Math 533 Winter 2021, Lecture 16: Multivariate polynomials

website: `https://www.cip.ifi.lmu.de/~grinberg/t/21w/`

1. Polynomials II

We fix a commutative ring R. We shall now resume the study of polynomials.

1.1. Multivariate polynomials again

Recall from Lecture 13 the following fact:

Theorem 1.1.1. Let $m \in \mathbb{N}$. Let $b \in R[x]$ be a polynomial of degree m such that its leading coefficient $[x^m] b$ is a unit. Then, each element of $R[x]/b$ can be uniquely written in the form

$$a_0\overline{x^0} + a_1\overline{x^1} + \cdots + a_{m-1}\overline{x^{m-1}} \qquad \text{with } a_0, a_1, \ldots, a_{m-1} \in R.$$

Equivalently, the m vectors $\overline{x^0}, \overline{x^1}, \ldots, \overline{x^{m-1}}$ form a basis of the R-module $R[x]/b$. Thus, this R-module $R[x]/b$ is free of rank $m = \deg b$. If $m > 0$, then the ring $R[x]/b$ contains "a copy of R".

Thus we understand quotients of univariate polynomials rings rather well when the leading coefficient is a unit. They are less predictable when it is not a unit. If R is a field, however, then the leading coefficient of a nonzero polynomial $b \in R[x]$ is always a unit, so we don't need to worry about this issue.

But can we do this with multivariate polynomials?

Consider, for example, the two-variable polynomial ring $R[x,y]$. How does $R[x,y]/b$ look like for a polynomial $b \in R[x,y]$? Keep in mind that the "idea" behind quotienting out b is that we are setting b to 0. So $R[x,y]/b$ is "the ring of polynomials in x and y subject to the assumption that $b(x,y) = 0$".

Let us first try to answer this question for some special polynomials b; we will then look for a pattern. There is a lot to be learned from the examples.

1.1.1. Example 1: $R[x,y]/y$

What is $R[x,y]/y$? We expect this to be isomorphic to $R[x]$, because setting y to 0 in a polynomial $f(x,y)$ should give $f(x,0) \in R[x]$.

This is indeed true, and the formal proof is essentially just a formalization of this informal argument:

Proposition 1.1.2. We have $R[x,y]/y \cong R[x]$ as R-algebras.

Proof. Define a map

$$\begin{aligned} \alpha : R[x,y]/y &\to R[x], \\ \overline{f} &\mapsto f(x,0). \end{aligned}$$

First, we need to check that this map α is well-defined. In other words, we need to check the following:

> *Claim 1:* If $f, g \in R[x,y]$ are two polynomials satisfying $\overline{f} = \overline{g}$ in $R[x,y]/y$, then $f(x,0) = g(x,0)$.

[*Proof of Claim 1:* Let $f, g \in R[x,y]$ be two polynomials satisfying $\overline{f} = \overline{g}$ in $R[x,y]/y$. Then, $\overline{f} = \overline{g}$ means that $f - g \in yR[x,y]$; in other words, $f - g = yp$ for some polynomial $p \in R[x,y]$. Consider this p. Now, evaluating both sides of the equality $f - g = yp$ at $(x,0)$ (that is, substituting 0 for y) yields $f(x,0) - g(x,0) = 0p(x,0) = 0$ and thus $f(x,0) = g(x,0)$. This proves Claim 1.]

Having proved Claim 1, we thus know that the map α is well-defined. It is straightforward to see that α is an R-algebra morphism (because the map $R[x,y] \to R[x]$, $f \mapsto f(x,0)$ is an R-algebra morphism[1]).

In the opposite direction, define a map

$$\begin{aligned} \beta : R[x] &\to R[x,y]/y, \\ g &\mapsto \overline{g[x]}. \end{aligned}$$

It is again clear that this is an R-algebra morphism.

Now, we shall show that the maps α and β are mutually inverse. To prove this, we need to check that $\alpha \circ \beta = \mathrm{id}$ and $\beta \circ \alpha = \mathrm{id}$. Checking $\alpha \circ \beta = \mathrm{id}$ is the easy part. The "hard part" is showing that $\beta \circ \alpha = \mathrm{id}$. There are two ways to do this:

[*First proof of $\beta \circ \alpha = \mathrm{id}$:* To show this, we need to prove that $(\beta \circ \alpha)\left(\overline{f}\right) = \overline{f}$ for each $f \in R[x,y]$. So let us fix an $f \in R[x,y]$. Then,

$$\begin{aligned} (\beta \circ \alpha)\left(\overline{f}\right) &= \beta\left(\alpha\left(\overline{f}\right)\right) = \beta(f(x,0)) \qquad \left(\text{since } \alpha\left(\overline{f}\right) \text{ was defined to be } f(x,0)\right) \\ &= \overline{(f(x,0))[x]} \qquad \text{(by the definition of } \beta) \\ &= \overline{f(x,0)} \qquad (\text{since } (f(x,0))[x] = f(x,0)). \end{aligned}$$

Thus, it remains to show that $\overline{f(x,0)} = \overline{f}$ (because we want to show that $(\beta \circ \alpha)\left(\overline{f}\right) = \overline{f}$). In other words, it remains to show that $f - f(x,0) \in yR[x,y]$.

[1] We saw this (in a more general setting) in Lecture 11.

