CCS Discrete Math I — Professor: Padraic Bartlett

Lecture 6: Finite Fields

Week 6 — *UCSB 2014*

It ain't what they call you, it's what you answer to.

W. C. Fields

1 Fields

In the next two weeks, we're going to study **fields**; a mathematical object that on one hand is a relatively simple generalization of the ideas behind groups, but on the other will allow us to understand a variety of beautiful mathematical concepts and applications. We start here with the basics:

1.1 Definitions and properties.

Definition. A **field** is any set $\mathbb{F}$ along with two binary operations $\cdot, + : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$, that satisfy the following properties:

- **Closure**$(+)$: $\forall a, b \in \mathbb{F}$, we have $a+b \in \mathbb{F}$.
- **Identity**$(+)$: $\exists 0 \in \mathbb{F}$ such that $\forall a \in \mathbb{F}$, $0 + a = a$.
- **Commutativity**$(+)$: $\forall a, b \in \mathbb{F}, a + b = b + a$.
- **Associativity**$(+)$: $\forall a, b, c \in \mathbb{F}, (a + b) + c = a + (b + c)$.
- **Inverse**$(+)$: $\forall a \in \mathbb{F}, \exists (-a) \in \mathbb{F}$ such that $a + (-a) = 0$.
- **Closure**$(\cdot)$: $\forall a, b \in \mathbb{F}$, we have $a \cdot b \in \mathbb{F}$.
- **Identity**$(\cdot)$: $\exists 1 \neq 0 \in \mathbb{F}$ such that $\forall a \in \mathbb{F}$, $1 \cdot a = a$.
- **Commutativity**$(\cdot)$: $\forall a, b \in \mathbb{F}, a \cdot b = b \cdot a$.
- **Associativity**$(\cdot)$: $\forall a, b, c \in \mathbb{F}, (a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- **Inverse**$(\cdot)$: $\forall a \neq 0 \in \mathbb{F}, \exists a^{-1} \in \mathbb{F}$ such that $a \cdot a^{-1} = 1$.
- **Distributivity**: $(+, \cdot) : \forall a, b, c \in \mathbb{F}, (a + b) \cdot c = (a \cdot c) + (b \cdot c)$

In a sense, a field is pretty much a set $\mathbb{F}$ that is a commutative group in two ways at the same time: that is, it is a group with respect to addition, and it is also a group with respect to multiplication if you ignore the additive identity 0!

One question that you might naturally ask here is **why** 0 is considered special: that is, when we're writing up our axioms, why did we make the "inverses$(\cdot)$" property only need to hold for nonzero elements? There are two natural answers here. The first, which comes from more of a utilitarian approach, is to ask **where** our notions for any of these properties

for a field come from! In practice, much like how the group axioms came naturally from looking at objects like $\langle \mathbb{Z}/n\mathbb{Z}, + \rangle$, $\langle S_n, \circ \rangle$, and $\langle \mathbb{Z}, + \rangle$ and noticing certain nice properties those objects had, we could have derived the **field axioms** from looking at some of the most commonly-occurring number systems we work with: $\langle \mathbb{R}, +, \cdot \rangle$, $\langle \mathbb{Q}, +, \cdot \rangle$, and $\langle \mathbb{C}, +, \cdot \rangle$ are all fields! Furthermore, in all three of those objects, we have that 0 is some sort of an "annihilating element" with respect to multiplication: that is, $\forall a$, $0 \cdot a = 0$. So it would be unreasonable to expect 0 to have a multiplicative inverse!

Another perspective to take here, however, would be to question whether it is even possible for 0 to have a multiplicative inverse. That is: suppose that we had any object that satisfies the axioms above. Is it possible for such an object to **also** have a multiplicative inverse for 0? Or is this something that we can prove is impossible?

The answer here turns out to be yes:

Claim. Suppose that $\langle \mathbb{F}, +, \cdot \rangle$ is a field. Then, for all $a \in \mathbb{F}, 0 \cdot a = 0$. Consequently, because $0 \neq 1$ (as stated in our definition of the multiplicative inverse,) 0 does not have a multiplicative inverse.

