisms in which the objects have underlying sets and the morphisms are given by certain functions on those sets, we can give a covariant functor to the category of sets: it assigns to each object its underlying set, and to each morphism the corresponding function. Functors of this sort are called *forgetful* functors. The category of rings and ring homomorphisms and the category of topological spaces and continuous maps both have forgetful functors as well.

(f) A category $\mathcal{C}$ is said to be a *full subcategory* of the category $\mathcal{D}$ if $\mathrm{Ob}\,(\mathcal{C}) \subseteq \mathrm{Ob}\,(\mathcal{D})$, for all $X,\, Y \in \mathrm{Ob}\,(\mathcal{C})$, $\mathrm{Mor}_{\,\mathcal{C}}(X,\, Y) = \mathrm{Mor}_{\,\mathcal{D}}(X,\, Y)$, and composition in $\mathcal{C}$ is the same as in $\mathcal{D}$. Thus, finite sets yield a full subcategory of sets, abelian groups is a full subcategory of the category of groups, finitely generated R-modules is a full subcategory of the category of R-modules and R-linear maps, and Hausdorff topological spaces give a full subcategory of topological spaces. The category of rings with identity is, however, not a full subcategory of the category of rings (where there need not be a multiplicative identity): in the former, the identity is required to map to the identity. Thus $0 \to \mathbb{Z}$ is a homomorphism of rings but not of commutative rings with identity.

We next want to give a contravariant functor from commutative rings to topological spaces.

We first want to review some terminological conventions. All rings, unless otherwise specified, are commutative with multiplicative identity 1. We use 1_R for the identity in the ring R if greater precision is needed. We recall that $1 = 0$ is allowed, but this forces every element of the ring to be 0. Up to unique isomorphism, there is a unique ring with one element, which we denote 0.

By a *domain* or *integral domain* we mean a commutative ring such that $1 \neq 0$ and such that if $ab = 0$ then either $a = 0$ or $b = 0$. It is an arbitrary convention to exclude the ring in which every element is zero, but this turns out to be convenient. By a field we mean a ring in which $1 \neq 0$ and in which every nonzero element has an inverse under multiplication. A field K has only two ideals: $\{0\}$ and K. A field is an integral domain, although the converse is not true in general.

An ideal P in R is called *prime* if R/P is an integral domain. This means that P is prime in R if and only if $1 \notin P$ and for all $a,\, b \in R$, if $ab \in P$ then either $a \in P$ or $b \in P$.

An ideal $m \in R$ is called *maximal* if, equivalently, either R/m is a field or m is maximal among all proper ideals of R. A maximal ideal is prime.

Every proper ideal is contained in a maximal ideal. To see this, we first recall Zorn's lemma, which we shall not prove. It is equivalent to the axiom of choice in set theory. A subset of a partially ordered set is called a *chain* if it is linearly ordered, i.e., if any two of its elements are comparable.

(3.1) Zorn's lemma. *Let P be a non-empty partially ordered set in which every chain has an upper bound. Then P has a maximal element.*

(3.2) Corollary. *Let I be a proper ideal of the commutative ring R. Then I is contained in a maximal ideal m.*

Proof. We apply Zorn's lemma to the partially ordered set of proper ideals containing I. Given any chain containing I, its union is a proper ideal containing I and is an upper bound for the chain. Thus there are maximal elements in the set of proper ideals containing I, and these will be maximal ide als. □

It is also true that the existence of maximal ideals in commutative rings implies Zorn's lemma see [W. Hdoges, *Krull implies Zorn*, J. London Math. Soc. **19** (1979), 285–287], or [B. Banaschewski, *A new proof that "Krull implies Zorn"*, Math. Log. Quart. **40** (1994), 478–480].

We are now ready to introduce our functor, Spec , from commutative rings to topological spaces. If R is a ring, let Spec (R) denote the set of all prime ideals of R. Note that Spec (R) is empty if and only if R is the 0 ring. We place a topology, *the Zariski topology*, on Spec (R) as follows. For any subset I of R, let $V(I)$ denote the set $\{P \in \operatorname{Spec}(R) : I \subseteq P\}$. If the set I is replaced by the ideal it generates, $V(I)$ is unaffected. The Zariski topology has the subsets of Spec (R) of the form $V(I)$ as its closed sets. Note that $V(0) = \operatorname{Spec}(R)$, that $V(R) = \emptyset$, and that for any family of ideals $\{I_\lambda\}_{\lambda\in\Lambda}$,

$$\bigcap_{\lambda\in\Lambda} V(I_\lambda) = V(\sum_{\lambda\in\Lambda} I_\lambda).$$

It remains only to show that the union of two closed sets (and, hence, any finite number) is closed, and this will follow if we can show that for any two ideals I, J, $V(I) \cup V(J) = V(I \cap J) = V(IJ)$. It is clear that the leftmost term is smallest. Suppose that a prime P contains IJ but not I, so that $u \in I$ but $u \notin P$. For every $v \in J$, $uv \in P$, and since $u \notin P$, we have $v \in P$. Thus, if P does not contain I, it contains J. It follows that $V(IJ) \subseteq V(I) \cup V(J)$, and the result follows.

The Zariski topology is T_0. If P and Q are distinct primes, one of them contains an element not in the other. Suppose, say, that $u \in P$ and $u \notin Q$. The closed set $V(u)$ contains P but not Q.

It is easy to show that the closure of the one point set $\{P\}$, where P is prime, is the set $V(P)$. The closure has the form $V(I)$, and is the smallest set of this form such that $P \in V(I)$, i.e., such that $I \subseteq P$. As I gets smaller, $V(I)$ gets larger. It is therefore immediate that the smallest closed set containing P is $V(P)$.

It follows that $\{P\}$ is closed if and only if P is maximal. In general, Spec (R) is not T_1.

Spec becomes a contravariant functor from the category of commutative rings with identity to the category of topological spaces if, given a ring homomorphism $f : R \to S$, we define Spec (f) by having it send $Q \in \operatorname{Spec}(S)$ to $f^{-1}(Q) = \{r \in R : f(r) \in Q\}$. There is an induced ring homomorphism $R/f^{-1}(Q) \to S/Q$ which is injective. Since S/Q is an integral domain, so is its subring $R/f^{-1}(Q)$. (We are also using tacitly that the inverse image of a proper ideal is proper, which is a consequence of our convention that $f(1_R) = 1_S$.) $f^{-1}(Q)$ is sometimes denoted Q^c and called the *contraction* of Q to R. This is a highly ambiguous notation.

We want to talk about when two functors are isomorphic and to do that, we need to have a notion of morphism between two functors. Let F, G be functors from $\mathcal{C} \to \mathcal{D}$ with the same variance. For simplicity, we shall assume that they are both covariant. The case where they are both contravariant is handled automatically by thinking instead of the case of covariant functors from $\mathcal{C}$ to $\mathcal{D}^{\mathrm{op}}$. A *natural transformation* from F to G assigns to every object $X \in \mathrm{Ob}\,(\mathcal{C})$ a morphism $T_X : F(X) \to G(X)$ in such a way that for all morphisms $f : X \to Y$ in $\mathcal{C}$, there is a commutative diagram:

$$\begin{tikzcd}
F(X) \arrow[r, "F(f)"] \arrow[d, "T_X"'] & F(Y) \arrow[d, "T_Y"] \\
G(X) \arrow[r, "G(f)"'] & G(Y)
\end{tikzcd}$$

The commutativity of the diagram simply means that $T_Y \circ F(f) = G(f) \circ T_X$.

This may seem like a complicated notion at first glance, but it is actually very "natural," if you will forgive the expression.

This example may clarify. If V is a vector space write V^* for the space of linear functionals on V, i.e., for $\mathrm{Hom}_K(V,\, K)$, the K-vector space of K-linear maps from $V \to K$. Then * is a contravariant functor from K-vector spaces and K-linear maps to itself. (If $\theta : V \to W$ is linear, $\theta^* : W^* \to V^*$ is induced by composition: if $g \in W^*$, so that $g : W \to K$, then $\theta^*(g) = g \circ \theta$.)

The composition of * with itself gives a covariant functor **: the double dual functor. We claim that there is a natural transformation T from the identity functor to **. To give T is the same as giving a map $T_V : V \to V^{**}$ for every vector space V. To specify $T_V(v)$ for $v \in V$, we need to give a map from V^* to K. If $g \in V^*$, the value of $T_V(v)$ on g is simply $g(v)$. To check that this is a natural transformation, one needs to check that for every K-linear map $f : V \to W$, the diagram

$$\begin{tikzcd}
V \arrow[r, "f"] \arrow[d, "T_V"'] & W \arrow[d, "T_W"] \\
V^{**} \arrow[r, "f^{**}"'] & W^{**}
\end{tikzcd}$$

commutes. This is straightforward. Note that the map $V \to V^{**}$ is not necessarily an isomorphism. It is always injective, and is an isomorphism when V is finite-dimensional over K.

Lecture of September 13

Here is another example of a natural transformation: in this case, the functors are contravariant. Let F and G be the functors from topological spaces to rings such that $F(X)$ (respectively, $G(X)$) is the ring of continuous real-valued (respectively, complex-valued) functions on X. (The values on continuous maps are both induced by composition.) The inclusions $F(X) \subseteq G(X)$ give a natural transformation from F to G.

Le $\mathcal{C}$ be the category of pairs $(X,\, x)$ where X is a non-empty topological space and $x \in X$, i.e., of topological spaces with basepoint. A morphism from $(X,\, x)$ to (Y, y) is a continuous function from X to Y such that $f(x) = y$. For every X there is a group homomorphism from $T_X : \pi_1(X, x) \to H_1(X,\, \mathbb{Z})$ where the former is the fundamental group and the latter is singular homology with integer coefficients. (Let S^1 be a circle and fix a generator θ of $H_1(S^1, \mathbb{Z}) \cong \mathbb{Z}$. Every element of $\pi_1(X, x)$ is represented by (the homotopy class of) a continuous map $f : S^1 \to X$. $T_X([f]) = f_*(\theta) \in H_1(X,\, \mathbb{Z})$.) These T_X give a natural transformation from π_1 to the functor $H_1(\_\,, \mathbb{Z})$, both regarded as functors from $\mathcal{C}$ to the category of groups. There are also natural transformations $H_1(\_\,, \mathbb{Z}) \to H_1(\_\,, \mathbb{Q}) \to H_1(\_\,, \mathbb{R}) \to H_1(\_\,, \mathbb{C})$.

In giving definitions for natural transformations, we will stick with the case of covariant functors: the contravariant case may be handled by replacing $\mathcal{D}$ by $\mathcal{D}^{\mathrm{op}}$.

Given functors F, G, H from $\mathcal{C} \to \mathcal{D}$, a natural transformation $S : F \to G$, and a natural transformation $T : G \to H$, we may define a natural transformation $T \circ S$ from F to H by the rule $(T \circ S)_X = T_X \circ S_X$.

There is an identity natural transformation, 1_F, from the functor $F : \mathcal{C} \to \mathcal{D}$ to itself: $1_{F,X} : F(X) \to F(X)$ is $1_{F(X)}$. It behaves as an identity should under composition. Given two functors F and G from $\mathcal{C} \to \mathcal{D}$, we can now define them to be isomorphic if there are natural transformations $T : F \to G$ and $T' : G \to F$ such that $T' \circ T = 1_F$ and $T \circ T' = 1_G$. In fact, T is an isomorphism of functors if and only if all the morphisms T_X are isomorphisms, and in that case the unique way to define T' is by the rule $T'_X = (T_X)^{-1}$.

Once we have a notion of isomorphism of functors we can define two categories $\mathcal{C}$ and $\mathcal{D}$ to be *equivalent* if there are functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ such that $G \circ F$ is isomorphic to the identity functor on $\mathcal{C}$ and $F \circ G$ is isomorphic to the identity functor on $\mathcal{D}$. If $\mathcal{C}$ is equivalent to $\mathcal{D}^{\mathrm{op}}$ it is said to be *antiequivalent* to $\mathcal{D}$. Roughly speaking, equivalence is like isomorphism, but there may not be the same number of objects in an isomorphism class in one of the two equivalent categories as there are in the other. For example, suppose that we have a category $\mathcal{D}$ and another $\mathcal{C}$ in which there is exactly one object of $\mathcal{D}$ from each isomorphism class of objects in $\mathcal{D}$. Also suppose that the morphisms from one object in $\mathcal{C}$ to another are the same as when they are considered as objects of $\mathcal{D}$, and likewise for composition. Then one can show, with a suitably strong form of the axiom of choice, that $\mathcal{C}$ and $\mathcal{D}$ are equivalent categories.

Another application of the notion of isomorphism of functors is the definition of a *representable* functor. This is a point of view that unifies numerous constructions, both in commutative algebra and in many other parts of mathematics. If we fix an object Z in a category $\mathcal{C}$ then we get a covariant functor h_Z mapping $\mathcal{C}$ to the category of sets by letting $h_Z(X) = \mathrm{Mor}\,(Z, X)$. If $f : X \to Y$ we let $h_Z(f) : \mathrm{Mor}\,(Z, X) \to \mathrm{Mor}\,(Z, Y)$ be the map induced by composition — it sends g to $f \circ g$. A covariant functor G from $\mathcal{C}$ to sets is called *representable* in $\mathcal{C}$ if it is isomorphic to h_Z for some $Z \in \mathrm{Ob}\,(\mathcal{C})$. We say that Z *represents* G. Similarly, we can define a contravariant functor h^Z to sets by $h^Z(X) = \mathrm{Mor}\,(X, Z)$ while $h^Z(f) : \mathrm{Mor}\,(Y, Z) \to \mathrm{Mor}\,(X, Z)$ sends g to $g \circ f$. A contravariant functor is *representable* in $\mathcal{C}$ if it is isomorphic with h^Z for some Z.

Examples. (a) Let $\mathcal{C}$ be the category of abelian groups and group homomorphisms. Let G be any group. We can define a functor F from abelian groups to sets by letting $F(A) = \mathrm{Hom}(G, A)$, the set of group homomorphisms from G to A. Can we represent F in the category of abelian groups? Yes! Let $\overline{G} = G/G'$, the abelianization of G. Then every homomorphism $G \to A$ factors uniquely $G \to \overline{G} \to A$, giving a bijection of $F(A)$ with $\mathrm{Hom}(\overline{G}, A)$. This yields an isomorphism of $F \cong h_{\overline{G}}$.

(b) Let R be a ring and and I be an ideal. Define a functor from the category of commutative rings with identity to the category of sets by letting $F(S)$ be the set of all ring homomorphisms $f : R \to S$ such that f kills I. Every homomorphism $R \to S$ such that f kills I factors uniquely $R \twoheadrightarrow R/I \to S$, from which it follows that the functor F is representable and is $\cong h_{R/I}$.

(c) In this example we want to define products in an arbitrary category. Our motivation is the way the Cartesian product $Z = X \times Y$ behaves in the category of sets. It has product projections $\pi_X : Z \to X$ sending (x, y) to x and $\pi_Y : Z \to Y$ sending (x, y) to y. To give a function from $W \to X \times Y$ is equivalent to giving a pair of functions, one $\alpha : W \to X$ and another $\beta : W \to Y$. The function $f : W \to X \times Y$ then sends w to $(\alpha(w), \beta(w))$. The functions α and β may be recovered from f as $\pi_X \circ f$ and $\pi_Y \circ f$, respectively.

Now let $\mathcal{C}$ be any category. Let $X, Y \in \mathrm{Ob}\,(C)$. An object Z together with morphisms $\pi_X : Z \to X$ and $\pi_Y : Z \to Y$ (called the *product projections* on X an Y, respectively) is called a *product* for X and Y in $\mathcal{C}$ if for all objects W in $\mathcal{C}$ the function $\mathrm{Mor}\,(W, Z) \to \mathrm{Mor}\,(W, X) \times Mor\,(W, Y)$ sending f to $(\pi_X \circ f, \pi_Y \circ f)$ is a bijection. This means that the functor sending W to $\mathrm{Mor}\,(W, X) \times \mathrm{Mor}\,(W, Y)$ is representable in $\mathcal{C}$. Given another product Z', π'_X, π'_Y, there are unique mutually inverse isomorphisms $\gamma : Z \to Z'$ and $\delta : Z' \to Z$ that are compatible with the product projections, i.e., such that $\pi_X = \gamma \circ \pi'_X$ $\pi_Y = \gamma \circ \pi'_Y$ (the existence and uniqueness of γ are guaranteed by the defining property of the product) and similarly for δ. The fact that the compositions are the appropriate identity maps also follows from the defining property of the product.

Products exist in many categories, but they may fail to exist. In the categories of sets, rings, groups, abelian groups, R-modules over a given ring R, and topological spaces, the product turns out to be the Cartesian product with the usual additional structure (in the algebraic examples, operations are performed coordinate-wise; in the case of topological spaces, the product topology works: the open sets are unions of Cartesian products of open sets from the two spaces). In all of these examples, the product projections are the usual set-theoretic ones. In the category associated with a partially ordered set, the product of two elements x and y is the greatest lower bound of x and y, if it exists. The point is that w has (necessarily unique) morphisms to both x and y iff $w \le x$ and $w \le y$ iff w is a lower bound for both x and y. For z to be a product, we must have that z is a lower bound for x, y such that every lower bound for x, y has a morphism to z. This says that z is a greatest lower bound for x, y in the partially ordered set. It is easy to give examples of partially ordered sets where not all products exist: e.g., a partially ordered set that consists of two mutually incomparable elements (there is no lower bound for the two), or one in which there are four elements a, b, x, y such that a and b are incomparable, x and y are incomparable, while both a and b are strictly less than both x and y. Here, a and b are both lower bounds for the x, y, but neither is a greatest lower bound.

The product of two objects in $\mathcal{C}^{\mathrm{op}}$ is called their *coproduct* in $\mathcal{C}$. Translating, the coproduct of X and Y in $\mathcal{C}$, if it exists, is given by an object Z and two morphisms $\iota_X : X \to Z$, $\iota_Y : Y \to Z$ such that for every object W, the map $\mathrm{Mor}\,(Z, W) \to \mathrm{Mor}\,(X, W) \times \mathrm{Mor}\,(Y, W)$ sending f to $(f \circ \iota_X, f \circ \iota_Y)$ is bijective. This means that the functor sending W to $\mathrm{Mor}\,(X, W) \times \mathrm{Mor}\,(Y, W)$ is representable in $\mathcal{C}$. Coproducts have the same sort of uniqueness that products do: they *are* products (in $\mathcal{C}^{\mathrm{op}}$).

In the category of sets, coproduct corresponds to disjoint union: one takes the union of disjoint sets X' and Y' set-isomorphic to X and Y respectively. The function ι_X is an isomorphism of X with X' composed with the inclusion of X' in $X' \cup Y'$, and similarly for ι_Y. To give a function from the disjoint union of two sets to W is the same as to give two functions to W, one from each set.

In the category of R-modules over a commutative ring R, coproduct corresponds to direct sum. We shall discuss the existence of coproducts in the category of commutative rings later on. In the category associated with a partially ordered set, it corresponds to the least upper bound of the two elements.

Lecture of September 15

Let R be a commutative ring with identity. An R-module F is said to be *free* with *free basis* $\mathcal{B} \subseteq F$ if every element of F is uniquely an R-linear combination of elements in $\mathcal{B}$. The uniqueness statement is very important: it implies that if $b_1, \ldots, b_n$ are distinct elements of $\mathcal{B}$ and $r_1 b_1 + \cdots + r_n b_n = 0$ then $r_1 = \cdots = r_n = 0$, which says that the elements of the free basis are linearly independent over R.

A word about degenerate cases: the 0 module is considered free on the empty set of generators.

In case R is a field, an R-module is just a vector space, and a free basis is the same thing as a vector space basis. (The term "basis" for a module is sometimes used to mean a set of generators or spanning set for the module. I will try not to use this term in this course, to avoid ambiguity.) By Zorn's lemma, every set of independent vectors in a vector space is contained in a maximal such set (one can start with the empty set), and a maximal independent set must span the whole space: any vector not in the span could be used to enlarge the maximal independent set. Thus, over a field, every module is free (i.e., every vector space has a basis).

Freeness is equivalent to the statement that for every $b \in \mathcal{B}$, $Rb \cong R$ in such a way that rb corresponds to r, and that F is the direct sum of all these copies of R, i.e., $F \cong \bigoplus_{b \in B} Rb$. The free R-module on the free basis $b_1, \ldots, b_n$ is isomorphic with R^n, the module of n-tuples of elements of R under coordinate-wise addition and scalar multiplication. Under the isomorphism, the element $r_1 b_1 + \cdots + r_n b_n$ corresponds to $(r_1, \ldots, r_n)$. The element b_i corresponds to $e_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ where the unique nonzero entry (which is 1) occurs in the i th coordinate. In particular, the e_i give a free basis for R^n.

In general, if F is free on $\mathcal{B}$, F is isomorphic with the set of functions $\mathcal{B} \to R$ which are 0 on all but finitely many elements of $\mathcal{B}$. Under the isomorphism, the element $r_1 b_1 + \cdots + r_n b_n$ corresponds to the function that assigns every b_i the value r_i, while assigning the value 0 to all elements of $\mathcal{B} - \{b_1, \ldots, b_n\}$. When $\mathcal{B}$ is infinite, this is strictly smaller than the set of all functions from $\mathcal{B}$ to R: the latter may be thought of as the product of a family of copies of R indexed by $\mathcal{B}$.

When M and N are R-modules, the set of R-linear maps from M to N is denoted $\mathrm{Hom}_R(M, N)$ or $\mathrm{Hom}\,(M, N)$: this is $\mathrm{Mor}\,(M, N)$ in the category of R-modules. It is not only a set: it is also an R-module, since we may define $f + g$ and rf for $r \in R$ by the rules $(f + g)(m) = f(m) + g(m)$ and $(rf)(m) = r\big(f(m)\big)$.

We next want to define the notion of an A-algebra, where A is a commutative ring. We shall say that R is an *A-algebra* if R itself is a commutative ring and is also a (unital) A-module in such a way that for all $a \in A$ and $r, s \in R$, $a(rs) = (ar)s$. (Note that the for all a, $b \in A$ and $r \in R$, we also have that $a(br) = (ab)r$, but we don't need to assume it separately: it is part of the definition of an A-module.) In this situation we get a ring homomorphism from $A \to R$ that sends $a \in A$ to $a \cdot 1_R$. Conversely, given a ring homomorphism $\theta : A \to R$, the ring R becomes an A-algebra if we define ar as $\theta(a)r$. That is, to give a ring R the structure of an A-algebra is exactly the same thing as to give a ring homomorphism $A \to R$. When R is an A-algebra, the homomorphism $\theta : A \to R$ is called the *structural* homomorphism of the algebra. A-algebras form a category: the A-algebra morphisms (usually referred to as *A-algebra homomorphisms*) from R to S are the A-linear ring homomorphisms. If f and g are the structural homomorphisms of R and S respectively over A and $h : R \to S$ is a ring homomorphism, it is an A-algebra homomorphism if and only if $h \circ f = g$.

Note that every commutative ring R with identity is a $\mathbb{Z}$-algebra in a unique way, i.e., there is a unique ring homomorphism $\mathbb{Z} \to R$. To see this, observe that 1 must map to 1_R. By repeated addition, we see that n maps to $n \cdot 1_R$ for every nonnegative integer n. It follows by taking inverses that this holds for negative integers as well. This shows uniqueness, and it is easy to check that the map that sends n to $n \cdot 1_R$ really is a ring homomorphism for every ring R.

By a *semigroup S* we mean a set together with an associative binary operation that has a two-sided identity. (The existence of such an identity is not always assumed. Some people use the term “monoid” for a semigroup with identity.) We shall assume the semigroup operation is written multiplicatively and that the identity is denoted 1_S or simply 1. A group is a semigroup in which every element has a two-sided inverse.

By a *homomorphism* of semigroups $h : S \to S'$ we mean a function on the underlying sets such that for all $s,\, t \in S$, $h(st) = h(s)h(t)$ and such that $h(1_S) = 1_{S'}$.

The elements of a commutative ring with identity form a commutative semigroup under multiplication.

The set of vectors $\mathbb{N}^n$ with nonnegative integer entries forms a semigroup under addition with identity $(0, \ldots, 0)$. We want to introduce an isomorphic semigroup that is written multiplicatively. If $x_1, \ldots, x_n$ are distinct elements we can introduce *formal monomials* $x_1^{k_1} \cdots x_n^{k_n}$ in these elements, in bijective correspondence with the elements $(k_1, \ldots, k_n) \in \mathbb{N}^n$. (We can, for example, make all this precise by letting $x_1^{k_1} \cdots x_n^{k_n}$ be an alternate notation for the function whose value on x_i is k_i, $1 \le i \le n$.) These formal monomials form a multiplicative semigroup that is isomorphic as a semigroup with $\mathbb{N}^n$: to multiply two formal monomials, one adds the corresponding exponents. It is also innocuous to follow the usual practices of omitting a power of one of the x_i from a monomial if the exponent on x_i is 0, of replacing x_i^1 by x_i, and of writing 1 for $x_1^0 \cdots x_n^0$. With these conventions, x_i^k is the product of x_i with itself k times, and $x_i^{k_1} \cdots x_n^{k_n}$ is the product of n terms, of which the ith term is $x_i^{k_i}$.

We can likewise introduce the multiplicative semigroup of formal monomials in the elements of an infinite set: it can thought of as the union of what one gets from its various finite subsets. Only finitely many of the elements occur with nonzero exponents in any given monomial.

Not every commutative semigroup is isomorphic with the multiplicative semigroup of a ring: for one thing, there need not be an element that behaves like 0. But even if we introduce an element that behaves like 0, this still need not be true. The infinite multiplicative semigroup of monomials in just one element, $\{x^k : k \in \mathbb{N}\}$, together with 0, is not the multiplicative semigroup of a ring. To see this, note that the ring must contain an element to serve as -1. If that element is x^k for $k > 0$, then $x^{2k} = 1$, and the multiplicative semigroup is not infinite after all. Therefore, we must have that $-1 = 1$, i.e., that the ring has characteristic 2. But then $x + 1$ must coincide with x^k for some $k > 1$, i.e., the equation $x^k - x - 1 = 0$ holds. This implies that every power of x is in the span of $1,\ x,\ \ldots,\ x^{k-1}$, forcing the ring to be a vector space of dimension at most k over $\mathbb{Z}_2$, and therefore finite, a contradiction.

Given a commutative semigroup S and a commutative ring A we can define a functor G from the category of A-algebras to sets whose value on R is the set of semigroup homomorphisms from S to R. If we have a homomorphism $R \to R'$ composition with it gives a function from $G(R)$ to $G(R')$. In this way, G is a covariant functor to the category of sets. We want to see that G is representable in the category of A-algebras. The construction is as follows: we put an A-algebra structure on the free A-module with free basis S by defining the product of $\sum_{i=1}^h a_i s_i$ with $\sum_{j=1}^k a'_j s'_j$, where the $a_i, a'_j \in A$ and the $s_i,\ s'_j \in S$, to be $\sum_{i,\,j}(a_i a'_j)(s_i s_j)$ where $a_i a'_j$ is calculated in A and $s_i s'_j$ is calculated in S. It is straightforward to check that this is a commutative ring with identity $1_A 1_S$ This ring is denoted $A[S]$ and is called the *semigroup ring of S with coefficients in A*. We identify S with the set of elements of the form $1_A s,\ s \in S$. It turns out that every semigroup homomorphism $\phi : S \to R$ (using R for the multiplicative semigroup of R), where R is an A-algebra, extends uniquely to an A-algebra homomorphism $A[S] \to R$. It is clear that to perform the extension one must send $\sum_{i=1}^h a_i s_i$ to $\sum_{i=1}^h a_i \phi(s_i)$, and it is straightforward to verify that this is an A-algebra homomorphism. Thus, restriction to S gives a bijection from $\mathrm{Hom}_A(A[S],\ R)$ to $G(R)$ for every A-algebra R, and so $A[S]$ represents the functor G in the category of A-algebras.

We can now define the polynomial ring in a finite or infinite set of variables $\{x_i : i \in I\}$ over A as the semigroup ring of the formal monomials in the x_i with coefficients in A.

We can also view the polynomial ring $A[\mathcal{X}]$ in a set of variables $\mathcal{X}$ as arising from representing a functor as follows. Given any A-algebra R, to give an A-homomorphism from $A[\mathcal{X}] \to R$ is the same as to give a function from $\mathcal{X} \to R$, i.e., the same as simply to specify the values of the A-homomorphism on the variables. Clearly, if the homomorphism is to have value r_i on x_i for every $x_i \in \mathcal{X}$, the monomial $x_{i_1}^{k_1} \cdots x_{i_n}^{k_n}$ must map to $r_{i_1}^{k_1} \cdots r_{i_n}^{k_n}$, and this tells us as well how to map any A-linear combination of monomials. If for example, only the indeterminates $x_1,\ \ldots,x_n$ occur in a given polynomial (there are always only finitely many in any one polynomial) then the polynomial can be written uniquely as $\sum_{\underline{k} \in E} a_{\underline{k}} x^{\underline{k}}$ where E is the finite set of n-tuples of exponents corresponding to monomials occurring with nonzero coefficient in the polynomial, $\underline{k} = (k_1,\ \ldots,k_n)$ is a n-tuple varying in E, every $a_{\underline{k}} \in A$, and $x^{\underline{k}}$ denotes $x_1^{k_1} \cdots x_n^{k_n}$. If the value that x_i has is r_i, this polynomial must map to $\sum_{\underline{k} \in E} a_{\underline{k}} r^{\underline{k}}$, where $r^{\underline{k}}$ denotes $r_1^{k_1} \cdots r_n^{k_n}$. It is straightforward to check that this does give an A-algebra homomorphism. In the case where there are n variables $x_1,\ \ldots,x_n$, and every x_i is to map to r_i, the value of a polynomial P under this homomorphism is denoted $P(r_1,\ \ldots,r_n)$, and we refer to the homomorphism as *evaluation* at $(r_1, \ldots, r_n)$. Let H denote the functor from A-algebras to sets whose value on R is the set of functions from $\mathcal{X}$ to R. Then the polynomial ring $A[\mathcal{X}]$ represents the functor H in the category of A-algebras: the map from $\mathrm{Hom}_A(A[\mathcal{X}], R)$ to $\mathrm{Mor}_{\,(\mathrm{sets})}(\mathcal{X}, R)$ that simply restricts a given A-homomorphism $A[\mathcal{X}] \to R$ to the set $\mathcal{X}$ gives a bijection, and this gives the required natural isomorphism of functors.

By a *multiplicative system* S in a ring R we mean a non-empty subset of R that is closed under multiplication. Given such a set S we next want to consider the problem of representing the functor L_S in the category of rings, where $L_S(T)$ denotes the set of ring homomorphisms $R \to T$ such the image of every element of S is invertible in T. We shall show that this is possible, and denote the ring we construct $S^{-1}R$. It is called the *localization* of R at S. It is constructed by enlarging R to have inverses for the elements of S while changing R as little as possible in any other way.

Lecture of September 18

We give two constructions of the localization of a ring R at a multiplicative system $S \subseteq R$. In the first construction we introduce an indeterminate x_s for every element of S. Let $A = R[x_s : s \in S]$, the polynomial ring in all these indeterminates. Let I be the ideal of A generated by all of the polynomials $sx_s - 1$ for $s \in S$. The composition of the homomorphisms $R \to R[x_s : s \in S] = A \twoheadrightarrow A/I$ makes A/I into an R-algebra, and we take $S^{-1}R$ to be this R-algebra. Note that the polynomials we killed force the image of x_s in $S^{-1}R$ to be an inverse for the image of s.

Now suppose that $g : R \to T$ is any ring homomorphism such that $g(s)$ is invertible in T for every element $s \in S$. We claim that $R \to T$ factors uniquely $R \to S^{-1}R \to T$, where the first homomorphism is the one we constructed above. To obtain the needed map, note that we must give an R-homomorphism of $A = R[x_s : s \in S] \to T$ that kills the ideal I. But there is one and only one way to specify values for the x_s in T so that all of the polynomials $sx_s - 1$ map to 0 in T: we must map x_s to $g(s)^{-1}$. This proves that the map does, in fact, factor uniquely in the manner specified, and also shows that $S^{-1}R$ represents the functor

$$L_S = \{g \in \mathrm{Hom}_R(R,\, T) : \text{for all } s \in S,\ g(s) \text{ is invertible}\}$$

in the category of rings, as required. Note that $x_{s_1} \cdots x_{s_k} = x_{s_1 \cdots s_k} \bmod I$, since both sides represent inverses for the image of $s_1 \cdots s_k$ in $S^{-1}T$. This means that every element of $S^{-1}R$ is expressible as an R-linear combination of the x_s. But we can manipulate further: it is easy to check that the images of $rs_2x_{s_1s_2}$ and rx_{s_1} are the same, since they are the same after multiplying by the invertible element which is the image of s_1s_2, and so $r_1x_{s_1} + r_2x_{s_2} = r_1s_2x_{s_1s_2} + r_2s_1x_{s_1s_2} = (r_1s_2 + r_2s_1)x_{s_1s_2} \bmod I$. Therefore every element of $S^{-1}R$ can be written as the image of rx_s for some $r \in R$ and $s \in S$. This representation is still not unique.

We now discuss the second construction. An element r of the ring R is called a *zerodivisor* if $ru = 0$ for $u \in R - \{0\}$. An element that is not a zerodivisor is a called a nonzerodivisor. The second construction is slightly complicated by the possibility that S contains zerodivisors. Define an equivalence relation $\sim$ on $R \times S$ by the condition that $(r_1, s_1) \sim (r_2, s_2)$ if there exists $s \in S$ such that $s(r_1s_2 - r_2s_1) = 0$. Note that if S contains no zerodivisors, this is the same as requiring that $r_1s_2 - r_2s_1 = 0$. In the case where S contains zerodivisors, one does not get an equivalence relation from the simpler condition. The equivalence class of (r, s) is often denoted r/s, but we stick with $[(r, s)]$ for the moment. The check that one has an equivalence relation is straightforward, as is the check that the set of equivalence classes becomes a ring if we define the operations by the rules $[(r_1,\, s_1)] + [(r_2,\, s_2)] = [(r_1s_2 + r_2s_1,\, s_1s_2)]$ and $[(r_1,\, s_1)][(r_2,\, s_2)] = [(r_1r_2,\, s_1s_2)]$. One needs to verify that the operations are well-defined, i.e., independent of choices of equivalence class representatives, and that the usual ring laws are satisfied. This is all straightforward. The zero element is $[(0,\, 1)]$, the multiplicative identity is $[(1,\, 1)]$, and the negative of $[(r,\, s)]$ is $[(-r,\, s)]$. Call this ring B for the moment. It is an R-algebra via the map that sends r to $[(r,\, 1)]$. The elements of S have invertible images in B, since $[(s,\, 1)][(1,\, s)] = [(s,\, s)] = [(1,\, 1)]$.

This implies that we have an R-algebra homomorphism $T \to B$. Note that it maps x_s to $[(1,\, s)]$, and, hence, it maps rx_s to $[(r,\, s)]$. Now one can prove that T is isomorphic with B by showing that the map $R \times S \to T$ that sends $(r,\, s)$ to rx_s is well-defined on equivalence classes. This yields a map $B \to T$ that sends $[(r,\, s)]$ to rx_s. It is then immediate that these are mutually inverse ring isomorphisms: since every element of T has the form rx_s, it is clear that the composition in either order gives the appropriate identity map.

It is easy to calculate the kernel of the map $R \to S^{-1}R$. By the definition of the equivalence relation we used, $(r, 1) \sim (0, 1)$ means that for some $s \in S$, $sr = 0$. The set $I = \{r \in R : \text{for some } s \in S, sr = 0\}$ is therefore the kernel. If no element of s is a zerodivisor in R, then the map $R \to S^{-1}R$ is injective. One can think of localization at S as being achieved in two steps: first kill I, and then localize at the image of S, which will consist entirely of nonzerodivisors in R/I.

If R is an integral domain then $S = R - \{0\}$ is a multiplicative system. In this case, $S^{-1}R$ is easily verified to be a field, the *fraction field* of R. Localization may be viewed as a generalization of the construction of fraction fields.

Localization and forming quotient rings are related operations. Both give R-algebras that represent functors. One corresponds to homomorphisms that kill an ideal I, while the other to homomorphisms that make every element in a multiplicative system S invertible. But the resemblance is even greater.

To explain this further similarity, we introduce the notion of an epimorphism in an arbitrary category. In the category of sets it will turn out that epimorphisms are just surjective maps. But this is not at all true in general. Let $\mathcal{C}$ be a category. Then $f : X \to Y$ is an *epimorphism* if for any two morphisms $g,\, h : Y \to Z$, whenever $g \circ f = h \circ f$ then $g = h$. In the case of functions, this says that if g and h agree on $f(X)$, then they agree on all of Y. This is obviously true if $f(X) = Y$, i.e., if f is surjective. It is almost as obvious that it is not true if f is not surjective: let Z have two elements, say 0 and 1. Let g be constantly 0 on Y, and let h be constantly 0 on $f(X)$ and constantly 1 on its complement. Then $g \neq h$ but $g \circ f = h \circ f$.

In the category of R-modules an epimorphism is a surjective homomorphism. In the category of Hausdorff topological spaces, any continuous function $f : X \to Y$ is an epimorphism provided that $f(X)$ is dense in Y: it need not be all of Y. Suppose that $g : Y \to Z$ and $h : Y \to Z$ agree on $f(X)$. We claim that they agree on all of Y. For suppose we have $y \in Y$ such that $g(y) \neq h(y)$. Then $g(y)$ and $h(y)$ are contained in disjoint open sets, U and V respectively, of Z. Then $g^{-1}(U) \cap h^{-1}(V)$ is an open set in Y, and is non-empty, since it contains y. It follows that it contains a point of $f(X)$, since $f(X)$ is dense in Y, say $f(x)$, where $x \in X$. But then $g\big(f(x)\big) \in U$, and $h\big(f(x)\big) \in V$, a contradiction, since $g\big(f(x)\big) = h\big(f(x)\big)$ is in $U \cap V$, while U and V were chosen disjoint.

The category of rings also provides some epimorphisms that are not surjective: both surjective maps and localization maps $R \to S^{-1}R$ are epimorphisms. We leave it as an exercise to verify that if two homomorphisms $S^{-1}R \to T$ agree on the image of R, then they agree on $S^{-1}T$.

By the way, an epimorphism in $\mathcal{C}^{\mathrm{op}}$ is called a *monomorphism* in $\mathcal{C}$. Said directly, $f : X \to Y$ is a monomorphism if whenever $g,\ h : W \to X$ are such that $f \circ g = f \circ h$ then $g = h$. We leave it as an exercise to verify that a monomorphism in the category of sets is the same as an injective function. This is also true in the category of R-modules, and in the category of rings.

Here is an example of a fairly simple category in which there are underlying sets and functions and a monomorphism that is not injective. An abelian group $(A, +)$ is called *divisible* if for every integer $n \neq 0$ and every $b \in A$, there exists an element $a \in A$ (it need not be unique) such that $na = b$. That, is multiplication by $n \neq 0$ gives a surjection of A onto itself. Vector spaces over the rational numbers are divisible, and arbitrary homomorphic images of divisible abelian groups are divisible. In the full subcategory of abelian groups whose objects are the divisible groups, the surjective map $\pi : \mathbb{Q} \twoheadrightarrow Q/\mathbb{Z}$ is a monomorphism. To see this, suppose that f and g are homomorphisms from a divisible abelian group D to $\mathbb{Q}$ such that $\pi \circ f = \pi \circ g$. We need to show that $f = g$. But then $\pi \circ (f - g) = 0$, which implies that the image of $f - g$ is in the kernel $\mathbb{Z}$ of π. This image must be a divisible subgroup of the integers. Hence, the image of $f - g$ is $\{0\}$, and so $f - g = 0$ and $f = g$, as required.

An ideal of a ring R is prime if and only if its complement is a multiplicative system. (Note that our multiplicative systems are required to be non-empty.) If P is a prime, the localization of R at P is denoted R_P. We shall soon see that R_P has a unique maximal ideal, which is generated by the image of P. A ring with a unique maximal ideal is called a *quasilocal* ring. Some authors use the term *local*, but we shall reserve that term for a Noetherian quasilocal ring. A major theme in commutative algebra is to use localization at various primes to reduce problems to the case where the ring is quasilocal.

We want to make a detail comparison of the ideals of a ring R with the ideals of the ring $S^{-1}R$. But we first want to explain why rings with just one maximal ideal are called "(quasi)local."

Let X be a topological space and x a point of X. Consider the set of functions from an open set containing x to $\mathbb{R}$. We define two such functions to be equivalent if they agree when restricted to a sufficiently small open set containing x. The equivalence classes are referred to as *germs* of continuous functions at x, and they form a ring. In this ring, the value of a germ of a function at a specific point is not well-defined, with the exception of the point x. A germ that does not vanish at x will, in fact, not vanish on an open set containing x, by continuity, and therefore has an inverse (given by taking the reciprocal at each point) on an open set containing x. Thus, the germs that do not vanish at x are all invertible, while the complementary set, consisting of germs that do vanish at x, is an ideal. This ideal is clearly the unique maximal ideal in the ring of germs. The ring of germs clearly reflects only geometry "near x." It makes sense to think of this as a "local" ring.

An entirely similar construction can be made for C^∞ $\mathbb{R}$-valued functions defined on an open set containing a point x of a C^∞ manifold. The rings of germs is again a ring with a unique maximal ideal, which consists of the germs that vanish at x. One can make an entirely analogous construction of a ring of germs at a point for other sorts of differentiable manifolds, where a different level of differentiability is assumed. These are all quasilocal rings.

If X is $\mathbb{C}^n$ (or an analytic manifold — there are also more general kinds of analytic sets) the ring of germs of holomorphic $\mathbb{C}$-valued functions on an open set containing x again has a unique maximal ideal consisting of the functions that vanish at x. In the case of the origin in $\mathbb{C}^n$, the ring of germs of holomorphic functions may be identified with the convergent power series in n variables, i.e., the power series that converge on a neighborhood of the origin. This ring is even Noetherian (this is not obvious), and so is a local ring in our terminology, not just a quasilocal ring.

We now return to the problem of comparing ideals in R with those in $S^{-1}R$. Given any ring homomorphism $f : R \to T$ we may make a comparison using two maps of ideals that always exist. Given an ideal $I \subseteq R$, IT denotes the ideal of T generated by the image of I, which is called the *expansion* of I to T. The image of I is not usually an ideal. One must take T-linear combinations of images of elements of I. For example, if we consider $\mathbb{Z} \subseteq \mathbb{Q}$, then $2\mathbb{Z}$ is a proper ideal of $\mathbb{Z}$, but it is not an ideal of $\mathbb{Q}$: the expansion is the unit ideal. The highly ambiguous notation I^{e} is used for the expansion of I to T. This is sometimes problematic, since T is not specified and their may be more than one choice. Also, if e might be denoting an integer, I^{e} might be taken for a power of I. Nonetheless, it is traditional, and convenient if the choice of T is clear.

If J is an ideal of T, we have already mentioned, at least in the case of primes, that $f^{-1}(J) = \{r \in R : f(r) \in J\}$ is denoted J^{c} and called the *contraction* of J to R. This notation has the same sorts of flaws and merits as the notation above for expansions. It is always the case that f induces an injection of $R/J^{\mathrm{c}} \hookrightarrow T/J$. It is trivial that $I \subseteq (I^{\mathrm{e}})^{\mathrm{c}} = I^{\mathrm{ec}}$, the contracted expansion, and that $J^{\mathrm{ce}} = (J^{\mathrm{c}})^{\mathrm{e}} \subseteq J$.

We now want to consider what happens when $T = S^{-1}R$. In this case, in general one only knows that $I \subseteq I^{\mathrm{ec}}$, but one can characterize I^{ec} as $\{r \in R : \text{for some } s \in S, sr \in I\}$. We leave this as an exercise. On the other hand, if $J \subseteq S$ is an ideal, $J = J^{\mathrm{ce}}$. That is, every ideal of $S^{-1}R$ is the expansion of its contraction to R. The reason is quite simple: if $r/s \in J$, then $r/1 \in J$, and r will be in the contraction of J. But then $r(1/s) = r/s$ will be in the expanded contraction. Call an ideal $I \subseteq R$ *contracted* with respect to the multiplicative system S if whenever $s \in S$ and $sr \in I$ then $r \in I$. Expansion and contraction give a bijection between ideals of R contracted with respect to S and ideals of $S^{-1}R$.

Lecture of September 20

Notice that the algebra map $R \to S^{-1}R$ provides a simple way of getting from modules over $S^{-1}R$ to R-modules: in fact, whenever one has any R-algebra T with structural homomorphism $f : R \to T$, a T-module M becomes an R-module if we define $r \cdot m = f(r)m$. This gives a covariant functor from T-modules to R-modules, and is referred to as *restriction of scalars.* The functions that give homomorphisms literally do not change at all, nor does the structure of each module as an abelian group under $+$.

A sequence of modules

$$\cdots \to M' \xrightarrow{\alpha} M \xrightarrow{\beta} M'' \to \cdots$$

is said to be *exact* at M if the image of α is equal to the kernel of β. A functor from R-modules to T-modules is called *exact* if it preserves exactness: the functor may be either covariant or contravariant. Restriction of scalars is an exact functor. Later, we shall consider the problem of making a transition (i.e., of defining a functor) from R-modules to T-modules when T is an R-algebra. This is more difficult: one makes use of tensor products, and the functor one gets is no longer exact.

It is easy to see that $S^{-1}R = 0$ iff $0 \in S$ iff some nilpotent element is in S. The issue is whether 1 becomes equal to 0 after localization, and this happens if and only if $s \cdot 1 = 0$ for some $s \in S$.

Prime ideals of $S^{-1}R$ correspond bijectively, via expansion and contraction, with primes of R that do not meet S. The key point is that if P is a prime not meeting S, it is automatically contracted with respect to S: if $su \in P$ with $s \in S$, then since $s \notin P$, we have that $u \in P$. The primes that do meet S all expand to the unit ideal.

In particular, when $S = R - P$, for P prime, the prime ideals of $R_P = (R_P)^{-1}R$ correspond bijectively with the prime ideals of R that are contained in P (this is equivalent to not meeting $R - P$) under contraction and expansion. This implies that PR_P is the unique maximal ideal of R_P, which was asserted earlier without proof. In particular, R_P is quasilocal.

It is straightforward to show that the map $\mathrm{Spec}\,(S^{-1}R)$ to $\mathrm{Spec}\,(R)$ is a homeomorphism of $\mathrm{Spec}\,(S^{-1}R)$ with $Y \subseteq \mathrm{Spec}\,(R)$ where

$$Y = \{P \in \mathrm{Spec}\,(R) : S \cap P = \emptyset\}.$$

This has some similarities to the situation when one compares ideals of R and ideals of R/I. Expansion and contraction give a bijection between ideals J of R that contain I and ideals of R/I. The ideal J corresponds to $J(R/I)$, which may be identified with J/I. This bijection preserves the property of being prime, since R/J is a domain if and only if $(R/I)/(J/I) \cong R/J$ is a domain. Thus, the map $\mathrm{Spec}\,(R/I) \to \mathrm{Spec}\,(R)$ is a bijection of the former onto $V(I)$. It is easy to verify that it is, in fact, a homeomorphism of $\mathrm{Spec}\,(R/I)$ with $V(I)$.

The notation R_a is used for $S^{-1}R$ where $S = \{1, a, a^2, \dots\}$, the multiplicative system of all powers of a. If R is a domain we may think of this ring as $R[1/a] \subseteq L$, where L is the field of fractions of R.

The notation R_S for $S^{-1}R$ is in use in the literature, but we shall not use it in this course.

Suppose that S and T are two multiplicative systems in S. Let ST be the multiplicative system $\{st : s \in S,\, t \in T\}$. Note that the image of st has an inverse in an R-algebra if and only if both the images of s and of t have inverses. Let S' be the image of S in $T^{-1}R$ and T' be the image of T in $S^{-1}R$. Then ${T'}^{-1}(S^{-1}R) \cong (ST)^{-1}R \cong {S'}^{-1}(T^{-1}R)$. All three represent the functor from rings to sets whose value on a ring A is the set of homomorphisms from R to A such that the images of the elements of both S and T are invertible in A.

Let $\overline{S}$ be the image of S in R/I, and use bars over elements to indicate images modulo I. Then $S^{-1}R/I^{\mathrm{e}} \cong \overline{S}^{\,-1}(R/I)$. The isomorphism takes the class of r/s to $\overline{r}/\overline{s}$. Both represent the functor from rings to sets whose value on T is the set of ring homomorphisms $g : R \to T$ that kill I and such that for all $s \in S$, $g(s)$ is invertible in T.

In the case of a prime ideal P, one has in particular that R_P/PR_P is the localization of the domain R/P at the multiplicative system of all nonzero elements (this is the image of $R - P$), which is the same as the fraction field of the domain R/P.

If S is a multiplicative system that does not meet I, then $S^{-1}R/I^{\mathrm{e}}$ has maximal ideals. Their contractions to R are certainly prime: they are precisely the ideals of R that contain I and are maximal with respect to not meeting S. Thus, if S does not meet I, there is a prime ideal of R that contains I and does not meet S.

In particular, if $a \in R$ is not nilpotent, then the multiplicative system of powers of a does not contain 0, and there is a prime that does not meet this multiplicative system. In particular, there is a prime ideal of R that does not contain a. From this we see at once:

Corollary. *The intersection of the prime ideals of R is the ideal of all nilpotent elements of R.* □

Corollary. *The intersection of all the prime ideals of R that contain I is the same as the ideal $\{a \in R : \text{for some } n \geq 1, a^n \in I\}$. This ideal is called the radical of I.*

Proof. If $a^n \in I$ and P is prime with $I \subseteq P$, then $a^n \in P$ and so $a \in P$. If $a^n \notin I$ for all n then the image of a is not nilpotent in R/I. Therefore some prime ideal P/I of R/I does not contain a. But this means that P is a prime ideal of R that contains I but not a. More briefly put, simply apply the immediately preceding Corollary to R/I. □

The radical of I is denoted $\mathrm{Rad}\,(I)$ or $\sqrt{I}$.

A ring is called *reduced* if the only nilpotent element is 0, i.e., if the radical of the ideal (0) is the ideal (0). If N denotes the ideal of all nilpotent elements of R, then R/N is called R *reduced*, and denoted R_{red}. Notice that $\mathrm{Spec}\,(R/N) \to \mathrm{Spec}\,(R)$ has image $V(N)$, i.e., the map is surjective as well as injective and is, in fact a homeomorphism.

The intersection of a nonempty chain of prime ideals is easily verified to be a prime ideal (the same is true for the union, by the way). By Zorn's lemma, every prime ideal of R contains a minimal prime ideal of R, one that does not contain any strictly smaller prime ideal. It follows that the intersection of all prime ideals of R is the same as the intersection of the minimal prime ideals of R.

Corollary. *The intersection of the minimal prime ideals of R is the ideal of all nilpotent elements of R.*

A prime that is minimal in the partially ordered set $V(I)$ is called a *minimal prime* of I. We also have:

Corollary. *The intersection of the minimal primes of I is* $\mathrm{Rad}\,(I)$*.*

Proof. Apply the preceding Corollary to R/I. $\square$

Note that the ideal of nilpotents in R is not necessarily a prime ideal: R may have many minimal primes. E.g., in $R = \mathbb{Z}/36\mathbb{Z}$, the ideal of nilpotents is generated by [6], and $R_{\mathrm{red}} \cong \mathbb{Z}/6Z$, which has two minimal primes, generated by the classes of 2 and 3 respectively.

More generally, suppose that T is a unique factorization domain and that $f_1, \ldots, f_n$ are irreducible elements of T generating mutually distinct prime ideals f_iT. Let $g = f_1^{k_1} \cdots f_n^{k_n}$ where the k_i are positive integers, and let $f = f_1 \cdots f_n$. Let $R = T/gT$. Then $R_{\mathrm{red}} \cong T/fT$, and there are n minimal primes: they are generated by the respective images of the f_i.

Corollary. *There is an order-reversing bijection between the closed sets of the Zariski topology on* $\mathrm{Spec}\,(R)$ *and the set of radical ideals of R.*

Proof. Every closed set has the form $V(II)$ for some ideal I, from the definition of the topology, and I may be replaced by its radical. Therefore, it suffices to see that the radical ideal I may be recovered from $V(I)$. This is clear, since it is the intersection of the primes in $V(I)$. $\square$

Lecture of September 22

I want to emphasize the difference between being finitely generated as an algebra and being finitely generated as a module. When S is finitely generated as an R-algebra it means that there are elements $s_1, \ldots, s_n \in S$ such that every element of S is an R-linear combination of finitely many monomials in the elements s_i. Each monomial has the form $s_1^{k_1} \cdots s_n^{k_n}$. The monomials include 1_S (when all the $k_i = 0$). When S is generated as an R-algebra by $s_1, \ldots, s_n$ there is no smaller ring that contains the image of R and all of the s_i. It also means that the the R-linear ring homomorphism of the polynomial ring $R[x_1, \ldots, x_n]$ to S that sends $x_i \mapsto s_i$ for every i is surjective. Note that in the polynomial ring $R[x, y]$ the module generated by 1, x, y is just $R+Rx+Ry$: it is missing all monomials of higher degree. If $s_1, \ldots, s_n$ generate S as an R-module, then every element of S can be written in the form $r_1s_1 + \cdots + r_ns_n$: there are no higher degree monomials in the s_j in the representation. When this happens, it is always true that the s_i generate R as an S-algebra as well, i..e., generators of S as an R-module always generate S as an R-algebra.

The ring $Z[1/2]$ is finitely generated as a $\mathbb{Z}$-algebra by $1/2$. It contains $1/2^k$ for every integer k. But it is not finitely generated as a $\mathbb{Z}$-module: any finitely generated submodule consists of fractions whose denominators can be simultaneously cleared by a single power of 2.

The polynomial ring in infinitely many variables over R is not finitely generated over R: any purported finite set of generators only involves finitely many of the variables, and the other variables cannot be obtained.

The field of rational numbers $\mathbb{Q}$ is not finitely generated as a $\mathbb{Z}$-algebra: any finitely generated subalgebra contains fractions involving only finitely many primes in the denominator (those occurring in the denominators of the generators), and will not contain the reciprocals of other primes.

The ring $\mathbb{Z}[\sqrt{2}]$ is finitely generated over $\mathbb{Z}$ not only as an algebra but also as a $\mathbb{Z}$-module. Every element can be written in the form $a + b\sqrt{2}$, where a and b are integers, and so 1, $\sqrt{2}$ generate it as a $\mathbb{Z}$-module.

Let X be any non-empty set, let K be a field, and let R be the ring of functions from X to K. We shall assume the result from the first problem set that asserts a bijection of ideals of R with filters on X. We want to observe that every prime ideal of R is maximal. Suppose that $\mathcal{F}$ is the filter corresponding to a prime ideal P. If Y and Y' are complementary subsets of X, i.e., if $Y \cap Y' = \emptyset$ while $Y \cup Y' = X$, let f be the function that is 1 on Y and 0 on Y' and and let g be the function $1_R - f$, which is 0 on Y and 1 on Y'. Then $fg = 0$, so that either $f \in P$ or $g \in P$. Thus, for every subset Y of X, either Y or $X - Y$ is already in $\mathcal{F}$, but not both, since $\emptyset \notin \mathcal{F}$. But this implies that $\mathcal{F}$ is maximal: it cannot be contained in a larger filter that does not contain $\emptyset$, for if $Y \notin \mathcal{F}$ then $X - Y \in \mathcal{F}$, and a filter that contains both Y and $\mathcal{F}$ must contain $Y \cap (X - Y) = \emptyset$. This shows that every prime ideal of R is maximal. But then every prime ideal of R is minimal! If X is infinite, the set of minimal primes is infinite — there is at least one for every point of X, the functions that vanish at that point. But further analysis of the case where X is infinite shows that the number of minimal primes is uncountable, even when X is countable: they correspond to the ultrafilters, which are the points of the Stone-Cech compactification.

The following question is problem 14. on p. 308 in an undergraduate abstract algebra text, *Abstract Algebra* by W. E. Deskins, MacMillan, New York, 1964. The first part of the problem asks the reader to show that if $R \cong S$ as rings, then $R[x] \cong S[x]$, where these are polynomial rings in one variable. This is easy. The second part of this question asks whether the converse is true. (Deskins was a professor at Michigan State University, by the way.) I have wondered on many occasions whether Deskins knew the answer. To avoid mistakes it may be better to ask, if $R[x] \cong S[y]$ is $R \cong S$? I have changed the letters to emphasize that an isomorphism between $R[x]$ and $S[y]$ might not take x to y. If it does, then one does immediately get an induced isomorphism $R[x]/xR[x] \cong S[y]/yS[y]$, and this implies that $R \cong S$. Without the extra hypothesis, the problem does not seem easy to me. But I have sometimes been wrong about such things. What do you think? Can you prove that $R \cong S$ or give a counterexample? The question remains difficult (I think) even if both rings are assumed to be finitely generated algebras over a field, and the isomorphism is assumed to preserve the field.

.

We now return to the study of the properties of the prime spectrum of a commutative ring.

Theorem. *If R is any commutative ring, then $X = \mathrm{Spec}\,(R)$ is quasicompact.*

Proof. . Let $\mathcal{I}$ be a family of ideals in R. Then $\mathcal{U} = \{X - V(I) : I \in \mathcal{I}\}$ is an open cover of X iff there is no prime P that is in all of the $V(I)$, i..e, that contains all the $I \in \mathcal{I}$. This mean that $\mathcal{U}$ is an open cover iff the I in $\mathcal{I}$ generate the unit ideal, which implies that $1 = i_1 + \cdots + i_n$ for some choice of $I_1, \ldots, I_n$ in $\mathcal{I}$ and $i_1 \in I_1, \ldots, i_n \in I_n$. But this means that the $\{X - V(I_i)\colon 1 \le i \le n\}$ is a finite subcover of $\mathcal{U}$. □

We write $D(A) := X - V(A)$ for $A \subseteq R$ and $D(a) := X - V(a)$, the open set of primes not containing $a \in R$. We have that

$$X - V(I) = \bigcup_{a \in I} D(a),$$

so that every open set is a union of sets $D(a)$, i.e., the sets $D(a)$ are a base for the Zariski topology. (A family of sets is a base for the open sets of a topology if the open sets coincide with the unions, finite and infinite, of the sets in the base.) Moreover, $D(a) \cap D(b) = D(ab)$, so this base is closed under finite intersection. Since $D(a) \approx \mathrm{Spec}\,(R_a)$, every open set of the form $D(a)$ is quasicompact, and this means that the quasicompact open subsets of X form a base. Each quasicompact open set will be a union of sets $D(a)$, and since it is quasicompact, the union can be taken to be finite. A finite union of quasicompact sets is quasicompact, and so the quasicompact open sets are precisely the sets that are finite unions of sets of the form $D(a)$. It follows that the intersection of two quasicompact open subsets is quasicompact and open.

A non-empty topological space X is called *irreducible* if it is not the union of two proper closed subsets, which is equivalent to the assumption that it is not the union of finitely many proper closed subsets. Another equivalent statement is that any two nonempty open sets meet, and this in turn is equivalent to the property that every nonempty open set is dense. (Check all these equivalences.)

This does not happen much in Hausdorff spaces: an irreducible Hausdorff space has exactly one point. If there were two, they would have disjoint open neighborhoods U and V, and the complements would be proper closed sets whose union is the whole space. But there are, typically, irreducible sets in $\mathrm{Spec}\,(R)$. A topological space X is said to have a *generic point* x if there is a point x such that the closure of $\{x\}$ is all of X. That is, $\{x\}$

is dense! Said yet another way, every nonempty open set contains x. In a T_0 space like $\operatorname{Spec}(R)$, a generic point, if it exists, is unique. If X has a generic point, it is irreducible: every non-empty open set contains the generic point, and is therefore dense.

The converse is true in $\operatorname{Spec}(R)$.

Proposition. $\operatorname{Spec}(R)$ *is irreducible if and only if the ideal of all nilpotents* N *is prime, in which case* N *is the unique minimal prime of* R*, and is consequently a generic point for* $\operatorname{Spec}(R)$*.* R *is a domain if and only if it is reduced and* $\operatorname{Spec}(R)$ *is irreducible. In* $\operatorname{Spec}(R)$*,* $V(I)$ *is irreducible if and only if the radical* P *of* I *is prime, in which case* $V(I) = V(P)$ *has a generic point, namely,* P*.*

Proof. The issues raised in the first and second sentences are really the same, since killing the ideal of nilpotents N does not affect the question: $\operatorname{Spec}(R/N) \approx \operatorname{Spec}(R)$. Likewise, the statement about $V(I)$ follows from applying the first two statements to R/I. We may therefore assume that R is reduced. If R is a domain, then it is clear that (0) is the unique minimal prime ideal of R. Now suppose instead that $\operatorname{Spec}(R)$ is irreducible, but that R is not a domain. Choose nonzero elements $a, b \in R$ such that $ab = 0$. Then every prime ideal contains a or contains b, and so $\operatorname{Spec}(R)$ is the union of the closed sets $V(a)$ and $V(b)$. Since neither a nor b is nilpotent, both of these closed sets are proper closed sets. This contradicts the assumption that $\operatorname{Spec}(R)$ is irreducible. □

We have already observed that there is an order-reversing bijection between the radical ideals of R and the closed sets in $\operatorname{Spec}(R)$. This bijection restricts to give an order-reversing bijection between $\operatorname{Spec}(R)$, the set of prime ideals of R, and the irreducible closed subsets of $\operatorname{Spec}(R)$. The prime ideal P corresponds to $V(P)$.

