Math 5111 (Algebra 1)

Lecture #1 $\sim$ September 10, 2020

Integers and Polynomials:

- Welcome to Math 5111 + Course Logistics
- The Integers
- Polynomials Over A Field

This material represents §1.1-1.2.4 from the course notes.

Welcome!

Welcome to Math 5111 (Algebra 1)! Here are some course-related locations to bookmark:

- The course webpage is here: `https://web.northeastern.edu/dummit/teaching_fa20_5111.html` . Most course-related information is posted there.
- Course-related discussion will be done via Piazza: `https://piazza.com/class/kbwve0ywnsm2yb` .

Course Topics: Algebra

As you might expect based on the title, we will be covering algebra in this course. The catalog description is quite out of date. Here are the course topics:

- Polynomials and rings.
- Fields and field extensions.
- Groups.
- Galois theory and its applications.

The main theme of the course is to study fields and field extensions. However, since fields are so tied to the other fundamental topics in algebra (namely, groups and rings), we will develop most of algebra as we go.

Lectures + Office Hours

The course lectures will be conducted via Zoom. All lectures are recorded for later viewing. For security reasons (since these lecture slides are posted publicly) the links to upcoming and past lectures are only available via the Canvas page or via the Piazza page.

- The course meets Mon/Thu from 6:00pm-7:30pm Eastern time. All lectures are recorded.
- I have office hours Mon/Wed from 3:00pm-4:15pm Eastern time, or by appointment. Office hours are not required.

Lecture attendance is not required. However, I would prefer if you attended each lecture live, and (if possible) turn your camera on and participate, because otherwise the lectures are not nearly as valuable.

Grades

Your course grade consists of 1/3 homework and 2/3 exams.

- There will be a take-home midterm and a take-home final. These are not timed.
- The homeworks are assigned weekly.
- Assignments are due via Canvas. This is to make it easier to record grading comments.

This is a graduate course, and as such a number of the secondary topics are developed via the homework assignments. You will miss out on a lot of the content if you do not devote appropriate time to working on the homework every week. (It is also not likely you will do very well in the course!)

Miscellaneous Info

Here is some other miscellaneous information:

- I will write lecture notes for the course (in lieu of an official textbook) as the semester progresses. The course will roughly follow the presentation in Dummit and Foote's "Abstract Algebra" (3rd edition), a truly excellent book if ever there was one, but it is not necessary to purchase the textbook.
- Course prerequisites: A basic comfort level with groups, polynomials, and linear algebra is expected. If you have not taken linear algebra and at least one semester of undergraduate algebra, you should consult the instructor.
- Collaboration: You are allowed to work on, and discuss, homework assignments together, as long as the actual submissions are your own work. Collaboration is, of course, not allowed on exams.

Other Boilerplate, I

- Statement on Academic Integrity: A commitment to the principles of academic integrity is essential to the mission of Northeastern University. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University. Violations of academic integrity include (but are not limited to) cheating on assignments or exams, fabrication or misrepresentation of data or other work, plagiarism, unauthorized collaboration, and facilitation of others' dishonesty. Possible sanctions include (but are not limited to) warnings, grade penalties, course failure, suspension, and expulsion.

Other Boilerplate, II

- Statement on Accommodations: Any student with a disability is encouraged to meet with or otherwise contact the instructor during the first week of classes to discuss accommodations. The student must bring a current Memorandum of Accommodations from the Office of Student Disability Services.
- Statement on Classroom Behavior: Disruptive classroom behavior will not be tolerated. In general, any behavior that impedes the ability of your fellow students to learn will be viewed as disruptive.
- Statement on Inclusivity: Faculty are encouraged to address students by their preferred name and gender pronoun. If you would like to be addressed using a specific name or pronoun, please let your instructor know.

Other Boilerplate, III

- Statement on Evaluations: Students are requested to complete the TRACE evaluations at the end of the course.
- Miscellaneous Disclaimer: The instructor reserves the right to change course policies, including the evaluation scheme of the course (e.g., in the event of natural disaster or global pandemic). Notice will be given in the event of any substantial changes.

Algebra Is Fun!

Pause here for questions about course logistics.

Note to self: don't read this slide out loud.

Overview of §1: Polynomials and Rings

Our goal in this course is to study fields and Galois theory. To do this, we will need some preliminary facts about polynomials and polynomial rings.

- As motivation, we will first review (very briskly) some facts about the integers $\mathbb{Z}$, and then develop some basic facts about polynomials (this lecture) and then rings (next week's lectures).
- It is not expected that you will know all of the material from these first few lectures, but many of the topics should be at least passingly familiar.

The Integers, I

We all probably know what the integers $\mathbb{Z}$ are, in the functional sense that we understand what $2 + 3$ means.

- Nonetheless, it is not quite so easy to axiomatize $\mathbb{Z}$.
- If this were an introductory-level course, I would probably spend the rest of the lecture discussing properties of arithmetic, give the axioms for $\mathbb{Z}$ and explain how to construct a set satisfying the axioms using set theory.
- However, since nobody after 1920 really needs to bother themselves with this, I will just put the axioms on the next slide, mention the one actually important one (the well-ordering axiom), and move on.