Proof. Take any $a \in \mathbb{F}$. Because of the closure$(\cdot)$ property, we know that $0 \cdot a$ is also a field element. Trivially, we know that

$$0 \cdot a = 0 \cdot a.$$

We also know that 0 is an additive identity: therefore, in specific, we know that $0 = 0 + 0$, and therefore that

$$0 \cdot a = (0 + 0) \cdot a.$$

Applying the distributive property then tells us that

$$0 \cdot a = (0 + 0) \cdot a = (0 \cdot a) + (0 \cdot a).$$

Now, we can use the inverse$(+)$ property to tell us that because $0 \cdot a$ is a field element, we also know that there is some other field element$-(0 \cdot a)$ such that $(0 \cdot a) + (-(0 \cdot a)) = 0$. Then, if we add this to both sides of our equality above, we get

$$(0 \cdot a) + (-(0 \cdot a)) = ((0 \cdot a) + (0 \cdot a)) + (-(0 \cdot a)).$$

Applying the inverse property to the left hand side tells us that it's 0; applying the associative property to the right side tells us that

$$0 = ((0 \cdot a) + (0 \cdot a)) + (-(0 \cdot a)) = (0 \cdot a) + ((0 \cdot a) + (-(0 \cdot a))) = (0 \cdot a) + 0 = (0 \cdot a),$$

by applying first the inverse property and then the additive identity property to make the $+0$ go away. Therefore, we've proven that for any $a \in \mathbb{F}$, we have

$$0 = 0 \cdot a.$$

□

Using similar techniques, we can prove other results, like the following:

Claim. For any $a \in \mathbb{F}$, we have that $(-a) = (-1) \cdot a$. In other words, we can create the additive inverse of any element by multiplying it by the additive inverse of 1.

Proof. By the multiplicative identity property, we know that $1 \in \mathbb{F}$; by the additive inverse property, we then also know that $-1 \in \mathbb{F}$ and that

$$0 = 1 + (-1).$$

Using closure, distributivity, and the multiplicative identity property, we can take any a and multiply it by the left and right hand sides above:

$$0 \cdot a = (1 + (-1)) \cdot a = 1 \cdot a + (-1) \cdot a = a + (-1) \cdot a).$$

Using our result above, we know that $0 \cdot a = 0$, and therefore that

$$0 = a + (-1) \cdot a).$$

Using the additive inverse property and closure, we know that $-a$ is an field element and that we can add it to the left and right hand sides above:

$$(-a) + 0 = (-a) + (a + (-1) \cdot a)\,.$$

Using the additive identity property at left and associativity/inverses/the additive identity at right gives us

$$(-a) = ((-a) + a) + (-1) \cdot a = 0 + (-1) \cdot a = (-1) \cdot a,$$

which is what we claimed. □

1.2 Finite fields: first examples.

This, however, is not what this class is focused on. As combinatorialists, our first question about an object once we understand it even slightly is always "How many of them are there?", perhaps followed-up by the question "Are there any finite examples?" We try to answer these questions in this section.

Let's start with the second one. Do we know of any objects that are finite commutative groups with respect to two operations $+, \cdot$ at the same time, if we ignore 0 for the multiplicative operation? If you think for a while, you'll see that the answer is yes:

Theorem. $\langle \mathbb{Z}/n\mathbb{Z}, +, \cdot \rangle$ is a field if and only if n is a prime number.

Proof. In week 4, we proved that $\langle \mathbb{Z}/n\mathbb{Z}, + \rangle$ satisfied the axioms for a commutative group for any n, and also saw that $\langle (\mathbb{Z}/n\mathbb{Z})^\times, \cdot \rangle$ satisfied the axioms for a commutative group if and only if n is prime. As a result, we know that $\langle \mathbb{Z}/n\mathbb{Z}, +, \cdot \rangle$ is not a field if n is not a prime number, and only need to check distributivity to see that $\langle \mathbb{Z}/n\mathbb{Z}, +, \cdot \rangle$ is a field if n is prime.

Distributivity, however, is not hard to check! In particular, we know that the integers are distributive (as discussed in Intro to Higher Math / is not a difficult thing to check,

once you figure out what plus and times really mean!) Therefore, we know that for any $a, b, c \in \mathbb{Z}$, we have

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

But if any two numbers are equal, they are certainly equal mod n for any n, as their difference is 0 (which is always a multiple of n. Therefore, for any $a, b, c \in \mathbb{Z}/n\mathbb{Z}$, we have

$$a \cdot (b + c) \equiv (a \cdot b) + (a \cdot c) \mod n,$$

which is precisely the distributive property.

As a side note: this proof method, where we argue that a certain property is "inherited" from a larger structure, is a very common one. Whenever you can in mathematics, look for shortcuts like this — they save time in your proofs, make your proofs clearer, and actually make them more enlightening as well (because the reader now knows **where** this property came from, instead of just believing that it magically held true for some random reason!) □

If we call the number of elements in a field its **order**, just like with groups, we now have the following result:

Theorem. There is a field of order p for any prime p.

