University of Wyoming
Math 3200—Spring 2012
$f(X,Y) \in \mathbb{Z}[X,Y]$
Polynomials
Department of Mathematics
$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Numbers and Polynomials (Handout January 20, 2012)

We are ready to define polynomials (these being rather like ‘numbers on steroids’) and to establish some of the relevant basic terminology. First we should talk about numbers. The plan here is for an informal presentation—this means that much of our terminology will be explained using examples rather than actual definitions. The terminology is very basic, and we present it at a fast pace, postponing details until later in the course. Please *read this through twice*—once quickly, to get an overview; and a second time, to make sure you have learned the main terminology. If anything is still unclear (such as notation from set theory, etc.) then please ask about it.

There is no universal concept of what a number is. Rather, there are many different number systems, each appropriate in its own context. So the term ‘number’ is context-dependent, as it refers to an element of whatever number system we are considering at a given moment. Among these the most familiar number systems are

> the set of **integers** $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots, \}$;
> the set of **natural numbers** $\mathbb{N} = \{1, 2, 3, \ldots\}$;
> the set of **rational numbers** $\mathbb{Q} = \left\{\frac{a}{b} \;:\; a, b \in \mathbb{Z},\; b \neq 0\right\}$;
> the set of **real numbers** $\mathbb{R}$; and
> the set of **complex numbers** $\mathbb{C} = \{a + bi \;:\; a, b \in \mathbb{R}\}$ where $i = \sqrt{-1}$; and
> for each positive integer n, we have the set $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$ of ‘integers mod n’.

Every integer is either positive, negative, or zero; and zero itself is neither positive nor negative. In this course, we will understand the natural numbers to be simply the positive integers, as indicated above. (*Warning:* Some sources instead take the natural numbers to be $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$. This is merely a matter of convention, not anything to get hung up on.) Although the most familiar number systems fit in a sequence under containment as $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$, there are many other important number systems. Some fit in between these; some contain $\mathbb{C}$; and others bear little or no relation to any of the number systems we have listed. We cannot strictly regard $\mathbb{Z}_n$ as a subset of any of these number systems, regardless of the similarity in symbols used, since its operations are very different; for example $2 + 2 = 1$ in $\mathbb{Z}_3$, although this is not true in any of the number systems $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.

The real number system $\mathbb{R}$ needs very little introduction. Two familiar ways to view real numbers are

(i) as numbers expressible in decimal form. These include rational numbers (such as $\frac{11}{4} = 2.75$, $\frac{5}{3} = 1.6666666666...$ and $-\frac{8}{19} = -0.4210526315...$) but also irrational numbers (such as $\sqrt{2} = 1.4142135623...$ and $\pi = 3.1415926535...$). An **irrational** number is a real number that is not rational. Also

(ii) real numbers are thought of as corresponding to points on the 'real number line'.

Both of these interpretations lend some useful intuition to the study of real numbers; but both interpretations are of limited value. Regarding (i), virtually none of the main properties of real numbers are easy consequences of the decimal expression. (Even to prove the distributive law $a(b+c) = ab+ac$ using decimal expansions would be an incredibly difficult task, to say nothing of the deeper properties such as the basic theorems of calculus.) And regarding (ii), it must be said that points on a physical line are not strictly in one-to-one correspondence with real numbers. For example, given any two *distinct* real numbers x and y, one of them is less than the other (i.e. either $x < y$ or $y < x$). However, points on an 'actual' line in physical space are not totally ordered in this way: Given two distinct points P and Q on a horizontal physical line, we imagine that one of them is always to the left of the other; but this is only approximately true—when two points are close enough together. (Within about $10^{-33}cm$ of each other, points are no longer even approximately ordered from left to right.) Physical space is more bizarre than is generally realized by most people. Moreover, the real numbers (an abstract mathematical concept) are also much more bizarre than most people realize. To top it off, the bizarre features of physical quantities (position, time, etc.) are *different* from the bizarre features of the real number system. So then what exactly *do* we mean by a real number? That's a story for another day—for now, just try to make do with (i) and (ii).

Remarkably, everyone has indicated (in class) some previous experience with the complex numbers. Nevertheless, we shall review some of their basic properties and applications.

