Math 615: Lecture of January 10, 2007

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

$$\begin{aligned}
&1 < x_2 < x_2^2 < \cdots < x_2^n < \cdots \\
&< x_1 < x_1x_2 < x_1x_2^2 < \cdots < x_1x_2^n < \cdots \\
&< x_1^2 < x_1^2x_2 < x_1^2x_2^2 < \cdots < x_1^2x_2^n < \cdots \\
&\qquad\qquad \cdots \\
&< x_1^m < x_1^mx_2 < x_1^mx_2^2 < \cdots < x_1^mx_2^n < \cdots \\
&\qquad\qquad \cdots
\end{aligned}$$

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

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

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

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

There is a double reversal of sorts here, since one is using the *opposite* of what lexicographic order gives when the variables are *numbered backwards.* One still has that

$x_1 > x_2 > \cdots > x_n$, and in the two variable case hlex and revlex are the same. In the three variable case one has that $x_1x_3 >_{\text{lex}} x_2^2$ while $x_2^2 >_{\text{revlex}} x_1x_3$. For the latter, the variable with the highest index for which the two monomials have different exponents is x_3, and the first monomial has the smaller exponent. The difference between the two conditions might be paraphrased by saying that if two monomials have the same degree, for hlex the greater involves more of the low index variables while for revlex the greater involves fewer of the high index variables. This statement is quite misleading, however, since it is only the first spot (for hlex) and the last spot (for revlex) where the monomials have differerent exponents that governs which monomial is greater. E.g., with 1000 variables,

$$x_1x_{999}^{10000000} > x_2^{10000000}x_{1000}$$

for both hlex and revlex.

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

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

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

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

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

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

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

(1) Every nonempty subset has a least element.

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

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

The following is a critical property of monomial orders.

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

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

Initial terms and the division algorithm

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

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

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

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

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

Let $M \subseteq F$ be an arbitrary submodule. The submodule of F spanned by the initial terms of all elements of M is a monomial submodule: instead of using $c\nu$ as a generator, where $c \in K - \{0\}$ and ν is a monomial, we can use ν itself. This submodule of F is denoted

$\mathrm{in}_>(M)$ or $\mathrm{in}(M)$ and is called the *initial module* of M. It is typically not contained in M (unless M itself is a monomial module). If $F = R$ and I is an ideal, $\mathrm{in}(I)$ is called the *initial ideal* of I. Just as in the case of individual elements, we may indicate that the monomial order used is lexicographic, homogeneous lexicographic, or reverse lexicographic order with the respective notations $\mathrm{in}_{\mathrm{lex}}(M)$, $\mathrm{in}_{\mathrm{hlex}}(M)$, or $\mathrm{in}_{\mathrm{revlex}}(M)$.

With these notations in place, we want to discuss an analogue of the division algorithm for polynomial rings in one variable over a field. However, in our case, instead of dividing by one polynomial to get a quotient and remainder, we may be "dividing" by several. Furthermore, instead of working with polynomials, we may be working with elements of F. However, for heuristic reasons, the reader may want to think at first only about the case where $F = R$.

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

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

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

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

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

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

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

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

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

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

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

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

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

and monomials ν_n in F (except that ν_n is not defined for the final value of n) as follows. If $n = 0$ the expression is simply given by taking all $q_{i,0} = 0$. If f_n has no term divisible by any of the $\mathrm{in}(g_i)$ the procedure stops, and we have that $(\#_n)$ is a standard expression for f with remainder $h = f_n$. Otherwise, once f_n and the corresponding expression $(\#_n)$ are known, let $c_n\nu_n$ be the largest term of f_n that is a multiple of one or more of the elements $\mathrm{in}(g_i)$. (The procedure that we are describing will eventually terminate no matter which of the g_i with $\mathrm{in}(g_i)$ dividing ν we choose, but we want to make it deterministic.) Let i_n be the least integer such that $\mathrm{in}(g_{i_n})$ divides $c_n\nu_n$, and choose $c'_n\mu_n$ such that $c\nu_n = c'\mu_n\mathrm{in}(g_{i_n})$. Finally, we let

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

and then we may take

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

for $j \neq i_n$ while

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

A straightforward induction then shows the following:

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

(b) The sequence

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

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

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

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

is

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

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

We have proved:

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

In case $F = R = K[x]$, with $r = 1$, so that we are dividing f by $g_1 = g$ in $K[x]$, the standard expression we get must be $f = qg + h$, where $\deg(h) < \deg(g)$ or $h = 0$. Here, if

$\deg(g) = d$, $\mathrm{in}(g) = cx^d$ for some $c \neq 0$ in K, and the condition that h has no term divisible by $\mathrm{in}(g)$ is equivalent to the condition that $\deg(h) < \deg(g)$ or $h = 0$. The individual steps in the algorithm are exactly the steps in the usual division algorithm for polynomials in one variable.

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

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

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

and

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

which is the remainder. The standard expression we get is

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

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

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

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