This raises a natural question: are there other sizes of fields possible? Let's check!

It's immediate that there are no finite fields of order 1, because any finite field must contain two distinct elements $0 \neq 1$, as stated in our axioms. So the first case to actually consider is whether a finite field of order 4 exists.

Before we decide whether or not this is possible, let's try to think about what any such object $\mathbb{F}_4$ should look like! At first, we can describe its elements without losing any generality as $\{0, 1, a, b\}$ for two nonidentity elements a, b, because we have four elements total and two are identities for our two operations.

What else can we say? On one hand, we know that $\mathbb{F}_4$ should be a group with respect to multiplication, if we remove the 0 element from it. Recall, however, Lagrange's theorem from our past lectures:

Theorem. For any finite group G and subgroup H, we have that $|H|$ divides G.

In particular, suppose that G is a group of prime order and that H is the subgroup generated by any nonidentity element a of G. Then the statement that $|\langle a \rangle|$ divides $|G|$ is just the statement that $|\langle a \rangle| = |G|$, because the only number that divides a prime that is not 1 is that prime itself! Consequently, we have the following corollary to Lagrange's theorem:

Corollary. If G is a finite group of order p, for some prime p, and a is any nonidentity element in G, then $G = \langle a \rangle = \{id, a, a^2, \ldots a^{p-1}\}$.

Apply this corollary to $\mathbb{F}_4$: because $\mathbb{F}_4 \setminus \{0\}$ is a three-element set, it is in particular a group under multiplication of prime order! Therefore, it has the form $\{1, a, a^2\}$, and thus

we actually know $\mathbb{F}_4$'s multiplication table!

$\langle \mathbb{F}_4, \cdot \rangle$	0	1	a	a^2
0	0	0	0	0
1	0	1	a	a^2
a	0	a	a^2	1
a^2	0	a^2	1	a

Checking this lets us verify the multiplication-field properties of $\mathbb{F}_4$. So it suffices to figure out how to fill in the addition table as well, and then to use this to check our additive properties! Let's start by filling in what we know:

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1			
a	a			
a^2	a^2			

We know that addition is a group operation on $\mathbb{F}_4$. Therefore, we know in particular that we cannot have any repetitions in any of our rows or columns! As well, we know that if our field is to be distributive, we need to insure that

$$a + b = c \Rightarrow \forall d, d(a + b) = d \cdot c.$$

But what does this mean for our table as currently constructed? Well: if I determine a value for $(1 + 1)$, say $(1 + 1) = x$, then distributivity tells me that I've set $(a + a) = ax$ and $(a^2 + a^2) = a^2x$! As well, if I tell you that $(1 + a) = y$, you could conclude that $(a + a^2) = ay$ and $(a^2 + 1) = a^2y$. In general, if I tell you what any of the cells corresponding to non-identity elements are in this table, then distributivity tells us how to fill in the **diagonal** that contains that entire cell! For example, suppose that we decided that $(1 + 1) = a$; then we would have

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	a		
a	a			
a^2	a^2			

$\Rightarrow$

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	a		
a	a		a^2	
a^2	a^2			1

Is this possible? Well: consider the column corresponding to a. It must contain a 1, because every symbol shows up in each column/row exactly once. But it cannot contain a 1 in its second cell, as there's already a 1 in that row; as well, it cannot contain a 1 in its last cell, as there's also a 1 in that row!

Therefore, we can conclude that $1 + 1 \neq a$! Similarly, if we try out $1 + 1 = a^2$, we run

into similar problems:

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	a^2		
a	a			
a^2	a^2			

$\Rightarrow$

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	a^2		
a	a		1	
a^2	a^2			a

Here, there is no way to place the symbol a^2 into a's column, which again creates a problem. Therefore, we know that the only thing that might work is $1 + 1 = 0$;

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	0		
a	a			
a^2	a^2			

$\Rightarrow$

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	0		
a	a		0	
a^2	a^2			0

If we use the observation that we cannot have any repetitions in any row or column, we can conclude that the other two cells in this row must be a^2 and a, in that order:

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	0	a^2	a
a	a		0	
a^2	a^2			0

Now, if we want to insure that we have distributivity through our entire array, we can use this "diagonals" trick on these two entries to fill in the rest of our array:

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	0	a^2	a
a	a		0	
a^2	a^2			0

$\Rightarrow$

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	0	a^2	a
a	a	a^2	0	1
a^2	a^2	a	1	0

Because we filled in this array via the distributive property, we know that we have satisfied distributivity; as well, we made 0 the identity and can see that 0 is in every row and column (and this that we have inverses!) As well, our array is symmetric over the main diagonal, so we have commutativity; this leaves just associativity to check!