The Integers, II

Definition

The integers are a set $\mathbb{Z}$ with two binary operations $+$ and $\cdot$ where

[I1] *$+$ is associative: $a + (b + c) = (a + b) + c$ for any $a, b, c \in \mathbb{Z}$.*

[I2] *$+$ is commutative: $a + b = b + a$ for any $a, b \in \mathbb{Z}$.*

[I3] *There is an additive identity 0 with $a + 0 = a$ for all $a \in \mathbb{Z}$.*

[I4] *Every integer a has an inverse $-a$ with $a + (-a) = 0$.*

[I5] *$\cdot$ is associative: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for any $a, b, c \in \mathbb{Z}$.*

[I6] *The operation $\cdot$ is commutative: $a \cdot b = b \cdot a$ for any $a, b \in \mathbb{Z}$.*

[I7] *There is a $1 \neq 0$ satisfying $1 \cdot a = a$ for all $a \in \mathbb{Z}$.*

[I8] *$\cdot$ distributes over $+$: $a \cdot (b + c) = a \cdot b + a \cdot c$ for any $a, b, c \in \mathbb{Z}$.*

Furthermore, there is a subset of $\mathbb{Z}$, called $\mathbb{N}$, such that

[N1] *For all $a \in \mathbb{Z}$, precisely one of $a \in \mathbb{N}$, $a = 0$, and $-a \in \mathbb{N}$ holds.*

[N2] *$\mathbb{N}$ is closed under $+$ and $\cdot$: if $a, b \in \mathbb{N}$ then $a + b, a \cdot b \in \mathbb{N}$.*

[N3] *Every nonempty subset S of $\mathbb{N}$ contains a smallest element: that is, an element $x \in S$ such that if $y \in S$, either $y = x$ or $y - x \in \mathbb{N}$.*

The Integers, II

The well-ordering axiom is what makes the integers special.

- It is equivalent to the inductive principle, which says that if S is a nonempty subset of the positive integers such that $1 \in S$ and $n \in S$ implies $(n+1) \in S$, then $S = \mathbb{N}$.
- This is the foundational idea for how proof by induction works.

Using the axiomatic description of $\mathbb{Z}$, one can establish all of the standard properties of arithmetic. We will discuss the interesting ones and write everything in normal language.

The Integers, III

The interesting part of the story with $\mathbb{Z}$ starts with division:

Definition

If $a \neq 0$, we say that a divides b (equivalently, b is divisible by a), written $a|b$, if there is an integer k with $b = ka$.

There are a bunch of properties of divisibility that are immediate from the definition:

- If $a|b$, then $a|bc$ for any c.
- If $a|b$ and $b|c$, then $a|c$.
- If $a|b$ and $a|c$, then $a|(xb + yc)$ for any x and y.
- If $a|b$ and $b|a$, then $a = b$ or $a = -b$.
- If $a|b$, and $a, b > 0$, then $a \leq b$.
- For any $m \neq 0$, $a|b$ is equivalent to $(ma)|(mb)$.

The Integers, IV

We can also do division with remainder (i.e., "long division"):

Proposition (Division Algorithm)

If a and b are positive integers, then there exist unique integers q and r such that $a = qb + r$ with $0 \leq r < b$. Furthermore, $r = 0$ if and only if $b|a$.

- Example: For $a = 18591$ and $b = 2291$, we have $18591 = 8 \cdot 2291 + 263$, so that $q = 8$ and $r = 263$.
- The proof of the existence of q and r relies on the well-ordering principle, and can be shown using induction.
- Uniqueness follows by rearranging $qb + r = a = q'b + r'$ to obtain $r - r' = b(q' - q)$: since $-b < r - r' < b$, this means $q' - q$ is an integer between -1 and 1, and hence must be 0.

The Integers, V

Of substantial utility are common divisors:

Definition

If $d|a$ and $d|b$, then d is a common divisor of a and b.
If a and b are not both zero, then there are only a finite number of common divisors: the largest one is called the greatest common divisor, or gcd, and denoted by $\gcd(a, b)$.
If the gcd is 1, we say a and b are relatively prime.

Some basic facts about greatest common divisors:

- If $m > 0$, then $m \cdot \gcd(a, b) = \gcd(ma, mb)$.
- If $d|a$ and $d|b$ with $d > 0$, then $\gcd(a/d, b/d) = \gcd(a, b)/d$.
- If a and b are both relatively prime to m, then so is ab.
- For any integer x, $\gcd(a, b) = \gcd(a, b + ax)$.
- If $c|ab$ and b, c are relatively prime, then $c|a$.

The Integers, V

The easiest way to compute gcds is using the Euclidean algorithm:

Theorem (Euclidean Algorithm)

Given integers $0 < b < a$*, repeatedly apply the division algorithm as follows, until a remainder of zero is obtained:*

$$\begin{aligned} a &= q_1 b + r_1 \\ b &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\vdots \\ r_{k-1} &= q_k r_k + r_{k+1} \\ r_k &= q_{k+1} r_{k+1}. \end{aligned}$$

Then $\gcd(a, b)$ *is equal to the last nonzero remainder,* r_{k+1}*.*

The algorithm terminates by the well-ordering axiom, and the gcd of any two consecutive remainders does not change, so the result follows by an easy induction.

The Integers, VI

As an immediate corollary of the Euclidean algorithm, we can also see that the GCD is a linear combination of the original integers:

Corollary (GCD as a Linear Combination)

If $d = \gcd(a, b)$, then there exist integers x and y with $d = xa + yb$.

Proof:

- By rearranging each equation in the Euclidean algorithm, we see that the newest remainder is a linear combination of the two previous terms.
- By an easy induction, every remainder can be written as an explicit linear combination of a and b (since the first two remainders clearly can be so written). In particular, $r_{k+1} = xa + yb$ for some integers x and y.

The Integers, VII

Example: Find the gcd of 1598 and 4879 using the Euclidean algorithm, and write it explicitly as a linear combination.