We do this directly: Write f in the form $f = \sum_{i,j \in \mathbb{N}} a_{i,j} x^i y^j$ (with $a_{i,j} \in R$). Then,

$$f(x,0) = \sum_{i,j \in \mathbb{N}} a_{i,j} x^i 0^j = \sum_{i \in \mathbb{N}} a_{i,j} x^i \underbrace{0^0}_{=1} + \sum_{\substack{i,j \in \mathbb{N}; \\ j>0}} a_{i,j} x^i \underbrace{0^j}_{\substack{=0 \\ (\text{since } j>0)}}$$

$$\left(\begin{array}{c} \text{here, we have split the sum into two parts:} \\ \text{one that contains all terms with } j = 0 \\ \text{and one that contains all terms with } j > 0 \end{array} \right)$$

$$= \sum_{i \in \mathbb{N}} a_{i,j} x^i = \sum_{\substack{i,j \in \mathbb{N}; \\ j=0}} a_{i,j} x^i y^j \qquad \left(\text{since } y^j = 1 \text{ for } j = 0 \right).$$

Subtracting this from $f = \sum_{i,j \in \mathbb{N}} a_{i,j} x^i y^j$, we find

$$f - f(x,0) = \sum_{i,j \in \mathbb{N}} a_{i,j} x^i y^j - \sum_{\substack{i,j \in \mathbb{N}; \\ j=0}} a_{i,j} x^i y^j = \sum_{\substack{i,j \in \mathbb{N}; \\ j>0}} a_{i,j} x^i \underbrace{y^j}_{\substack{=yy^{j-1} \\ (\text{we can do this} \\ \text{because } j>0)}}$$

$$= \sum_{\substack{i,j \in \mathbb{N}; \\ j>0}} a_{i,j} x^i y y^{j-1} = y \sum_{\substack{i,j \in \mathbb{N}; \\ j>0}} a_{i,j} x^i y^{j-1} \in yR[x,y],$$

as we wanted to prove. Thus, $\overline{f(x,0)} = \overline{f}$, so that $(\beta \circ \alpha)\left(\overline{f}\right) = \overline{f(x,0)} = \overline{f}$. This proves $\beta \circ \alpha = \mathrm{id}$.]

[*Second proof of $\beta \circ \alpha = \mathrm{id}$:* Here is a more "cultured" proof. We know that β and α are R-algebra morphisms, hence are R-linear maps. Thus, $\beta \circ \alpha$ and id are two R-linear maps from $R[x,y]/y$ to $R[x,y]/y$. Our goal is to prove that these two R-linear maps $\beta \circ \alpha$ and id are equal. As we have learned in Lecture 10, there is a shortcut for proving that two R-linear maps are equal: It suffices to pick a family of vectors that spans the domain (in our case, the R-module $R[x,y]/y$), and to show that the two maps agree on the vectors of this family. In our case, there is a rather natural choice of such a family: the family of monomials, or rather of their cosets. That is, we choose the family $\left(\overline{x^i y^j}\right)_{i,j \in \mathbb{N}}$. This family spans the R-module $R[x,y]/y$ (since the family $\left(x^i y^j\right)_{i,j \in \mathbb{N}}$ spans the R-module $R[x,y]$, and since the canonical projection onto $R[x,y]/y$ clearly preserves their spanning property). Thus, we only need to show that the two maps $\beta \circ \alpha$ and id agree on the vectors of this family – i.e., to show that

$$(\beta \circ \alpha)\left(\overline{x^i y^j}\right) = \mathrm{id}\left(\overline{x^i y^j}\right) \qquad \text{for any } i,j \in \mathbb{N}.$$

But this is straightforward: We fix $i,j \in \mathbb{N}$, and set out to show that $(\beta \circ \alpha)\left(\overline{x^i y^j}\right) = \mathrm{id}\left(\overline{x^i y^j}\right)$. If $j > 0$, then $\overline{x^i y^j} = 0$ (since $x^i y^j \in yR[x,y]$ in this case) and therefore

both $(\beta \circ \alpha)\left(\overline{x^i y^j}\right)$ and $\mathrm{id}\left(\overline{x^i y^j}\right)$ must be 0 in this case (since R-linear maps always send 0 to 0). If, on the other hand, $j = 0$, then $\overline{x^i y^j} = \overline{x^i y^0} = \overline{x^i}$ and therefore $\alpha\left(\overline{x^i y^j}\right) = \alpha\left(\overline{x^i}\right) = x^i$ (since substituting 0 for y does not change the monomial x^i) and thus $(\beta \circ \alpha)\left(\overline{x^i y^j}\right) = \beta\left(x^i\right) = \overline{x^i} = \overline{x^i y^j} = \mathrm{id}\left(\overline{x^i y^j}\right)$. Hence, in both cases, we have shown that $(\beta \circ \alpha)\left(\overline{x^i y^j}\right) = \mathrm{id}\left(\overline{x^i y^j}\right)$. This completes the proof of $\beta \circ \alpha = \mathrm{id}$.]

Either way, we have now shown that $\beta \circ \alpha = \mathrm{id}$. Combined with $\alpha \circ \beta = \mathrm{id}$, this yields that the two maps α and β are mutually inverse. Thus, α is an invertible R-algebra morphism, hence an R-algebra isomorphism. This proves Proposition 1.1.2. □

We can easily generalize this to multiple variables:

Proposition 1.1.3. For any $n > 0$, we have

$$R\left[x_1, x_2, \ldots, x_n\right] / x_n \cong R\left[x_1, x_2, \ldots, x_{n-1}\right] \text{ as } R\text{-algebras.}$$

Proof. Same idea as for Proposition 1.1.2, but requiring more subscripts to juggle. □

1.1.2. Example 2: $R\left[x, y\right] / \left(x^2 + y^2 - 1\right)$

How does $R\left[x, y\right] / \left(x^2 + y^2 - 1\right)$ look like?

This is a fairly useful R-algebra; it can be viewed as the algebra of polynomial functions on the unit circle. Indeed, any element $\overline{f} \in R\left[x, y\right] / \left(x^2 + y^2 - 1\right)$ can be "evaluated" at a point (a, b) on the unit circle (meaning, a pair of elements $a, b \in R$ with $a^2 + b^2 = 1$).

There are various interesting ring-theoretical questions to be asked about the quotient ring $R\left[x, y\right] / \left(x^2 + y^2 - 1\right)$; however, let us restrict ourselves to studying it as an R-module. As an R-module, is $R\left[x, y\right] / \left(x^2 + y^2 - 1\right)$ free? What is a basis? This boils down to asking whether (and how) we can divide polynomials with remainder by $x^2 + y^2 - 1$.