Note also that:

$V(I) = V(I')$ if and only if
I and I' are contained in all the same primes if and only if
the primes in $V(I)$ have the same intersection as those in $V(J)$ if and only if
the radical of I and the radical of J are equal.

Putting all this together, we now know the following about $\operatorname{Spec}(R)$: it is a quasicompact T_0 space in which the quasicompact open sets are closed under finite intersection and form a base for the topology. Moreover, every irreducible closed subset has a generic point. The converse is true, i.e., every topological space with these properties occurs as $\operatorname{Spec}(R)$ for some commutative ring R. See [M. Hochster, *Prime ideal structure in commutative rings*, Trans. of the Amer. Math. Soc. **142** (1969) 43–60].

Lecture of September 25

An infinite product indexed by a set may be thought of as functions from the index set such that the value at an element u of the index set is taken in the factor corresponding to u. When the product consists of rings, a ring structure is introduced using coordinate-wise addition and multiplication.

Let Y denote either the set of all primes of R or the set of minimal primes of R. Suppose that R is reduced, so that the intersection of the primes in Y is the zero ideal. There is a ring homomorphism $R \to \displaystyle\prod_{P \in Y} R/P$ that sends the element $r \in R$ to the element in the product whose P-coordinate is the image $r + P$ of r in R/P for all $P \in Y$. Since the intersection of the primes in Y is (0), this homomorphism is injective. We therefore have:

Corollary. *R is reduced if and only if it is isomorphic with a subring of a product (which may be infinite) of integral domains.* $\square$

Each of these integral domains R/P may be enlarged to field, $\mathrm{frac}\,(R/P)$, where $\mathrm{frac}\,(D)$ denotes the fraction field of the integral domain D. Thus, in the Corollary, we can replace "integral domain" by "field." A ring is reduced if and only if it is a subring of a product of fields.

The partial ordering of the prime ideals of R can be recovered from the topology of $\mathrm{Spec}\,(R)$, because $P \subseteq Q$ if and only if Q is in the closure of $\{P\}$ if and only if the closure of P contains the closure of Q. We may also recover the partially ordered set of primes under $\subseteq$ as the poset of irreducible closed subsets under $\supseteq$.

The next segment of the course will deal with the interactions between the notion of an integral extension of a ring and the theory of Krull dimension.

We shall use $\subset$ to indicate *strict* containment of sets. Let $P_0 \subset P_1 \subset \cdots \subset P_d$ be a chain of primes in a ring R. By the *length* of the chain we mean the integer d. This is the number of strict inclusions and is one smaller than the number of distinct prime ideals in the chain. By the *Krull dimension* of the ring R we mean the supremum of lengths of finite strictly ascending chains of prime ideals of R. Note that this is the same as the supremum of lengths of finite strictly descending chains of irreducible closed sets in $\mathrm{Spec}\,(R)$. (This is not so different from one characterization of dimension of a finite-dimensional vector space: it is the supremum of lengths of chains of strictly descending vector subspaces.) It may be $+\infty$. We need a convention for the case where R is the 0 ring: in that case we are taking the least upper bound of an empty set of integers. We make the convention that the dimension is -1 in that case. Another possible convention would be to make the dimension $-\infty$.

Note that a ring has dimension 0 if it is nonzero and any two distinct prime ideals are incomparable. The latter condition is equivalent to the condition that every prime ideal is maximal, and also to the condition that every prime ideal is minimal. A field has dimension 0. A principal ideal domain that is not a field has dimension 1: the ideal (0) is the unique minimal prime. All other prime ideals are maximal.

In exploring the notion of dimension, we shall prove that every ring that is finitely generated over a field or a PID has finite dimension. We shall prove that every local ring (Noetherian quasilocal ring) has finite dimension. In both these cases we shall characterize dimension in other ways. We shall show that the polynomial ring in n variables over a field has dimension n.

There exist Noetherian rings of infinite Krull dimension. They do not arise readily: one has to work at giving an example.

An important tool in the study of dimension is the theory of integral ring extensions. We shall also use this theory to prove Hilbert's Nullstellensatz.

Let S be an R-algebra with structural homomorphism $f : R \to S$. An element $s \in S$ is called *integral* over R if for some positive integer d we have that

$$s^d = r_{d-1}s^{d-1} + \cdots + r_1 s + r_0 \cdot 1_S$$

for suitable elements r_j of r, i.e., $s^d \in Rs^{d-1} + \cdots + R1_S$. If we multiply by s, we see that s^{d+1} is in the R-span of $s^d, \ldots, 1_S$, and s^d is not needed, because it is in the R-span of its predecessors. Thus s^{d+1} is in the R-span of $s^{d-1}, \ldots, 1_S$. We may continue in this way to prove by a straightforward induction that s^t is in the R-span of $s^{d-1}, \ldots, 1_S$ for all t.

Thus, the fact that s is integral over R is equivalent to the assertion that the R-submodule of S spanned by the powers of s (included 1_S as the 0 th power) is finitely generated. (Note that any set of generators will involve only finitely many powers of s, and that these powers of s will lie among the elements $s^{d-1}, \ldots, 1$ for any $d \gg 0$.) Let A denote the image of R in S. Then another equivalent statement is that the ring $A[s]$ is a finitely generated A-module, and yet another is that s satisfies a monic polynomial (i.e., one with leading coefficient 1) with coefficients in A, say $s^d + a_{d-1}s^{d-1} + \cdots + a_1 s + a_0 = 0$ where every a_i has the form $f(r_i)$ for some element $r_i \in R$. From this definition, it is clear that s is integral over R if and only if it is integral over the image $A = f(R)$ of R in S. Thus, questions about integrality reduce, for the most part, to the case where $R \subseteq S$, and we usually assume this without much comment in the proofs.

Note that $1/2$ is not integral over $\mathbb{Z}$: its dth power is not a $\mathbb{Z}$-linear combination of lower powers for any d. On the other hand in $\mathbb{Z}[\sqrt{2}]$ the element $\sqrt{2}$ is integral over $\mathbb{Z}$: it satisfies the monic polynomial equation $x^2 - 2 = 0$. Note that $\mathbb{Z}[\sqrt{2}] = \mathbb{Z} + \mathbb{Z}\sqrt{2}$ is spanned over $\mathbb{Z}$ by 1 and $\sqrt{2}$.

S is said to be *integral over* R if every element of S is integral over R. If $R \subseteq S$ and S is integral over R then S is called an *integral extension* of R. S is said to be module-finite over R if S is finitely generated as an R-module. This is much stronger than the requirement that S be finitely generated as an R-algebra. If $R \subseteq S$ and S is module-finite over R, then S is called a *module-finite extension* of R. We want to explore the connection between module-finite extensions and integral extensions.

We need to extend aspects of the theory of determinants to arbitrary commutative rings. If (r_{ij}) is an $n \times n$ matrix with entries in R, we define

$$\det{(r_{ij})} = \sum_{\pi \in S_n} \mathrm{sgn}\,(\pi) r_{1,\pi(1)} r_{2,\pi(2)} \cdots r_{n,\pi(n)}$$

where S_n is the set of permutations of $\{1, 2, \ldots, n\}$ and $\mathrm{sgn}\,(\pi)$ is 1 if π is an even permutation -1 if π is an odd permutation.

Certain facts about determinants follow from polynomial identities in the entries. To prove them for any ring, it suffices to prove them for polynomial rings over the integers, and since the problem remains the same if we think over the fraction field, we see that it is enough to prove the result over a field of characteristic 0. For example, suppose we want to prove that A and its transpose have the same determinant. If one knows this when A is matrix of indeterminates over $\mathbb{Z}$, one gets the general case by taking a homomorphism from $\mathbb{Z}[x_{ij}] \to R$ that maps x_{ij} to r_{ij} for all choices of i, j. The result that $\det(AB) = \det(A)\det(B)$ can be proved similarly: one starts with the case where A and B are two matrices of indeterminates. One can similarly prove that if two rows (or columns) are identical the determinant is 0, and that switching two rows or columns reverses the sign.

Let A_{ij} denote the submatrix of A obtained by deleting the i th row and j th column. The determinant of A_{ij} is called the i, j minor of A, and $(-1)^{i+j}\det(A_{ij})$ is called the i, j *cofactor*. The *classical adjoint* of A is the matrix whose i, j entry is the j, i cofactor of A: it is also referred to as the transpose of the cofactor matrix. We denote it $\mathrm{adj}(A)$. The determinant of a matrix can be found by multiplying each element of the i th row by its cofactor and summing: this called *expansion by minors* with respect to the i th row. There is a similar expansion with respect to any column. Then $A\ \mathrm{adj}(A) = \det(A)I_n$, where I_n is the $n \times n$ identity matrix. Each entry of the product on the left is the determinant of a matrix obtained by expanding with respect to a row. If the entry is off diagonal, the matrix whose determinant is being expanded has two rows equal. If the entry is on the diagonal, one gets one of the expansions for $\det(A)$ by minors. A similar argument using columns shows that $\mathrm{adj}(A)\ A = \det(A)I$.

These results are valid for any commutative ring R. If the case of a field of characteristic 0 is taken as known, they can be deduced from that case by the type of argument discussed above, using maps of polynomial rings.

Here is another illustration of this principle. We can prove the Cijayley-Hamilton theorem, that a matrix satisfies its characteristic polynomial, is valid over any commutative ring with identity. It suffices to show this for an $n \times n$ matrix of indeterminates $X = (x_{ij})$ over the integers, working in the domain $\mathbb{Z}[x_{ij}]$, and since we may think over the fraction field F of this domain, the result follows for all commutative rings is one knows the field case: we can map $\mathbb{Z}[x_{ij} : i, j]$ to the ring in which we are interested so that the xij map to the entries of a given matrix. Without assuming the field case, we can now prove the theorem as follows: assume, for the moment, that over an algebraic closure of the field F the matrix X has n distinct eigenvalues. Then the matrix is diagonalizable, and it suffices to check that a similar diagonal matrix satisifies its characteristic polynomial. But this follows from the fact that all the eigenvalues satisfy the characteristic polynomial of the matrix.

It remains to prove that the eigenvalues of X are distinct. For an arbitrary matrix over a domain, one may test whether the eigenvalues are distinct over an algebraic closure of the fraction field as follows. Consider the product of the squares of the differences of the eigenvalues. This is a symmetric function of the eigenvalues, and so may be formally expressed as a polynomial in the elementary symmetric functions of the eigenvalues. These are the same as the coefficients of the characteristic polynomial, and so are polynomials over $\mathbb{Z}$ in the entries of the matrix. Hence, there is a polynomial D in the variables x_{ij} with integer coefficients that has the following property: if one substitutes elements a_{ij} of a domain A for the x_{ij}, then the matrix (a_{ij}) has distinct eigenvalues over the algebraic closure of the fraction field of A if and only if $D(a_{ij}) \neq 0$. It now follows that $D(x_{ij}) \neq 0$ for our matrix of indeterminates X. In fact, even if we replace the off-diagonal entries x_{ij}, $i \neq j$ by 0, we get a matrix that evidently has the distinct eigenvalues $x_{11}, \, \dots \, , {}_{nn}$.

Note that if M is an R-module, then M^n, written as column vectors of size n, is a left module over the ring of $n \times n$ matrices over R: in particular, if A, B are $n \times n$ matrices over R and $V \in M^n$, $(AB)M = A(BM)$.

The fact that for an $n \times n$ matrix A over a commutative ring R one has $\mathrm{adj}(A)\ A = \det(A)I_n$ has the following consequence: om

Lemma. *Let $A = (r_ij)$ be an $n \times n$ matrix over R and let V be an $n \times 1$ column matrix such that $AV = 0$. Then $\det(A)$ kills every entry of V, i.e., $\det(A)V = 0$.*

Proof. $\det(A)V = \det(A)I_nV = \mathrm{adj}(A)AV = \mathrm{adj}(A)0 = 0$. □

We note that if x is an indeterminate over the ring R and B is an $n \times n$ matrix over R, then $\det(xI_n - B) \in R[x]$ is a monic polynomial of degree n in x with coefficients in R. The product of the entries of the main diagonal provides a unique term of degree n in x, namely, x^n, while the product of any other n entries can involve x at most to the $n-1$st power. As in the case of elementary linear algebra, this polynomial is called the *characteristic polynomial* of the matrix B. We can now prove:

Theorem. *Let S be module-finite over the ring R. Then every element of S is integral over R.*

Proof. We may replace R by its image in S, and so assume that $R \subseteq S$. Let $s_1, \, \dots \, , s_n$ be a finite set of generators for S as an R-module. Since we may enlarge this set of generators as we please, we may assume that $s_1 = 1$. Let $s \in S$ be any element. Then for every i we have an equation

$$ss_i = \sum_{j=1}^{n} r_{ij} s_j$$

with coefficients r_{ij} in R, simply because ss_j is some element of S and so can be written as an R-linear combination of elements of $s_1, \, \dots \, , s_n$. Let I_n be the $n \times n$ identity matrix, let V be the $n \times 1$ column vector whose entries are $s_1, \, \dots \, , s_n$, and let $B = (r_{ij})$. Then these equations can be written in matrix form as $sIV = BV$ or $(sI - B)V = 0$. Applying the preceding Lemma with $A = sI - B$, we find that $\det(sI - B)$ kills all the entries of V, one of which is $s_1 = 1$, and so $\det(sI - B) = 0$. This implies that s is a root of the characteristic polynomial of B over R, and so s is integral over R. □

Proposition. *Let $R \to S \to T$ be ring homomorphisms such that S is module-finite over R with generators $s_1, \ldots, s_m$ and T is module-finite over S with generators $t_1, \ldots, t_n$. Then the composition $R \to T$ is module-finite with the mn generators $s_i t_j$, $1 \leq i \leq m$, $1 \leq j \leq n$.*

Proof. Every element of t can be written as $\sum_{j=1}^{n} \sigma_j t_j$ for suitable elements $\sigma_j \in S$, and each σ_j can be written as $\sum_{i=1}^{m} r_{ij} s_i$ for suitable elements r_{ij} of R. Substituting in the expression for t shows that the elements $s_i t_j$ span T as an R-module. □

Corollary. *The elements of S integral over R form a subring of S.*

Proof. Replace R by its image in S and so assume $R \subseteq S$. Let s, s' be elements of S integral over R. Then $R[s]$ is module-finite over R and, since s' is integral over R it is certainly integral over $R[s]$: use the same monic polynomial to see this. Thus, $(R[s])[s'] = R[s, s']$ is module-finite over $R[s]$, and so, by the preceding Corollary, it is module-finite over R. Thus, $s \pm s'$ and ss', which are in $R[s, s']$, are integral over R. □

This depends on the characteristic polynomial method that was used to prove the Theorem above. A bit of further analysis of the proof shows that if s, s' satisfy monic polynomial equations of degrees m and n over R, the every element of $R[s, s']$ satisfies a monic polynomial equation of degree mn over R. It can be shown that, in general, one cannot do better.

If F is a finite algebraic field extension of the rational numbers the elements of F that are integral over $\mathbb{Z}$ are referred to as the *algebraic integers* of F, and form a ring $\mathfrak{o}$. The study of such rings is the branch of mathematics known as *algebraic number theory.*

Lecture of September 27

We next observe:

Theorem. *Let S be an R-algebra. Then S is module-finite over R if and only if S is finitely generated as an R-algebra and integral over R. For S to be module-finite over R, it suffices if S is generated over R by finitely many elements each of which is integral over R.*

Proof. We have already seen that module-finite extensions are integral, and it is clear that they are finitely generated as R-algebras.

For the other half, it suffices to prove the final statement, and we may suppose that $R \subseteq S$ and that $S = R[s_1, \ldots, s_n]$. $R[s_1]$ is module-finite over R by one of our characterizations of when an element is integral, and S is module-finite over $R[s_1]$ by induction on n. The result now follows because a module-finite extension of a module-finite extension of R is module-finite over R. □

A poset is called *directed* if for any two elements x, y of the poset there exists z in the poset such that $x \leq z$ and $y \leq z$. A union of a family of sets, subgroups, submodules, subrings or subalgebras is called a *directed union* if any two of them are contained in a third: the underlying sets form a directed poset under $\subseteq$. Then any finite union of them is contained in one of them.

Corollary. *S is integral over R if and only if it is a directed union of module-finite extensions of R.*

Proof. "If" is clear, since every element of S will be in one of the module-finite extensions and therefore integral over R. For "only if," note that S is the directed union of its finitely generated R-subalgebras, each of which will be module-finite over R. □

Observe that $\mathbb{Z}[\sqrt{p} : p > 1 \text{ is prime}]$ is integral over $\mathbb{Z}$ but not module-finite (and hence not finitely generated as a $\mathbb{Z}$-algebra). In fact, adjoining the square roots of the several primes to even to $\mathbb{Q}$ does not introduce the square roots of any other primes. Similarly, if K is a field and x is an indeterminate, the ring $K[x^{1/2^n} : n \in \mathbb{N}]$ is integral over $K[x]$ but is neither module-finite nor finitely generated as an algebra over $K[x]$.

If $R \subseteq S$ are rings, a prime Q of S that contracts to a prime P of R is said to *lie over* P.

Lemma. *Let $R \subseteq S$ be domains and let $s \in S - \{0\}$ be integral over R. Then s has a nonzero multiple in R.*

Proof. Consider an equation of integral dependence for s on R of degree n. Since $s \neq 0$, we must have that one of the lower coefficients r_i is not 0: let h be the least value of i such that $r_h \neq 0$, so that $r_i = 0$ for $i < h < n$. Then the equation can be rewritten as $s^h(s^{n-h} + \cdots + r_{h+1}s + r_h) = 0$. Since $s \neq 0$ and S is a domain, we have that $s^{n-h} + \cdots + r_{h+1}s + r_h = 0$, so that $r_h = s(-s^{n-h_1} - \cdots - r_{h+1})$, which shows that r_h is a nonzero multiple of s in R. □

Theorem. *Let S be an integral extension of R, $I \subseteq R$ an ideal, and $u \in IS$. Then u satisfies a monic polynomial equation $u^n + i_1u^{n-1} + \cdots + i_{n-1}u + i_n = 0$ where $i_t \in I^t$ for $1 \leq t \leq n$.*

Proof. We have that $u = \sum_{t=1}^{n} s_t i_t$, with the $s_t \in S$ and the $i_t \in I$. We may therefore replace S by the smaller ring generated over R by u and the elements s_t. This ring is module-finite over R. Thus, there is no loss of generality in assuming that S is module-finite over R, with generators $s_1, \ldots, s_n$, and, as earlier, we may enlarge the set of generators so that we may assume that $s_1 = 1$. It is easy to see that $IS = Is_1 + \cdots + Is_n$, the set of linear combinations of $s_1, \ldots, s_n$ with coefficients in I: each element is for $i \in I$ and $s \in S$ has this form because each element of S is an R-linear combination of $s_1, \ldots, s_n$. If $u \in IS$, then every us_j is in IS, and so there are n equations

$$us_j = \sum_{t=1}^{n} i_{jk}s_k.$$

Let V be the $n \times 1$ column matrix with entries $s_1, \ldots, s_n$ and let B be the $n \times n$ matrix (i_{jk}). Then the same argument that we gave earlier shows that u satisfies the characteristic polynomial of B, which has the form

$$x^n + i_1x^{n-1} + i_2x^{n-2} + \cdots + i_n$$

where i_t is in $I^t \subseteq R$ for every t, $1 \leq t \leq n$. □

Lying over theorem. *Let S be an integral extension of R. Then for every prime P of R, there are primes of S that contract to P, and they are mutually incomparable. In particular, the map* $\mathrm{Spec}\,(S) \to \mathrm{Spec}\,(R)$ *is onto. For every ideal I of R, the contraction of IS to R is contained in* $\mathrm{Rad}\, I$*, and so if I is radical, $IS \cap R = I$.*

Proof. We prove the last statement first. Let $u \in IS \cap R$, Consider the monic equation that u satisfies given by the preceding theorem. After we substitute u for x, the leftmost term of the equation is u^n while the other terms are in I. This implies that $u^n \in I$ and so $u \in \mathrm{Rad}\, I$, as required.

In particular, if $I = P$ is prime then $R - P$ is a multiplicative system in $R \subseteq S$, and PS does not meet it, since $PS \cap R = P$. Therefore there is a prime ideal Q of S that contains PS and is disjoint from $R - P$. Since $P \subseteq PS$, we see that $Q \cap R = P$.

It remains only to show that two primes lying over $P \subseteq R$ cannot be comparable. Suppose to the contrary that $Q_0 \subset Q$ both lie over P in R. The trick here is to pass to $R/P \subseteq S/Q_0$. This extension is still integral: given $s \in S$, it satisfies a monic equation over R, and $s + Q$ satisfies the same equation with coefficients considered mod P. Now the nonzero prime ideal Q/Q_0 lies over the prime ideal (0) in R/P. Thus, it suffices to show that if $R \subseteq S$ are domains, then a nonzero prime ideal Q of S cannot lie over (0) in R. This is immediate from the preceding Lemma: any nonzero element of Q has a nonzero multiple in R. □

Example. The ring of functions from an infinite set X to $\mathbb{Z}/2\mathbb{Z}$ is integral over $\mathbb{Z}/2\mathbb{Z}$: every element satisfies $x^2 - x = 0$. It has uncountably minimal primes, mutually incomparable and all lying over (0) in $\mathbb{Z}/2\mathbb{Z}$.

Lecture of September 29

We give another proof of the lying over theorem that does not involve the eigenvalue trick. Suppose that $R \subseteq S$ is integral and that $P \in \mathrm{Spec}\,(R)$.

Quite generally, suppose $R \subseteq S$ is integral and let W be a multiplicative system in R. It is easy to check that $W^{-1}R \subseteq W^{-1}S$ and that the extension is still integral: $W^{-1}S$ is generated over $W^{-1}R$ by the elements $s/1$, and these satisfy monic polynomials over the image of R in $W^{-1}R$.

We apply this here with $W = R - P$. Let $S_1 = W^{-1}S$. If Q_1 is a prime of S_1 lying over PR_P, then the contraction Q of Q_1 to S will lie over P, since PR_P lies over P. Thus, we have reduced to the case where R is quasilocal with maximal ideal P. It now suffices to show that $PS \neq S$, for then any maximal ideal of S containing PS will be prime, and its contraction to R will contain the maximal ideal P but not 1, forcing the contraction to be P. Consider the family of ideals of R contained in P whose expansion to S is not all of S. This family contains (0), and the union of a chain in the family is again in the family: if $1 \in S$ is a linear combination of finitely many elements from the union, these elements will come from just finitely many of the ideals in the family, and will all lie in the largest of them. Therefore this family has a maximal element I. Consider $IS \cap R = J$. Then $I \subseteq J$, and we must have $J = I$ or else $JS \subseteq IS \neq S$ contradicts the maximality of I. Then $R/I \to S/IS$ is injective and still integral, and R/I is quasilocal. Therefore we may replace $R \subseteq S$ by $R/I \subseteq S/IS$. If $P = (0)$ we are done. If not, then choose $a \in P - \{0\}$. Then the maximality of I implies that $aS = S$ (or else we could have enlarged $I \subseteq R$ using a preimage of a). This means that there is an element b of S such that $ab = 1$. But b is integral over R, so that there is an equation

$$b^n = r_{n-1}b^{n-1} + r_{n-2}b^{n-2} + \cdots + r_1 b + r_0$$

Since $b = a^{-1}$, when we multiply both sides by a^{n-1} we get that

$$b = r_{n-1} + r_{n-2}a + \cdots + r_1 a^{n-2} + r_0 a^{n-1}$$

which shows that $a^{-1} = b \in R$. Thus, a has an inverse in R, contradicting the assumption that $a \in P - \{0\}$. □

Corollary (Going up theorem). *Let $R \hookrightarrow S$ be an integral extension and let*

$$P_0 \subset P_1 \subset \cdots \subset P_d$$

be a chain of prime ideals of R. Let Q_0 be a prime ideal of S lying over P_0. Then there is a chain of prime ideals

$$Q_0 \subset Q_1 \subset \cdots \subset Q_d$$

of S such that for all t, Q_t lies over P_t.

Proof. It suffices to construct $Q_1 \supset Q_0$ lying over P_1: the result then follows by a straightforward induction on d. Consider $R/P_0 \subseteq S/Q_0$. This is an integral extension, and P_1/P_0 is a prime ideal of R/P_0, so there is a prime ideal of S/Q_0 that lies over it: it will have the form Q_1/Q_0 for some prime ideal Q_1 of S. It is clear that $Q_0 \subset Q_1$, and it is easy to verify that that Q_1 lies over P_1 in R. □

Corollary. *If* $R \hookrightarrow S$ *is an integral extension then* $\dim R = \dim S$.

Proof. Let $Q_0 \subset \cdots \subset Q_d$ be a chain of prime ideals of S. Their contractions will give a chain of prime ideals of the same length in R: they will be distinct, because comparable primes cannot contract to the same prime ideal. This shows that $\dim S \leq \dim R$.

On the other hand, given a finite chain of primes in R, the going up theorem implies the existence of a chain of prime ideals of S of the same length, so that $\dim S \geq \dim R$. □

We next want to observe that if the functors h_X and h_Y are isomorphic, this yields and isomorphism $X \cong Y$. In fact, we prove much more. Note that any morphism $f : Y \to X$ gives a natural transformation T^f from h_X to h_Y: the needed map from $h_X(Z) = \mathrm{Mor}\,(X, Z) \to \mathrm{Mor}\,(Y, Z) = h_Y(Z)$ sends $g : X \to Z$ to $g \circ f$. Notice that the map $h_X(X) = \mathrm{Mor}\,(X, X) \to \mathrm{Mor}\,(Y, X) = h_Y(X)$ associated with T^f sends 1_X to f. The key point is that *every* natural transformation $T : h_X \to h_Y$ arises in this way, uniquely (uniqueness will be obvious, since we have already seen how to recover f from T^f). Given T, let $f = T_X(1_X) \in \mathrm{Mor}\,(Y,\, X)$. For all Z, $T_Z : h_X(Z) = \mathrm{Mor}\,(X, Z) \to h_Y(Z) = \mathrm{Mor}\,(Y, Z)$. Fix $g : X \to Z$. Then the definition of a natural transformation yields a commutative diagram:

$$\begin{tikzcd}
\mathrm{Mor}\,(X, X) \arrow[r, "h_X(g)"] \arrow[d, "T_X"'] & \mathrm{Mor}\,(X, Z) \arrow[d, "T_Z"] \\
\mathrm{Mor}\,(Y, X) \arrow[r, "h_Y(g)"'] & \mathrm{Mor}\,(Y, Z)
\end{tikzcd}$$

We can compute the image of 1_X in $\mathrm{Mor}\,(Y, Z)$ two ways. Using the top and right arrows, we get $T_Z(g \circ 1_x) = T_Z(g)$. Using the left and bottom arrows, we get $g \circ f$. Thus, $T_Z(g) = g \circ f$ always, which is exactly what we wanted to show. But then natural transformations $h_X \to h_Y$ and $h_Y \to h_X$ whose composition in either order is the identity natural transformation (from $h_X \to h_X$ or from $h_Y \to h_Y$ must come from morphisms $f : Y \to X$ and $f' : X \to Y$ whose composition in either order is the identity on X or Y.

A very useful consequence of this discussion is that the object representing a functor is unique, up to isomorphism. This establishes literally hundreds of isomorphisms. For example, if S is a multiplicative system in R with image $\overline{S}$ in R/I, the isomorphism $S^{-1}R/IS^{-1}R \cong \overline{S}^{-1}(R/I)$ is a consequence of the fact that both represent, in the category of rings, the functor that assigns to the ring T all homomorphisms from $R \to T$ such that I maps to 0 in T and S maps into the units of T.

Let $f : R \to S$ be a ring homomorphism, and let $f^* = \mathrm{Spec}\,(f) : \mathrm{Spec}\,(S) \to \mathrm{Spec}\,(R)$ be the usual map given by contraction. Let $Y = \mathrm{Spec}\,(S)$ and $X = \mathrm{Spec}\,(R)$. Given a map of sets $g : Y \to X$, and a point $x \in X$, the set $g^{-1}(x)$ is called the *fiber* of g over x: it is simply the set of points of Y that map to x. Thus, the fiber of the function $f^* = \mathrm{Spec}\,(f)$ over $P \in \mathrm{Spec}\,(R)$ is precisely the set of primes of S lying over P in R. This set of primes is homeomorphic with Spec of

$$(R - P)^{-1}S/P^{\mathrm{e}} \cong (\overline{R - P})^{-1}(S/PS),$$

where $\overline{R-P}$ is the image of $R-P$ in S/PS. The ring $(R-P)^{-1}S/P^{\mathrm{e}}$ is called the *fiber* of $R \to S$ over P. (This is really terminology from the theory of schemes, and the term *scheme-theoretic fiber* is also used.) Alternatively, it may be defined as the canonically isomorphic ring $(\overline{R-P})^{-1}(S/PS)$. Note that it is an S-algebra. Its primes correspond exactly to primes of S that contain PS and are disjoint from $R-P$, which is exactly the condition for them to lie over P in R. $(R-P)^{-1}S/P^{\mathrm{e}}$ is also an algebra over R_P/PR_P (which may be identified with fraction field of the domain R/P).

If $R \to S$ is integral (respectively, module-finite), then $R_P/PR_P \to (R-P)^{-1}S/P^{\mathrm{e}}$ is also integral (respectively, module-finite). Up to multiplication by elements coming from units of R, every element of the $(R-P)^{-1}S/P^{\mathrm{e}}$ comes from S, and for the image of an element of S we may use the same equation of integral dependence that it satisfied over R, taking the images of the coefficients in R_P/PR_P. In the case where S is spanned over R by $s_1, \ldots, s_n$, the images of $s_1, \ldots, s_n$ span $(R-P)^{-1}S/P^{\mathrm{e}}$ over R_P/PR_P.

We want to obtain a bound for the number of primes lying over P in the case of a module-finite extension.

We first prove two preliminary results.

Two ideals I, J of a ring R are called *comaximal* if $I+J=R$. Ideals $I_1, \ldots, I_n$ of R are called *pairwise comaximal* if for all $j \neq k$, $I_j + I_k = R$. Note that if $m_1, \ldots, m_n$ are mutually distinct maximal ideals of R, then they are pairwise comaximal.

We recall that the *product ideal* IJ is the ideal generated by all the elements ij for $i \in I$ and $j \in J$. Each element of IJ is a sum of the form $i_1j_1 + \cdots + i_kj_k$ for some positive integer k and elements $i_1, \ldots, i_k \in I$ and $j_1, \ldots, j_k \in J$.

Lemma (Chinese remainder theorem). *If $I_1, \ldots, I_n$ are pairwise comaximal in the ring R, then*

$$I_1 \cdots I_n = I_1 \cap \cdots \cap I_n.$$

Let $J = I_1 \cdots I_n$. The ideals

$$I_1I_2,\, I_3,\, \ldots,\, I_n$$

are also pairwise comaximal. Moreover, the map

$$R/J \to R/I_1 \times \cdots \times R/I_n$$

that sends $r+J$ to $(r+I_1, \ldots, r+I_n)$ is a ring isomorphism.

Proof. First consider the case where $n=2$. Choose $i_1 \in I_1$ and $i_2 \in I_2$ such that $i_1+i_2=1$. If $u \in I \cap J$ then $u = u \cdot 1 = u(i_1+i_2) = ui_1 + ui_2$. But $ui_1 \in I_1I_2$ because $u \in I_2$, and $ui_2 \in I_1I_2$ because $u \in I_1$. Thus, $u \in I_1I_2$. The map $R \to R/I_1 \times R/I_2$ that sends r to $(r+I_1,\, r+I_2)$ is a ring homomorphism that clearly has kernel $I_1 \cap I_2 = I_1I_2$. It therefore induces an injection $R/I_1I_2 \hookrightarrow R/I_1 \times R_2$. To see that this map is surjective, let $(r_1+I_1,\, r_2+I_2)$ in the image be given. Then $r_1i_2 + r_2i_1$ maps to this element: mod I_1, $r_1i_2 + r_2i_1 \equiv r_1 \cdot 1 + r_2 \cdot 0 \equiv r_1$, and the calculation mod I_2 is exactly similar.

To prove the second statement, it clearly suffices to show that I_1I_2 is comaximal with I_j for $j \geq 3$. Choose $i_1 \in I_1$ and $u \in I_j$ such $i_1 + u = 1$, and choose $i_2 \in I_2$ and $v \in I_j$ such that $i_2 + v = 1$. Multiply these equations. Then $i_1i_2 + i_1v + ui_2 + uv = 1$, and $i_1i_2 \in I_1I_2$ while $i_1v + ui_2 + uv \in I_j$.

The general case of the ring isomorphism now follows by induction on n. By the induction hypothesis,

$$R/J = R/\big((I_1I_2)I_3 \cdots I_n\big) \cong \big(R/(I_1I_2)\big) \times R/I_3 \times \cdots \times R/I_n$$

and $R/(I_1I_2) \cong R/I_1 \times R/I_2$ by the case $n = 2$ already established. □

If $R = \mathbb{Z}$, the principal ideals $a_1\mathbb{Z}$, ... $a_n\mathbb{Z}$ are pairwise comaximal if and only if the integers $a_1, \ldots, a_n$ are relatively prime in pairs, and we get the classical Chinese remainder theorem.

Theorem. *Let R be a reduced K-algebra that is module-finite over the field K. This simply means that R is a finite-dimensional vector space over K. Then R is a product of finite algebraic field extensions $L_1 \times \cdots \times L_n$ of K. R has n maximal ideals, the kernels of the n product projections $R \twoheadrightarrow L_i$, $1 \leq i \leq n$, and n, the number of maximal ideals, is at most the dimension of R as K-vector space.*

Proof. Since K has dimension 0 and R is integral over K, R has dimension 0. Thus, every prime ideal is maximal. Let $m_1, \ldots, m_h$ be any subset of the maximal ideals of R. By the Chinese remainder theorem, $R/(m_1 \cdots m_h) \cong R/m_1 \times \cdots \times R/m_h$. Let $L_i = R/m_i$. L_i is a field and finite-dimensional as a K-vector space, and so it is a finite algebraic extension of K. As a K-vector space, $R/m_1 \times \cdots \times R/m_h$ is the direct sum over K of the L_i, which shows that h is at most the K-vector space dimension of $R/(m_1 \cdots m_h)$, and therefore is also at most the K-vector space dimension of R. This means that the number of maximal ideals of R is at most the K-vector space dimension of R. Now suppose that $m_1, \ldots, m_n$ are all the maximal ideals of R. Since R is reduced, the intersection of the m_i is (0). Thus, $R \cong R/(0) \cong R/m_1 \times \cdots \times R/m_n$. □

Corollary. *Let S be module-finite over R with n generators. The number of prime ideals of S lying over a prime P of R is at most n.*

Proof. By our earlier remarks, we may replace $R \to S$ by $R_P/PR_P \to (R_P)^{-1}S/P^{\mathrm{e}}$, and n does not increase. But now $R = K$ is a field, and S is a finite-dimensional K-vector space of dimension at most n. Passing to S_{red} can only decrease its K-vector space dimension, while the number of prime ideals (which are all maximal) does not change, and now we may apply the preceding result. □

Lecture of October 2

If P is a prime ideal of R, by the *height* of P we mean the supremum of lengths of finite strictly ascending chains of primes contained in P. It is immediate that the height of P is the same as the Krull dimension of the quasilocal ring R_P. It should be clear that the dimension of R is the same as the supremum of heights of all prime ideals, and that this will be the same as the supremum of heights of all maximal ideals.

Corollary. *If $R \subseteq S$ is an integral extension and Q is a prime ideal of S lying over a prime P in R, then the height of P is bigger than or equal to the height of Q.*

Proof. A chain of distinct primes contained in Q will contract to a chain of distinct primes contained in P. □

A much harder problem is this: suppose that S is integral over R and we are given a chain

$$P_n \supset P_{n-1} \supset \cdots \supset P_0$$

of primes in R, and a prime Q_n of S lying over P_n. Can we find a chain

$$Q_n \supset Q_{n-1} \supset \cdots \supset Q_0$$

of S such that Q_i lies over P_i for every i? This turns out to need additional hypotheses even when R is a domain. In order to formulate the correct hypothesis on R needed here, we must discuss the notion of an integrally closed domain.

The set of elements of $S \supseteq R$ that are integral over R was shown earlier to be a ring. This ring is called the *integral closure of* R *in* S.

We shall say that a domain R is *integrally closed* or *normal* if every element of the fraction field of R that is integral over R is in R. The integral closure of a domain R in its fraction field is called the *the integral closure* or *normalization* of R.

A unique factorization domain is normal. To see this, suppose that a/b is a fraction integral over R but not in R. We may assume that it has been written in lowest terms, so that a and b have no common divisor other than units, and b is not a unit. If it satisfies the equation

$$(a/b)^d + r_{n-1}(a/b)^{d-1} + \cdots + r_0 = 0$$

with the $r_i \in R$ we may multiply through by b^d to get the equation

$$a^d + r_{n-1}a^{d-1}b + \cdots + r_0 b^d = 0.$$

Every term other than the leftmost is divisible by b, and so $b \,|\, a^d$. Any prime factor of b must divide a^d and therefore a, a contradiction, since a/b is in lowest terms. □

In particular, any principal ideal domain, as well as any polynomial ring over a field or a principal ideal domain, is normal.

If K is a field, $R = K[x^2, x^3]$ is not normal. $x = x^3/x^2$ is in the fraction field, and is integral over $K[x^2, x^3]$, since $z = x$ is a root of $z^2 - x^2 = 0$. The integral closure of R is $K[x]$.

The ring $\mathbb{Z}[\sqrt{5}]$ is not integrally closed. The element $\tau = \dfrac{1+\sqrt{5}}{2}$ is in the fraction field, and is integral, since it is a root of $x^2 - x - 1 = 0$. It is not obvious but not difficult to show that $\mathbb{Z} + \mathbb{Z}\tau$ is integrally closed, and is the integral closure of $\mathbb{Z}[\sqrt{5}]$. (Suppose that $a + b\sqrt{5}$ is integral over $\mathbb{Z}[\sqrt{5}]$ and hence over $\mathbb{Z}$, where $a,\, b \in \mathbb{Q}$. It follows that $a - b\sqrt{5}$

will satisfy the same monic polynomial over $\mathbb{Z}$ that $a + b\sqrt{5}$ does, and so is also integral over $\mathbb{Z}$. Adding, we find that $a + b\sqrt{5} + a - b\sqrt{5} = 2a$ is integral over $\mathbb{Z}$, and therefore in $\mathbb{Z}$. Thus, a is either k or $k + 1/2$, where k is an integer. By subtracting a suitable integer linear combination of $\sqrt{5}$ and τ, we get an element of the form $c\sqrt{5}$, integral over $\mathbb{Z}$, such that c is an rational. It will therefore suffice to show that if c is rational and $c\sqrt{5}$ is integral over $\mathbb{Z}$, then c is an integer. Write $c = m/n$ in lowest terms. Then $5c^2$ is rational and is integral over $\mathbb{Z}$ and therefore is an integer, i.e., $n^2 \mid 5m^2$. If $5 \mid n$ then it does not divide m, and this is impossible. If 5 does not divide n, then $n^2 \mid m^2$, so that c is a rational number whose square is an integer, and it follows that c is an integer. $\square$)

If $R \subseteq S$ are domains and R is a direct summand of S as an R-module, then R is normal whenever S is. For Suppose that $a, b \in R$, $b \neq 0$, but that a/b is integral over R. Then it is integral over S, and therefore $a/b = s \in S$, i.e., $a = bs$. But there is an R-linear map f from $S = R \oplus_R W$ (where W is an R-submodule of S) that kills W and is the identity on R. It follows that $a = f(a) = f(bs) = bf(s)$, and so $a/b = f(s) \in R$.

Let K be a field. Then the ring R generated over K by all monomials of degree d in $S = K[x_1, \ldots, x_n]$ is integrally closed: we shall show that it is a direct summand of $K[x_1, \ldots, x_n]$. Note that every monomial of degree divisible by d, say degree dk, is the product of k monomials of degree d. Let W be the K-span of all monomials whose degree is not divisible by d. The product of an element of R and an element of W is in W: when we multiply and distribute in all possible ways, we get a sum of terms each of which is the product of a monomial of degree divisible by d and a monomial of degree not divisible by d, and that product is in W. Thus, $S = R \oplus_R W$. If the number of variables is greater than one and $d > 1$, these rings are not unique factorization domains. For example, if $n = 2$ and $d = 2$, $S = K[x_1, x_2]$ and $R = K[x_1^2, x_1x_2, x_2^2]$. The fact that $(x_1x_2)^2 = (x_1^2)(x_2^2)$ shows that R is not a UFD.

We can now state the result we aim to prove:

Theorem (Going down theorem). *Let R be a normal integral domain, and let S be integral over R. Suppose that no nonzero element of R is a zerodivisor in S, i.e., that S is torsion-free as an R-module. Let*

$$P_n \supset P_{n-1} \supset \cdots \supset P_0$$

be a chain of primes in R, and let Q_n be a prime ideal of S lying over P_n. Then there is a chain of primes

$$Q_n \supset Q_{n-1} \supset \cdots \supset Q_0$$

of S such that Q_i lies over P_i for every i.

We need some preliminaries before we can prove this.

Proposition. *Let A be a ring and $A[x]$ the polynomial ring in one variable over A.*

(a) *If f and g are nonzero polynomials of $A[x]$ with degrees n and d and leading coefficients a and b respectively, then if either a or b is not a zerodivisor in A, the degree of fg is $d + n$ and its leading coefficient is ab. In particular, the conclusion holds if f or g is monic.*

(b) **(Division algorithm)** *Let g be any polynomial and f a monic polynomial in $R[x]$ of degree d. Then one can write $g = qf + r$, where $q, r \in A[x]$ and either $r = 0$ or the degree of r is $< d$. This representation is unique.*

(c) *Let $R \subseteq S$ be a ring extension and let f, g be as in (b), with f monic. Then g is a multiple of f in $R[x]$ if and only if it is a multiple of f in $S[x]$.*

Proof. It is clear that fg has at most one term of degree $d+n$, namely abx^{d+n}, with all other terms of lower degree, and that it has such a term provided that $ab \neq 0$, which is true if either a or b is not a zerodivisor. This proves part (a).

To prove existence in part (b), we perform long division in the usual way. To make this precise, first note that if $g = 0$ or has degree $< d$, we may take $q = 0$ and $r = g$. Otherwise, let ax^n be the highest degree term in g, where $a \neq 0$ is in R. Then $g_1 = g - ax^{n-d}g$ has smaller degree than f, and so can be written in the form q_1g+r by induction on the degree of f. But then $f = (ax^{n-d} + q_1)g + r$, as required.

It remains to prove uniqueness. But if $qf+r = q'f+r'$ both satisfy the condition, then $(q - q')f = r' - r$ is 0 or has degree smaller than that of f, which is impossible from part (a) unless $q - q' = 0$, in which case $r' - r = 0$ as well.

To prove part (c), note that we can perform the division algorithm thinking in $R[x]$ or in $S[x]$. By uniqueness, the result is the same. If g is a multiple of f in $S[x]$ the remainder must be zero, and then the same holds in $R[x]$. □

Note in connection with part (a) that if $A = \mathbb{Z}/(4)$ and $\overline{2}$ denotes the image of 2 in A, then $(\overline{2}x + 1)(\overline{2}x + 1) = 1$ in $A[x]$.

Lecture of October 4

Proposition. *Let R be an integrally closed domain with fraction field K and let S be a domain containing R. Suppose that $s \in S$ is integral over R. Let $f(x) \in K[x]$ be the minimal monic polynomial of s over K. Then $f(x) \in R[x]$, and for any polynomial $g(x) \in R[x]$ such that $g(s) = 0$, $f(x) \,|\, g(x)$ in $R[x]$.*

Proof. Choose an algebraically closed field L that contains the fraction field of S. Thus, $K \subseteq L$ as well. s satisfies some monic polynomial $h(x)$ with coefficients in R. It follows that $g(x) \,|\, h(x)$ in $K[x]$. Therefore, every root of g in L is a root of $h(x)$. It follows that all the roots of g are integral over R. The coefficients of g are elementary symmetric functions of the roots of g. Therefore, the coefficients of g are elements of K that are integral over R. Since R is normal, they are in R. Now suppose that $g(x)$ is any polynomial of $R[x]$ such that $g(s) = 0$. We know that $f(x) \,|\, g(x)$ in $K[x]$. The fact that $f(x) \,|\, g(x)$ in $R[x]$ follows from part (c) of the preceding proposition. □

We are now ready for:

Proof of the going down theorem. We have an integrally closed domain $R \subseteq S$ where S is integral over R and the nonzero elements of R are not zerodivisors in S. We are given a prime Q of S lying over P in R, and a prime P_0 of R with $P_0 \subset P$. We want to show that there is a prime $Q_0 \subset Q$ such that Q_0 lies over P_0. The general case of the going down theorem then follows by a straightforward induction.

We begin by showing that there is a prime ideal $q \subseteq S$ such that $q \subset Q$ and q lies over the prime ideal (0) in R. To see this, consider the multipicative system $W = (R - \{0\})(S - Q)$ in S. Because the elements of $R - \{0\}$ are not zerodivisors in S and the elements of $S - Q$ are not zero, the multiplicative system W does not contain 0. This means that there is a prime ideal q of S disjoint from W. In particular, since $R - \{0\} \subseteq W$, we must have that $q \cap R = (0)$, and since $S - Q \subseteq W$, we must have that $q \subseteq Q$. Since Q lies over P and $P_0 \subset P$, $P \neq (0)$, and this means that $q \subset Q$. We now replace S by S/q. Since q does not meet R, we still have an injection $R \hookrightarrow S/q$, and we may replace R by its image in S/q and so assume that $R \subseteq S/q$. This extension is obviously still integral: the monic equation over R satisfied by $s \in S$ is also satisfied by its image in S/q. We replace Q by Q/q, which still lies over P. If we can find a prime of S/q contained in Q/q that lies over P_0, it will have the form Q_0/q for some prime Q_0 of S with $Q_0 \subseteq Q$. Then Q_0 will lie over P_0 in R and we will also have $Q_0 \subseteq Q$. Since $P_0 \subset P$, we actually have that $Q_0 \subset Q$.

Therefore, we may assume without loss of generality that $R \subseteq S$ is an extension of domains and that S is integral over R. This stronger condition replaces the assumption that nonzero elements of R are not zerodivisors in S. Let $A = R - P_0$ and $B = S - Q$. To complete the proof, we shall show that the multiplicative system AB does not meet the ideal P_0S. This implies that there is a prime ideal Q_0 of S containing P_0S and disjoint from $AB \supseteq A \cup B$, so that $P_0 \subseteq Q_0$ and Q_0 meets neither $R - P_0$ nor $S - Q$. But this means that Q_0 lies over P_0 and is contained in Q, as required.

Suppose that $a \in A$ and $b \in B$ are such that $ab \in P_0S$. The argument used in the proof of the lying over theorem (see the lecture notes from September 27) shows that ab satisfies a monic polynomial equation $g_1(x)$ in one variable x such that all coefficients of the equation except the leading coefficient are in P_0 (not just in P_0S).

This means that b is a root of the polynomial $g(x) = g_1(ax)$ over b. Note that the leading coefficient of $g(x)$ is a power of a, and that all other coefficients are in P_0.

Think of $K = \text{frac}\,(R)$ as contained in $\text{frac}\,(S) = L$. Since b satisfies the algebraic equation $g(b) = 0$, it is algebraic over K, and has a monic minimal polynomial $f(x)$ with coefficients in K that is irreducible in $K[x]$. By the preceding Lemma, this polynomial has coefficients in R, since R is normal. It divides $g(x)$ in $K[x]$, because $g(x)$ has coefficients in $R \subseteq K$, and $f(x)$ is the minimal polynomial of b.

Since $f(x)$ is monic, our result on the division algorithm implies that $f(x)$ divides $g(x)$ in $R[x]$ as well: let us say that $g(x) = f(x)q(x)$, where all three have coefficients in R. We now consider coefficients mod P_0, which means, in effect , that we are working in $\overline{R}[x]$, where $\overline{R} = R/P_0$. Let $\overline{a}$ be the image of a in $\overline{R}$: since $a \in R - P$, $\overline{a} \neq 0$ in R/P. Then, mod P_0, $g(x)$ has the form $\overline{a}^d x^d$, since all lower coefficents are in P_0. This implies that the monic polynomial f must become x^k mod P_0, where k is its degree. This means, thinking over R, that $f(x)$ is monic of degree k with all lower coefficients in P_0: say $f(x) = x^k + p_{k-1}x^{k-1} + \cdots + p_0$, where the $p_j \in P_0$.

Since b is a root of $f(x)$, we have that $b^k = -p_{k-1}b^{k-1} - \cdots - p_0 \in P_0S \subseteq Q$, and so $b \in Q$, which is a contradiction! Thus, AB does not meet P_0S, and we are done. □

Corollary. *Let R be an integrally closed domain, S an integral extension of R that is torsion free over R, and Q a prime ideal of S that lies over P in R. Then the height of Q is equal to the height of P.*

Proof. We have already seen that the height of Q is at most the height of P. Conversely, given a chain of primes contained in P we may use the going down theorem, starting with the largest prime in the chain, to construct a chain of primes in S that lies over it and is contained in Q, and this shows that the height of Q is at least as big as the height of P. □

Let's look at two examples. Consider $R = K[x] \subseteq K[x, y]/(y^2 - y, xy) = S$. This is integral, since y satisfies a monic equation. It is an extension: we can map this larger algebra back to $K[x]$ by sending $x \mapsto x$ and $y \mapsto 0$, and the composition is the identity on $K[x]$. The element $1 - y$ generates a minimal prime Q of the larger ring containing x and not y: we can see that it is minimal, because a smaller prime cannot contain $(1 - y)$ and cannot contain y either (or else Q would contain both y and $1 - y$), while $y(1 - y) = 0$ in the quotient. But $(1 - y)S$ contracts to $xK[x]$, which has height one. The problem here is that x is a zerodivisor in S, which shows that one cannot omit the hypothesis that S be torsion-free over R in the statement of the going down theorem.

In the example above, R is normal. We next consider an example where both rings are domains but R is not normal: in fact, S is the integral closure of R. Let K be a field, let

$S = K[x,\, y]$, and let

$$R = K[x(1-x),\, x^2(1-x),\, y,\, xy] \subseteq S.$$

S is integral over R since it is generated over $K[y] \subseteq R$ by x, and $z = x$ satisfies the monic polynomial $z^2 - z - x(1-x) = 0$, which has coefficients in R. x is in the fraction field of R, since it is equal to xy/y or $x^2(1-x)/\bigl(x(1-x)\bigr)$. Let $Q = (1-x,\, y)S$, which is easily seen to lie over

$$P = \bigl(x(1-x),\ y,\, xy\bigr)R,$$

a maximal ideal of R, and let P_0 be the contraction of xS to R. Then

$$P_0 = \bigl(x(1-x),\, xy\bigr)R.$$

We claim that no prime Q_0 contained in Q lies over P_0. For any prime of S contained in Q cannot contain x, for $x \notin Q$. But since Q_0 must contain both $x(1-x)$ and xy (these elements are in P_0) and it does not contain x, it must contain both $1-x$ and y, which forces it to be equal to Q. But then it lies over P, not P_0. This shows that one cannot omit the hypothesis that R be normal in the statement of the going down theorem.

Lecture of October 6

The following result implies that, after a change of variables, any nonzero polynomial in $R = K[x_1, \ldots, x_n]$, the polynomial ring in in n variables over a ring A, becomes a nonzero scalar times a polynomial that is monic in x_n with coefficients in $A = K[x_1, \ldots, x_{n-1}] \subseteq R$, where we think of R as $A[x_n]$. We may also do this with any one of the other variables. This simple trick, or method, provides a wealth of information about algebras finitely generated over a field. It will be the key to our proofs of the Noether normalization theorem and Hilbert's Nullstellensatz.

Consider this example: the polynomial x_1x_2 is not monic in either variable. But there is an automorphism of the polynomial ring in two variables that fixes x_2 and maps x_1 to $x_1 + x_2$. (Its inverse fixes x_2 and maps x_1 to $x_1 - x_2$.) The image of x_1x_2 is $(x_1 + x_2)x_2 = x_2^2 + x_1x_2$. As a polynomial in x_2 over $K[x_1]$, this is monic. Note that we may also think of the effect of applying an automorphism as a change of variables.

More generally, if D is any ring, $R = D[x_1, \ldots, x_n]$ is a polynomial ring over D, and $g_1(x_n), \ldots, g_{n-1}(x_n)$ are arbitrary elements of $D[x_n] \subseteq R$, then there is a K-automorphism ϕ of R such that $x_i \mapsto y_i = x_i + g_i(x_n)$ for $i < n$ and while $x_n = y_n$ is fixed. The inverse automorphism is such that $x_i \mapsto x_i - g_i(x_n)$ while x_n is again fixed. This means that the elements y_i are algebraically independent and generate $D[x_1, \ldots, x_n]$. They are "just as good" as our original indeterminates.

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 K-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 is a polynomial whose sole highest degree term in x_n is a nonzero element c of D times a power of x_n. Hence, the image of f is a unit of D_c times a monic polynomial in x_n over $D_c[x_1, \ldots, x_{n-1}]$.*

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 scalar 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 in x_n: it is the product of the highest degree terms coming from all the factors, and it is

$$c_\alpha(x_n^N)^{a_1}(x_n^{N^2})^{a_2} \cdots (x_n^{N^{n-1}})^{a_{n-1}} x_n^{a_n} = c x_n^{a_n + a_1N + a_2N^2 + \cdots + a_{n-1}N^{n-1}}.$$

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

$$a_n + a_1N + a_2N^2 + \cdots + a_{n-1}N^{n-1}$$

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 $c_{\alpha_0}^{-1}\phi(f)$ is monic of degree m in x_n when viewed as a polynomial in x_n over $D_c[x_1, \ldots, x_{n-1}]$, as required. □

Let R be an A-algebra and $z_1, \ldots, z_d \in R$. We shall say that the elements $z_1, \ldots, z_d$ are *algebraically independent* over A if the unique A-algebra homomorphism from the polynomial ring $A[x_1, \ldots, x_d] \to R$ that sends x_i to z_i for $1 \leq i \leq n$ is an isomorphism. An equivalent statement is that the mononomials $z_1^{a_1} \cdots z_d^{a_d}$ as $(a_1, \ldots, a_d)$ varies in $\mathbb{N}^d$ are all distinct and span a free A-submodule of R: of course, this free A-submodule is $A[z_1, \ldots, z_d]$. The failure of the z_j to be algebraically independent means precisely that there is some nonzero polynomial $f(x_1, \ldots, x_d) \in A[x_1, \ldots, x_d]$ such that $f(z_1, \ldots, z_d) = 0$.

Note that when $D \subseteq R$ are rings and W is a multiplicative system in D, we have that $W^{-1}D \hookrightarrow W^{-1}R$. (The element d/w maps to 0 if and only if $d/1$ maps to 0, in which case $w'd = 0$ for some element of $w' \in W$ thinking in R. But this means $w'd = 0$ in D as well, and so $d/1 = 0$ and $d/w = 0$ in $W^{-1}D$.)

We can now show:

Noether normalization theorem. *Let D be a domain and let R be any finitely generated D-algebra. Then there exist a nonzero element $c \in D$ and algebraically independent elements $z_1, \ldots, z_d$ in R_c over D_c such that R_c is module-finite over its subring $D_c[z_1, \ldots, z_d]$, which is isomorphic to a polynomial ring (d may be zero).*

Hence, if D is a field K, every finitely generated K-algebra is isomorphic with a module-finite extension of a polynomial ring!

Proof. We use induction on the number n of generators of R over D. If $n = 0$ then $R = D$. We may take $d = 0$. Now suppose that $n \geq 1$ and that we know the result for algebras generated by $n - 1$ or fewer elements. Suppose that $R = K[\theta_1, \ldots, \theta_n]$ has n generators. If the θ_i are algebraically independent over D then we are done: we may take $d = n$ and $z_i = \theta_i$, $1 \leq i \leq n$. Therefore we may assume that we have a nonzero polynomial $f(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ 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. Let $c \in D - \{0\}$ be the coefficient of the highest degree term of g in x_n. Then, over D_c, $c^{-1}g$ is an equation of integral dependence for $\theta_n'/1$ over $D_c[\theta_1, \ldots, \theta_{n-1}] \subseteq R_c$, Thus, θ_n' is integral over $D_c[\theta_1', \ldots, \theta_{n-1}'] = S$, and so R_c is module-finite over S. Note that we can invert an element c'/c^t of D_c by inverting c'. Since S has $n-1$ generators over D_c, we have by the induction hypothesis for some $c' \in D - \{0\}$, that $S_{c'}$ is module-finite over a polynomial ring $D_{cc'}[z_1, \ldots, z_d] \subseteq S_{c'}$, and then $R_{cc'}$ is module-finite over $D_{cc'}[z_1, \ldots, z_d]$ as well. □

Note that if $K \subseteq L$ are fields, the statement that L is module-finite over K is equivalent to the statement that L is a finite-dimensional vector space over K, and both are equivalent to the statement that L is a finite algebraic extension of K.

Also notice that the polynomial ring $R = K[x_1, \dots, x_d]$ for $d \geq 1$ has dimension at least d: $(0) \subset (x_1)R \subset (x_1, x_2)R \subset \cdots \subset (x_1, \dots, x_d)R$ is a strictly increasing chain of prime ideals of length d. Later we shall show that the dimension of $K[x_1, \dots, x_d]$ is exactly d. But for the moment, all we need is that $K[x_1, \dots, x_d]$ has dimension at least one for $d \geq 1$.

Corollary. *Let R be a finitely generated algebra over a field K, and suppose that R is a field. Then R is a finite algebraic extension of K, i.e., R is module-finite over K.*

Proof. By the Noether normalization theorem, R is module-finite over some polynomial subring $K[z_1, \dots, z_d]$. If $d \geq 1$, the polynomial ring has dimension at least one, and then R has dimension at least one, a contradiction. Thus, $d = 0$, and R is module-finite over K. Since R is a field, this means precisely that R is a finite algebraic extension of K. □

Corollary. *Let K be an algebraically closed field, let R be a finitely generated K-algebra, and let m be a maximal ideal of R. Then the composite map $K \to R \twoheadrightarrow R/m$ is an isomorphism.*

Proof. R/m is a finitely generated K-algebra, since R is, and it is a field. Thus, $K \to R/m$ gives a finite algebraic extension of K. Since K is algebraically closed, it has no proper algebraic extension, and so $K \to R/m$ must be an isomorphism.

Corollary (Hilbert's Nullstellensatz, weak form). *Let $R = K[x_1, \dots, x_n]$ be a polynomial ring over and algebraically closed field K. Then every maximal ideal m of R is the kernel of a K-homomorphism $K[x_1, \dots, x_n] \to K$, and so is determined by the elements $\lambda_1, \dots, \lambda_n \in K$ to which $x_1, \dots, x_n$ map. This maximal ideal is the kernel of the evaluation map $f(x_1, \dots, x_n) \mapsto f(\lambda_1, \dots, \lambda_n)$. It may also be described as the ideal $(x_1 - \lambda_1, \dots, x_n - \lambda_n)R$.*

Proof. Since $\gamma : K \cong R/m$, the K-algebra map $R \to R/m$, composed with γ^{-1}, gives a map $R \twoheadrightarrow K$ whose kernel is m. □

Thus, when K is algebraically closed, we have a bijection between the points of K^n and the maximal ideals of $K[x_1, \dots, x_n]$.

Corollary (Hilbert's Nullstellensatz, alternate weak form). *Let $f_1, \dots, f_n$ be polynomials in $K[x_1, \dots, x_n]$, where K is algebraically closed. Then the f_i generate the unit ideal (i.e., we have $1 = \sum_t g_t f_t$ for suitable polynomials g_t) if and only if the polynomials f_i do not vanish simultaneously, i.e., if and only if the algebraic set $V(f_1, \dots, f_n) = \emptyset$.*

Proof. If the f_i do not generate the unit ideal, the ideal they generate is contained in some maximal ideal of $K[x_1, \dots, x_n]$. But the functions in that maximal ideal all vanish at one point of K^n, a contradiction. On the other hand, if the f_i all vanish simultaneously at a point of K^n, they are in the maximal ideal of polynomials that vanish at that point: this direction does not need that K is algebraically closed. □

We have two uses of the notation $V(S)$: one is for any subset S of any ring, and it is the set of all primes containing S. The other use is for polynomial rings $K[x_1, \dots, x_n]$, and then it is the set of points where the given polynomials vanish. For clarity, suppose that we use $\mathcal{V}$ for the second meaning. If we think of these points as corresponding to a subset of the maximal ideals of the ring (it corresponds to all maximal ideals when the field is algebraically closed), we have that $\mathcal{V}(S)$ is the intersection of $V(S)$ with the maximal ideals corresponding to points of K^n, thought of as a subset of K^n. Suppose that for every $y \in K^n$ we let $m_y = \{f \in K[x_1, \dots, x_n] : f(y) = 0\}$. Then m_y is a maximal ideal of $K[x_1, \dots, x_n]$ whether K is algebraically closed or not. When K is algebraically closed, we know that all maximal ideals have this form. This gives an injection $K^n \to \operatorname{Spec}(R)$ that sends y to m_y. The closed algebraic sets of K^n are simply the closed sets of $\operatorname{Spec}(R)$ intersected with the image of K^n, if we identify that image with K^n. Thus, the algebraic sets are the closed sets of a topology on K^n, which is called the *Zariski topology.* It is the inherited Zariski topology from $\operatorname{Spec}(R)$. Note that $\mathcal{V}(I) = \{y \in Y : m_y \in V(I)\}$.