- First, we use the Euclidean algorithm:

$$\begin{aligned} 4879 &= 3 \cdot 1598 + 85 \\ 1598 &= 18 \cdot 85 + 68 \\ 85 &= 1 \cdot 68 + 17 \\ 68 &= 4 \cdot 17 \end{aligned}$$

  and so the gcd is 17.

- For the linear combination, we solve for the remainders:

$$\begin{aligned} 85 &= && = && 1 \cdot 4879 - 3 \cdot 1598 \\ 68 &= 1598 - 18 \cdot 85 && = && -18 \cdot 4879 + 55 \cdot 1598 \\ 17 &= 85 - 1 \cdot 68 && = && 19 \cdot 4879 - 58 \cdot 1598 \end{aligned}$$

  so we obtain $17 = 19 \cdot 4879 - 58 \cdot 1598$.

The Fundamental Theorem of Arithmetic

The other fundamental fact about the integers is that they possess unique prime factorization.

Definition

If $p > 1$ is an integer, we say it is prime if there is no d with $1 < d < p$ such that $d|p$: in other words, if p has no positive divisors other than 1 and itself. If $n > 1$ is not prime, meaning that there is some $d|n$ with $1 < d < n$, we say n is composite. (The integer 1 is neither prime nor composite.)

Theorem (Fundamental Theorem of Arithmetic)

Every integer $n > 1$ can be factored into a product of primes, and this factorization is unique up to reordering of the factors.

Both existence and uniqueness follow by induction arguments.

Modular Arithmetic, I

The other important construction using $\mathbb{Z}$ that we will discuss is the integers modulo m.

Definition

If m is a positive integer and m divides $b - a$, we say that a and b are congruent modulo m (or equivalent modulo m), and write "$a \equiv b$ (modulo m)".

Examples:

- $3 \equiv 9 \pmod 6$, since 6 divides $9 - 3 = 6$.
- $-2 \equiv 28 \pmod 5$, since 5 divides $28 - (-2) = 30$.
- $0 \equiv -666 \pmod 3$, since 3 divides $-666 - 0 = -666$.

Modular Arithmetic, II

Various properties of congruence follow (more or less immediately) from properties of divisibility:

- $a \equiv a \pmod{m}$.
- $a \equiv b \pmod{m}$ if and only if $b \equiv a \pmod{m}$.
- If $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$, then $a \equiv c \pmod{m}$.
- If $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a + c \equiv b + d$ and $ac \equiv bd \pmod{m}$.

The first three properties show that congruence mod m is an equivalence relation. The fourth shows that congruence respects addition and multiplication.

Modular Arithmetic, III

Our main interest is to discuss arithmetic modulo m, which we do via residue classes:

Definition

If a is an integer, the residue class of a modulo m, denoted $\overline{a}$, is the collection of all integers congruent to a modulo m. Observe that $\overline{a} = \{a + km,\ k \in \mathbb{Z}\}$.

Examples:

- The residue class of 2 modulo 4 is the set $\{\ldots, -6, -2, 2, 6, 10, 14, \ldots\}$.
- The residue class of 2 modulo 5 is the set $\{\ldots, -8, -3, 2, 7, 12, 17, \ldots\}$.

Note that the residue class of a modulo m is the equivalence class of a under the equivalence relation of congruence.

Modular Arithmetic, IV

Here are a few fundamental properties of residue classes:

Proposition

Proposition[Properties of Residue Classes] Suppose m is a positive integer. Then

1. *If a and b are integers with respective residue classes $\overline{a}$, $\overline{b}$ modulo m, then $a \equiv b \pmod{m}$ if and only if $\overline{a} = \overline{b}$.*
2. *Two residue classes modulo m are either disjoint or identical.*
3. *There are exactly m distinct residue classes modulo m, given by $\overline{0}$, $\overline{1}$, $\ldots$, $\overline{m-1}$.*

(1) follows from the definition, (2) is a general property of equivalence classes, and (3) follows from the division algorithm.

Modular Arithmetic, V

The main idea is that the addition and multiplication operations in $\mathbb{Z}$ also give rise to well-defined addition and multiplication operations modulo m:

Definition

The collection of residue classes modulo m is denoted $\mathbb{Z}/m\mathbb{Z}$ (read as "$\mathbb{Z}$ modulo $m\mathbb{Z}$").

Note that $\mathbb{Z}/m\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$.

Proposition (Modular Arithmetic)

The operations $\overline{a} + \overline{b} = \overline{a+b}$ and $\overline{a} \cdot \overline{b} = \overline{ab}$ are well defined on $\mathbb{Z}/m\mathbb{Z}$.

The well-definedness follows from the properties of congruences: if $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a + c \equiv b + d \pmod{m}$ and $ac \equiv bd \pmod{m}$.

Modular Arithmetic, VI

Here are the addition and multiplication tables for $\mathbb{Z}/5\mathbb{Z}$:

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{1}$	$\overline{3}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{1}$	$\overline{4}$	$\overline{2}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Note that, for example, the statement $\overline{2} + \overline{4} = \overline{1}$ is now perfectly acceptable, and correctly stated with the equals sign; we do not need to write it as $\overline{2} + \overline{4} \equiv \overline{1} \pmod 5$. Likewise, $\overline{3} \cdot \overline{3} = \overline{4}$ is correct.