Here we will be helped by the following fact:

Proposition 1.1.4. We have

$$R\left[x, y\right] \cong \left(R\left[x\right]\right)\left[y\right] \qquad \text{as } R\text{-algebras.}$$

More concretely, the map

$$\varphi : R\left[x, y\right] \to \left(R\left[x\right]\right)\left[y\right],$$

$$\sum_{i,j \in \mathbb{N}} a_{i,j} x^i y^j \mapsto \sum_{j \in \mathbb{N}} \left(\sum_{i \in \mathbb{N}} a_{i,j} x^i\right) y^j \qquad (\text{where } a_{i,j} \in R)$$

is an R-algebra isomorphism.

Proof. First of all, you are excused for wondering what the deal is: Isn't the above map φ just the identity map, since $\sum_{j\in\mathbb{N}} \left(\sum_{i\in\mathbb{N}} a_{i,j} x^i \right) y^j$ is the same polynomial as $\sum_{i,j\in\mathbb{N}} a_{i,j} x^i y^j$ (just rewritten)?

Essentially yes, but there is a technical difference between the rings $R[x,y]$ and $(R[x])[y]$. The former is a polynomial ring in two indeterminates x, y over R, whereas the latter is a polynomial ring in one indeterminate y over the ring $R[x]$. Hence,

- the elements of $R[x,y]$ are polynomials in two variables x, y with coefficients in R, whereas
- the elements of $(R[x])[y]$ are polynomials in one variable y with coefficients in $R[x]$ (that is, their coefficients themselves are polynomials in one variable x over R).

Thus, even if a polynomial in $R[x,y]$ and a polynomial in $(R[x])[y]$ look exactly the same (such as, for example, the polynomials $2x^2y^3$ in both rings), they are technically different. (The polynomial $2x^2y^3$ in $R[x,y]$ has the monomial x^2y^3 appear in it with coefficient 2, whereas the polynomial $2x^2y^3$ in $(R[x])[y]$ has the monomial y^3 appear in it with coefficient $2x^2$.) The map φ thus sends each polynomial in $R[x,y]$ to the identically-looking polynomial in $(R[x])[y]$.

This being said, the claim we are proving is saying precisely that the difference between $R[x,y]$ and $(R[x])[y]$ is only a technicality; in essence the two rings are the same. The proof is rather straightforward. The simplest way is as follows: The map φ defined in the proposition is easily seen to be well-defined and an R-**module** isomorphism. Thus, it remains to prove that this map φ respects multiplication and respects the unity. It is clear enough that φ respects the unity (since the unities of both rings equal x^0y^0), so we only need to check that φ respects multiplication. According to the lemma from Lecture 11, it suffices to prove this on a family of vectors that spans the R-module $R[x,y]$; in other words, we only need to find a family $(m_i)_{i\in I}$ of vectors in $R[x,y]$ that spans $R[x,y]$, and show that

$$\varphi(m_i m_j) = \varphi(m_i)\,\varphi(m_j) \qquad \text{for all } i,j \in I.$$

Fortunately, the family of monomials $\left(x^i y^j\right)_{(i,j)\in\mathbb{N}^2}$ is such a family of vectors (even better, it is a basis of the R-module $R[x,y]$); thus, we only need to prove that

$$\varphi\left(x^i y^u \cdot x^j y^v\right) = \varphi\left(x^i y^u\right) \cdot \varphi\left(x^j y^v\right) \qquad \text{for all } (i,u),(j,v) \in \mathbb{N}^2.$$

But this is easy (the left and right hand sides both equal $x^{i+j}y^{u+v} \in (R[x])[y]$). Thus, we conclude that φ respects multiplication; as we said above, this completes the proof of Proposition 1.1.4. $\square$

Now, in view of Proposition 1.1.4, we have

$$R[x,y] / \left(x^2+y^2-1\right) \cong (R[x])[y] / \left(x^2+y^2-1\right) \tag{1}$$

(since the isomorphism φ from Proposition 1.1.4 sends the polynomial $x^2+y^2-1 \in R[x,y]$ to the identically-looking polynomial $x^2+y^2-1 \in (R[x])[y]$).

The ring on the right hand side of (1) is a quotient ring of the **univariate** polynomial ring $(R[x])[y]$ modulo the **monic** polynomial $x^2+y^2-1 = y^2 + \underbrace{\left(x^2-1\right)}_{\text{constant term in } R[x]}$ in the variable y. Thus, Theorem 1.1.1 (applied to 2, $R[x]$, y and x^2+y^2-1 instead of m, R, x and b) shows that this quotient ring $(R[x])[y] / \left(x^2+y^2-1\right)$ has a basis $\left(\overline{y^0}, \overline{y^1}\right)$ as an $R[x]$-module. This means that any element of $(R[x])[y] / \left(x^2+y^2-1\right)$ can be uniquely written as

$$\alpha\overline{y^0} + \beta\overline{y^1} \qquad \text{for some } \alpha, \beta \in R[x].$$

Since elements of $R[x]$ themselves can be uniquely written as R-linear combinations of powers of x, we thus conclude that any element of $(R[x])[y] / \left(x^2+y^2-1\right)$ can be uniquely written as

$$\begin{aligned}
&\left(\alpha_0x^0+\alpha_1x^1+\alpha_2x^2+\cdots\right)\overline{y^0} + \left(\beta_0x^0+\beta_1x^1+\beta_2x^2+\cdots\right)\overline{y^1} \\
&= \overline{\left(\alpha_0x^0+\alpha_1x^1+\alpha_2x^2+\cdots\right)y^0 + \left(\beta_0x^0+\beta_1x^1+\beta_2x^2+\cdots\right)y^1} \\
&= \alpha_0\overline{x^0y^0} + \alpha_1\overline{x^1y^0} + \alpha_2\overline{x^2y^0} + \cdots + \beta_0\overline{x^0y^1} + \beta_1\overline{x^1y^1} + \beta_2\overline{x^2y^1} + \cdots
\end{aligned}$$

for some $\alpha_0, \alpha_1, \alpha_2, \ldots, \beta_0, \beta_1, \beta_2, \ldots \in R$ (with all but finitely many of these coefficients $\alpha_0, \alpha_1, \alpha_2, \ldots, \beta_0, \beta_1, \beta_2, \ldots$ being 0). Thus, as an R-module, $(R[x])[y] / \left(x^2+y^2-1\right)$ has a basis

$$\left(\overline{x^0y^0}, \overline{x^1y^0}, \overline{x^2y^0}, \ldots, \overline{x^0y^1}, \overline{x^1y^1}, \overline{x^2y^1}, \ldots\right).$$

