Lecture 7

Polynomial rings

7.1 Goals

(a) Become familiar with computations of products and sums of polynomials.

(b) See polynomial rings in many variables.

(c) Begin to utilize properties of special rings to prove statements about polynomial rings (by doing the exercises.)

7.2 Coming attraction: Algebraic geometry

We've seen the following examples of commutative rings:

$$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, M_n(R),$$

where the last example—n-by-n matrices with entry in a ring R—is never commutative when $n \geq 2$. We've studied most $M_2(\mathbb{R})$, because it acts in a geometrically interesting way on $\mathbb{R}^2$.

Soon, we're going to start seeing a completely *different* way that rings and geometry interact.

The general philosophy is that any kind of shape—which we often call a space in mathematics—can be understood by the collection of functions on the shape. And where do rings come in? The point is that the most basic collection of functions on a shape always forms a ring—we can add and multiply functions. And, in fact, it is most common to be in a setting where the ring of function is commutative. So commutative rings emerge very naturally in geometry.

It's fair to say that we can classify different branches of math by classifying what kinds of shapes we study, and what kinds of functions we study between shapes. Today we're going to foray into *algebraic geometry*, where "algebraic" is a proxy for "polynomial." So we'll want to understand what we mean by polynomials.

Moreover, polynomial rings are great examples of rings. They are simple to define, and easy to ask questions about—but some questions are very hard to answer. These make for the best math problems. (See Remark 8.3.7.)

7.3 Polynomials in one variable

Let's begin by fixing a commutative ring R. We call this the *base ring.* For visualization purposes, we'll often choose our base ring to be $\mathbb{R}$.

Remark 7.3.1. In algebraic geometry, the most common base ring is $\mathbb{C}$, and sometimes $\mathbb{Z}$. But it takes some sophistication to understand the sense in which there is geometry in studying $\mathbb{Z}$. As for $\mathbb{C}$, it's harder to visualize $\mathbb{C}^2$ (which is equivalent to $\mathbb{R}^4$) because we don't all have practice understanding 4-dimensional space. On the other hand, a nice way to think about algebraic geometry is that it gives us tools for understanding shapes in four-dimensional space without having to visualize them. We probably won't get a chance to study what I mean by this.

Definition 7.3.2. A *polynomial over R*, or a *polynomial with coefficients in R*, is an expression

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

where each $a_i \in R$. As usual, we say that a_i is the *ith degree coefficient* of the polynomial.

Notation 7.3.3. We may write the above polynomial using summation notation as:

$$\sum_{i=0}^{n} a_i x^i.$$

Finally, we will often write p instead of $p(x)$, for brevity.

Remark 7.3.4. Books often say that x is a "formal variable," but this is a meaningless phrase. We use x as a useful visual placeholder to remind us what degree the coefficient a_i corresponds to, and also to remind us of the use of x in high school algebra.

Notation 7.3.5. Fix a commutative base ring R. We let $R[x]$ denote the set of all polynomials over R.

Example 7.3.6. Here are examples of elements of $\mathbb{Z}[x]$:

$$-3+2x-9x^4, \qquad 9x^3, \qquad 13x^4-x^{1000}, \qquad 0.$$

Because $\mathbb{Z} \subset \mathbb{R}$, the above are all also elements of $\mathbb{R}[x]$. Here are examples of elements in $\mathbb{R}[x]$ that are not in $\mathbb{Z}[x]$:

$$\pi, \qquad 1-\pi x^3, \qquad \sqrt{7}x^4, \qquad \frac{1}{e}x.$$

Finally, here are examples of elements in $\mathbb{C}[x]$ that are not in $\mathbb{R}[x]$:

$$\sqrt{5}i+7x, \qquad -i, \qquad 9x^2-ix^3+19x^5.$$

Remark 7.3.7. Note that, even though we tend to write an undetermined polynomial as $a_0 + a_1x + \ldots + a_nx^n$, when we write out a polynomial, we often write $1-3x$ rather than $1+(-3)x$. This is a common convention you are used to, and we do it often to save space. As we will see, this is notation that is also consistent with the fact that $(-3)x$ is in fact the additive inverse to $3x$. So $-(3x)$ is the same thing as $(-3)x$.

Remark 7.3.8. You can think of a polynomial over R just as an ordered collection of elements of R: The data of $a_0 + a_1x + \ldots a_nx^n$ is equivalent to the data of the ordered n-tuple $(a_0, a_1, \ldots, a_n) \in R^n$.