Modular Arithmetic, VII

The arithmetic in $\mathbb{Z}/m\mathbb{Z}$ inherits many nice properties from $\mathbb{Z}$:

- $+$ is associative: $\overline{a} + (\overline{b} + \overline{c}) = (\overline{a} + \overline{b}) + \overline{c}$ for any $\overline{a}$, $\overline{b}$, and $\overline{c}$.
- $+$ is commutative: $\overline{a} + \overline{b} = \overline{b} + \overline{a}$ for any $\overline{a}$ and $\overline{b}$.
- $\overline{0}$ is an additive identity: $\overline{a} + \overline{0} = \overline{a}$ for any $\overline{a}$.
- Any $\overline{a}$ has an additive inverse $-\overline{a}$ satisfying $\overline{a} + (-\overline{a}) = \overline{0}$.
- $\cdot$ is associative: $\overline{a} \cdot (\overline{b} \cdot \overline{c}) = (\overline{a} \cdot \overline{b}) \cdot \overline{c}$ for any $\overline{a}$, $\overline{b}$, and $\overline{c}$.
- $\cdot$ is commutative: $\overline{a} \cdot \overline{b} = \overline{b} \cdot \overline{a}$ for any $\overline{a}$ and $\overline{b}$.
- $\cdot$ distributes over $+$: $\overline{a} \cdot (\overline{b} + \overline{c}) = \overline{a} \cdot \overline{b} + \overline{a} \cdot \overline{c}$ for any $\overline{a}$, $\overline{b}$, $\overline{c}$.
- $\overline{1}$ is a multiplicative identity: $\overline{1} \cdot \overline{a} = \overline{a}$ for any $\overline{a}$.

Modular Arithmetic, VIII

One very important difference, however, is that $\mathbb{Z}/m\mathbb{Z}$ can have zero divisors, and thus multiplicative cancellation does not always work.

- For example, if a, b, c are integers with $ab = ac$ and $a \neq 0$, then we can "cancel" a from both sides to conclude that $b = c$.
- However, this does not always work in $\mathbb{Z}/m\mathbb{Z}$: for example, $2 \cdot 1 = 2 \cdot 4$ modulo 6, but $1 \neq 4$ modulo 6.
- The issue here is that 2 and the modulus 6 are not relatively prime: 6 divides $2(4-1)$, but 6 does not divide $4-1$.

Modular Arithmetic, IX

We can characterize exactly what is happening here using gcds:

Proposition (Invertible Elements in $\mathbb{Z}/m\mathbb{Z}$)

If $m > 0$, then the residue class $\overline{a}$ has a multiplicative inverse in $\mathbb{Z}/m\mathbb{Z}$ if and only if a and m are relatively prime.

Proof:

- First suppose that a and m are relatively prime. Then by our analysis of the Euclidean algorithm, there exist integers x and y such that $xa + ym = 1$: then $xa \equiv 1 \pmod{m}$, which is to say $\overline{x} \cdot \overline{a} = \overline{1}$, so that $\overline{a}$ has a multiplicative inverse as claimed.
- Conversely, suppose $\overline{a}$ were invertible in $\mathbb{Z}/m\mathbb{Z}$ with inverse $\overline{x}$. Then $\overline{x} \cdot \overline{a} = \overline{1}$, or equivalently $xa \equiv 1 \pmod{m}$, and this is in turn equivalent to saying there exists an integer y with $xa + ym = 1$. But then the common divisor d would divide $xa + ym$ hence divide 1, and so a and m are relatively prime.

Modular Arithmetic, X

The proof on the previous slide shows how we can compute the inverse of an invertible residue class using the Euclidean algorithm.

Example: Find the multiplicative inverse of $\overline{9}$ in $\mathbb{Z}/11\mathbb{Z}$.

- Using the Euclidean algorithm, we can obtain $1 = 5 \cdot 11 - 6 \cdot 9$.
- Considering both sides modulo 11 yields $\overline{1} = \overline{-6} \cdot \overline{9}$.
- Thus, $\overline{-6}$ is the multiplicative inverse of $\overline{9}$ in $\mathbb{Z}/11\mathbb{Z}$.
- We could also write the inverse as $\overline{5}$, since $\overline{-6} = \overline{5}$, if we wanted to.

Modular Arithmetic, XI

The case where the modulus is prime is of particular importance:

Corollary

If p is a prime number, then every nonzero residue class in $\mathbb{Z}/p\mathbb{Z}$ has a multiplicative inverse.

Proof:

- If p is prime, then p is relatively prime to each of 1, 2, ... , $p-1$ and hence all of the nonzero residue classes modulo p are invertible.

This corollary tells us that $\mathbb{Z}/p\mathbb{Z}$ is a field, because every nonzero element has a multiplicative inverse.

To emphasize the field structure, we will often write this field as $\mathbb{F}_p$ ("the field with p elements").

Speaking of Fields....

We will discuss fields in much more detail later. But here are the fields you should keep in the back of your mind whenever I mention the word "field" over the next few lectures:

- $\mathbb{Q}$, the rational numbers.
- $\mathbb{R}$, the real numbers.
- $\mathbb{C}$, the complex numbers.
- $\mathbb{F}_p$, the field with p elements (p being a prime), also known as $\mathbb{Z}/p\mathbb{Z}$, the integers modulo p.

Polynomials, I

With these basics in hand, we can now start our discussion of polynomials.

- Polynomials with real coefficients (like $p(x) = 1 + x^2$ or $q(x) = 3 + \pi x^2$) are likely familiar from elementary algebra.
- Unlike in elementary algebra, however, our polynomials will be "formal symbols" rather than functions.
- We will soon exploit the connection between polynomials and functions, but there are very important reasons for us to take a more abstract approach to polynomials than simply viewing them as functions.