In view of the isomorphism (1) (which is an R-algebra isomorphism, and sends each $\overline{x^iy^j}$ to $\overline{x^iy^j}$), we can thus conclude that, as an R-module, $R[x,y] / \left(x^2+y^2-1\right)$ has a basis

$$\left(\overline{x^0y^0}, \overline{x^1y^0}, \overline{x^2y^0}, \ldots, \overline{x^0y^1}, \overline{x^1y^1}, \overline{x^2y^1}, \ldots\right). \tag{2}$$

The order in which we list the variables doesn't matter much in a polynomial ring; thus, Proposition 1.1.4 has the following analogue (which is proved similarly):

Proposition 1.1.5. We have

$$R[x,y] \cong (R[y])[x] \qquad \text{as } R\text{-algebras.}$$

More concretely, the map

$$\varphi : R[x,y] \to (R[y])[x],$$
$$\sum_{i,j\in\mathbb{N}} a_{i,j}x^i y^j \mapsto \sum_{i\in\mathbb{N}} \left(\sum_{j\in\mathbb{N}} a_{i,j}y^j\right) x^i \qquad (\text{where } a_{i,j} \in R)$$

is an R-algebra isomorphism.

Proposition 1.1.4 can also be generalized:

Proposition 1.1.6. For any $n > 0$, we have

$$R[x_1, x_2, \ldots, x_n] \cong (R[x_1, x_2, \ldots, x_{n-1}])[x_n] \text{ as } R\text{-algebras.}$$

Proof. Generalize the proof of Proposition 1.1.4 (same idea, more subscripts). □

1.1.3. More examples?

Having understood the R-modules $R[x,y]/y$ and $R[x,y]/\left(x^2+y^2-1\right)$, we move on to further examples.

How does $R[x,y]/(xy)$ look like? We cannot answer this using the methods used above, since the polynomial xy is neither monic in y when considered as a polynomial in $(R[x])[y]$ nor monic in x when considered as a polynomial in $(R[y])[x]$.

What about $R[x,y]/(xy(x-y))$? Can we divide $(x+y)^3$ by $xy(x-y)$ with remainder? What is the remainder? Should we replace x^2y by xy^2 or vice versa?

To make things more complicated (but also more useful), let's not forget that we can quotient a ring by an ideal, not just by a single element. Even if R is a field, the polynomial ring $R[x,y]$ is not a PID (unlike $R[x]$ for a field R), so not every ideal is principal.

The following shorthand will be useful:

Definition 1.1.7. Let S be a commutative ring. Let $a_1, a_2, \ldots, a_k$ be elements of S. Then, the ideal $a_1S + a_2S + \cdots + a_kS$ (this is the set of all S-linear combinations of $a_1, a_2, \ldots, a_k$) is called **the ideal generated by** $a_1, a_2, \ldots, a_k$. The quotient ring $S/(a_1S + a_2S + \cdots + a_kS)$ will be denoted by $S/(a_1, a_2, \ldots, a_k)$.

(Many authors actually write $(a_1, a_2, \ldots, a_k)$ for the ideal $a_1S + a_2S + \cdots + a_kS$, but this risks confusion since $(a_1, a_2, \ldots, a_k)$ also means the k-tuple.)

Informally, $S/(a_1, a_2, \ldots, a_k)$ is what is obtained from S if you set all of $a_1, a_2, \ldots, a_k$ to 0.

For an example, we can look at $R[x,y]/(x+y, x-y)$. This behaves differently depending on R:

- If $R = \mathbb{Q}$, then

$$R[x,y]/(x+y,x-y) = \mathbb{Q}[x,y]/(x+y,x-y) = \mathbb{Q}[x,y]/(x,y)$$

(since it is easy to see that the $\mathbb{Q}[x,y]$-linear combinations of $x+y$ and $x-y$ are precisely the $\mathbb{Q}[x,y]$-linear combinations of x and y), and thus

$$R[x,y]/(x+y,x-y) = \mathbb{Q}[x,y]/(x,y) \cong \mathbb{Q}.$$

- If $R = \mathbb{Z}/2$, then

$$\begin{aligned} R[x,y]/(x+y,x-y) &= (\mathbb{Z}/2)[x,y]/\left(\underbrace{x+y}_{\substack{=x-y \\ \text{(since we are in} \\ \text{characteristic 2)}}}, x-y\right) \\ &= (\mathbb{Z}/2)[x,y]/(x-y,x-y) \\ &= (\mathbb{Z}/2)[x,y]/(x-y) \cong (\mathbb{Z}/2)[x]. \end{aligned}$$

We can easily come up with more complicated examples:

What is $R[x,y,z]/\left(x^2-yz, y^2-zx, z^2-xy\right)$? What lies in the ideal $\left(x^2-yz\right)R[x,y,z] + \left(y^2-zx\right)R[x,y,z] + \left(z^2-xy\right)R[x,y,z]$?

What is $R[x,y,z]/\left(x^2+xy, y^2+yz, z^2+zx\right)$? What lies in the ideal $\left(x^2+xy\right)R[x,y,z] + \left(y^2+yz\right)R[x,y,z] + \left(z^2+zx\right)R[x,y,z]$? For example, I claim that z^4 lies in this ideal, but z^3 does not. How do I know? How can you tell?