In this course, I will continue from here on to use the alternate notation $\mathcal{V}$ when discussing algebraic sets. However, people often use the same notation for both, depending on the context to make clear which is meant.

Theorem (Hilbert's Nullstellensatz, strong form. *Let K be an algebraically closed field and let $R = K[x_1, \dots, x_n]$ be the polynomial ring in n variables over K. Suppose that $g, f_1, \dots, f_s \in R$. Then $g \in \operatorname{Rad}(f_1, \dots, f_s)$ if and only if $\mathcal{V}(g) \supseteq \mathcal{V}(f_1, \dots, f_s)$, i.e., if and only if g vanishes at every point where the f_i vanish simultaneously.*

Proof. It is clear that $g^N = \sum_{i=1}^{s} g_i f_i$ implies that g vanishes wherever the all of the f_i vanish: at such a point y, we have that $g(y)^N = 0$ and so $g(y) = 0$.

The more interesting implication is the statement that if g does vanish whenever all the f_i vanish then g has a power that is in the ideal generated by the f_i. The following method of proof is called Rabinowitsch's trick. Introduce an extra variable z and consider the polynomials $f_1, \dots, f_s, 1 - gz \in K[x_1, \dots, x_n, z]$. There is no point of K^{n+1} where these all vanish: at any point where the f_i vanish (this only depends on what the first n coordinates of the point are), we have that g vanishes as well, and therefore $1 - gz$ is $1 - 0 = 1$. This means that $f_1, \dots, f_s, 1 - gz$ generate the unit ideal in $K[x_1, \dots, x_n, z]$, by the weak form of Hilbert's Nullstellensatz that we have already established. This means that there is an equation

$$1 = H_1(z)f_1 + \cdots + H_s(z)f_s + H(z)(1 - gz)$$

where $H_1(z), \dots, H_s(z)$ and $H(z)$ are polynomials in $K[x_1, \dots, x_n, z]$: all of them may involve all of the variables x_j and z, but we have chosen a notation that emphasizes their dependence on z. But note that $f_1, \dots, f_s$ and g do not depend on z. We may assume that $g \neq 0$ or the result is obvious. We now define a $K[x_1, \dots, x_n]$-algebra map ϕ from $K[x_1, \dots, x_n, z]$, which we think of as $K[x_1, \dots, x_n][z]$, to the ring $K[x_1, \dots, x_n][1/g] = K[x_1, \dots, x_n]_g$, which we may think of as a subring of the fraction field of $K[x_1, \dots, x_n]$. This ring is also the localization of $K[x_1, \dots, x_n]$ at the multiplicative system $\{1, g, g^2, \dots\}$ consisting of all powers of g. Note that every element of $K[x_1, \dots, x_n]_g$ can be written in the form u/g^h, where $u \in K[x_1, \ldots, x_n]$ and h is some nonnegative integer. We define the $K[x_1, \ldots, x_n]$-algebra map ϕ simply by specifying that the value of z is to be $1/g$. Applying this homomorphism to the displayed equation, we find that

$$1 = H_1(1/g)f_1 + \cdots + H_s(1/g)f_s + H(1/g)(1-1)$$

or

$$1 = H_1(1/g)f_1 + \cdots + H_s(1/g)f_s.$$

Since each of the $H_i(1/g)$ is in $K[x_1, \ldots, x_n]_g$, we can choose a positive integer N so large that each of the $g_i = g^N H_i(1/g) \in K[x_1, \ldots, x_n]$: there are only finitely many denominators to clear. Multiplying the most recently displayed equation by g^N gives the equation $g^N = g_1 f_1 + \cdots + g_n f_n$ with $g_i \in K[x_1, \ldots, x_n]$, which is exactly what we wanted to prove. □

Lecture of October 9

Corollary. *Let $R \to S$ be a homomorphism of finitely generated K-algebras. Then every maximal ideal of S contracts to a maximal ideal of R.*

Proof. Suppose that the maximal ideal n of S contracts to the prime P in R, so that $K \subseteq R/P \subseteq S/n$. Then S/n is a finite algebraic extension of K, i.e., a finite dimensional K-vector space, and so the domain R/P is a finite-dimensional K-vector space, i.e., it is module-finite over K, and therefore it is a domain of dimension 0, which forces it to be a field. □

An element $x \neq 0$ of a ring R is called *prime* if it generates a prime ideal. This means that x is not a unit and if $x \,|\, (rr')$ with r, $r' \in R$, then $x \,|\, r$ or $x \,|\, r'$. An element $x \neq 0$ is called *irreducible* if it is not a unit and cannot be written as the product of two elements neither of which is a unit. In a domain, prime elements are always irreducible. (If x is prime and $x = fg$, then x divides f or g, say $f = xf_1$, and then $x(1 - f_1g) = 0$. Since $x \neq 0$, g is a unit.) In a UFD, irreducible elements are prime, so that the two notions agree, and every element factors uniquely as a product of finitely many primes, where "uniquely" means up to the order of the factors and adjustments for multiplication by units: one may alter a factorization by multiplying one of the factors by a unit and another by its inverse. Thus if $f = f_1 \cdots f_n = g_1 \cdots g_n$ then there is a permutation π of $\{1, \ldots, n\}$ and there are units $\alpha_1, \ldots, \alpha_n$ of R such that $g_{\pi(j)} = \alpha_j f_j$, $1 \leq j \leq n$, and $\alpha_1 \cdots \alpha_n = 1$. Note also that if a non-unit f divides an irreducible g, so that $g = fu$, then u must be a unit. In particular, if one irreducible divides another, they are associates, i.e., each is a unit times the other.

Proposition. *Let R be a UFD. Then every nonzero prime ideal of R contains a prime ideal generated by an irreducible element, and a prime ideal of R has height one if and only if it is generated by an irreducible element.*

Proof. Let Q be any nonzero prime ideal, and let $f \in Q - \{0\}$. The f can be factored into irreducible factors, say $f = f_1 \cdots f_k$, and since this product is in Q, at least one of the factors, say f_i, is in Q. Then f_i generates a prime ideal contained in Q. This shows that a prime ideal cannot possibly have height one unless it is generated by an irreducible element. Finally, if $P = fR$ is generated by an irreducible element but contains a smaller nonzero prime ideal, that prime will in turn contain a prime generated by a nonzero irreducible element g. But then $f \,|\, g$, which implies that they are the same, up to unit factors. □

In any ring R, a chain of prime ideals $P_0 \subset P_1 \subset \cdots \subset P_n$ is called *saturated* if for all i, $0 \leq i < n$, there is no prime strictly between P_i and P_{i+1}.

Theorem. *Let R be a finitely generated integral domain over the field K. Choose $z_1, \ldots, z_d \in R$ such that R is module-finite over the polynomial ring $K[z_1, \ldots, z_d]$. Then* $\dim R = d$. *In fact, the height of every maximal ideal of R is d. In particular, the height of every maximal ideal in $K[z_1, \ldots, z_d]$ is d. Moreover, every saturated chain of primes in R from (0) to a maximal ideal m has length d.*

Proof. We first prove that the dimension of R is d by induction on d. We know at once that $\dim R = \dim A$, where $A = K[z_1, \ldots, z_d]$, and we have already seen that $\dim A \geq d$. It will suffice to show that $\dim A \leq d$. Consider any chain of primes of A. We can assume that the two smallest primes in it are P and 0, where P is a height one prime generated by an irreducible element f. There will be a chain of length one less in A/P. Therefore, it suffices to show that $\dim A/P = d - 1$.

But after a change of variables we may assume that f is monic in z_d over $K[z_1, \ldots, z_{d-1}]$, and therefore A/P is integral over $K[z_1, \ldots, z_{d-1}]$.

Thus, the dimension of R is d. We can use almost the same argument to show by induction that every saturated chain from (0) to a maximal ideal m of R has length d, but we must make use of the going down theorem for this. Note that this statement evidently implies that the height of m is d. Fix a maximal ideal m of R and consider a saturated chain contained in m, say

$$(0) \subset Q_1 \subset \cdots \subset Q_k = m.$$

We want to show that $k = d$. Since the chain is saturated, we know that Q_1 has height one in R. But the contraction of Q_1 to $A = K[z_1, \ldots, z_d]$ must have height one as well (this uses the going down theorem), and so must be generated by a single irreducible polynomial f. As before, we may assume, after a change of variables, that f is monic in z_d over $K[z_1, \ldots, z_{d-1}]$. Now,

$$(0) = Q_1/Q_1 \subset Q_2/Q_1 \subset \cdots Q_k/Q_1 = m/Q_1$$

is a saturated chain of primes in the domain R/Q_1 from (0) to the maximal ideal m/Q_1. But R/Q_1 is module-finite over A/fA, which in turn is module-finite over $K[z_1, \ldots, z_{d-1}]$, and so has dimension $d - 1$. It follows from the induction hypothesis that $k - 1 = d - 1$, and so $k = d$. □

We review the notions of transcendence basis and transcendence degree. Let $K \subseteq L$ be fields. By Zorn's lemma, any set of elements of L algebraically independent over K can be enlarged to a maximal such set, which is called a *transcendence basis* for L over K. Such a basis will be empty if and only if L is algebraic over K. If $\{x_\lambda : \lambda \in \Lambda\}$ is a transcendence basis, then L contains a subring $K[x_\lambda : \lambda \in \Lambda]$ which is isomorphic with a polynomial ring in variables corresponding to the x_λ, and it also contains the fraction field, denoted $K(x_\lambda : \lambda \in \Lambda)$ of that polynomial ring, which is called a *pure transcendental extension* of K. It is easy to see that L is algebraic over $K(x_\lambda : \lambda \in \Lambda)$ (a transcendental element could be used to enlarge the transcendence basis), and so every field extension can be obtained in two steps: a pure transcendental extension followed by an algebraic extension. Either step might just consist of a trivial field extension. The key fact that we have not yet proved but will prove in the sequel is that any two transcendence bases have the same cardinality, which is called the *transcendence degree* of L over K. We are primarily interested in the case where the transcendence degree is finite, which it always is when L is finitely generated as a field over K. However, we treat the general case.

An alternative characterization of a transcendence basis for the field L over its subfield K is that it is a set of algebraically independent elements in L over K generating a subfield L_0 of L such that L is algebraic over L_0.

We sketch the proof that any two transcendence bases for L over K have the same cardinality. (The reader already familiar with this material or not interested in the proof may skip this and the next two paragraphs.) It suffices to show that if X is a set of algebraically independent elements of L and Y is a transcendence basis, then there is an injection $f : X \hookrightarrow Y$, such that $X \cup \bigl(Y - f(X)\bigr)$ is a transcendence basis for L over K. That is, one may replace a certain subset of Y with the same cardinality as X with the elements of X and still have a transcendence basis. This will imply that the cardinality of X is at most that of Y. Given two transcendence bases, it follows that the cardinality of each is at most that of the other, so that they have the same cardinality.

Consider all injections $g : X_0 \to Y$, where X_0 is a (possibly empty) subset of X, such that $X_0 \cup \bigl(Y - g(X_0)\bigr)$ is a transcendence basis for L. These are partially ordered by the rule $(X_0, g_0) \leq (X_1, g_1)$ if $X_0 \subseteq X_1$ and the restriction of g_1 to X_0 is g_0. Every chain $(X_i, g_i)_{i \in I}$ has an upper bound: there is a unique function g on $X = \bigcup_i X_i$ which extends all of the given functions. It is easy to see that one has algebraic independence for the elements of $X \cup \bigl(Y - g(X)\bigr)$ and that L is algebraic over the field that they generate. (Any element of L is algebraic over a field generated by finitely many of the $y_j \in Y$. Those that get replaced by x_k when we take the union have already been replaced for some sufficiently large X_i in the union.) By Zorn's Lemma, there is a maximal pair $(X_0,\, g)$ with the specified property.

We want to see that X_0 is all of X. If not, choose $x \in X - X_0$. Then x is algebraic over $K\bigl(X_0 \cup \bigl(Y - g(X_0)\bigr)\bigr)$, and so satisfies a polynomial equation over this field. We may clear denominators to obtain a polynomial F over K in x and finitely many of the variables in $X_0 \cup \bigl(Y - g(X_0)\bigr)$ such that x actually occurs in F. The polynomial F must involve at least one element of $Y - g(X_0)$, or else X would not be an algebraically independent set. This means that we can choose $y \in Y - g(X_0)$ that occurs in F. But then, using F, we see that y is algebraic over the field generated over K by $X_0 \cup \{x\} \cup \bigl(Y - g(X_0) - y\bigr)$, and we extend g to g' on $X_1 = X \cup \{x\}$ by letting $g'(x) = y$. We still have algebraic independence: if we omit x, that is clear, while if an algebraic relation involves x, then x is algebraic over the field generated by the others, and that implies that y is as well, a contradiction. L is algebraic over the field generated by these new elements, because y is. □

From the definition of transcendence degree and what we have already proved, we have at once:

Corollary. *Let R be a domain finitely generated over a field K. Then* $\dim R$ *is the transcendence degree of* $\mathrm{frac}\,(R)$ *over K.*

Proof. R is module-finite over a polynomial ring $K[z_1, \, \ldots \, , z_d]$ for some integer d, which means that $\mathrm{frac}\,(R)$ is algebraic over the pure transcendental extension $K(z_1, \, \ldots \, , z_d)$ of K. Thus, the transcendence degree is d, which we already know to be equal to $\dim R$. □

It was an open question for a considerable time whether, in any commutative ring, there could be saturated chains of distinct finite lengths joining two primes. M. Nagata gave the first counter-example: he constructed a Noetherian domain of dimension 3 having a unique maximal ideal m with saturated chains $(0) \subset P_1 \subset P_2 \subset m$ of length 3 and also $0 \subset Q \subset m$ of length 2. Cf. [M. Nagata, *Local rings*, Interscience, New York, 1962], Appendix A1., *Examples of bad Noetherian rings.* In [M. Hochster, *Prime ideal structure in commutative rings*, Trans. Amer. Math. Soc. **142** (1969), 43–60] it is shown that the spectrum of a commutative ring can be any finite partially ordered set. However, examples of such behavior in Noetherian rings are not easily come by. The Noetherian rings that arise in algebraic geometry, number theory, and several complex variables all have a property introduced by A. Grothendieck called *excellence*, which implies that saturated chains joining P to Q when $P \subset Q$ all have the same length.

However, one does not need to look at pathological examples to find instances where maximal ideals have different heights: this happens in the polynomial ring in one variable over a PID, if the PID has a unique maximal ideal. Such PIDs are of great importance, and we digress briefly to discuss them.

Let V be a principal ideal domain with just one maximal ideal P. Such rings are called *discrete rank one valuation domains*, but it is common practice to refer to them more briefly as *discrete valuation rings*, and, unless otherwise specified, we shall do that here. Note that if S is any principal ideal domain and Q is any nonzero prime ideal of S, then S_Q is a discrete valuation ring. The acronym *DVR* is used for *discrete valuation ring.* In a DVR, the maximal ideal is principal. Let t be the generator. This is the only prime element (up to multiplication by units). Thus, every nonzero element f can be written uniquely as αt^n, where α is a unit. The non-negative integer n is called the *order* of f, often written $\mathrm{ord}\, f$. Note that if $f,\ g \neq 0$, then

$$(1) \qquad \mathrm{ord}\,(fg) = \mathrm{ord}\, f + \mathrm{ord}\, g,$$

and that and if $f + g$ is not zero as well, then

$$(2) \qquad \mathrm{ord}\,(f+g) \geq \min\,\{\mathrm{ord}\, f,\ \mathrm{ord}\, g\}$$

with equality if $\mathrm{ord}\, f \neq \mathrm{ord}\, g$. Localizing V at any nonzero element in the maximal ideal gets rid of the only nonzero prime in V, and produces the fraction field of V. In particular, V_t is the fraction field. We can extend the order function from $V - \{0\}$ to $V_t - \{0\}$ by letting $\mathrm{ord}\,(f/t^n) = \mathrm{ord}\,(f) - n$. This is easily checked to be independent of the representation of the element as a fraction, and the displayed properties (1), (2) of ord continue to hold. The function ord from $\mathrm{frac}\,(V) - \{0\} \twoheadrightarrow \mathbb{Z}$ is called the *valuation* associated to V. Note that V is 0 together with the set of elements of $\mathrm{frac}\,(V)$ of nonnegative order, and that the maximal ideal of V is 0 together with the set of elements of positive order, while the set of units of V coincides with the subset of $\mathrm{frac}\,(V)$ of elements of order 0.

Conversely, given a field F and a surjective function $\mathrm{ord} : F - \{0\} \twoheadrightarrow \mathbb{Z}$ such that for all $f, g \in F - \{0\}$,

$$(1) \qquad \mathrm{ord}\,(fg) = \mathrm{ord}\, f + \mathrm{ord}\, g,$$

and for all $f, g \in F - \{0\}$ such that $f + g \neq 0$,

$$(2) \qquad \mathrm{ord}\,(f+g) \geq \min\,\{\mathrm{ord}\, f,\ \mathrm{ord}\, g\}$$

with equality if $\operatorname{ord} f \neq \operatorname{ord} g$, the set of elements in F on which ord is nonnegative together with 0 is a subring of F. The elements of positive order together with 0 form a unique maximal ideal, which is generated by any element of order 1, and every nonzero element is a unit times a power of that generator. Thus, every such function determines a unique DVR for which it is the associated valuation.

One can consider functions on a field with the properties displayed above taking values in a totally ordered abelian group other than $\mathbb{Z}$. When the set of values is the group $\mathbb{Z}^{\oplus r}$ (with a suitable order: we are not giving all the details here) one refers to a *discrete rank r valuation ring*. When the set of values is, for example the rational numbers, the valuation is no longer discrete. In these lectures, unless otherwise specified, we shall assume that any given valuation has $\mathbb{Z}$ as the set of values, and that all given discrete valuation rings are rank one discrete valuation domains.

The ring of formal power series $K[[t]]$ in one variable over a field K is a discrete valuation ring with maximal ideal generated by t. The key point is that a power series with a nonzero constant term has an inverse. This comes down to the case where the constant term is 1. The point is that if the power series is $1 + tf$ then the formal expression $1 - tf + t^2f^2 - t^3f^3 + \cdots$ can be given a meaning as a power series, because although the sum looks infinite, there are only finitely many terms involving a given power of t, and this gives the inverse of $1 + tf$.

The localization of the integers at the prime ideal generated by p, where p is a prime integer, is also a DVR, with maximal ideal generated by p. This ring is the set

$$\{m/n : m,\, n \in \mathbb{Z}, p \nmid n\} \subseteq \mathbb{Q}.$$

Lecture of October 11

Examples. If V is a DVR with maximal ideal tV, then in $V[x]$, which is a UFD, the element $tx-1$ generates a maximal ideal: $V[x]/(tx-1) \cong V[1/t] = V_t = \mathrm{frac}\,(V)$, a field. On the other hand, the chain $(0) \subset (x)V[x] \subset (x,\, t)V[x]$ shows that $(x,\, t)V[x]$ is a maximal ideal of height at least 2 (and we shall see later that the height is exactly 2).

Also, consider $R = K[x,y,z]/I$, where $I = (xy,\, xz) = (x) \cap (y,z)$. In this ring, every prime contains either the image $\overline{x}$ of x or both of the images $\overline{y}$, $\overline{z}$ of y and z. Then $P = (\overline{x})R$ is a minimal prime of R with $R/(\overline{x})R \cong K[y,z]$, and $P' = (\overline{y},\, \overline{z})R$ is a minimal prime of R with $R/(\overline{y},\, \overline{z})R \cong K[x]$. Saturated chains from P to a maximal ideal correspond to saturated chains from (0) to a maximal ideal in $K[y,\, z]$ and have length two while saturated chains from P' to a maximal ideal correspond to saturated chains from (0) to a maximal ideal in $K[x]$, and have length one.

We do have the following:

Theorem. *Let R be a finitely generated algebra over the field K.*

(a) *The dimension of R is the same as the maximum cardinality of a set of elements of R that is algebraically independent over K.*

(b) *If $P \subseteq Q$ are primes of R, all saturated chains of primes from P to Q have the same length.*

(c) *Suppose that R is a domain. Then all saturated chains from 0 to a prime ideal P have length equal to height P, and all saturated chains from P to a maximal ideal have length equal to* $\dim(R/P)$. *Moreover* $\mathrm{height}\, P + \dim(R/P) = \dim R$. *For any two primes $P \subseteq Q$, every saturated chain from P to Q has length* $\mathrm{height}\, Q - \mathrm{height}\, P$.

Proof. We first prove (c). Choose any saturated chain from (0) to P: suppose that it has length k. Also choose any saturated chain from P to a maximal ideal m: this corresponds to a saturated chain from $(0) = P/P$ to m/P in R/P, and so has length $\dim(R/P)$. Putting these two chains together gives a saturated chain in R from 0 to m of length $k + \dim(R/P)$, and this saturated chain has length equal to $\dim(R)$. Thus, $k + \dim(R/P) = \dim R$, and so all saturated chains from (0) to P have the same length, $\dim R - \dim(R/P)$, which must be the same as the height of P. Finally, a saturated chain from P to Q corresponds to a saturated chain from $(0) = P/P$ to Q/P in R/P. Its length is therefore $\dim(R/P) - \dim\bigl((R/P)/(Q/P)\bigr) = \dim(R/P) - \dim(R/Q)$ which we may rewrite as $\bigl(\dim R - \mathrm{height}\, P\bigr) - (\dim R - \mathrm{height}\, Q) = \mathrm{height}\, Q - \mathrm{height}\, P$, as required. (b) is obvious because saturated chains from P to Q in R correspond to saturated chains from $(0) = P/P$ to Q/P in the domain R/P.

Finally, to prove (a), first note that, by Noether normalization, R is module-finite over a polynomial ring $K[z_1,\, \ldots\, , z_d]$, and then $d = \dim R$. This shows that there exist $\dim R$ algebraically independent elements in R. To see that there cannot be more, suppose that $K[x_1,\, \ldots\, , x_h] \subseteq R$, where $x_1,\, \ldots\, , x_h$ are algebraically independent. Then the set $K[x_1,\, \ldots\, , x_h] - \{0\}$ is a multiplicative system of R not containing 0, and so there exists a prime ideal P of R disjoint from it. This means that $P \cap K[x_1,\, \ldots\, , x_h] = (0)$, and this implies that the composite map $K[x_1, \dots, x_h] \hookrightarrow R \twoheadrightarrow R/P$ is injective. Then $\dim R \geq \dim(R/P)$, which is the transcendence degree of $\mathrm{frac}\,(R/P)$ over K, and since $K[x_1, \dots, x_h] \hookrightarrow R/P$, the transcendence degree is $\geq h$, which shows that $\dim R \geq h$, as required. □

Because height $P = \dim R - \dim(R/P)$ in a domain R that is a finitely generated K-algebra, the height of a prime is also called its *codimension*.

Theorem. *Let R be any finitely generated algebra over the field K. The every prime ideal and, hence, every radical ideal is the intersection of the maximal ideals that contain it. It follows at once that for any ideal I, the intersection of the maximal ideals containing I is* $\mathrm{Rad}\,(I)$.

Proof. Since every radical ideal is the intersection of the primes that contain it, it is clear that we need only prove this for prime ideals P. Suppose that $u \notin P$ is in every maximal ideal that contains P. Then the image of u in R/P is a nonzero element that is in every maximal ideal. We therefore may reduce at once to the case where R is a domain, $P = (0)$, and we need only show that there is no element u that is in every maximal ideal. By Noether normalization, R is then module-finite over a polynomial ring $A = K[x_1, \dots, x_d]$. The nonzero element u will have a nonzero multiple in A (this was shown in the course of the proof of the lying over theorem), and so we may assume without loss of generality that $u \in A - \{0\}$. Since every maximal ideal of R lies over a maximal ideal of A, it suffices to show that a nonzero element u of a polynomial ring A cannot be in every maximal ideal.

If K is infinite we need only consider maximal ideals that arise as the set of polynomials that vanish at a point of K^d: if u were in all of these, it would be a nonzero polynomial that vanishes everywhere. (This does not happen. One can use induction on the number of variables. In the case of one variable, the degree bounds the number of roots. In the case of d variables, view the polynomial as a polynomial in x_d with coefficients in $K[x_1, \dots, x_{d-1}]$. At least one coefficient is nonzero, and by the induction hypothesis will not vanish at some point $(\lambda_1, \dots, \lambda_{d-1}) \in K^{d-1}$. Substitute these λ_i for the x_i, $1 \leq i \leq d-1$. This produces a nonzero polynomial in x_n, and there will be values for x_n for which it does not vanish.)

If the field K is finite, pick a point $(\lambda_1, \dots, \lambda_d)$ of the algebraic closure L of K at which u does not vanish: the algebraic closure is infinite. Then $K[\lambda_1, \dots, \lambda_d]$ is a finite algebraic extension of K, and so a field, and evaluation at $(\lambda_1, \dots, \lambda_d)$ gives a surjection $K[x_1, \dots, x_d] \twoheadrightarrow K[\lambda_1, \dots, \lambda_d]$ that does not kill u. The kernel is a maximal ideal not containing u. □

We noted earlier that when working with finitely generated K-algebras, maximal ideals contract to maximal ideals. For any commutative ring R, we may let $\mathrm{MaxSpec}\,(R)$ denote the space of maximal ideals of R in the inherited Zariski topology. This is not a functor, in that maximal ideals need not contract to maximal ideals (the ideal $(0) \subseteq \mathbb{Q}$ is maximal, but its contraction to $\mathbb{Z}$ is not). But when both rings are finitely generated K-algebras and one has $f : R \to S$, the restriction of $\mathrm{Spec}\,(f)$ to $\mathrm{MaxSpec}\,(S)$ gives a map into $\mathrm{MaxSpec}\,(R)$.

It is worth noting that for a finitely generated K-algebra R, there is bijection of the closed sets in $\mathrm{Spec}\,(R)$ with the closed sets in $\mathrm{MaxSpec}\,(R)$ that sends $V(I)$ to its intersection with $\mathrm{MaxSpec}\,(R)$. The reason is that the set of maximal ideals in $V(I) \in \mathrm{Spec}\,(R)$

is dense, and so determines $V(I)$. The closure of a set of primes $\{P_\sigma\}_{\sigma\in\Sigma}$ is the smallest closed set $V(J)$ that contains them all, which is given by the largest ideal J such that $J \subseteq P_\sigma$ for all σ, and thus the closure is $V(\bigcap_{\sigma\in\Sigma} P_\sigma)$. In a finitely generated K-algebra, the intersection of the maximal ideals containing I is $\mathrm{Rad}\,(I)$, by the result we just proved, and so the closure of the set of maximal ideals containing I is $V\bigl(\mathrm{Rad}\,(I)\bigr) = V(I)$.

Thus, $\mathrm{Spec}\,(R)$ and $\mathrm{MaxSpec}\,(R)$ are *very* closely related when R is a finitely generated K-algebra. They have "the same" closed sets, but there are "extra points" thrown in when one looks at $\mathrm{Spec}\,(R)$.

A partially ordered set is said to satisfy the *ascending chain condition* or *ACC* if, equivalently:

(1) Every strictly ascending chain is finite.
(2) Every infinite non-decreasing chain is eventually constant.
(3) Every non-empty subset has a maximal element.

That (2) implies (1) is obvious, and (1) implies (2) because in a counter-example to (2) one may omit the duplicated terms. (3) implies (1) is clear because because an infinite strictly ascending chain is a non-empty subset with no maximal element. The fact that (1) implies (3) is slightly more subtle, and actually uses a weak version of the axiom of choice. If one has a non-empty subset with no maximal element one can construct a strictly ascending sequence of elements recursively as follows. Let x_1 be any element in the set. Assume that $x_1, \ldots, x_n$ have been chosen and form a strictly ascending chain. Then choose x_{n+1} to be any element of the subset strictly larger than x_n. This must be possible, or else x_n would be a maximal element. Note that in this process we need to make countably many choices.

A partially ordered set is said to satisfy the *descending chain condition* or *DCC* if, equivalently:

(1) Every strictly descending chain is finite.
(2) Every infinite non-increasing chain is eventually constant.
(3) Every non-empty subset has a minimal element.

Of course, a poset satisfies DCC if and only if the poset obtained by reversing the order has ACC. A linearly ordered set with DCC is the same thing as a well-ordered set.

A module M over a ring R is said to satisfy *ACC* or to be *Noetherian* (after Emmy Noether) if its partially ordered set of submodules under $\subseteq$ has ACC, and M is said to have *DCC* or to be *Artinian* (after Emil Artin) if its partially order set of submodules under $\subseteq$ has DCC.

Lecture of October 13

Proposition. *The following conditions on a module M over a ring R are equivalent:*

(a) *M has ACC, i.e., M is Noetherian.*
(b) *Every nonempty family of submodules of M has a maximal element*
(c) *Given any set S of elements of M spanning a submodule N of M, there is a finite subset of S spanning N.*
(d) *Given any infinite sequence of elements of M spanning a submodule N, some finite initial segment of the sequence spans N.*
(e) *Every submodule of M is finitely generated.*

Proof. We already know that (a) and (b) are equivalent, while (c) follows from (b) applied to the family of submodules generated by finite subsets of S (the empty subset spans 0), for if N_0 is spanned by the finite set $S_0 \subseteq S$ is maximal among these but different from N, we can choose $s \in S$ not in N_0 and then $S \cup \{s\}$ spans a larger submodule than N_0. It is clear that (c) implies (d), since any finite subset of the sequence is contained in some initial segment. To see that (d) implies (e), let $N \subseteq M$ be any submodule, and suppose that it is not finitely generated. We construct an infinite sequence recursively as follows. Choose a nonzero element $u_1 \in N$. If $u_1, \ldots, u_n$ have been chosen such that for every i, $1 < i \leq n$, u_i is not in the span of its predecessors, note that since $Ru_1 + \cdots + Ru_n = N_n \neq N$, we can choose $u_{n+1} \in N - N_n$. We have now constructed a sequence that contradicts condition (d). Finally, to see that (e) implies (a), note that if M has a non-decreasing chain of submodules N_i, the union N is finitely generated. Then for all sufficiently large i, all of the generators are in N_i, and so the sequence is constant from some point on. □

Recall that $0 \to N \to M \to Q \to 0$ is a short exact sequence of R-modules if N injects into M and is the kernel of $M \to Q$, which is surjective. In studying short exact sequences we might as well replace N by its image and assume that $N \subseteq M$. The hypothesis then means that the induced map $M/N \to Q$ is an isomorphism, so that one might as well assume that $Q = M/N$. We may make this transition in a proof without comment.

Lemma. *Let $0 \to N \to M \to Q \to 0$ be a short exact sequence of R-modules.*

(a) *Let $M_0 \subseteq M_1 \subseteq M$ be submodules, and suppose that $M_1 \cap N = M_0 \cap N$ and that the images of M_0 and M_1 in Q are the same. Then $M_0 = M_1$,*
(b) *M is Noetherian if and only if both N and Q are.*
(c) *M is Artinian if and only if both N and Q are.*
(d) *A finite direct sum of Noetherian (respectively, Artinian) modules is Noetherian (respectively, Artinian).*

Proof. To prove (a), suppose that $u \in M_1$. Then some element $v \in M_0$ has the same image as u in Q. It follows that $v - u = w$ maps to 0 in Q, and so is in $M_1 \cap N = M_0 \cap N$. Thus, $u = v - w \in M_0$, as required.

To prove (b), suppose first that M is Noetherian. An increasing chain in N is an increasing chain in M, and so N is Noetherian. The inverse images in M of the modules in an increasing chain in Q form an increasing chain in M, and so Q is Noetherian. Suppose, conversely, that N and Q are both Noetherian, and that one has an increasing chain in M. The intersections of the modules in the chain with N are eventually constant, and the images of the modules in Q are eventually constant. It follows from part (a) that the chain in M is eventually constant.

The proof of (c) is exactly the same, with the word "increasing" replaced throughout by the word "decreasing." (d) follows by induction from the case of a direct sum of two modules, which in turn follows from the (b) or (c) applied to the short exact sequence $0 \to M_1 \to M_1 \oplus_R M_2 \to M_2 \to 0$. □

A ring R is called *Noetherian* (respectively, *Artinian* or *Artin*) if it has that property as a module over itself. Since the R-submodules of R are the ideals of R, this is equivalent to assuming that the ideals of R satisfy ACC (respectively, DCC). Also, a ring is Noetherian iff every ideal is finitely generated.

Note that in part (a) of the Lemma, the condition that $M_0 \subseteq M_1$ is needed. To see why, let K be an infinite field and consider the short exact sequence

$$0 \to Ke_1 \to Ke_1 \oplus Ke_2 \to Ke_2 \to 0$$

where $Ke_1 \oplus Ke_2 \cong K^2$ is a two-dimensional vector space with basis e_1, e_2. Let M_λ be the span of the vector $e_1 + \lambda e_2$, where $\lambda \in K - \{0\}$. The M_λ are mutually distinct lines in K^2 (and they are mutually incomparable), but they all intersect Ke_1 in 0 and they all have image Ke_2 in Ke_2.

Proposition. *A module M over a Noetherian ring R is Noetherian iff it is finitely generated. A module M over a ring R is Noetherian if and only if it is finitely generated and $R/\mathrm{Ann}_R M$ is a Noetherian ring.*

Proof. If R is Noetherian then so is each finitely generated free module, since such a module is a finite direct sum of copies of R, and every finitely generated module is a homomorphic image of a finitely generated free module.

If M is finitely generated and $R/\mathrm{Ann}_R M$ is Noetherian, we may think of M as a module over $R/\mathrm{Ann}_R M$, and then it is clear from the first part that M is Noetherian.

Now suppose that M is Noetherian. It is obviously finitely generated: call the generators $m_1, \dots, m_n$. Then $M^{\oplus n}$ is Noetherian, and we can map $R \to M^{\oplus n}$ by sending $r \in R$ to $(rm_1, \dots, rm_n)$. The element r is in the kernel if and only if it kills all the generators of M, which is equivalent to killing M. Thus, there is an injection of $R/Ann_R M$ into the Noetherian module $M^{\oplus n}$, and so $R/\mathrm{Ann}_R M$ is Noetherian, as required. □

We next want to prove that polynomial rings over a field are Noetherian. We shall give two proofs: the first is not standard. We observe:

Lemma. *If R is Noetherian and S is a module-finite extension of R, then every intermediate ring $R \subseteq B \subseteq S$ is module-finite over R, and is a Noetherian ring.*

Proof. S is a Noetherian R-module, and B is an R-submodule of S and therefore finitely generated. It is a Noetherian R-module. Since any ideal of B is an R-submodule of B, the fact that B has ACC for R-submodules implies that it has ACC for ideals. □

Theorem (Hilbert basis theorem). *The polynomial ring R in n variables over a field K is Noetherian. Hence, every finitely generated K-algebra is Noetherian.*

Proof. The second statement follows from the first because every finitely generated K-algebra is a homomorphic image of a polynomial ring.

We use induction on n. Let I be a nonzero ideal of R and $f \in I - \{0\}$. To show that I is finitely generated, it suffices to show that I/fR is finitely generated in R/fR: if $g_1, \dots, g_k$ are elements of I whose images $\overline{g_i}$ in I/fR generate I/fR, then $g_1, \dots, g_k$ together with f generate I. But we may assume that f is monic in x_n when viewed as an element of $K[x_1, \dots, x_{n-1}][x_n]$, so that R/fR is module-finite over $K[x_1, \dots, x_{n-1}]$, which is Noetherian by the induction hypothesis. It follows that R/fR is Noetherian. □

Our second proof has the advantage that it works whenever K is a Noetherian ring, not necessarily a field.

Theorem (Hilbert basis theorem). *Let R be a Noetherian ring. Then every finitely generated R-algebra is Noetherian.*

Proof. Since the rings considered are homomorphic images of polynomial rings in finitely many variables over R, we need only consider the case of a polynomial ring. By induction on the number of variables, it suffices to prove that if R is Noetherian, then $R[x]$ is Noetherian.

Let $J \subseteq R[x]$ be an ideal. For $t \in \mathbb{N}$, let $I_t \subseteq R$ be the set of elements of R that occur as leading coefficient of a polynomial of degree t in J, together with 0. It is easy to see that I_t is an ideal of R, and that $I_t \subseteq I_{t+1}$ since the leading coefficient of xf is the same as the leading coefficient of f. Thus, we can choose k such that $I_k = I_{k+1} = \cdots = I_{k+m} = \cdots$. For each t, $0 \le t \le k$, choose polynomials $f_{t,1}, \dots f_{t,h_t} \in J$ of degree t whose leading coefficients generate I_t. We claim that the $f_{t,s}$ generate J. Let J_0 be the ideal they generate, and suppose that $g \in J - J_0$ is of smallest possible degree. If g is of degree $t \le k$ we may subtract an R-linear combination j_0 of the f_{ts} (thus, $j_0 \in J_0$), that will cancel the leading term of g, and this will not introduce any terms of degree larger than t. Since $g - j_0 \in J_0$ (since g has minimum degree for elements in $J - J_0$), we have that $g \in J_0$, a contradiction.

If the degree of g is $d > k$, we can give essentially the same argument: now we subtract off an R-linear combination of the polynomials $x^{d-k} f_{k,s}$ to cancel the highest degree term. □

Lecture of October 18

Note that while a monic polynomial of degree d over a field or domain has at most d roots, nothing like this is true in rings with zerodivisors. For example, consider the ring of functions from an arbitrary set X taking values in a field K. This ring is reduced: the only nilpotent is the 0 function. But the functions on X taking on only the values 0 and 1 all satisfy the degree 2 monic equation $z^2 - z = 0$. There is one such function for every subset of X (the function that is 1 on that subset and 0 elsewhere). If X is countably infinite, the number of solutions of $z^2 - z = 0$ is uncountable.

From the Hilbert basis theorem (second version) we have at once:

Corollary. *A finitely generated algebra over a PID is Noetherian.* □

Proposition. *A localization of a Noetherian ring at any multiplicative system is Noetherian.*

Proof. The ideals of $S^{-1}R$ are in bijective order-preserving correspondence with the ideals of R that are contracted with respect to S. □

Since fields and principal ideal domains are Noetherian and the class of Noetherian rings is closed under taking homomorphic images, localizations, and finitely generated algebras, we have quite a few examples. Later we shall see that formal power series rings in finitely many variables over Noetherian rings are Noetherian.

Suppose that we want to prove a theorem about Noetherian modules (or Noetherian rings). One can assume that one has a counter-example M. Consider the family of all submodules N of M such that M/N is a counterexample, i.e., satisfies the hypothesis but not the conclusion of the theorem. This family contains the 0 submodule, and so is non-empty. Therefore it has a maximal element. One may therefore work with M/N instead of N, and now one may assume that every proper quotient of M satisfies the theorem. In case R is a ring, one is looking at quotients R/I and they are also rings. This method of proof is called *Noetherian induction.* Here is an example:

Theorem. *Every Noetherian ring has only finitely many minimal primes. Hence, every ideal of a Noetherian ring has only finitely many minimal primes.*

Proof. The second statement follows from the first by passing to the ring R/I. By Noetherian induction, we may assume that every proper quotient of R has only finitely many minimal primes. If R is a domain, we are done: the only minimal prime is (0). If R is not a domain we can choose nonzero elements x, y such that $xy = 0$. Every minimal prime of R either contains x or contains y. If the former holds it corresponds to a minimal prime of R/xR, and there are only finitely many of these by the hypothesis of Noetherian induction. Likewise, if it contains y it corresponds to a minimal prime of R/yR, and again, there are only finitely many minimal primes in R/yR by the hypothesis of Noetherian induction. □

We next return to the discussion of algebraic sets, and give another strong form of Hilbert's Nullstellensatz.

We now have available the theorem that the polynomial ring $K[x_1, \dots, x_n]$ is Noetherian. For every set of polynomials $S \subseteq K[x_1, \dots, x_n]$, $\mathcal{V}(S) = \mathcal{V}(I)$, where I is the ideal generated by S, and $\mathcal{V}(I) = \mathcal{V}(\mathrm{Rad}\, I)$, since $\mathcal{V}(f^n) = \mathcal{V}(f)$, always. Since every ideal is finitely generated, we may choose finitely many elements $f_1, \dots, f_m$ that generate I, or any ideal with the same radical as I, and then $\mathcal{V}(S) = \mathcal{V}(f_1, \dots, f_m) = \mathcal{V}(f_1) \cap \cdots \cap \mathcal{V}(f_m)$. We are now ready to prove another strong form of Hilbert's Nullstellensatz. If X is any subset of K^n, we write $\mathcal{I}(X) = \{f \in K[x_1, \dots, x_n] : \text{for all } x \in X, f(x) = 0\}$. Note that if $X = \{x\}$ has one point, then $\mathcal{I}(\{x\}) = m_x$, the maximal ideal consisting of all functions that vanish at x. Also note that $\mathcal{I}(X) = \cap_{x \in X} m_x$, and is always a radical ideal. These statements are all valid even without the assumption that K is algebraically closed. When K is algebraically closed, we can also state the following:

Theorem (Hilbert's Nullstellensatz, second strong form). *Let K be an algebraically closed field, and consider the polynomial ring $R = K[x_1, \dots, x_n]$ and algebraic sets in K^n. The functions $\mathcal{V}$ and $\mathcal{I}$ give a bijective order-reversing correspondence between radical ideals of R and closed algebraic sets in K^n.*

Proof. Let I be a radical ideal. We may write $I = (f_1, \dots, f_m)R$ for suitable f_j. We must show that $\mathcal{I}\big(\mathcal{V}(I)\big) = I$. The left hand side consists of all polynomials that vanish everywhere that the f_i vanish, and the earlier strong form of Hilbert's Nullstellensatz that we proved says precisely that if g vanishes on $\mathcal{V}(f_1, \dots, f_m)$, then $g \in \mathrm{Rad}\,(f_1, \dots, f_m) = (f_1, \dots, f_m)$ in this case, since we assumed that $I = (f_1, \dots, f_m)$ is radical.

What remains to be shown is that if X is an algebraic set then $\mathcal{V}\big(\mathcal{I}(X)\big) = X$. But since X is an algebraic set, we have that $X = \mathcal{V}(I)$ for some radical ideal I. Consequently, $\mathcal{V}\big(\mathcal{I}(X)\big) = \mathcal{V}\big(\mathcal{I}\big(\mathcal{V}(I)\big)\big) = \mathcal{V}(I)$, since $\mathcal{I}\big(\mathcal{V}(I)\big) = I$, by what we proved just above, and $\mathcal{V}(I) = X$. $\square$

Lecture of October 20

Proposition. *In $X = \operatorname{Spec}(R)$ where R is Noetherian, every closed set Z has finitely many maximal closed irreducible subsets, and it is the union of these. This union is irredundant, i.e., none of the maximal closed irreducible sets can be omitted. The maximal closed irreducible subsets of Z are the same as the maximal irreducible subsets of Z.*

If K is an algebraically closed field, the same statements apply to the closed algebraic sets in K^n.

Proof. The maximal irreducible closed subsets of Z correspond to the minimal primes $P_1, \ldots, P_n$ of the radical ideal I such that $V(I) = Z$, and this shows that Z is the union of the maximal irreducible closed sets $Z_i = V(P_i)$ contained in Z.

On the other hand, if Z is a finite union of mutually incomparable irreducible closed sets Z_i, then every irreducible subset W of Z is contained in one of them, for W is the union of the closed subsets $W \cap Z_i$, and so we must have $W = W \cap Z_i$ for some i, and thus $W \subseteq Z_i$. This proves that the Z_i are maximal irreducible subsets, and that none of them can be omitted from the union: if Z_j could be omitted it would be contained in the union of the others and therefore contained in one of the others.

The proof for the case of algebraic sets in K^n is the same. $\square$

In both contexts, the maximal irreducible closed subsets in Z are called the *irreducible components* of Z.

Irreducible closed algebraic sets in K^n, when K is algebraically closed, are called *algebraic varieties*. (To be precise, they are called *affine* algebraic varieties, but we shall not be dealing in this course with the other kinds. These include the irreducible closed algebraic sets in a projective space over K, which are called *projective varieties*, irreducible open sets in an affine variety, which are called *quasi-affine* varieties, and irreducible open sets in a projective variety, which are called *quasi-projective* varieties. The last type includes the others already mentioned. There is also an abstract notion of variety which is more general, but the most important examples are quasi-projective.)

The notation $\mathbb{A}^n_K$ is used for K^n to emphasize that is being thought of as an algebraic set (rather than as, say, a vector space).

Examples. In $\mathbb{A}^2_K$, $\mathcal{V}(x_1x_2) = \mathcal{V}(x_1) \cap \mathcal{V}(x_2)$ gives the representation of the algebraic set which is the union of the axes as an irredundant union of irreducible algebraic sets. This corresponds to the fact that in $K[x,\, y]$, $(xy) = (x) \cap (y)$. Now consider $\mathbb{A}^6_K$ where the variables are $x_1,\, x_2,\, x_3, y_1,\, y_2,\, y_3$, so that our polynomial ring is $R = K[x_1,\, x_2,\, x_3, y_1,\, y_2,\, y_3]$. Instead of thinking of algebraic sets as lying in $\mathbb{A}^6_K$, we shall think instead of them as sets of 2×3 matrices, where the values of the variables x_i and y_j are used to create a matrix as shown: $\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$. Let $\Delta_1 = x_2y_3 - x_3y_2$, $\Delta_2 = x_1y_3 - x_3y_1$ and $\Delta_3 = x_1y_2 - x_2y_1$ be the three 2×2 minors of this matrix. Consider the algebraic set $\mathcal{V}(\Delta_2, \Delta_3)$. We may think of this as the algebraic set of 2×3 matrices such that the minor formed from the first two columns and the minor formed from the first and third columns vanish. If a matrix is in this set, there are two possibilities. One is that the first column is zero, in which case the two minors involved do vanish. The second case is that the first column is not zero. In this case, the second and third columns are multiples of the first column, and this implies that Δ_1 vanishes. From this we obtain that $\mathcal{V}(\Delta_2,\, \Delta_3) = \mathcal{V}(x_1,\, y_1) \cup \mathcal{V}(\Delta_1,\, \Delta_2,\, \Delta_3)$. This does turn out to be the decomposition of $\mathcal{V}(\Delta_2,\, \Delta_3)$ as an irredundant union of irreducible components. The hardest part here is to show that $\mathcal{V}(\Delta_1,\, \Delta_2,\, \Delta_3)$ is irreducible.

A topological space is called *Noetherian* if it satisfies DCC on closed sets. Thus, Spec (R) is Noetherian iff the radical ideals of R have ACC, which is, of course true if R is Noetherian.

Proposition. *A subspace Y of a Noetherian topological space X is Noetherian. A Noetherian space is quasicompact. A topological space X is Noetherian if and only if every open subspace is quasicompact, in which case every subspace is quasicompact. In a Noetherian topological space, every closed subset is the finite irredundant union of its maximal closed irreducible subsets, which are the same as its irreducible subsets.*

Proof. For the first statement, it suffices to show that a non-increasing sequence of closed sets Y_i in Y is stable, and we can write $Y_i = Z_i \cap Y$, where Z_i is closed in X. Then the sequence $Z_1,\, Z_1 \cap Z_2,\, \ldots\,,\, Z_1 \cap \cdots \cap Z_n,\, \ldots$ is eventually stable in X, and the intersection of the n th term with Y is $Y_1 \cap \cdots \cap Y_n = Y_n$.

Consider next a family of closed sets in X with FIP. We must show the intersection is non-empty. We may assume without loss of generality that the family is closed under intersection. But it has a minimal element, and this must be contained in all of the sets, or we could intersect further, contradicting minimality.

Clearly, if X is Noetherian, then every subset is Noetherian and hence quasicompact, and so is every open subset. It suffices to show that if every open subset is quasicompact, then X is Noetherian. If not, let $Z_1 \supset Z_2 \supset \cdots \supset Z_n \supset \cdots$ be a strictly decreasing sequence of closed sets. Call the intersection Z. Then $X - Z$ is open, and is the strictly increasing union of the open sets $X - Z_n$. This gives an open cover with no finite sub-cover, contradicting the quasicompactness of X.

Finally, let Z be any closed set in X. If it is not a finite union of irreducibles, take a minimal counter-example. If Z itself is irreducible, we are done. If not then $Z = Z_1 \cup Z_2$, where these are proper closed subsets, and hence each is a finite union of irreducibles, since Z is a minimal counterexample. Once we have Z as a finite union of irreducibles, we can omit terms until we have Z as an irredundant finite union of irreducibles, say $Z = Z_1 \cup \cdots \cup Z_n$. Now, if Y is an irreducible set contained in Z, it must be contained in one of Z_i, since it is the union of its intersections with the Z_i, which shows that the Z_i are the maximal irreducible sets contained in Z, as well as the maximal irreducible closed sets contained in Z. □

We next want to make the closed algebraic sets over an algebraically closed field K into a category. Suppose we are given $X \subseteq K^n$ and $Y \subseteq K^m$. We could write $\mathbb{A}^n_K$ instead of K^n and $\mathbb{A}^m_K$ instead of K^m. We define a function $f : X \to Y$ to be *regular* if it there exist polynomials $g_1, \ldots, g_m \in K[x_1, \ldots, x_n]$ such that for all points $x \in X$,

$f(x) = \big(g_1(x), \, \dots \, , g_m(x)\big)$. Thus, the function f can be given by a polynomial formula in the coordinates. It is easy to verify that the identity function is regular and that the composition of two regular functions is regular. The closed algebraic sets over K become a category if we define $\mathrm{Mor}\,(X, Y)$ to be the set of regular functions from X to Y.

It may seem a bit artificial to require that a map of $X \subseteq \mathbb{A}^n_K$ to $Y \subseteq \mathbb{A}^m_K$ be induced by a map from $\mathbb{A}^n_K$ to $\mathbb{A}^m_K$ (the polynomials g_j in the definition of regular map actually give a map $K^n \to K^m$ that happens to take X into Y). However, this is not much different from the situation in topology.

Most of the objects of interest in topology (compact manifolds or compact manifolds with boundary) are embeddable as closed sets in $\mathbb{R}^n$ for some n. If $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$, then every continuous function from X to Y is the restriction of a continuous function from $\mathbb{R}^n \to \mathbb{R}^m$. To see this, think about the composition $X \to Y \subseteq \mathbb{R}^m$. The function $X \to \mathbb{R}^m$ is given by an m-tuple of continuous functions from X to $\mathbb{R}$. But a continuous function from a closed set $X \subseteq \mathbb{R}^n$ to $\mathbb{R}$ does extend to a continuous function from $\mathbb{R}^n$ to $\mathbb{R}$: this is the Tietze extension theorem, and uses only that $\mathbb{R}^n$ is a normal topological space.

We now enlarge the category of algebraic sets slightly. Given an algebraic set X and mutually inverse set bijections $\alpha : X' \to X$ and $\beta : X \to X'$ we shall think of these maps as giving X' the structure of an algebraic set. We define a map $f : X' \to Y$ to be *regular* if $f \circ \beta$ is regular, and a map $g : Y \to X'$ to be *regular* if $\alpha \circ g$ is regular.

Of course if we have also given, say, Y', the structure of an algebraic set via mutually inverse set isomorphisms $\gamma : Y' \to Y$ and $\delta : Y \to Y'$ with an algebraic set Y, then $f : X' \to Y'$ is *regular* if $\gamma \circ f \circ \beta$ is a regular function from X to Y, while $g : Y' \to X'$ is *regular* if $\alpha \circ g \circ \delta$ is a regular function from Y to X.

More generally, given any category in which the objects have underlying sets and the morphisms are functions on the underlying sets with, possibly, some further restrictive property (groups and group homomorphisms, rings and ring homomorphisms, and topological spaces and continuous maps are examples), one can make an entirely similar construction: given a bijection $\alpha : X' \to X$ one can introduce an object with underlying set X' into the category in such a way that α is an isomorphism of that new object with X. In the case of rings, one uses the bijection to introduce addition and multiplication on X': one adds elements of X' by taking the images of the elements in X, adding them in X, and then applying the inverse bijection to the sum to get an element of X'. One introduces multiplication in X' in an entirely similar way.

Given a closed algebraic set $X \subseteq \mathbb{A}^n_K$, the regular functions to K (i.e., to $\mathbb{A}^1_K$) have the structure of a K-algebra: the restrictions of polynomials g_1 and g_2 to X have a sum (respectively, a product) that is regular because it is the restriction of $g_1 + g_2$ (respectively, $g_1 g_2$). This ring is called the *coordinate ring* of X and is denoted $K[X]$. It is a reduced finitely generated K-algebra: if a power of a function is 0, all of its values are nilpotent in K and therefore 0 in K, so that the function is identically zero. The coordinate ring is generated over K by the images of the n functions represented by the variables $x_1, \, \dots \, , x_n$. The function x_i assigns to a point in X its ith coordinate, and so the functions x_i are referred to as *coordinate functions*, which explains why the K-algebra they generate is called the coordinate ring.

$K[X]$ is a homomorphic image of $K[x_1, \dots, x_n]$ under the K-algebra homomorphism that sends the function given by a polynomial $g \in K[x_1, \dots, x_n]$ to its restriction to X. The kernel of this K-algebra homomorphism is the ideal $\mathcal{I}(X)$ of all polynomial functions that vanish on X, and so we have a K-algebra isomorphism $K[x_1, \dots, x_n]/\mathcal{I}(X) \cong K[X]$.

In fact, $\mathcal{F} = \mathrm{Mor}\,(\_\,, \mathbb{A}^1_K)$ is a contravariant functor from algebraic sets to reduced finitely generated K-algebras. Given a map of algebraic sets $f : X \to Y$ there is a K-algebra homomorphism $f^* : K[Y] \to K[X]$ induced by composition; for each $g : Y \to \mathbb{A}^1_K$, we let $f^*(g) = g \circ f : X \to \mathbb{A}^1_K$.

Now consider the functor $\mathcal{G} = \mathrm{Hom}_{K\text{-alg}}(\_\,, K)$ from reduced finitely generated K-algebras to algebraic sets. Here the subscript indicates that we are dealing with K-algebra homomorphisms. For this to make sense, we have to give $\mathrm{Hom}_{K\text{-alg}}(R, K)$ the structure of an algebraic set: we do this by choosing a finite set of algebra generators $r_1, \dots, r_n$ for R over K, and then mapping $\mathrm{Hom}_{K\text{-alg}}(R, K)$ to $\mathbb{A}^n_K$ by sending $\phi \in \mathrm{Hom}_{K\text{-alg}}(R, K)$ to the n-tuple $\bigl(\phi(r_1), \dots, \phi(r_n)\bigr) \in \mathbb{A}^n_K$. We shall see below that the set of values of this map is a closed algebraic set in $\mathbb{A}^n_K$, and that, up to isomorphism, this algebraic set is independent of the choice of a finite set of generators for R over K. Thus, $\mathrm{Hom}_{K\text{-alg}}(R, K)$ has the structure of an algebraic set. Moreover, $\mathrm{Hom}_{K\text{-alg}}(\_\,, K)$ is a contravariant functor: if $h : R \to S$ is a K-algebra homomorphism, we get a map $h^* : \mathrm{Hom}_{K\text{-alg}}(S, K)$ to $\mathrm{Hom}_{K\text{-alg}}(R, K)$ induced by composition: $h^*(\theta) = \theta \circ h$. We shall see that this makes $\mathcal{G} = \mathrm{Hom}_{K\text{-alg}}(\_\,, K)$ into a contravariant functor from reduced finitely generated K-algebras to closed algebraic sets over K.

Note that the elements of $\mathrm{Hom}_{K\text{-alg}}(R, K)$ correspond bijectively with the maximal ideals of R: the maximal ideal is recovered from a given homomorphism as its kernel. On the other hand, we have already seen that for any maximal ideal m, $K \to R/m$ is an isomorphism μ when K is algebraically closed, and we may compose $R \to R/m$ with μ^{-1} to obtain a K-algebra homomorphism $R \twoheadrightarrow K$ whose kernel is the specified maximal ideal m. Note that if we have $\theta : S \twoheadrightarrow K$ and we compose with $f : R \to S$, the kernel of the composition $R \to S \twoheadrightarrow K$ is the same as the contraction of the kernel of θ to R. Thus, the functor MaxSpec is isomorphic with $\mathcal{G} = \mathrm{Hom}_{K\text{-alg}}(\_\,, K)$, and so we could have worked with this functor instead of $\mathcal{G}$. In particular, we can give every MaxSpec$\,(R)$ the structure of an algebraic set.

Our main result in this direction is:

Theorem. *The procedure for giving Hom*$_{K\text{-alg}}(R, K)$ *the structure of an algebraic set described above does produce a bijection with an algebraic set, and changing the choice of the finite set of generators for R produces an isomorphic algebraic set. $\mathcal{F}$ and $\mathcal{G}$ as described above are contravariant functors such that $\mathcal{G} \circ \mathcal{F}$ is isomorphic with the identity functor on closed algebraic sets over K, and $\mathcal{F} \circ \mathcal{G}$ is isomorphic with the identity functor on reduced finitely generated K-algebras. Thus, the category of closed algebraic sets and regular functions over the algebraically closed field K is anti-equivalent to the category of reduced finitely generated K-algebras.*

Lecture of October 23

Proof. We first note that the points of the closed algebraic set X correspond bijectively in an obvious way with the elements of $\mathrm{Hom}_{K\text{-alg}}(K[X],\,K)$, and, likewise, with the maximal ideals of $K[X]$. Think of $K[X]$, as usual, as $K[x_1,\,\ldots\,,x_n]/\mathcal{I}(X)$. The maximal ideals of this ring correspond to maximal ideals of $K[x_1,\,\ldots\,,x_n]$ containing $\mathcal{I}(X)$. Each such maximal ideal has the form m_y for some $y \in \mathbb{A}^n_K$, and the condition that y must satisfy is that $\mathcal{I}(X) \subseteq m_y$, i.e., that all functions in $\mathcal{I}(X)$ vanish at y, which says that $y \in \mathcal{V}\bigl(\mathcal{I}(X)\bigr)$. By our second strong version of Hilbert's Nullstellensatz (Lecture of October 18), $\mathcal{V}\bigl(\mathcal{I}(X)\bigr) = X$.

We next note that our procedure for assigning the structure of an algebraic set to $\mathrm{Hom}_{K\text{-alg}}(R,\,K)$ really does give an algebraic set, which is independent, up to isomorphism, of the choice of the set of generators of R as a K-algebra. To see this, let $r_1,\,\ldots\,,r_n$ be one set of generators of R. Map $K[x_1,\,\ldots\,,x_n] \twoheadrightarrow R$ using the unique K-algebra homomorphism that sends $x_i \mapsto r_i$, $1 \leq i \leq n$. Let I be the radical ideal which is the kernel of this homomorphism, so that $R \cong K[x_1,\,\ldots\,,x_n]/I$. The set we assigned to $\mathrm{Hom}_{K\text{-alg}}(R,\,K)$ is $\{\bigl(h(r_1),\,\ldots\,,\,h(r_n)\bigr) : h \in \mathrm{Hom}_{K\text{-alg}}(R,\,K)\}$. Each K-homomorphism h is uniquely determined by its values on the generators $r_1,\,\ldots\,,r_n$. An n-tuple $(\lambda_1,\,\ldots\,,\lambda_n)$ can be used to define a K-homomorphism if and only if the elements of I vanish on $(\lambda_1,\,\ldots\,,\lambda_n)$, i.e., if and only if $(\lambda_1,\,\ldots\,,\lambda_n) \in \mathcal{V}(I)$. This shows that our map from $\mathrm{Hom}_{K\text{-alg}}(R,\,K)\}$ to K^n gives a bijection of $\mathrm{Hom}_{K\text{-alg}}(R,\,K)\}$ with the algebraic set $\mathcal{V}(I)$.

Now suppose that $r'_1,\,\ldots\,,r'_m$ are additional elements of R. For every r'_j we can choose $g_j \in K[x_1,\,\ldots\,,x_n]$ such that $r'_j = g_j(r_1,\,\ldots\,,r_n)$. The new algebraic set that we get by evaluating every element $h \in \mathrm{Hom}_{K\text{-alg}}(R,\,K)\}$ on $r_1,\,\ldots\,,r_m\,,r'_1,\,\ldots\,,r'_m$ is precisely $X' = \{\bigl(\lambda_1,\,\ldots\,,\lambda_n,\,g_1(\underline{\lambda}),\,\ldots\,,\,g_m(\underline{\lambda})\bigr) : \underline{\lambda} \in X\}$, where $\underline{\lambda} = (\lambda_1,\,\ldots\,,\lambda_n)$. The map $X \to X'$ that sends $\underline{\lambda} = (\lambda_1,\,\ldots\,,\lambda_n)$ to $\bigl(\lambda_1,\,\ldots\,,\lambda_n,\,g_1(\underline{\lambda}),\,\ldots\,,g_m(\underline{\lambda})\bigr)$ is given in coordinates by the polynomials $x_1,\,\ldots\,,x_n,\,g_1,\,\ldots\,,g_m$, and so is a morphism in the category of algebraic sets. Likewise, the map $X' \to X$ which is simply projection on the first n coordinates is given by polynomials in the coordinates, and these are mutually inverse morphisms of algebraic sets. Thus, $X \cong X'$, as required.

This handles the case where one set of generators is contained in another. But now, if $r_1,\,\ldots\,,r_n$ and $r'_1,\,\ldots\,,r'_m$ are two sets of generators, we may compare the algebraic set given by $r_1,\,\ldots\,,r_n$ with that given by $r_1,\,\ldots\,,r_n,\,r'_1,\,\ldots\,,r'_m$, and then the latter with the algebraic set given by $r'_1,\,\ldots\,,r'_m$. This completes the proof of the independence of the algebraic set structure that we are assigning to $\mathrm{Hom}_{K\text{-alg}}(R,\,K)$ from the choice of K-algebra generators for R.