Polynomials, II

Definition (Useful Definition of Polynomials)

Let F be a field and x be an indeterminate. A polynomial in x with coefficients in F consists of a formal sum $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, for an integer $n \geq 0$ and where each element $a_i \in F$.

If you want the entirely rigorous definition (i.e., the foundationally correct one that nobody uses), here it is:

Definition (Technical Definition of Polynomials)

Let F be a field and C be the Cartesian product $\prod_{\mathbb{Z}_{\geq 0}} F = (a_0, a_1, a_2, \dots)$ indexed by the nonnegative integers. Then the polynomials with coefficients in F are the sequences in C all but finitely many of whose entries are zero. We interpret the sequence $(a_0, a_1, a_2, \dots, a_n, 0, 0, \dots)$ as the formal sum $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$.

Polynomials, III

Some notation and terminology:

- If $a_n \neq 0$, we say that the polynomial has degree n and if $a_n = 1$ we say the polynomial is monic. (By convention, the degree of the zero polynomial 0 is $-\infty$.)
- The leading term of the polynomial is its highest-degree term (i.e., $a_n x^n$) and its leading coefficient is the corresponding coefficient (i.e., a_n).
- I will use function notation for polynomials (e.g., by writing a polynomial as $p(x) = x^2 + 5$), and also often drop the variable portion (e.g., "the polynomial p") when convenient.
- To reiterate, however, our polynomials are *not* functions, but rather formal sums.

Polynomials, IV

Polynomials have natural arithmetic operations:

- Addition is defined termwise:
  $(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0) + (b_n x^n + b_{n-1} x^{n-1} + \cdots + b_0)$
  $= (a_n + b_n) x^n + (a_{n-1} + b_{n-1}) x^{n-1} + \cdots + (a_0 + b_0)$.
- Multiplication is defined first on monomials via $(a x^n) \cdot (b x^m) = ab x^{n+m}$, and then extended to arbitrary polynomials via the distributive laws:
  $(a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n) \cdot (b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m) =$
  $a_0 b_0 + (a_1 b_0 + a_0 b_1) x + (a_2 b_0 + a_1 b_1 + a_0 b_2) x^2 + \cdots + a_n b_m x^{n+m}$
  where the coefficient of x^j in the product is $\sum_{k=0}^{j} a_k b_{j-k}$.
- It is tedious (but not difficult) to verify the basic properties of arithmetic for $F[x]$: the associative, commutative, and distributive laws for $+$ and $\cdot$, that 0 is an additive identity, that 1 is a multiplicative identity, and so forth.

Polynomials, V

Degrees behave quite well under addition and multiplication:

Proposition (Properties of Degree)

If p and q are any polynomials in $F[x]$, then $\deg(p+q) \leq \max(\deg p, \deg q)$, *and* $\deg(p \cdot q) = \deg p + \deg q$.

Proof:

- Each claim clearly holds if p or q is zero (in which case the left side of each inequality is $-\infty$). Now assume $p, q \neq 0$.
- For $p+q$, observe that if there are no terms of degree n or higher in p or q, then there are no terms of degree n or higher in $p+q$ either.
- For $p \cdot q$, observe that if the leading terms of p and q are $a_n x^n$ and $b_m x^m$ respectively, then the leading term of $p \cdot q$ is $a_n b_m x^{m+n}$, and $a_n b_m \neq 0$ since F is a field.

Polynomial Division, I

We can define divisibility of polynomials:

Definition

If $a, b \in F[x]$, we say that a divides b (written $a|b$), if there is a $k \in F[x]$ with $b = ka$.

Examples:

- We see that $x - 1$ divides $x^2 - 1$ in $\mathbb{Q}[x]$, since $x^2 - 1 = (x - 1)(x + 1)$.
- We see that $x - i\sqrt{2}$ divides $x^4 - 4$ in $\mathbb{C}[x]$, since $x^4 - 4 = (x - i\sqrt{2})(x^3 + i\sqrt{2}x^2 - 2x - 2i\sqrt{2})$.

Polynomial Division, II

$F[x]$ possesses a long division algorithm, where we measure the size of a polynomial via its degree.

Theorem (Division Algorithm in $F[x]$)

If F is a field, and $a(x)$ and $b(x)$ are any polynomials in $F[x]$ with $b(x) \neq 0$, then there exist unique polynomials $q(x)$ and $r(x)$ such that $a(x) = b(x)q(x) + r(x)$, where $\deg(r) < \deg(b)$.
Furthermore, $b|a$ if and only if $r = 0$.

- We require F to be a field to be able to divide by arbitrary nonzero coefficients. (Over $\mathbb{Z}$, for instance, we cannot divide x^2 by $2x$ and get a remainder that is a constant polynomial.)
- For example, when we divide the polynomial $x^3 + x^2 + 3x + 5$ by the polynomial $x^2 + 3x + 1$ in $\mathbb{R}[x]$, we obtain the quotient $q(x) = x - 2$ and remainder $r(x) = 8x + 7$: indeed, we have $x^3 + x^2 + 3x + 5 = (x - 2)(x^2 + 3x + 1) + (8x + 7)$.