In theory, you could imagine that there are ideals that do not even have a finite list of elements generating them. There are rings that have such ideals. For example, the polynomial ring $\mathbb{Z}[x_1, x_2, x_3, \ldots]$ in infinitely many variables has such ideals. But polynomial rings in finitely many variables over a field are not this bad. Indeed:

Theorem 1.1.8 (Hilbert's basis theorem)**.** Let F be a field. Let S be the polynomial ring $F[x_1, x_2, \ldots, x_n]$ for some $n \in \mathbb{N}$. Then, any ideal I of S is finitely generated (this means that there is a finite list $(a_1, a_2, \ldots, a_k)$ of elements of I such that $I = a_1 S + a_2 S + \cdots + a_k S$).

Proof. See [DF, §9.6, Corollary 22]. □

Warning: If $n = 1$, then the ideal I in Theorem 1.1.8 is principal (since $F[x_1]$ is a PID), so you can get by with a length-1 list (i.e., with $k = 1$). However, if $n = 2$, then the list can be arbitrarily large. You cannot always find a length-2 list. For example, in the polynomial ring $F[x, y]$, the ideal generated by all monomials of degree p (that is, by $x^p, x^{p-1}y, x^{p-2}y^2, \ldots, y^p$) cannot be generated by p or fewer elements.

1.2. Degrees and the deg-lex order

Let us now attempt a more general approach.

From now on, for the rest of this chapter, we fix a commutative ring R and an $n \in \mathbb{N}$.

We let P denote the polynomial ring $R[x_1, x_2, \ldots, x_n]$.

As we recall, a **monomial** is an element of the free abelian monoid $C^{(n)}$ with n generators $x_1, x_2, \ldots, x_n$; it has the form $x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ for some $(a_1, a_2, \ldots, a_n) \in \mathbb{N}^n$.

Our first goal is to define the degree of a polynomial in n variables. We begin by defining the degree of a monomial:

Definition 1.2.1. The **degree** of a monomial $\mathfrak{m} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \in C^{(n)}$ is defined to be the number $a_1 + a_2 + \cdots + a_n \in \mathbb{N}$. It is denoted by $\deg \mathfrak{m}$.

For example, the monomial $x_1^5 x_2 x_4^2 = x_1^5 x_2^1 x_3^0 x_4^2$ has degree $5 + 1 + 0 + 2 = 8$.

Definition 1.2.2. A monomial $\mathfrak{m}$ is said to **appear** in a polynomial $f \in P$ if $[\mathfrak{m}] f \neq 0$. (Recall that $[\mathfrak{m}] f$ means the coefficient of $\mathfrak{m}$ in f.)

For example, the monomial $x^2 y$ appears in $(x + y)^3 \in R[x, y]$ (if $3 \neq 0$ in R), but the monomial xy does not.

Definition 1.2.3. The **degree** (or **total degree**) of a nonzero polynomial $f \in P$ is the largest degree of a monomial that appears in f.

For example, the polynomial $(x + y + 1)^3$ has degree 3.

Definition 1.2.3 generalizes our old definition of degree for nonzero univariate polynomials.

The following proposition generalizes a fact that we proved for univariate polynomials in Lecture 12:

Proposition 1.2.4 (Degree-of-a-product formula)**.** Let R be a commutative ring. Let $p, q \in P$ be nonzero.

(a) We have $\deg(pq) \leq \deg p + \deg q$.

(b) We have $\deg(pq) = \deg p + \deg q$ if R is an integral domain.

Part **(a)** of this proposition is pretty clear. (The reason is that $\deg(\mathfrak{mn}) = \deg\mathfrak{m} + \deg\mathfrak{n}$ for any monomials $\mathfrak{m}, \mathfrak{n}$.)

What about part **(b)**? We proved this for univariate polynomials using leading coefficients. What is a leading coefficient when several monomials can have the same degree? In order to define it, we need to break ties (i.e., establish an ordering on monomials of equal degrees) in a way that will be compatible with products[2]. To that aim, we shall introduce a total order on the set $C^{(n)}$ of all monomials.

Recall that a **total order** (or, to be more precise, a **strict total order**) on a set S is a binary relation $\prec$ on S that is

- **asymmetric** (meaning that no two elements a and b of S satisfy $a \prec b$ and $b \prec a$ at the same time);
- **transitive** (meaning that if $a, b, c \in S$ satisfy $a \prec b$ and $b \prec c$, then $a \prec c$);
- **trichotomous** (meaning that for any two elements a and b of S, we have $a \prec b$ or $a = b$ or $b \prec a$).

Examples:

- The relation $<$ on the set $\mathbb{N}$ or on the set $\mathbb{Z}$ or on the set $\mathbb{R}$ is a total order.
- So is the relation $>$ on each of these three sets.
- If S is a finite set, and if $(s_1, s_2, \ldots, s_k)$ is a list of all elements of S, with each element of S appearing exactly once in this list, then we can define a total order $\prec$ on S as follows: We declare that two elements $u, v \in S$ satisfy $u \prec v$ if and only if u appears prior to v in this list $(s_1, s_2, \ldots, s_k)$ (that is, if $u = s_i$ and $v = s_j$ for some $i < j$).

If $\prec$ is a total order on a set S, then we view relations of the form $a \prec b$ as saying that a is in some sense smaller than b. We will use the notations $\preccurlyeq$, $\succ$ and $\succcurlyeq$ accordingly; this means that

- we write "$a \preccurlyeq b$" for "$a \prec b$ or $a = b$".
- we write "$a \succ b$" for "$b \prec a$".
- we write "$a \succcurlyeq b$" for "$a \succ b$ or $a = b$".

So we all know a total order on the set $\mathbb{R}$ of all real numbers. But what about other sets? For example, how can we find a total order on the set of words in the English language? A long time ago, creators of dictionaries and encyclopedias were faced with this very problem, because it would be hard

[2] I will explain what this means later.

to look a word up in a dictionary if there was no well-known total order in which the words appeared in the dictionary. The total order commonly used in dictionaries is known as the **lexicographic order** (or **dictionary order**): Words are ordered by their first letter (e.g., "ant" $\prec$ "bear"); ties are broken using the second letter ("ant" $\prec$ "armadillo"); remaining ties are broken using the third letter ("camel" $\prec$ "cat"); and so on; absent letters are treated as being smaller than present letters (e.g., "ant" $\prec$ "anteater"). We use this as an inspiration for defining a total order on $C^{(n)}$, but we shall use the degree as the first level of comparison.