This, ordinarily, would be difficult. To check it quicker, we can observe that we already have worked with this group! Consider the map $\varphi : \mathbb{F}_4 \to (\mathbb{Z}2\mathbb{Z})^2$ that sends $0 \to (0,0), 1 \to (1,1), a \to (0,1), a^2 \to (1,0)$:

$\langle \mathbb{F}_4, + \rangle$	0	1	a	a^2
0	0	1	a	a^2
1	1	0	a^2	a
a	a	a^2	0	1
a^2	a^2	a	1	0

$\overset{\varphi}{\longleftrightarrow}$

$\langle (\mathbb{Z}2\mathbb{Z})^2, + \rangle$	$(0,0)$	$(1,1)$	$(0,1)$	$(1,0)$
$(0,0)$	$(0,0)$	$(1,1)$	$(0,1)$	$(1,0)$
$(1,1)$	$(1,1)$	$(0,0)$	$(1,0)$	$(0,1)$
$(0,1)$	$(0,1)$	$(1,0)$	$(0,0)$	$(1,1)$
$(1,0)$	$(1,0)$	$(0,1)$	$(1,1)$	$(0,0)$

Using this map, we can see that $\langle \mathbb{F}_4, + \rangle$ is isomorphic to $\langle (\mathbb{Z}2\mathbb{Z})^2, + \rangle$: this is because ϕ is a bijection that is compatible with the group tables of our two objects! In particular, this tells us that because $\langle (\mathbb{Z}2\mathbb{Z})^2, + \rangle$ is a group (as shown before in class,) it is in particular associative, and therefore that $\langle \mathbb{F}_4, + \rangle$ is also associative and a group!

This finishes checking all of the properties needed to be a field; therefore, we have proven the following result:

Theorem. There is a field of order 4.

Using similar methods, we can explore the next nonprime value, 6, that comes up. We can again note that if such a field were to exist, its five nonzero elements would have to form a multiplicative group. Because 5 is prime, we can conclude that this group is of the form $\{1, a, a^2, a^3, a^4\}$ for some nonzero and non-one element a! This gives us our multiplication operation, and leaves us with just the task of determining how addition might work out:

$\langle \mathbb{F}_6, + \rangle$	0	1	a	a^2	a^3	a^4
0	0	1	a	a^2	a^3	a^4
1	1					
a	a					
a^2	a^2					
a^3	a^3					
a^4	a^4					

Once again, we can consider what values can go in the cell corresponding to $(1+1)$. Initially, it may seem like we can place any value we want here. However, notice the following observations:

$$a^4(1+a) = 1 + a^4, \qquad a^3(1+a^2) = 1 + a^3.$$

Consequently, we know that if the $1 + a$ cell in the first row is 0, so is the $(1 + a^4)$ cell! Because values are not repeated in any row, this is impossible. Similarly, we know that it is impossible for the $(1 + a^2), (1 + a^3), (1 + a^4)$ cells in this row to be nonzero, which leaves us with the $(1, 1)$ cell as the only possible zero-containing cell. Therefore, it **must** contain zero, as zero must show up in every row and column!

Distributivity then gives us the following structure:

$\langle \mathbb{F}_6, + \rangle$	0	1	a	a^2	a^3	a^4
0	0	1	a	a^2	a^3	a^4
1	1	0				
a	a		0			
a^2	a^2			0		
a^3	a^3				0	
a^4	a^4					0

You could then try to fill in the rest of this array, but I claim that we already have a contradiction to our field axioms! To see it, consider the following set of elements:

$$H = \{0, 1, a, 1 + a\}.$$

We claim that if $\langle \mathbb{F}_6, + \rangle$ is a commutative group, then H is a subgroup of this group.

To see why, we just need to check that this subset contains the identity, inverses, and is closed. It contains 0, so we have the identity; similarly, because $1 + 1 = 0, a + a = 0$ and we're associative/commutative, we have that $(1 + a) + (1 + a) = 1 + 1 + a + a = 0 + 0 = 0$. Therefore we contain inverses! Similarly, this set is closed under addition, which we can check by casework; we already know that 0+ anything stays in our set, and that anything plus itself is 0 and thus stays in our set. By commutativity, we only have the three remaining cases to check:

- $(1) + (a) = 1 + a$,
- $(1 + a) + (a) = 1 + a + a = 1$,
- $(1 + a) + (1) = 1 + 1 + a = 1$.

