Exactness, Graded Rings

April 7, 2007

We begin with a short review of quotient rings and modules. Recall that we have a surjective map

$$\pi : R \to R/I \quad \pi(r) = r + I.$$

This is clearly surjective. To compute its kernel, we note that the zero element in R/I corresponds to the set $0 + I$. Recall that $0 + I = r + I$ so long as $r - 0 = r \in I$. Thus the kernel of π is exactly I. This proves the following theorem:

Theorem 1. *Let R be a ring, and I an ideal. Then there exists a ring S and a ring homomorphism $\phi : R \to S$ so that $I = ker\ \phi$.*

Proof. Just let $S = R/I$, and $\phi = \pi$. □

Note that the same results hold if we replace R, I by any pair $N \subseteq M$ of submodules. This result gives as nice characteriziation of ideals and submodules, but upon looking at the proof, it's clear that not much significant is happening.

1 Exactness

Definition 1. *Let R be a ring and L, M, N be R-modules. Consider the sequence of maps:*

$$L \xrightarrow{f} M \xrightarrow{g} N.$$

We say that this sequence is exact at M if ker g =im f. A longer sequence

$$\cdots \xrightarrow{f_{n+2}} M_{n+1} \xrightarrow{f_{n+1}} M_n \xrightarrow{f_n} M_{n-1} \xrightarrow{f_{n-1}} M_{n-2} \xrightarrow{f_{n-2}} \cdots$$

is called exact if it is exact at every M_i. In other words, ker f_i = im f_{i+1} for all i.

Exact sequences are extremely useful all over mathematics, from algebra to topology, to geometry. We illustrate the basic cases below. (If no map is given, it is because the map is obvious)

- $M \xrightarrow{g} N \to 0$ is exact iff g is surjective.

- $0 \to M \xrightarrow{f} N$ is exact iff f is injective.
- $0 \to M \xrightarrow{f} N \to 0$ is exact iff f is injective and surjective (an isomorphism of modules).
- $0 \to M \xrightarrow{f} N \xrightarrow{g} P \to 0$ is exact iff f is injective, g is surjective, and ker $g = \text{im} f$. Sequences of this form, with five terms and zeros on the ends are called short exact sequences (SES). Longer ones are called long exact sequences.

These special sequences are the most important to remember, and it can be shown in fact that all long exact seqeuences can be interpretted in terms of short exact sequences.

Let's now do any example with the modules being vector spaces. Consider the sequence

$$0 \to \mathbb{R} \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R} \to 0$$

$$f = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad g = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

as matrices. It is then straightforward to check that g is surjective, f is injective, and ker $g =$im f.

To see that this is not the standard case, we'll use the same final map g but now change what occurs to the left:

$$0 \to \mathbb{R} \xrightarrow{h} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R} \to 0$$

$$h = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad f = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad g = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

It is straightforward to check that this is exact. One could easily imagine how to extend this example, so that for example $\mathbb{R}^n$ appeared for any n, just by shuffling around the entries of the matrix. Thus the function $g : \mathbb{R}^2 \to \mathbb{R}$ leads to several different exact sequences. These sequences however are quite redundant, since most of the matrices will involve rows and columns of zeros. The only real important one is the one in the first example. We will discuss this notion later when we talk about free resolutions.

Another prime example of an exact sequence is as follows. Recall that if R is a ring, and I an ideal, then R, I and R/I are R-modules. Thus the following sequence is a sequence of R-modules which can be seen to be exact by the remarks at the beginning of this section.

$$0 \to I \xrightarrow{i} R \xrightarrow{\pi} R/I \to 0$$

where i is the inclusion map $i(r) = r$ and π is the projection map.