If $R = K[X]$ and we choose as generators r_i the restrictions of the coordinate functions x_i to R, then the algebraic set we get from $\mathrm{Hom}_{K\text{-alg}}(K[X],\,K)$ is X itself, and this is the same identification of X with $\mathrm{Hom}_{K\text{-alg}}(K[X],\,K)$ that we made in the first paragraph. Thus, if we let $S_X : X \to \mathrm{Hom}_{K\text{-alg}}(K[X],\,K)$ as in that paragraph, we get an isomorphism of algebraic sets, for we may use the restricted coordinate functions as the generators to place the algebraic set structure on $\mathrm{Hom}_{K\text{-alg}}(K[X],\, K) = (\mathcal{G}\circ\mathcal{F})(X)$. We claim that S_X is a natural transformation from the identity functor on the category of algebraic sets over K to $\mathcal{G}\circ\mathcal{F}$. We need to see that if $\theta : X \to Y$ is a morphism of algebraic sets, then $(\mathcal{G}\circ\mathcal{F})(\theta)$ is the same as θ once we identify $\mathrm{Hom}_{K\text{-alg}}(K[X],\, K)$ with X and $\mathrm{Hom}_{K\text{-alg}}(K[X],\, K)$ with Y. Let ϕ_x (resp., ϕ'_y) denote evaluation as at $x \in X$ (resp., $y \in Y$). We need to show that $\bigl((\mathcal{G}\circ\mathcal{F})(\theta)\bigr)(\phi_x) = \phi'_{\theta(x)}$ for all $x \in X$. Now, $\mathcal{F}(\theta)$ acting on $v \in K[Y]$ is $v \circ \theta$, and $\mathcal{G}$ applied to $\mathcal{F}(\theta)$ acts by composition as well, so that its value on ϕ_x is the map that sends $v \in K[Y]$ to $(v\circ\theta)(x) = v\bigl(\theta(x)\bigr)$, which is evaluation at $\theta(x)$, as required.

Finally, we need to see that $\mathcal{F}\circ\mathcal{G}$ is isomorphic to the identity functor on finitely generated reduced K-algebras. The map sends R to $K[\mathrm{Hom}_{K\text{-alg}}(R,\, K)]$ where $\mathrm{Hom}_{K\text{-alg}}(R,\, K)$ is viewed as a closed algebraic set as discussed above. Each element r of R maps to a function f_r on the set $\mathrm{Hom}_{K\text{-alg}}(R,\, K)$ by the rule $f_r(u) = u(r)$. It is immediate that this is a K-algebra homomorphism: call it T_R. We shall show that the T_R give an isomorphism of the identity functor with $\mathcal{F}\circ\mathcal{G}$. We first need to show that every T_R is an isomorphism. We use the fact that $R \cong K[x_1,\, \dots, x_n]/I$ for some radical ideal I, with the coordinate functions as generators, and it suffices to consider the case where $R = K[x_1,\, \dots, x_n]/I$. This identifies $\mathrm{Hom}_{K\text{-alg}}(R,\, K)$ with $\mathcal{V}(I)$, and the needed isomorphism follows from the fact that $K[\mathcal{V}(I)] \cong K[x_1,\, \dots, x_n]/\mathcal{I}\bigl(\mathcal{V}(I)\bigr) = K[x_1,\, \dots, x_n]/I$, again by the second strong version of Hilbert's Nullstellensatz (Lecture of October 18).

The last step is to check that T is a natural transformation. Consider a K-algebra homomorphism $\alpha : R \to S$. Choose a K-algebra homomorphism γ of polynomial ring $A = K[y_1,\, \dots, y_m]$ onto R with kernel I and a K-algebra homomorphism δ of a polynomial ring $B = K[x_1,\, \dots, x_n]$ onto S with kernel J. Without loss of generality, we may assume that $R = A/I$, $S = B/J$. Choose $g_1,\, \dots, g_m \in K[x_1,\, \dots, x_n]$ such that the image of y_j in R maps to the image of g_j in B, $1 \leq j \leq m$, so that α is induced by the K-algebra map $A \to B$ that sends y_j to g_j, $1 \leq j \leq m$. The corresponding map of algebraic sets $\mathcal{V}(J) \to \mathcal{V}(I)$ is given in coordinates by the g_j. Finally, the induced map $K[V(I)] \cong A/\mathcal{I}\bigl(\mathcal{V}(I)\bigr) = A/I$ to $K[V(J)] \cong B/\mathcal{I}\bigl(\mathcal{V}(J)\bigr) = B/J$ is induced by composition with the map given by the polynomials $g_1,\, \dots, g_m$. This means that the image of an element of A/I represented by $P(y_1,\, \dots, y_m) \in A$ is represented by the coset in B/J of $P(g_1,\, \dots, g_m) \in B$, and this shows that with the identifications we are making, $\mathcal{F}\circ\mathcal{G}(\alpha)$ is α, which is exactly what we need. □

Given an algebraic set X over an algebraically closed field K, we define $\dim(X)$ to be the same as $\dim(K[X])$. The dimension of a ring is the supremum of the dimensions of its quotients by minimal primes. Thus, $\dim(X)$ is the same as the supremum of the dimensions of the irreducible components of X. Evidently, $\dim(X)$ is also the same as the supremum of lengths of chains of irreducible closed subsets of X. We define the dimension of X near a point $x \in X$ to the be the supremum of the dimensions of the irreducible components of X that contain x. If the corresponding maximal ideal of $R = K[X]$ is $m = m_x$, this is also the dimension of R_m: it has minimal primes P corresponding precisely to the irreducible components $V(P)$ that contain x, and the length of any saturated chain from P to m $= \dim(R_m/PR_m) = \dim(R/P) =$ the dimension of the irreducible component $V(P)$, from which the result follows.

There are at least three ways to think of an algebra R over a commutative ring ring K. It is worth considering all three points of view. One is purely algebraic: R is an abstract algebraic environment in which one may perform certain sorts of algebraic manipulations.

A second point of view is to think of R, or rather some topological space associated with R, as a geometric object. We have seen explicitly how to do this when R is a finitely generated reduced K-algebra and K is an algebraically closed field. But a geometric point of view, introduced by A. Grothendieck, can be taken in great generality, when R is *any* commutative ring. In Grothendieck's theory of schemes, a geometric object $\mathrm{Spec}\,(R)$, is introduced that has more structure than just the topological space of prime ideals of R that we have talked about here. The geometric point of view has been very effective as a tool in commutative algebra, even if one is only interested in seemingly purely algebraic properties of rings.

The third point of view is simplest when R is a finitely generated algebra over a Noetherian ring K (and it simplest of all when K is a field). In this case one has that $R = K[x_1, \ldots, x_n]/(f_1, \ldots, f_m)$. Now let S be any K-algebra. Then $\mathrm{Hom}_{K\text{-alg}}(R, S)$ is in bijective correspondence with the set of solutions of the set of m simultaneous equations

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

in S^n, for to give a K-homomorphism from R to S is the same as to give an n-tuple of elements of S (which will serve as the values of the homomorphism on the images of the variables $x_1, \ldots, x_n$) that satisfy these equations. The set of homomorphisms $\mathrm{Hom}_{K\text{-alg}}(R, S)$ is called the *set of S-valued points* of the scheme $\mathrm{Spec}\,(R)$ in scheme theory: since we don't have that theory available, we shall simply refer to it as the set of S-valued points of R. Recall again that K can be any Noetherian ring here. This point of view can be extended: we do not need to assume that R is finitely generated over K, nor that K is Noetherian, if we allow infinitely many variables in our polynomial ring, and infinite families of polynomial equations to solve. Thus, very generally, a K-algebra may be thought of as an encoded system of equations. When one takes homomorphisms into S, one is solving the equations in S. A different way to say this is the following: suppose that we start with a system of equations over K, and define a functor from K-algebras to sets that assigns to every K-algebra S the set of solutions of the family of equations such that the values of the variables are in S. If one forms the polynomial ring in the variables occurring and then the quotient by the ideal generated by the polynomials set equal to 0 in the equations, the resulting K-algebra represents this functor.

Here is an example. Let $B = \mathbb{R}[X, Y, Z]/(X^2 + Y^2 + Z^2 - 1) = \mathbb{R}[x, y, z]$, and let $S = B[U, V, W]/(xU + yV + zW) = \mathbb{R}[x, y, z, u, v, w]$. We can also form B in a single step as $\mathbb{R}[X, Y, Z, U, V, W]]/(X^2+Y^2+Z^2-1, XU+YV+ZW)$. The $\mathbb{R}$-homomorphisms from B or $\mathbb{R}$-valued points of B correspond to the set $\{(a, b, c) \in \mathbb{R}^3 : a^2 + b^2 + c^2 = 1\}$: the real 2-sphere of radius one centered at the origin in $\mathbb{R}^3$. The $\mathbb{R}$-valued points of S correspond to pairs (a, b, c), (d, e, f) such that $(a, b, c) \in S^2$ and $(a, b, c) \cdot (d, e, f) = 0$, which means that the vector (d, e, f) represents a tangent vector to the sphere at the point (a, b, c). That is, the $\mathbb{R}$-valued points of S correspond to the points of the tangent bundle to the real 2-sphere. It turns out that if T is a new indeterminate over S and T_1, T_2, T_3 are three new indeterminates over A, then $S[T] \cong A[T_1, T_2, T_3]$, but that S is not isomorphic with $A[T_1, T_2]$. This answers the question raised by the exercise in the book of Deskins discussed during the Lecture of September 22. One key point is that the direct sum of the tangent bundle to the 2-sphere and a trivial line bundle is a trivial vector bundle of rank 3, but that the tangent bundle to the 2-sphere is non-trivial: in fact, it has no non-vanishing section. This last statement amounts to the assertion that there is no non-vanishing continuous field of tangent vectors on a 2-sphere. Sometimes this is expressed by saying "You can't comb the hair on a billiard ball." The question as to whether $S[T] \cong S'[T']$ implies $S \cong S'$ appears to be purely algebraic. There may be a moral in the fact that the simplest counter-example requires some substantial knowledge from topology.

We next want to explore the notion of a (formal) power series ring in finitely many variables over a ring R, and show that it is Noetherian when R is. But we begin with a definition in much greater generality.

Let S be a commutative semigroup (which will have identity $1_S = 1$) written multiplicatively. We shall assume that S has the following property:

(#) For all $s \in S$, $\{(s_1, s_2) \in S \times S : s_1 s_2 = s\}$ is finite.

Thus, each element of S has only finitely many factorizations as a product of two elements. For example, we may take S to be the set of all monomials $\{x_1^{k_1} \cdots x_n^{k_n} : (k_1, \ldots, k_n) \in \mathbb{N}^n\}$ in n variables. We construct a ring denoted $R[[S]]$: we may think of this ring formally as consisting of all functions from S to R, but we shall indicate elements of the ring notationally as (possibly infinite) formal sums $\sum_{s \in S} r_s s$, where the function corresponding to this formal sum maps s to r_s for all $s \in S$. Addition is performed by adding corresponding coefficients, while $(\sum_{s \in S} r_s s)(\sum_{s' \in S} r_{s'} s')$ is defined to be

$$\sum_{t \in S} (\sum_{s, s' \in S, ss' = t} r_s r_{s'}) t.$$

Heuristically, this is what one would get by distributing the product in all possible ways, and then "collecting terms": this is possible because, by (#), only finitely many terms $r_s r_{s'} ss'$ occur for any particular $t = ss'$. The ring has identity corresponding to the sum in which 1_S has coefficient $1 = 1_R$ and all other coefficients are 0. It is straightforward to verify all the ring laws and the commutativity of multiplication. $R[S]$, the semigroup ring defined earlier, is a subring: it may be identified with the formal sums in which all but finitely many coefficients are 0. One frequently omits terms with coefficient 0 from the notation. If $S = \{x_1^{k_1} \cdots x_n^{k_n} : (k_1, \ldots, k_n) \in \mathbb{N}^N\}$, the notation $R[[x_1, \ldots, x_n]]$ is used instead of $R[[S]]$: one writes generators for S inside the double brackets instead of S itself.

If S and S' both satisfy (#), so does the product semigroup $S \times S'$, and one has the isomorphism $(R[[S]])[[S']] \cong R[[S \times S']]$. If the coefficient of s' in an element of the former is $\sum_{s \in S} r_{s,s'} s$ for every $s' \in S'$, one identifies $\sum_{s' \in S'} (\sum_{s \in S} r_{s,s'} s) s'$ with $\sum_{(s,s') \in S \times S'} r_{s,s'} (ss')$. It is straightforward to check that this is an isomorphism.

Lecture of October 25

The ring $R[[x_1, \ldots, x_n]]$ is referred to as a *(formal) power series* ring over R, and the x_i are called *formal* or *analytic* indeterminates to indicate that two power series agree if and only if their corresponding coefficients are all identical.

In the case of two finite semigroups of monomials, the fact that $R[[S \times S']] \cong (R[[S]])[[S']]$ implies that

$$(R[[x_1, \ldots, x_n]])[[y_1, \ldots, y_m]] \cong R[[x_1, \ldots, x_n, y_1, \ldots, y_m]].$$

In particular, for $n \geq 2$,

$$R[[x_1, \ldots, x_n]] \cong (R[[x_1, \ldots, x_{n-1}]])[[x_n]].$$

Of course, there is a completely analogous statement for polynomial rings, with single brackets replacing double brackets. However, note that while

$$(R[[x_1, \ldots, x_n]])[y_1, \ldots, y_m] \hookrightarrow (R[y_1, \ldots, y_m])[[x_1, \ldots, x_n]],$$

the opposite inclusion always fails when R is not 0 and m, $n \geq 1$. First, to see the inclusion, note that if one has a homomorphism $h : R \to T$ there is an induced homomorphism $R[[x_1, \ldots, x_n]] \to T[[x_1, \ldots, x_n]]$: apply h to every coefficient. Let $T = R[y_1, \ldots, y_m]$ and h be the inclusion $R \subseteq T$ to get an injection

$$R[[x_1, \ldots, x_n]] \to (R[y_1, \ldots, y_m])[[x_1, \ldots, x_n]].$$

Now extend this homomorphism of $R[[x_1, \ldots, x_n]]$-algebras to the polynomial ring

$$(R[[x_1, \ldots, x_n]])[y_1, \ldots, y_m]$$

by letting y_i map to $y_i \in (R[y_1, \ldots, y_m])[[x_1, \ldots, x_n]]$. To see that the inclusion is typically strict, note that $\sum_{t=0}^{\infty} y_1^t x_1^t$ is an element of $(R[y_1, \ldots, y_m])[[x_1, \ldots, x_n]]$ but is not in $(R[[x_1, \ldots, x_n]])[y_1, \ldots, y_m]$, where every element has bounded degree in the y_j. Both rings inject into $R[[x_1, \ldots, x_n, y_1, \ldots, y_m]]$.

Theorem. *If R is Noetherian ring then the formal power series ring $R[[x_1, \ldots, x_n]]$ is Noetherian.*

Proof. By induction on the number of variables one reduces at once to proving that $S = R[[x]]$ is Noetherian. Let $J \subseteq R[[x]]$ be an ideal. Let I_t denote the set of elements r of R such that rx^n is the term of least degree in an element of J, together with 0. This is easily verified to be an ideal of R. If $f \in J$ is not zero, and rx^n is the least degree term in f, then rx^{n+1} is the least degree term in $xf \in J$. This shows that $\{I_t\}_{t \geq 0}$ is a non-decreasing sequence of ideals of R. Since R is Noetherian, we may choose $k \in \mathbb{N}$ such that $I_k = I_{k+1} = \cdots = I_{k+m} = \cdots$, and then for $0 \leq t \leq k$ we may choose

$f_{1,t}, \dots, f_{h_t,t} \in J$ such that each $f_{i,t}$ has smallest degree term of the form $r_{i,t}x^t$ and the elements $r_{i,t}, \dots, r_{h_t,t}$ are a finite set of generators of I_t. We claim that the finite set of power series $f_{i,t}$, $0 \le t \le k$, $1 \le i \le h_t$, generates J. Let J_0 be the ideal they generate, and let $u \in J$ be given. We may subtract an R-linear combination of the $f_{i,0}$ from u to get an element of J whose lowest degree term is in degree at least one (or such that the difference is 0). We continue in this way so long as we have a lowest degree term of degree less than k: if the degree is $t < k$, we may increase it by subtracting an R-linear combination of the $f_{i,\,t}$. Thus, after subtracting an element of J_0 from u, we may assume without loss of generality that the lowest degree term in u occurs in degree $\ge k$ (or else u is 0, but then there is nothing to prove). It will suffice to prove that this new choice of u is in J_0. We claim more: we shall show that in this case, u is in the ideal generated by the $f_{i,k} = f_i$. Let $h = h_k$. We recursively construct the partial sums (which are polynomials) of power series g_i such that $u = \sum_{i=1}^h g_i f_i$.

Put slightly differently and more precisely, we shall construct, for every i, $1 \le i \le h$, by induction on $m \in \mathbb{N}$, a sequence of polynomials $g_{i,m}(x) \in R[x]$ with the following properties:

(1) Every $g_{i,m}$ has degree at most m.

(2) If $m_1 < m_2$ then g_{i,m_1} is the sum of the terms of degree at most m_1 that occur in g_{i,m_2}. Given (1), this is equivalent to the condition that for all $m \ge 0$, $g_{i,m+1} - g_{i,m}$ has the form rx^{m+1} for some $r \in R$, which may be 0.

(3) For every m, the lowest degree term in $u - \sum_{i=1}^h g_{i,m} f_i$ has degree at least $k + m + 1$ (or else the difference is 0).

Notice that conditions (1) and (2) together imply that for every i, the $g_{i,m}$ are the partial sums of a formal power series, where the m th partial sum of a power series $\sum_{j=0}^\infty r_j x^j$ is defined to be $\sum_{j=0}^m r_j x^j$.

To begin the induction, note that the least degree term of u occurs in degree k or higher. Therefore the coefficient of x^k in u is in the ideal generated by the lowest degree coefficients of $f_1, \dots, f_h$, and it follows that there are elements $r_{1,0}, \dots, r_{h,0}$ of R such that the lowest degree term of $u - \sum_{i=1}^h r_{i,0} f_i$ occurs in degree at least $k + 1$ (or the difference is 0). We take $g_{i,0} = r_{i,0}$, $1 \le i \le h$.

Now suppose that the $g_{i,s}$ have been constructed for $1 \le i \le h$, $0 \le s \le m$ such that conditions (1), (2), and (3) are satisfied. We shall show that we can construct $g_{i,m+1}$ so that (1), (2), and (3) are satisfied. Since $u' = u - \sum_{i=1}^h g_{i,m} f_i$ has lowest degree term of degree at least $m+k+1$, the coefficient of x^{m+k+1} is in the R-span of the coefficients of x^k in the polynomials f_i, and so we can choose elements $r_{i,m+1} \in R$ so that $u' - \sum_{i=1}^h r_{i,m+1} x^{m+1} f_i$ has lowest degree term in degree at least $m + k + 2$ (or is 0). It follows that if we take $g_{i,m+1} = g_{i,m} + r_{i,m+1} x^{m+1}$ for $1 \le i \le h$, then (1) and (2) are satisfied, and (3) is as well because $u - \sum_{i=1}^h g_{i,m+1} f_i = u' - \sum_{i=1}^h r_{i,m+1} x^{m+1} f_i$ has lowest degree term in degree at least $m + k + 2$ (or the difference is 0). For each i, $1 \le i \le h$, let g_i be the formal power series whose partial sums are the $g_{i,m}$.

We claim that $u = \sum_{i=1}^h g_i f_i$. It suffices to show that the coefficients on corresponding powers of x are the same on both sides. Neither side has a nonzero term involving x^t for

$t < k$. On the other hand, for all $m \geq 0$, the coefficient of x^{k+m} on the right will not change if we replace every g_i on the right by $g_{i,m}$, since $g_i - g_{i,m}$ involves only terms of degree strictly bigger than $m + k + 1$. Thus, it suffices to show that for all $m \geq 0$, the difference $u - \sum_{i=1}^{h} g_{i,m} f_i$ has coefficient 0 on x^{m+k}, and this is true by part (3). But the $f_i = f_{i,k}$ are in J_0, so that $u \in J_0$, as required. □

It is also true that the subring of $\mathbb{C}[[x_1, \ldots, x_n]]$ (respectively, $\mathbb{R}[[x_1, \ldots, x_n]]$) consisting of power series that converge on a neighborhood of the origin in $\mathbb{C}^n$ (respectively, $\mathbb{R}^n$) is a Noetherian ring with a unique maximal ideal, generated by $x_1, \ldots, x_n$. These rings are denoted $\mathbb{C}\langle\langle x_1, \ldots, x_n\rangle\rangle$ and $\mathbb{R}\langle\langle x_1, \ldots, x_n\rangle\rangle$, respectively.

The Noetherian property of the ring $\mathbb{C}\langle\langle x_1, \ldots, x_n\rangle\rangle$ is of considerable usefulness in studying functions of several complex variables: this is the ring of germs of holomorphic functions at a point in $\mathbb{C}^n$. We shall not give the proof of the Noetherian property for convergent power series rings here: proofs may be found in [O. Zariski and P. Samuel, *Commutative Algebra*, Vol. II, Van Nostrand, Princeton, 1960], pp. 142–148 or [M. Nagata, *Local Rings*, Interscience, New York, 1962], pp. 190–194.

The operations of taking quotients, localization, forming polynomial rings in finitely many variables, and forming formal power series rings in finitely many variables have immensely increased the class of examples of Noetherian rings available to us. We may perform several iterations of these operations on a known Noetherian ring to create new examples, and the order in which the operations are done usually matters, although two operations of the same kind can be combined into one, and localization commutes with formation of quotient rings.

Here is a simple example of a ring that we have not looked at yet: $V_p = \mathbb{Z}[[x]]/(x - p)$, where p is a positive prime integer. We shall see later that this ring is a PID with a unique maximal ideal, i.e., a discrete valuation ring, in which p generates the maximal ideal. In this ring, we can make sense of a formal power series in p with integer coefficients: the same power series can be written down with x replacing p, and so has a meaning in $\mathbb{Z}[[x]]$, and it therefore represents an element of the quotient. For example, $1 + p + p^2 + p^3 + \cdots$ can be interpreted as the image of $1 + x + x^2 + x^3 + \cdots$ in the quotient. It turns out that the value of $1 + p + p^2 + p^3 + \cdots$ is an inverse for $1 - p$, just as if p were a small real number and we were using the formula for the sum of an infinite geometric progression. The ring V_p is called the ring of *p-adic integers*: these rings have considerable importance in number theory. It turns out that every element of V_p can be represented uniquely in the form $\sum_{t=0}^{\infty} a_t p^t$ where where every a_i is an integer satisfying $0 \leq a_i \leq p - 1$. We shall return to this example later when we study complete local rings.

In the problem dealing with Cohen's theorem, it is useful to consider colon ideals. We give the definition here. Let $I \subseteq R$ be an ideal in the ring R, and let S be an arbitrary subset of R. Then $I :_R S$ (or simply $I : S$) is, by definition, $\{r \in R : \text{for all } s \in S, rs \in I\}$, which is easily verified to be an ideal of R. If J is the ideal of R generated by the set S, it is straightforward to verify that $I : J = I : S$. We shall make use of colon ideals later when we study primary decomposition of ideals.

We next want to review the basic facts about tensor products of modules over a ring R. We shall use tensor products for several purposes. When S is an R-algebra and M is an R-module, then $S \otimes_R M$ is an S-module, and is finitely generated if M is. This gives us a method of passing from R-modules to S-modules that is called *extension of scalars.* When S is a localization of R, this gives a method of localizing modules as well, although there are alternative constructions of the localization of a module.

If S and T are both R-modules it turns out that $S \otimes_R T$ has the structure of an R-algebra, and is a coproduct for S and T in the category of R-algebras! Both extension of scalars and this method of constructing coproducts are of great importance, both in commutative algebra and in algebraic geometry.

We first recall the notion of a bilinear map. If M, N and W are R-modules, a *bilinear* map $B : M \times N \to W$ is a function such that for each fixed $v \in N$, the map $B_v : M \to W$ via $B_v(u) = B(u, v)$ is R-linear, and for each fixed $u \in M$, the map $B^u : N \to W$ via $B^u(v) = B(u, v)$ is R-linear. We can express all this at once by the requirement that for all $u_1, u_2 \in M$, for all $v_1, v_2 \in N$, and for all $r_1, r_2, s_1, s_2 \in R$, we have that

$$\begin{aligned}&B(r_1u_1 + r_2u_2, s_1v_1 + s_2v_2) =\\ &r_1s_1B(u_1, v_1) + r_1s_2B(u_1, v_2) + r_2s_1B(u_2, v_1) + r_2s_2B(u_2, v_2).\end{aligned}$$

One of the simplest and most important examples of a bilinear map is the map from $R \times R \to R$ that sends (r, r') to rr': bilinearity is a consequence of the left and right distributive laws in R (which, of course, imply each other in a commutative ring R).

The composition $T \circ B$ of a bilinear map $M \times N \to W$ and an R-linear map $T : W \to W'$ is easily verified to be a bilinear map $M \times N \to W'$. For fixed R-modules M, N this enables us to define a covariant functor from R-modules to sets whose value $\mathrm{Bil}(M, N; W)$ on the R-module W is the set of bilinear maps from $M \times N$ to W. Given $T : W \to W'$ the map $\mathrm{Bil}(M, N; W) \to \mathrm{Bil}(M, N; W')$ is induced by composition with T, as just described.

It will turn out that the tensor product $M \otimes_R N$ is an R-module that represents this functor, so that for all W we get a bijection $\mathrm{Hom}_R(M \otimes_R N,\, W) \cong \mathrm{Bil}(M, N; W)$: these bijections give an isomorphism of functors of W. We have not yet shown the existence and uniqueness of the tensor product, but we want to state first the key property that it has in somewhat greater detail. We shall show that there is an R-module $M \otimes_R N$ together with a bilinear map $\beta : M \times N \to M \otimes_R N$ with the following property: for every R-module W and bilinear map $B : M \times N \to W$ there is a unique linear map $T : M \otimes_R N \to W$ such that $B = T \circ \beta$. The tensor product $M \otimes_R N$ together with the bilinear map $\beta : M \times N \to M \otimes_R N$ give a universal bilinear map from $M \times N$, in the sense that every other bilinear map from $M \times N$ arises from β uniquely, by composition of a linear map with β.

We shall show next that tensor products exist, and are unique up to isomorphism that is also unique if the universal bilinear map β is taken into account.

Lecture of October 27

Let M and N be any two R-modules. To construct β and $M \otimes_R N$, take the free module F on a basis $b_{m,n}$ indexed by the elements of $M \times N$. Let G be the submodule of F spanned by the elements of the following forms as u, u' vary in M, v, v' vary in N, and r varies in R:

(1) $b_{u+u',v} - b_{u,v} - b_{u',v}$
(2) $b_{ru,v} - rb_{u,v}$
(3) $b_{u,v+v'} - b_{u,v} - b_{u,v'}$
(4) $b_{u,rv} - rb_{u,v}$

We define $M \otimes_R N$ to be F/G, and the map β by the rule $\beta(u,v) = [b_{u,v}]$, where the brackets [ ] indicate images mod G. The four types of elements that we killed by placing them in G precisely guarantee that β is bilinear.

Proposition. *For any bilinear map $B : M \times N \to W$, there is a unique linear map $f : M \otimes_R N \to W$ such that $B = f \circ \beta$.*

If $\gamma : M \times N \to T$ is another bilinear map with the same universal property, there is are unique isomorphism $\phi : M \otimes_R N \to T$ such that $\gamma = \phi \circ \beta$ (and, of course, $\beta = \phi^{-1} \circ \gamma$).

Proof. In order that $B = f \circ \beta$, we must have $f([b_{u,v}]) = f\big(\beta(u,v)\big) = B(u,v)$, which shows that f is unique. To show that it exists, define f_0 on F by the rule $f_0(b_{u,v}) = B(u,v)$. Then f_0 kills G, simply because B is bilinear: this is a straightforward check. Thus, f_0 induces an R-linear map $f : F/G \to W$ with the required properties.

If $\gamma : M \times N \to T$ also has this property, then there is a unique linear map $\psi : T \to M \otimes_R N$ such that $\beta = \psi \circ \gamma$ and a unique linear map $\phi : M \otimes_R N \to T$ such that $\gamma = \phi \circ \beta$ because of the property just proved for $M \otimes_R N$. The composition $\psi \circ \phi : T \to T$ has the property that its composition with γ is $\phi \circ \psi \circ \gamma = \phi \circ \beta = \gamma$, and the identity map on T has the same property. By the uniqueness property asserted for T and γ, $\phi \circ \psi$ is the identity map on T. By an exactly similar argument, $\psi \circ \phi$ is the identity map on $M \otimes_R N$. □

The image of (u,v) in $M \otimes_R N$ is denoted $u \otimes v$. This symbol has the following properties:

(1) $(u + u') \otimes v = u \otimes v + u' \otimes v$.
(2) $u \otimes (v + v') = u \otimes v + u \otimes v'$.
(3) $(ru) \otimes v = r(u \otimes v) = u \otimes (rv)$,

These properties are implied by the bilinearity of the map β.

Since the $b_{u,v}$ span F over R, the elements $u \otimes v$ span $M \otimes_R N$ over R. However, not every element has this form. It is however true that if two R-linear maps on $M \otimes N$ to W agree on all elements of the form $u \otimes v$, then they are equal.

Proposition. *If $\{u_i : i \in I\}$ spans M and $\{v_j : j \in J\}$ spans N, then $\{u_i \otimes v_j : (i,j) \in I \times J\}$ spans $M \otimes N$. Hence, if M and N are finitely generated, so is $M \otimes N$.*

Proof. $M \otimes N$ is spanned by the elements $u \otimes v$ and $u = \sum_{s=1}^{h} r_s u_{i_s}$ while $v = \sum_{t=1}^{k} r'_t v_{j_t}$. But then $u \otimes v = (\sum_{s=1}^{h} r_s u_{i_s}) \otimes (\sum_{t=1}^{k} r'_t v_{j_t}) = \sum_{s,t} (r_s r'_t)(u_{i_s} \otimes v_{j_t})$. □

Maps from tensor products are almost always constructed by giving a bilinear map first. The proofs of the following results give examples:

Proposition. *Let M, M' and N, N' be modules over the ring R.*

(a) *There is a unique isomorphism $M \otimes N \cong N \otimes M$ under which $u \otimes v$ is mapped to $v \otimes u$ for all $u \in M$, $v \in M$.*

(b) *There is an isomorphism $M \cong R \otimes M$ that maps u to $1 \otimes u$; its inverse maps $r \otimes m$ to rm.*

(c) *If $f : M \to M'$ and $g : N \to N'$ are R-linear, there is a unique R-linear map, $f \otimes g : M \otimes M' \to N \otimes N'$ such that $(f \otimes g)(u \otimes u') = f(u) \otimes g(u')$.*

(d) *There is a unique isomorphism $(M \oplus M') \otimes N \cong (M \otimes N) \oplus (M' \otimes N)$ that sends $(u \oplus u') \otimes v \cong (u \otimes v) \oplus (u' \otimes v)$. This extends at once, by induction, to direct sums of finitely many modules, and there is a corresponding fact when the second module is a direct sum.*

(e) *If $M = \bigoplus_{i \in I} M_i$ and $N = \bigoplus_{j \in J} N_j$ are arbitrary direct sums, then*

$$M \otimes N = \bigoplus_{(i,j) \in I \times J} M_i \otimes N_j.$$

(f) *If F is free over R on the free basis b_i, $i \in I$ and F' is free on the free basis b'_j, $j \in J$, then $F \otimes F'$ is free on the the free basis $b_i \otimes b'_j$, $(i, j) \in I \times J$.*

Proof. (a) follows from the fact that there is a bilinear map $M \times N \to N \otimes M$ taking (u, v) to $v \otimes u$: the check of bilinearity is straightforward. The construction of the map $N \otimes M \to M \otimes N$ is the same. Since the maps interchange $u \otimes v$ and $v \otimes u$, it is clear that each composition is the relevant identity map, since that is true on a spanning set.

(b) The check that the specified map $M \to R \otimes M$ is linear is easy, and it is likewise easy to check that there is a bilinear map $R \times M \to M$ sending (r, m) to rm, and hence a linear map $R \otimes M \to M$ sending $r \otimes m$ to rm. One of the compositions is sending $r \otimes m$ first to rm and then to $1 \otimes rm = r(1) \otimes m$, and so is the identity. The other check is easier.

(c) A linear map as specified exists because the map $M \times N \to M' \otimes N'$ that sends $(u, v) \to f(u) \otimes g(v)$ is readily checked to be bilinear.

(d) There is a bilinear map $(M \oplus M') \times N \to (M \otimes N) \oplus (M' \otimes N)$ that sends $(u \oplus u', n)$ to $(u \otimes v) \oplus (u' \otimes v)$. By (c), the injections $\iota : M \hookrightarrow M \oplus M'$, $\iota' : M' \hookrightarrow M \oplus M'$ induce maps $\iota \otimes 1_N : M \otimes N \to (M \oplus M') \otimes N$ and $\iota' \otimes 1_N : M' \to (M \oplus M') \otimes N$. These together give a map $(\iota \otimes 1_N) \oplus (\iota' \otimes 1_N) : (M \otimes N) \oplus (M' \otimes N) \to (M \oplus M') \otimes N$. It is completely straightforward to check that the compositions are the relevant identity maps, working with elements of the form $(u \oplus u') \otimes v$ in one case, and with elements of the forms $(u \otimes v) \oplus 0$ and $0 \oplus (u' \otimes v)$ in the other case.

(e) We first consider the case where there is just one module N on the right. Consider any finite number of the summands on the left: call them $M_{i_1}, \ldots, M_{i_n}$, and let M' be the direct sum of all the others. Then by part (d), we have $M \otimes N \cong \bigoplus_{t=1}^{n} M_t \otimes N \oplus M' \otimes N$. It follows that all of the modules $M_i \otimes N$ inject into $M \otimes N$ (as direct summands) and that any one of them is disjoint from a finite sum of the others. Since the M_i span M, it follows that the $M_i \otimes N$ have sum $M \otimes N$, and thus we have the required direct sum decomposition of $M \otimes N$. Obviously, there is a similar result for direct sum decompositions of N. Thus,

$$M \otimes N \cong (\bigoplus M_i) \otimes N \cong \bigoplus_i (M_i \otimes N) \cong \bigoplus_i (M_i \otimes (\bigoplus_j N_j)) \cong \bigoplus_i (\bigoplus_j M_i \otimes N_j)),$$

and the result follows.

(f) F (respectively, G) has the form $\bigoplus_i Rb_i$ (respectively, $\bigoplus_j Rb'_j$), where each $Rb_i \cong R$ (respectively, $Rb'_j \cong R$), and the result now follows from (e) and (b) in the special case $M \cong R$. □

Given two categories $\mathcal{C}$ and $\mathcal{D}$ one may define a product carry $\mathcal{C} \times \mathcal{D}$ whose objects are pairs of objects (X, Y) where X is an object of $\mathcal{C}$ and Y is an object of $\mathcal{D}$. The morphisms of (X, Y) to (X', Y') are pairs of morphisms (f, g) where $f : X \to X'$, $g : Y \to Y'$. Composition is performed coordinate-wise. From part (c) we deduce that $\otimes_R$ is a covariant functor of two variables, i.e., if $\mathcal{C}$ is the category of R-modules, it is a functor $\mathcal{C} \times \mathcal{C} \to \mathcal{C}$. One frequently considers this functor when one of the modules is fixed: thus, there is a functor $\_ \otimes_R N$ from R modules to R-modules that maps M to $M \otimes_R N$: it takes the map $f : M \to M'$ to the map $f \otimes 1_N$, so that $u \otimes v$ maps to $f(u) \otimes v$.

Proposition. *If* $0 \to M' \xrightarrow{f} M \xrightarrow{g} M'' \to 0$ *is a short exact sequence of modules, then* $M' \otimes N \to M \otimes N \to M'' \otimes N \to 0$ *is exact, i.e.,* $g \otimes 1_N$ *is surjective, and the image of* $f \otimes 1_N$ *is the kernel of* $g \otimes 1_N$. *The latter fact implies, when* $M' \subseteq M$ *and* $M'' = M/M'$, *that* $(M/M') \otimes N \cong (M \otimes N)/Im\,(M' \otimes N)$.

The conclusion remains correct if one only has that $M' \xrightarrow{f} M \xrightarrow{g} M'' \to 0$ *is exact.*

Proof. Since $M'' \otimes N$ is spanned by elements $u'' \otimes v$ for $u'' \in M''$ and $v \in N$, to prove that $g \otimes 1_N$ is surjective it suffices to observe that each such element is the image of $u \otimes v$, where $u \in M$ is chosen so that $g(u) = u''$. We clearly have a surjection of $M \otimes N/\mathrm{Im}\,(M' \otimes N)$ onto $M'' \otimes N$ mapping $u \otimes v$ to $g(u) \otimes v$. To complete the proof, we show that this is an isomorphism by constructing an inverse map h. There is a bilinear map $M'' \times N \to M \otimes N/\mathrm{Im}\,(M' \otimes N)$ sends (u'', v) to the class of (u, v), where $g(u) = u''$. This must be checked to be independent of the choice of u. But if we choose a different element u_1 that maps to u'', it differs from u by an element in the image of M', from which it follows that $u \otimes v - u_1 \otimes v = (u - u_1) \otimes v$ is in the image of $M' \otimes N$. Once the map is known to be well-defined, it is straightforward to check that it is bilinear, and it is clear the the compositions of h and $g \otimes 1_N$ give the appropriate identity map in either order, since one need only check what happens on the spanning elements such as $[u \otimes v]$ and $g(u) \otimes v$, and each maps to the other. □

The result of the preceding Proposition is referred to as the *right exactness of tensor*.

Applying $\_ \otimes N$ does not preserve injectivity in general. For example, consider the injection $2\mathbb{Z} \subseteq \mathbb{Z}$ of $\mathbb{Z}$-modules and apply $\_ \otimes \mathbb{Z}/2\mathbb{Z}$. We have that $2\mathbb{Z} \cong \mathbb{Z}$, and so the first module becomes $2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z}$ but the induced map $\mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ is 0. It might be better to think of the left hand copy of $\mathbb{Z}/2\mathbb{Z}$ as $2\mathbb{Z}/4\mathbb{Z}$. Note that when we look at the element $2 \otimes [1]$ in $2\mathbb{Z} \otimes \mathbb{Z}/2\mathbb{Z}$, we may not move the 2 that occurs to the left of the tensor symbol across the tensor symbol, because the element 1 is "missing" from $2\mathbb{Z}$.

Similarly, if a is a nonzero element of the domain A, we have $aA \subseteq A$, but applying $\_ \otimes A/aA$ gives the zero map from $aA/a^2A \cong A/aA$ to A/aA.

Corollary. *If $M' \subseteq M$ and $N' \subseteq N$ are submodules of the R-modules M, N, then $(M/M') \otimes (N/N') \cong (M \otimes N)/\big(Im\,(M \otimes N') + Im\,(M' \otimes N)\big)$.*

Proof. By the preceding result, this is $(M/M') \otimes N$ mod the image of $(M/M') \otimes N'$, and the former may be identified with $(M \otimes N)/\mathrm{Im}\,(M' \otimes N)$. The image of $(M/M') \otimes N'$ in this module is the same as the image of $M \otimes N'$, and the result follows. □

Corollary. *$(R/I) \otimes M \cong M/IM$ while $(R/I) \otimes (R/J) \cong R/(I + J)$.*

Proof. The image of $I \otimes M$ in $R \otimes M \cong M$ under the map that sends $r \otimes m$ to rM is IM. This proves the first statement. The second statement is immediate from the preceding Corollary. □

Lecture of October 30

Notice that if K is a field, V has finite basis $v_1, \ldots, v_m$, and and W has finite basis $w_1, \ldots, w_m$, then $V \otimes_K W$ has finite basis $v_i \otimes W_j$, and so every vector in $V \otimes W$ can be written uniquely in the form $\sum_{i,j} a_{ij}(v_i \otimes w_j)$ where $1 \le i \le m$ and $1 \le j \le n$. This gives a vector space isomorphism of $V \otimes W$ with $m \times n$ matrices (a_{ij}) over K which is not canonical: it depends on the choices of basis for V and for W. But it is useful. For example, an element of the form $v \otimes w$ in $V \otimes W$ can be written as

$$(\sum_{i=1}^{m} b_i v_i) \otimes (\sum_{j=1}^{n} c_j w_j) = \sum_{i,j} (b_i c_j)(v_i \otimes w_j),$$

so that the corresponding matrix $(b_i c_j)$ factors as the product of the $m \times 1$ matrix with entries b_i and the $1 \times n$ row matrix $(c_1 \ \ldots c_n)$. A matrix has such a factorization if and only if it has rank at most one. Thus, an element of $V \otimes W$ is decomposable as $v \otimes w$ if and only if the corresponding matrix has rank at most one. This condition is independent of the choices of basis. Such matrices are rather special among all $m \times n$ matrices (unless $m \le 1$ or $n \le 1$).

If $M_1, \ldots, M_k, W$ are R-modules a map $M_1 \times \cdots \times M_k \to W$ is called *k-multilinear* or simply *multilinear* over R if, for every i, it becomes an R-linear function of u_i when all the other entries $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots u_k$ of $u_1, \ldots, u_k$ are held fixed. An example is the map $R \times R \times \cdots \times R \to R$ that sends $(r_1, \ldots, r_k) \to r_1 r_2 \cdots r_k$. As a function of k variables it is a polynomial of degree k, but it is linear in each variable if all of the others are held fixed. If $k = 2$, this means that the map is bilinear. When $k = 3$ we may use the term *trilinear*.

We next note that the map τ on $M_1 \times M_2 \times M_3$ sending (u_1, u_2, u_3) to $(u_1 \otimes u_2) \otimes u_3$ in $(M_1 \otimes M_2) \otimes M_3$ is trilinear, and in fact is universal, in the sense that any trilinear map $T : M_1 \times M_2 \times M_3 \to W$ factors uniquely as the composition of τ with a linear map $f : (M_1 \otimes M_2) \otimes M_3 \to W$. Uniqueness is clear, since given T, the value of f on $(u_1 \otimes u_2) \otimes u_3$ must be $T(u_1, u_2, u_3)$. To show that such a map f exists, note that for each fixed u_3, T defines a bilinear map $B_{u_3} : M_1 \times M_2 \to W$ such that $B_{u_3}(u_1, u_2) = T(u_1, u_2, u_3)$, and therefore a linear map $g_{u_3} : (M_1 \times M_2) \to W$. We can then define a bilinear map $B : (M_1 \otimes M_2) \times M_3 \to W$ by the rule by the rule $B(v, u_3) = g_{u_3}(v)$. It is straightforward to check bilinearity, and that the map $f : (M_1 \otimes M_2) \otimes M_3 \to W$ induced by B satisfies $f\big((u_1 \otimes u_2) \otimes u_3\big) = T(u_1, u_2, u_3)$ for all $u_i \in M_i$.

An entirely similar argument shows that we could have used $M_1 \otimes (M_2 \otimes M_3)$ instead, i.e., that the map $M_1 \times M_2 \times M_3 \to M_1 \otimes (M_2 \otimes M_3)$ that sends (u_1, u_2, u_3) to $u_1 \otimes (u_2 \otimes u_3)$ is also a universal trilinear map. This gives a map $(M_1 \otimes M_2) \otimes M_3 \to M_1 \otimes (M_2 \otimes M_3)$ and also a map in the other direction such that the first takes every $(u_1 \otimes u_2) \otimes u_3$ to $u_1 \otimes (u_2 \otimes u_3)$ while the second takes every $u_1 \otimes (u_2 \otimes u_3)$ to $(u_1 \otimes u_2) \otimes u_3$. These are evidently mutually inverse maps, and the isomorphism $(M_1 \otimes M_2) \otimes M_3 \cong M_1 \otimes (M_2 \otimes M_3)$

just described is referred to as the *associativity of tensor*, although we shall also soon see that there is a stronger version.

It then follows that $M_1 \otimes M_2 \otimes \cdots \otimes M_k$ has a meaning independent of how one inserts parentheses, and that the map $\mu : M_1 \times \cdots \times M_k \to M_1 \otimes M_2 \otimes \cdots \otimes M_k$ that sends $(u_1, \cdots, u_k)$ to $u_1 \otimes \cdots \otimes u_k$ is a universal k-multilinear map: every k-multilinear map from $M_1 \times \cdots \times M_k$ arises from μ by composing it with a linear map, and the linear map that can be used is unique. The only step of interest in the proof is to show that given a k-multilinear map $T : M_1 \otimes M_2 \otimes \cdots \otimes M_k \to W$ there is a map $f : M_1 \otimes M_2 \otimes \cdots \otimes M_k \to W$ such that $f(u_1 \otimes \cdots \otimes u_k) = T(u_1, \ldots, u_k)$ for all $(u_1, \ldots, u_k) \in M_1 \times \cdots M_k$. One uses induction. Again, for each fixed $u_k \in M_k$, T yields a $(k-1)$-multilinear map g_{u_k} of the first $k-1$ variables to W, which in turn, by the induction hypothesis, induces a map $g_{u_k} : M_1 \otimes \cdots M_{k-1} \to W$. The map $(M_1 \otimes \cdots M_{k-1}) \times M_k$ whose value on $(u_1 \otimes \cdots u_{k-1}, u_k)$ is $g_{u_k}(u_1 \otimes \cdots u_{k-1})$ may be easily checked to be bilinear, and this induces the map we want.

We next want to show that if M_1 is an S-module and M_2 is an R-module, then $M_1 \otimes_R M_2$ (where M_1 is regarded as an R-module via restriction of scalars) is an S-module in such a way that for all $s \in S$, $u_1 \in M_1$ and $u_2 \in M_2$, $s(u_1 \otimes u_2) = (su_1) \otimes u_2$. First note that we have an R-trilinear map $S \times M_1 \times M_2 \to M_1 \otimes_R M_2$ that sends (s, u_1, u_2) to $(su_1) \otimes u_2$. This yields an R-linear map $S \otimes_R (M_1 \otimes_R M_2) \to M_1 \otimes_R M_2$ and therefore an R-bilinear map $S \times (M_1 \otimes_R M_2) \to M_1 \otimes_R M_2$. This is the map we shall use for multiplication by s: for $z \in M_1 \otimes M_2$, sz is the image of (s, z) under this map. This has the stated property that $s(u_1 \otimes u_2) = (su_1) \otimes u_2$. The R-bilinearity also implies most of the conditions that we need for this action of S to make $M_1 \otimes_R M_2$ into an S-module. However, we still need to check that for all $u \in M_1 \otimes_R M_2$ and $s, s' \in S$, $(ss')u = s(s'u)$. Since multiplication by an element of S is R-linear, it suffices to check this for elements u that generate $M_1 \otimes_R M_2$ as an R-module, and so we may assume that $u = u_1 \otimes u_2$. But then $(ss')(u_1 \otimes u_2) = \big((ss')u_1\big) \otimes u_2 = \big(s(s'u)_1\big) \otimes u_2 == s\big((s'u)_1 \otimes u_2\big) = ss'(u_1 \otimes u_2)$, as required.

Similarly, if M_1 is an R-module and M_2 is an S-module, we can give an S-module structure to $M_1 \otimes_R M_2$ such that $s(m_1 \times m_2) = m_1 \otimes (sm_2)$.

A word of caution: if M_1 and M_2 are both S-modules, then $M_1 \otimes_R M_2$ has *two* S-module structures, one that comes from M_1, and one that comes from M_2, and *these are almost always different*. The point is that when we tensor over R, scalars from S cannot be "passed through" the tensor symbol (although scalars from R can), and so $sm_1 \otimes m_2$ and $m_1 \otimes sm_2$ are usually distinct.

Something analogous happens with $\mathrm{Hom}_R(M_1, M_2)$ when S is an R-algebra. If M_2 is an S-module, then $\mathrm{Hom}_R(M_1, M_2)$ becomes an S-module if one defines sf by the rule $(sf)(u_1) = s\big(f(u_1)\big)$. Likewise, if M_1 is an S-module then $\mathrm{Hom}_R(M_1, M_2)$ becomes an S-module if one uses the rule $(sf)(u_1) = f(su_1)$. However, when M_1 and M_2 are both S-modules, these two S-module structures are usually different: we do not have $f(su_1) = sf(u_1)$ because f is being assumed R-linear but not necessarily S-linear.

If M, N are S-modules and Q is an R-module then $(M \otimes_S N) \otimes_R Q \cong M \otimes_S (N \otimes_R Q)$ as S-modules: this is also called *associativity of tensor*: it strengthens our previous result, which was the case where $S = R$. As before, $(u \otimes v) \otimes w$ and $u \otimes (v \otimes w)$ correspond under the two isomorphisms. We construct maps as follows. For each fixed $w \in Q$ there is an S-bilinear map $B_w : M \times N \to M \otimes_S (N \otimes_R Q)$ by the rule $B_w(u, v) = u \otimes (v \otimes w)$. This gives an S-linear map $g_v : M \otimes N \to W$. We can then define an R-bilinear map $(M \otimes_S N) \times_S Q \to M \otimes_S (N \otimes_R Q)$ that sends (y, v) to $g_v(y)$, and this induces an R-linear map $(M \otimes_S N) \otimes_S Q \to M \otimes_S (N \otimes_R Q)$ that is easily checked to send $(u \otimes v) \otimes w$ to $u \otimes (v \otimes w)$. This map turns out to be S-linear: we need only check this on the generators $(u \otimes v) \otimes w$, and $s\big((u \otimes v) \otimes w)\big) = (su \otimes v) \otimes w$ has image $(su) \otimes (v \otimes w) = s\big(u \otimes (v \otimes w)\big)$, as required.

To get a map in the other direction, for each fixed $u \in M$ define an R-bilinear map $B'_u : N \times Q \to (M \otimes_S N) \otimes_R Q$ by the rule $B'_u(v, w) = (u \otimes v) \otimes w$, which yields an R-linear map $g'_u : N \otimes_R M \to (M \otimes_S N) \otimes_R Q$. It is easy to check that this map is actually S-linear, because $(u \otimes sv) \otimes w = s\big(u \otimes (v \otimes w)\big)$. We then define an S-bilinear map $M \times (N \otimes_R Q) \to (M \otimes_S N) \otimes_R Q$ that sends (u, z) to $g'_u(z)$. We now have S-linear maps in both directions that on generators interchange $(u \otimes v) \otimes w$ and $u \otimes (v \otimes w)$, as required.

By these remarks, if S is an R-algebra we have a covariant right exact functor from R-modules to S-modules given by $S \otimes_R \_$. This operation is referred to as *extension of scalars* or *base change*, because the "base ring" R is being replaced by the base ring S. R-modules get converted to S-modules. This turns out to be an extraordinarily useful technique. The method of studying real vector spaces and real matrices by enlarging the field to the complex numbers and taking complex linear combinations and so forth is actually an example of this method being used in a tacit way.

Note that since $S \otimes_R M$ is an S-module, it is also an R-module by restriction of scalars. The map $M \to S \otimes M$ that takes u to $1 \otimes u$ is R-linear. In general, it need not be injective, however.

Because of the canonical isomorphism $S \otimes_R R \cong S$ as S-modules, and the fact that tensor product commutes with direct sum, base change converts free R modules with free basis $\{b_i\}_{i \in I}$ to free S-modules with free basis $\{1 \otimes b_i\}_{i \in I}$. If $f : R \to S$ is the structural homomorphism for S as an R-algebra, the map $R \to R$ given by multiplication by $r \in R$ becomes the map $R \to S$ given by multiplication by $f(r) \in S$.

To understand what base change does to an arbitrary module it may be helpful to think in terms of presentations. Given an R-module M one may choose generators $\{u_i\}_{i \in I}$ where I is a suitable index set (I may be infinite), and then form a free module $R^{\oplus I}$ with free basis $\{b_i\}_{i \in I}$ indexed by I. We then have a surjection $R^{\oplus I} \twoheadrightarrow M$ that sends b_i to u_i for every i. We may then choose generators for the kernel $M' \subseteq R^{\oplus I}$, and so construct another surjection $R^{\oplus J} \twoheadrightarrow M'$. This then yields an exact sequence $R^{\oplus J} \to R^{\oplus I} \twoheadrightarrow M \to 0$, and the map $R^{\oplus J} \to R^{\oplus I}$ determines M. We may think of the map as given by a matrix indexed by $I \times J$ such that each column has only finitely many nonzero entries. Thus, the columns represent vectors in $R^{\oplus I}$ that span M', and we get M by killing the R-span of these columns. In other words, M is simply the cokernel of the map $R^{\oplus J} \to R^{\oplus I}$. The sequence $R^{\oplus J} \to R^{\oplus I} \twoheadrightarrow M \to 0$ is called a *presentation* of M.

This is all more standard when M is finitely generated and the module M' is also finitely generated. Then I and J are finite sets, and M is said to be *finitely presented.* If R is Noetherian, every finitely generated module has a finite presentation: if I is finite, $R^{\oplus I}$ is finitely generated, and, hence Noetherian. This implies that M' is finitely generated.

In this case, where I and J are finite, we can think of the presentation sequence as having the form $R^n \to R^m \twoheadrightarrow M \to 0$, where we think of R^m as $m \times 1$ column vectors. The matrix of the map is then an $m \times n$ matrix (r_{ij}) over R. M is the cokernel of the map, which is the same as the quotient of R^m by the submodule spanned by the n columns. When we apply $S \otimes_R \_$, we get an exact sequence $S^n \to S^m \twoheadrightarrow S \otimes_R M \to 0$, and so we get a presentation of the module $S \otimes_R M$. The new matrix $\bigl(f(r_{ij})\bigr)$ is obtained by applying f to the entries of the original matrix. One corollary of this point of view is that if M is finitely generated by, say, $u_1, \ldots, u_m$, then $S \otimes_R M$ is finitely generated by the elements $1 \otimes u_1, \ldots, 1 \otimes u_m$. The same remark applies to arbitrary sets of generators of M.

We can think essentially the same way even when I and J are infinite: we still have a matrix in the form of an R-valued function on $I \times J$ (subject to the additional condition that for every $j \in J$ it is nonzero for only finitely many $i \in I$), and the new matrix for the map $S^{\oplus J} \to S^{\oplus I}$ in the presentation sequence $S^{\oplus J} \to S^{\oplus I} \twoheadrightarrow S \otimes_R M \to 0$ is obtained by applying f to each entry of the original matrix.

We next note that the module $S \otimes_R M$ has a certain universal mapping property, and can be thought of as representing a functor:

Theorem. *If S is an R-algebra, M is an R-module, and N is an S-module, there is a canonical isomorphism $\theta_N : \mathrm{Hom}_R(M, N) \cong \mathrm{Hom}_S(S \otimes_R M, N)$. This isomorphism takes $f : M \to N$ to the map $S \otimes M \to N$ induced by the R-bilinear map $S \times M \to N$ that sends (s, m) to $sf(m)$ for all $s \in S$ and $m \in M$. The inverse σ_N of θ_N is obtained by composing $g : S \otimes_R M \to N$ with the map $M \to S \otimes_R M$ described earlier (which maps $m \in M$ to $1 \otimes m$). The isomorphisms θ_N together give an isomorphism between the covariant functors $\mathrm{Hom}_R(M, \_)$ and $\mathrm{Hom}_S(S \otimes_R M, \_)$ viewed as functors to sets (or as functors to S-modules). Thus, $S \otimes_R M$ represents $\mathrm{Hom}_R(M, \_)$ in the category of S-modules.*

Proof. There are only a few things to check. One is that the maps θ_N and σ_N are inverses. Given $f : M \to N$, $\theta_N(f)$ maps $s \otimes u$ to $sf(u)$, and so composing with $M \to S \otimes M$ gives a map that takes u to $1 \cdot f(u) = f(u)$, as required. On the other hand, given a map $g : S \otimes M \to N$ that is S-linear, it suffices to see that when we apply $\theta_N \circ \sigma_N$ we get a map that agrees with g on elements of the form $s \otimes g$. The effect of applying σ_N is to give a map $M \to N$ that sends u to $g(1 \otimes u)$, and then the further effect of applying θ_N gives a map that sends $s \otimes u$ to $sg(1 \otimes u) = g(s \otimes u)$, as required. All other checks needed are at least as easy. $\square$

We shall next use base change to develop a theory of localization of modules, although it will prove useful to have an alternative characterization of localization of a module.

Lecture of November 1

We give one construction for localization of R-modules with respect to a multiplicative system, and then show that it is really an instance of base change.

Suppose that $S = W^{-1}R$, where W is a multiplicative system in R. Given an R-module M we can construct an S-module $W^{-1}M$ as follows. Define an equivalence relation on $M \times W$ via $(u, w) \sim (u', w')$ iff there exists $w'' \in W$ such that $w''(w'u - wu') = 0$. The equivalence classes form an abelian group with the addition $[(u, w)] + [(u', w')] = [(w'u + wu', ww')]$, and an S-module via the multiplication $(r/w)[(u, w')] = [(ru, ww')]$. These operations are easily checked to be independent of the choices of equivalence class representatives. The class $[(u,\, w)]$ is often denoted u/w. $W^{-1}M$ also remains an R-module, by restriction of scalars. In fact, $r(u/w) = (ru)/w$. There is a map $M \to W^{-1}M$ that is R-linear, sending u to $[(u, 1)]$. The kernel of this map is $\{u \in M : \text{for some } w \in W, wu = 0\}$. This is easily checked because, by the definition of the equivalence relation, $[(u, 1)] \sim [(0, 1)]$ if and only if for some $w \in W$, $w(1 \cdot u - 1 \cdot 0) = 0$, i.e., $wu = 0$.

This localization operation can also be defined on maps: given an R-linear map $M \to N$ there is a unique S-linear map $W^{-1}f : W^{-1}M \to W^{-1}N$ such that the diagram

$$\begin{tikzcd}
W^{-1}M \arrow[r, "W^{-1}f"] & W^{-1}N \\
M \arrow[u] \arrow[r, "f"] & N \arrow[u]
\end{tikzcd}$$

commutes. This map is defined to take $[(u, w)]$ to $[f(u)/w]$, i.e., to take u/w to $f(u)/w$. It is easily checked that the equivalence class of the value is independent of the choice of representative of the equivalence class $[(u, w)]$, and that the map is S-linear. It is clear that it does make the diagram commute. Uniqueness follows, because the value of the map on $w(u/w) = u/1$ must be $f(u)$ for the diagram to commute, and if we multiply by $1/w$ we see that the map must take u/w to $f(u)/w$.

The map $M \to W^{-1}M$ induces a unique map $S \otimes M \to W^{-1}R$ such that $(r/w) \otimes u$ maps to ru/w for all $r \in R$, $w \in W$, and $u \in M$. This map is clearly surjective, and is an isomorphism: to give a map in the other direction, simply send $[(u, w)]$ to $(1/w) \otimes u$, which is easily checked to be independent of the choice of representatives and to be a ring homomorphism. All of the elements $s \otimes u$ for $s = r/w \in S$ can be rewritten as $(1/w) \otimes ru$. Therefore to see that the two maps are mutually inverse it suffices to note that they interchange $(1/w) \otimes u$ and u/w for all $w \in W$ and $u \in M$.

The identification reconfirms that localization at W gives a covariant functor from R-modules to S-modules: we have already described it directly.

A module M over R is called *flat* if $M \otimes_R \_$ is an exact functor, i.e., if whenever $N \subseteq Q$ are R-modules the map $M \otimes N \to M \otimes Q$ is injective. Then all exact sequences are preserved. An R-algebra S is called *flat* if it is flat as an R-module. If S is flat over R, base change from R-modules to S-modules is an exact covariant functor.

Theorem. *If R is a ring and W a multiplicative system, $W^{-1}R$ is R-flat.*

Proof. This comes down to the assertion that if $N \subseteq M$ then the induced map $W^{-1}N \to W^{-1}M$ is injective. But n/w maps to 0 if and only if n/w thought of in $W^{-1}M$ is 0, and the definition of the equivalence relation tells us that n/w is 0 if and only if $w'n = 0$ for some $w' \in W$. But this condition implies that n/w is already 0 in $W^{-1}N$. □

It is a straightforward exercise to show that a direct sum of modules is flat if and only if all the summands are flat. R itself is obviously flat, since $R \otimes_R \_$ is isomorphic to the identity functor, and it follows that free modules are flat. Thus, when R is a field, every R-module is flat.

Direct summands of free modules are called *projective* modules. They need not be free, but they are flat.

We recall some facts about splitting. Suppose that we have an R-linear surjection $f : M \twoheadrightarrow P$ with kernel Q. If there is an R-linear map $g : P \to M$ such that $f \circ g$ is the identity map on P, then it is clear that g is injective, with image $P' \subseteq M$ that is isomorphic to P. Moreover, M is the internal direct sum of P and Q. Give $u \in M$, $p' = g(f(u)) \in P'$, and $f(u - p') = f(u) - (f \circ g)\bigl(f(u)\bigr) = f(u) - f(u) = 0$, so that $u - p' \in Q$. Thus, $M = P' + Q$. If $u \in P' \cap Q$ then $f(u) = 0$, since $u \in Q$. But $u = g(p)$ for some $p \in P$, and so $0 = f(u) = (fg)(p) = p$, and then $u = g(p) = g(0) = 0$. Thus, $M = P' \oplus_R Q$ internally, and $M \cong P \oplus_R Q$. With this in mind, we give some examples of projective modules that are not free.