So we're a subgroup!

However, we are a subgroup of size **4**, for a group of size 6. We proved that the size of any subgroup must divide the order of our group: consequently, we must have a contradiction somewhere! The only assumption we made thus far was that $\langle \mathbb{F}_6, + \rangle$ is a commutative group; therefore this cannot be possible.

Therefore, we have found a problem with any possible finite field of order 6! Consequently, we have the following result:

Theorem. There is no finite field of order 6.

While this is satisfying, we will want stronger and more general methods to get further results on finite fields. We develop those in the next section:

1.3 Polynomials and finite fields.

Consider the collection of all polynomials with real-valued coefficients, which we denote as $\mathbb{R}[x]$. Take any polynomial $h(x) \in \mathbb{R}[x]$.

Consider the following relation:

$$\equiv_h := \{(f(x), g(x)) \mid \exists q(x) \in \mathbb{R}[x], f(x) - g(x) = q(x)h(x)\}.$$

For example, if $h(x) = x - 2$, we would say that $f(x) = x^2 - 4$ is equivalent to $g(x) = x^2 - 3x + 2$, because

$$f(x) - g(x) = x^2 - 4 - (x^2 - 3x + 2) = 3x - 6 = 3(x - 2) = 3h(x).$$

This is an equivalence relation! To see this, we just check for reflexivity, symmetry and transitivity.

Reflexivity: For any $f(x)$, we want $f(x) \equiv_h f(x)$ to hold. But this is equivalent to asking that for any $f(x)$, $f(x) - f(x) = 0$ is a multiple of $h(x)$: this is true!

Symmetry: We want to show that if $f(x) \equiv_h g(x)$, then $g(x) \equiv_h f(x)$. But this is easy to check: if $f(x) \equiv_h g(x)$, then we can write $f(x) - g(x)$ as some multiple $q(x)h(x)$ of $h(x)$. But this means that $g(x) - f(x) = -q(x)h(x)$ is also a multiple of $h(x)$: in other words, that $g(x) \equiv_h f(x)$.

Transitivity: Suppose that $f(x) \equiv g(x)$ and $g(x) \equiv_h j(x)$. We want to show that $f(x) \equiv_h j(x)$. This is similar to the above. Notice that by definition, there must be two polynomials $q(x), r(x)$ such that $f(x) - g(x) = q(x)h(x)$ and $g(x) - j(x) = r(x)h(x)$. Consequently, we have $f(x) - g(x) + g(x) - j(x) = f(x) - j(x) = (q(x) + r(x))h(x)$. In other words, $f(x) - j(x)$ is a multiple of $h(x)$. So we have $f(x) \equiv_h j(x)$, as desired!

Given any polynomial $h(x) = h_0 + h_1x + \ldots + h_nx^n$ of degree n, we can use this equivalence relation to form the algebraic object $\langle \mathbb{R}[x]/h(x), +, \cdot \rangle$, defined as follows:

- The set here is the collection of all polynomials with degree at most $n - 1$. Notice that **any** element $p(x)$of $\mathbb{R}[x]$ is equivalent to some polynomial with degree at most $n - 1$, which we can prove by induction on the degree of $p(x)$:
  - Base case: if the degree of $p(x)$ is at most $n - 1$, we're trivially done.
  - Inductive step: assume that we can do this for any polynomial of degree at most $m - 1$, and take any polynomial $p(x) = p_0 + p_1x + \ldots p_mx^m$ of degree m, for $m \geq n$. This polynomial is equivalent to $p(x) - \frac{p_m}{h_n}x^{m-n}h(x)$, because they only differ by a multiple of $h(x)$! However, this $p(x) - \frac{p_m}{h_n}x^{m-n}h(x)$ has degree at most $m - 1$; therefore by induction it is equivalent to some polynomial of degree at most $n-1$. Therefore, by transitivity, $p(x)$ itself is equivalent to a polynomial of degree at most $n - 1$, as claimed.