2 Resolutions

To give you a taste of what we'll do in the future, here's a problem statement: Let $R = k[x, y, z]$ and let $I = (x, y, z)$ be an ideal. Then suppose you wanted to create an exact sequence with the last maps $R \to R/I \to 0$. Well the very natural answer might be

$$0 \to I \to R \to R/I \to 0.$$

This is of course correct, but now we make the problem more interesting by requiring that each module appearing is of the form R^n for some n. We note the following: Since $I = (x, y, z)$ there exists a (surjective) map

$$\phi : R^3 \to I \quad (f, g, h) \mapsto fx + gy + hz.$$

Now looking at the following

$$\begin{array}{ccccccc} R^3 & \xrightarrow{\phi_1} & R & \longrightarrow & R/I & \longrightarrow & 0 \\ & \searrow^{\phi} & \uparrow{\scriptstyle i} & & & & \\ & & I & & & & \end{array}$$

where the map ϕ_1 is just the composition $i \circ \phi$. Convince yourself now that the image of ϕ_1 is the same as the image of i (since ϕ is surjective). Thus the top row of this sequence remains exact and we have successfully gotten rid of I and replaced it with a copy of R^3. Now if J is the kernel of the map ϕ_1 then we have an exact sequence

$$0 \to J \to R^3 \to R \to R/I \to 0.$$

So to continue we'll try to do the same thing we did before. But first we must compute the kernel of ϕ_1. Note that in terms of a matrix, $\phi_1 = (x\ y\ z)$. (The same as the matrix of ϕ). Note that

$$\phi_1 \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix} = \phi_1 \begin{pmatrix} z \\ 0 \\ -x \end{pmatrix} = \phi_1 \begin{pmatrix} 0 \\ z \\ -y \end{pmatrix} = 0$$

and in fact these elements generate the whole kernel so we have

$$J = \operatorname{span}_R \left\{ \begin{pmatrix} y \\ -x \\ 0 \end{pmatrix}, \begin{pmatrix} z \\ 0 \\ -x \end{pmatrix}, \begin{pmatrix} 0 \\ z \\ -y \end{pmatrix} \right\} = \operatorname{span}_R \{v_1, v_2, v_3\}.$$

As before we have a map $\psi : R^3 \to J$ sending $(f, g, h) \to fv_1 + gv_2 + hv_3$. Thus as before, we can replace our J with R^3 from the following diagram

$$\begin{array}{ccccccccc} R^3 & \xrightarrow{\psi_1} & R^3 & \xrightarrow{\phi_1} & R & \longrightarrow & R/I & \longrightarrow & 0 \\ & \searrow^{\psi} & \uparrow{\scriptstyle i} & & & & & & \\ & & J & & & & & & \end{array}$$

where since the matrix of ψ_1 is the matrix of ψ which is just the v_i in columns.

$$\psi_1 = \begin{pmatrix} y & z & 0 \\ -z & 0 & z \\ 0 & -x & -y \end{pmatrix}$$

We now call the kernel of this map K and we try to compute it. Again, note that

$$\psi_1 \begin{pmatrix} z \\ -y \\ x \end{pmatrix}$$

And trust us again when we say that this generates all of K. Thus

$$K = \text{span}_R \left\{ \begin{pmatrix} z \\ -y \\ x \end{pmatrix} \right\}$$

and finally, we have a surjective map $\xi : R \to K$ as before, leading to the exact sequence.

$$R \xrightarrow{\xi_1} R^3 \xrightarrow{\psi_1} R^3 \xrightarrow{\phi_1} R \longrightarrow R/I \longrightarrow 0 \ .$$

Finally, since the matrix of ξ_1 is just

$$\xi_1 = \begin{pmatrix} z \\ -y \\ x \end{pmatrix}$$

we have that the kernel of ξ_1 is 0. Thus in conclusion our exact sequence for R/I is

$$0 \longrightarrow R \xrightarrow{\xi_1} R^3 \xrightarrow{\psi_1} R^3 \xrightarrow{\phi_1} R \longrightarrow R/I \longrightarrow 0 \ .$$

This preceding example is a particular case of the Koszul Complex. I have given a sequence of talks on the Koszul complex in general, and notes on these talks can be found on the main page of my website. The level of the notes starts rather low but escalates somewhat as they progress.

In any event, we should now talk about what we just did, and why we did it. Recall that in linear algebra, when we studied vector spaces over a field, there was a notion of a generating set, or spanning set. Furthermore, we had the theorem that if you had a minimal spanning set, then the set was also linearly independent. In other words, if removing any vector causes you not to be able to span the same set, then the original set of vectors was indpendent. This is not the case when we are talking about modules.