In fact, it's even healthy to think about the set $R[x]$ as the subset of R^∞ for which only finitely many coordinates are non-zero.

(Take a moment to think about this if you haven't yet. By R^∞, we mean the collection of all functions from the set of non-negative integers $\{0, 1, 2, \ldots\}$ to the set R. This is just as R^n can be thought of as the collection of all functions from the set $\{1, 2, \ldots, n\}$ to R. Finally, note that for an element of R^∞ to have only finitely many non-zero coordinates means that we can find a largest n so that, whenever $i > n$, $a_i = 0$.)

Definition 7.3.9. Let p be a polynomial over R. Then the *degree* of p is $\max\{i \,|a_i \neq 0\}$. In other words, the degree of p is the largest index i for which the ith coefficient is non-zero.

Warning 7.3.10. By convention, we will declare the maximum of the empty set to be $-\infty$. In particular, the degree of the 0 polynomial is declared to be $-\infty$. All other polynomials have some non-negative integer degree.

Example 7.3.11. The first four polynomials in Example 7.3.6 have degrees 4, 3, 1000, and $-\infty$.

7.3.1 Addition and multiplication of polynomials

Definition 7.3.12 (Addition of polynomials)**.** One can add two polynomials. If $q(x) = b_0 + b_1x + \ldots + b_mx^m$ is another polynomial and $m \leq n$ we define

$$p(x)+q(x) = (a_0+b_0)+(a_1+b_1)x+\ldots+(a_m+b_m)x^m+a_{m+1}x^{m+1}+\ldots+a_nx^n.$$

In summation notation,

$$p(x) + q(x) = \sum_{i=0}^{n}(a_i + b_i)x^i$$

where it's understood that $b_i = 0$ for all $i > m$.

Remark 7.3.13. Note that the definition of polynomial addition depends only on the definition of addition in R. In fact, you can think of it as vector addition in R^∞. (See Remark 7.3.8.) For example, if you think of two polynomials as encoding the ordered tuples $(a_0, a_1, \ldots)$ and $(b_0, b_1, \ldots)$, then the sum of polynomials is indeed the vector sum

$$(a_0 + b_0, a_1 + b_1, \ldots).$$

Example 7.3.14. Here are some examples of polynomial addition:

1. $p(x) = 7x + 9x^3, q(x) = 8x \implies p(x) + q(x) = 15x + 9x^3.$
2. $p(x) = 7x + 9x^3, q(x) = -7x - 9x^3 \implies p(x) + q(x) = 0.$
3. $p(x) = 7x+9x^3, q(x) = \pi+\sqrt{2}x^5 \implies p(x)+q(x) = \pi+7x+9x^3+\sqrt{2}x^5.$

Note that the first two additions could be construed as taking place in $\mathbb{Z}[x], \mathbb{Q}[x], \mathbb{R}[x]$, or $\mathbb{C}[x]$. However, the last addition takes place in either $\mathbb{R}[x]$ or $\mathbb{C}[x]$.

Definition 7.3.15 (Multiplication of polynomials)**.** If $p(x) = a_0 + a_1x + \ldots a_nx^n$ and $q(x) = b_0 + b_1x + \ldots + b_mx^m$ are polynomials over R, we define the product of p and q to be

$$p(x)q(x) = (a_0b_0) + (a_0b_1 + a_0b_1)x + \ldots + a_nb_mx^{m+n}.$$

In summation notation,

$$p(x)q(x) = \sum_{i=0}^{m+n} \left(\sum_{j+k=i} (a_jb_k) \right) x^i$$

where it's understood that $b_i = 0$ for all $i > m$.

Remark 7.3.16. Note that multiplication depends both on multiplication and addition in the base ring R. For example, the ith coefficient is given by multiplying a_j and b_k, then summing the terms a_jb_k over all j, k satisfying $j + k = i$.

Remark 7.3.17. We have so far seen examples of rings where addition looks just like vector addition—$\mathbb{C}$ and $M_n(\mathbb{R})$—but where multiplication is interesting. The same is true for polynomials. (You may have never thought of polynomial multiplication as interesting, but it's certainly more interesting than polynomial addition.)