We will soon learn to distinguish number systems based on their properties. Most of the number systems listed above ($\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Z}_n$) are examples of *rings*. I'd rather wait until later before formally defining a ring. However, let me point out that a **ring** has two basic operations of addition and multiplication; it has an identity element for addition (called 0) and an identity element for multiplication (called 1); there is also an operation of subtraction (the inverse of addition); and various familiar commutative, associative and distributive properties are satisfied. The number system $\mathbb{N}$ is not a ring because subtraction is not defined in $\mathbb{N}$ (for example, for the two elements $3, 11 \in \mathbb{N}$, there is no element of $\mathbb{N}$ of the form $3 - 11$).

Some rings have non-commutative multiplication; an example is the set of all 2×2 matrices with real entries, where $AB \neq BA$ for most choices of matrices A and B. (However, in order to qualify as a ring, addition is required to be commutative, i.e. $x + y = y + x$ for all ring elements x, y.) We do not foresee any use for non commutative multiplication in this course; so all our rings will be commutative.

If there is also an operation of division (the inverse of multiplication), then the number system is called a **field**. More precisely, in a field, we divide any element by any *nonzero* element, and get an element of the field. Thus $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields, but $\mathbb{N}$ and $\mathbb{Z}$ are not. (Try to divide 3 by 5 in $\mathbb{Z}$; this does not give an element of $\mathbb{Z}$, so $\mathbb{Z}$ is not a field.) Note that every field is a ring; but not every ring is a field. Further examples of fields include $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_5$, $\mathbb{Z}_7$, ... But the rings $\mathbb{Z}_4$, $\mathbb{Z}_6$, $\mathbb{Z}_8$, $\mathbb{Z}_9$, ... are *not* fields; for example, we cannot divide 3 by 2 in $\mathbb{Z}_4$. As we will explain later, $\mathbb{Z}_p$ is a field whenever p is prime. We prefer the notation $\mathbb{F}_p = \{0, 1, 2, \ldots, p-1\}$ in this case (where '$\mathbb{F}$' stands for 'field').

Fields are very important when studying linear algebra. This is because in trying to solve linear equations, one of the fundamental steps involved (in Gaussian elimination) is the process of division. For this reason, a course in linear algebra starts by designating some field F (such as $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{F}_2 = \{0, 1\}$) as the field of scalars; and then all vectors and matrices have entries chosen from F.

Now for polynomials. As examples, consider $f(x) = 1 + 4x - 7x^3$ and $g(x) = 1.8x + x^2 - 7\pi x^9 + 7.62x^{11}$. These are polynomials in x with real **coefficients**; accordingly, we write $f(x), g(x) \in \mathbb{R}[x]$. In fact since $f(x)$ has integer coefficients, it is correct to say that $f(x) \in \mathbb{Z}[x]$. We say that

the **coefficient** of x^3 in $1 + 4x - 7x^3$ is -7;
the coefficient of x^2 (or of x^4) in $1 + 4x - 7x^3$ is 0;
the coefficient of x^2 in $(1+x)^3$ is 3. (To see this, first expand $(1+x)^3 = 1+3x+3x^2+x^3$;
the coefficient of $x^0 = 1$ (i.e. the constant term) in $5 - 8x - 9x^2$ is 5.

More generally, if R is any commutative ring (such as $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{Z}_n$), then a **polynomial** in x with coefficients in R is defined to be an expression of the form

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

where $a_0, a_1, \ldots, a_n \in R$. We denote by $R[x]$ the set of all such polynomials. It's a ring (the **polynomial ring** in x with coefficients in R). We say that the polynomial $f(x)$, as shown, has **degree** n (assuming $a_n \neq 0$; otherwise, its degree would actually be less than n.) It's better to say that the degree of $f(x)$ is the largest integer d such that x^d has nonzero coefficient in $f(x)$. We write $\deg(f(x)) = d$ in this case. It is not too hard to see that for two polynomials $f(x)$ and $g(x)$, the product has degree

$$(*) \qquad \deg\big(f(x)g(x)\big) = \deg(f(x)) + \deg(g(x)).$$

Note that $2 + 7x$ has degree 1; and the constant polynomial 5 has degree 0. What is the degree of the constant polynomial 0 (i.e. the zero polynomial)? Either it is undefined (as some sources say), or we define it as $\deg(0) = -\infty$. This convention is chosen so that the rule (*) works (try it!).

Why did we define a polynomial as an *expression*, rather than as a *function*?