As we say in the example above, if $R = k[x, y, z]$ then $I = (x, y, z)$ is an R-module which is generated by x, y, z. Removing any one of them changes the ideal, so this set is minimal. But as we saw earlier, $yx - xy = 0$ so even x, y is not independent over R. This prompts the following definition.

Definition 2. *An R-module M is called a free module if there exists an independent spanning set of M over R. We call such a set a basis.*

It is clear from the above, that (x, y, z) or even (x, y) are not free R-modules. (Are they free k-modules?) But note that (x) *is* a free R-module since $fx = 0$ implies that $f = 0$ so that $\{x\}$ is an independent set. Free modules are very nicely classified, which we illustrate in the following theorem.

Theorem 2. *Let M be a n R-module. Then M is a free R-module if and only if $M \cong R^n$ for some n.*

Proof. We first prove that R^n is always a free module. Indeed, R^n has the standard basis $e_1, \ldots, e_n$ where e_i is the vector $(0, \ldots, 1, \ldots, 0)$ with a 1 in the ith spot. If $\sum r_i e_i = 0$ then $(r_1, \ldots, r_n) = 0$ which implies that $r_i = 0$ for all i, and this e_i is a basis. Conversely, if M is free, then M has some basis $b_1, \ldots, b_k$. Thus define a map $\phi : M \to R^k$ by $\phi(b_i) = e_i$ extended by linearity. This map is clearly an isomorphism. □

Thus all that business we were doing above with writing our exact sequence could be rephrased as follows: Write an exact sequence with rightmost maps $R \to R/I \to 0$ so that every module appearing other than R/I is a free module. In general we have the following definition.

Definition 3. *Let M be an R-module. An exact sequence of the form*

$$\cdots F_n \to F_{n-1} \to \cdots \to F_0 \to M \to 0$$

where each F_i is free is called a free resolution of M.

Above we saw a free resolution of the $k[x, y, z]$-module $R/(x, y, z)$. This resolution had only a finite number of nonzero terms. The following example shows this need not be the case.

Example: Let $R = k[x]/(x^2)$. Then if $\overline{x}$ represents the image of x in R, that is, $\overline{x} = x + (x^2)$, then let $M = R/(\overline{x})$. Then a free resolution of M is

$$\cdots \to R \to R \to \cdots \to R \to R/(\overline{x}) \to 0$$

where each map is multiplication by $\overline{x}$. To see this is exact, we check it. R can be identified with the set of all $a + bx$ $a, b \in k$. Then $x(a + bx) = xa$. So the image of this map is all multiples of x, and the kernel of the map is also all multiples of x. Exactness follows immediately noting that the last map is a projection and thus surjective.

The phenomena illustrated in this example, is somewhat abnormal, and is due to the fact that the ring R we were using was somewhat complicated. Soon we will learn the Hilbert Syzygy theorem which states that when the ring R is a polynomial ring, and the module is sufficiently "nice" the resolution will always have finite length.

3 Graded Rings and Modules

One of the first things that we all learned about polynomials in high school was that they have a particular degree. Degree was an important notion and had many useful properties. For example, the degree of a product was equal to the sum of the degrees, and the degree of a sum was less than the max of original degrees. We recall that we assigned a degree of zero to all constant functions, and the pesky 0 polynomial caused us problems. Some texts don't give it a degree, while others assign it the degree $-\infty$. In this section we will give a more abstract notion of degree that will coincide with the one we are familiar with. We begin by reviewing the notion of a direct sum.

Definition 4. *Let R_i with $i \in I$. be abelian groups. (e.g. rings, ignoring the multiplication) We define the direct sum $\bigoplus_{i\in I} R_i$ to be the ring whose elements are all finite sums $r_{i_1} + \cdots + r_{i_n}$ with $r_{i_j} \in R_{i_j}$. Equivalently, it is the set of all sums*

$$\sum_{i\in I} r_i \quad r_i \in R_i$$

where all but finitely many of the r_i are 0. The addition on this new group is as follows. If $\sum r_i$ and $\sum s_i$ are two elements, then we add all terms r_i and s_i that are in the same R_i to simplify the formal sum $\sum r_i + s_i$.