Polynomial Division, III

Proof:

- We induct on the degree n of $a(x)$.
- The base case is trivial, as we may take $q = r = 0$ if $a = 0$.
- Now suppose the result holds for all polynomials $a(x)$ of degree $\leq n-1$. If $\deg(b) > \deg(a)$ then we can simply take $q = 0$ and $r = a$, so now also assume $\deg(b) \leq \deg(a)$.
- Write $a(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ and $b(x) = b_m x^m + \cdots + b_0$, where $b_m \neq 0$ since $b(x) \neq 0$.
- Observe that $a^\dagger(x) = a(x) - \frac{a_n}{b_m} x^{n-m} b(x)$ has degree less than n, since we have cancelled the leading term of $a(x)$. (Here we are using the fact that F is a field, so that $\frac{a_n}{b_m}$ also lies in F.)
- By the induction hypothesis, $a^\dagger(x) = q^\dagger(x)b(x) + r^\dagger(x)$ for some $q^\dagger(x)$ and $r^\dagger(x)$ with $r^\dagger = 0$ or $\deg(r^\dagger) < \deg(b)$.
- Then $a(x) = [q^\dagger(x) + \frac{a_n}{b_m} x^{n-m}]b(x) + r^\dagger(x)$, so $q(x) = q^\dagger(x) + \frac{a_n}{b_m} x^{n-m}$ and $r(x) = r^\dagger(x)$ work as claimed.

Polynomial Division, IV

Proof (continued):

- We have shown that there exist q and r with $a = qb + r$ and $\deg(r) < \deg(b)$.
- For the uniqueness, suppose that $a = qb + r = q'b + r'$: then $r - r' = b(q' - q)$ has degree less than $\deg(b)$ but is also divisible by b, hence must be zero.
- Finally, by definition if $r = 0$ then $b|a$, and conversely if $b|a$ then since r is unique we must have $r = 0$.

Polynomial Division, V

The existence of this division algorithm in $F[x]$ allows us to adapt many results that hold in $\mathbb{Z}$ into this setting.

We could, in fact, do all of this more generally in the context of Euclidean domains. However, we will only need the results for polynomials, so in the interest of brevity, we will just stick with polynomials.

Divisors and Euclid, I

First is the idea of a common divisor:

Definition

If a and b are polynomials in $F[x]$, we say a polynomial d is a common divisor if $d|a$ and $d|b$.

Example:

- The polynomials $x+1$ and $2x+2$ are both divisors of x^2-1 and x^2+3x+2 in $\mathbb{R}[x]$.

We would next want to define the greatest common divisor to be the polynomial of largest degree dividing both a and b.

- However, this polynomial is not unique: in the example above, it is easy to see that x^2-1 and x^2+3x+2 do not have a common divisor of degree 2 (or larger), so both $x+1$ and $2x+2$ are common divisors of maximal degree.

Divisors and Euclid, II

The situation on the previous slide is easy to rectify, since $x + 1$ and $2x + 2$ only differ by a constant factor.

Definition

If p and q are polynomials in $F[x]$ and there exists a nonzero constant c such that $p = cq$, we say p and q are associate.

Our next claim is that the gcd of any two polynomials is unique up to associates. Of course, this requires defining the gcd properly (and then proving its uniqueness).

Divisors and Euclid, III

We adopt the following definition:

Definition

If a and b are polynomials in $F[x]$, we say the polynomial d is a greatest common divisor of a and b if it a common divisor of a and b with the property that if d' is any other common divisor, then $d'|d$.

This definition does not immediately imply that a gcd actually exists. To establish this fact, we adapt the Euclidean algorithm to this setting, which will also give us a procedure for computing the gcd and for writing it as a linear combination.

Divisors and Euclid, IV

Algorithm (Euclidean Algorithm in $F[x]$)

Given a and b in $F[x]$, not both zero, repeatedly apply the division algorithm as follows, until a remainder of zero is obtained:

$$\begin{aligned} a &= q_1 b + r_1 \\ b &= q_2 r_1 + r_2 \\ r_1 &= q_3 r_2 + r_3 \\ &\vdots \\ r_{k-1} &= q_k r_k + r_{k+1} \\ r_k &= q_{k+1} r_{k+1}. \end{aligned}$$

Then the last nonzero remainder r_{k+1} is a gcd of a and b.

Just as in $\mathbb{Z}$, by successively solving for the remainders and plugging in the previous equations, the gcd r_{k+1} can be written as a linear combination of a and b.

Divisors and Euclid, VI

The uniqueness of the gcd up to associates follows from the Euclidean algorithm:

- Explicitly, if d_1 and d_2 are both greatest common divisors of a and b, then $d_1|d_2$ and $d_2|d_1$, so that $d_1 = sd_2$ and $d_2 = td_1$ for some polynomials s and t.
- By comparing degrees, we see that $\deg(s) = \deg(t) = 0$, meaning that s and t must both be constants, and thus d_1 and d_2 are associates.
- Since the gcd of any two polynomials exists by the Euclidean algorithm, the gcd is unique up to associates as claimed.
- If a and b are not both zero, we can make the gcd unique by additionally requiring that it be monic (i.e., have leading coefficient 1).

Divisors and Euclid, VII

Example: Find the monic gcd $d(x)$ of the polynomials $p = x^6 + 2$ and $q = x^8 + 2$ in $\mathbb{F}_3[x]$, and then write it as a linear combination of p and q.

- We apply the Euclidean algorithm: we have

$$\begin{aligned} x^8 + 2 &= x^2(x^6 + 2) + (x^2 + 2) \\ x^6 + 2 &= (x^4 + x^2 + 1)(x^2 + 2) \end{aligned}$$

  and so the last nonzero remainder is $x^2 + 2$.
- By back-solving, we see that $x^2 + 2 = 1 \cdot (x^8 + 2) - x^2(x^6 + 2)$.

Of course, most applications will require more than one step, in which case we would solve the equations for the remainders from the top down.