  Moreover, notice that no two distinct polynomials $p(x), q(x)$ of degree at most $n-1$ are equivalent. This is because the claim that $p(x) \equiv_{h(x)} q(x)$ is equivalent to $p(x) - q(x)$ being a multiple of $h(x)$. Because the degree of $p(x) - q(x)$ is less than that of $h(x)$, the only multiple possible of $h(x)$ is 0. Therefore we have $p(x) = q(x)$ and thus that no two distinct polynomials of degree at most $n - 1$ are equivalent, as claimed.

- We define addition for any three polynomials $p(x), q(x), r(x)$ in our set as follows: set $p(x) + q(x) \equiv_{h(x)} r(x)$ if and only if $p(x) + q(x)$, $r(x)$ differ by a multiple of $h(x)$. Because any polynomial has a unique representative in our set, as proven above, this operation is defined for any two polynomials and has a unique and well-defined output in our set.

- Multiplication is defined similarly.

This raises a natural question: is this structure a field? We explore this in two examples:

Example. The structure $\langle \mathbb{R}[x]/h(x), +, \cdot \rangle$, for $h(x) = (x^2 - 2x + 1)$, is not a field.

Proof. We first notice that many of the properties of a field come to us for "free." For example, because $\mathbb{R}[x]$, the collection of polynomials with real coefficients, satisfies associativity and commutativity for both addition and multiplication, has an additive identity 0 and a multiplicative identity 1, and has distributivity, we get to "inherit" all of these properties! This is because (just like with modular arithmetic) equality implies equivalence mod anything: that is, if we know two polynomial expressions are equal, then they are certainly equivalent up to any $h(x)$! (We actually proved this; this was reflexivity!)

So, because 0, 1 are polynomials of degree at most 1, we have everything except for perhaps additive and multiplicative inverses in $\langle \mathbb{R}[x]/(x^2 - 2x + 1), +, \cdot \rangle$.

Additive inverses are immediate: take any element of $\langle \mathbb{R}[x]/h(x), +, \cdot \rangle$. It looks like $a+bx$ for some $a, b \in \mathbb{R}$; therefore, because $(-a) + (-b)x$ is also a degree-at-most-1 polynomial, we know that we have $a + bx$'s inverse in our set, as $(a + bx) + (-a - bx) = 0$. In fact, this proof holds in general: if we look at the collection of all polynomials of degree at most n with coefficients in some field, we will always have additive inverses, because our coefficients have additive inverses!

Multiplicative inverses are the only interesting property to check. We first note that in fact, some elements **do** have multiplicative inverses! Take $1 + 2x$ as an example. Notice that for any $a + bx$ in our set, we have

$$(1 + 2x)(a + bx) = a + (2a + b)x + 2bx^2.$$

However, because $x^2 - 2x + 1 \equiv_{h(x)} 0$ in our set (because for any $h(x)$, $h(x)$ is a multiple of itself!), we have that in fact $x^2 \equiv_{h(x)} 2x - 1$, and therefore that

$$(1 + 2x)(a + bx) = a + (2a + b)x + 2bx^2 \equiv_{h(x)} (a - 2b) + (2a + 5b)x.$$

So: if we want this to be equivalent to 1, we want $2a + 5b = 0$ and $a - 2b = 1$. Solving these equations gives us $a = \frac{5}{9}, b = -\frac{2}{9}$; and indeed, we can check that

$$(1 + 2x)\left(\frac{5}{9} - \frac{2}{9}x\right) = \frac{5}{9} + \left(\frac{10}{9} - \frac{4}{9}\right)x - \frac{4}{9}x^2 \equiv_{h(x)} \frac{5}{9} + \left(\frac{10}{9} - \frac{4}{9}\right)x - \frac{4}{9}(2x - 1) = 1.$$

However, other elements do not have inverses! Namely, consider $(x-1)$, and in particular notice that $h(x) = (x - 1)^2$. Consequently, we know that

$$(x - 1)^2 \equiv_{h(x)} 0.$$

So can $(x - 1)$ have an inverse? Well: if it did have some inverse $(a + bx)$, then we would have

$$(a + bx)^2(x - 1)^2 \equiv_{h(x)} 1 \cdot 1 = 1.$$

But

$$(a + bx)^2(x - 1)^2 \equiv_{h(x)} (a + bx)^2 \cdot 0,$$

and $1 \neq 0$! So this is impossible, and therefore this is not a field. □

However, it is possible for this construction to yield a field:

Example. The structure $\langle \mathbb{R}[x]/h(x), +, \cdot \rangle$, for $h(x) = (x^2 + 1)$, is a field.

Proof. As before, the only interesting thing to check is whether or not we have multiplicative inverses. Unlike last time, $h(x)$ doesn't factor into linear polynomials — therefore, we might hope that we have a chance!