Note here that if I is finite then it is standard to write the direct sum in vector format: For example $R \oplus S = (r, s)$ such that $r \in R, s \in S$. In fact, to better understand the group operation on the direct sum, it is perhaps best to think of elements in $\oplus R_i$ not as sums but as infinite vectors with almost all entries zero.

Note that just as in the finite case, $R_i \cap R_j = 0$ if $i \neq j$. This is even true if say the R_i are the same groups, because $(r, 0, \ldots) \neq (0, r, \ldots)$.

Definition 5. *A graded ring is a ring R so that R has a decomposition*

$$R = \bigoplus_{i=0}^{n} R_i = R_0 \oplus R_1 \oplus \cdots$$

where the R_i are abelian groups and the multiplication maps satisfy $R_i R_j \subset R_{i+j}$.

These definitions were seemingly painful so we now take the time to translate them into more tractable terms. We illustrate both of these definition using the polynomial ring $k[x]$.

Note that $R_n = \{0\} \cup \{\text{monomials of degree } n\}$ is an abelian group for all n. Indeed, it is closed under addition and if taking inverses under addition. By the direct sum of the R_i we mean the set of finite sums of monomials - polynomials! It is now clear that

$$k[x] = \bigoplus_{i\geq 0} R_i$$

since any polynomial can be broken down into a sum of monomials of each degree. We now notice that this decomposition also shows that $k[x]$ is a graded ring. Indeed, what happens when we multiply an element of R_i by an element in R_j? This amounts to multiplying two monomials of degrees i and j which is a polynomial of degree $i+j$. Thus we have just shown that $k[x]$ is a graded ring.

This idea can be generalized to rings in any number of variables, if we let the abelian groups R_i be the sets of homogeneous polynomials of degree i.

Note that any ring R can be given a *grading* by setting $R_0 = R$ and $R_i = 0$ if $i > 0$. This is called the trivial grading of a ring.

In a graded ring, we call the elements of R_i the homogeneous elements of degree i. This agrees with our notion of homogeneity of polynomials. Note that any polynomial can be split into homogeneous parts (terms of like degree). For example

$$x^3 + x^2z + y^2 + xyz + 5x + y + z = (x^3 + x^2z + xyz) + (y^2) + (5x + y + z).$$

We say an ideal I is homogenous if for each $f \in I$ each homogeneous component of f is also in I. Ideals which are generated by homogeneous polynomials will be homogeneous.

Definition 6. *Let R be a graded ring. We say that M is a graded R-module if there exists a decomposition of M into abelian groups $M = \oplus_{i \geq 0} M_i$ so that the scalar multiplication by R satisfies $R_i M_j \subset M_{i+j}$.*

If R is a graded ring, then R itself is a graded module. If I is a homogeneous ideal, then R/I is a graded module (prove this!).

To see that the ideal needs to be homogeneous, let $R = k[x]$ and let $I = x^2 - 1$. Then R/I is an R-module. Suppose it were graded, then the element $x + I$ has some degree d, $x + I \in M_d$. Then $x^2(x + I) = x^3 + I = x + I$ should be in M_{d+2}. But then $x + I \in M_d \cap M_{d+2}$. But this is a contradiction since only the 0 element can appear in different elements of a direct sum.

We now begin to discuss Hilbert Functions again, only this time, in terms of graded modules.

Definition 7. *Let M be a finitely graded module over $k[x_0, \ldots, x_n]$. Then M is a k-vector space and the function*

$$h_M(i) = dim_k M_i$$

is called the Hilbert function of M.

Let's do an example where $M = (x^3)$ the ideal generated by x^3 in $k[x]$. We give M the standard grading where M_i is just the set of monomials of degree i. Thus

$$H_M(0) = H_M(1) = H_M(2) = 0$$

since there are no element of degree 0, 1 or 2 in M. After that $H_M(k) = 1$ for $k \geq 1$.

We now do another example, this time, with a quotient ring. Let $R = k[x, y]$ and $I = (x^2, y^3)$. To aid in counting the dimension, we note that the basis elements of R/I are exactly the monic monomials that are not in I. We do this example in both ways to help convince you of this.

