Math 615: Lecture of January 24, 2007

Hilbert functions

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

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

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

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

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

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

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

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

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

which is

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

We can get the same formula from a purely combinatorial argument. $\mathrm{Hilb}(d)$ is the number of monomials x^α where $\alpha = (a_1, \ldots, a_n)$ where the $a_i \in \mathbb{N}$ and $a_1 + \cdots + a_n = d$. Each such monomial can be represented by a string containing d blanks _ interspersed with $n-1$ slashes /, where there are first a_1 blanks, then a slash as a separator, then a_2 blanks, then a slash as a separator, and so forth. The string will end with a slash, then a_{n-1} blanks, then a slash, and, finally a_n blanks. (For example, if $n = 4$ and $d = 8$, the string corresponding to $x_1^3x_3x_4^5$ is

_ _ _ / / _ / _ _ _ _ _ .

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

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

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

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

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

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

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

it follows that

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

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

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

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

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

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

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

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

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

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

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

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

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

where every

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

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

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

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

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

Theorem (Schreyer). *Let notation be as above. Then the standard relations ρ_{ij} generate the module of relations on the Gröbner basis $g_1, \ldots, g_r$. What is more, the relations ρ_{ij} form a Gröbner basis for the module of relations on the $g_1, \ldots, g_r$ with respect to the monomial order on R^r defined above.*

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

$$\mathrm{in}(f_1e_1 + \cdots + f_re_r)$$

for an arbitrary relation on $g_1, \ldots, g_r$. All we need to do is show that each such initial term is a multiple of one of the $\mathrm{in}(\rho_{ij})$. Each $\nu_i = \mathrm{in}(g_i)$ involves one element of the free basis $b_1, \ldots, b_s$ for the original free module R^e: call this element $b_{L(i)}$. Then the monomial μ in f_i that gives rise to the largest term of f_ie_i after multiplying out is the same monomial μ that gives the largest term in f_ig_i, and this is $\mathrm{in}_{>_{L(i)}}(f_i)\nu_i$ by the displayed formula (†) on p. 2 of the Lecture Notes of January 19. It follows that the largest term in f_ie_i is $\mathrm{in}_{>_{L(i)}}e_i$. Thus, $\mathrm{in}(f_1e_1 + \cdots f_re_r)$ may be described as follows. Consider the largest initial term for any f_ig_i, call it ν, and choose the smallest i such that ν is $\mathrm{in}(f_ig_i)$, up to a nonzero scalar multiple. Then $\mathrm{in}(f_1e_1 + \cdots + f_re_r)$ is $\mathrm{in}(f_ie_i) = \mathrm{in}_{>_{L(i)}}(f_i)e_i$ for this smallest value of i.

This is precisely the same use of ν as in the proof of the Buchberger criterion in the Lecture Notes of January 19.

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

Since $f_1g_1 + \cdots + f_rg_r = 0$, the initial terms of products f_jg_j that are, up to a nonzero scalar multiple, equal to ν must cancel. Suppose the products that have $c\nu$ as initial term for $c \in K - \{0\}$ are indexed by $j_1, \ldots, j_h$ where $j_1 < \cdots < j_h$. Let $\mu_j = \mathrm{in}_{>_{L(j)}}(f_j)$.

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

$$\mathrm{in}(f_1e_1 + \cdots + f_re_r) = \mu_{j_1}e_{j_1}.$$

We also have the relation $\sum_{t=1}^{h} \mu_t\nu_t = 0$. Exactly as in the proof of the Buchberger criterion, this means that $(\mu_1, \ldots, \mu_h)$ is a homogeneous linear combination, with coefficients that are terms in R, of the relations θ_{ij}: see the displayed line (#) near the top of p. 4 of the Lecture Notes of January 19 and the preceding discussion. However, in fact, we only need those θ_{ij} such that $i = j_a < j_b = j$. This means that μ_{j_1} must be a multiple, by a term in R, of the coefficient of e_{j_1} in some $\theta_{j_1 j_t}$ for $t > 1$. But this also means precisely that $\mu_{j_1} e_1$ is a multiple of $\mathrm{in}(\rho_{j_1 j_t})$ for some $t > 1$. $\square$

Finding the relations on elements that are not a Gröbner basis

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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