Irreducibility and Factorization, I

We next develop the polynomial analogue of the prime factorization of an integer: namely, writing a polynomial as a product of irreducible factors, and showing that this factorization is essentially unique.

Definition

A nonzero polynomial $p \in F[x]$ is irreducible if it is not a constant, and for any "factorization" $p = bc$ with $b, c \in F[x]$, one of b and c must be a constant polynomial. If p is not a constant and possesses a factorization $p = bc$ where neither b nor c is constant, then p is reducible.

A polynomial is irreducible if it cannot be written as a product of two polynomials of smaller positive degree, and is reducible if it can be so written.

Irreducibility and Factorization, II

Examples:

- Any polynomial of degree 1 is irreducible.
- The polynomial $x^2 + x + 1$ is irreducible in $\mathbb{F}_2[x]$, since the only possible factorizations would be $x \cdot x$, $x \cdot (x+1)$, or $(x+1) \cdot (x+1)$, and none of these is equal to $x^2 + x + 1$.
- The polynomial $x^4 + 4$ is reducible in $\mathbb{Q}[x]$, since we can write $x^4 + 4 = (x^2 + 2x + 2)(x^2 - 2x + 2)$.
- The polynomial $x^2 + 1$ is irreducible in $\mathbb{R}[x]$, since there is no way to write it as the product of two linear polynomials with real coefficients.

We warn that whether a given polynomial is irreducible in $F[x]$ depends on the field F. For example, $x^2 + 1$ is irreducible in $\mathbb{R}[x]$ but not in $\mathbb{C}[x]$, since we can write $x^2 + 1 = (x + i)(x - i)$ in $\mathbb{C}[x]$.

Irreducibility and Factorization, III

Proposition (Factorization into Irreducibles)

Every polynomial of positive degree in $F[x]$ can be written as a product of irreducible polynomials (where a "product" is allowed to have only one term).

Proof:

- We use strong induction on $n = \deg(p)$. The result clearly holds if $n = 1$, since any polynomial of degree 1 is irreducible.
- Now suppose $n \geq 2$. If p is irreducible, we are done, so otherwise assume that p is reducible.
- By definition, there exist polynomials a, b with $0 < \deg(a), \deg(b) < n$ with $p = ab$.
- By the strong induction hypothesis, both a and b can be written as a product of irreducibles; multiplying these two products then gives p as a product of irreducibles.

Irreducibility and Factorization, IV

We will also need the following divisibility property:

Proposition (Irreducibles are Prime in $F[x]$)

If $p \in F[x]$ is irreducible and $p|ab$, then $p|a$ or $p|b$.

Proof:

- Suppose $p|ab$. If $p|a$, we are done, so suppose $p \nmid a$, and let d be a gcd of p and a.
- By hypothesis, d divides p, so (since p is irreducible) either d is a constant, or $d = up$ for some constant u; the latter cannot happen, because then up (hence p) would divide a.
- Hence d is a constant, say with inverse e.
- By the Euclidean algorithm, there exist x, y with $xp + ya = d$.
- Multiplying by be and regrouping the terms yields $(bce)p + ey(ab) = (de)b = b$. Since p divides both terms on the left-hand side, we conclude $p|b$.

Irreducibility and Factorization, V

Now we can address the uniqueness of irreducible factorizations:

- There is one additional wrinkle to address, however, which involves constant factors.
- To illustrate, note that in $\mathbb{C}[x]$, we can write $x^2 + 1 = (x + i)(x - i) = (ix + 1)(-ix + 1)$.
- It would seem that these are two different factorizations, but we should really consider them the same, because all we have done is moved some units around: $x + i = i(-ix + 1)$ and $x - i = (-i)(ix + 1)$.
- We should declare that two factorizations are equivalent if the only differences between them are by reordering terms or moving constant factors around, which is equivalent to replacing elements with associates.

Irreducibility and Factorization, VI

Now we prove our fundamental result about unique factorization:

Theorem (Unique Factorization in $F[x]$)

Every polynomial of positive degree in $F[x]$ can be written as a product of irreducible polynomials. Furthermore, this factorization is unique up to reordering and associates: if $p = r_1 r_2 \cdots r_d = q_1 q_2 \cdots q_k$, then $d = k$ and there is some reordering of the factors such that p_i and q_i are associate for each $1 \leq i \leq k$.

Proof:

- We already showed existence, so we only need uniqueness.
- Induct on the number of irreducible factors of $p = r_1 r_2 \cdots r_d$.
- If $d = 0$, then p is a constant. If p had some other factorization $p = rc$ with r irreducible, then q would divide a constant, hence be a constant (impossible).

Irreducibility and Factorization, VI

Proof (continued):

- Now suppose $d \geq 1$ and that $r = r_1 r_2 \cdots r_k = q_1 q_2 \cdots q_d$ has two factorizations into irreducibles.
- Since $r_1 | (q_1 \cdots q_d)$ and r_1 is irreducible, repeatedly applying the fact that r_1 irreducible and $r_1 | ab$ implies $r_1 | a$ or $r_1 | b$ shows that r_1 must divide q_i for some i.
- Then $q_i = r_1 u$ for some u: then since q_i is irreducible (and r_1 is not a constant), u must be a constant, and thus q_i and r_1 are associates.
- Cancelling r_1 from both sides then yields the equation $r_2 \cdots r_d = (uq_2) \cdots q_k$, which is a product of fewer irreducibles. By the induction hypothesis, such a factorization is unique up to associates. This immediately yields the desired uniqueness result for p as well.