The first way, note that the elements of degree 0 are just constant polynomials, of the form $c + I$. This is one dimensional. The set of elements of degree 1 are of the form $ax + I$ and $by + I$ - two dimensional. For dimension 2, we have the following $xy + I, y^2 + I$ - two dimensional. (Note that $x^2 + I = 0 + I$). In dimension three we have xy^2 - one dimensional. Finally, in dimension four or higher, all monomials involve either x^2 or y^3, so everything is zero. Thus the Hilbert function is $1, 2, 2, 1, 0, 0, \ldots$.

The second way, note that the monomials not in I are $1, x, y, y^2, xy, xy^2$ and counting, the result follows.

Clearly the second method is a lot quicker. We'll illustrate the technique again with $R = k[x, y, z]$ and $I = (xy^2, z^3)$. We list the monomials not in I.

$$1, (x, y, z), (x^2, y^2, z^2, xy, xz, yz), (x^3, x^2y, x^2z, xyz, xz^2, y^3, y^2z, yz^2), \ldots$$

In degree four and higher, we see that the only monomials of degree d in I are ones of the form xy^2f or z^3f for f of degree $d - 3$. There are $\binom{d-3+2}{2}$ such f in each case (since f can be any monomial of degree $d - 3$). And hence there are $2\binom{d-1}{2}$ such monomials. We have to subtract monomials we counted twice, which are precisely monomials of the form xy^2z^3g with g of degree $d - 6$. Thus the total number of monomials not in I is given by

$$\text{number in } I \;-\; \text{number of multiples of } (xy^2, z^3) \;+\; \text{multiples of both.}$$

$$\binom{d+2}{2} - 2\binom{d-1}{2} + \binom{d-4}{2} = 9.$$

Since our formula only involves $d - 4$ it is valid for all $d \geq 4$ and thus this shows that the Hilbert function is constantly 9 after degree 4. Thus $H_{R/I}(i) = 1, 3, 6, 8, 9, 9, \ldots$.

Important Remark: This last example contains a good deal of information, so we pause for a moment to recap. In these situations it is often possible to write down formulas for the number of elements in an ideal of a specified degree, just as we have in the previous example. This usually amounts to worrying about counting things twice or three times, and the inclusion-exclusion principle should be employed. The basic idea is this. The Hilbert function of R/I in degree d is the number of elements in R_d but that are not in I. This is a number we can count.

Since our formulas are only valid when all numbers appearing in the binomial coefficients are nonnegative, anything of degree that violates this must be checked by hand as we have done in the previous example.

We now take a break from these tedious examples, and try to develop the theory that will make these computations a bit easier, and in doing so, will show the relationship between Hilbert functions and resolutions.

4 Shifts of Modules

Let $R = k[x, y, z]$, and consider the ideal (x^2). What is the dimension of this ideal in degree k? Simple computations show that the dimensions are (starting in degree 0) $0, 0, 1, 3, 6, 10, 15, \ldots$ which are exactly the dimensions of $k[x, y, z]$ only shifting to the right 2 units. We can understand why this is, by just noting there is a bijection between elements of degree d in I (henceforth denoted I_d) and R_{d-2} given by

$$\phi : I_d \to R_{d-2} \qquad x^2 f \mapsto f.$$

Thus since we have formula for the dimension R_d this now gives us a formula for the dimensions of I_d. We formalize this idea.

Definition 8. *Let $R = k[x_0, \ldots, x_n]$. Define $R(t)$, the shifted module, by $R(m)_i = R_{m+i}$.*

For example, $R(-2)$ is the module which in degree 0, is $R_{-2} = 0$ and in degree 1, is $R_{-1} = 0$, but in degree 3, $R(-3)_3 = R_0$ which is one dimensional. This is exactly the same as our ideal (x^2), so we can identify $(x^2) = R(-2)$. Computations with these shifted modules is exceedingly easy, as the next theorem suggests.

Theorem 3. *Let $R = k[x_0, \ldots, x_n]$. Then*

$$H_{R(m)}(d) = \binom{m+d+n}{n}.$$

Proof. The proof follows from the following string of equalities

$$H_{R(m)}(d) = \dim_k R(m)_d = \dim_k R_{m+d} = \binom{m+d+n}{n}.$$

□

One major use of shifts of modules is in computing Hilbert functions of ideals generated by an element of a certain degree. Indeed, if $I = (f)$ and f is of degree m then $I \cong R(-m)$