Let me ask you: what is a *spoon*? Most people would define a spoon with reference to its usual function (of eating soups, etc.). I could however give examples of spoons that have other purposes: spoons as percussion instruments, measuring spoons, and decorative souvenir spoons (like your grandmother used to collect) that arguably have no practical value other than to look interesting and remind us of some of the places we have traveled!

Likewise a polynomial may serve in many ways: sometimes as a function, sometimes not, and in many different roles; and sometimes the focus is not on any particular application—the polynomial just sits there and looks pretty. (OK, beauty is in the eye of the beholder.)

Let's point out that a polynomial can have any finite number of terms, but only a finite number of terms. An expression like

$$\sin x = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \frac{x^9}{362880} - \frac{x^{11}}{39916800} + \cdots$$

is *not* a polynomial; it is a **power series**. The set of all power series in x, with coefficients in R, is denoted $R[\![x]\!]$; thus, for example, $\sin x \in \mathbb{Q}[\![x]\!]$. We have $R[x] \subset R[\![x]\!]$. Like polynomials, power series have many applications (not always as functions). For example

$$\begin{aligned} g(x) &= 1 + x + 2x^2 + 6x^3 + 24x^4 + 120x^5 + 720x^6 + 5040x^7 + \cdots \in \mathbb{R}[\![x]\!] \text{ and} \\ h(x) &= 1 + x + (2x)^2 + (3x)^3 + (4x)^4 + (5x^5) + (6x)^6 + (7x)^7 + \cdots \in \mathbb{R}[\![x]\!] \end{aligned}$$

are two different power series, even though they both give the same very trivial function. As functions $\mathbb{R} \to \mathbb{R}$, we have $g(0) = 1$ and $h(0) = 1$; and $g(a)$ and $h(a)$ are both undefined for any *nonzero* real number a since the series only converge at 0. The reason that $g(x)$ and $h(x)$ are different power series is that they have different coefficients. To emphasize when we are talking about a power series as opposed to a function, many mathematicians say that $g(x)$ and $h(x)$ are **formal power series**. But this terminology suggests that formal power series are a special kind of power series, which is rather backwards from the truth: in fact only a very special kind of power series can be interpreted as any kind of function; since to represent a function, we would require convergence of the series. Our formal viewpoint means that we are usually not concerned with convergence or with functions; and this makes life easier.

Adding, or multiplying, two polynomials gives a polynomial. The same can be said for power series. In fact, both $R[x]$ and $R[\![x]\!]$ are rings.

The variable x is known as an **indeterminate**. It is merely a symbol, not a number: its purpose in life (like a souvenir spoon) is just to sit there and look pretty. Actually, it does a little more: without it, the polynomial is just a list of coefficients, and $f(x) = 1+4x-7x^3$ would be written simply as $f = (1, 4, 0, -7, 0, 0, 0, \ldots)$ which would be fine, except that the rule for multiplying polynomials:

$$(1, 4, 0, -7, 0, 0, 0, \ldots)(2, -3, 0, 0, 0, \ldots) = (2, 5, -12, -14, 21, 0, 0, 0, \ldots)$$

is more easily remembered and implemented when written in the form

$$(1 + 4x - 7x^3)(2 - 3x) = 2 + 5x - 12x^2 - 14x^3 + 21x^4.$$

The choice of letter x for the indeterminate is not so important; we could just as well write $f(t) = 1+4t-7t^3 \in \mathbb{Z}[t]$ or $f(r) = 1+4r-7r^3 \in \mathbb{Z}[r]$. But for now, we are not considering polynomials in more than one indeterminate; expressions like

$$h(x, y) = 4 + 2x - 3y + 7x^2 - 8xy - 10y^2 \in \mathbb{Z}[x, y]$$

will come up somewhat later in our course.

I have already used the term **term**. Please note that an expression like $U+V+W$ has three *terms*. An expression like uvw has only one term; it has three *factors*. When several quantities are added, they are called terms. When several quantities are multiplied, they are called factors. An expression like $5x^3$ is a **monomial**. An expression like $4x - 7x^5$ is a **binomial**. An expression like $1 + 4x - 7x^3$ is a **trinomial**. Note that the prefix (mono, bi, tri) in each case indicates the number of terms. The prefix 'poly' means 'many' or 'several'; hence a polynomial has any finite number of terms.