Example 7.3.18. For explicitness, here are the degree 2, 3, and 4 terms of the product pq:

$$(a_0b_2 + a_1b_1 + a_2b_0)x^2 + (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0)x^3 \\ + (a_0b_4 + a_1b_3 + a_2b_2 + a_3b_1 + a_4b_0)x^4.$$

Example 7.3.19. Let $p(x) = 1 - 3x + 7x^2$ and $q(x) = 8 - 5x^3$. Then pq is the polynomial

$$8 - 24x + 56x^2 - 5x^3 + 15x^4 - 35x^5.$$

Let $r(x) = 2 - 2x$. Then

$$p(x)r(x) = 2 - 8x + 20x^2 - 14x^3.$$

7.3.2 $R[x]$ as a ring

We wouldn't talk about the "polynomial ring" if it weren't a ring, would we?

Theorem 7.3.20. Fix a commutative ring R. Then $R[x]$—with the operations of addition and multiplication defined above—is a commutative ring.

Remark 7.3.21 (Notation)**.** To be explicit:

1. The additive identity of $R[x]$ is the polynomial whose coefficients are all 0. (So $a_i = 0$ for all i.) We write this polynomial as 0, as we have already done in earlier examples.

2. The multiplicative identity of $R[x]$ is the polynomial whose 0th degree coefficient is 1, and whose other coefficients are all 0. (So $a_0 = 1$, and for all $i \geq 1$, we have $a_i = 0$.) We write this polynomial as 1.

3. Given a polynomial $p(x) = a_0 + a_1 x + \ldots a_n x^n$, the additive inverse $-p$ is the polynomial $-a_0 + (-a_1)x + \ldots + (-a_n)x^n$, otherwise written as $-a_0 - a_1 x - a_2 x^2 - \ldots - a_n x^n$.

We'll also note that proving commutativity of multiplication, associativity of multiplication, and distributivity, all come down to (i) the corresponding properties of R, and (ii) clear use of summation notation.

Remark 7.3.22. When you see an element like $53x \in R[x]$, you can also interpret it as the product of the element 53 (which is a degree 0 polynomial) and x (which is a degree 1 polynomial). Likewise, $12x^2$ can be interpreted also as the product of 12 and x^2, or as the product of $3x$ with $4x$.

Similarly for addition: While we have said that "a polynomial is an expression" like $4 + 3x - 7x^2$, we can also interpret this expression literally: 4, $3x$, and $-7x^2$ are all elements of $R[x]$, and this polynomial is the sum of these three elements.

7.3.3 Degrees

The following notion turns out to be very useful:

Definition 7.3.23. Let R be a commutative ring. R is called an *integral domain* if "$a \neq 0$ and $b \neq 0 \implies ab \neq 0$. Equivalently, R is an integral domain if whenever the product ab equals 0, one may conclude that at least one of a or b equals zero.

Example 7.3.24. Every commutative ring we have encountered so far is an integral domain. (See Exercise 7.5.4.)

$M_2(\mathbb{R})$ is not an integral domain for at least two reasons. First, it is not commutative. Second, there exists non-zero matrices A, B for which $AB = 0$.

The degree of a polynomial (Definition 7.3.9) acts in a fairly controlled way with respect to addition and multiplication, especially when R is an integral domain:

Proposition 7.3.25. Fix a commutative base ring R. Let p and q be degree n and m polynomials, respectively, over R. Then

1. $\deg(p+q) \leq \max(\deg(p), \deg(q))$. In other words, the degree of $p+q$ is at most the bigger of the degree of p and q.

2. Suppose that R is an integral domain. Then $\deg(pq) = \deg(p)+\deg(q)$.

Example 7.3.26. If $p = 0$, then $pq = 0$, so $\deg(pq) = -\infty$. By convention, we will say that $-\infty + m = -\infty$ for any value of m.

If $p = 2 - x^2$ and $q = 2 + 3x + x^2$, then $\deg(p + q) = \deg(4 + 3x) = 1$, and 1 is less than or equal to $\max(\deg(p), \deg(q)) = \max(2, 2) = 2$.

7.4 Polynomials in finitely many variables

Do you remember that, for *any* ring R, $M_n(R)$ is a ring again? (Theorem 3.5.2.) This allows us to contemplate rings like $M_2(M_2(M_2(\mathbb{R})))$ (though we haven't yet).

Theorem 7.3.20 lets us do the same thing: Fix a commutative ring R. Then $R[x]$ is a commutative ring. So we can iterate the construction to obtain a new commutative ring!