Examples. If R has a non-trivial idempotent e, then R is the direct sum of eR and $(1-e)R$ as R-modules. These are projective R-modules that are not free (each is the annihilator of the other, while a nonzero free module has annihilator (0)).

Here is a much more intriguing example. Let $T = \mathbb{R}[X, Y, Z]/(X^2 + Y^2 + Z^2 - 1) = \mathbb{R}[x, y, z]$. The elements of this ring are represented by polynomials and these may be restricted, as $\mathbb{R}$-valued functions, to the unit 2-sphere centered at the origin in $\mathbb{R}^3$, and so give continuous functions on the 2-sphere S^2. Different representatives of the same class give the same function, since they differ by a multiple of $X^2 + Y^2 + Z^2 - 1$, which vanishes on S^2. Consider the T-linear map $f : T^3 \to T$ with matrix $(x \;\; y \;\; z)$. We have a map $g : T \to T^3$ given by the 3×1 column matrix u with entries x, y, z. The composition $f \circ g$ is given by the 1×1 matrix whose single entry is $x^2 + y^2 + z^2 = 1$, and so $f \circ g$ is the identity on T. The image of g is the free module Tu with the generator u. By the discussion just preceding these examples, $T^3 = Tu \oplus_T Q$ where Q is the kernel of the map. Thus, Q is a projective module.

If we make a base change by tensoring over T with the fraction field $\mathcal{K}$ of T, we see that $\mathcal{K}^3 \cong \mathcal{K} \oplus_{\mathcal{K}} (\mathcal{K} \otimes_T Q)$. It follows that if Q is free over T, it must be free on two generators. But Q *is not* $\cong T^2$. To see this, suppose Q has a free basis consisting of column vectors v and w. Then u, v and w give the columns of a 3×3 matrix A. Since u, v, and w span T^3, there cannot be any linear relation on them, for then they will span $\mathcal{K}^3$ over $\mathcal{K}$, and so they are a vector space basis for $\mathcal{K}^3$ over $\mathcal{K}$ and have no linear relation even over $\mathcal{K}$. Thus, they give a new free basis for T^3. It follows that the map given by the matrix A is an automorphism of T^3, and its inverse will be given by a matrix B such that $AB = BA = I$, the 3×3 identity matrix over T. We then have that $\det(A)\det(B) = 1$, and so $\det(A)$ is a unit α of T, and we can multiply the second (or third) column of A by α^{-1}. We therefore see that u is part of free basis for T^3 if and only if it is the first column of a 3×3 matrix over T with determinant one (we only proved one direction, which is the direction that we are using, but the other direction is quite straightforward). However, it is impossible to give such a matrix even if the entries are allowed to be arbitrary continuous functions on the 2-sphere!

Suppose that the second column of the matrix A is $(f,\, g,\, h)^{\mathrm{tr}}$, where we are using the superscript $^{\mathrm{tr}}$ to indicate the transpose of a matrix. We may think of both the first and second columns as continuous vector-valued functions on S^2: the value of the first column is a unit vector that is the position vector of the point of S^2 that we are considering. At each point (a, b, c) of S^2 the vectors which are the values of the first and second columns are linearly independent, because the determinant of the matrix A, when evaluated at (a, b, c), is constantly 1. We can subtract off the component of the second column in the direction of the unit vector (a, b, c): we obtain a continuous non-vanishing vector-valued function that gives a non-vanishing vector field of tangent vectors to the 2-sphere, a contradiction. More explicitly, consider $V = (f,\, g,\, h)^{\mathrm{tr}} - (xf + yg + zh)(x,\, y,\, z)^{\mathrm{tr}}$. This is a continuous vector-valued function on S^2. It does not vanish on S^2, because at each point the values of $(f,\, g,\, h)^{\mathrm{tr}}$ and $(x,\, y,\, z)^{\mathrm{tr}}$ are linearly independent. At every point $(a, b, c) \in S^2$, the value of V is orthogonal to the unit vector (a, b, c): the dot product vanishes. Thus, as already asserted, we have constructed a non-vanishing vector-valued function whose value at every point of S^2 is a tangent vector of S^2. This completes the proof that Q is not free! $\square$

We next note the following exactness properties of Hom:

Proposition. *Let $0 \to M' \xrightarrow{\alpha} M \xrightarrow{\beta} M''$ be an exact sequence of R-modules, and let N be any R-module. Then $0 \to Hom_R(N,\, M') \to Hom_R(N, M) \to Hom_R(N,\, M'')$ is exact, and so $Hom(N,\, \_\,)$ is a left exact covariant functor from R-modules to R-modules.*

What is more, if $M' \xrightarrow{\alpha} M \xrightarrow{\beta} M'' \to 0$ is an exact sequence of R-modules, then $0 \to Hom_R(M'',\, N) \to Hom_R(M,\, N) \to Hom(M',\, N)$ is exact. Thus, $Hom_R(\_\,, N)$ is a contravariant left exact functor from R-modules to R-modules.

Proof. For the statement in the first paragraph, note that α obviously induces an injection, and a map from N to M is killed if and only if all its values are, which means that it is taking values in the image of M'.

For the statement in the second paragraph, note that the map induced by β is obviously injective: if $f(u'') \neq 0$, there exists $u \in M$ such that $\beta(u) = u''$, and then $(f \circ \beta)(u) \neq 0$. A map from M to N is killed if and only if its restriction to M' is the zero map, i.e., if and only if it induces a map from M'' to N, and it will be the image of this map. $\square$

Note that $0 \to 2\mathbb{Z} \subseteq \mathbb{Z} \twoheadrightarrow \mathbb{Z}/2\mathbb{Z} \to 0$ is exact, but that if we apply $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z},\, \_\,)$ the map $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z},\, \mathbb{Z}) = 0$ to $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z},\, \mathbb{Z}/2\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}$ is not onto, while if we apply $\mathrm{Hom}_{\mathbb{Z}}(\_\,,\, \mathbb{Z})$ the map $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z},\, \mathbb{Z}) \to \mathrm{Hom}_{\mathbb{Z}}(2\mathbb{Z},\, \mathbb{Z})$ is not onto: the map $2\mathbb{Z} \to \mathbb{Z}$ that sends $2 \mapsto 1$ does not extend to $\mathbb{Z}$.

Lecture of November 3

It is worth noting that $\mathrm{Hom}_R(R,\, M) \cong M$ for all R-modules M: to give an R-linear map $R \to M$ is the same as to specify its value on the generator 1 of R as an R-module, and so we have a map $\mathrm{Hom}_R(R,\, M) \to M$ sending $f \mapsto f(1)$ and a map $M \to \mathrm{Hom}_R(R,\, M)$ sending $u \in M$ to the map whose value on r is ru for all $r \in R$. These mutually inverse isomorphisms establish an isomorphism of $\mathrm{Hom}_R(R,\, \_)$ with the identity functor on R-modules. Also note that we have a similar isomorphism $\mathrm{Hom}_R(R/I,\, M) \cong \mathrm{Ann}_M I$, the set of all elements of M that are killed by I. We use a bar to indicate classes mod I. Giving a map from R/I is the same as specifying its value on $\overline{1}$, but the fact that $r(\overline{1}) = \overline{r} = 0$ for $r \in I$ implies that $rf(\overline{1}) = f(\overline{r}) = f(0) = 0$ implies that only elements of $\mathrm{Ann}_M I$ are available as values for $f(\overline{1})$.

We note the following alternative characterization of projective modules: an R-module P is projective if and only if $(*)$ $\mathrm{Hom}_R(P,\, \_)$ is an exact functor, which means that for every surjective map $M \twoheadrightarrow M''$, $\mathrm{Hom}_R(P,\, M) \to \mathrm{Hom}_R(P,\, M'')$ is surjective. If $P = R$ one gets the identity functor, so that R has property $(*)$, and a direct sum of modules has property $(*)$ if and only if they all do, so that free modules have it and direct summands of free modules also have it. But if P has property $(*)$ we can map a free module F onto it, say $f : F \twoheadrightarrow P$, and then $\mathrm{Hom}_R(P,\, F) \to \mathrm{Hom}_R(P,\, P)$ is onto, and so there is a map $g : P \to F$ whose composition with surjection $f : F \twoheadrightarrow P$ is the identity on P. This implies that P is isomorphic with a direct summand of F: F is the internal direct sum of the kernel Q of f and $g(P)$, which is isomorphic with P.

We next want to construct coproducts in the category of A-algebras for any arbitrary commutative ring A: the category of all commutative rings is included, since this is identical with the category of $\mathbb{Z}$-algebras.

If R and S are A-algebras then there is an A-bilinear map $\mu : (R \otimes_A S) \times (R \otimes_A S) \to R \otimes_A S$ such that $\big(r \otimes s, r' \otimes s'\big) \mapsto (rr') \otimes (ss')$. To give this map is the same as giving a linear map $(R \otimes_A S) \otimes_A (R \otimes_A S) \to R \otimes_A S$. But this in turn is the same as giving a 4-linear map of A-modules $R \times S \times R \times S \to R \otimes_A S$ and we can simply send (r, s, r', s') to $(rr') \otimes (ss')$. This gives a multiplication on $R \otimes_A S$ that makes it into an A-algebra. The distributive law follows from the bilinearity of μ. There are a number of things to check, for example, that the multiplication one gets is commutative and associative and that $1 \otimes 1$ is an identity. It suffices to check this on the A-generators, e.g., for associativity that $\big((r \otimes s)(r' \otimes s')\big)(r'' \otimes s'') = (r \otimes s)\big((r' \otimes s')(r'' \otimes s'')\big)$. This comes down to $\big((rr')r''\big) \otimes \big((ss')s''\big) = \big(r(r'r'')\big) \otimes \big(s(s's'')\big)$, which is immediate from the associativity of the respective multiplications in R and S. The other checks are equally easy. Notice that $\iota_1 : R \to R \otimes_A S$ sending $r \to r \otimes 1$ is an A-algebra homomorphism, and $\iota_2 S \to R \otimes_A S$ sending s to $1 \otimes s$ is as well.

Theorem. *$R \otimes_A S$ together with the maps $\iota_1 : R \to R \otimes_A S$ and $\iota_2 : S \to R \otimes_A S$ is a coproduct for R and S in the category of A-algebras: for every A-algebra T there is a bijection $Hom_{A\text{-alg}}(R \otimes_A S,\, T) \cong Hom_{A\text{-alg}}(R,\, T) \times Hom_{A\text{-alg}}(S,\, T)$ that sends f to $(f \circ \iota_1,\, f \circ \iota_2)$.*

Proof. Because $R \otimes_A S$ is generated by the images of R and S over A (note that $r \otimes s = (r \otimes 1)(1 \otimes s)$), it is obvious that any A-algebra homomorphism $f : R \otimes_A S \to T$ is determined by $f \circ \iota_1$ and $f \circ \iota_2$. Therefore, the specified map is one-to-one. To show that it is onto we need to show that given $g : R \to T$ and $h : S \to T$ as A-algebras we can construct an A-algebra homomorphism $f : R \otimes_A S \to T$ such that $f(r \otimes 1) = g(r)$ and $f(1 \otimes s) = h(s)$ for all $r \in R$ and $s \in S$, and for this it suffices to construct f such that $f(r \otimes s) = f(r)g(s)$ for all $r \in R$ and $s \in S$. That there is such a map of $R \otimes_A S \to T$ simply as a map of A-modules follows from the A-bilinearity of the map $R \times S \to T$ that sends (r, s) to $g(r)h(s)$. To check that the induced A-linear map from $R \otimes_A S \to T$ is a ring homomorphism, we only need to check that it preserves multiplication. Since the elements $r \otimes s$ span over A, by virtue of the distributive law it suffices to check that $f\big((r \otimes s)(r' \otimes s')\big) = f(r \otimes s)g(r' \otimes s')$. The left side is $f\big((rr') \otimes (ss')\big) = g(rr')h(ss') = g(r)g(r')h(s)h(s') = \big(g(r)h(s)\big)\big(g(r')h(s')\big) = f(r \otimes s)f(r' \otimes s')$, as required. □

Consider the coproduct of two polynomial rings $R = A[x_i : i \in I]$ and $B = A[y_j : j \in J]$ over A, where I and J are (possibly infinite) index sets. The monomials in the x_i are a free basis for R over A, and the monomials in the y_j are a free basis for S over A. Thus, the set of elements $\mathcal{M} \otimes \mathcal{M}'$ where $\mathcal{M}$ is a monomial in the x_i and $\mathcal{M}'$ is a monomial in the y_j is a free basis for $R \otimes_A S$ over A, and monomials are multiplied by the rule $(\mathcal{M}_1 \otimes \mathcal{M}_1')(\mathcal{M}_2 \otimes \mathcal{M}_2') = (\mathcal{M}_1\mathcal{M}_2 \otimes \mathcal{M}_1'\mathcal{M}_2')$. It follows that the ring $R \otimes_A S$ is a polynomial in ring in variables indexed by the disjoint union of I and J.

This may seem odd at first when the sets of variables overlap or are even equal: however, because the variables cannot pass through the tensor symbol, in the tensor product they have become disjoint sets of variables.

Thus, $A[x] \otimes_A A[x]$ is a polynomial ring in two variables over A: the elements $x \otimes 1$ and $1 \otimes x$ have no algebraic relation over A in the tensor product, because x does not pass through the tensor symbol. This can be a bit confusing. Note, however, that if we tensor over $A[x]$, we have instead that $A[x] \otimes_{A[x]} A[x] \cong A[x]$.

One can give an alternative description of the coproduct of R and S over A. Map a polynomial ring $A[x_\sigma : \sigma \in \Sigma]$ onto R (one can even introduce one indeterminate for every element of R, and map that indeterminate to the specified element of R), and call the kernel I. Map a polynomial ring $A[y_\tau : \tau \in T]$ onto S, and call the kernel J. Assume for simplicity that the indeterminates are all mutually distinct, and that the union of the two sets of indeterminates is an algebraically independent set. If we form the polynomial ring in the union of the two sets of indeterminates, then a coproduct may be constructed as the quotient of the polynomial ring in all the variables mod the sum of the expansions of I and J.

The map $M \otimes_R N \twoheadrightarrow M \otimes_S N$, defined whenever S is an R-algebra and M and N are S-modules, is not, in general, an isomorphism. For example, for any choice of the ring A, $A[x] \otimes_A A[x] \twoheadrightarrow A[x] \otimes_{A[x]} A[x] \cong A[x]$ is a surjection of a polynomial ring in two variables, $x \otimes 1$ and $1 \otimes x$, onto the polynomial ring in one variable.

However, there are two important cases where the two tensor products are the same: one is when $S = R/I$ is a homomorphic image of R, and the other is when $S = W^{-1}R$

is a localization of R. In the first case, the point is that any scalar in R/I is the image of a scalar in r which can be passed through the tensor symbol: if $r \in R$ maps to $\overline{r}$ in R/I, we have that $\overline{r}u \otimes v = ru \otimes v = u \otimes rv = u \otimes \overline{r}v$. In the case where $S = W^{-1}R$, the scalars have the form r/w with $r \in R$ and $w \in W$. Because N is an S-module we can write $v = w(1/w)v$, and in $M \otimes_R N$, $(r/w)u \otimes v = (r/w)u \otimes w(1/w)v = w(r/w)u \otimes (1/w)v$ (since we may pass w through the tensor symbol) $= ru \otimes (1/w)v = u \otimes (r/w)v$, as required, because we can pass r through the tensor symbol.

Note that base change from R to R/I sends M to $(R/I) \otimes_R M \cong R/IM$, which is, of course, an (R/I)-module. This is particularly useful when m is a maximal ideal of R, for then M/mM is a vector space over the field $K = R/m$. If M and N are R-modules and m is a maximal ideal of R, we have surjections $f : M \twoheadrightarrow M/mM$ and $N \twoheadrightarrow N/mN$, and, hence a surjection $f \otimes g : M \otimes_R N \twoheadrightarrow (M/mM) \otimes_R (N/mN) \cong (M/mM) \otimes_K (N/mN)$, which is the tensor product of two vector spaces over a field, and is more readily understood than a tensor product over a base ring that is not a field: in particular, we can get a K-basis for this last tensor product by tensoring together pairs from a K-basis for M/mM and a K-basis for N/mN, and this makes it easy to understand whether an element of the tensor product either is or is not zero.

In particular, if $R = K[x, y]$ and $m = (x, y)$ thought of as an R-module, it is clear that no nonzero K-linear combination of x and y is in m^2, from which it follows that the images of x and y are a K-basis for m/m^2, which yields information about when elements of $(m/m^2) \otimes_K (m/m^2)$ are zero.

In trying to understand the tensor product of two finitely presented modules over R, one can use the fact that $M/N \otimes_R M'/N' \cong (M \otimes_R M')/\bigl(\mathrm{Im}\,(N \otimes_R M') + \mathrm{Im}\,(M \otimes_R N')\bigr)$ to give a finite presentation of the tensor product. Suppose that M is free on a free basis $\{v_i\}_i$, that M' is free on a free basis $\{w_j\}_j$, that N is the span of vectors $\{y_h\}_h$ and that N' is the span of vectors $\{z_k\}_k$. Then the tensor product $(M/N) \otimes_R (M'/N')$ is the quotient of the free R-module with free basis $\{v_i \otimes w_j\}_{i,j}$ by the R-span of all the vectors $y_h \otimes w_j$ together with all the vectors $v_i \otimes z_k$.

We next want to consider what happens to intersections of submodules when we tensor with a flat R-module F. Let N_1 and N_2 be any two R-submodules of M. Then $F \otimes_R N_1$, $F \otimes_R N_2$ and $F \otimes_R (N_1 \cap N_2)$ all inject canonically into $F \otimes_R M$. We identify each of these modules with its image in $F \otimes_R M$. We claim that $F \otimes_R (N_1 \cap N_2) = (F \otimes N_1) \cap (F \otimes N_2)$. *A priori,* we only have an inclusion. Consider the exact sequence of modules

$$0 \to N_1 \cap N_2 \xrightarrow{\alpha} N_1 \oplus N_2 \xrightarrow{\beta} M$$

where $\alpha(u) = u \oplus u$ and $\beta(u_1 \oplus u_2) = u_1 - u_2$. It is quite easy to see that this sequence is exact. The key point here is that when we apply $F \otimes_R \_$, this exactness is preserved, so that we get an exact sequence:

$$0 \to F \otimes_R (N_1 \cap N_2) \xrightarrow{1_F \otimes \alpha} F \otimes_R N_1 \oplus F \otimes_R N_2 \xrightarrow{1_F \otimes \beta} F \otimes M$$

and the map $1_F \otimes \beta$ sends $v_1 \oplus v_2$ to $v_1 - v_2$. It follows that the kernel of $1_F \otimes \beta$ is the image of $(F \otimes N_1) \cap (F \otimes N_2)$ under $1_F \otimes \alpha$, but because F is R-flat, we also that the kernel is the image of $F \otimes_R (N_1 \cap N_2)$ under $1_F \otimes \alpha$. □

Of course, this result extends by a straightforward induction to intersections involving finitely many submodules N_i of M, and it applies to flat base change when $F = S$ is a flat R-algebra. An important special case is when $S = W^{-1}R$, and we see that localization commutes with finite intersection of submodules.

Lecture of November 6

Note that if $I \subseteq R$ and S is an R-flat algebra then $I \otimes_R S$ injects into $R \otimes_S S \cong S$ with image IS: that means that $I \otimes_R S$ may be identified with IS.

Recall that if I and J are ideals of R then $I :_R J = \{r \in R : rJ \subseteq I\}$. If J is finitely generated, this colon operation commutes with flat base change:

Proposition. *If S is a flat R-algebra, and I, J are ideals of R with J finitely generated, then $(I :_R J)S = IS :_S JS$. In particular, this holds when S is a localization of R.*

Proof. If $J = fR$ is principal, we have an exact sequence

$$0 \to (I :_R fR)/I \to R/I \xrightarrow{f} R/I \to 0.$$

When we tensor with S and make obvious identifications, we get an an exact sequence

$$0 \to \big((I :_R fR)S\big)/IS \to S/IS \xrightarrow{f} S/IS \to 0.$$

But the kernel of multiplication by f on S/IS (this is the same as multiplication by the image of f in S) is $(IS :_S fS)/IS$, from which we can conclude that $IS :_S fS = (I :_R fR)S$. In the general case, where $J = (f_1, \ldots, f_h)R$, we use the obvious fact that $I :_R J = \bigcap_t (I :_R f_tR)$, and the fact that flat base change commutes with finite intersection. But we then have

$$(I :_R J)S = (I :_R J) \otimes_R S = (\bigcap_{t=1}^{h} I :_R f_tR) \otimes_R S = \bigcap_{t=1}^{h} \big((I :_R f_tR)S\big)$$

and by the case where $J = fR$, which we have already done, this becomes

$$\bigcap_{t=1}^{h} (IS :_S f_tS) = IS :_S JS,$$

as required. □

Example. This fails even for localization when J is not finitely generated. Let

$$S = K[y, x_1, x_2, x_3, \ldots]$$

be the polynomial ring in countably many variables over the field K. Let W be the multiplicative system of all powers of y. Let I be the ideal $(x_t y^t : t = 1, 2, 3, \ldots)S$, and let $J = (x_t : t = 1, 2, 3, \ldots)$. Then before localization at W, $I :_R J$ is an ideal not containing any power of y, since y^t fails to multiply x_{t+1} into I. Thus, with $S = W^{-1}R = R_y$, we have that $(I :_R J)S$ is a proper ideal. But $IS = JS$, and so $IS :_S JS = S$.

There are many simple examples where localization fails to commute with infinite intersection, even when the ring is Noetherian. E.g., If $R = \mathbb{Z}$ or $R = K[x]$ where K is a field, and I is generated by a prime element, then the intersection of the ideals I^n is (0). But if we localize at a generator of I, then all the ideals I^n expand to the unit ideal, and their intersection is the unit ideal.

Theorem. *Let R be a ring, M, M' be R-modules, let $f : M \to M'$ be R-linear, let $u \in M$, and let N and various N_i be submodules of M. The statements below hold when the phrase "for all P" is interpreted either to mean "for all prime ideals P of R" or "for all maximal ideals P of R."*

(a) *Let $f : M \to M'$ be R-linear. Then for all P, the formation of the kernel, cokernel and image of f commute with localization. E.g., $\bigl(\mathrm{Ker}\,(f)\bigr)_P \cong \mathrm{Ker}\,(f_P)$, where $f_P : M_P \to M'_P$ is the map induced by f.*

(b) *$u/1 \in M_P$ is nonzero if and only if $P \supseteq I = Ann_R u$. The element $u = 0$ in M if and only if $u/1 \in M_P$ is 0 for all P.*

(c) *$M = 0$ iff $M_P = 0$ for all P.*

(d) *$f : M \to M'$ is injective (respectively, surjective, respectively bijective) if and only if f_P is for all P.*

(e) *$u \in M$ is in N if and only if $u/1 \in M_P$ is in N_P for all P.*

(f) *$N_1 \subseteq N_2$ (respectively, $N_1 = N_2$) if and only if $(N_1)_P \subseteq (N_2)_P$ for all P.*

(g) *$0 \to M' \to M \to M'' \to 0$ is exact if and only if $0 \to M'_P \to M_P \to M''_P \to 0$ is exact for all P, and $M' \to M \to M''$ is exact if and only if $M'_P \to M_P \to M''_P$ is exact for all P.*

Proof. (a) follows from the exactness of localization at P, and is also valid for localization at an arbitrary multiplicative system and, in fact, for arbitrary flat base change.

To prove (b), note that the surjection $R \to Ru$ sending r to ru has kernel I, so that $Ru \cong R/I$. Now $(R/I)_P \neq 0$ iff I is disjoint from the multiplicative system $R - P$, which is equivalent to $P \supseteq I$. The last statement follows because if $u \neq 0$ then I is proper and we can choose P maximal containing I. Part (c) is immediate: if $u \neq 0$ in M, then $Ru \hookrightarrow M$, and this is preserved when we localize at P containing $I = \mathrm{Ann}_R u$.

(d) follows from parts (a) and (c): f is injective iff $\mathrm{Ker}\, f = 0$ iff $\bigl(\mathrm{Ker}\,(f)\bigr)_P = 0$ for all P iff $\mathrm{Ker}\,(f_P) = 0$ for all P iff f_P is injective for all P. The argument for surjective is the same with the cokernel replacing the kernel. A map is bijective if and only if it is both injective and surjective.

(e) follows from (b) applied to the class of u in M/N.

For (f), note that $N_1 \subseteq N_2$ iff $N_1/(N_1 \cap N_2) = 0$. Now use the fact that localization commutes with intersection coupled with (c). The second part follows since $N_1 = N_2$ iff $N_1 \subseteq N_2$ and $N_2 \subseteq N_1$ (an alternative is to use the fact that $N_1 = N_2$ iff the module $(N_1 + N_2)/(N_1 \cap N_2) = 0$.

The first statement in part (g) follows from the second (applied repeatedly), and the second statement follows from the fact that the calculations of image and kernel commute with localization, which is part (a), together with the fact that exactness holds iff the image of $M' \to M$ is equal to the kernel of $M \to M''$, together with part (f). □

Next note that given an R-algebra S there is an S-linear map

$$\theta_M : S \otimes_R \mathrm{Hom}_R(M,\, N) \to \mathrm{Hom}_S(S \otimes_R M,\, S \otimes_R N)$$

that sends $s \otimes f$ to $s(1_S \otimes f)$: the map is well-defined because $(s,\, f) \mapsto s(1_S \otimes f)$ is R-bilinear, and S-linear because the image of $s'(s \otimes f) = (s's) \otimes f$ is $(s's)(1_S \otimes f) = s'\bigl(s(1_S \otimes f)\bigr)$. Moreover, given a map $g : M \to M'$ there is a commutative diagram:

$$\begin{tikzcd}
S \otimes_R \operatorname{Hom}_R(M,\, N) \arrow[r, "\theta_M"] & \operatorname{Hom}_S(S \otimes_R M,\, S \otimes_R N) \\
S \otimes_R \operatorname{Hom}_R(M',\, N) \arrow[u, "1_S \otimes g^*"] \arrow[r, "\theta'_M"'] & \operatorname{Hom}_S(S \otimes_R M',\, S \otimes_R N) \arrow[u, "(1_S \otimes g)^*"']
\end{tikzcd}$$

so that the θ_M taken together give a natural transformation of contravariant functors from $S \otimes_R \operatorname{Hom}_R(\_,\, N)$ to $\operatorname{Hom}_S(S \otimes_R \_,\, S \otimes_R N)$. The commutativity of the diagram may be checked on elements of the form $s \otimes f$, where $f : M \to M'$. Applying the map in the leftmost column first and then the map in the top row, we get first $s \otimes (f \circ g)$ and then $s(1_S \otimes (f \circ g)) = s\bigl((1_S \otimes f) \circ (1_S \otimes g)\bigr)$, while going around the square the other way one first gets $s(1_S \otimes f)$ and then $\bigl(s(1_S \otimes f)\bigr) \circ (1_S \otimes g)$ $\square$

Proposition (Hom commutes with flat base change). *If S is a flat R-algebra and M, N are R-modules such that M is finitely presented over R, then the canonical homomorphism*

$$\theta_M \colon S \otimes_R Hom_R(M, N) \to Hom_S(S \otimes_R M, S \otimes_R N)$$

sending $s \otimes f$ to $s(1_S \otimes f)$ is an isomorphism.

Proof. It is easy to see that θ_R is an isomorphism and that $\theta_{M_1 \oplus M_2}$ may be identified with $\theta_{M_1} \oplus \theta_{M_2}$, so that θ_G is an isomorphism whenever G is a finitely generated free R-module.

Since M is finitely presented, we have an exact sequence $H \to G \twoheadrightarrow M \to 0$ where G, H are finitely generated free R-modules. In the diagram below the right column is obtained by first applying $S \otimes_R \_$ (exactness is preserved since $\otimes$ is right exact, and then applying $\operatorname{Hom}_S(\_, S \otimes_R N)$, so that the right column is exact. The left column is obtained by first applying $\operatorname{Hom}_R(\_, N)$, and then $S \otimes_R \_$ (exactness is preserved because of the hypothesis that S is R-flat). The squares commute because the θ_M give a natural transformation.

$$\begin{tikzcd}
S \otimes_R \operatorname{Hom}_R(H, N) \arrow[r, "\theta_H"] & \operatorname{Hom}_S(S \otimes_R H, S \otimes_R N) \\
S \otimes_R \operatorname{Hom}_R(G, N) \arrow[u] \arrow[r, "\theta_G"] & \operatorname{Hom}_S(S \otimes_R G, S \otimes_R N) \arrow[u] \\
S \otimes_R \operatorname{Hom}_R(M, N) \arrow[u] \arrow[r, "\theta_M"] & \operatorname{Hom}_S(S \otimes_R M, S \otimes_R N) \arrow[u] \\
0 \arrow[u] \arrow[r] & 0 \arrow[u]
\end{tikzcd}$$

From the fact, established in the first paragraph, that θ_G and θ_H are isomorphisms and the exactness of the two columns, it follows that θ_M is an isomorphism as well (kernels of isomorphic maps are isomorphic). □

Corollary. *If W is a multiplicative system in R and M is finitely presented, we have that $W^{-1}Hom_R(M, N) \cong Hom_{W^{-1}R}(W^{-1}M, W^{-1}N)$.* □

The result fails when M is not finitely generated, even if it is free. Let M be the free $\mathbb{Z}$-module on countably many generators b_i, and let $N = \mathbb{Z}$. Giving an element of $\mathrm{Hom}_{\mathbb{Z}}(M,\,\mathbb{Z})$ is equivalent to specifying its values on the free generators, i.e., to giving a sequence of integers n_i, where n_i is the value of the homomorphism on b_i. Let $S = \mathbb{Z}[1/p]$. Any element of $S \otimes_{\mathbb{Z}} \mathrm{Hom}_{\mathbb{Z}}(M,\,\mathbb{Z})$ then corresponds to a sequence of elements in S such that the denominators are bounded: in this module, we can clear denominators. However, $\mathrm{Hom}_S(S \otimes M,\,S)$ is larger: elements correspond to arbitrary sequences in S. In particular, the homomorphism whose value on $b_i/1$ is $1/p^i$ for all i is not in the image of $S \otimes_{\mathbb{Z}} \mathrm{Hom}_{\mathbb{Z}}(M,\,\mathbb{Z})$. When M is finitely generated, even cyclic, the result still fails if M is not finitely presented. Let R, I, and J be as in the example immediately following the Proposition at the beginning of this lecture, let $M = R/J$ let $N = R/I$, let $W = \{y^t : t \in \mathbb{N}\}$ and $S = W^{-1}R = R_y$. Then $\mathrm{Hom}_R(M,\,N) \cong \mathrm{Ann}_N J \cong (I :_R J)/I$, while, similarly, $\mathrm{Hom}_S(W^{-1}M,\,W^{-1}N) \cong (IS :_S JS)/IS$. Thus, the issue becomes whether $S \otimes_R \bigl((I :_R J)/I\bigr) \cong (IS :_S JS)/IS$, and since the left hand side is $(I :_R J)S/IS$, the issue is simply whether $(I :_R J)S = IS :_S JS$. We have seen in the earlier Example following the Proposition about expansions of colon ideals to flat algebras that this is false.

Finally, the result also fails without flatness. For example, let $R = \mathbb{Z}$ and $S = \mathbb{Z}/p\mathbb{Z}$ for some prime integer p. Then $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z},\,\mathbb{Z}) = 0$, and so $(\mathbb{Z}/p\mathbb{Z}) \otimes_{\mathbb{Z}} \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z},\,\mathbb{Z}) = 0$, while $\mathrm{Hom}_{\mathbb{Z}/p\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z},\,\mathbb{Z}/p\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z} \neq 0$. We next note:

Proposition. *If $Q = M/N$ is finitely presented over R, then the short exact sequence of R-modules $0 \to N \to M \xrightarrow{f} Q \to 0$ splits (i.e. there exists $g : Q \to M$ such that $g \circ f = 1_Q$) if and only if $0 \to N_P \to M_P \to Q_P \to 0$ splits for all prime (respectively, maximal) ideals P in R.*

Proof. The sequence splits if and only if $\mathrm{Hom}_R(Q,\,M) \to \mathrm{Hom}_R(Q,\,Q)$ is onto. (If M is $N \oplus Q$ then $\mathrm{Hom}_R(V,\,M) \cong \mathrm{Hom}_R(V,\,N) \oplus \mathrm{Hom}_R(V,\,Q)$ for all R-modules V, which implies the surjectivity. On the other hand if the map is surjective then 1_Q is the image of some $g : Q \to M$, and this means that $f \circ g = 1_Q$.) It is clear that if the map splits it continues to do so after localization (or any base change, whether flat or not). If the map does not split, then the map $\mathrm{Hom}_R(Q, M) \to \mathrm{Hom}_R(Q, Q)$ is not onto, and this will be preserved when we localize at a suitable P. By the theorem, flat base change commutes with Hom in this case, and so we have that $\mathrm{Hom}_{R_P}(Q_P,\,M_P) \to \mathrm{Hom}_{R_P}(Q_P,\,Q_P)$ is not onto as well. □

We next note that if $R \to S \to T$ are homomorphisms of rings and M is any R-module, then there is a bijective S-linear map $T \otimes_S (S \otimes_R M) \to T \otimes_R M$: the left side may be identified with $(T \otimes_S S) \otimes_R M$ by the second form of the associativity of tensor, and $T \otimes_S S \cong T$. These isomorphisms are easily checked to be isomorphisms of T-modules, and together these isomorphisms give a natural transformation, showing that $T \otimes_S (S \otimes_R \_)$ and $T \otimes_R \_$ are isomorphic functors from R-modules to T-modules. Put more briefly, an iterated base change can be done instead with a single base change.

Corollary. *If S is flat over R and T is flat over S, then T is flat over R.*

Proof. Given an injection $N \hookrightarrow M$ we have an injection $S \otimes_R N \hookrightarrow S \otimes_R M$, since S is R-flat, and then an injection $T \otimes_S (S \otimes_R N) \hookrightarrow T \otimes_S (S \otimes_R M)$, since T is S-flat, and this is the same map as $T \otimes_R N \to T \otimes_R N$. $\square$

We also note:

Proposition. *If F is flat, free or projective over R, then $S \otimes_R M$ has the corresponding property over S.*

Proof. We know this for free modules G, and if $G = P \oplus Q$, then $S \otimes_R G = (S \otimes_R P) \oplus (S \otimes_R Q)$. Now suppose that F is R-flat. The fact that $S \otimes_R F$ is S-flat is immediate from the fact that tensoring with this module over S is isomorphic as functor with tensoring with F over R: the identification $(S \otimes_R F) \otimes_S M \cong F \otimes_R M$ follows from the associativity of tensor if we rearrange the terms: $M \otimes_S (S \otimes_R F) \cong (M \otimes_S S) \otimes_R F = M \otimes_R F$. $\square$

Lemma. *If A and B are any two R-modules, W is a multiplicative system in R, and $S = W^{-1}R$, then $W^{-1}(A \otimes_R B) \cong W^{-1}A \otimes_S W^{-1}B$ in such a way that $(a \otimes b)/1$ maps to $(a/1) \otimes (b/1)$.*

Proof. We have already seen that if U, V are S-modules, then $U \otimes_R V \cong U \otimes_S V$, so that $S \otimes_R S \cong S \otimes_S S \cong S$. Thus, $W^{-1}(A \otimes_R B) \cong S \otimes_R (A \otimes_R B) \cong (S \otimes_R S) \otimes_R (A \otimes_R B) \cong (S \otimes_R A) \otimes_R (S \otimes_R B) \cong W^{-1}A \otimes_R W^{-1}B \cong W^{-1}A \otimes_S W^{-1}B$ $\square$

Proposition. *F is R-flat if and only if F_P is R_P-flat for all prime (respectively, maximal ideals) P.*

Proof. We have already seen "only if." Now suppose that M_P is R_P-flat for all maximal ideals P. Suppose that $N \subseteq M$ but that $F \otimes_R N \to F \otimes_R M$ has a nonzero kernel V. We can choose P maximal such that V_P is not 0. Then $(F \otimes_R N)_P \to (F \otimes_R M)_P$ is not injective. By the preceding Lemma, this may be identified with $F_P \otimes_{R_P} N_P \to F_P \otimes_{R_P} M_P$. Since $N \to M$ is injective, so is $N_P \to M_P$, and this contradicts the flatness of F_P over R_P. $\square$

We define the *support* of the module M to be $\{P \in \mathrm{Spec}\,(R) : M_P \neq 0\}$. We have seen earlier that every nonzero module has nonempty support. If M is finitely generated, we can say a lot more.

Proposition. *Let M be a finitely generated R-module with annihilator I in R. Then the support of M is closed, and is equal to $V(I)$.*

Proof. Let $u_1, \ldots, u_h$ generate M, and let I_t be the annihilator of u_t. An element kills M if and only if it kills all the generators of M, and so $I = \bigcap_t I_t$. Since M is the sum of the Ru_t, and each Ru_t injects into M, we have that $M_P \neq 0$ if and only if some $(Ru_t)_P \neq 0$, i.e., if and only if $P \supseteq I_t$ for some t. This shows that the support of M is $\bigcup_t V(I_t) = V(\bigcap_t I_t) = V(I)$, as required. $\square$

Lecture of November 8

Discussion: product decompositions. We want to examine when a ring R has a decomposition $R \cong S \times T$ as the product of two rings: we call a decomposition *trivial* when S or T is 0 and the other factor is isomorphic to R. Note that $S \times T$ has two idempotents, $e = (1, 0)$ and $f = 1 - e = (0, 1)$. In general, if e is idempotent we call $1 - e$ its *complementary* idempotent. Complementary idempotents are characterized by the equations $e + f = 1$ and $ef = 0$. The latter then implies $e(1 - e) = 0$ or $e = e^2$. In the product situation note that $S \cong S \times \{0\}$ with the identity corresponding to e. Also, $R/fR \cong S$ and the localization $R_e \cong S$. $\mathrm{Spec}\,(S) \approx V(f) = D(e)$ and $\mathrm{Spec}\,(T) \cong V(e) = D(f)$. Note also that $\mathrm{Spec}\,(R)$ is the disjoint union of the sets $V(f)$ and $V(e)$ (no prime contains both, since $e + f = 1$, and every prime contains one of them, since $ef = 0$). Thus, $\mathrm{Spec}\,(R)$ is the disjoint union of $\mathrm{Spec}\,(S)$ and $\mathrm{Spec}\,(T)$, each of which is both closed and open, i.e., *clopen*, in $\mathrm{Spec}\,(R)$. Thus, a non-trivial product decomposition of R gives a disconnection of $\mathrm{Spec}\,(R)$, i.e., a way of writing it as a the disjoint union of two non-empty closed sets (which are then automatically open as well, since each is the complement of the other).

Conversely, giving complementary idempotents e, f in R gives an essentially unique product decomposition that gives rise to them as above: each $r \in R$ can be written as $r(e + f) = re + rf$, so that $R = Re + Rf$. The sum is direct, for if $re = r'f$ we may multiply by e to get $re = re^2 = r'ef = r'(0) = 0$, so that $Re \cap Rf = (0)$. Then Re is a ring S with identity e, Rf is a ring T with identity f, and $R \cong S \times T$: note that $(ae + bf)(a'e + b'f) = aea'e + bfb'f$, because the two terms not shown are multiples of ef and so are 0.

Remarkably, giving a disconnection of $\mathrm{Spec}\,(R)$ as the union of disjoint closed sets X, Y yields unique nontrivial complementary idempotents e, f such that $X = V(f)$ and $Y = V(e)$. Thus, the issue of whether R is a non-trivial product of two rings is topological. Let $X = V(J)$ and $Y = V(I)$. Since $X \cap Y = \emptyset$, $I + J = R$. Since $X \cup Y = \mathrm{Spec}\,(R)$, IJ is contained in every prime and so is contained in the ideal $N = \mathrm{Rad}\,(0)$ of all nilpotent elements. Choose $u \in J$, $v \in I$ such that $u + v = 1$. Then uv is nilpotent, so that u, v have images that are complementary idempotents in R_{red}. It is shown below that these lift uniquely to idempotent elements e, $f \in R$ such that $e - u$, $f - v \in N$. Now, since $e \in J + N$, $X' = V(e) \supseteq V(J + N) = V(J) = X$, and, similarly, $Y' = V(f) \supseteq Y$. Since $X' \supseteq X$ and is disjoint from $Y' \supseteq Y$, $X' = X$. Similarly $Y' = Y$. This completes the proof once we have shown that idempotents lift uniquely.

To see this, suppose that $u + v = 1$ and $(uv)^n = 0$. Then $(u + v)^{2n-1} = 1$, and when we expand the left side using the binomial theorem the first n terms are divisible by u^n (and all but the first are divisible by v) and the next (last) n terms are divisible by v^n. Call the sum of the first n terms e and the sum of the last n terms f. Then $e + f = 1$ and ef is a multiple of $u^n v^n$, and so 0. Thus, $e(1 - e) = 0$ and e is idempotent. But e has the form $u^{2n-1} + u^n vr$, and, mod nilpotents, $u^{2n-1} \equiv u$ and $u^n vr \equiv 0$, so that $e \equiv u$. This shows the existence of e lifting u. Suppose e' is another idempotent that such that $e - e' \neq 0$ is nilpotent. Then we can localize at a maximal ideal that contains the annihilator of $e - e'$, and obtain an example in a quasilocal ring by localizing at m. But in a quasilocal ring, $e(1-e)=0$ and either e or $1-e$ must be a unit (they cannot both be in the maximal ideal), so that e must be 0 or 1. Since e' must also be 0 or 1 and e, e' must be 0 and 1 in some order. But then $e-e'$ is not nilpotent. □

We note that if (R-mod) (respectively (R-alg)) denotes the category of R-modules (respectively, R-algebras) and R-linear (respectively, R-algebra) maps, then $\bigl((S\times T)\text{-mod}\bigr)$ is equivalent to the product catgeory (S-mod)×(T-mod) (respectively, $\bigl((S\times T)\text{-alg}\bigr)$ is equvialent to the product category (S-alg)×(T-alg).

We continue our discussion of properties that can be checked locally. If R is reduced, note that $W^{-1}R$ is reduced: if r/w is nilpotent, then $r^n/w^n=0$ for some n,i., and then $r^n/1=0$ so that $w'r^n=0$ for some $w'\in W$, which implies that $(w'r)^n=0$. Since R is reduced, this yields that $w'r=0$, and so $r/1=0$ and $r/w=0$.

Note also that if R is a domain with fraction field $\mathcal{F}$, and W is a multiplicative system in $R-\{0\}$, then $W^{-1}R\cong R[1/w : w\in W]\subseteq\mathcal{F}$, and so $W^{-1}R\subseteq\mathcal{F}$ is a domain. The following theorem allows us to che certain properties of R locally.

Proposition. *Let R be a ring. The statements below are valid if "for all P" is interpreted to mean either "for all prime ideals P of R" or "for all maximal ideals P of R."*

(a) *R is reduced iff R_P is reduced for all P.*

(b) *If R is a domain, then R is normal if and only if R_P is normal for all P.*

(c) *If R is Noetherian, or, more generally, if R has only finitely many minimal primes, then R is a domain if and only if* Spec$\,(R)$ *is connected and R_P is a domain for all P.*

Proof. (a) We have already seen that if R is reduced then so are all of its localizations. But if R is not reduced and $r\neq 0$ is a nilpotent, we can choose P so that $r/1\in R_P$ is not 0, and it will still be nilpotent.

(b) Let D' indicate the integral closure of the domain D in its fraction field. Let $\mathcal{F}$ be the fraction field of R, which is also the fraction field of $W^{-1}R$ for any multiplicative system $W\subseteq R-\{0\}$. By problems **1.** and **2.** of Supplementary Problem Set #2, we know that the integral closure of $W^{-1}R$ in $W^{-1}\mathcal{F}=\mathcal{F}$ is $W^{-1}R'$. In particular, it follows that for all P, $(R_P)'=(R')_P$. But R is normal iff $R'=R$ iff $R'/R=0$ (as an R-module) iff $(R'/R)_P=0$ for all P, and $(R'/R)_P\cong(R')_P/R_P=(R_P)'/R_P$, so that R' is normal if and only if $(R_P)'=R_P$ for all P, i.e., every R_P is normal.

(c) It is clear that if R is a domain then every R_P is and Spec$\,(R)$ is connected: Spec$\,(R)$ is not connected iff R is a product in a non-trivial way iff R contains an idempotent e other than 0, 1, and the equation $e(1-e)=0$ in a domain implies $e=0$ or $e=1$. Now suppose that Spec$\,(R)$ is connected and that R_P is a domain for all P. By part (a), R is reduced. Let $P_1,\,\ldots\,,P_k$ be the minimal primes of R. The union of the closed sets $V(P_t)$ is Spec$\,(R)$, since every prime contains a minimal prime, and they are mutually disjoint, for if Q contains both P_i and P_j (we may assume that Q is maximal, replacing it by a maximal ideal that contains it if necessary), we have that R_Q has at least two minimal primes, corresponding to P_i and P_j, contradicting the assumption that R_Q is a domain.

But then these sets are all open as well as closed, and since Spec (R) is connected it follows that there is a unique minimal prime P. But in any commutative ring, the intersection of the minimal primes is the same as the intersection of all primes: it is the ideal of all nilpotents. Thus, if there is a unique minimal prime, all of its elements are 0, since R is reduced. But this means that (0) is prime, so that R is a domain. □

The statement that (R, m) is quasilocal means that R has unique maximal ideal m. The statement that (R, m, K) is quasilocal means that R is quasilocal with maximal ideal m and residue class field $K \cong R/m$. If R is Noetherian and quasilocal one says that R, or $(R,\, m)$ or $(R,\, m,\, K)$ is *local* instead. Let M be an R-module. Then $M/mM \cong (R/m) \otimes_R M \cong K \otimes M$ is a K-vector space. In a quasilocal ring, if $r \in m$, then $1 - r \notin m$. which implies that $1 - r$ is a unit: otherwise it would generate a proper ideal, which then must be contained in the unique maximal ideal m.

Theorem (Nakayama's lemma). . *Let M be a finitely generated module over the quasilocal ring $R = (R,\, m,\, K)$. If $M = mM$, i.e., if $K \otimes_R M = 0$, then $M = 0$.*

Proof. We use induction on the number of generators n of M. If $M = Ru$, then $mM = mu$, and so if $M = mM$ we must have $u = ru$ for some $r \in m$. Then $(1 - r)u = 0$, and since $1 - r$ is a unit, we have that $u = 0$. At the inductive step suppose that M is generated by $u_1, \ldots, u_n$. Let $M_1 = M/Ru_n$, which is generated by $n - 1$ elements. We still have that $mM_1 = M_1$, and so, by the induction hypothesis, $M_1 = 0$, which says that $M = Ru_n$. But we have already done the case where $n = 1$. □

Corollary (Nakayama's lemma, second form). *If J is contained in every maximal ideal of R, M is finitely generated, and $M = JM$, then $M = 0$.*

Proof. It suffices to show that $M_P = 0$ for all maximal P. But $M = JM \Rightarrow M_P = JM_P \Rightarrow M_P = PM_P = (PR_P)M_P$, and so $M_P = 0$. □

We leave it to the reader to see that j is in every ideal maximal ideal of R if and only if $1 - jr$ is a unit for every element r of R. The intersection of the maximal ideals of R is called the *Jacobson radical* of R.

Corollary (Nakayama's lemma, third form). *Let M be a finitely generated module over a quasilocal ring (R, m, K). Then $u_1, \ldots, u_n$ generate M if and only their images in the K-vector space M/mM span M/mM over K. Hence, from any set of generators of M, one may choose a subset that is a minimal set of generators, and all such minimal sets of generators have the same cardinality, which is the K-vector space dimension of M/mM. Any set of elements of M whose images are linearly independent in M/mM can be extended to a minimal set of generators of M/mM.*

Proof. Let $N = Ru_1 + \cdots + Ru_n$. Then $M = N$ iff $M/N = 0$ iff $M/N = m(M/N)$ iff $M = N + mM$ iff the image of N in M/mN is all of M/mM iff the images of the u_i span M/mM. A set of elements of M is a minimal set of generators for M iff its image in M/mM is a K-vector space basis. The remaining statements in the theorem are a consequence of the fact that every set of vectors that spans a vector space has a subset that is a basis, and every independent set of vectors can be enlarged to a basis. □

We want to use Nakayama's lemma to investigate the support of the tensor product of two finitely generated modules. The following fact comes up frequently enough that it is worth isolating:

Lemma. *If M, N are arbitrary R-modules and P is a prime of the ring R, then there is an isomorphism $(M \otimes_R N)_P \cong M_P \otimes_{R_P} N_P$.*

Proof. We use that the tensor product of two modules over a localization of R is independent of whether it is taken over R or over the localization. In particular, $(R_P \otimes_R \otimes R_P) \cong R_P \otimes_{R_P} R_P \cong R_P$. Thus, $(M \otimes_R N)_P = R_P \otimes_R (M \otimes_R N) \cong (R_P \otimes_R R_P) \otimes_R (M \otimes_R N) \cong (R_P \otimes M) \otimes_R (R_P \otimes_R N) \cong M_P \otimes_R N_P \cong M_P \otimes_{R_P} N_P$. □

Proposition. *Over any ring R, if M and N are finitely generated R-modules, then the support of $M \otimes_R N$ is the intersection of the supports of M and N.*

Proof. First suppose that (R, m, K) is quasilocal and that M, N are nonzero. We claim that $M \otimes_R N$ is nonzero: we have surjection $M \otimes_R N \to (M/mM) \otimes_R (N/mN) \cong (M/mM) \otimes_K (N/mN)$. By Nakayama's lemma, M/mM and N/mN are nonzero K-vector spaces, and so their tensor product over K is nonzero.

In the general case, $(M \otimes N)_P \cong M_P \otimes_{R_P} N_P$, and so vanishes if and only if $M_P = 0$ or $N_P = 0$. □

If a module M if finitely presented there is an exact sequence $0 \to N \to R^n \twoheadrightarrow M \to 0$ with N finitely generated. It turns out that if one chooses a different surjection of $R^{n'}$ to M, the kernel N' will also be finitely generated: this is a problem in Supplementary Problem Set #5. The idea is to compare each of two sets of generators with their union, and then to reduce to the case where one has two sets of generators, one of which is obtained from the other by enlarging it with a single redundant element. We shall assume this fact here.

Note that if (R, m, K) is quasilocal, and M has $u_1, \ldots, u_n$ as a minimal set of generators, we may map $R^n \twoheadrightarrow M$ so that $(r_1, \ldots, r_n) \mapsto r_1 u_1 + \cdots + r_n u_n$. If one tensors with K, one gets a surjection $K^n \to M/mM$, and since the images of the u_i are vector space basis for M/mM, this surjection is actually an isomorphism.

If $f : A \to B$ and $g : C \to D$ are maps of R-modules then there is a diagram:

$$\begin{tikzcd}
A \otimes_R C \arrow[r, "1_A \otimes g"] \arrow[d, "f \otimes 1_C"'] & A \otimes_R D \arrow[d, "f \otimes 1_D"] \\
B \otimes_R C \arrow[r, "1_B \otimes g"'] & B \otimes_R D
\end{tikzcd}$$

Note that $(f \otimes 1_D) \circ (1_A \otimes g) = f \otimes g = (1_B \otimes g) \circ (f \otimes 1_C)$, so the diagram commutes.

We are now ready to prove:

Theorem. *Let M be a finitely presented module over a quasilocal ring (R, m, K). Then the following conditions are equivalent:*
(a) *M is free.*
(b) *M is projective.*
(c) *M is flat.*
(d) *The map $m \otimes M \to M$ that sends $r \otimes u$ to ru is injective.*

Proof. We already know that (a) ⇒ (b) ⇒ (c) ⇒ (d): the last implication comes from applying $\_ \otimes_R M$ to the injection $0 \to m \subseteq R$. We only need to show that (d) ⇒ (a).

Choose a minimal set of generators $u_1, \ldots, u_n$ for M and map R^n onto M 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 $R^n \twoheadrightarrow M$, so that we have a short exact sequence $0 \to N \to R^n \to M \to 0$. 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):

$$\begin{tikzcd}
 & & & 0 \arrow[d] & \\
 & m \otimes N \arrow[r] \arrow[d] & m \otimes R^n \arrow[r] \arrow[d] & m \otimes M \arrow[r] \arrow[d, "\alpha"] & 0 \\
0 \arrow[r] & N \arrow[r] \arrow[d] & R^n \arrow[r] \arrow[d] & M \arrow[r] \arrow[d] & 0 \\
 & K \otimes N \arrow[r, "f"] \arrow[d] & K \otimes R^n \arrow[r, "g"] \arrow[d] & K \otimes M \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

The rows are obtained by applying $m \otimes \_$, $R \otimes \_$, and $K \otimes \_$, respectively to the short exact sequence $0 \to N \to R^n \to M \to 0$, and the columns are obtained by applying $\_ \otimes N$, $\_ \otimes R^n$, and $\_ \otimes M$, 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 R is free, and the injectivity of α is the hypothesis in (d). (We also have an injection at the top of the middle column because R^n is free, but we don't need this.)

The four squares in the diagram commute: each has the form described in the remark just preceding the statement of the theorem.

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! □

Lecture of November 10

Note that a finitely generated projective R-module P is automatically finitely presented: let Q be the kernel of a surjection $R^n \twoheadrightarrow P$. As we have already seen, this surjection splits, so that $P \oplus Q \cong R^n$. But then $Q \cong R^N/P$ is finitely generated as well.

Also note that if $P \subseteq Q$ are primes then $R_P \cong (R_Q)_{P^{\mathrm{e}}}$ where $P^{\mathrm{e}} = PR_Q$, so that if M_Q is R_Q-free for all maximal ideals Q then M_P is *a priori* R_P-free for all prime ideals P as well, since freeness is preserved by arbitrary base change.

We next give a global version of what we just proved for the quasilocal case:

Theorem. *Let M be a finitely presented R-module. The following conditions are equivalent:*
(a) *M is projective.*
(b) *M is flat.*
(c) *M is locally free, i.e., for all maximal ideals (respectively, for all prime ideals) P of R, M_P is R_P-free.*

Proof. We know (a) ⇒ (b). If M is flat over R, M_P is flat over R_P, and so M_P is free over R_P by the preceding result. It remains to show that (c) ⇒ (a). Map a finitely generated free module R^n onto M. We get an exact sequence $0 \to N \to R^n \to M \to 0$. If we localize at any prime P, M_P is free over R_P, and then the sequence splits. By the Proposition on p. 4 of the Lecture Notes for November 6, the sequence splits over R, so that M is projective. □

We next want to use localization as a tool to study the structure of the ideals of a Noetherian ring R. We want to show that every ideal is a finite intersection of rather special ideals called *primary ideals.* In the integers, the primary ideals are the same as the ideal (0) and the ideals generated by a power of a prime element, and this is true more generally in any principal ideal domain. In the general case the situation is much more complicated. Every ideal is a finite intersection of primary ideals, and if the intersection is irredundant in a sense that we shall make precise, then it satisfies certain uniqueness statements. However, for many ideals the so-called *primary decomposition* is not unique.

This theory was first developed by the chess champion Emmanuel Lasker for polynomial rings finitely generated over a field, and then for arbitrary Noetherian rings by Emmy Noether. The irredundant primary decomposition of an ideal is also called the *Noether-Lasker decomposition.*

An ideal I in a ring R is called *primary* if whenever $ab \in I$ then either $a \in I$ or $b \in \mathrm{Rad}\,(I)$. If $ab \in \mathrm{Rad}\,(I)$, then $a^n b^n \in I$, so that either $a^n \in I$ or $b^n \in \mathrm{Rad}\,(I)$. But then either $a \in \mathrm{Rad}\,(I)$ or $b \in \mathrm{Rad}\,(I)$. Thus, if I is primary, $\mathrm{Rad}\,(I)$ is prime, say P, and one says that I is *primary* to P.

It is not true that I must be primary simply because its radical is prime. Let $I = (x^2,\, xy) \subseteq R = K[x,\, y]$, a polynomial ring in two variables. Then $\mathrm{Rad}\,(I) = xR$, which is prime. However, $xy \in I$, while x is not in I and y is not in $\mathrm{Rad}\,(I)$. On the other hand, by part (a) of the result that follows, it is true that an ideal is primary if its radical is a maximal ideal. Moreover, a prime ideal P is primary to itself.

Proposition. *Let R be a ring and I an ideal of R with radical P.*

(a) *If P is maximal, then I is primary to P.*

(b) *I is primary if and only if P is prime and I is contracted with respect to $R-P$. Thus, the ideals primary to P are in bijective correspondence with the ideals primary to the maximal ideal PR_P of R_P.*

(c) *I is primary to P if and only if P/I is prime and the elements of $R - P$ are not zerodivisors on R/I, that is, if and only if the nilpotent elements in R/I form a prime ideal (which will necessarily be the unique minimal prime) and the elements that are not nilpotent in R/I are not zerodivisors.*

(d) *If $J \subseteq I$, then I is primary to P if and only if I/J is primary to P/J in R/J.*

Proof. (a) Suppose $ab \in I$ and b has no power in I. Then $b \notin \mathrm{Rad}\, I$, which is maximal. It follows that $\mathrm{Rad}\,(I) + Rb = R$, so that $V\big(\mathrm{Rad}\,(I) + Rb\big) = \emptyset$, and this is the same as $V(I + Rb)$, so that $I + Rb = R$, say $i + rb = 1$ with $i \in I$ and $r \in R$. Then $a = a(i + rb) = ai + rab \in I$, as required, since $i,\, ab \in I$.

(b) We already know that if I is primary then $P = \mathrm{Rad}\,(I)$ is prime, so we may assume this. The definition of primary ideal then says precisely that if $ab \in I$ and $b \in R - P$, then $a \in I$, which is the definition of being contracted with respect to $R - P$. The second statement then follows from the general fact that ideals of R_P are in bijective correspondence with ideals of R contracted with respect to $R - P$, restricted to the case where the radical of the ideal is P.

The statement in (c) is equivalent to the statement in (b), since P/I is prime iff P is prime, and since the image of $b \in R - P$ in R/I is a not a zerodivisor if and only if for all $a \in R$, $ab \in I$ implies $a \in I$.

(d) Part (c) characterizes when I is primary in terms of the quotient ring R/I: the nilpotent elements from a unique minimal prime, and the elements that are not nilpotent are not zerodivisors. Part (d) follows that once, since $(R/J)/(I/J) \cong R/I$. $\square$

Proposition. *Let R be a ring and P a prime ideal of R.*

(a) *The intersection of finitely many P-primary ideals is P-primary.*

(b) *If $R \to S$ is a ring homomorphism, and J is an ideal of S primary to a prime ideal Q lying over P in R, then the contraction I of J to R is primary to P.*

Proof. (a) Suppose that I_1 and I_2 are primary to P. Since every element of P has a power in I_1 and a power in I_2, the higher of these two powers will be in $I_1 \cap I_2$, and so $\mathrm{Rad}\,(I_1 \cap I_2) = P$. Suppose that $ab \in I_1 \cap I_2$ and $a \notin I_1 \cap I_2$. But if $a \notin I_t$ for $t = 1$ or $t = 2$ then $b \in P = \mathrm{Rad}\,(I_1 \cap I_2)$. The general case follows by an obvious induction on the number of ideals.

(b) We have an injection of $R/I \hookrightarrow S/J$, since I is the contraction of J to R. The elements of P/I map into Q/J, and are nilpotent in S/J. Therefore, they are nilpotent in R/I. The elements of $R/I - P/I$ map into $S/J - Q/J$, and are therefore not zerodivisors in S/J. It follows that they are not zerodivisors in the subring R/I. The result follows from part (c) of the preceding Proposition. □

A *primary decomposition* of an ideal I is a representation of I as a finite intersection of of primary ideals. Given such a decomposition, if several of the ideals have the same radical, we may intersect them, and so give a decomposition that involves intersecting fewer ideals. If some proper subset of the ideals has the same intersection, we may work with that proper subset instead of the original set of ideals. Therefore, if an ideal has a primary decomposition it has a primary decomposition satisfying:

(1) The radicals of the mutually distinct ideals occurring are mutually distinct primes.

(2) No term may be omitted without strictly increasing the intersection.

Such a primary decomposition is called *irredundant.*

We shall prove that every ideal of a Noetherian ring has an irredundant primary decomposition, and that it has some uniqueness properties: the number of ideals occurring in such a decomposition and the set of primes occurring are unique. Some of the primes occurring are minimal in the set of primes occurring. These turn out to be the same as the minimal primes of the original ideal. The primary ideals corresponding to minimal primes of I occurring in an irredundant primary decomposition are unique. The other primes that occur are called *embedded primes.* Note that if Q is an embedded prime and it contains a minimal prime P, then $V(Q) \subset V(P)$, which may help to explain the terminology.

Before proving this statement we consider an example in which primary decomposition is not unique. Let $I = (x^2, xy)R$ in $R = K[x,\, y]$, the polynomial ring in two variables over a field K. It is easy to check that

$$(x^2, xy)R = xR \cap (x^2,\, y)R$$

is an irredundant primary decomposition. Note that xR is prime, and the radical $(x, y)R$ of $(x^2, y)R$ is maximal. Observe also that $\mathrm{Rad}\,(I) = xR$ is the unique minimal prime of I, and that $(x, y)R$, which contains xR, is an embedded prime.