In fact, we do. Take any $a + bx \neq 0$, and notice that

$$(a + bx) \cdot \left(\frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}x\right) = \frac{a^2 - b^2x^2}{a^2 + b^2}.$$

But $x^2 + 1 \equiv_{h(x)} = 0$ implies that $x^2 \equiv_{h(x)} -1$; so we have that

$$(a + bx) \cdot \left(\frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} x \right) = \frac{a^2 - b^2 x^2}{a^2 + b^2} \equiv_{h(x)} \frac{a^2 + b^2}{a^2 + b^2} = 1,$$

and therefore that $a + bx$ has an inverse, for any nonzero $a + bx$!

Therefore, this is a field! □

I claim that this is actually a field you've seen before: to see why, consider when you've interacted with some object of the form

$$\{a + bx \mid a, b \in \mathbb{R}, x^2 = -1\}.$$

. . . it's the complex numbers! That is:

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}, i^2 = -1\}$$

is precisely this set. (In fact, this is what motivated us to derive the inverse formula we came up with above, though you could have certainly just solved the equation $(a+bx)(c+dx) \equiv_{h(x)} 1$ for c, d in terms of a, b if you didn't see this.)

The difference between our successful example and our nonsuccessful example is like the difference between $\mathbb{Z}/14\mathbb{Z}$ and $\mathbb{Z}/11\mathbb{Z}$ in being fields!

In the first case, even though some elements have inverses (i.e $9 \cdot 11 \equiv 1 mod 14$), others do not (7 has no multiplicative inverse mod 14) — in particular, the elements that are factors of 14 don't have inverses! Therefore, $\mathbb{Z}/14\mathbb{Z}$ is not a field, and in fact in general $\mathbb{Z}/n\mathbb{Z}$ for any composite n is not a field.

This is like the issues with $\langle \mathbb{R}[x]/h(x), +, \cdot \rangle$ for $h(x) = (x^2 - 2x + 1)$; this failed to be a field because $h(x)$ had nontrivial factors!

Meanwhile, in the second case, there are no nontrivial factors of 11 that occur in $\mathbb{Z}/11\mathbb{Z}$, because 11 is prime! Consequently, as shown before $\mathbb{Z}/11\mathbb{Z}$ is a field!

In a sense, this is why $\langle \mathbb{R}[x]/h(x), +, \cdot \rangle$, for $h(x) = (x^2 + 1)$ was a field: $x^2 + 1$ had no factors!

Motivated by this, we would hope that in general, "modding out by factor-less polynomials" is a process that will help us make fields! Because "factor-less" is an awkward word, let's define this concept:

Definition. Call a polynomial $p(x)$ **irreducible** if its only factors are 1 and itself.

In the next section, we use this concept to make finite fields. To do so, we will need to borrow one result from abstract algebra, that is not hard to prove but doesn't really belong in a discrete mathematics class:

Theorem. Suppose that F is a field, and $F[x]$ is the collection of all polynomials with coefficients in that field. Then any polynomial in this field can be decomposed into irreducible polynomial factors; moreover, this factorization is unique!

In this sense, irreducible polynomials really are like primes; we can build any polynomial out of them, and any polynomial uniquely factors into a product of such irreducibles!

1.4 Polynomials over $\mathbb{Z}/p\mathbb{Z}$.

Definition. For ease of notation, let $\mathbb{F}_p$ denote $\mathbb{Z}/p\mathbb{Z}$; with this notation, we get to emphasize that we are thinking of $\mathbb{Z}/p\mathbb{Z}$ as a field (and also get to type a lot less.)

Definition. Let $\mathbb{F}_p[x]$ denote the collection of all polynomials with coefficients in $\mathbb{F}_p$, where arithmetic is all done mod p.

Example. For example, in $\mathbb{F}_2[x]$, we have 8 polynomials of degree at most 2:

$$0, 1, x, 1+x, x^2, 1+x^2, x+x^2, 1+x+x^2,$$

because each of the powers of x has two choices of coefficient (0 or 1.)

Arithmetic here is done mod 2: that is,

$$(1+x)+(1+x+x^2) = x^2,$$

because $2 + 2x = 0 + 0x = 0 \bmod 2$. Similarly,

$$(1+x)^2 = 1 + 2x + x^2 = 1 + x^2,$$

because again $2x = 0x$ in mod 2.