Roots and Factorization, I

It is reasonable to ask how one actually factors polynomials or proves that they are irreducible. (The glib answer is: "by making a computer do it".)

There is much to say about factorization algorithms, and we will not go very far in this direction: we will content ourselves with some basic facts about irreducibility and roots.

Roots and Factorization, II

In elementary algebra, polynomials are examples of functions. We would like to extend this idea of "plugging values in" to a general polynomial in $F[x]$.

Definition

If F is a field and $p = a_0 + a_1x + \cdots + a_nx^n$ is an element of $F[x]$, for any $r \in F$ we define the value $p(r)$ to be the element $a_0 + a_1r + \cdots + a_nr^n \in F$.

In this way, we can view a polynomial $p \in F[x]$ as a function $p : F \to F$, with $p(r) = a_0 + a_1r + \cdots + a_nr^n$.

We will remark that the polynomial notation $p(x)$ is somewhat ambiguous: we may be considering $p(x)$ as an element in $F[x]$ (in which case "x" represents an indeterminate), or we may be viewing it as a function from F to F (in which case "x" represents the variable of the function).

Roots and Factorization, III

Here's a pair of observations from elementary algebra:

Proposition (Remainder/Factor Theorem)

Let F be a field. If $p \in F[x]$ is a polynomial and $r \in F$, then the remainder upon dividing $p(x)$ by $x - r$ is $p(r)$. In particular, $x - r$ divides $p(x)$ if and only if $p(r) = 0$; i.e., if r is a zero (or root) of p.

Proof:

- Let $p(x) = a_0 + \cdots + a_n x^n$. Observe that $x - r$ divides $x^k - r^k$ since $(x^k - r^k) = (x - r)(x^{k-1} + x^{k-2}r + \cdots + xr^{k-2} + r^{k-1})$.
- Now write $p(x) - p(r) = \sum_{k=0}^{n} a_k(x^k - r^k)$: since $x - r$ divides each term in the sum, it divides $p(x) - p(r)$.
- Since $p(r)$ is a constant, it is therefore the remainder after dividing $p(x)$ by $x - r$. The other statement is immediate from the uniqueness of the remainder in the division algorithm.

Roots and Factorization, IV

We also bound the number of zeroes that a polynomial can have:

Proposition (Number of Roots)

Let F be a field. If $p \in F[x]$ is a polynomial of degree d, then p has at most d distinct roots in F.

Proof:

- We induct on the degree d. The base case $d = 1$ is easy.
- Now suppose the result holds for all polynomials of degree $\leq d$ and let p be a polynomial of degree $d + 1$.
- If p has no zeroes we are obviously done, so suppose otherwise and let $p(r) = 0$. We can then factor to write $p(x) = (x - r)q(x)$ for some polynomial $q(x)$ of degree d.
- By the induction hypothesis, $q(x)$ has at most d roots: then $p(x)$ has at most $d + 1$ roots, because $(a - r)q(a) = 0$ only when $a = r$ or $q(a) = 0$ (since F is a field).

Roots and Factorization, V

Here is a useful result for irreducibility in low degree:

Proposition (Irreducibility in Degrees 2 and 3)

If F is a field and $p \in F[x]$ has degree 2 or 3 and has no zeroes in F, then p is irreducible.

Proof:

- If $p(x) = a(x)b(x)$ then $\deg(p) = \deg(a) + \deg(b)$.
- Suppose a and b are not constant. Then since both have positive degree and $\deg(p)$ is 2 or 3, at least one of a and b must have degree 1.
- Its root is then also a root of $p(x)$. Taking the contrapositive gives the desired statement.

Roots and Factorization, VI

Examples:

- Over $\mathbb{R}$, the polynomial $x^2 + 2x + 11$ has no roots (it is always positive, as can be seen by completing the square), so it is irreducible.
- Over $\mathbb{F}_2$, the polynomial $q(x) = x^3 + x + 1$ is irreducible: it has no roots since $q(0) = q(1) = 1$.
- Over $\mathbb{F}_5$, the polynomial $q(x) = x^3 + x + 1$ is irreducible: it has no roots since $q(0) = 1$, $q(1) = 3$, $q(2) = 1$, $q(3) = 1$, and $q(4) = 4$.

Note of course that a polynomial of larger degree can be reducible without having any zeroes: for example, $x^4 + 3x^2 + 2$ has no zeroes in $\mathbb{R}$, but it is still reducible: $x^4 + 3x^2 + 2 = (x^2 + 1)(x^2 + 2)$.

The Fundamental Theorem of Algebra

For certain particular fields, we can say more about the structure of the irreducible polynomials.

Theorem (Fundamental Theorem of Algebra)

Every polynomial of positive degree in $\mathbb{C}[x]$ has at least one root. Therefore, the irreducible polynomials in $\mathbb{C}[x]$ are precisely the polynomials of degree 1, and so every polynomial in $\mathbb{C}[x]$ factors into a product of degree-1 polynomials.

Despite the fact that this is known as the Fundamental Theorem of Algebra, it is really more of an analytic statement. The usual proofs involve either complex analysis or topology.

A standard argument is as follows: $|p|$ is a continuous map from $\mathbb{R}^2$ to $\mathbb{R}$, by compactness $|p|$ must have a global minimum on $\mathbb{R}^2$, and then by the Taylor expansion the only possible global minimum of $|p|$ is 0. (Another approach is to use Rouché's theorem.)

Summary

We discussed the logistics for Math 5111.

We discussed $\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$.

We discussed polynomials, polynomial operations, irreducibility, unique factorization, and roots.

Next lecture: More with polynomials, rings