Definition 1.2.5. We define a total order $\prec$ (called the **degree-lexicographic order**, or – for short – the **deg-lex order**) on the set $C^{(n)}$ of all monomials as follows:

For two monomials $\mathfrak{m} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ and $\mathfrak{n} = x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n}$, we declare that $\mathfrak{m} \prec \mathfrak{n}$ if and only if

- **either** $\deg \mathfrak{m} < \deg \mathfrak{n}$;
- **or** $\deg \mathfrak{m} = \deg \mathfrak{n}$ and the following holds: There is an $i \in \{1, 2, \ldots, n\}$ such that $a_i \neq b_i$, and the **smallest** such i satisfies $a_i < b_i$.

In words:

- If two monomials have different degrees, then we declare the monomial with smaller degree to be the smaller one.
- If they have equal degrees, then we look at the first variable that has different exponents in the two monomials, and we declare the monomial with the smaller exponent on this variable to be smaller.

For example:

- We have $x_1^2 \prec x_2 x_3^2$, since $\deg\left(x_1^2\right) = 2 < 3 = \deg\left(x_2 x_3^2\right)$.
- We have $x_1^5 x_2 x_3 x_4^2 \prec x_1^5 x_2 x_3^2 x_4$, since the two monomials have the same degree, and the first variable that has different exponents in these two monomials is x_3, and this variable appears with a smaller exponent in $x_1^5 x_2 x_3 x_4^2$ (namely, with exponent 1) than in $x_1^5 x_2 x_3^2 x_4$ (namely, with exponent 2).
- We have $x_1 x_3^2 \prec x_1 x_2 x_3$, since the first variable that has different exponents in these two monomials is x_2, and this variable appears with a smaller exponent in $x_1 x_3^2$ (namely, with exponent 0) than in $x_1 x_2 x_3$ (namely, with exponent 1).

- The reader may easily check that $x_3^3 \prec x_1x_2x_3^2 \prec x_1x_2^2x_3 \prec x_1^2x_2x_3 \prec x_3^5 \prec x_1^2x_2^2x_3^2$.

The deg-lex order has several good properties:

Proposition 1.2.6. **(a)** The deg-lex order really is a total order on $C^{(n)}$.

(b) If $\mathfrak{m}, \mathfrak{n}, \mathfrak{p} \in C^{(n)}$ satisfy $\mathfrak{m} \prec \mathfrak{n}$, then $\mathfrak{mp} \prec \mathfrak{np}$.

(c) We have $1 \preccurlyeq \mathfrak{m}$ for any $\mathfrak{m} \in C^{(n)}$.

(d) Let $\mathfrak{m} \in C^{(n)}$ be any monomial. Then, there are only finitely many monomials $\mathfrak{p}$ such that $\mathfrak{p} \prec \mathfrak{m}$.

(e) There are no infinite decreasing chains $\mathfrak{m}_0 \succ \mathfrak{m}_1 \succ \mathfrak{m}_2 \succ \cdots$ of monomials.

(f) If T is a nonempty finite set of monomials, then T has a largest element with respect to $\prec$ (that is, an element $\mathfrak{t} \in T$ such that $\mathfrak{m} \preccurlyeq \mathfrak{t}$ for all $\mathfrak{m} \in T$).

(g) If T is a nonempty set of monomials, then T has a smallest element with respect to $\prec$ (that is, an element $\mathfrak{t} \in T$ such that $\mathfrak{m} \succcurlyeq \mathfrak{t}$ for all $\mathfrak{m} \in T$).

Note that we require T to be finite in Proposition 1.2.6 **(f)** but not in Proposition 1.2.6 **(g)**. This is similar to the situation for sets of nonnegative integers (viz., any nonempty set of nonnegative integers has a smallest element, but only finite nonempty sets of nonnegative integers have largest elements).

Hints to the proof of Proposition 1.2.6. **(a)**, **(b)**, **(c)**, **(d)** LTTR.

(e) This follows from **(d)**.

(f) This holds for any total order on any set.

(g) This is easily proved using **(d)** (or, less easily, using **(e)**). LTTR. □

(Proposition 1.2.6 **(b)** is what I meant when I said that the deg-lex order is "compatible with products".)

Now, we can define leading coefficients of multivariate polynomials:

Definition 1.2.7. Let $f \in P$ be a nonzero polynomial.

(a) The **leading monomial** of f means the largest (with respect to $\prec$) monomial that appears in f. It is denoted by $\operatorname{LM} f$.

(b) The **leading coefficient** of f means the coefficient $[\operatorname{LM} f] f$. It is denoted by $\operatorname{LC} f$.

(c) The **leading term** of f means the product $\operatorname{LC} f \cdot \operatorname{LM} f$. It is denoted by $\operatorname{LT} f$.

For example, if $3 \neq 0$ in R, then

$$\operatorname{LM}\left(\left(x_1 + x_2 + 1\right)^3 - x_1^3\right) = x_1^2 x_2;$$
$$\operatorname{LC}\left(\left(x_1 + x_2 + 1\right)^3 - x_1^3\right) = 3;$$
$$\operatorname{LT}\left(\left(x_1 + x_2 + 1\right)^3 - x_1^3\right) = 3x_1^2 x_2.$$

Two simple consequences of this definition are:

Proposition 1.2.8. Let $f \in P$ be a nonzero polynomial. Then, $f - \operatorname{LT} f = 0$ or else $\operatorname{LM}(f - \operatorname{LT} f) \prec \operatorname{LM} f$.