For any scalar $c \in K$, the elements x, $cx + y$ also generate the polynomial ring R and can be used as "new indeterminates," while

$$(x^2, xy)R = \bigl(x^2, x(cx + y)\bigr)R.$$

Thus, we also have that

$$(x^2,\, xy)R = xR \cap (x^2, x + cy)R$$

for all $c \in K$. If $K = \mathbb{C}$, say, is uncountable, this gives uncountably many distinct irredundant primary decompositions of $(x^2,\, xy)R$. (If we had $(x^2, x+cy)R = (x^2, x+c'y)R$ for $c \neq c'$, then the difference $(x + cy) - (x + c'y) = (c - c')y$ would be in both ideals, and so y would be in both ideals, and then $x = (x + cy) - cy$ would be in both ideals as well, a contradiction.)

We now want to start on proving that primary decompositions exist in a Noetherian ring. A proper ideal of a ring R is called *irreducible* if it is not the intersection of two (equivalently, finitely many) strictly larger ideals. We shall show that every proper ideal of a Noetherian ring is the intersection of finitely many irreducible ideals, and then we shall show that every irreducible ideal is primary. This will give a primary decomposition, which, by the comments made above, implies the existence of an irredundant primary decomposition.

Proposition. *Every proper ideal of a Noetherian ring is the intersection of a finite family of irreducible ideals (if the ideal is irreducible, the family has just one element).*

Proof. If this is false, the set of ideals that give counterexamples has a maximal element I. If I is irreducible, we are done. Thus, we must have $I = J \cap J'$, where J and J' are strictly larger ideals. It follows that J and J' are proper (if $J = R$, then $J' = I$ and vice versa). By the maximality of I among counterexamples, each of J and J' is the intersection of a finite family of irreducible ideals. But then I is the intersection of the ideals in the union of these two finite families, a contradiction. $\square$

Lecture of November 13

Note that a linear map from a formal power series ring need not commute with "infinite" addition, which is a formal operation. For example, $\mathbb{Q}[[x]]$ is uncountable and has an uncountable basis over $\mathbb{Q}$, while $\mathbb{Q}[x]$ is countable. Thus, there are uncountably many $\mathbb{Q}$-linear maps of $\mathbb{Q}[[x]]/\mathbb{Q}[x]$ to $\mathbb{Q}$: these can be specified arbitrarily on the uncountable basis. It follows that there are uncountably many composite maps $Q[[x]] \twoheadrightarrow \mathbb{Q}[[x]]/\mathbb{Q}[x] \to \mathbb{Q}$ that kill 1 and every power of x. Therefore, giving the values of a $\mathbb{Q}$-linear map $\mathbb{Q}[[x]] \to \mathbb{Q}$ on the powers of x does not come anywhere near determining the map.

The next result guarantees the existence of irredundant primary decompositions for every proper ideal in every Noetherian ring.

Theorem. *Let R be a Noetherian ring and I an irreducible ideal of R. Then I is primary.*

Proof. Let $ab \in I$, and suppose, to the contrary, that $a \notin I$ and $b \notin \mathrm{Rad}\,(I)$, so that $b^n \notin I$ for all n. Then the sequence of ideals $I :_R b^n$ is obviously non-decreasing. Since R is Noetherian this sequence stabilizes, and so we may choose n so that $R : b^n = R : b^N$ for all $N \geq n$. In particular, we may choose n so that $R : b^n = R : b^{2n}$. Since $b^n \notin I$, we have that $I + Rb^n$ is strictly larger than I, and since $ab \in I$, we have that $ab^n \in I$, so that $I :_R b^n$, which contains a, is strictly larger than I. To complete the proof, we shall show that

$$(I + Rb^n) \cap (I :_R b^n) = I,$$

contradicting the irreducibility of I. Suppose that $u = i + rb^n$ is in the intersection, where $i \in I$ and $r \in R$. Then it multiplies b^n into I, so that $ub^n = ib^n + rb^{2n} \in I$, which implies that $rb^{2n} \in I$ and so $r \in I : b^{2n} = I : b^n$. But then $rb^n \in I$, and so $u = i + rb^n \in I$, as required. □

Putting this together with the results of the previous lecture, we have:

Theorem (existence of irredundant primary decompositions in the Noetherian case). *Every proper ideal I of an arbitrary Noetherian ring has an irredundant primary decomposition.* □

The uniqueness statements that one can make about primary decomposition are independent of the Noetherian hypotheses. We state the uniqueness result, although we postpone the proof briefly.

Theorem (uniqueness statements for primary decomposition). *If a proper ideal $I \subseteq R$ has a primary decomposition, it has an irredundant one, say $I = \mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n$. In this case the prime ideals $P_i = \mathrm{Rad}\,(\mathfrak{A}_i)$ are distinct, by the definition of irredundant, and are uniquely determined. In fact, a prime Q occurs if and only if it has the form* $\mathrm{Rad}\,(I :_R r)$ *for some $r \in R$. The number of terms n is therefore uniquely determined as well. The minimal elements among $P_1, \ldots, P_n$, when intersected, give an irredundant primary decomposition of* $\mathrm{Rad}\,(I)$*, and are the same as the minimal primes of I. The primary ideal $\mathfrak{A}$ in the decomposition corresponding to P, where P is one of the minimal* *primes among* $\{P_1, \ldots, P_n\}$*, is the contraction of* IR_P *to* R*, and so is uniquely determined as well.*

Before proving this, we want to establish:

Lemma. *Let* R *be a ring.*

(a) *If* $I_1, \ldots, I_n$ *are ideals of* R*, then*

$$\mathrm{Rad}\,(I_1 \cap \cdots \cap I_n) = \mathrm{Rad}\,(I_1) \cap \cdots \mathrm{Rad}\,(I_n).$$

(b) *If* $P_1, \ldots, P_k$ *are finitely many mutually incomparable prime ideals, then the* P_i *are the minimal primes of* $P_1 \cap \cdots \cap P_k$*.*

(c) *If* $\mathfrak{A}$ *is primary to* P*, then* $\mathrm{Rad}\,(\mathfrak{A} : r) = R$ *if* $r \in \mathfrak{A}$ *and* $\mathrm{Rad}\,(\mathfrak{A} : r) = P$ *if* $r \notin \mathfrak{A}$*. Moreover, if* $r \notin P$*, then* $\mathfrak{A} : r = \mathfrak{A}$*.*

Proof. (a) has been discussed before: the harder part is that if an element has a power in each of the ideals intersected, the highest power used is in all of them. For (b) we must show that if a prime $Q \supseteq P_1 \cap \cdots \cap P_k$ then it must contain some P_i. If not choose $r_i \in P_i - Q$ for every i. Then the product of the r_i is in all the P_i but not in Q, a contradiction.

For part (c), the only statement that is not immediate is that if an element r is not in $\mathfrak{A}$ then $\mathrm{Rad}\,(\mathfrak{A} : r) = P$. Since $\mathfrak{A} \subseteq \mathfrak{A} : r$, we have that $P = \mathrm{Rad}\,(\mathfrak{A}) \subseteq \mathrm{Rad}\,(\mathfrak{A} : r)$. Therefore, it suffices to show that if $u \in R - P$ and $r \notin \mathfrak{A}$, then $u \notin \mathrm{Rad}\,(\mathfrak{A} : r)$, i.e., $u^t \notin \mathfrak{A} : r$, or $ru^t \notin \mathfrak{A}$. But since $u \notin P$, we have $u^t \notin P$, and so $ru^t \in \mathfrak{A}$ implies $r \in \mathfrak{A}$ since $\mathfrak{A}$ is P-primary, which gives the needed contradiction. □

Proof of the uniqueness statements for primary decomposition. Since $I = \mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n$, from part (a) of the Lemma we have that $\mathrm{Rad}\,(I) = P_1 \cap \cdots \cap P_n$. Suppose that the P_i have been numbered so that $P_1, \ldots, P_k$ are the minimal elements of $\{P_1, \ldots, P_n\}$. Then we also have that $\mathrm{Rad}\,(I) = P_1 \cap \cdots \cap P_k$, and it follows from the Lemma that $P_1, \ldots, P_k$, which are clearly mutually incomparable, are the minimal primes of $\mathrm{Rad}\,(I)$ and, hence, of I. Now suppose that $P = P_i$ is one of these minimal primes, and that we localize at P. Note that for any ideal $J \subseteq R$, we have $J_P \subseteq R_P$ and J_P may be identified with JR_P. Since $I = \mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n$ and localization commutes with finite intersection, we have that $I_P = (\mathfrak{A}_1)_P \cap \cdots (\mathfrak{A}_n)_P$. If $j \neq i$, then $P_j = \mathrm{Rad}\,(\mathfrak{A}_j)$ is not contained in $P = P_i$, and so some element of P_j is in $R - P$. This element has a power in $\mathfrak{A}_j$. Therefore, $(\mathfrak{A}_j)_P = R_P$. We therefore get that $(\mathfrak{A}_i)_P = I_P = IR_P$. Since $\mathfrak{A}_i$ is P-primary, if we expand to R_P and then contract, we get $\mathfrak{A}_i$. Thus, $\mathfrak{A}_i$ is the contraction of $(\mathfrak{A}_i)_P = IR_P$ to R.

Finally, if r is any element of R, then

$$\mathrm{Rad}\,(I : r) = \mathrm{Rad}\,\big((\mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n) : r\big) = \bigcap_i \mathrm{Rad}\,(\mathfrak{A}_i : r) = \bigcap_{i \text{ such that } r \notin \mathfrak{A}_i} P_i$$

by part (c) of the Lemma, where the intersection over the empty set is defined to be R. Therefore we get the intersection of a certain subset S of the P_i, which is the same as the intersection of the primes of S that are minimal elements of S. This intersection can only be prime if it is equal to one of the P_i. To see that we actually do get each of the P_i, notice that the intersection of $\mathfrak{A}_j$ for $j \neq i$ cannot be contained in $\mathfrak{A}_i$, or $\mathfrak{A}_i$ could be omitted and the intersection would not be irredundant. Choose r in the intersection of the $\mathfrak{A}_j$ for $j \neq i$, but not in $\mathfrak{A}_i$. By the calculation above, for this choice of r we have that $\mathrm{Rad}\,(I : r) = P_i$. □

In the Noetherian case the primes that occur as radicals for an irredundant primary decomposition have an alternative characterization. In order to give this characterization, we introduce the set of associated primes of a module M. We do not need finiteness conditions to give the definition.

A prime ideal P of R is called an *associated prime* of the R-module M if equivalently:

(1) There is an element $u \in M$ whose annihilator is P.
(2) There is an injection $R/P \hookrightarrow M$.

These two conditions are equivalent because the submodule of M generated by u is isomorphic with R/P if and only if the annihilator of u is P. Note that the element u with prime annihilator can never be 0, since the annihilator of 0 is the unit ideal.

The set of associated primes of M is denoted $\mathrm{Ass}\,(M)$ and is sometimes called the *assassinator* of M. When M is not Noetherian there may be no primes in $\mathrm{Ass}\,(M)$.

We shall soon show that in the Noetherian case $\mathrm{Ass}\,(M)$ is finite, and non-empty if $M \neq 0$. Moreover, it will turn out that $\mathrm{Ass}\,(R/I)$ is the same as the set of primes that occurs as radicals of primary ideals in an irredundant primary decomposition of I. The primes that occur in a primary decomposition are sometimes called *associated primes of* I, which is ambiguous because I may also be considered as an R-module. But there should be no problem if they are referred to as the *associated primes of I as an ideal.* Then, in the Noetherian case, the associated primes of I as an ideal are the same as the associated primes of the module R/I.

The following facts hold quite generally:

Proposition. *Let R be a ring.*
(a) *If P is prime in R, then* $\mathrm{Ass}\,(R/P) = \{P\}$.
(b) *If* $0 \to M' \to M \to M'' \to 0$ *is exact, then* $\mathrm{Ass}\,(M) \subseteq \mathrm{Ass}\,(M') \cup \mathrm{Ass}\,(M'')$.

Proof. (a) Given any nonzero element of R/P represented by $r \in R/P$, its annihilator is P, precisely because P is prime: if $s \notin P$, rs is not 0 in R/P.

For the second part, we may assume without loss of generality that $M' \subseteq M$ and $M'' = M/M'$. Suppose that $u \in M$ has annihilator P, so that $Ru \cong R/P$. If $Ru \cap M' \neq 0$, some nonzero element v of Ru is in M', and, as observed in the proof of part (a), the annihilator of v is P, so that $P \in \mathrm{Ass}\,(M')$. On the other hand, if $Ru \cap M' = 0$, then $Ru \cong R/P$ embeds into $M/M' = M''$, and so $P \in \mathrm{Ass}\,(M'')$, as required. □

Lecture of November 15

Lemma. *Let M be an R-module and let $u \in M - \{0\}$. Suppose that M or R is Noetherian. Then we may choose $r \in R$ such that $ru \neq 0$ and $P = \mathrm{Ann}_R ru$ is maximal among ideals that are annihilators of nonzero multiples of u. For such a choice of r, P is a prime ideal.*

Proof. Without loss of generality we may replace M by Ru and then R by $R/\mathrm{Ann}_R M$, so that we may assume that M and R are Noetherian. The set of ideals $\{\mathrm{Ann}_R ru : ru \neq 0\}$ is a non-empty family in a Noetherian ring. Therefore, we may choose an element $ru \in Ru - \{0\}$ whose annihilator P is maximal in this set. Suppose that $ab \in P$, but $a \notin P$. Then $aru \neq 0$, and is killed by $P + Rb$, so that we must have $b \in P$, or else P would not be a maximal annihilator. $\square$

By a *finite ascending filtration* of an R-module M we mean a sequence

$$0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_{n-1} \subseteq M_n = M$$

of submodules of n. The filtration is said to have *length n*. The modules M_{i+1}/M_i, $0 \leq i \leq n-1$ are called the *factors of the filtration.*

If N is a submodule of M the problem of giving a finite ascending filtration of M that contains N is equivalent to that of giving such filtrations for N and M/N. Suppose that we have a filtration

$$0 = M_0 \subseteq \cdots \subseteq M_k = N$$

of N. Any filtration of M/N has the form

$$0 \subseteq M_{k+1}/N \subseteq \cdots \subseteq M_n/N$$

where $M_n = M$, since the submodules of M/N correspond bijectively with the submodules of M containing N in such a way that Q/N corresponds to its inverse image Q in M. Note that the 0 occurring initially on the left may be thought of as N/N. Then

$$0 \subseteq M_1 \subseteq \cdots \subseteq M_k \subseteq M_{k+1} \subseteq \cdots \subseteq M_n = M$$

is the required filtration of M. The factors from this filtration are the union of the two sets of factors. The length of this filtration of M is the sum of the lengths of the filtrations of N and M/N.

Proposition. *Let $0 = M_0 \subseteq \cdots \subseteq M_i \subseteq \cdots \subseteq M_n = M$ be a finite ascending filtration of M. Then*

$$\mathrm{Ass}\,(M) \subseteq \bigcup_{i=0}^{n-1} \mathrm{Ass}\,(M_{i+1}/M_i).$$

Proof. This is obvious if there is only one factor, and we may use induction on n. Because of the short exact sequence $0 \to M_{n-1} \to M \to M/M_{n-1} \to 0$ we have that

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

and we may apply the induction hypothesis to the filtration

$$0 \subseteq M_1 \subseteq \cdots \subseteq M_{n-1}$$

of M_{n-1}. $\square$

Theorem. *Every Noetherian module $M \neq 0$ has a finite ascending filtration in which the factors are prime cyclic modules, R/P_i. Therefore* $\mathrm{Ass}\,(M)$ *is finite, and is contained in the set $\{P_1, \ldots, P_n\}$ of primes that occur. Thus,* $\mathrm{Ass}\,(M) = \emptyset$ *if and only if $M = 0$.*

Proof. By Noetherian induction we may assume that the result holds for every quotient of M by a nonzero submodule. (If M is a counterexample, the family of submodules N of M such that M/N is counterexample is non-empty, since it contains 0, and therefore has a maximal element N_1. Work with M/N_1 instead of M.) If $M \neq 0$ we can choose $u \neq 0$ in M and r as in the Lemma so that $P = \mathrm{Ann}_R ru$ is prime. Then $R/P \cong Ru \subseteq M$, so that $P \in \mathrm{Ass}\,(M)$. Let $N = Ru$. By the hypothesis of Noetherian induction, M/N has a filtration of the specified type, and, hence, so does M. □

A cyclic module with prime annihilator P (which will then be isomorphic with R/P) is called a *prime cyclic* module. A finite ascending filtration in which all the factors are prime cyclic modules is called a *prime cyclic filtration.*

Proposition. *Let M be an R-module and W a multiplicative system in R. If R is Noetherian, then* $\mathrm{Ass}\,(W^{-1}M)$ *over $W^{-1}R$ is*

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

More generally, for any R, if $P \in \mathrm{Ass}\,(M)$ and $P \cap W = \emptyset$, then $PW^{-1}R \in \mathrm{Ass}\,(W^{-1}M)$ over $W^{-1}R$. If P is finitely generated, then $PW^{-1}R \in \mathrm{Ass}\,(W^{-1}M)$ over $W^{-1}R$ if and only if $P \in \mathrm{Ass}\,(M)$ and $P \cap W = \emptyset$.

Proof. Since, quite generally, the primes of $W^{-1}R$ have the form $PW^{-1}R$ for a unique choice of prime $P \subseteq R$ disjoint from W, the results in the last two sentences imply the result stated for the Noetherian case. If $P \in \mathrm{Ass}\,(M)$ we have an injection $R/P \hookrightarrow M$, and localizing gives an injection $W^{-1}(R/P) \hookrightarrow W^{-1}M$, where $W^{-1}(R/P) \cong W^{-1}R/PW^{-1}R$. Since $P \cap W = \emptyset$, $PW^{-1}R$ is a prime ideal of $W^{-1}R$, and we are done.

Now suppose that $P = (f_1, \ldots, f_s)R$ is finitely generated and that $PW^{-1}R$ is an element of $\mathrm{Ass}\,(W^{-1}M)$. We can choose a nonzero element of $W^{-1}M$ that has $PW^{-1}R$ as annihilator, and after multiplying an element in the image of W, we may assume this element has the form $u/1$ for $u \in M$. Since each $f_i u/1$ is 0 in $W^{-1}M$, for each i we can choose $w_i \in W$ such that $w_i f_i u = 0$ in M. Let w be the product of the w_i. Then each of $f_1, \ldots, f_n$ kills wu, and so P kills wu. We claim that P is $\mathrm{Ann}_R wu$ which will show that $P \in \mathrm{Ass}\,(M)$, as required. Let $w' \in R - P$. We need only check that $w'wu \neq 0$ in R. But this is clear, since otherwise $u/1$ would be 0 in $W^{-1}M$. □

Corollary. *Let M be a finitely generated module over a Noetherian ring, and suppose that* $\mathrm{Ass}\,(M) = \{P_1, \ldots, P_n\}$. *Then* $\mathrm{Rad}\,(\mathrm{Ann}_R M) = \bigcap_i P_i$. *Thus,* $\mathrm{Rad}\,(\mathrm{Ann}_R M)$ *is the intersection of the minimal elements of* $\mathrm{Ass}\,(M)$*: these are the minimal primes of* $\mathrm{Rad}\,(\mathrm{Ann}_R M)$*, and also the minimal primes of* $\mathrm{Ann}_R(M)$*. Since the support* $\mathrm{Supp}\,(M)$ *of M is $V\big(\mathrm{Ann}_R(M)\big)$, the minimal primes in $P_1, \ldots, P_n$ are also the minimal elements of* $\mathrm{Supp}\,(M)$.

Proof. Let $u_1, \ldots, u_n$ generate M. Let $r \in R$. We know that $u_i/1 = 0$ in $M_r = W^{-1}M$, where $W = \{1, r, r^2, r^3, \cdots\}$ if and only if some power of r kills u_i. Now $M_r = 0$ iff each of $u_i/1$ is 0 in M_r iff some power of r kills each of the u_i iff some power of r kills all the u_i iff some power of r kills M iff $r \in \mathrm{Rad}\,(\mathrm{Ann}(M))$. But $M_r = 0$ iff $\mathrm{Ass}\,(M_r) = \emptyset$, and $\mathrm{Ass}\,(M_r)$ is the set of minimal primes in $\mathrm{Ass}\,(M)$ not containing r, so that $M_r = 0$ iff r is in every prime in $\mathrm{Ass}\,(M)$. □

When the ring R and the module M are Noetherian, the minimal primes of $\mathrm{Ass}\,(M)$ (equivalently, of $\mathrm{Ann}_R(M)$) are called the *minimal primes* of M.

Note that if J is an ideal of R, then $\mathrm{Rad}\,(JW^{-1}R) = \big(\mathrm{Rad}\,(J)\big)W^{-1}R$. Clearly, it suffices to prove $\subseteq$. But if u/w has a power in $JW^{-1}R$, where $u \in R$ and $w_0 \in W$, then $u/1$ does as well, and so $u^n/1 = j/w_1$ for some n, $j \in J$ and $w_1 \in W$. It follows that for some $w_2 \in W$, $w_2(w_1u^n - j) = 0$, from which we have that $wu^n \in J$ with $w = w_1w_2$, and so $(wu)^n \in J$. But then $u \in W^{-1}\mathrm{Rad}\,(J)$. We shall use this fact in analyzing the effect of localization on primary decomposition.

Proposition. *Let I have irredundant primary decomposition*

$$\mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n,$$

and let W be a multiplicative system in R. Let $P_i = \mathrm{Rad}\,(\mathfrak{A}_i)$. Then the intersection of those $\mathfrak{A}_iW^{-1}R$ such that P_i does not meet W is an irredundant primary decomposition of $IW^{-1}R$. In particular, if $\mathfrak{A}$ is primary with radical P, then $\mathfrak{A}W^{-1}R$ is the unit ideal if W meets P and is primary to $PW^{-1}R$ otherwise.

Proof. We establish the final statement first. If W meets P, then some element of W has a power in $\mathfrak{A}$, and so $\mathfrak{A}W^{-1}R$ is the unit ideal. If not, $\mathfrak{A}W^{-1}R$ has radical $PW^{-1}R$, and it suffices to show that if r, $s \in R$, v, $w \in W$, and $(r/v)(s/w) \in \mathfrak{A}W^{-1}R$ then $r/v \in \mathfrak{A}W^{-1}R$ or $s/w \in \mathrm{Rad}\,(\mathfrak{A}W^{-1}R)$. Since $rs/vw \in \mathfrak{A}W^{-1}R$, we find that $w'(rs) \in \mathfrak{A}$ for some $w' \in W$. Since $W \subseteq R - P$, this implies that $rs \in \mathfrak{A}$, so $r \in \mathfrak{A}$ or $s \in \mathrm{Rad}\,(\mathfrak{A})$, from which the desired result follows.

We recall that we have an identification $W^{-1}J \cong JW^{-1}R$ for every ideal J of R and make free use of it. Since localizing commutes with finite intersection, we have that $W^{-1}I = \bigcap_i W^{-1}\mathfrak{A}_i$, and we may omit those terms such that W meets P_i, since for those, $W^{-1}\mathfrak{A}_i$ is the unit ideal. This gives a primary decomposition involving distinct primes. To see that it is irredundant, let P_i be a fixed one of the primes occurring that is disjoint from W. We know that $P_i = \mathrm{Rad}\,(I :_R r)$ for some element of R. Then $W^{-1}P_i = W^{-1}\big(\mathrm{Rad}\,(I :_R r)\big) = \mathrm{Rad}\,\big(W^{-1}(I :_R r)\big) = \mathrm{Rad}\,\big(W^{-1}I :_{W^{-1}R} (r/1)\big)$, which shows, by our earlier criterion for when a prime must occur as the radical of some term in a primary decomposition, that all of the terms are needed. □

The contraction of P^nR_P to R is a P-primary ideal that contains P^n. It is the smallest P-primary ideal containing P^n, and is called the nth *symbolic power* of P, and denote $P^{(n)}$. Note that P is the radical of P^n, and so it is the unique minimal prime of P^n. If P^n has a primary decomposition, the P-primary ideal that is used must be $P^{(n)}$. We also have the description

$$P^{(n)} = \{r \in R : \text{for some } w \in R - P, wr \in P^n\}.$$

In general, $P^{(n)} \supseteq P^n$, and the containment is often strict, even when the ambient ring R is a polynomial ring. The behavior of symbolic powers is very subtle, and has engendered a huge literature.

Example (F. S. Macaulay). Let $R = K[x, y, z]$, the polynomial ring in three variables over a field, and map R by a K-algebra homomorphism onto $K[t^3,\, t^4, t^5] \subseteq K[t]$, where t is another variable, by sending $x \mapsto t^3$, $y \mapsto t^4$ and $z \mapsto t^5$. The kernel P of this homomorphism is a prime ideal of $K[x,\, y,\, z]$. We leave it to the reader to show that $P = (f, g, h)R$ where $f = xz - y^2$, $g = x^3 - yz$, and $h = x^3y - z^2$: these elements are the 2×2 minors of the matrix

$$\begin{pmatrix} x & y & z \\ y & z & x^2 \end{pmatrix}$$

which maps to the rank one matrix

$$\begin{pmatrix} t^3 & t^4 & t^5 \\ t^4 & t^5 & t^6 \end{pmatrix}$$

(the rank is one because the second row is t times the first row). Of course, it is clear that f, g, and h are in the kernel: the problem is to show that they generate the entire kernel. Assuming that we have these generators, it is not difficult to see that there is an element of $P^{(2)}$ that is not in P^2. We assign degrees to the variables in a somewhat non-standard way, so that x , y, and z have degrees 3, 4, and 5, respectively. Then $x^iy^iz^k$ has degree $3i + 4j + 5k$. The elements f, g and h are homogeneous with respect to this grading, of degrees 8, 9, and 10 respectively. Now consider $fh - g^2$. Working mod x, this is $(-y^2)(-z^2) - (-yz)^2 = y^2z^2 - (yz)^2 = 0$. That is, x divides $fh - g^2$, and we can write $fh - g^2 = xq$. Note that $fh - g^2 \neq 0$, since, for example, g^2 has an x^6 term while fh does not. Thus, $q \neq 0$. Now $fh - g^2$ is homogeneous of degree 18, and x is homogeneous of degree 3. It therefore follows without computation that q is homogeneous of degree 15. Since $xq = fh - g^2 \in P^2$, while $x \notin P$, it follows that $q \in P^{(2)}$. But q cannot be in P^2: its degree is 15, while the generators $f^2, g^2, h^2, fg, fh, gh$ of P^2 all have degree 16 or more.

Lecture of November 17

Note that if $J \subseteq I \subseteq R$ then the problem of giving an (irredundant) primary decomposition for I in R is equivalent to the problem of giving an (irredundant) primary decomposition for I/J in R/J. In particular, it is the same problem as giving an (irredundant) primary decomposition of 0 in R/I. The same remark applies to studying whether I is irreducible in R.

Also note that while every prime in $\mathrm{Ass}\,(M)$ must occur as the annihilator of a factor in any finite filtration of M with prime cyclic factors, it may be impossible to give such a filtration of M in which only primes of $\mathrm{Ass}\,(M)$ occur. If M is torsion-free over a domain R, then the annihilator of any nonzero element is (0) in R: thus, $\mathrm{Ass}\,(M) = \{(0)\}$. Consider any torsion-free module over R that is not free. A prime cyclic filtration cannot consist only of factors that are $\cong R = R/(0)$. (If one has a finite filtration in which all the factors are R, there is a surjection of M onto the last factor, $M \twoheadrightarrow R$, which will split, so that $M = M_0 \oplus_R R$, where M_0 has such a filtration with one fewer copy of R. By induction, induction on the number of factors, M is R-free.) One can start with several such factors, but eventually one will have a quotient which is a torsion module. For example, let $M = (x, y)R$ in the polynomial ring $K[x, y]$. M needs two generators, and after killing any copy of $R = Rf$ where f is an element of M one has a torsion module and other primes are needed for the filtration. E.g., if one kills xR, the quotient is $\cong y(R/xR) \cong R/xR$, and one has a prime cyclic filtration that involves (0) and xR.

If P is any prime occurring in a prime cyclic filtration of M, then R/P is a homomorphic image of a submodule of M, and therefore if I kills M, then I kills R/P, so that $I \subseteq P$. Thus, $\mathrm{Ann}_R M \subseteq P$, and this implies that P contains a minimal prime of M. Thus, even the "extraneous" primes occurring in a prime cyclic filtration of M (by which we mean the primes occurring that are not associated primes of M) must contain a minimal prime of M.

Examples. (1) Let $R = K[X_1, X_2, X_3, \ldots]/J$ where $J = (X_t^{t+1} : t \geq 1)$. The ideal m of R generated by the images x_t of the X_t is maximal: $R/m \cong K$. Since every x_t is nilpotent, this maximal ideal is also the unique minimal prime of R. Thus, $\mathrm{Spec}\,(R) = \{m\}$. We claim that $\mathrm{Ass}\, R = \emptyset$. Since m is the only prime ideal of R, this amounts to the assertion that there is no element of $R - \{0\}$ that is killed by m. Note that m is spanned over K by the monomials $x_1^{k_1} \cdots x_n^{k_n}$, with n varying, such that for all t, $0 \leq k_t \leq t$. Suppose that $f \in R - \{0\}$, and that x_N does not occur in f. Then $x_N f \neq 0$, which establishes our claim. The theory that we have already developed shows that this does not happen if R is Noetherian.

(2) Let $R = K[y, x_1, x_2, x_3, \ldots]$ be a polynomial ring in infinitely many variables. Let $P = (x_i : i \geq 1)$ and let $M = R/J$, where $J = (y^i x_i : i \geq 1)R$. Let $W = \{y^t : t \in \mathbb{N}\}$. Then $PW^{-1}R \in \mathrm{Ass}\,(W^{-1}M)$: in fact, $W^{-1}M \cong W^{-1}R/PW^{-1}R$. But no element of $R - J$ is multiplied into J by P, so that $P \notin \mathrm{Ass}\,(M)$. This is another sort of behavior that cannot occur in the Noetherian case.

(3) Let $R = K[x,\,y,\,u,\,v,\,z,\,w]/(xy+uv+zw)$. It is easy to see that $xy+uv+zw$ is irreducible, and this ring can be shown to be UFD. Let $P = (x,\,y,\,u,\,v,\,w)R$. This is ideal is prime, with quotient ring $K[z]$, and of course $Q = P + zR$ is a maximal ideal with quotient ring K. Now, $P \subseteq Q$ and $P^2 \subseteq Q^2$, but $P^{(2)}$ is not contained in $Q^{(2)}$. Because Q is maximal, Q^2 is Q-primary and so $Q^2 = Q^{(2)}$. But since $zw = -xy - uv \in P^2$ while $z \notin P$, we have that $w \in P^{(2)}$. In a polynomial or power series ring over a field or a PID, it is true that if $P \subseteq Q$ are primes then $P^{(n)} \subseteq Q^{(n)}$ for all n: but this is a difficult theorem due to Nagata and Zariski independently. Cf. [M. Nagata, *Local Rings*, Interscience, New York, 1962], p. 143, for Nagata's proof. In the massive breakthrough paper [H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero*, Annals of Math. **79** (1964), pp. 205–326], Hironaka gives Zariski's proof: see Theorem 1. A more elementary argument is used to prove this fact in [M. Hochster, *Symbolic powers in Noetherian domains*, Illinois Math. J. (1971), pp. 9–27]: this paper has a proof for formal power series rings over a field that uses the Weierstrass preparation theorem. The case of formal power series over a field implies the case of polynomial rings over a field. Another subtle result on behavior of symbolic powers in polynomial and power series rings over fields is that if P has height h, then $P^{(n)} \subseteq P^{nh}$ for all n. This was proved over fields of characteristic 0 in [L. Ein, R. Lazarsfeld, and K. E. Smith, *Uniform bounds and symbolic powers on smooth varieties*, Inventiones Math. **144** (2001), pp. 241–252] and over arbitrary fields in [M. Hochster and C. Huneke, *Comparison of symbolic and ordinary powers of ideals*, Inventiones Math. **147** (2002),, pp. 349–369].

Note that if $Q_1,\,\ldots\,,Q_n$ are submodules of M, then the kernel of the map

$$M \to M/Q_1 \oplus_R \cdots \oplus_R M/Q_n$$

such that

$$u \mapsto (u+Q_1) \oplus \cdots \oplus (u+Q_n)$$

is precisely $Q_1 \cap \cdots \cap Q_n$, yielding an injection

$$M/(Q_1 \cap \cdots \cap Q_n) \hookrightarrow M/Q_1 \oplus_R \cdots \oplus_R M/Q_n.$$

A finite direct sum of modules $W_1 \oplus_R \oplus_R \cdots \oplus_R W_n$ has a filtration

$$0 \subseteq W_1 \subseteq W_1 \oplus_R W_2 \subseteq W_1 \oplus_R W_2 \oplus_R W_3 \subseteq \cdots \subseteq W_1 \oplus_R W_2 \oplus_R \cdots \oplus_R W_n$$

with factors W_i, and so

$$\mathrm{Ass}\,(W_1 \oplus_R \cdots \oplus_R W_n) \subseteq \bigcup_i \mathrm{Ass}\,(W_i).$$

Thus, when $Q_1,\,\ldots\,,Q_n$ are submodules of M with intersection N, we have that

$$\mathrm{Ass}\,(M/N) \subseteq \mathrm{Ass}\,\big(\oplus_i(M/Q_i)\big) \subseteq \bigcup_i \mathrm{Ass}\,(M/Q_i).$$

Theorem. *Let R be a Noetherian ring, M a finitely generated R-module, and let $I \subseteq R$ be an ideal.*

(a) *An element $r \in R$ is a zerodivisor on M (i.e., $ru = 0$ for some $u \in M - \{0\}$) if and only if it belongs to a prime $P \in \mathrm{Ass}\,(M)$. In other words, the set of zerodivisors on M in R is the same as the union of the associated prime ideals of M.*

(b) *I is primary if and only if $\mathrm{Ass}\,(R/I)$ contains just one element P, in which case I is primary to P.*

(c) *The associated primes of I as an ideal are the elements of $\mathrm{Ass}\,(R/I)$.*

Proof. For part (a), note that if $u \in P \in \mathrm{Ass}\,(M)$, then $P = Ann_R u$, $u \neq 0$, and so $ru = 0$ with $u \neq 0$. On the other hand, if $ru = 0$ with $u \neq 0$ then u has a multiple $r'u$ that is not 0 with prime annihilator P. Clearly $r \in P \in \mathrm{Ass}\,(M)$.

For parts (b) and (c), first observe that if I is primary to P then the zerodivisors on I are precisely the elements of P/I (which are nilpotent in R/I: by the definition of primary ideal, the elements of $R - P$ are not zerodivisors on the module R/I). Thus, $\mathrm{Ass}\,(R/I) = P$. This proves the "only if" part of (b).

Now suppose that $I = \mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n$ is an irredundant primary decomposition of I and that $\mathrm{Rad}\,(\mathfrak{A}_i) = P_i$. Then the remarks preceding the statement of the theorem show that $\mathrm{Ass}\,(R/I) \subseteq \bigcup_i \mathrm{Ass}\,(R/\mathfrak{A}_i) = \{P_1, \dots, P_n\}$ by the preceding paragraph. Now fix i and choose r in the intersection of the $\mathfrak{A}_j$ for $j \neq i$ but not in $\mathfrak{A}_i$, so that $\mathrm{Rad}\,(I :_R r) = P_i$. Let $N = (I + rR)/I \cong \overline{r}(R/I)$, where $\overline{r}$ denotes the class of r in R/I. Then $\mathrm{Ass}\,(N) \subseteq \mathrm{Ass}\,(M)$. But the annihilator of the cyclic module N is $I :_R r$, whose radical is P_i. Since P_i is a minimal prime of $I :_R rR$, $P_i \in \mathrm{Ass}\,\big(R/(I :_R rR)\big) = \mathrm{Ass}\,(N) \subseteq \mathrm{Ass}\,(M)$. This shows that every associated prime of I as an ideal is in $\mathrm{Ass}\,(R/I)$.

Finally, if $\mathrm{Ass}\,(R/I) = P$, then there is only one term in the primary decomposition of R, and so I is primary with $\mathrm{Rad}\,(I) = P$, which proves the "if" part of (b). $\square$

Next, note that if a finitely generated ideal $I = (f_1, \dots, f_h)$ is contained in the radical of J, then $I^N \subseteq J$ for sufficiently large N. Each f_t has a power in J: say the $f_t^{a_t} \in J$. Take $N \geq a_1 + \cdots a_h - h + 1$. Then I^N is generated by the monomials of degree N in the f_t, and, in any such monomial, the exponent on some f_t must be at least a_t, or else the sum of the exponents is at most $(a_1 - 1) + \cdots + (a_h - 1) = a_1 + \cdots + a_h - h$.

As an application of primary decomposition, we prove the following beautiful result.

Theorem. *Let (R, m) be a local ring, i.e., a Noetherian ring with a unique maximal ideal m. Then $\bigcap_n m^n = (0)$.*

Proof. Let $J = \bigcap_n m^n$. Let $mJ = \mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n$ be a primary decomposition for mJ. We shall show that $J \subseteq \mathfrak{A}_i$ for every I. But this proves that $J \subseteq mJ$, so that $J = mJ$. But then $J = (0)$ by Nakayama's Lemma.

To prove that $J \subseteq \mathfrak{A}_i$ we consider two cases. First suppose that $P_i = \mathrm{Rad}\,(\mathfrak{A}_i)$ is different from m. Choose $x \in m - P_i$. Then $xJ \subseteq mJ \subseteq \mathfrak{A}_i$, but x is not in $\mathrm{Rad}\,(\mathfrak{A}_i)$. This implies that $J \subseteq \mathfrak{A}_i$. The remaining case is where $\mathfrak{A}_i$ is primary to m. But then each generator of m has a power in $\mathfrak{A}_i$, and since m is finitely generated, $m^N \subseteq \mathfrak{A}_i$ for all $N \gg 0$. But $J \subseteq m^N$ for all N, and so $J \subseteq \mathfrak{A}_i$ in this case as well. $\square$

We want to show that $\bigcap_n m^n M = 0$ for any finitely generated R-module M as well. There are at least three methods of doing this: one is extend the theory of primary decomposition to modules, and we shall do this in these notes shortly. A second method involves a result called the Artin-Rees theorem, and we shall also give that proof eventually.

The third method is to deduce the result for all modules from the ring case by a trick: Nagata's idealization trick (or method). The key point is that if R is any ring and M is any R-module, then $R \oplus_R M$ becomes a commutative ring with identity if we define multiplication by the rule

$$(r \oplus u)(s \oplus v) = rs \oplus (rv + su).$$

This ring is an R-algebra. M is an ideal in which the square of every element is 0, and the product of any two elements of M is 0. Killing M gives R back. Every prime ideal of $R \oplus_R M$ has the form $P \oplus_R M$ for some prime P of R: the same is true for maximal ideals. If R is quasilocal, then $R \oplus_R M$ is quasilocal, and if R is local and M is finitely generated as an R-module then $R \oplus M$ is module-finite over R, hence, Noetherian, and is a local ring.

Theorem. *If (R, m) is local and M is a finitely generated R-module, then $\bigcap_n m^n M = 0$.*

Proof. Consider the local ring $R \oplus_R M$ described just above. Its unique maximal ideal is $m \oplus M$, and

$$(m \oplus M)^{n+1} = m^{n+1} \oplus m^n M.$$

Any element of $\bigcap m^n M$ is therefore in every power of the maximal ideal of $R \oplus_R M$, and is therefore 0. □

Theorem. *Let R be a Noetherian ring, M a finitely generated R-module, and I an ideal of R. Then $u \in \bigcap_n I^n M$ if and only if there exists an element $i \in I$ such that $u = iu$.*

Proof. The "if" part is trivial, for if $u = iu$ then $u = iu = i(iu) = i^2 u$, and by a straightforward induction, $u = i^n u \in I^n M$ for all n.

For the other direction, suppose that $I + \mathrm{Ann}_R u$ is a proper ideal of R, let m be a maximal ideal of R containing it. Then $u/1$ is nonzero in M_m, and $I_m \subseteq \mathcal{M} = mR_m$, the maximal ideal of R_m. But then $u \in I^n M$ for all n implies that $u/1 \in \mathcal{M}^n M$ for all n, and this is a contradiction. Thus, we can choose $i \in I$ and $z \in \mathrm{Ann}_R u$ such that $i + z = 1$. But then $iu = iu + zu = (i + z)u = 1u = u$. □

Notice that this is a global result obtained by reduction to the local case. Our next main objectives are first, to classify all rings with DCC, and second to make use of the theory of Noetherian rings and modules that we have developed to analyze dimension theory in an arbitrary Noetherian ring.

However, before leaving the topic of primary decomposition, we extend the theory to an arbitrary submodule N of a module M over a Noetherian ring R. We define $Q \subseteq M$ to be *primary* to a prime ideal P of R if $\mathrm{Ass}\,(M/Q) = \{P\}$. We shall say that M/Q is *P-coprimary.* A proper submodule $Q \subseteq M$ is called *irreducible* in M if it is not the intersection of two (equivalently, finitely many) strictly larger submodules of M.

Theorem (primary decomposition for modules). *Let R be Noetherian, let M be a finitely generated R-module, and let N, $Q \subseteq M$ be finitely generated R-submodules.*

(a) *Q is primary if and only if $\mathrm{Rad}\,\big(\mathrm{Ann}_R(M/Q)\big)$ is a prime ideal P (whose elements then act nilpotently on M/Q) and the elements of $R-P$ are not zerodivisors on M/Q.*

(b) *Every irreducible submodule of Q of M is primary.*

(c) *Every proper submodule N of M is a finite intersection of of irreducible submodules.*

(d) *If $Q_1, \,\ldots\,, Q_n$ are primary to P, so is their intersection.*

(e) *Every proper submodule N of M is an irredundant intersection of primary submodules, where irredundant means that the primes to which they are primary are mutually distinct, and that no term can be omitted. The set of primes that occur are the associated primes of M/N, and so are unique, the minimal primes among them are the minimal primes of $\mathrm{Ann}(M/N)$, and if P is one of these minimal primes, then the primary submodule to P that occurs is unique, and is the submodule of all elements of M whose images in M_P are in N_P.*

(f) *If N has irredundant primary decomposition $Q_1 \cap \cdots \cap Q_n$ where Q_i is primary to P_i, $1 \leq i \leq n$, and W is a multiplicative system in R, then the $W^{-1}R$-module $W^{-1}M$ has an irredundant primary decomposition as the intersection of the $W^{-1}Q_i$ for those i such that P_i does not meet W.*

Proof. (a) $\mathrm{Ass}\,(M/Q) = \{P\}$ implies that $\mathrm{Rad}\,(\mathrm{Ann}_R m) = P$, and this in turn implies that elements of P act nilpotently on M. Since the union of the associated primes is the set of zerodivisors on M/Q, we also have that an element of $R-P$ is not a zerodivisor on M/Q. Conversely, if $\mathrm{Rad}\,(\mathrm{Ann}_R M) = P$ then P consists of zerodivisors on M, and no prime strictly smaller than P can be in $\mathrm{Ass}\,(M)$, while no prime strictly larger than P can be in $\mathrm{Ass}\,(M)$, since elements of $R-P$ are not zerodivisors on M/Q.

For part (b), let Q be irreducible. We replace M by M/Q. We want to show that $\mathrm{Ass}\,(M)$ contains just one element. Suppose that there are two elements P and P': then R/P and R/P' both embed into M, and their images can meet only in 0, because any nonzero element in either has annihilator P and also annihilator P'. Thus, 0 is the intersection of two larger submodules, a contradiction.

By Noetherian induction, if some proper submodule is not the intersection of a finite family of strictly larger submodules, there is a maximal such submodule. Then either it is irreducible, or it is the intersection of two larger proper submodules, each of which is a finite intersection of irreducible submodules. This proves (c).

For part (d), if $Q_1, \,\ldots\,, Q_n$ are all primary P then $M/(\bigcap_i Q_i)$ embeds in $\bigoplus_i M/Q_i$ and it follows that $\mathrm{Ass}\,\big(M/(\bigcap_i Q_i)\big) = \{P\}$, as required.

The existence of an irredundant primary decomposition for a submodule is now obvious. If $N = Q_1 \cap \cdots \cap Q_n$ is an irredundant primary decomposition, then we have an embedding of M/N into the direct sum of the M/Q_i. Therefore, $\mathrm{Ass}\,(M) \subseteq \{P_1, \,\ldots\,, P_n\}$, where $\mathrm{Ass}\,(M/Q_i) = \{P_i\}$. Now, for fixed i, we know that the intersection of the Q_j for $j \neq i$ is not contained in Q_i, or Q_i would be redundant. Pick $u \in Q_j$ for $j \neq i$ such that $u \notin Q_i$. Then $u \notin N$. Consider $R\overline{u} \in M/N$. We claim that P_i is a minimal prime of this module: in fact the support of this module is $V(P_i)$. To see this, note that for a prime P, $(R\overline{u})_P \neq 0$

iff $u/1 \notin N_P$ iff $u/1 \notin (Q_t)_P$ for some t iff $u/1 \notin (Q_i)_P$ (since u was chosen in all the other Q_j). Since M/Q_i is primary to P_i, it has support $V(P_i)$, and this shows $P \supseteq P_i$. But it is clear that localizing at $P \supseteq P_i$ will not kill u mod Q_i, since elements of $R - P \subseteq R - P_i$ are not zerodivisors on M/Q_i. Thus, P_i is the unique minimal prime of $R\overline{u} \subseteq M/N$. But then $P_i \in \mathrm{Ass}\,(R\overline{u}) \subseteq \mathrm{Ass}\,(M/N)$, as required. This completes the proof of part (e), except for the very last statement, which we shall give in the next paragraph when we prove part (f).

(f) Since localization commutes with finite intersection, we have that $W^{-1}N$ is the intersection of all the $W^{-1}Q_i$. If W meets P_i, then since $P_i = \mathrm{Ass}\,(M/Q_i)$, we have that $P_i = \mathrm{Rad}\,\bigl(\mathrm{Ann}_R(M/Q_i)\bigr)$, and then some element of W will have a power that annihilates M/Q_i, and if follows that $W^{-1}(M/Q_i) = 0$ for such i, i.e., that $W^{-1}Q_i = W^{-1}M$. Evidently, these terms may be omitted, and we know that $W^{-1}N$ is the intersection of the others, which are primary to various distinct primes $P_iW^{-1}R$ or $W^{-1}R$. All of the terms are needed, because we know that these primes are precisely the ones in $\mathrm{Ass}\,(W^{-1}M)$. This also shows that if P_i is minimal among the $\{P_1, \ldots, P_n\}$, then $N_{P_i} = (Q_i)_{P_i}$, and since elements of $R - P_i$ are not zerodivisors on M/Q_i, we find that $u \in M$ is in Q_i if and only if $u/1 \in (Q_i)_{P_i} = N_{P_i}$, as required. □

It is worth noting that the problem of giving an (irredundant) primary decomposition of $N \subseteq M$ (and also the issue of whether N is an irreducible submodule of M) is unaffected by replacing the pair $N \subseteq M$ by the pair $N/N_0 \subseteq M/N_0$, where N_0 is a submodule of N. In particular, one may as well study the problem for $0 \subseteq M/N$.

Lecture of November 20

We need to understand one more case where the ring is not necessarily Noetherian, but one knows nonetheless that there is a primary decomposition.

Theorem. *Let R be any ring and I an ideal such that $V(I)$ is a finite set of ideals, all of which are maximal. Then I has a primary decomposition, $I = \mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n$, which is unique except for the order of the terms. In this case I is also the product $\mathfrak{A}_1 \cdots \mathfrak{A}_n$, and R/I is isomorphic with the product of the rings $R/\mathfrak{A}_i$.*

Proof. It is an equivalent problem to find a primary decomposition for (0) in the ring R/I. Therefore we may assume that R is a ring such that every prime ideal is maximal, and such that there are only finitely many maximal ideals, say $m_1, \ldots, m_n$. We seek a primary decomposition for the ideal (0). Let $\mathfrak{A}_n$ be the contraction of the (0) ideal from R_{m_i}, i.e., the set of elements of R which map to 0 in R_{m_i}: these are the elements that are killed by an element not in m_i. Then $\mathfrak{A}_i$ is m_i-primary, since it is the contraction of an $m_iR_{m_i}$-primary ideal, the zero ideal, of R_{m_i}. Moreover, $0 = \mathfrak{A}_1 \cap \cdots \cap \mathfrak{A}_n$, for any element of this intersection vanishes no matter at which prime ideal of R we localize: the m_i constitute all of the prime ideals of R. This gives a primary decomposition of (0), and since all of the primes occurring as radicals are maximal, they are all minimal, and so the primary decomposition is unique.

Since the $\mathfrak{A}_i$ have radicals that are mutually distinct maximal ideals, they are pairwise comaximal: if $i \neq j$, $\mathrm{Rad}\,(\mathfrak{A}_i + \mathfrak{A}_j) \supseteq \mathrm{Rad}\,(\mathfrak{A}_i) + \mathrm{Rad}\,(\mathfrak{A}_j) \supseteq m_i + m_j = R$, and the remaining statements now follow from the Chinese Remainder Theorem. □

A nonzero module over a ring R is called *simple* if, equivalently, (1) it has no nonzero proper submodule or (2) it is isomorphic with R/m for some maximal ideal m. Note that a module satisfying (1) must be generated by any nonzero element, and is therefore cyclic and of the form R/I for some proper ideal I. The statement that there are no proper submodules except (0) is the equivalent to the statement that every nonzero ideal of R/I is the unit ideal, which forces R/I to be a field.

A module is said to have *finite length* if it has a filtration in which every factor is simple. Recall that a refinement of a filtration is chain of submodules that contains the original chain: what happens is that between pairs of modules $M_i \subseteq M_{i+1}$ in the original chain, additional modules $M'_{i,t}$ may be inserted, so that one has $M_i \subseteq M'_{i,1} \subseteq \cdots \subseteq M'_{i,k} \subseteq M_{i+1}$. In any refinement of a filtration with simple factors, every factor is 0 or simple. If M has finite length, by the Jordan-Hölder theorem, any finite filtration can be refined to one in which every factor is simple or 0. In any two filtrations such that all factors are 0 or simple, the simple factors are the same in some order, counting multiplicities, because, again by the Jordan-Hölder theorem, the two filtrations have isomorphic refinements. If M has finite length the *length* $\ell(M)$ is defined to be the number of simple factors in any finite filtration such that all factors are simple or 0. If $0 \to M' \to M \to M'' \to 0$ is a short exact sequence of modules, then M has finite length if and only if both M' and M'' have finite length, and then $\ell(M) = \ell(M') + \ell(M'')$. (If M' and $M/M' \cong M''$ have finite filtrations with simple factors, these can be conjoined to give such a filtration for M, and the lengths add. On the other hand, if M has such a filtration, the filtration $0 \subseteq M' \subseteq M$ can be refined to a filtration where all factors are simple or 0, and it follows that both M' and $M'' \cong M/M'$ have finite length.)

If M has finite length and $M_1 \subseteq M_2$ are submodules, then $\ell(M_2) = \ell(M_1) + \ell(M_2/M_1)$, and so M_1 and M_2 are equal if and only if they have the same length. Thus, any chain of distinct submodules of M has length at most equal to $\ell(M)$, and a finite length module has both ACC and DCC. If M is killed by a maximal ideal m, then it has finite length if and only if it is a finite-dimensional vector space over R/m, in which case its length is the same as its dimension over R/m.

For a vector space W over a field K, we note that the conditions of having ACC, DCC, and finite length are all equivalent: they hold precisely when K has finite dimension, which we know is equal to its length over K. If W has finite length, we have already seen that ACC and DCC hold. On the other hand, if W contains an infinite set of linearly independent vectors $v_1,\ v_2,\ v_3, \ldots$ then there is an infinite strictly ascending chain of which the nth term is the span of $v_1, \ldots, v_n$, and an infinite strictly descending chain of which the nth term is the span of $v_n, v_{n+1}, v_{n+2}, \cdots$.

Over a principal ideal domain R, the length of R/f, where $f \neq 0$, is the same as the number n of irreducible factors in a factorization $f = f_1 \cdots f_n$ of f into irreducible factors (which are allowed to be repeated). Thus, $\ell(\mathbb{Z}/60\mathbb{Z}) = 4$, since $60 = 2 \cdot 2 \cdot 3 \cdot 5$, and $\ell\big(K[x]/(x^3 - x)\big) = 3$, since $x^3 - x = (x-1)x(x+1)$.

Note that $\mathbb{C}$ has length 1 as a $\mathbb{C}$-module and length 2 as an $\mathbb{R}$-module.

A module M has finite length iff M is Noetherian and $\mathrm{Ass}\,(M)$ consists entirely of maximal ideals. This is clear because if M has a prime cyclic filtration by modules R/m with m maximal, it is immediate that M is Noetherian and that $\mathrm{Ass}\,(M)$ is contained in the set of maximal ideals occurring. Conversely, if M is Noetherian and $\mathrm{Ass}\,(M)$ consists entirely of maximal ideals, then $\mathrm{Ann}_R M$ is the intersection of these. Any prime occurring in a finite prime cyclic filtration of M must contain $\mathrm{Ann}_R M$ and therefore must be in $\mathrm{Ass}\,(M)$. It follows that a finite prime cyclic filtration has only factors of the form R/m, where m is maximal.

A local ring (R, m, K) of dimension 0 has finite length as a module over itself: if m is the maximal ideal, then every element of m is nilpotent. Since m is finitely generated, some power of m is 0. Say that $m^n = 0$. Then $0 = m^n \subseteq m^{n-1} \subseteq \cdots \subseteq m^2 \subseteq m \subseteq R$ is a filtration of R, and each factor has the form m^i/m^{i+1}, is a vector space over R/m, and is finite dimensional, since m^i is finitely generated. The length of R is the same as the sum of these dimensions.

Theorem. *The following conditions on a ring R are equivalent.*
(1) *R is Noetherian of Krull dimension 0.*
(2) *R is a finite product of local rings of Krull dimension 0.*
(3) *R has finite length as a module over itself.*
(4) *R has DCC.*

Proof. To see that (1) $\Rightarrow$ (2), note that when (1) holds, all prime ideals of R are minimal as well as maximal: thus, R has only finitely many maximal ideals $m_1, \ldots, m_n$, and we may use the preceding theorem to write R as the product of rings $R/\mathfrak{A}_i$ where $\mathfrak{A}_i$ is primary to m_i.

That (2) $\Rightarrow$ (3) is obvious, since we have already seen that a local ring of dimension 0 has finite length as a module over itself, and (3) $\Rightarrow$ (4) has already been noted.

It remains only to prove that (4) $\Rightarrow$ (1). Since R has DCC, so does every quotient. Let P be prime in R. If $A = R/P$ is not a field, we may choose $a \in A$ that is not 0 and not a unit. The sequence of ideals a^nA must stabilize. But then $a^n \in a^{n+1}A$ for some n, say $a^n = a^{n+1}b$. But since $a \neq 0$ and A is a domain, we get $1 = ab$, a contradiction. Thus, every prime ideal of A is maximal.

If there were infinitely many maximal ideals m_1, m_2, m_3, $\ldots$ the chain

$$m_1 \supseteq m_1 \cap m_2 \supseteq m_1 \cap m_2 \cap m_3 \supseteq \cdots$$

would have to stabilize, yielding $m_{n+1} \supseteq m_1 \cap \cdots \cap m_n$ for some sufficiently large n. But then $m_{n+1} \supseteq m_i$ for some $i \leq n$, a contradiction.

Therefore R has Krull dimension 0, and has only finitely many maximal ideals. The preceding theorem on primary decomposition in this situation enables us to write R as a finite product of quasilocal rings of dimension 0. We have therefore reduced to studying the case where R is quasilocal, with a unique prime ideal (which is necessarily its maximal ideal).

Now suppose that (R, m, K) is quasilocal of Krull dimension 0, and has DCC. Then the sequence of ideals

$$m \supseteq m^2 \supseteq m^3 \supseteq \cdots \supseteq m^n \supseteq \cdots$$

is eventually stable, and we may assume that n has been chosen such that $m^n = m^{n+1}$. Each of the vector spaces m^i/m^{i+1} has DCC, and therefore each is finite-dimensional over K. We want to show that $m^n = 0$. Assume otherwise. Consider the family of ideals $\{I \subseteq m : Im^n \neq 0\}$. Then m is in this family. Therefore, the family has a minimal element $J \subseteq m$. Clearly, we can choose $x \in J$ such that $xm^n \neq 0$, and so $Rx \subseteq J$ is in the family. Therefore, $J = Rx$. Now, $xm(m^n) = xm^{n+1} = xm^n \neq 0$, and so $xm \subseteq Rx$ is also in the family, and we get that $Rx = mx = mRx$. By Nakayama's lemma, $Rx = 0$, a contradiction. Thus, $m^n = 0$ for some n, and then R has a finite filtration

$$0 = m^n \subseteq m^{n-1} \subseteq \cdots \subseteq m^2 \subseteq m \subseteq R$$

whose factors are finite-dimensional vector spaces. Thus, R has finite length as a module over itself and, therefore, R is Noetherian. □

A word of caution: although a ring with DCC has finite length, a module with DCC over a Noetherian ring need not have finite length. Let $V = \mathbb{Z}_P$, where P is the prime ideal generated by the prime integer p. Then $\mathbb{Q}/V$ has DCC as a V-module, but not finite length. What happens is that every proper submodule of $\mathbb{Q}/V$ has finite length, but $\mathbb{Q}/V$ itself does not. It is also true that if V is any discrete valuation ring with fraction field $\mathcal{F}$, then $\mathcal{F}/V$ has DCC but not finite length as a V-module.

Lecture of November 22

We note one more fact about the behavior of primary ideals.

Proposition. *Let I be primary to P in R. And let S be a polynomial ring over R (the number of variables may be infinite). Then IS is primary to PS, which is prime, in S.*

Proof. Let x denote all the variables being adjoined. Then $PR[x]$ is the kernel of the obvious surjection $R[x] \twoheadrightarrow (R/P)[x]$ that replaces every coefficient of a given polynomial with its image in R/P. Thus, $PR[x]$ is prime, and is certainly the radical of $IR[x]$. We replace R by R/I and P by P/I and henceforth assume that $I = (0)$. Thus, P consists of nilpotents, and elements of $R - P$ are not zerodivisors in R.

Now suppose that f is a polynomial with a coefficient that is not in P. It suffices to see that f does not kill any nonzero polynomial in $R[x]$. Suppose that $fg = 0$ where $g \neq 0$. Consider the subring R_0 of R generated over the image of $\mathbb{Z}$ in R by the coefficients of f and g. Let $P_0 = P \cap R_0$, which is the same as the (prime) ideal of all nilpotents in R_0. Notice that $f,\, g \in R_0[x]$, and that f has a coefficient not in P_0. Elements of $R_0 - P_0$ are in $R - P$ and therefore are not zerodivisors even in R. Of course, we still have that $fg = 0$. Thus, we may replace R and P by R_0 and P_0, and we have reduced to the Noetherian case. We change notation and write R for R_0 and P for P_0.

We may also omit adjoining any indeterminates not occurring in f or g, and we may therefore assume that the number of indeterminates is finite. By induction on the numberof indeterminates, we may assume that there is only one indeterminate.

Elements of $R - P$ are clearly not zerodivisors in $R[x]$ as well. We may therefore replace R by R_P, and assume that R is local with nilpotent maximal ideal P, f has a coefficient not in P, hence, a unit, and $g \in R[x] - \{0\}$. We want to show that $fg = 0$ leads to a contradiction. Now, $P^N = 0$ in R for sufficiently large N. We can replace g by a nonzero multiple all of whose coefficients are killed by P: if all coefficients of g are killed by P we are done: if not, multiply by some element of P that does not kill g. This procedure can be repeated at most N times, since $P^N = 0$. Thus, we may assume without loss of generality that every coefficient of g is killed by P.

All terms of f whose coefficients are in P kill g. Therefore, if $fg = 0$ and we omit all terms from f with coefficients in P, we still have that $fg = 0$. Thus, we may assume that the highest degree term in f has a coefficient that is a unit. Multiplying this term by the highest degree nonzero term in g produces a nonzero term in the product that cannot be canceled. □

Our next objective is to study dimension theory in Noetherian rings. There was initially amazement that the results that follow hold in an arbitrary Noetherian ring.

Theorem (Krull's principal ideal theorem). *Let R be a Noetherian ring, $x \in R$, and P a minimal prime of xR. Then the height of $P \leq 1$.*

Before giving the proof, we want to state a consequence that appears much more general. The following result is also frequently referred to as *Krull's principal ideal theorem*, even though no principal ideals are present. But the heart of the proof is the case $n = 1$, which is the principal ideal theorem. This result is sometimes called *Krull's height theorem.* It follows by induction from the principal ideal theorem, although the induction is not quite straightforward, and the converse also needs a result on prime avoidance.

Theorem (Krull's principal ideal theorem, strong version, alias Krull's height theorem). *Let R be a Noetherian ring and P a minimal prime ideal of an ideal generated by n elements. Then the height of P is at most n. Conversely, if P has height n then it is a minimal prime of an ideal generated by n elements. That is, the height of a prime P is the same as the least number of generators of an ideal $I \subseteq P$ of which P is a minimal prime. In particular, the height of every prime ideal P is at most the number of generators of P, and is therefore finite. For every local ring R, the Krull dimension of R is finite.*

Proof of the first version of the principal ideal theorem. If we have a counterexample, we still have a counterexample after we localize at P. Therefore we may assume that $(R,\, P)$ is local. Suppose that there is a chain of length two or more. Then there is a strict chain

$$P \supset Q \supset Q_0$$

in R. We may replace R, P, Q, Q_0 by R/Q_0, P/Q_0, Q/Q_0, (0). We may therefore assume that $(R,\, P)$ is a local domain, that P is a minimal prime of xR, and that there is a prime Q with $0 \subset Q \subset P$, where the inclusions are strict. We shall get a contradiction.