We'll clean up some notation to do this. While we've used $R[x]$ to denote a polynomial ring whose formal variable is x, we could just as well define a ring $R[y]$ whose formal variable is y. The idea that we can use different variables is useful if we want to iteratively study polynomial rings—$R[x][y]$ is much better notation than $R[x][x]$.

Remark 7.4.1. That we use x and y for variables is a historical convention. One could easily have used a symbol like $\heartsuit$ and written a polynomial as $a_0 + a_1\heartsuit + a_2\heartsuit^2 + \ldots + a_n\heartsuit^n \in R[\heartsuit]$.

Let's study what $(R[x])[y]$ is, then. An element of this set is a polynomial (whose formal variable is denoted y) with coefficients in $R[x]$. In other words, an element can be written as

$$p_0 + p_1 y + p_2 y^2 + \dots p_n y^n \tag{7.4.0.1}$$

where each $p_i \in R[x]$. For every polynomial p_i, let's write out its coefficients as follows:

$$p_i(x) = a_{0,i} + a_{1,i}x + a_{2,i}x^2 + \dots.$$

Then we can write (7.4.0.1) explicitly as

$$\begin{aligned}(a_{0,0} + a_{1,0}x + a_{2,0}x^2 + \dots) &+ (a_{0,1} + a_{1,1}x + a_{2,1}x^2 + \dots)y \\ &+ (a_{0,2} + a_{1,2}x + a_{2,2}x^2 + \dots)y^2 \\ &+ \dots \\ &+ (a_{0,n} + a_{1,n}x + a_{2,n}x^2 + \dots)y^n.\end{aligned}$$

Distributing out the variables, and re-ordering the addition (remember that $(R[x])[y]$ is a ring by Theorem 7.3.20), we end up with the expression

$$a_{0,0} + a_{1,0}x + a_{0,1}y + a_{2,0}x^2 + a_{1,1}xy + a_{0,2}y^2 + \dots$$

which is far more neatly organized using summation notation:

$$\sum a_{i,j} x^i y^j.$$

(I am being lazy and not writing the bounds of the summation. This is to reduce clutter—I don't want to make new notation for the upper bounds. But, do be careful: The summation is indexed over both different values of i, and over different values of j.)

The upshot : One can interpret an element of $(R[x])[y]$ as a polynomial in *two* variables. Note also that we see a natural bijection between $(R[y])[x]$ and $(R[x])[y]$. So let's introduce the following notation:

Notation 7.4.2. Let R be a commutative ring. We let

$$R[x, y]$$

denote the set of polynomials in two variables x and y, with coefficients in R. More generally, we let

$$R[x, y, z]$$

denote the set of three-variable polynomials, and when we would rather use symbols like $x_1, x_2, \ldots$ rather than $x, y, \ldots$, we let

$$R[x_1, x_2, \ldots, x_n]$$

denote the set of polynomials in n variables (with coefficients in R). Note that all of these are commutative rings by iterated applications of Theorem 7.3.20.

Example 7.4.3. The following are elements of $\mathbb{Z}[x, y]$:

$$0, \qquad 1, \qquad x+x^2-x^3, \qquad y^2-y^3, \qquad 3+x+y-8xy, \qquad 16x-9y^2+17xy^3.$$

The following are elements of $\mathbb{Z}[x, y, z]$:

$$6, \qquad z+z^2, \qquad y^2-y^3, \qquad 3+x+y-8xz, \qquad 16xyz-9y^2.$$

And the following are elements of $\mathbb{Z}[x_1, x_2, x_3, x_4]$:

$$8-x_1+x_4, \qquad x_1x_2x_3x_4+x_1x_3^3-x_4^9, \qquad x_1x_2^2x_3^9.$$

Remark 7.4.4. Fix a commutative ring R. By induction on n, the collection of polynomials $R[x_1, \ldots, x_n]$ forms a commutative ring. (The base case $n = 0$ is the assumption that R is commutative. The inductive step uses Theorem 7.3.20 and the identification

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

(which we exposited in the discussion preceding Notation 7.4.2). Using this identification, you can confirm that addition and multiplication is what you think it is. For example, we have that

$$(x_1 - 3x_1x_2 + x_3^2)(1 + x_2) = x_1 - 2x_1x_2 - 3x_1x_2^2 + x_3^2 + x_2x_3^2.$$

7.5 Exercises

Exercise 7.5.1. For each polynomial p below, state what the degree 4 coefficient of p^2 is.

(a) $p(x) = 8 - 2x + 6x^2 - 5x^3 + x^4 - 35x^5$.