Proof. By Definition 1.2.7, we have

$$f = \operatorname{LT} f + (\text{an } R\text{-linear combination of monomials } \mathfrak{m} \text{ with } \mathfrak{m} \prec \operatorname{LM} f).$$

Hence, $f - \operatorname{LT} f$ is an R-linear combination of monomials $\mathfrak{m}$ with $\mathfrak{m} \prec \operatorname{LM} f$. Therefore, $f - \operatorname{LT} f = 0$ or else $\operatorname{LM}(f - \operatorname{LT} f) \prec \operatorname{LM} f$. □

Proposition 1.2.9. Let $f, g \in P$ be nonzero polynomials such that $\operatorname{LC} f$ is not a zero-divisor in R. Then,

$$\operatorname{LM}(fg) = \operatorname{LM} f \cdot \operatorname{LM} g \qquad \text{and} \qquad \operatorname{LC}(fg) = \operatorname{LC} f \cdot \operatorname{LC} g.$$

Proof. LTTR. (Use Proposition 1.2.6 **(b)**.) □

Now we can easily prove Proposition 1.2.4 **(b)**. (The details are LTTR.) From Proposition 1.2.4, we obtain the following:

Corollary 1.2.10. If R is an integral domain, then the polynomial ring $P = R[x_1, x_2, \ldots, x_n]$ is an integral domain.

1.3. Division with remainder and Gröbner bases

By defining leading monomials and leading coefficients, we have recovered one piece of the nice theory of univariate polynomials in the multivariate case. Can we do more? Can we define division with remainder?

1.3.1. The case of principal ideals

We **can** divide with remainder by a single polynomial[3]:

Theorem 1.3.1. Let $b \in P$ be a nonzero polynomial whose leading coefficient $\operatorname{LC} b$ is a unit of R. Let $a \in P$ be any polynomial.

Then, there is a **unique** pair (q, r) of polynomials in P such that

$$a = qb + r \qquad \text{and} \qquad r \text{ is } \operatorname{LM} b\text{-reduced}.$$

Here, a polynomial $r \in P$ is said to be $\mathfrak{m}$**-reduced** (where $\mathfrak{m}$ is a monomial) if no monomial divisible by $\mathfrak{m}$ appears in r.

[3]Recall that $P = R[x_1, x_2, \ldots, x_n]$.

This generalizes the division-with-remainder theorem for univariate polynomials; indeed, if $n = 1$, then the condition "r is $\operatorname{LM} b$-reduced" is equivalent to "$\deg r < \deg b$" (which is familiar from the case of univariate polynomials).

Let us illustrate Theorem 1.3.1 on an example:

- Let $n = 2$ and $R = \mathbb{Z}$, and let us rename the indeterminates x_1, x_2 as x, y. Thus, $P = \mathbb{Z}[x, y]$. Let $b = xy(x-y) \in P$. Thus, $\operatorname{LM} b = x^2 y$ and $\operatorname{LC} b = 1$.

  Let $a = (x+y)^4$. We want to divide a by b with remainder. That is, we want to find the pair (q, r) in Theorem 1.3.1.

  Theorem 1.3.1 says that "a can be written as a multiple of b plus some $\operatorname{LM} b$-reduced polynomial". In other words, it says that by subtracting an appropriate multiple of b from a, we can obtain an $\operatorname{LM} b$-reduced polynomial. How do we find the right multiple to subtract?

  In the univariate case, "$\operatorname{LM} b$-reduced" was simply saying that $\deg r < \deg b$, and we achieved this by repeatedly subtracting multiples of b from a in order to chip away at the leading term (reducing the degree by at least 1 in each step). We can do this similarly in the multivariate case: We simply check whether a is already $\operatorname{LM} b$-reduced. As long as it isn't, we find some monomial divisible by $\operatorname{LM} b$ that appears in a, and we clear it out by subtracting an appropriate multiple of b (so that this monomial no longer appears in a). More precisely, we clear out the highest such monomial that appears in a. We keep doing this until no such monomials remain (which means that a has become $\operatorname{LM} b$-reduced).

  Let us actually do this in our above example: We start with

$$a = (x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4.$$

  Two monomials that are multiples of $\operatorname{LM} b = x^2 y$ appear on the right hand side: $x^3 y$ and $x^2 y^2$. The highest of them is $x^3 y$, so we clear it out by subtracting an appropriate multiple of b. This appropriate multiple is $4xb$, since we want to clear out a $4x^3 y$ term. So we get

$$\begin{aligned} a - 4xb &= \left(x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4\right) - 4x \cdot xy(x-y) \\ &= x^4 + 10x^2y^2 + 4xy^3 + y^4. \end{aligned}$$

  Now we still have one monomial left that is a multiple of $\operatorname{LM} b = x^2 y$, namely $x^2 y^2$. We clear it out by subtracting $10yb$, and we end up with

$$\begin{aligned} a - 4xb - 10yb &= \left(x^4 + 10x^2y^2 + 4xy^3 + y^4\right) - 10y \cdot xy(x-y) \\ &= x^4 + 14xy^3 + y^4. \end{aligned}$$

The right hand side of this equality is $\operatorname{LM} b$-reduced, so it is the remainder we were looking for. That is, the r in our pair (q, r) is $x^4 + 14xy^3 + y^4$. The q in this pair we find by collecting the multiples of b that we have subtracted; thus, we get $q = 4x + 10y$. Hence, our pair (q, r) is

$$(q, r) = \left(4x + 10y, x^4 + 14xy^3 + y^4\right).$$

Hints to the proof of Theorem 1.3.1. The existence of the pair (q, r) is proved by the same idea as in the example we just did. All we need to do is to explain why our procedure terminates (rather than running forever). This is not hard: We observe that, as we keep subtracting appropriate multiples of b from a, the **highest** monomial that is a multiple of $\operatorname{LM} b$ and appears in a becomes smaller and smaller (because each subtraction clears out the highest such monomial, and can only introduce lower such monomials). Thus, if our procedure would run forever, then we would obtain an infinite decreasing chain $\mathfrak{m}_0 \succ \mathfrak{m}_1 \succ \mathfrak{m}_2 \succ \cdots$ of monomials; but this would contradict Proposition 1.2.6 **(e)**. Thus, the algorithm eventually terminates, and this proves the existence of (q, r).