Recall that $Q^{(n)} = Q^n R_Q \cap R$, the nth symbolic power of Q. It is Q-primary. Now, the ring R/xR has only one prime ideal, P/xR. Therefore it is a zero dimensional local ring, and has DCC. In consequence the chain of ideals $Q^{(n)}R/xR$ is eventually stable. Taking inverse images in R, we find that there exists N such that

$$Q^{(n)} + xR = Q^{(n+1)} + xR$$

for all $n \geq N$. For $n \geq N$ we have $Q^{(n)} \subseteq Q^{(n+1)} + xR$. Let $u \in Q^{(n)}$. Then $u = q + xr$ where $q \in Q^{(n+1)}$, and so $xr = u - q \in Q^{(n)}$. But $x \notin Q$, since P is the only minimal prime of xR in R. Since $Q^{(n)}$ is Q-primary, we have that $r \in Q^{(n)}$. This leads to the conclusion that $Q^{(n)} \subseteq Q^{(n+1)} + xQ^{(n)}$, and so

$$Q^{(n)} = Q^{(n+1)} + xQ^{(n)}.$$

But that means that with $M = Q^{(n)}/Q^{(n+1)}$, we have that $M = xM$. By Nakayama's lemma, $M = 0$, i.e., $Q^n/Q^{n+1} = 0$.

Thus, $Q^{(n)} = Q^{(N)}$ for all $n \geq N$. If $a \in Q - \{0\}$, it follows that $a^N \in Q^N \subseteq Q^{(N)}$ and is hence in the intersection of all the $Q^{(n)}$. But then, since $Q^{(n)} \subseteq Q^n R$ for all n, in the local domain R_Q, the intersection of the powers of the maximal ideal QR_Q is not 0, a contradiction. □

Before proving the strong version of the principal ideal theorem, we want to record the following result on prime avoidance. In applications of part (b) of this result, W is frequently a K-algebra R, while the other subspaces are ideals of R. This shows that if there is an infinite field in the ring R, the assumptions about ideals being prime in part (a) are not needed.

Theorem (prime avoidance). *Let R be a ring. Let $V \subseteq W$ be vector spaces over an infinite field K.*

(a) *Let $\mathfrak{A}$ be an ideal of R (or a subset of R closed under addition and multiplication). Given finitely many ideals of R all but two of which are prime, if $\mathfrak{A}$ is not contained in any of these ideals, then it is not contained in their union.*

(b) *Given finitely many subspaces of W, if V is not contained in any of these subspaces, then V is not contained in their union.*

(c) *(Ed Davis) Let $x \in R$ and $I, P_1, \ldots, P_n$ be ideals of R such that the P_i are prime. If $I + Rx$ is not contained in any of the P_t, then for some $i \in I$, $i + x \notin \bigcup_t P_t$.*

Proof. (a) We may assume that no term may be omitted from the union, or work with a smaller family of ideals. Call the ideals I, J, $P_1, \ldots, P_n$ with the P_t prime. Choose elements $i \in I \cap \mathfrak{A}$, $j \in J \cap \mathfrak{A}$, and $a_t \in P_t \cap \mathfrak{A}$, $1 \le t \le n$, such that each belongs to only one of the ideals I, J, $P_1, \ldots, P_n$, i.e., to the one it is specified to be in. This must be possible, or not all of the ideals would be needed to cover $\mathfrak{A}$. Let $a = (i + j) + ijb$ where

$$b = \prod_{t \text{ such that } i+j \notin P_t} a_t,$$

where a product over the empty set is defined to be 1. Then $i + j$ is not in I nor in J, while ijb is in both, so that $a \notin I$ and $a \notin J$. Now choose t, $1 \le t \le n$. If $i + j \in P_t$, the factors of ijb are not in P_t, and so $ijb \notin P_t$, and therefore $a \notin P_t$. If $i + j \notin P_t$ there is a factor of b in P_t, and so $a \notin P_t$ again.

(b) If V is not contained in any one of the finitely many vector spaces V_t covering V, for every t choose a vector $v_t \in V - V_t$. Let V_0 be the span of the v_t. Then V_0 is a finite-dimensional counterexample. We replace V by V_0 and V_t by its intersection with V_0. Thus, we need only show that a finite-dimensional vector space K^n is not a finite union of proper subspaces V_t. (When the field is algebraically closed we have a contradiction because K^n is irreducible. Essentially the same idea works over any infinite field.) For each t we can choose a linear form $L_t \neq 0$ that vanishes on V_t. The product $f = L_1 \cdots L_t$ is a nonzero polynomial that vanishes identically on K^n. This is a contradiction, since K is infinite.

(c) We may assume that no P_t may be omitted from the union. For every t, choose an element p_t in P_t and not in any of the other P_k. Suppose, after renumbering, that $P_1, \ldots, P_k$ all contain x while the other P_t do not (the values 0 and n for k are allowed). If $I \subseteq \bigcup_{j=1}^k P_j$ then it is easy to see that $I + Rx \subseteq \bigcup_{j=1}^k P_j$, and hence in one of the P_j by part (a), a contradiction. Choose $i' \in I$ not in any of $P_1, \ldots, P_k$. Let q be the product of the p_t for $t > k$ (or 1, if $k = n$). Then $x + i'q$ is not in any P_t, and so we may take $i = i'q$. $\square$

Lecture of November 27

Examples. Let $K = \mathbb{Z}/2\mathbb{Z}$ and let $V = K^2$. This vector space is the union of the three subspaces spanned by (1, 0), (0, 1) and (1, 1), respectively. This explains why we need an infinite field in part (b) of the preceding theorem. Now consider the K-algebra $K \oplus_K V$ where the product of any two elements of V is 0. (This ring is isomorphic with $K[x,\, y]/(x^2,\, xy,\, y^2)$, where x and y are indeterminates.) Then the maximal ideal is, likewise, the union of the three ideals spanned by its three nonzero elements. This shows that we cannot replace "all but two are prime" by "all but three are prime" in part (a) of the preceding theorem.

Proof of Krull's principal ideal theorem, strong version. We begin by proving by induction on n that the first statement holds. If $n = 0$ then P is a minimal prime of (0) and this does mean that P has height 0. Note that the zero ideal is the ideal generated by the empty set, and so constitutes a 0 generator ideal. The case where $n = 1$ has already been proved. Now suppose that $n \geq 2$ and that we know the result for integers $< n$. Suppose that P is a minimal prime of $(x_1,\, \ldots\, ,x_n)R$, and that we want to show that the height of P is at most n. Suppose not, and that there is a chain of primes

$$P = P_{n+1} \supset \cdots \supset P_0$$

with strict inclusions. If $x_1 \in P_1$ then P is evidently also a minimal prime of $P_1 + (x_2,\, \ldots\, ,\, x_n)R$, and this implies that P/P_1 is a minimal prime of the ideal generated by the images of $x_2,\, \ldots\, ,\, x_n$ in R/P_1. The chain

$$P_{n+1}/P_1 \supset \cdots \supset P_1/P_1$$

then contradicts the induction hypothesis. Therefore, it will suffice to show that the chain

$$P = P_{n+1} \supset \cdots \supset P_1 \supset 0$$

can be modified so that $x = x_1$ is in P_1. Suppose that $x \in P_k$ but not in P_{k-1} for $k \geq 2$. (To get started, note that $x \in P = P_{n+1}$.) It will suffice to show that there is a prime strictly between P_k and P_{k-2} that contains x, for then we may use this prime instead of P_{k-1}, and we have increased the number of primes in the chain that contain x. Thus, we eventually reach a chain such that $x \in P_1$.

To find such a prime, we may work in the local domain

$$D = R_{P_k}/P_{k-2}R_{P_k}.$$

The element x has nonzero image in the maximal ideal of this ring. A minimal prime P' of xR in this ring cannot be $P_kR_{P_k}$, for that ideal has height at least two, and P' has height at most one by the case of the principal ideal theorem already proved. Of course, $P' \neq 0$ since it contains $x \neq 0$. The inverse image of P' in R gives the required prime.

Thus, we can modify the chain

$$P = P_{n+1} \supset \cdots \supset P_1 \supset P_0$$

repeatedly until $x_1 \in P_1$. This completes the proof that the height of P is at most n.

We now prove the converse. Suppose that P is a prime ideal of R of height n. We want to show that we can choose $x_1, \ldots, x_n$ in P such that P is a minimal prime of $(x_1, \ldots, x_n)R$. If $n = 0$ we take the empty set of x_i. The fact that P has height 0 means precisely that it is a minimal prime of (0). It remains to consider the case where $n > 0$. We use induction on n. Let $q_1, \ldots, q_k$ be the minimal primes of R that are contained in P. Then P cannot be contained in the union of these, or else it will be contained in one of them, and hence be equal to one of them and of height 0. Choose $x_1 \in P$ not in any minimal prime contained in P. Then the height of P/x_1R in R/x_1R is at most $n-1$: the chains in R descending from P that had maximum length n must have ended with a minimal prime of R contained in P, and these are now longer available. By the induction hypothesis, P/x_1R is a minimal prime of an ideal generated by at most $n-1$ elements. Consider x_1 together with pre-images of these elements chosen in R. Then P is a minimal prime of the ideal they generate, and so P is a minimal prime of an ideal generated by at most n elements. The number cannot be smaller than n, or else by the first part, P could not have height n. □

If (R, m) is a local ring of Krull dimension n, a *system of parameters* for R is a sequence of elements $x_1, \ldots, x_n \in m$ such that, equivalently:

(1) m is a minimal prime of $(x_1, \ldots, x_n)R$.
(2) $\operatorname{Rad}(x_1, \ldots, x_n)R$ is m.
(3) m has a power in $(x_1, \ldots, x_n)R$.
(4) $(x_1, \ldots, x_n)R$ is m-primary.

The theorem we have just proved shows that every local ring of Krull dimension n has a system of parameters.

One cannot have fewer than n elements generating an ideal whose radical is m, for then $\dim(R)$ would be $< n$. We leave it to the reader to see that $x_1, \ldots, x_k \in m$ can be extended to a system of parameters for R if and only if

$$\dim\bigl(R/(x_1, \ldots, x_k)R\bigr) \le n-k,$$

in which case

$$\dim\bigl(R/(x_1, \ldots, x_k)R\bigr) = n-k.$$

In particular, $x = x_1$ is part of a system of parameters iff x is not in any minimal prime P of R such that $\dim(R/P) = n$. In this situation, elements $y_1, \ldots, y_{n-k}$ extend $x_1, \ldots, x_k$ to a system of parameters for R if and only if their images in $R/(x_1, \ldots, x_k)R$ are a system of parameters for $R/(x_1, \ldots, x_k)R$.

The following statement is now immediate:

Corollary. *Let $(R,\, m)$ be local and let $x_1,\, \dots\, , x_k$ be k elements of m. Then the dimension of $R/(x_1,\, \dots\, , x_k)R$ is at least* $\dim(R) - k$.

Proof. Suppose the quotient has dimension h. If $y_1,\, \dots\, , y_h \in m$ are such that their images in $R/(x_1,\, \dots\, , x_k)R$ are a system of parameters in the quotient, then m is a minimal prime of $(x_1,\, \dots\, , x_k,\, y_1,\, \dots\, , y_h)R$, which shows that $h + k \geq n$. □

We are now almost ready to address the issue of how dimension behaves for Noetherian rings when one adjoins either polynomial or formal power series indeterminates.

We first note the following fact:

Lemma. *Let x be an indeterminate over R. Then x is in every maximal ideal of $R[[x]]$.*

Proof. If x is not in the maximal ideal $\mathcal{M}$ it has an inverse mod $\mathcal{M}$, so that we have $xf \equiv 1$ mod $\mathcal{M}$, i.e., $1 - xf \in \mathcal{M}$. Thus, it will suffice to show that $1 - xf$ is a unit. The idea of the proof is to show that

$$u = 1 + xf + x^2f^2 + x^3f^3 + \cdots$$

is an inverse: the infinite sum makes sense because only finitely many terms involve any given power of x. Note that

$$u = (1 + xf + \cdots + x^nf^n) + x^{n+1}w_n$$

with

$$w_n = f^{n+1} + xf^{n+2} + x^2f^{n+3} + \cdots\,,$$

which again makes sense since any given power of x occurs in only finitely many terms. Thus:

$$u(1 - xf) - 1 = (1 + xf + \cdots + x^nf^n)(1 - xf) + x^{n+1}w_n(1 - xf) - 1.$$

The first of the summands on the right is $1 - x^{n+1}f^{n+1}$, and so this becomes

$$1 - x^{n+1}f^{n+1} + x^{n+1}w_n(1 - xf) - 1 = x^{n+1}\bigl(-f^{n+1} + w_n(1 - xf)\bigr) \in x^{n+1}R[[x]],$$

and since the intersection of the ideals $x^tR[[x]]$ is clearly 0, we have that $u(1-xf)-1 = 0$, as required. □

Theorem. *Let R be a Noetherian ring and let $x_1,\, \dots\, , x_n$ be indeterminates. Then $S = R[x_1,\, \dots\, , x_k]$ and $T = R[[x_1,\, \dots\, , x_k]]$ both have dimension* $\dim(R) + k$.

Proof. By a straightforward induction we may assume that $k = 1$. Write $x_1 = x$. If P is a prime ideal of R then PS and PT are both prime, with quotients $(R/P)[x]$ and $(R/P)[[x]]$, and $PS + xS$, $PS + xT$ are prime as well. If $P_0 \subset \cdots \subset P_n$ is a chain of primes in R, then their expansions P_i^{e} together with $P_n^{\mathrm{e}} + (x)$ give a chain of primes of length one greater in S or T. This shows that the dimensions of S and T are at least $\dim(R) + 1$.

If R has infinite dimension, so do S and T. Therefore let $\dim(R) = n$ be finite. We want to show that S and T have dimension at most $n+1$. We first consider the case of $S = R[x]$. Let Q be a prime ideal of this ring and let P be its contraction to R. It suffices to show that the height of Q is at most one more than the height of P. To this end we can replace R by R_P and S by $R_P[x]$: $QR_P[x]$ will be a prime ideal of this ring, and the height of Q is the same as the height of its expansion. We have therefore reduced to the local case. Let $x_1, \ldots, x_n$ be a system of parameters for R (which is now local). It suffices to show that we can extend it to a system of parameters for $R[x]_Q$ using at most one more element. It therefore suffices to show that $R[x]_Q/(x_1, \ldots, x_n)$ has dimension at most 1. This ring is a localization of $(R/(x_1, \ldots, x_n))[x]$, and so it suffices to see that this ring has dimension at most 1. To this end, we may kill the ideal of nilpotents, which is the expansion of P, producing $K[x]$. Since this ring has dimension 1, we are done.

In the case of T we first note that, by the Lemma, every maximal ideal of T contains x. Choose Q maximal in T. Since $x \in Q$, Q corresponds to a maximal ideal m of R, and has the form $m^{\mathrm{e}} + (x)$. If m is minimal over $(x_1, \ldots, x_n)$, then Q is minimal over $(x_1, \ldots, x_n, x)$. This proves that the height of $Q \leq n+1$, as required. □

If R is not Noetherian but has finite Krull dimension n, it is true that $R[x]$ has finite Krull dimension, and it lies between $n+1$ and $2n+1$. The upper bound is proved by showing that in a chain of primes in $R[x]$, at most two (necessarily consecutive) primes lie over the same prime P of R. This result is sharp.

Lecture of November 29

Let K be an algebraically closed field. Given two algebraic sets $X = \mathcal{V}(I) \in K^m = \mathbb{A}_K^m$, where we use $x_1, \ldots, x_m$ for coordinates, and $Y = \mathcal{V}(J) \subseteq K^n = \mathbb{A}_K^n$, where we use $y_1, \ldots, y_n$ for coordinates, the set $X \times Y \subseteq K^{m+n} = \mathbb{A}_K^{m+n}$ is an algebraic set defined by the expansions of I and J to $K[x_1, \ldots, x_m, y_1, \ldots, y_n] \cong K[x_1, \ldots, x_m] \otimes_K K[y_1, \ldots, y_n]$. It is obvious that a point satisfies both the conditions imposed by the vanishing of I and of J if and only if its first m coordinates give a point of X and its last n coordinates give a point of Y.

Let $S = K[x_1, \ldots, x_m]$ thought of as $K[\mathbb{A}_K^m]$ and $Y = K[y_1, \ldots, y_n]$ thought of as $K[\mathbb{A}_K^n]$. Then

$$K[X \times Y] \cong (S \otimes_K T)/\mathrm{Rad}\,(I^{\mathrm{e}} + J^{\mathrm{e}}),$$

where the superscript $^{\mathrm{e}}$ indicates expansion of ideals. Since

$$(S \otimes_K T)/(I^{\mathrm{e}} + J^{\mathrm{e}}) \cong (S \otimes_K T)/(I \otimes_K T + S \otimes_K J) \cong (S/I) \otimes_K (T/J),$$

we have that

$$K[X \times Y] \cong \bigl((S/I) \otimes_K (T/J)\bigr)_{\mathrm{red}} \cong (K[X] \otimes_K K[Y])_{\mathrm{red}}.$$

It is not necessary to kill the nilpotents, because of the following fact:

Theorem. *Let R and S be algebras over an algebraically closed field K.*
(a) *If R and S are domains, then $R \otimes_K S$ is a domain.*
(b) *If R and S are reduced, then $R \otimes_K S$ is reduced.*

Proof. For part (a), let $\mathcal{F}$ denote the fraction field of R. Since K is a field, every K-module is free, and, therefore, flat. We have an injection $R \hookrightarrow \mathcal{F}$. Thus, $R \otimes_K S \hookrightarrow \mathcal{F} \otimes_K S$. By Supplementary Problem Set #4, problem **6.**, this ring is a domain, and so its subring $R \otimes_K S$ is a domain.

For part (b), note that R is a the directed union of its finitely generated K-subalgebras R_0. Thus, $R \otimes_K S$ is the directed union of its subalgebras $R_0 \otimes_K S$ where $R_0 \subseteq R$ is finitely generated. Similarly, this ring is the directed union of its subalgebras $R_0 \otimes_K S_0$, where both $R_0 \subseteq R$ and $S_0 \subseteq S$ are finitely generated. We can therefore reduce to the case where R and S are finitely generated.

Let $P_1, \ldots, P_m$ be the minimal primes of R. Since R is reduced, their intersection is 0. Therefore, R injects into $\prod_i (R/P_i)$. Thus,

$$R \otimes_K S \hookrightarrow \bigl(\prod_i (R/P_i)\bigr) \otimes_K S \cong \prod_i \bigl((R/P_i) \otimes_K S\bigr)$$

(if we think of the products as direct sums, we have an obvious isomorphism of K-vector spaces: the check that multiplication is preserved is straightforward), and so it suffices to show that each factor ring of this product is reduced. Thus, we need only show that if R is a domain and S is reduced, where these are finitely generated K-algebras, then $R \otimes_K S$ is reduced. But now we may repeat this argument using the minimal primes $Q_1, \ldots, Q_n$ of S, and so we need only show that each ring $R \otimes_K (S/Q_j)$ is reduced, where now both R and S/Q_j are domains. By part (a), these tensor products are domains. □

One may also show that the tensor product of two reduced rings over an algebraically closed field is reduced using an equational argument and Hilbert's Nullstellensatz, similar to the argument for Supplementary Problem Set #4, **6.**

We return to the study of algebraic sets over an algebraically closed field. We have now established an isomorphism $K[X \times Y] \cong K[X] \otimes_K K[Y]$. Moreover, it is easy to see that the product projections $X \times Y \to X$, $X \times Y \to Y$ correspond to the respective injections $K[X] \to K[X] \otimes_K K[Y]$ and $K[Y] \to K[X] \otimes_K K[Y]$, where the first sends $f \mapsto f \otimes 1$ and the second sends $g \mapsto 1 \otimes g$.

From the fact that $K[X] \otimes_K K[Y]$ is a coproduct of $K[X]$ and $K[Y]$ in the category of K-algebras, it follows easily that $X \times Y$ (with the usual product projections) is a product of X and Y in the category of algebraic sets. That is, giving a morphism from Z to $X \times Y$ is equivalent to giving a pair of morphisms, one from Z to X and the other from $Z \to Y$. This is simply because giving a morphism from Z to $X \times Y$ is equivalent to giving a K-homomorphism $K[X] \otimes_K K[Y]$ to $K[Z]$, which we know is equivalent to giving a K-homomorphism $K[X] \to K[Z]$ and a K-homomorphism $K[Y] \to K[Z]$: as already mentioned, $K[X] \otimes_K K[Y]$ is a coproduct for $K[X]$ and $K[Y]$ in the category of K-algebras. Notice also that since $K[X] \otimes_K K[Y]$ is a domain whenever $K[X]$ and $K[Y]$ are both domains, we have:

Corollary. *The product of two varieties (i.e., irreducible algebraic sets) in $\mathbb{A}^n_K$ over an algebraically closed field K is a variety (i.e., irreducible).*

We also note:

Proposition. *If X and Y are algebraic sets over the algebraically closed field K, then*

$$\dim(X \times Y) = \dim(X) + \dim(Y).$$

Proof. $K[X]$ is module-finite over a polynomial ring A in d variables where $d = \dim(X)$, say with module generators $u_1, \ldots, u_s$, and $K[Y]$ is module-finite, say with module generators $v_1, \ldots, v_t$, over a polynomial ring B in d' variables. Hence, $K[X] \otimes_K K[Y]$ is module-finite (with module generators $u_i \otimes v_j$) over a polynomial ring in $d + d'$ variables. Note that $A \otimes_K B$ injects into $A \otimes_K K[Y]$ because A is K-flat, and the latter injects into $K[X] \otimes_K K[Y]$ because $K[Y]$ is K-flat. □

We next prove a result that was promised long ago:

Theorem. *Let X and Y be irreducible algebraic sets meeting at a point $x \in \mathbb{A}^n_K$, where K is an algebraic closed field. Then*

$$\dim(X \cap Y) \geq \dim(X) + \dim(Y) - n.$$

In fact every irreducible component of $X \cap Y$ has dimension $\geq \dim(X) + \dim(Y) - n$.

Proof. Let $X = \mathcal{V}(P)$ and $Y = \mathcal{V}(Q)$, where P and Q are prime ideals of $K[x_1, \ldots, x_n]$. Then $X \cap Y = \mathcal{V}(P+Q)$, although $P+Q$ need not be radical, and

$$K[X \cap Y] = \bigl(K[x_1, \ldots, x_n]/(P+Q)\bigr)_{\mathrm{red}}.$$

Now

$$K[x_1, \ldots, x_n]/(P+Q) \cong \bigl((K[x_1, \ldots, x_n]/P) \otimes_K (K[y_1, \ldots, y_n]/Q')\bigr)/I_\Delta,$$

where I_Δ is the ideal generated by the $x_i - y_i$ for $1 \leq i \leq n$, which is the ideal that defines the diagonal Δ in $\mathbb{A}^n_K \times_K \mathbb{A}^n_K$. The point is that once we kill the generators $x_i - y_i$ of I_Δ, the ring $K[y_1, \ldots, y_n]$ is identified with $K[x_1, \ldots, x_n]$, and the image of Q' is Q. (Geometrically, we are identifying $X \cap Y$ with $(X \times Y) \cap \Delta$ in $\mathbb{A}^n_K \times \mathbb{A}^n_K$, via the map $z \mapsto (z, z)$.) Let $R = (K[x_1, \ldots, x_n]/P) \otimes_K (K[y_1, \ldots, y_n]/Q')$. The dimension of $R = K[X \times Y]$ is $\dim(X) + \dim(Y)$. Since the intersection $X \cap Y$ is non-empty, we know that I_Δ expands to a proper ideal. The dimension of the quotient will be the supremum of the heights of the m/I_Δ as m runs through maximal ideals containing I_Δ, and this will be the supremum of the dimensions of the local rings $\dim(R_m/I_\Delta R_m)$. Each R_m has dimension equal to that of R, i.e., $\dim(X) + \dim(Y)$. But I_Δ is generated by n elements, and killing n elements in the maximal ideal of a local ring drops the dimension of the local ring by at most n. Thus, every $R_m/I_\Delta R_m$ has dimension at least $\dim(X) + \dim(Y) - n$, and the result follows. To get the final statement, let x be a point of the irreducible component considered not in any other irreducible component of $X \cap Y$, and let m be the corresponding maximal ideal of R. We have that $R_m/I_\Delta R_m$ has dimension at least $\dim(X) + \dim(Y) - n$ as before, but now there is a unique minimal prime P in this ring, corresponding to the fact that only one irreducible component of $X \cap Y$ contains x. It follows that this irreducible component has dimension at least $\dim(X) + \dim(Y) - n$. □

Note that the argument in the proof shows that the map $X \cap Y \to (X \times Y) \cap \Delta$ that sends z to (z, z) is an isomorphism of algebraic sets.

Recall that $\dim_x(X)$ is the largest dimension of an irreducible component of X that contains x. It follows at once that:

Corollary. *Let X and Y be algebraic sets in K^n, where K is an algebraically closed field, and suppose $x \in X \cap Y$. Then*

$$\dim_x(X \cap Y) \geq \dim_x(X) + \dim_x(Y) - n.$$

Proof. Let X_0 be an irreducible component of X containing x of largest dimension that contains x and Y_0 be such a component of Y with $x \in Y_0$. Then $\dim_x(X) = \dim(X_0)$ and $\dim_x(Y) = \dim(Y_0)$. Apply the result for the irreducible case to X_0 and Y_0. □

The theorem we have just proved may be thought of as an existence theorem for solutions of equations: given two sets of equations in n variables over an algebraically closed field, if the two sets of equations have a common solution x, and the solutions of the first set have dimension d near x while the solutions of the second set have dimension d' near x, then the set of simultaneous solutions of the two sets has dimension at least $d + d' - n$ near x. This is well known for solutions of linear equations, but surprising for algebraic sets!

Lecture of December 1

A subset of a topological space is called *locally closed* if it is, equivalently, (1) the intersection of an open set with a closed set, (2) a closed subset of an open set, or (3) an open subset of a closed set. Let $X \subseteq \mathbb{A}^n_K$ be a closed algebraic set. Let $f \in K[X] = R$, and let $X_f = \{x \in X : f(x) \neq 0\}$. Then X_f corresponds bijectively to the set of maximal ideals in R_f. Therefore, X_f has the structure of a closed algebraic set (*a priori*, it is only a locally closed algebraic set). If we think of R as $K[x_1, \dots, x_n]/I$ where $I = \mathcal{I}(X)$, we can map $K[x_1, \dots, x_{n+1}] \twoheadrightarrow R_f$, extending the map $K[x_1, \dots, x_n] \twoheadrightarrow R$ by mapping $x_{n+1} \to 1/f$. X now corresponds bijectively to a closed algebraic set in $\mathbb{A}^{n+1}_K$: the bijection sends x to $\bigl(x, 1/f(x)\bigr)$. The closed algebraic set in question may be described as $\{(x, \lambda) \in \mathbb{A}^{n+1}_K : x \in X \text{ and } \lambda = 1/f(x)\}$. The new defining ideal is $I + (fx_{n+1} - 1)$.

We define a function $X_f \to K$ to be *regular* if it is regular with respect to the closed algebraic set structure that we have placed on X_f. This raises the following question: suppose that we have a cover of a closed algebraic set X by open sets X_{f_i} and a function $g : X \to K$ such that the restriction of g to each X_{f_i} is regular in the sense just specified. Is g regular? We shall show that the answer is "yes," and this shows that regularity is a local property with respect to the Zariski topology. Let g_i denote the restriction of g to $X_i = X_{f_i}$. Note that $g_i|_{X_j} = g_j|_{X_i}$ for all i, j, since they are both restrictions of g.

The following fact gives a generalization to arbitrary modules over an arbitrary commutative ring, and underlies the theory of schemes.

Theorem. *Let R be any ring and M any R-module. Let $X = \mathrm{Spec}\,(R)$, and let f_i be a family of elements of R such that the open sets $X_i = X_{f_i} = D(f_i)$ cover X. Suppose that for every i we are given an element $u_i \in M_{f_i} = M_i$, and suppose that $(*)$ for all choices of i and j, the images of u_i and u_j in $M_{f_i f_j}$ agree. Then there is a unique element $u \in M$ such that for all i, the image of u in M_{f_i} is u_i.*

The result says, informally, that "constructing" an element of a module is a local problem: one can solve it on an open cover, provided the solutions "fit together" on overlaps. This turns many problems into local problems: for example, if M is finitely presented, the problem of constructing a map of modules $M \to N$ amounts to giving an element of the module $\mathrm{Hom}_R(M, N)$. Since localization commutes with Hom when M is finitely presented, the problem of doing the construction becomes local.

Note that if we apply this result in the case of the algebraic set X, we find that there is an element $g_0 \in K[X]$ whose image in $K[X_i]$ is g_i for all i. This implies that g_0 agrees with g on X_i. Since the X_i cover X, $g_0 = g$. Thus, $g \in K[X]$. Consequently, the theorem stated above does show that regularity is a local property.

Proof of the theorem. Uniqueness is obvious: if u and u' are two such elements, then they agree after localizing at any f_i. When one localizes at a prime P, since P cannot contain all the f_i, u and u' have the same image in M_P. It follows that $u = u'$. We focus on the existence of u.

The statement that the X_i cover is equivalent to the statement that the f_i generate the unit ideal. Then finitely many generate the unit ideal: call these $f_{i_1}, \ldots, f_{i_n}$. Suppose that we can construct $u \in M$ such that the image of u is $u_{i_t} \in M_{i_t}$, $1 \leq t \leq n$. We claim that the image u'_j of u in M_j is u_j for any j. To see this, it suffices to show that $u'_j - u_j$ vanishes in $(M_j)_P$ for any $P \in X_j$. But X_j is covered by the sets $X_j \cap X_{i_t}$, $1 \leq t \leq n$. If $P \in X_{i_t}$, it suffices to show that u'_j and u_j have the same image in $M_{f_{i_t} f_j}$. The image of u'_j is the same as the image of u, and hence the same as the image of u_{i_t}, and the result follows from our assumption $(*)$.

Therefore, it suffices to work with the cover by the $X_{f_{i_t}}$, and we simplify notation: we let the index set be $\{1, \ldots, n\}$ and so the f_is are simply $f_1, \ldots, f_n$, the cover is $X_1, \ldots, X_n$, and $M_i = M_{f_i}$. We use induction on n. If $n = 1$, $X_1 = X$ and the result is clear: $u = u_1$.

We next consider the case where $n = 2$. This is the core of the proof. Let $u_1 = v_1/f_1^s$ and $u_2 = v_2/f_2^t$ where $v_1, v_2 \in M$. Since these agree in $M_{f_1 f_2}$ there exists an integer N such $f_1^N f_2^N (f_2^t v_1 - f_1^s v_2) = 0$. Then $u_1 = f_1^N v_1/f_1^{N+s}$, $u_2 = f_2^N v_2/f_2^{N+t}$, and

$$f_2^{N+t} f_1^N v_1 - f_1^{N+s} f_2^N v_2 = (f_1 f_2)^N (f_2^t v_1 - f_1^s v_2) = 0.$$

Thus, if we replace f_1 by f_1^{N+s}, f_2 by f_2^{N+t}, v_1 by $f_1^N v_1$ and v_2 by $f_2^N v_2$, then $u_1 = v_1/f_1$, $u_2 = v_2/f_2$, and $f_2 v_1 - f_1 v_2 = 0$ (Note that the original f_1^{N+s} and f_2^{N+t} generate the unit ideal, since any maximal ideal containing both would have to contain both f_1 and f_2, a contradiction: thus, the new f_1 and f_2 still generate the unit ideal).

Choose r_1, r_2 such that $r_1 f_1 + r_2 f_2 = 1$. Let $u = r_1 v_1 + r_2 v_2$. Then

$$f_1 u = r_1 f_1 v_1 + r_2 (f_1 v_2) = r_1 f_1 v_1 + r_2 (f_2 v_1) = (r_1 f_1 + r_2 f_2) v_1 = v_1,$$

so that $u = v_1/f_1$ in M_1, and $u = v_2/f_2$ in M_2 by symmetry.

We now assume that $n > 2$ and that the result has been established for integers $< n$. Suppose that

$$r_1 f_1 + \cdots + r_n f_n = 1.$$

Let

$$g_1 = r_1 f_1 + \cdots + r_{n-1} f_{n-1}$$

and $g_2 = f_n$. Evidently, g_1 and g_2 generate the unit ideal, since $g_1 + r_n g_2 = 1$. Consider the images of $f_1, \ldots, f_{n-1}$ in R_{g_1}. Because g_1 is invertible, they generate the unit ideal. We now apply the induction hypothesis to M_{g_1}, using the images of the f_i for $1 \leq i \leq n-1$ to give the open cover of Spec$\,(R_{g_1})$. Let u'_i denote the image of u_i in $M_{g_1 f_i}$, $1 \leq i \leq n-1$. It is straightforward to verify that condition $(*)$ continues to hold here, using cases of the original condition $(*)$. By the induction hypothesis, there is an element of M_{g_1}, call it w_1, such that the image of w_1 in each $M_{g_1 f_i}$ is the same as the image of u_i, $1 \leq i \leq n-1$. We claim that the images of w_1 and u_n agree in $M_{g_1 f_n}$. It suffices to show that they agree after localizing at any prime P, and P cannot contain the images of all of $f_1, \ldots, f_{n-1}$. If P does not contain f_i, $1 \leq i \leq n-1$, the result follows because the images of u_i and u_n agree in $M_{f_i f_n}$. We can now apply the case where $n = 2$ to construct the required element of M. □

Corollary. *Let X and Y be closed algebraic sets over an algebraically closed field K. Then a function $h : X \to Y$ is regular if and only if (#) it is continuous and for all $x \in X$ there is an open neighborhood Y_g of $y = h(x)$ and an open neighborhood $X_f \subseteq h^{-1}(y)$ such that the restriction of h mapping X_f to Y_g is regular.*

Proof. $Y \subseteq \mathbb{A}^n_K$ (with coordinates $x_1, \ldots, x_n$ in the latter), and we will reduce to showing that the composite map $X \to \mathbb{A}^n_K$ is regular. Let h_i be the composition of this map with the i th coordinate projection. It suffices to show that every h_i is regular. Let X_f be a neighborhood of $x \in X$ such that h maps into an open neighborhood Y_g of $h(x)$. It will correspond to a K-algebra homomorphism $K[Y]_g \to K[X]_f$. Note that g is the restriction of a function g' on $\mathbb{A}^n_K$, and $(\mathbb{A}^n_K)_{g'}$ meets Y in Y_g. The inclusion $Y \subseteq \mathbb{A}^n_K$ corresponds to a surjection $K[x_1, \ldots, x_n] \to K[Y]$. The map $Y_g \to (\mathbb{A}^n_K)_{g'}$ corresponds to the ring map $K[x_1, \ldots, x_n]_{g'} \to K[Y]_g$ induced by localization at the multiplicative system generated by g' (recall that g' maps to g). Thus, the map $X_f \to (\mathbb{A}^n_K)_{g'}$ is regular, and so is the map $X_f \to \mathbb{A}^n_K$, which corresponds to the composite ring map

$$K[x_1, \ldots, x_n] \to K[x_1, \ldots, x_n]_{g'} \to K[Y]_g \to K[X]_f.$$

It follows that the composition of the map $X_f \to \mathbb{A}^n_K$ with the i th coordinate projection is regular: this is the restriction of h_i to X_f. Since the X_f cover X, it follows that every h_i is regular, and so h is regular. □

We can now define when a function between open subsets of algebraic sets (i.e., locally closed algebraic sets) is a morphism: simply use the condition (#) in the Corollary.

A set has the structure of a *reduced scheme of finite type* over an algebraically closed field K if it is a topological space X with a finite open cover by sets X_i together with, for every i, a bijection $f_i : X_i \cong Y_i$ where Y_i is a closed algebraic set over K, satisfying the additional condition that if

$$f_{ij} : X_i \cap X_j \cong f_i(X_i \cap X_j) = Y_{ij} \subseteq Y_i,$$

then the for all i, j the composite

$$f_{ji} \circ f_{ij}^{-1} : Y_{ij} \to Y_{ji}$$

is an isomorphism of (locally closed) algebraic sets.

Roughly speaking, a reduced scheme of finite type over K is the result of pasting together finitely many closed algebraic sets along overlaps that are isomorphic in the category of locally closed algebraic sets. This is analogous to the definitions of topological, differentiable and analytic manifolds by pasting open subsets having the same structure as an open set in $\mathbb{R}^n$ (or $\mathbb{C}^n$ in the case of an analytic manifold).

One can use condition (#) to define when a function between two reduced schemes of finite type over K is a morphism: thus, we require that f be continuous, and that for all $x \in X$, if $y = f(x)$, then when we choose an open neighborhood V of y with the structure of a closed algebraic set, and and an open neighborhood U of x with the structure of a closed algebraic set such that $f(U) \subseteq V$, then restriction of f to a map from U to V is a morphism of algebraic sets. Our results on the local character of morphisms show that when X and Y are closed algebraic sets, we have not enlarged the set of morphisms from X to Y.

A major failing of this theory is that while the category of finitely generated K-algebras has rings with nilpotents, our reduced schemes never have any. It turns out that the presence of nilpotents can carry geometric information! Even if one detests nilpotents and never wants them around, it is very useful on occasion to be able to say that there really aren't any because of a suitable theorem (as opposed to saying that there aren't any because we were forced by our definitions to kill them all). For example, one cannot express the fact that the tensor product of two reduced K-algebras is reduced in the category of reduced schemes. While there is an object corresponding to the reduced tensor product, there is no object corresponding to the tensor product. The remedy is the theory of schemes: the category of schemes contains the opposite of the category of rings as a subcategory, and contains the category of reduced schemes of finite type over an algebraically closed field as well.

When one does the full theory of schemes, the definition of a reduced scheme of finite type over an algebraically closed field K is somewhat different, but the category of reduced schemes of finite type over K introduced here is equivalent to the category one gets from the more general theory of schemes.

Lecture of December 4

We briefly discuss the notion of the Krull dimension of a Noetherian module M over a Noetherian ring R. Let I be the annihilator of M in R. Then we define $\dim(M) = \dim(R/I)$, which is the same as

$$\sup\{\dim(R/P) : P \text{ is a minimal prime of } I\} = \sup\{\dim(R/P) : P \in \mathrm{Ass}\,(M)\}.$$

It follows easily that if

$$0 \to M' \to M \to M'' \to 0$$

is exact then

$$\dim(M) = \sup\{\dim(M'),\, \dim(M'')\}$$

and, by induction on n that if

$$0 = M_0 \subseteq M_1 \subseteq \cdots \subseteq M_{n-1} \subseteq M_n = M$$

is a finite filtration of M then

$$\dim(M) = \sup\{\dim(M_{i+1}/M_i) : 0 \le i \le n-1\}.$$

If $(R,\, m)$ is local with $x_1,\, \ldots\, , x_n \in m$ and $M \neq 0$ is finitely generated over R, then $\ell\big(M/(x_1,\, \ldots\, , x_n)M\big)$ is finite iff

$$M/(x_1,\, \ldots\, , x_n)M = \big(R/(x_1,\, \ldots\, , x_n)R\big) \otimes_R M$$

is supported precisely at m iff

$$\mathrm{Supp}\,\big(R/(x_1,\, \ldots\, , x_n)R\big) \cap \mathrm{Supp}\,(M) = \{m\}$$

iff

$$V\big((x_1,\, \ldots\, , x_n)R\big) \cap V(I) = \{m\}$$

iff $(x_1,\, \ldots\, , x_n)R + I$ has radical m iff $(x_1,\, \ldots\, , x_n)(R/I)$ is (m/I)-primary. The least integer n such that $M/(x_1,\, \ldots\, , x_n)M$ has finite length for $x_1,\, \ldots\, , x_n \in m$ is therefore the same as $\dim(R/I) = \dim(M)$, and the elements $x_1,\, \ldots\, , x_n \in m$ are called a *system of parameters* for M if $n = \dim(M)$ and $\ell\big(M/\big(x_1,\, \ldots\, , x_n\big)M\big)$ is finite. Clearly, $x_1,\, \ldots\, , x_n \in R$ form a system of parameters for M iff their images in R/I are a system of parameters for R/I.

Our next objective is to give an important characterization of normal Noetherian domains, and then apply it to the study of normal Noetherian domains of Krull dimension one. A normal Noetherian domain of Krull dimension one is called a *Dedekind* domain. Every PID that is not a field is a Dedekind domain. The integral closure of $\mathbb{Z}$ in a finite algebraic extension F of $\mathbb{Q}$ also turns out to be a Dedekind domain. F is called an *algebraic number field* and the integral closure of $\mathbb{Z}$ in F is called the *ring of algebraic integers* of F. It is module-finite over $\mathbb{Z}$, as we shall see below.

Before giving our characterization of normal Noetherian domains, we need:

Proposition. *Let R be a Noetherian ring, and suppose that P is an associated prime of the ideal xR, where $x \in R$ is not a zerodivisor. Then P is an associated prime of y for every nonzerodivisor $y \in P$.*

Proof. The issues are unaffected by passing to R_P. Therefore we may assume that $(R,\, P)$ is local. Then we may choose $a \in P - xR$ such that $Pa \subseteq xR$. Then $ya = xb$ for some $b \in R$. Note that $b \notin yR$, or else $b = yr$, and then $ya = xyr \Rightarrow y(a - xr) = 0 \Rightarrow a = xr$, since y is not a zerodivisor. This is a contradiction, since $a \notin xR$, which completes the proof that $b \notin yR$. We claim that $Pb \subseteq yR$, for if $u \in P$, we have that $ua = xs$ for some $s \in R$. Since $ya = xb$, we find that $uxb = uya = yxs$ and so $x(ub - ys) = 0 \Rightarrow ub = ys$, since x is not a zerodivisor. □

Proposition. *A local domain $(R,\, P)$ not a field is a DVR iff $P = yR$ is principal.*

Proof. The condition is clearly necessary. Now suppose that $P = yR$. Consider any nonzero element r of P. Since it is not in every power of P, there is a largest integer $n \geq 1$ such that $r = uy^n$ for $u \in R$. Then y does not divide u, which shows that $u \in R - P$ is a unit. That is, every nonzero non-unit is a unit times a power of y. It follows at once that any proper nonzero ideal is generated by the least power of y that it contains. □

Theorem. *Let R be a Noetherian domain. Then R is normal if and only if (1) every associated prime of any nonzero principal ideal has height one, and (2) the localization of R at every height one prime is a DVR. In particular, if R is one-dimensional and local, then R is normal if and only if R is a DVR.*

Proof. First suppose that R is normal. Let x be any nonzero element of R and let P be an associated prime of xR: note that any height one prime will be a minimal prime (and, hence, an associated prime) of a principal ideal. We may localize at P. We shall show that R_P is a DVR, which evidently implies that the height of P is one. This establishes both (1) and (2) in the characterization of normality.

Since $P \neq 0$, $P \neq P^2$, by Nakayama's lemma. Choose $y \in P - P^2$. We shall prove that $yR = P$, which shows that P has height one and that $R = R_P$ is a DVR. Note that P is an associated prime of yR, by the first Proposition above. Thus, we may choose $a \in R$ such that $a \notin yR$ but $Pa \subseteq yR$. If a is a unit we find that $P \subseteq yR$ and so $P = yR$ and we are done. Suppose $a \in P$. We shall obtain a contradiction. We claim that $Pa \subseteq yP$. For if $r \in P$ and $ra = yu$, if u were a unit we would have that $yu \in P^2$, and then $y \in P^2$, a contradiction.

Let $f_1, \, \ldots \, , f_n$ generate P. Then for every i we have an equation

$$af_i = y \sum_j r_{ij} f_j.$$

If we make $f_1, \, \ldots \, , f_n$ into the entries of a column vector V and let A be the matrix (r_{ij}), this says that $AV = (a/y)V$. We are working over a domain R, so that we may think over the fraction field of R. The entries of V generate the nonzero prime P, and so $V \neq 0$, and is an eigenvector of A for the eigenvalue a/y. It follows as usual that $\det\bigl((a/y)I - A\bigr) = 0$

i.e., that a/y satisfies the characteristic polynomial of the matrix $A = (r_{ij})$, which is a monic polynomial with coefficients in R. Since $a/y \in \mathrm{frac}\,(R)$ and R is normal, this implies that $a/y \in R$, and so $a \in yR$, a contradiction. This concludes the proof of the necessity of conditions (1) and (2).

Now suppose that associated primes of nonzero principal ideals are height one and that the localization of R at any height one prime is a DVR. We must show that R is normal. If not, choose a fraction α that is integral over R. Then $M = R[\alpha]/R$ is a finitely generated nonzero R-module: choose a prime $P \in \mathrm{Ass}\,(M)$. Now replace R by R_P. Our hypotheses are preserved. Moreover, $PR_P \in \mathrm{Ass}\,(M_P)$ and so $M_P \neq 0$, which means that $R_P[\alpha]$ is strictly bigger than R_P. We change notation and assume that (R, P) is local and that $\beta \in R[\alpha]$ is such that P is the annihilator of the image of β in $R[\alpha]/R$, that is, $\beta \notin R$ but $P\beta \subseteq R$. We may write $\beta = a/x$ where $x \in R - \{0\}$ and $a \notin xR$. Then $P(a/x) \subseteq R$, which implies that $Pa \subseteq xR$. This implies that P is an associated prime of the ideal xR. Therefore, P has height one. But then $R = R_P$ is a DVR, and is normal, since a DVR is a PID and therefore a UFD. This is a contradiction. □

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. 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 $[P]$ corresponding bijectively to the height one prime ideals P of R. 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 $\mathrm{div}\,(a)$. 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 $\mathrm{div}\,(I)$ associated with I, namely $\sum_{i=1}^n k_i[P_i]$. If $I = R$ is the unit ideal, we define $\mathrm{div}\,(I) = 0$.

Theorem. *Let R be a Noetherian normal domain. If I has pure height one, then so is fI for every nonzero element f of R, and* $\mathrm{div}\,(fI) = \mathrm{div}\,(f) + \mathrm{div}\,(I)$*. For any two ideals I and J of pure height one,* $\mathrm{div}\,(I) = \mathrm{div}\,(J)$ *iff $I = J$, while the images of* $\mathrm{div}\,(I)$ *and* $\mathrm{div}\,(J)$ *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* $\mathrm{div}\,(I)$ *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* $J \cong Hom_R(I, R)$*. 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. Note that $0 \subseteq fR/fI \subseteq R/fI$ and that the quotient is $\cong R/fR$ 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, $\operatorname{Hom}_R(I, R) \cong \{f \in R : (f/g)I \subseteq R\}$, 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, $\operatorname{Hom}_R(I, R) \cong gR :_R I = J$. We claim that J has pure height one (even if I does not) and that if I has pure height one then $\operatorname{div}(J) + \operatorname{div}(I) = \operatorname{div}(g)$, which shows that $\operatorname{div}(J) = -\operatorname{div}(I)$ in $\mathcal{C}\ell(R)$. If not, let P be an associated prime of J of height two or more, and localize at P. Then there is an element $u \notin J$, i.e., such that $uI \not\subseteq gI$, but such that $Pu \subseteq J$, i.e. $PuI \subseteq gR$. Choose $r \in I$ such that $ur \notin gR$. Then $Pur \subseteq gR$, which shows that P is an associated prime of g, a contradiction, since R is normal. Thus, J has pure height one. Now localize at any height one prime P to check that $\operatorname{div}(J) + \operatorname{div}(I) = \operatorname{div}(g)$. 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 needed to show that the coefficients of P in $\operatorname{div}(I)$ and $\operatorname{div}(J)$ sum to the coefficient of P in $\operatorname{div}(g)$.

It remains only to show that every element of $\mathcal{C}\ell(R)$ is represented by $\operatorname{div}(I)$ 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. □

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 $\operatorname{Hom}(\_\,, R)$, 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.)

Computing the divisor class group is extremely difficult, even for rings of algebraic integers in an algebraic number field: such calculations constitute a branch of mathematics in its own right. The problem is amazingly hard even for quadratic extensions of $\mathbb{Q}$.

We next want to comment further on the normal Noetherian domains of Krull dimension one: these are called *Dedekind domains.*

Theorem. *The following conditions on a domain R of Krull dimension one are equivalent:*

(1) *R is normal, i.e., is a Dedekind domain.*
(2) *For every maximal ideal P in R, R_P is a discrete valuation ring.*

In a Dedekind domain, every nonzero ideal other than R is uniquely a product of powers of maximal ideals.

Proof. We know that the property of being normal is local, and a local domain of Krull dimension one is normal if and only if it is a DVR. The final statement corresponds to primary decomposition for principal ideals in a normal Noetherian ring R: symbolic powers of maximal ideals agree with ordinary powers, since the ordinary powers are primary, and we can replace intersection with product because the ideals involved are pairwise comaximal. □

We shall come back to the study of Dedekind domains but we first want to observe some other results about normal rings that need not have Krull dimension one.

Lecture of December 6

Theorem. *A Noetherian ring is normal if and only if it is the intersection of discrete valuation rings, and these may be taken to be its localizations at height one primes.*

Proof. A DVR is normal, and hence an intersection of DVRs is normal. Thus, it suffices to show that a normal Noetherian ring is the intersection of its localizations at height one primes. Let $f = a/x$ be a fraction supposedly in all these localizations. Let $M = (R + R(a/x)/R)$, which is a nonzero module. Then a/x has some nonzero multiple b/x with prime annihilator P mod R. Localize at P, and change notation, replacing R by R_P.

It follows that $b/x \notin R$ but $P(b/x) \subseteq R$, which says that $b \notin xR$ but $Pb \subseteq xR$. This implies that P is an associated prime of the ideal xR, and since R is normal we have that P has height one and R is a DVR. But then $b/x \in R$, a contradiction.

Theorem. *A polynomial ring (even in infinitely many variables) over a normal Noetherian ring is normal.*

Proof. It is easy to check that a directed union of normal domains is normal. (An element *alpha* in the fraction field integral over the union will be in the fraction field of one of these domains, and likewise will be integral over one of them. But some domain D in the family will contain both, and since D is normal, α is integral over D and in the fraction field of D, and so must be in D.) Therefore, it suffices to consider the case of finitely many variables. By a straightforward induction we need only consider the case of one variable. Since the ring R is the intersection of discrete valuation rings $V \subseteq \mathcal{F} = \operatorname{frac}(R)$, it follows that $R[x]$ is the intersection of the rings $V[x] \subseteq \mathcal{F}[x]$, and every $V[x]$ is a UFD, and, hence, normal. Thus, $R[x]$ is normal. □

The same proof applies to the power series ring in finitely many variables over a normal Noetherian ring, although we are missing a step in the proof, since we do not know that the ring $V[[x]]$ is a UFD, but this is true. This is part of the theory of regular local rings. For polynomial rings, the Noetherian restriction on R can be removed. One method of proof is to represent R is an intersection of non-Noetherian valuation domains: these are domains in which the ideals are totally ordered. One thus reduces to proving the result for valuation domains. Another method is to show that a domain finitely generated over $\mathbb{Z}$ or over any $\mathbb{Z}/p\mathbb{Z}$ has the property that its integral closure is a finite module over it. It then follows that every normal ring is a directed union of Noetherian normal rings, and one can reduce to the Noetherian case. A direct argument is also possible.

However it is *not* true that when R is normal but not Noetherian, that the formal power series ring $R[[x]]$ must be normal: this is not true even when R is a valuation domain, although it is erroneously asserted to be true in the first edition of Nagata's book *Local Rings*. It is also worth noting that in the Noetherian case, the formal power series ring in one variable over a UFD need not be a UFD, by independent examples of D. Buchsbaum and P. Samuel, although the formal power series ring in any number of variables over a field or a discrete valuation ring is a UFD.

We next want to prove that certain integral closures are module-finite:

Theorem. *Let R be a normal Noetherian domain and let $\mathcal{L}$ be a finite separable algebraic extension of the fraction field $\mathcal{K}$ of R (separability is automatic if $\mathcal{K}$ has characteristic 0). Then the integral closure S of R in L is a module-finite over R, and, hence, a Noetherian normal ring.*

When $\mathcal{K} \subseteq \mathcal{L}$ is a finite algebraic extension of fields, for any $\lambda \in \mathcal{L}$, we define $\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(\lambda)$ to be trace of the $\mathcal{K}$-linear map $\mathcal{L} \to \mathcal{L}$ given by λ: it may be computed by choosing a basis for L over K, finding the matrix of the map given by multiplication by λ, and summing the entries of the main diagonal of this matrix. It is independent of the choice of basis. If the characteristic polynomial is $x^n - cx^{n-1}+$ lower degree terms, where $n = [\mathcal{L} : \mathcal{K}]$, the trace is c. It is also the sum of the eigenvalues of the matrix (calculated in a splitting field for f or any larger field, such as an algebraic closure of $\mathcal{K}$), i.e., the sum of the roots of f (where if a root has multiplicity k, it is used a summand k times in the sum of the roots). We give a further discussion of the properties of trace following the proof of the theorem.

A key element of the proof is that a finite algebraic extension $\mathcal{L}$ of $\mathcal{K}$ is separable if and only if some element of $\mathcal{L}$ has nonzero trace in $\mathcal{K}$. This fact is quite trivial in characteristic 0, since the trace of the identity element is $[\mathcal{L} : \mathcal{K}] \neq 0$. This implies that the function $B : L \times L \to K$ that maps (a, b) to the trace of ab is a non-degenerate symmetric bilinear form: it is non-degenerate because if c has nonzero trace, given $a \in L - \{0\}$, $B(a, c/a)$ is the trace of c, and so is not 0. Here, $n = [\mathcal{L} : \mathcal{K}]$. This non-degeneracy tells us that if $b_1, \ldots, b_n$ is any basis for $\mathcal{L}$ over $\mathcal{K}$, then the matrix $(Tr_{\mathcal{L}/\mathcal{K}} b_i b_j)$ is invertible over $\mathcal{K}$, and we shall assume this in proving the theorem. After we give the proof we discuss further the facts about bilinear forms and about trace that we are using, including the characterization of separability using trace in positive characteristic.

We next prove a preliminary result of great importance in its own right.

Theorem. *Let R be a normal domain with fraction field $\mathcal{K}$, and let $\mathcal{L}$ be a finite algebraic extension of $\mathcal{K}$. Let $s \in \mathcal{L}$ be integral over R. Multiplication by s defines a $\mathcal{K}$-linear map of $\mathcal{L}$ to itself. The coefficients of the characteristic polynomial of this $\mathcal{K}$-linear map are in R. In particular, $\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(s) \in R$.*

Proof. We first consider the case where $\mathcal{L} = \mathcal{K}[s]$. Let f be the minimal polynomial of s over $\mathcal{K}$, which has degree d. We showed earlier that f has all of its coefficients in R: this is the first Proposition in the Lecture Notes from Octboer 1. Suppose that f has degree d. Then $[\mathcal{L} : \mathcal{K}] = d$, and the characteristic polynomial of the matrix of multiplication by s has degree d. Since the matrix satisfies this polynomial, so does s. It follows that the characteristic polynomial is equal to the minimal polynomial of s over $\mathcal{K}$.

In the general case, let $\mathcal{L}_0 = \mathcal{K}[s] \subseteq L$, and let $v_1, \ldots, v_d$ be a basis for $\mathcal{L}_0$ over $\mathcal{K}$, and let $w_1, \ldots, w_h$ be a basis for $\mathcal{L}/\mathcal{L}_0$. Let A be the matrix of multiplication by s on $\mathcal{L}_0$ with respect to $v_1, \ldots, v_d$. Then the span of $v_1 w_j, \ldots, v_d w_j$ is $\mathcal{L}_0 w_j$ and is stable under multiplication by s, whose matrix with respect to this basis is also A. Therefore, the matrix of multiplication by s with respect to the basis

$$v_1 w_1, v_2 w_1, \ldots, v_d w_1, \ldots, v_1 w_h, v_2 w_1, \ldots, v_d w_h$$

is the direct sum of h copies of A, and its characteristic polynomial is f^h, where f is the characteristic polynomial of A. We already know that f has coefficients in R. □

Corollary. *Let R be a normal domain that contains $\mathbb{Q}$, and let S be a module-finite extension of R. Then R is a direct summand of S as an R-module. Hence, for every ideal I of R, $IS \cap R = I$.*

Proof. $R - \{0\}$ is a multiplicative system in S, and so there is a prime ideal P of S disjoint from $R - \{0\}$. Then R embeds in the domain S/P, which is still module-finite over R. It suffices to show that $R \hookrightarrow S/P$ splits, for if $\phi : S/P \to R$ is R-linear and restricts to the identity map on R, then the composition of ϕ with $S \twoheadrightarrow S/P$ will be an R-linear map $S \to R$ that restricts to the identity on R. Thus, we have reduced to the case where S is a module-finite extension domain of R. Let $\mathcal{K}$ and $\mathcal{L}$ be the fraction fields of R and S, respectively, and let $n = [\mathcal{L} : \mathcal{K}]$. Then $(1/n)\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}$, when restricted to S, takes values in R (by the preceding Theorem), is R-linear, and is the identity when restricted to R. □

Proof of the Theorem. Consider $\mathcal{K} \otimes_R S = (R - \{0\})^{-1}S \subseteq (S - \{0\})^{-1}S = \mathcal{L}$. This domain is module-finite over $\mathcal{K}$ and so has dimension 0. Therefore, it is a field containing S, and so must be $\mathcal{L}$. It follows that every element of $\mathcal{L}$ can multiplied in S by an element of $R - \{0\}$. Choose a basis for $\mathcal{L}$ over $\mathcal{K}$, and multiply each basis element by a nonzero element of R so as to get a basis for $\mathcal{L}$ over $\mathcal{K}$ consisting of elements of S. Call this basis $b_1, \ldots, b_n$, where $n = [\mathcal{L} : \mathcal{K}]$. Because the field extension is separable, the matrix $A = \big(\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(b_ib_j)\big)$ is invertible. By the preceding theorem, each entry of this matrix is in R, and so the determinant D of this matrix is a nonzero element of R. We shall prove that $DS \subseteq Rb_1 + \cdots + Rb_n = G$. Since R is a Noetherian ring, G is a Noetherian R-module, and this implies that DS is a Noetherian R-module. But $S \cong DS$ as R-modules, via $s \mapsto Ds$.

It remains to show that $DS \subseteq R$. Let $s \in S$. Then $s \in L$ and so can be written uniquely in the form $\alpha_1 b_1 + \cdots + \alpha_n b_n$. We may multiply by $b_i \in S$ and take the trace of both sides:

$$\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(sb_i) = \sum_{j=1}^{n} \alpha_j \mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(b_ib_j),$$

Let $r_i = \mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(sb_i)$, let W be the column vector whose entries are the r_i (which are in R, by the preceding Theorem), and let V be the column vector whose entries are the α_j. Then $W = AV$, where A and W have entries in R. Let B be the classical adjoint of A, i.e., the transpose of the matrix of cofactors. Then B also has entries in R, and $BA = D(I)$, where I is the size n identity matrix. It follows that $BW = BAV = DV$, so that each $D\alpha_j$ is in R. But then $Ds = (D\alpha_1)b_1 + \cdots + (D\alpha_n)b_n \in G$, as required. □

Corollary. *Let D be any Dedekind domain whose fraction field $\mathcal{K}$ has characteristic 0, such as the integers. Let $\mathcal{L}$ be a finite algebraic extension of $\mathcal{K}$. Then the integral closure of D in $\mathcal{L}$ is module-finite over $\mathcal{K}$, and is a Dedekind domain.*

Proof. It is module-finite by the Theorem we just proved, and therefore Noetherian. It is normal by construction,' and one-dimensional because it is an integral extension of a ring of Krull dimension one. □

We next backtrack and review some facts about bilinear forms. Let $\mathcal{K}$ be a field and V a vector space of finite dimension n over $\mathcal{K}$. A bilinear form is simply a bilinear map $B : V \times V \to \mathcal{K}$, and giving V is the same a giving a linear map $T : V \otimes V \to K$. If $v_1, \ldots, v_n$ is a basis for V, then the elements $v_i \otimes v_j$ are a basis for $V \otimes_{\mathcal{K}} V$, and so B is completely determined by the matrix $A = (B(v_i,\, v_j) = T(v_i \otimes v_j)$. If the matrix $A = (a_{ij})$. Suppose that we use this basis to identify V with $\mathcal{K}^n$, with standard basis $e_1, \ldots, e_n$, then $B(e_i, e_j) = a_{ij}$. If v and w are the $n \times 1$ column matrices with entries $c_1, \ldots, c_n$ and $d_1, \ldots, d_n$, respectively, then

$$B(v,\, w) = B(\sum_i c_i e_i, \sum_j d_j e_j) = \sum_{i,j} c_i d_j B(e_i,\, e_j) = \sum_{i,j} c_i a_{ij} d_j = \sum_i c_i (\sum_j a_{ij} d_j)$$

which is the unique entry of the 1×1 matrix $v^{\mathrm{tr}}\ Aw$.

To see the effect of change of basis, let C be an $n \times n$ matrix whose columns $w_1, \ldots, w_n$ are a possibly new basis for $V = K^n$. Then $w_i^{\mathrm{tr}}\ Aw_j$ is the i, j entry of the matrix $C^{\mathrm{tr}}\ AC$ (which is called *congruent* or *cogredient* to A). The invertibility of A is unaffected by the choice of basis. If A is invertible, the bilinear form is called *non-degenerate*.