(b) $p(x) = 5x^3 + 15x^4 - 35x^5$.

(c) $p(x) = 8 - 2x + 6x^2$.

Exercise 7.5.2. For each polynomial p below, state what the coefficient of the x^2y^3 term of p^2 is.

(a) $p(x, y) = 8 - x^2 - \pi xy + 7xy^2 + 5y^3 + 3x^2y^3 + x^6 - y^8 + x^2y^7$.

(b) $p(x, y) = 8 - x^2 - \pi xy + 5y^3$.

(c) $p(x, y) = 5y^3 + 3x^2y^3 + x^6 - y^8 + x^2y^7$.

Exercise 7.5.3. Prove Proposition 7.3.25.

Exercise 7.5.4. Show that if R is an integral domain, then $R[x]$ is an integral domain. (Hint: This may require some *degree* of contemplation.)

Exercise 7.5.5. Suppose R is an integral domain, and fix an element $a \in R$. Show that there are at most two elements in R that square to a. (Hint: For any two elements x, y in a commutative ring, the distributive property implies that $x^2 - y^2 = (x + y)(x - y)$.)

Exercise 7.5.6. A commutative ring is called a *field* if $0 \neq 1$, and if every non-zero element admits a multiplicative inverse.

Show that $R[x]$ is never a field (regardless of the choice of commutative ring R).

Exercise 7.5.7. In this exercise, we'll see that different equations can be solved in different rings.

(a) Is there an element w in the ring $\mathbb{Z}$ for which $3w = 4$?

(b) Is there an element w in the ring $\mathbb{Q}$ for which $3w = 4$?

(c) Is there an element p in the ring $\mathbb{Z}[x]$ for which $4p^2 = x^2 + 2x + 1$?

(d) Is there an element p in the ring $\mathbb{Q}[x]$ for which $4p^2 = x^2 + 2x + 1$?

Exercise 7.5.8. Let R be an integral domain. Suppose there exists a polynomial $p \in R[x]$ for which $p^2 = x^2 + 1$. What must be true of the element $1 + 1$ in R?

Exercise 7.5.9. Suppose that a_0, a_1, a_2 are rational numbers, and suppose that there is an element $p \in \mathbb{R}[x]$ for which $p^2 = a_0 + a_1x + a_2x^2$. Prove that, in fact, $p \in \mathbb{Q}[x]$.

Exercise 7.5.10. Composition is not an operation that has anything to do with the ring structure of $R[x]$, but I'll introduce it anyway because it's such an important operation.

Let p and q be polynomials over R. Then we define the *composition* $p \circ q$ to be the polynomial

$$p(q(x)) = a_0 + a_1 q(x) + a_2(q(x)^2) + \ldots + a_n(q(x)^n).$$

(a) Show that if p and q have degree less than or equal to 1, then so does their composition $p \circ q$.

(b) Given a polynomial p, let $\tau(p)$ denote its degree ≤ 1 truncation. Explicitly, if $p = a_0 + a_1 x + a_2 x^2 + \ldots a_n x^n$, then

$$\tau(p)(x) = a_0 + a_1 x.$$

If p and q are arbitrary polynomials, can you find any relationships between the following four polynomials?

(i) pq

(ii) $p \circ q$

(iii) $\tau(p)\tau(q)$

(iv) $\tau(p) \circ \tau(q)$

7.6 Extra credit: Universal property of polynomial rings

Let S be an arbitrary ring.

(a) Prove that there exists exactly one ring homomorphism from $\mathbb{Z}$ to S.

Now consider the polynomial ring $\mathbb{Z}[x]$.

(b) Let $\hom_{Ring}(\mathbb{Z}[x], S)$ denote the set of all ring homomorphisms f from $\mathbb{Z}[x]$ to S. Consider the function

$$\mathrm{ev}_x : \hom_{Ring}(\mathbb{Z}[x], S) \to S, \qquad f \mapsto f(x).$$

Show that ev_x is a bijection.

(What the above says is that to specify a ring homomorphism from $\mathbb{Z}[x]$ to S is the same thing as just specifying an element of S. This is called the *universal property* of $\mathbb{Z}[x]$.)

(c) Now assume that S is commutative. Consider the function

$$\mathrm{ev} : \hom_{Ring}(\mathbb{Z}[x_1, \ldots, x_n], S) \to S^n, \qquad f \mapsto (f(x_1), \ldots, f(x_n)).$$

Show that ev is a bijection.