To prove the uniqueness of (q, r), it suffices to show that no nonzero multiple of b is $\operatorname{LM} b$-reduced[4]. But this follows easily from Proposition 1.2.9. □

As a consequence of Theorem 1.3.1 (or, more precisely, of the algorithm for the construction of (q, r) that we demonstrated in the above example), we obtain an algorithmic way to tell whether a polynomial $a \in P$ is divisible by b or not (whenever $b \in P$ is a nonzero polynomial whose leading coefficient $\operatorname{LC} b$ is a unit of R). Namely, we compute the pair (q, r) from Theorem 1.3.1, and check whether $r = 0$. The uniqueness of this pair easily yields that $b \mid a$ if and only if $r = 0$.

Another consequence of Theorem 1.3.1 is the following theorem that characterizes the R-module P/b:

Corollary 1.3.2. Let $b \in P$ be a nonzero polynomial whose leading coefficient $\operatorname{LC} b$ is a unit of R. Then, each element of P/b can be uniquely written in the form

$$\sum_{\substack{\mathfrak{m} \text{ is a monomial} \\ \text{not divisible by } \operatorname{LM} b}} a_{\mathfrak{m}} \overline{\mathfrak{m}} \qquad \text{with } a_{\mathfrak{m}} \in R$$

(where all but finitely many $\mathfrak{m}$ satisfy $a_{\mathfrak{m}} = 0$). Equivalently, the family $\left(\overline{\mathfrak{m}}\right)_{\mathfrak{m} \text{ is a monomial not divisible by } \operatorname{LM} b}$ is a basis of the R-module P/b. If b is not constant, then the ring P/b contains "a copy of R".

Corollary 1.3.2 generalizes Theorem 1.1.1 (and is proved in the same way, except that we use Theorem 1.3.1 instead of the univariate division-with-remainder theorem). Here are some examples:

[4] Indeed, if (q_1, r_1) and (q_2, r_2) are two pairs (q, r) satisfying the claim of Theorem 1.3.1, then $r_1 - r_2 = (q_2 - q_1)\, b$ is a multiple of b that is $\operatorname{LM} b$-reduced (since r_1 and r_2 are $\operatorname{LM} b$-reduced).

- Let us take $P = R[x,y]$ and $b = y$ in Corollary 1.3.2. Then, $\operatorname{LM} b = y$, so that Corollary 1.3.2 yields that the family $\left(\overline{\mathfrak{m}}\right)_{\mathfrak{m} \text{ is a monomial not divisible by } y}$ is a basis of the R-module $P/b = R[x,y]/y$. Since the monomials not divisible by y are precisely the powers of x (that is, $x^0, x^1, x^2, \ldots$), we can rewrite this as follows: The family $\left(\overline{x^i}\right)_{i \in \mathbb{N}} = \left(\overline{x^0}, \overline{x^1}, \overline{x^2}, \ldots\right)$ is a basis of the R-module $P/b = R[x,y]/y$. This is in line with Proposition 1.1.2 (indeed, the isomorphism $R[x,y]/y \to R[x]$ sends this family to the standard basis $\left(x^0, x^1, x^2, \ldots\right)$ of $R[x]$).

- Let us take $P = R[x,y]$ and $b = x^2 + y^2 - 1$ in Corollary 1.3.2. Then, $\operatorname{LM} b = x^2$, so that Corollary 1.3.2 yields that the family $\left(\overline{\mathfrak{m}}\right)_{\mathfrak{m} \text{ is a monomial not divisible by } x^2}$ is a basis of the R-module $P/b = R[x,y]/\left(x^2+y^2-1\right)$. Since the monomials not divisible by x^2 are precisely the monomials $x^i y^j$ with $i < 2$, we can rewrite this as follows: The family

$$\left(\overline{x^i y^j}\right)_{(i,j) \in \mathbb{N}^2;\ i<2} = \left(\overline{x^0 y^0}, \overline{x^0 y^1}, \overline{x^0 y^2}, \ldots, \overline{x^1 y^0}, \overline{x^1 y^1}, \overline{x^1 y^2}, \ldots\right)$$

  is a basis of the R-module $P/b = R[x,y]/\left(x^2+y^2-1\right)$. This is not the basis that we obtained back in (2), but rather is obtained from the latter by interchanging x and y. Of course, it is no surprise that interchanging x and y turns a basis into a basis; indeed, the variables x and y clearly play symmetric roles in $R[x,y]/\left(x^2+y^2-1\right)$, so every basis that treats them unequally has a "mirror" version with x and y interchanged.

- Let us take $P = R[x,y]$ and $b = xy$ in Corollary 1.3.2. Then, $\operatorname{LM} b = xy$, so that Corollary 1.3.2 yields that the family $\left(\overline{\mathfrak{m}}\right)_{\mathfrak{m} \text{ is a monomial not divisible by } xy}$ is a basis of the R-module $P/b = R[x,y]/(xy)$. Since the monomials not divisible by xy are precisely the monomials $1, x^1, x^2, x^3, \ldots, y^1, y^2, y^3, \ldots$ (that is, the monomials that are powers of a single indeterminate), we can rewrite this as follows: The family

$$\left(\overline{1}, \overline{x^1}, \overline{x^2}, \overline{x^3}, \ldots, \overline{y^1}, \overline{y^2}, \overline{y^3}, \ldots\right)$$

  is a basis of the R-module $P/b = R[x,y]/(xy)$. This can be obtained in more direct ways, too.

- Likewise, applying Corollary 1.3.2 to $P = R[x,y]$ and $b = xy(x-y)$ yields that the family

$$\begin{aligned}&\left(\overline{\mathfrak{m}}\right)_{\mathfrak{m} \text{ is a monomial not divisible by } x^2 y} \\ &= \left(\overline{1}, \overline{x^1}, \overline{x^2}, \overline{x^3}, \ldots, \overline{y^1}, \overline{y^2}, \overline{y^3}, \ldots, \overline{xy^1}, \overline{xy^2}, \overline{xy^3}, \ldots\right)\end{aligned}$$

  is a basis of the R-module $R[x,y]/(xy(x-y))$.