Let B be a bilinear form and fix a basis $v_1, \ldots, v_n$ for V. Let V^* be the dual vector space. Then B gives a linear map $L : V \to V^*$ by the rule $L(v)(w) = B(v,\, w)$. fix a basis $v_1, \ldots, v_n$ for V. There is a *dual basis* for the dual vector space V^* of V, whose i th element f_i is the linear functional whose value on v_i is 1 and whose value on v_j is 0 for $j \neq i$. Since the value of $L(v_i)$ on $w = \sum_j c_j v_j$ is

$$B(v_i, \sum_j c_j v_j) = \sum_j c_j B(v_i, v_j) = \sum_j B(v_i, v_j) f_j(w),$$

we have that $L(v_i) = \sum_j B(v_i, v_j) f_j$. Thus, the matrix of B with respect to $c_1, \ldots, c_n$ is the same as the matrix of L with respect to the two bases $v_1, \ldots, v_n$ and $f_1, \ldots, f_n$. Hence, the matrix of B is invertible if and only if $L : V \to V^*$ is an isomorphism. This shows that B is non-degenerate if and only if L is one-to-one, which means that B is non-degenerate if and only if for all $v \in V - \{0\}$ there exists $w \in V$ such that $L(v,\, w) \neq 0$.

B is called *symmetric* if $B(v,\, w) = B(w,\, v)$ for all $v,\, w \in V$, and this holds if and only if its matrix A is symmetric.

We next give some further discussion of the notion of trace, and prove the trace characterization of separability discussed earlier.

Let R be any ring and $F \cong R^n$ a free R-module. Consider any R-linear endomorphism $T : F \to F$. We define the *trace* of T as follows: choose a free basis for F, let $M = (r_{ij})$ be a matrix for T, and let $\mathrm{Tr}(T)$ be the sum $\sum_{i=1}^n r_{ii}$ of the entries on the main diagonal of M. This is independent of the choice of free basis for F: if one has another free basis, the new matrix has the form AMA^{-1} for some invertible $n \times n$ matrix A over R, and the trace is unaffected.

If $S \neq 0$ is a free R-algebra that has finite rank as an R-module, so that $S \cong R^n$ as an R-module for some positive integer n, then for every element $s \in S$ we define $\mathrm{Tr}_{S/R} s$ to be the trace of the R-linear endomorphism of S given by multiplication by s. Then $\mathrm{Tr}_{S/R} : S \to R$ is an R-linear map. If $r \in R$, $\mathrm{Tr}_{S/R}(r) = nr$, since the matrix of multiplication by r is r times the $n \times n$ identity matrix. We are mainly interested in the case where R and S are both fields. We first note:

Lemma. *If T is a free S-algebra of finite rank $m \geq 1$ and S is free R-algebra of finite rank $n \geq 1$, then $\mathrm{Tr}_{T/R}$ is the composition $\mathrm{Tr}_{S/R} \circ \mathrm{Tr}_{T/S}$.*

Proof. Let $u_1, \ldots, u_n$ be a free basis for S over R, and let $v_1, \ldots, v_m$ be a free basis for $\frac{T}{S}$. Let $A = (s_{ij})$ be the $m \times m$ matrix over S for multiplication by $t \in T$ with respect to the free basis $v_1, \ldots, v_m$ over S. Let B_{ij} be the $n \times n$ matrix over R for multiplication by s_{ij} acting on S with respect to the basis $u_1, \ldots, u_n$ for S over R. Then

$$t(u_h v_k) = u_h(tv_k) = u_h(\sum_j s_{jk} v_k) = \sum_j (s_{jk} u_h) v_k$$

and $s_{jk} u_h$ is the dot product of the h column of B_{ij} with the column whose entries are $u_1, \ldots, u_n$. It follows that a matrix for multiplication by t acting on T over R with respect to the basis $u_h v_k$ is obtained, in block form, from (s_{ij}) by replacing the i, j entry by the block B_{ij}. Then $\mathrm{Tr}_{T/R}(t)$ is the sum of the diagonal entries of this matrix, which is sum over i of the sums of the diagonals of the matrices B_{ii}. Now, $Tr_{T/S}(t)$ is the sum of the s_{ii}, and when we apply $Tr_{S/R}$ be get the sum over i of the elements $\tau_i = Tr_{S/R}(s_{ii}$. But τ_i is the same as the sum of diagonal elements in B_{ii}, and the result follows. □

Theorem. *Let cL be a finite algebraic extension field of $\mathcal{K}$. Then the extension is separable if and only if there is a (nonzero) element $\lambda \in \mathcal{L}$ such that $\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(\lambda) \neq 0$.*

Proof. We have already observed that the trace of 1 is $n = [\mathcal{L} : \mathcal{K}]$ which will be nonzero if $\mathcal{K}$ has characteristic 0, and every finite algebraic extension is separable in characteristic 0. Now suppose that $\mathcal{K}$ (and, hence, $\mathcal{L}$) have positive prime characteristic p.

If the extension if not separable, let $\mathcal{F}$ be the largest separable extension of $\mathcal{K}$ within $\mathcal{L}$. Since we must have an element $\theta \in \mathcal{L}$ such that $\theta^p \in \mathcal{F}$ but $\theta \notin \mathcal{F}$. Let $\mathcal{G}$ be the field $\mathcal{F}[\theta]$. Since

$$\mathrm{Tr}_{\mathcal{L}/\mathcal{K}} = \mathrm{Tr}_{\mathcal{F}/\mathcal{K}} \circ \mathrm{Tr}_{\mathcal{G}/\mathcal{F}} \circ \mathrm{Tr}_{\mathcal{L}/\mathcal{G}}$$

it will suffice to show that $\mathrm{Tr}_{\mathcal{G}/\mathcal{F}}$ vanishes identically. We have therefore reduced to the case where $\mathcal{L}$ is purely inseparable over $\mathcal{K}$, generated by a single element θ such that $\theta^p \in cK$. For an element $c \in \mathcal{K}$, $\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(c) = pc = 0$. For an element $\lambda \in \mathcal{L} - \mathcal{K}$, we have that $\lambda^p = c \in \mathcal{K}$. Since $[\mathcal{L} : \mathcal{K}] = p$ is prime, there are no strictly intermediate fields, and so $\mathcal{K}[\lambda] = \mathcal{L}$, and λ has degree p over $\mathcal{K}$. It follows that the minimal polynomial of λ over $\mathcal{K}$ is $x^p - c$, and that the elements λ^t, $0 \leq t \leq p-1$, are a basis for $\mathcal{L}$ over $\mathcal{K}$. Multiplication by λ. maps each basis vector to the next, except for λ^{p-1}, which is mapped to $c \cdot 1$. The matrix for multiplication by λ therefore has only zero entries on the main diagonal, and so $Tr_{\mathcal{L}/\mathcal{K}}(\lambda) = 0$, as required. (The matrix has a string of entries equal to one just below the main diagonal, and the element c occurs in the upper right hand corner. All other entries are 0.)

It remains to show that if $\mathcal{L}/\mathcal{K}$ is separable, then some element has trace different from 0. By the theorem on the primitive element, we may assume that $\mathcal{L} = \mathcal{K}[\theta]$. (Even without knowing this theorem, we can think of $\mathcal{L}$ as obtained from $\mathcal{K}$ by a finite sequence of field extensions, each of which consists of adjoining just one element, and so reduce to the case where one has a primitive element.) Let f be the minimal polynomial of θ: the hypothesis of separability implies that the roots of f are n distinct elements of the algebraic closure $\overline{\mathcal{L}}$ of $\mathcal{L}$: call them $\theta_1, \, \ldots \, , \theta_n$. Let A be the matrix for multiplication by θ with respect to some basis for $\mathcal{L}$ over $\mathcal{K}$. Then for every t, A^t gives a matrix for multiplication by θ^t. We shall show that for some i, $0 \leq i \leq n-1$, $\mathrm{Tr}_{cL/\mathcal{K}}(\theta^i) \neq 0$. Assume otherwise.

Since A satisfies its characteristic polynomial, call it g, which is monic of degree n, θ satisfies g. Thus, $f \,|\, g$. Since f and g are monic of the same degree, $g = f$. Thus, the eigenvalues of A are distinct: they are the elements θ_j. Therefore, A is similar over $\overline{(}\mathcal{L})$ to diagonal matrix with the θ_j on the diagonal, and it follows that, for every i, A^i is similar to a diagonal matrix with the entries θ_j^i on the diagonal. Therefore,

$$\mathrm{Tr}_{\mathcal{L}/\mathcal{K}}(\theta^i) = \sum_{j=1}^{n} \theta_j^i = 0.$$

Thus, the sum of the columns of the matrix $\Theta = (\theta_j^{i-1})$ is 0, which implies that the determinant is 0. We conclude the proof by showing that the determinant cannot be zero. (This is the well-known Van der Monde determinant, and its value can be shown to be the product of the $\binom{n}{2}$ differences $\theta_j - \theta_i$ for $j > i$. It will not vanish because the θ_j are distinct. But we argue differently, without assuming this.) If the determinant is 0 there is an $\overline{\mathcal{L}}$-linear relation on the rows as well: suppose that $\gamma = (c_0 \;\; c_1 \;\; \ldots \;\; c_n)$ is a vector such that $\gamma\Theta = 0$, giving a relation on the rows. This simply says that for every j,

$$\sum_{i=0}^{n-1} c_i \theta_j^i = 0.$$

But if

$$h(x) = c_0 + c_1 x + \cdots + c_{n-1} x^{n-1},$$

this says that all of the θ_j are roots of $h(x)$, a polynomial of degree at most $n-1$. This is a contradiction unless all of the c_i are 0. $\square$

This completes our treatment of separability and trace.

Lecture of December 8

It is a fact that in any Dedekind domain R, $fg \in (f^2, g^2)R$. This has been left as an exercise. It is also true that in the polynomial ring in n variables over a field K, for any $n+1$ elements $f_1, \ldots, f_{n+1}$, we have that $(f_1 \cdots f_{n+1})^n \in (f_1^{n+1}, \ldots, f_n^{n+1})R$. In particular, in $R = \mathbb{C}[x, y]$, for any three elements f, g, h, we have that $(fgh)^2 \in (f^3, g^3, h^3)R$. I know of three proofs of this, all involving some difficult ideas. It would be of great interest to find an elementary proof, even in the case of $\mathbb{C}[x, y]$.

Here is another example where indeterminates cannot be canceled: let

$$R = \mathbb{C}[x,y,z]/\bigl(xy - (1-z^2)\bigr)$$

and let

$$S = \mathbb{C}[x,y,z]/\bigl(x^2y - (1-z^2)\bigr).$$

There is no obvious reason why R and S should be isomorphic, and they are not (although it is not easy to prove this). It may come as a surprise that $R[t] \cong S[t]$: they do become isomorphic when an indeterminate is adjoined. These are called *Danielewski surfaces*. Cf. [W. Danielewski, *On the cancellation problem and automorphism groups of affine algebraic varieties*, preprint, Warsaw, 1989] and [K.-H. Fieseler, *On complex affine surfaces with* $\mathbf{C}^+$*-action*, Comment. Math. Helv. **69** (1994), 5–27].

A Noetherian ring with only finitely many maximal ideals is called *semilocal*. (The term *quasisemilocal* is used for rings with finitely many maximal ideals if they need not be Noetherian.) Given finitely many mutually incomparable primes $P_1, \ldots, P_k$ of a ring R, if $W = R - \bigcap_{j=1}^k P_j$, then W is a multiplicative system in R. The ring $W^{-1}R$ has as its maximal ideals precisely the k ideals $P_jW^{-1}R$. Thus, if R is Noetherian, it is semi-local. It is referred to as the *localization* of R at the primes $P_1, \ldots, P_k$.

Theorem. *Let R be a Dedekind domain. Let M be a finitely generated R-module.*

(a) *If M is torsion-free, it is projective. In particular, every ideal of R is projective, and the product IJ of two ideals is $\cong I \otimes J$ as a module, and so its isomorphism class as a module depends only on the isomorphism classes of I and J.*

(b) *R is a UFD if and only if R is a PID.*

(c) *If R is semi-local, then R is a PID.*

(d) *Given finitely many maximal ideals $P_1, \ldots, P_k$ of R and an ideal $I \neq 0$, I is isomorphic with an ideal not contained in any of the P_i.*

(e) *M is a direct sum of a torsion module and a torsion-free module. The torsion submodule N is unique and may be viewed as a module over the localization of R at the set of finitely many maximal ideals in its support: the localization is a PID and the theory of modules over a PID applies. Thus, N is a direct sum of cyclic modules.*

(f) *If I and J are nonzero ideals of R, then $I \oplus J \cong I \cap J \oplus (I + J) \cong IJ \oplus R$, and if $I_1, \ldots, I_n$ are nonzero ideals then $I_1 \oplus \cdots \oplus I_n \cong (I_1 \cdots I_n) \oplus R^{n-1}$.*

(g) *If N is the torsion submodule of M, the torsion-free summand of M is isomorphic with M/N. Any finitely generated torsion-free R-module M is the direct sum of a free*

R-module R^{n-1} and an ideal $I \subseteq R$. The integer n is uniquely determined, and I is uniquely determined up to module isomorphism.

Proof. (a) Projective is equivalent to locally free. Locally, R is a DVR, and every finitely generated torsion-free module is free, since a DVR is a PID. When we apply $I \otimes_R \_$ to the injection $J \subseteq R$ we find that $I \otimes_R J \hookrightarrow I \otimes_R R \cong I$: the map is injective because I is projective and, therefore, flat. The image of this map is IJ.

(b) To prove "only if," note that in a UFD, height one primes are principal. Every maximal ideal of R is therefore principal, and every nonzero proper ideal is a finite product of powers of maximal ideals and so principal. But the "if" part is clear, since a PID is a UFD.

(c) Let $m = m_1, m_2, \ldots, m_k$ be the maximal ideals: it suffices to show that each is principal. m is not contained in any of $m^2, m_2, \ldots, m_k$: choose $x \in m$ not in the union of these. Then $xR = m$, because that is true if we localize at any m_i. If $i = 1$, this is because $x \in m - m^2$, and so $x \notin m^2 R_m$, since $m^2 R_m \cap R = m^{(2)} = m^2$, since m is maximal. For any other m_i, $xR_{m_i} = mR_{m_i}$: in fact, both are the unit ideal.

(d) After localization at $P_1, \ldots, P_k$, $IW^{-1}R$ becomes principal: we can choose $b \in I$ that generates. Thus, there exists w not in any P_k such that $wI \subseteq bR \subseteq I$, and so $J = (w/b)I \subseteq R$ is isomorphic with I as an R-module. If $J \subseteq P_i$, then $wI \subseteq bP_i$. When we localize at P_i, w becomes invertible, yielding $IR_{P_i} \subseteq bP_i R_{P_i}$. Since $IR_{P_i} = bR_{P_i}$, we have that $bR_{P_i} \subseteq bP_i R_{P_i}$, and so $R_{P_i} \subseteq P_i R_{P_i}$, a contradiction.

(e) The torsion submodule consists of all torsion elements in M and so is obviously unique. M/N is clearly torsion-free and so projective. Thus, $0 \to N \to M \to M/N \to 0$ splits. Choose finitely many generators for N. Each has a nonzero annihilator, and, hence, so does N. We may view N as a module over $R/\mathfrak{A}$, where $\mathfrak{A} = \mathrm{Ann}_R N$, and this is a zero-dimensional Noetherian ring. Clearly, it has only finitely many maximal ideals coming from maximal ideals $P_1, \ldots, P_k$ of R. Any element not in the union of the P_j acts invertibly on N, and so N is a module over the localization of R at $P_1, \ldots, P_k$.

(f) There is an exact sequence $0 \to I \cap J \to I \oplus J \to I + J \to 0$ where the map $I \oplus J \twoheadrightarrow I + J$ sends $i \oplus j$ to $i - j$, and then map $I \cap J \to I \oplus J$ sends u to $u \oplus u$. Since $I + J$ is projective, the sequence is split exact, which shows that $I \oplus J \cong (I \cap J) \oplus (I + J)$. By (d), we can choose I', an ideal isomorphic with I as a module, but such that I' is not contained in any of the finitely many minimal primes of J. This means that I' and J are comaximal, i.e., $I' + J = R$, and then $I' \cap J = I'J$. Thus, $I \oplus J \cong I' \oplus J \cong (I' \cap J) \oplus (I' + J) = I'J \oplus R \cong (I' \otimes_R J) \oplus R \cong (I \otimes_R J) \oplus R \cong IJ \oplus R$, as required. The final statement is a straightforward induction on n, using the result just proved.

(g) Let $\mathcal{K} = \mathrm{frac}\,(R)$. Then $\mathcal{K} \otimes_R M \cong \mathcal{K}^n$ and we can therefore choose a nonzero map from $\mathcal{K} \otimes_R M$ onto $\mathcal{K}$. This is an element of $\mathrm{Hom}_{\mathcal{K}}(\mathcal{K} \otimes_R M, \mathcal{K} \otimes R) \cong \mathcal{K} \otimes_R \mathrm{Hom}_R(M, R)$, and so there must exist a nonzero R-linear map $f : M \to R$. Call the image I, which is an ideal of R. Then I is projective. The map $M \twoheadrightarrow I$ therefore splits, and $M \cong M_0 \oplus I$. Iterating, we see that M is a direct sum of ideals, $I_1 \oplus \cdots \oplus I_n$. By (f), this direct sum is isomorphic with $R^{n-1} \oplus (I_1 \cdots I_n)$. The integer n is the torsion-free rank of M, i.e., the vector space dimension over $\mathcal{K}$ of $\mathcal{K} \otimes_R M$. It remains only to see that the module isomorphism class of the ideal I is unique, which follows from the Lemma given immediately following. □

Lemma. *Let R be a Noetherian ring and let P, P' be finitely generated modules that are locally free of rank one. Suppose that $M = R^{n-1} \oplus P \cong R^{n-1} \oplus P'$. Then $P \cong P'$.*

This immediately yields the fact needed to complete the proof of part (g) of the preceding theorem.

The Lemma is proved by showing that $P \cong \bigwedge_R^n M \cong P'$. Before giving the details, we review the properties of exterior powers.

A multilinear map of R-modules $M^n \to W$ is called *alternate* or *alternating* if its value is 0 whenever two entries of an n-tuple are equal. (This implies that switching two entries negates the value. Making an even permutation of the entries will not change the value, while an odd permutation negates the value.) Let $\bigwedge_R^n(M) = \bigwedge^n(M)$ denote the quotient of $M^{\otimes n}$ by the submodule spanned by all n-tuples two of whose entries are equal. We make the convention that $\bigwedge^0 M \cong R$, and note that we may identify $M \cong \bigwedge^1 M$. Then $\bigwedge M = \bigoplus_n \bigwedge^n M$ is an associative $\mathbb{N}$-graded algebra with R in the center, with $\bigwedge^n(V)$ as the component in degree n. $\bigwedge(V)$ is called the *exterior algebra* of M over R, and $\bigwedge^n(M)$ is called the n th *exterior power* of M over k. The multiplication on $\bigwedge(M)$ is often denoted $\wedge$. If the elements u_i span M, then the elements $u_{i_1} \wedge \cdots \wedge u_{i_n}$ span $\bigwedge^n(V)$. If α has degree m and β has degree n, then $\alpha \wedge \beta = (-1)^{mn} \beta \wedge \alpha$. Thus, the even degree elements are all in the center, while any two odd degree elements anti-commute. If $u_1, \ldots, u_n$ is a free basis for M, then the elements $u_{i_1} \wedge \cdots \wedge u_{i_k}$, $1 \le i_1 < \cdots < i_k \le n$ form a free basis for $\bigwedge^k(M)$, and $\bigwedge^k(M)$ has dimension $\binom{n}{k}$. In particular, $\bigwedge^N(M) = 0$ if N exceeds the rank of the free module M (more generally, $\bigwedge^N M = 0$ whenever M is spanned by fewer than N elements).

If $f : M \to N$ is R-linear, there is an induced map $\bigwedge^n(f) : \bigwedge^n(M) \to \bigwedge^n(N)$, and $\bigwedge^n(f' \circ f) = \bigwedge^n(f') \circ \bigwedge^n(f)$ when the composition $f' \circ f$ is defined. Together these maps give a ring homomorphism of $\bigwedge(M) \to \bigwedge(N)$ that preserves degrees. Thus, $\bigwedge(\_)$ is a functor from R-modules to skew-commutative associative graded R-algebras, and every $\bigwedge^i(\_)$ is a covariant functor from R-modules to R-modules.

If M is free of rank n with basis $v_1, \ldots, v_n$ and $f : M \to M$ has matrix $A = (a_{ij})$, then $\bigwedge^n(f) : M \to M$ sends $v_1 \wedge \cdots \wedge v_n$ to $\det(A) v_1 \wedge \cdots \wedge v_n$. (We have that

$$\bigwedge^n(f)(v_1 \wedge \cdots \wedge v_n) = (a_{11}v_1 + \cdots + a_{1n}v_n) \wedge \cdots \wedge (a_{n1}v_1 + \cdots + a_{nn}v_n).$$

Expanding by the generalized distributive law yields n^n terms each of which has the form $a_{i_1,1} \cdots a_{i_n,n} v_{i_1} \wedge \cdots \wedge v_{i_n}$. If two of the i_t are equal, this term is 0. If they are all distinct, the v_{i_t} constitute all the elements $v_1, \ldots, v_n$ in some order: call the corresponding permutation σ. Rearranging the v_j gives $\operatorname{sgn}(\sigma) a_{i_1,1} \cdots a_{i_n,n} v_1 \wedge \cdots \wedge v_n$. The sum of all of the $n!$ surviving terms is $\det(A) v_1 \wedge \cdots \wedge v_n$, using one of the standard definitions of $\det(A)$). The fact that the determinant of a product of two $n \times n$ matrices is the product of the determinants may consequently be deduced from the fact that $\bigwedge^n$ preserves composition.

We note that if M and N are any two R-modules then there is a canonical isomorphism

$$\theta : \bigwedge^n(M \oplus N) \cong \bigoplus_{i+j=n} \bigwedge^i M \otimes_R \bigwedge^j N = W.$$

Here, i, j are restricted to be nonnegative integers. This isomorphism is suggested by the fact that $(y_1 \oplus z_1) \wedge \cdots \wedge (y_n \oplus z_n)$, where the y_t are in M and the z_t are in N, expands as the sum of 2^n terms of the form $u_1 \wedge \cdots \wedge u_n$, and in one of these terms, if $u_{t_1}, \ldots, u_{t_i}$ are from M and $u_{t'_1}, \ldots, u_{t'_j}$ are from N, the term can be rewritten as $\mathrm{sgn}\,(\sigma)(u_{t_1} \wedge \cdots \wedge u_{t_i}) \wedge (u_{t'_1} \wedge \cdots \wedge u_{t'_j})$, where σ is the permutation of $\{1, \ldots, n\}$ whose values on $1, \ldots, n$ are $t_1, \ldots, t_i, t'_1, \ldots, t'_j$. Thus, to construct θ, we give a multilinear map $(M \oplus N)^n \to W$ as follows. It is equivalent to give a linear map $(M \oplus N)^{\otimes n} \to W$, and $(M \oplus N)^{\otimes n}$ is the direct sum of 2^n terms of the form $U_1 \otimes \cdots \otimes U_n$ where every U_t is either M or N. It suffices to give a multilinear map on every $U_1 \times \cdots \times U_n$ to W. Suppose that we have $U_{t_1} = \cdots = U_{t_i} = M$, where $t_1 < \cdots < t_i$, and that we likewise have $U_{t'_1} = \cdots = U_{t'_j} = N$, where $t'_1 < \cdots < t'_j$. Here $i + j = N$. Let σ be the permutation whose values on $1, \ldots, n$ are $t_1, \ldots, t_i, t'_1, \ldots, t'_j$, as above. Then our map will send $(u_1, \ldots, u_n)$ to $\mathrm{sgn}\,(\sigma)(u_{t_1} \wedge \cdots \wedge u_{t_i}) \otimes (u_{t'_1} \wedge \cdots \wedge u_{t'_j})$. The direct sum of all these determines a multilinear map $(M \oplus N)^n \to W$, and it is straightforward to check that it is alternating, and so induces a map $\theta : \bigwedge^n(M \oplus N) \to W$.

Note that there is a multilinear map $M^i \times N^j \to \bigwedge^n(M \oplus N)$ such that

$$(u_1, \ldots, u_i, v_1, \ldots, v_j) \mapsto u_1 \wedge \cdots \wedge u_i \wedge v_1 \wedge \cdots \wedge v_j,$$

and so we have a map

$$M^{\otimes i} \otimes_R N^{\otimes j} \to \bigwedge^n(M \oplus N) = W.$$

It is easy to verify that this map factors through $\bigwedge^i(M) \otimes_R \bigwedge^j(N)$, and the direct sum of all these maps gives an inverse ϕ for θ: that these maps are mutually inverse is easy to check on suitable generators: these can be taken to be of the form $m_{t_1} \wedge \cdots \wedge m_{t_i} \otimes n_{t'_1} \wedge \cdots \wedge n_{t'_j}$ for the left hand module and of the form $m_{t_1} \wedge \cdots \wedge m_{t_i} \wedge n_{t'_1} \wedge \cdots \wedge n_{t'_j}$ for the right hand module.

The following result can now be used to complete the proof of the Lemma above.

Proposition. *Let $R \to S$ be a map of rings and let M be an R-module.*

(a) *For all i, $\bigwedge^i_S(S \otimes_R M) \cong S \otimes \bigwedge^i_R(M)$ in such a way that $(s_1 \otimes m_1) \wedge \cdots \wedge (s_i \otimes m_i)$ corresponds to $(s_1 \cdots s_i) \otimes (m_1 \wedge \cdots \wedge m_i)$. In particular, we may take S to be a localization of R, and, in this sense, exterior powers commute with localization.*

(b) *If for every prime (or maximal) ideal P of R, M_P has at most n generators over R_P, then $\bigwedge^i(M) = 0$ for $i > n$. In particular, if M is locally free of rank n, then $\bigwedge^i(M) = 0$ for $i > n$.*

(c) *If R is Noetherian and M is a finitely generated projective module, then every $\bigwedge^i(M)$ is a finitely generated projective module. If M is locally free of constant rank n, then for $0 \le i \le n$, $\bigwedge^i(M)$ is locally free of constant rank $\binom{n}{i}$.*

Proof. (a) There is an R-multilinear map $(S \times M)^i \to S \otimes_R \bigwedge^i_R(M)$ that sends the element $\big((s_1, m_1), \ldots, (s_i, m_i)\big) \mapsto (s_1 \cdots s_i) \otimes m_1 \wedge \cdots \wedge m_i$. This yields an R-linear map from $(S \otimes_R M)^i$ to $S \otimes_R \bigwedge^i(M)$ which is easily checked to be both S-multilinear and alternating, and so we have a map $\bigwedge_S^i(S \otimes_R M) \to S \otimes_R \bigwedge_R^i(M)$. On the other hand, there is an R-multilinear map $S \times M^i \to \bigwedge_S^i(S \otimes_R M)$ that sends the element $(s, m_1, \ldots, m_i)$ to $s\big((1 \otimes m_1) \wedge \cdots (1 \otimes m_i)\big)$, which induces an R-bilinear map $S \times \bigwedge_R^i M \to \bigwedge_S^i(S \otimes_R M)$ (for each fixed $s \in S$, the map one gets on M^i is alternating), and hence an R-linear map $S \otimes_R \bigwedge_R^i M \to \bigwedge_S^i(S \otimes_R M)$. But this map is easily checked to be S-linear. Moreover, this map and the map $\bigwedge_S^i(S \otimes_R M) \cong S \otimes \bigwedge_R^i(M)$ constructed earlier are readily checked to be inverses.

(b) The issue of whether a module is zero can be checked locally, and the hypothesis implies that all localizations of $\bigwedge^i(M)$ are 0.

(c) For finitely generated modules over a Noetherian ring, projective is equivalent to locally free, and so the statements reduce to the known case where the module is free. □

Proof of the Lemma. It suffices to prove that $\bigwedge^n(P \oplus R^{n-1}) \cong P$. This is the direct sum of terms $\bigwedge^i P \otimes_R \bigwedge^j(R^{n-1})$, where $i, j \geq 0$ and $i + j = n$. The term for $i = 0, j = n$ vanishes because $\bigwedge^n(R^{n-1}) = 0$. The terms for $i > 1$ vanish because $\wedge^i(P) = 0$, since P is locally free of rank one. The only summand that might not vanish is therefore $\bigwedge_R^1(P) \otimes_R \bigwedge_R^{n-1}(R^{n-1}) \cong P \otimes_R R \cong P$, as required. □

Recall that a partially ordered set $(\Lambda, \leq)$ is called *directed* if for any two elements λ, $\mu \in \Lambda$, there exists $\nu \in \Lambda$ with $\lambda \leq \nu$ and $\mu \leq \nu$. That is, any two elements of Λ have a common upper bound. Examples include any totally ordered set, the finite subsets of a given set under $\subseteq$, the finitely generated R-submodules of an R-module under $\subseteq$, and the finitely generated R-subalgebras of an R-algebra under $\subseteq$. Another example is given by the open neighborhoods of a point $x \in X$, where X is a topological space, under $\supseteq$. The nonnegative integers $\mathbb{N}$ and the positive integers are particularly important examples.

Recall that a partially ordered set $(\Lambda, \leq)$ becomes a category whose objects are the elements of Λ, and such that there is a morphism from λ to μ iff $\lambda \leq \mu$, in which case there is a unique morphism from λ to μ. By a *direct limit system* in a category $\mathcal{C}$ indexed by the partially ordered set Λ, we mean a covariant functor from Λ to $\mathcal{C}$. Explicitly, this means that for every element λ in Λ we have an object in $\mathcal{C}$, call it X_λ, and for all pairs λ, μ such that $\lambda \leq \mu$ a morphism $f_{\lambda,\mu} : X_\lambda \to X_\mu$ satisfying (1) every $f_{\lambda,\lambda}$ is the identity on X_λ and (2) whenever $\lambda \leq \mu \leq \nu$, we have that $f_{\lambda,\nu} = f_{\mu,\nu} \circ f_{\lambda,\mu}$.

By a *candidate* for the direct limit of a direct limit system we mean an object X together with a family of maps $g_\lambda : X_\lambda \to X$ for all $\lambda \in \Lambda$ such that whenever $\lambda \leq \mu$, $g_\mu = f_{\lambda,\mu} \circ g_\lambda$. (This can be expressed alternatively as follows. Given any object X we can construct a direct limit system in which the object assigned to every λ is X, and all the maps are the identity on X. We refer to this as a *one-object system.* A candidate for the direct limit is the same thing as a natural transformation of functors from the functor defining the system to a functor defining a one-object system.)

We say that a candidate $(Y,\, h_\lambda : X_\lambda \to Y)$ for the direct limit is the *direct limit* of the direct limit system if for every candidate X, $g_\lambda : X_\lambda \to X$ there is a unique morphism $k : Y \to X$ such that for all $\lambda \in \Lambda$, $g_\lambda = k \circ h_\lambda$. We write $Y = \varinjlim_\lambda X_\lambda$.

Lecture of December 11

Direct limits are automatically unique up to unique isomorphism compatible with their structures as candidates.

We focus on the categories of sets, groups, abelian groups, rings, R-modules, and R-algebras. In each case, there is an underlying set, and a morphism is a function possibly satisfying additional conditions. Consider an example where the objects are subobjects X_λ of a given object Z, and the maps are inclusion maps. The direct limit is simply the union of the subobjects, and is called a *directed union.*

Direct limits exist in general in the categories mentioned above. In the category of sets, one takes a disjoint union of the sets in the indexed family, and then for every $\lambda < \mu$ one identifies every $x \in X_\lambda$ with its image in X_μ. That is, one takes the smallest equivalence relation such that for $\lambda < \mu$, every element $x \in X_\lambda$ is equivalent to its image in X_μ, and then the direct limit is the set of equivalence classes. Every element in X_λ maps to its equivalence class.

If the sets have an additional structure such as group, abelian group, ring, R-module, or R-algebra, the same construction still works. To define the needed operations on the direct limit set, suppose, for example, that one wants to add or multiply two elements of the direct limit. They are images of elements from X_λ and X_μ for a certain λ and μ. These both map to elements in X_ν for some ν that is an upper bound for both λ and μ, and one can add or multiply these element in X_ν and then take the image in the direct limit.

In the case of abelian groups or R-modules, one can proceed alternatively by taking the direct sum over R of all the X_λ, and then killing the span of all elements of the form $f_{\lambda,\mu}(x) - x$, where $\lambda \leq \mu$ and $x \in X_\lambda$.

If X is a topological space, $x \in X$, Λ is the set of all open neighborhoods of x ordered by $\supseteq$, and R_U denotes the ring of all $\mathbb{R}$-valued continuous functions on U, we get a direct limit system if, whenever $U \supseteq V$, the map $R_U \to R_V$ is given by $f \mapsto f|_V$, the restriction of $f : U \to \mathbb{R}$ to V. then $\varinjlim_U R_U$ is the ring of germs of continuous functions at $x \in X$. If X is a C^∞ manifold or an analytic space, similar constructions lead to the rings of germs of C^∞ or analytic functions. In all these cases, the direct limit ring is a quasilocal ring.

We note that tensor product commutes with direct limit. Given a direct limit system of R-modules M_λ and an R-module N, we claim that there is an isomorphism

$$(\varinjlim_\lambda M_\lambda) \otimes_R N \cong \varinjlim_\lambda (M_\lambda \otimes_R N).$$

For each fixed $v \in N$ we have a map $f^v_\lambda : M_\lambda \to \varinjlim_\lambda (M_\lambda \otimes_R N)$ sending u to the image of $u \otimes v$. These induce a map $f^v : \varinjlim_\lambda M \to \varinjlim_\lambda (M_\lambda \otimes_R N)$, which gives an R-bilinear map $(\varinjlim_\lambda M_\lambda) \times N \to \varinjlim_\lambda (M_\lambda \otimes_R N)$ and, hence, we get an R-linear map $(\varinjlim_\lambda M_\lambda) \otimes_R N \to \varinjlim_\lambda (M_\lambda \otimes_R N)$. On the other hand, for each λ the map $M_\lambda \to \varinjlim_\lambda M_\lambda$ induces a map $M_\lambda \otimes N \to \varinjlim_\lambda M_\lambda \otimes_R N$, and this gives the required map

$$\varinjlim_\lambda (M_\lambda \otimes_R N) \to (\varinjlim_\lambda M_\lambda) \otimes_R N.$$

It is straightforward to check that these are mutually inverse.

Corollary. *A direct limit of flat R-modules is R-flat.*

Proof. Let $F = \varinjlim_\lambda F_\lambda$ where each F_λ is flat, and let $N \subseteq M$ be R-modules. Suppose that $u \in F \otimes_R N$ maps to 0 in $F \otimes_R M$. Then u is the image of $u_\lambda \in F_\lambda \otimes_R N$ for some λ. The fact that the image of u_λ in $F \otimes M \cong \varinjlim_\lambda (F_\lambda \otimes_R M)$ is 0 implies that it maps to 0 in $F_\mu \otimes_R M$ for some $\mu \geq \lambda$. Since the composite $F_\lambda \otimes_R N \to F_\mu \otimes_R N \to F_\mu \otimes_R M$ kills u_λ, while the right hand map is injective because F_μ is R-flat, it follows that the image of u_λ in $F_\mu \otimes N$ is 0. But that image maps to u, and so u is 0. □

Proposition. *Let R be a domain and F a flat R-module. Then F is torsion-free over R.*

Proof. Let $x \in R$ be nonzero. Then $0 \to R \xrightarrow{x} R$ is exact. Apply $\_ \otimes M$, we have that $0 \to M \xrightarrow{x} M$ is exact, i.e., that x is not a zerodivisor on M. □

More generally, if R is any ring, $x \in R$ is not a zerodivisor in R, and M is R-flat, then x is not a zerodivisor on M. We have the following consequence of the two preceding results.

Corollary. *A module F over a Dedekind domain R is flat if and only if it is torsion-free.*

Proof. If it is flat, it is torsion-free by the Proposition. Now suppose that it is torsion-free. It is the directed union of its finitely generated submodules, and so a direct limit of them. But a finitely generated module over a Dedekind domain is projective, and therefore flat. □

Let $(\Lambda, \leq)$ be a directed set. By an *inverse limit system* in a category $\mathcal{C}$ we mean a direct limit system in $\mathcal{C}^{\mathrm{op}}$. The notions of candidate for an inverse limit and inverse limit are then immediately given by applying the definitions for direct limit system to $\mathcal{C}^{\mathrm{op}}$. However, we briefly make all this more explicit. An inverse limit system consists of objects X_λ indexed by Λ and for all $\lambda \leq \mu$ a morphism $f_{\lambda,\mu} : X_\mu \to X_\lambda$. A candidate for the inverse limit consists of an object X together with maps $g_\lambda : X \to X_\lambda$ such that for all $\lambda \leq \mu$, $g_\lambda = f_{\lambda,\mu} \circ g_\mu$. A candidate Y together with morphisms $h_\lambda : Y \to X_\lambda$ is an inverse limit precisely if for every candidate (X, g_λ) there is a unique morphism $k : X \to Y$ such that for all λ, $g_\lambda = h_\lambda \circ k$. The inverse limit is denoted $\varprojlim_\lambda X_\lambda$ and, if it exists, it is unique up to canonical isomorphism compatible with the morphisms giving X and Y the structure of candidates.

We next want to see that inverse limits exist in the categories of sets, abelian groups, rings, R-modules, and R-algebras. The construction for sets also works in the other categories mentioned. Let $(\Lambda, \leq)$ be a directed partially ordered set and let $(X_\lambda, f_{\lambda,\mu})$ be an inverse limit system of sets. Consider the subset $X \subseteq \prod_\lambda X_\lambda$ consisting of all elements x of the product such that for $\lambda \leq \mu$, $f_{\lambda,\mu}(x_\mu) = x_\lambda$, where x_λ and x_μ are the λ and μ

coordinates, respectively, of x. It is straightforward to verify that X is an inverse limit for the system: the maps $X \to X_\lambda$ are obtained by composing the inclusion of X in the product with the product projections π_λ mapping the product to X_λ.

If each X_λ is in one of the categories specified above, notice that the Cartesian product is as well, and the set X is easily verified to be a subobject in the appropriate category. In every instance, it is straightforward to check that X is an inverse limit.

Suppose, for example, that X_λ is a family of subsets of A ordered by $\supseteq$, and that the map $X_\mu \to X_\lambda$ for $X_\lambda \supseteq X_\mu$ is the inclusion of $X_\mu \subseteq X_\lambda$. The condition for the partially ordered set to be directed is that for all λ and μ, there is a set in the family contained in $X_\lambda \cap X_\mu$. The construction for the inverse limit given above yields all functions on these sets with a constant value in the intersection of all of them. This set evidently may be identified with $\bigcap_\lambda X_\lambda$.

We are particularly interested in inverse limit systems indexed by $\mathbb{N}$. To give such a system one needs to give an infinite sequence of objects $X_0, X_1, X_2, \ldots$ in the category and for every $i \geq 0$ a map $X_{i+1} \to X_i$. The other maps needed can be obtained from these by composition. In the cases of the categories mentioned above, to give an element of the inverse limit is the same a giving a sequence of elements $x_0, x_1, x_2, \ldots$ such that for all i, $x_i \in X_i$, and x_{i+1} maps to x_i for all $i \geq 0$. One can attempt to construct an element of the inverse limit by choosing an element $x_0 \in X_0$, then choosing an element $x_1 \in X_1$ that maps to x_0, etc. If the maps are all surjective, then given $x_i \in X_i$ one can always find an element of the inverse limit that has x_i as its ith coordinate: for $h < i$, use the image of x_i in X_h, while for $i+1, i+2, \ldots$ one can choose values recursively, using the surjectivity of the maps.

We want to use these ideas to describe the I-adic completion of a ring R, where R is a ring and $I \subseteq R$ is an ideal. We give two alternative descriptions. Consider the set of all sequences of elements of R indexed by $\mathbb{N}$ under termwise addition under multiplication: this ring is the same as the product of a family of copies of R index by $\mathbb{N}$. Let $\mathfrak{C}_I(R)$ denote the subring of *Cauchy sequences for the I-adic topology*: by definition these are the sequences such that for all $t \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that for all i, $j \geq N$, $r_i - r_j \in I^t$. This is a subring of the ring of sequences. It is an R-algebra via the map $R \to \mathfrak{C}_I(R)$ that sends $r \in R$ to the constant sequence $r, r, r, \ldots$. Let $\mathfrak{C}_i^0(R)$ be the set of *Cauchy sequences that converge to 0*: by definition, these are the sequences such that for all $t \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that for all $i \geq N$, $r_i \in I^t$. These sequences are automatically Cauchy. Then $\mathfrak{C}_I^0(R)$ is an ideal of $\mathfrak{C}_I(R)$. It is easy to verify that every subsequence of a Cauchy sequence is again Cauchy, and that it differs from the original sequence by an element of $\mathfrak{C}_I^0(R)$.

Given an element of $\mathfrak{C}_I(R)$, say $r_0, r_1, r_2, \ldots$ we may consider the residue mod I^t for a given t. These are eventually all the same, by the definition of a Cauchy sequence. The stable value of these residues is an element of R/I^t, and we thus have a map $\mathfrak{C}_I(R) \twoheadrightarrow R/I^t$ that is easily seen to be a ring homomorphism that kills $\mathfrak{C}_I^0(R)$. Therefore, for all t we have a surjection $\mathfrak{C}_I(R)/\mathfrak{C}_I^0(R) \twoheadrightarrow R/I^t$. These maps make $\mathfrak{C}_I(R)/\mathfrak{C}_I^0(R)$ a candidate for $\varprojlim_t (R/I^t)$, and so induce a ring homomorphism $\mathfrak{C}_I(R)/\mathfrak{C}_i^0(R) \to \varprojlim_t R/I^t$.

This map is an isomorphism. Given a sequence of elements in the rings R/I^t that determine an element of the inverse limit, for each residue ρ_t choose an element r_t of R that represents it. It is straightforward to verify that the r_t form a Cauchy sequence in R and that it maps to the element of $\varprojlim_t R/I^t$ with which we started. Consider any other Cauchy sequence with the same image. It is again straightforward to verify that the difference of the two Cauchy sequences is in $\mathfrak{C}_i^0(R)$. This proves the isomorphism:

Theorem. *Let R be any ring and I any ideal. Then $\mathfrak{C}_I(R)/\mathfrak{C}_I^0(R) \to \varprojlim_t (R/I^t)$ is an isomorphism, and the kernel of the map from R to either of these isomorphic R-algebras is $\cap_t I^t$.* □

These isomorphic rings are denoted $\widehat{R}^I$ or simply $\widehat{R}$, if I is understood, and either is referred to as the *I-adic completion* of R. If $I \subseteq R$, then R is called I-adically separated if $\bigcap_t I^t = (0)$, and I-adically complete if $R \to \widehat{R}^I$ is an isomorphism: this holds iff R is I-adically separated, and every Cauchy sequence is the sum of a constant sequence r, r, r, ... and a sequence that converges to 0. The Cauchy sequence is said to *converge* to r.

Given a Cauchy sequence in R with respect to I, we may choose a subsequence such that the residues of all terms from the t th on are constant mod I^{t+1}. Call such a Cauchy sequence *standard.* Given a standard Cauchy sequence, let $s_0 = r_0$ and $s_{t+1} = r_{t+1} - r_t \in I^t$ for $t \geq 0$. Then the $s_0 + \cdots + s_t = r_t$. Thus, the partial sums of the "formal series" $s_0 + s_1 + s_2 + \cdots$ form a Cauchy sequence, and if the ring is complete it converges. Given any formal series $\sum_{t=0}^{\infty} s_t$ such that $s_t \in I^t$ for all t, the partial sums form a Cauchy sequence, and every Cauchy sequence is obtained, up to equivalence (i.e., up to adding a sequence that converges to 0) in this way.

Proposition. *Let J denote the kernel of the map from $\widehat{R}^I \twoheadrightarrow R/I$ (J consists of elements represented by Cauchy sequences all of whose terms are in I). Then every element of $\widehat{R}^I$ that is the sum of a unit and an element of J is invertible in $\widehat{R}^I$. Every maximal ideal of $\widehat{R}^I$ contains J, and so there is a bijection between the maximal ideals of $\widehat{R}^I$ and the maximal ideals of R/I. In particular, if R/I is quasilocal, then $\widehat{R}^I$ is quasilocal.*

Proof. If u is a unit and $j \in J$ we may write $u = u(1 + u^{-1}j)$, and so it suffices to to show that $1 + j$ is invertible for $j \in J$. Let $r_0, r_1, \ldots$ be a Cauchy sequence that represents j. Consider the sequence $1 - r_0, 1 - r_1 + r_1^2, \ldots 1 - r_n + r_n^2 - \cdots + (-1)^{n-1} r_n^{n+1}, \cdots$: call the n th term of this sequence v_n. If r_n and r_{n+1} differ by an element of I^t, then v_n and v_{n+1} differ by an element of $I^t + I^{n+2}$. From this it follows that v_n is a Cauchy sequence, and $1 - (1 + r_n)v_n = r_n^{n+2}$ converges to 0. Thus, the sequence v_n represents an inverse for $1 + j$ in $\widehat{R}^I$.

Suppose that m is a maximal ideal of $\widehat{R}^I$ and does not contain $j \in J$. Then j has an inverse v mod m, so that we have $jv = 1 + u$ where $u \in m$, and then $-u = 1 - jv$ is not invertible, a contradiction, since $jv \in J$. □

Suppose that $\bigcap_t I^t = 0$. We define the distance $d(r, s)$ between two elements $r, s \in R$ to be 0 if $r = s$, and otherwise to be $1/2^n$ (this choice is somewhat arbitrary), where n is the largest integer such that $r - s \in I^n$. This is a metric on R: given three elements

$r, s, t \in R$, the triangle inequality is clearly satisfied if any two of them are equal. If not, let n, p, q be the largest powers of I containing $r - s$, $s - t$, and $t - r$, respectively. Since $t - r = -(s - t) - (r - s)$, $q \geq \min\{n, p\}$, with equality unless $n = p$. It follows that in every "triangle," the two largest sides (or all three sides) are equal, which implies the triangle inequality. The notion of Cauchy sequence that we have given is the same as the notion of Cauchy sequence for this metric. Thus, $\widehat{R}^I$ is literally the completion of R as a metric space with respect to this metric.

Given a ring homomorphism $R \to R'$ mapping I into an ideal I' of R', Cauchy sequences in R with respect to I map to Cauchy sequences in R' with respect to I', and Cauchy sequences that converge to 0 map to Cauchy sequences that converge to 0. Thus, we get an induced ring homomorphism $\widehat{R}^I \to \widehat{R'}^{I'}$. This construction is functorial in the sense that if we have a map to a third ring R'', a ring homomorphism $R' \to R''$, and an ideal I'' of R'' such that I' maps into I'', then the induced map $\widehat{R}^I \to \widehat{R''}^{I''}$ is the composition $(\widehat{R'}^{I'} \to \widehat{R''}^{I''}) \circ (\widehat{R}^I \to \widehat{R'}^{I'})$. If $R \to R'$ is surjective and I maps onto I', then the map of completions is surjective: each element of $\widehat{R'}^{I'}$ can be represented as the partial sums of a series $s_0 + s_1 + s_2 + \cdots$, where $s_n \in (I')^n$. But I^n will map onto $(I')^n$, and so we can find $r_n \in I^n$ that maps to s_n, and then $r_0 + r_1 + r_2 \cdots$ represents an element of $\widehat{R}^I$ that maps to $s_0 + s_1 + s_2 + \cdots$.

Example. Let $S = R[x_1, \ldots, x_n]$ be the polynomial ring in n variables over R, and let $I = (x_1, \ldots, x_n)S$. An element of S/I^n is represented by a polynomial of degree $\leq n - 1$ in the x_i. A sequence of such polynomials will represent an element of the inverse limit if and only if, for every n, then nth term is precisely the sum of the terms of degree at most n in the $n+1$st term. It follows that the inverse limit ring $\widehat{S}^I$ is $R[[x_1, \ldots, x_n]]$, the formal power series ring. In consequence, we can prove:

Theorem. *If R is a Noetherian ring and I is an ideal of R, then $\widehat{R}^I$ is Noetherian.*

Proof. Suppose that $I = (f_1, \ldots, f_n)R$. Map the polynomial ring $S = R[x_1, \ldots, x_n]$ to R as an R-algebra by letting $x_j \mapsto f_j$. This is surjective, and $(x_1, \ldots, x_n)S$ maps onto I. Therefore we get a surjection $R[[x_1, \ldots, x_n]] \twoheadrightarrow \widehat{R}^I$. Since we already know that the formal power series ring is Noetherian, it follows that $\widehat{R}^I$ is Noetherian. □

We next want to form the I-adic completion of an R-module M. This will be not only an R-module: it will also be a module over $\widehat{R}^I$. Let R be a ring, $I \subseteq R$ an ideal and M an R-module. Let $\mathfrak{C}_I(M)$ denote the Cauchy sequences in M with respect to I: the sequence $u_0, u_1, u_2, \cdots$ is a *Cauchy sequence* if for all $t \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that $u_i - u_j \in I^t M$ for all $i, j \geq N$. These form a module over $\mathfrak{C}_I(R)$ under termwise multiplication, and set of Cauchy sequences, $\mathfrak{C}_I^0(M)$, that converge to 0, where this means that for all t, the terms of the sequence are eventually all in $I^t M$, is a submodule that contains $\mathfrak{C}_I^0(R)\mathfrak{C}_I(M)$. The quotient $\mathfrak{C}_I(M)/\mathfrak{C}_I^0(M)$ is consequently a module over $\widehat{R}^I$. Moreover, any homomorphism $h : M \to N$ induces a homomorphism from $\mathfrak{C}_I(M) \to \mathfrak{C}_I(N)$ that preserves convergence to 0, and hence a homomorphism $\widehat{h}^I : \widehat{M}^I \to \widehat{N}^I$. This is a covariant functor from R-modules to $\widehat{R}^I$-modules. There is an R-linear map $M \to \widehat{M}^I$

that sends the element u to the element represented by the constant Cauchy sequence whose terms are all u. The kernel of this map is $\bigcap_t I^t M$, and so it is injective if and only if $\bigcap_t I^t M = 0$, in which case M is called *I-adically separated.* If $M \to \widehat{M}^I$ is an isomorphism, M is called *I-adically complete.* The maps $M \to \widehat{M}^I$ give a natural transformation from the identity functor on R-modules to the I-adic completion functor. Moreover, by exactly the same reasoning as in the case where $M = R$, $\widehat{M}^I \cong \varprojlim_t M/I^t M$.

I-adic completion commutes in an obvious way with finite direct sums and products (which may be identified in the category of R-modules). The point is that $u_n \oplus v_n$ gives a Cauchy sequence (respectively, a sequence converging to 0) in $M \oplus N$ if and only if u_n and v_n give such sequences in M and N. Moreover if $f_1 : M_1 \to N$ and $f_2 : M_2 \to N$, we have that the I-adic completion of the map $f_1 \oplus f_2 : M_1 \oplus M_2 \to N$ is the direct sum of the completions, $\widehat{f_1} \oplus \widehat{f_2}$. A similar remark applies when we have $g_1 : M \to N_1$ and $g_2 : M \to N_2$, and we consider the map $(g_1, g_2) : M \to N_1 \times N_2$. The situation is the same for finite direct sums and finite direct products. Note also that if we consider the map given by multiplication by r on M, the induced endomorphism of $\widehat{M}^I$ is given by multiplication by r (or by the image of r in $\widehat{R}^I$).

If $M \to Q$ is surjective, the map $\widehat{M}^I \to \widehat{Q}^I$ is surjective: as in the case of rings, any element z of $\widehat{Q}^I$ can be represented using the Cauchy sequence of partial sums of a formal series $q_0 + q_1 + q_2 + \cdots$ where $q_t \in I^t Q$. To see this, take a Cauchy sequence that represents the element. Pass to a subsequence w_0, w_1, w_2, ... such that the residue of w_k in $M/I^t M$ is the same for all $k \geq t$. The element can be thought of as

$$w_0 + (w_1 - w_0) + (w_2 - w_1) + \cdots .$$

Thus, take $q_0 = w_0$ and $q_t = w_t - w_{t-1}$ for $t \geq 1$. For all t, $I^t M$ maps onto $I^t Q$. Therefore we can find $u_t \in I^t M$ such that u_t maps to q_t, and the partial sums of $u_0 + u_1 + u_2 + \cdots$ represent an element of $\widehat{M}^I$ that maps to z.

Note that because $\widehat{M}^I$ is an R-module and we have a canonical map $M \to \widehat{M}^I$ that is R-linear, the universal property of base change determines a map $\widehat{R}^I \otimes_R M \to \widehat{M}^I$. These maps give a natural transformation from the functor $\widehat{R}^I \otimes_R \_$ to the I-adic completion functor: these are both functors from R-modules to $\widehat{R}^I$-modules. If M is finitely generated over a Noetherian ring R, this map is an isomorphism: not only that: restricted to finitely generated modules, I-adic completion is an exact functor, and $\widehat{R}^I$ is flat over R.

In order to prove this, we need to prove the famous Artin-Rees Lemma. Let R be a ring and I an ideal of R. Let t be an indeterminate, and let $It = \{it : i \in I\} \subseteq R[t]$. Then $R[It] = R + It + I^2t^2 + \cdots$ is called the *Rees ring* of I. If $I = (f_1, \ldots, f_n)$ is finitely generated as an ideal, then $R[It] = R[f_1 t, \ldots, f_n t]$ is a finitely generated R-algebra. Therefore, the Rees ring is Noetherian if R is.

Before proving the Artin-Rees theorem, we note that if M is an R-module and t and indeterminate, then every element of $R[t] \otimes M$ can be written uniquely in the form

$$1 \otimes u_0 + t \otimes u_1 + \cdots + t^k \otimes u_k,$$

where the $u_j \in M$, for any sufficiently large k: if a larger integer s is used, then one has $m_{k+1} = \cdots = m_s = 0$. This is a consequence of the fact that $R[t]$ is R-free with the powers of t as a free basis. Frequently one writes $u_0 + u_1t + \cdots + u_kt^k$ instead, which looks like a polynomial in t with coefficients in M. When this notation is used, $M[t]$ is used as a notation for the module. Note that the $R[t]$-module structure is suggested by the notation: $(rt^j)(ut^k) = (ru)t^{j+k}$, and all other more general instances of multiplication are then determined by the distributive law.

We are now ready to prove the Artin-Rees Theorem, which is due independently to Emil Artin and David Rees.

Theorem (E. Artin, D. Rees). *Let $N \subseteq M$ be Noetherian modules over the Noetherian ring R and let I be an ideal of R. Then there is a constant positive integer c such that for all $n \geq c$, $I^nM \cap N = I^{n-c}(I^cM \cap N)$. That is, eventually, each of the modules $N_{n+1} = I^{n+1}M \cap N$ is I times its predecessor, $N_n = I^nM \cap N$.*

In particular, there is a constant c such that $I^nM \cap N \subseteq I^{n-c}N$ for all $n \geq c$. In consequence, if a sequence of elements in N is an I-adic Cauchy sequence in M (respectively, converges to 0 in M) then it is an I-adic Cauchy sequence in N (respectively, converges to 0 in N).

Proof. We consider the module $R[t] \otimes M$, which we think of as $M[t]$. Within this module,

$$\mathcal{M} = M + IMt + I^2Mt^2 + \cdots + I^kMt^k + \cdots$$

is a finitely generated $R[It]$-module, generated by generators for M as an R-module: this is straightforward. Therefore, $\mathcal{M}$ is Noetherian over $R[It]$. But

$$\mathcal{N} = N + (IM \cap N)t + (I^2M \cap N)t^2 + \cdots,$$

which may also be described as $N[t] \cap \mathcal{M}$, is an $R[It]$ submodule of $\mathcal{M}$, and so finitely generated over $R[It]$. Therefore for some $c \in \mathbb{N}$ we can choose a finite set of generators whose degrees in t are all at most c. By breaking the generators into summands homogeneous with respect to t, we see that we may use elements from

$$N, (IM \cap N)t, (I^2M \cap N)t^2, \ldots, (I^cM \cap N)t^c$$

as generators. Now suppose that $n \geq c$ and that $u \in I^nM \cap N$. Then ut^n can be written as an $R[It]$-linear combination of of elements from

$$N, (IM \cap N)t, (I^2M \cap N)t^2, \ldots, (I^cM \cap N)t^c,$$

and hence as an sum of terms of the form

$$i_ht^hv_jt^j = (i_hv_j)t^{h+j}$$

where $j \leq c$, $i_h \in I^h$, and

$$v_j \in I^jM \cap N.$$

Of course, one only needs to use those terms such that $h + j = n$. This shows that $(I^n M) \cap N$ is the sum of the modules

$$I^{n-j}(I^j M \cap N)$$

for $j \leq c$. But

$$I^{n-j}(I^j M \cap N) = I^{n-c} I^{c-j}(I^j M \cap N),$$

and

$$I^{c-j}(I^j M \cap N) \subseteq I^c M \cap N,$$

so that we only need the single term $I^{n-c}(I^c M \cap N)$. □

Theorem. *Let R be a Noetherian ring, $I \subseteq R$ an ideal.*

(a) *If $0 \to N \to M \to Q \to 0$ is a short exact sequence of finitely generated R-modules, then the sequence $0 \to \widehat{N}^I \to \widehat{M}^I \to \widehat{Q}^I \to 0$ is exact. That is, I-adic completion is an exact functor on finitely generated R-modules.*

(b) *The natural transformation θ from $\widehat{R}^I \otimes_R \_$ to the I-adic completion functor is an isomorphism of functors on finitely generated R-modules. That is, for every finitely generated R-module M, the natural map $\theta_M : \widehat{R}^I \otimes_R M \to \widehat{M}^I$ is an isomorphism.*

(c) *$\widehat{R}^I$ is a flat R-algebra. If (R, m) is local, $\widehat{R} = \widehat{R}^m$ is a faithfully flat local R-algebra.*

Proof. (a) We have already seen that the map $\widehat{M}^I \to \widehat{Q}^I$ is surjective. Let y be an element of $\widehat{M}^I$ that maps to 0 in $\widehat{Q}$. Choose a Cauchy sequence that represents z, say $u_0, u_1, u_2, \ldots$. After passing to a subsequence we may assume that $u_t - u_{t+1} \in I^t M$ for every t. The images of the u_t in $Q \cong M/N$ converge to 0. Passing to a further subsequence we may assume that the image of $u_t \in I^t(M/N)$ for all t, so that $u^t \in I^t M + N$, say $u_t = v_t + w_t$ where $v_t \in I^t M$ and $w_t \in N$. Then w_t is a Cauchy sequence in M that represents z: in fact, $w_t - w_{t+1} \in I^t M \cap N$ for all t. Each $w_t \in N$, and so the elements w_t form a Cauchy sequence in N, by the Artin-Rees Theorem. Thus, every element in $\operatorname{Ker}(\widehat{M}^I \to \widehat{Q}^I)$ is in the image of $\widehat{N}^I$.

Finally, suppose that $z_0, z_1, z_2, \ldots$ is a Cauchy sequence in N that converges to 0 in M. Then z_t already converges to 0 in N, and this shows that $\widehat{N}^I$ injects into $\widehat{M}^I$. This completes the proof of part (a).

(b) Take a presentation of M, say $R^n \xrightarrow{A} R^m \to M \to 0$, where $A = (r_{ij})$ is an $m \times n$ matrix over R. This yields a diagram:

$$\begin{tikzcd}
\widehat{R}^I \otimes_R R^n \arrow[r, "A"] & \widehat{R}^I \otimes_R R^m \arrow[r] & \widehat{R}^I \otimes_R M \arrow[r] & 0 \\
\widehat{R^n}^I \arrow[u, "\theta_{R^n}"] \arrow[r, "A"] & \widehat{R^m}^I \arrow[u, "\theta_{R^m}"] \arrow[r] & \widehat{M}^I \arrow[u, "\theta_M"] \arrow[r] & 0
\end{tikzcd}$$

where the top row is obtained by applying $\widehat{R}^I \otimes \_$, and is exact by the right exactness of tensor, the bottom row is obtained by applying the I-adic completion functor, and is exact by part (a). The vertical arrows are given by the natural transformation θ, and the squares commute because θ is natural. The map θ_{R^h} is an isomorphism for $h = m$ or $h = n$ because both functors commute with direct sum, and the case where the free module is just R is obvious. But then θ_M is an isomorphism, because cokernels of isomorphic maps are isomorphic.

(c) We must show that $\widehat{R}^I \otimes_R N \to \widehat{R}^I \otimes_R M$ is injective for every pair of R-modules $N \subseteq M$. We know this from parts (a) and (b) when the modules are finitely generated. The result now follows from the Lemma just below. Faithful flatness is clear, since the maximal ideal of R clearly expands to a proper ideal in $\widehat{R}^I$. □

Lemma. *Let F be an R-module, and suppose that whenever $N \subseteq M$ are finitely generated R-modules then $F \otimes_R N \to F \otimes_R M$ is injective. Then F is flat.*

Proof. Let $N \subseteq M$ be arbitrary R-modules. Then $F \otimes_R N$ is the directed union of the images of the modules $F \otimes_R N_0$ as F runs through the finitely generated submodules of M. Thus, if $z \in F \otimes N$ maps to 0 in $F \otimes M$, it will be the image of $z' \in N_0 \otimes M - \{0\}$, which implies that $z' \in F \otimes_R N_0$ maps to 0 in $F \otimes_R M$. But since M is the directed union of its finitely generated modules M_0 containing N_0, and since $F \otimes_R M$ is the direct limit of these, it follows that for some sufficiently large but finitely generated $M_0 \supseteq N_0$, the image of z' under the map $F \otimes N_0 \to F \otimes M_0$ is 0. But then $z' = 0$ and so $z = 0$, as required. □