This does something surprising, in that $x^2 + 1$ is no longer an irreducible polynomial! In fact, if we check the various degree-2 polynomials in $\mathbb{F}_2[x]$, we can see that there is only one irreducible polynomial: x^2+x+1, which cannot be written as the product of any other polynomials. (Check this by looking at all possible products of smaller-degree polynomials!)

We can define equivalence on $\mathbb{F}_p[x]$ up to some element $h(x) \in \mathbb{F}_p[x]$ in the same way as we did before:

$$\equiv_h := \{(f(x), g(x)) \mid \exists q(x) \in \mathbb{F}_p[x], f(x) - g(x) = q(x)h(x)\}.$$

As before, this is an equivalence relation, and as before to check if any $\mathbb{F}_p[x]/h(x)$ is a field, the only interesting thing to check is multiplicative inverses!

We give one concrete example as a warm-up:

Proposition. $\mathbb{F}_2[x]/h(x)$ is a field of order 4, for $h(x) = x^2 + x + 1$.

Theorem. Again, notice that because we've modded out by a a degree-2 polynomial, the only elements that remain are degree 0 or 1. In other words, we only have four elements in our set: $0, 1, x, 1 + x$.

By the exact same logic as in our $\mathbb{R}[x]$ arguments above, our only interesting task is to determine whether multiplicative inverses exist — everything else is inherited from $\mathbb{F}_2[x]$ "for free," via identical reasoning to what we did before.

To finish our proof, then, we just calculate the multiplication table. All calculations below are done both mod 2 (so $2x = 0x, 2 = 0$) and mod x^2+x+1 (so $x^2 = -x-1 = x+1$):

$\langle \mathbb{F}_2[x]/(x^2+x+1), \cdot \rangle$	0	1	x	$1+x$
0	0	0	0	0
1	0	1	x	$1+x$
x	0	x	$1+x$	1
$1+x$	0	$1+x$	1	x

There is a 1 in every nonzero row and column; thus we have inverses, as claimed!

We can in fact generalize this process much further:

Proposition. Suppose that $h(x)$ is an irreducible polynomial in $\mathbb{F}_p[x]$ of degree n. Then $\mathbb{F}_p[x]/h(x)$ is a field of order p^n.

Theorem. Once again, we need to simply show that every element has a multiplicative inverse, as everything else is done for us by our earlier arguments.

We proceed here by contradiction. Suppose not: that there was some $q(x) \neq 0$ in $\mathbb{F}_p[x]/h(x)$ that has no inverse. Then, if we look at the row corresponding to $q(x)$ in the multiplication table of $\mathbb{F}_p[x]/h(x)$, there is no 1-element! Consequently, because there are as many cells in this row as there are elements in $\mathbb{F}_p[x]/h(x)$, and we are omitting one element from this row, the pigeonhole principle tells us that some element in this row must be repeated at least twice! In other words, there are polynomials $f(x) \neq g(x)$ such that

$$f(x)q(x) \equiv_{h(x)} g(x)q(x).$$

But this is equivalent to asking that

$$(f(x) - g(x))q(x) \equiv_{h(x)} 0;$$

in other words, that $(f(x) - g(x))q(x)$ is a multiple of $h(x)$.

Is this possible? Well: we know that $h(x)$ is irreducible. Therefore, if $(f(x) - g(x))q(x)$ is a multiple of $h(x)$, then $h(x)$ is a factor of either $q(x)$ or $f(x)-g(x)$! However, $q(x), f(x)-g(x)$ are both polynomials of degree strictly smaller than $h(x)$; therefore the only way in which this is possible is if one of $q(x), f(x) - g(x)$ are zero. But we assumed that $q(x) \neq 0$ and that $f(x) \neq g(x)$; so neither is zero!

This gives us a contradiction; consequently, $q(x)$ must have had an inverse! Therefore, $\mathbb{F}_p[x]/h(x)$ is a field, as claimed.

This gives us lots of potential finite field sizes! In our next week's talks, in fact, we will attempt to prove the following theorem:

Theorem. For any n and p, there is an irreducible polynomial of degree n in $\mathbb{F}_p[x]$.

This theorem, if we can prove it, will give us the following corollary:

Corollary. For any prime p and positive integer n, there is a finite field of order p^n.

Once we get to linear algebra, we'll be able to strengthen this theorem as follows:

Theorem. There is a field of order k if and only if k is a prime power; that is, all finite fields have order p^n for some prime p and positive integer n.

But that will need to wait until week 10 (or perhaps next quarter, depending on how things go!)