Finite Fields:
An introduction through exercises

Jonathan Buss
Spring 2014

A typical course in "abstract algebra" starts with groups, and then moves on to rings, vector spaces, fields, etc. This sequence may give the impression that fields form an advanced and arcane subject. In fact, however, fields are rather simpler than other structures, by virtue of being more constrained. This holds particularly for the case of finite fields, which arise in many areas of computer science.

This sequence of exercises will take you through most of the mathematical theory one needs in order to use finite fields.[1] Although none of the steps is difficult by itself, we have quite a few details to go through. When you find yourself stuck, short of "aha!" moments, take a break and come back later. (Or phone a friend, or whatever.)

Fields

Roughly speaking, a "field" is a mathematical domain in which both addition and multiplication have the properties we expect from high-school mathematics. Specifically, a field has the following properties.

Closure: If a and b are elements of the field, then $a+b$ and $a\cdot b$ are also elements of the field.
When there is no danger of confusion, we sometimes write "ab" instead of $a\cdot b$.

The "-ivities": Addition and multiplication are associative, i.e.,

$$(a+b)+c = a+(b+c) \quad \text{and} \quad (a\cdot b)\cdot c = a\cdot(b\cdot c) \ ,$$

and commutative $a+b=b+a$ and $a\cdot b = b\cdot a$.

Also, multiplication distributes over addition: $a\cdot(b+c) = ab+ac$.

Identities: The two elements 0 and 1 are identities for addition and multiplication, respectively. Thus for any element a,

$$a+0=0+a=a \qquad \text{and} \qquad a\cdot 1 = 1\cdot a = a$$

Inverses: Every element a has an additive inverse (denoted $-a$), and every element except 0 has a multiplicative inverse (denoted a^{-1} or $1/a$):

$$a+(-a)=0 \qquad \text{and} \qquad a\cdot a^{-1} = 1 \ .$$

Due to the presence of inverses, we can define subtraction and division:

$$a-b =_{\text{def}} a+(-b) \qquad \text{and} \qquad a/b =_{\text{def}} a\cdot b^{-1} \text{ (if } b\neq 0) \ .$$

[1] It does not (yet) discuss their implementation on a computer. A simple implementation is rather straightforward. The problem of finding the most efficient implementations proves difficult; some aspects remain open.

1. Exercise. Many other familiar algebraic properties follow from the requirements on a field.

1. Show that the cancellation laws hold: if $a+b=a+c$, then $b=c$; and if $ab=ac$, then either $a=0$ or $b=c$.
   (Hint: if $a+b=a+c$, then $-a+(a+b)=-a+(a+c)$. Proceed from there.)

2. Show that $0\cdot a=0$ for every a in the field.
   (Hint: note that $0\cdot a=(0+0)\cdot a=0\cdot a+0\cdot a$. Then use a cancellation law.)

3. Show that a field has no "zero-divisors"; that is, if two elements a and b satisfy $a\cdot b=0$, then either $a=0$ or $b=0$.
   (Hint: if a is not 0, it has an inverse a^{-1}.)

4. Select other rules from high-school arithmetic, and convince yourself that they hold in any field. Some possibilities:

   - $-(-a)=a$ and $(a^{-1})^{-1}=a$ (if $a\neq 0$),
   - $-(a+b)=(-a)+(-b)=-a-b$ and $(a\cdot b)^{-1}=a^{-1}\cdot b^{-1}$,
   - $(-a)(-b)=ab$,
   - etc.

   (Select purely algebraic rules. A rule like "if $a>0$ then $-a<0$," which refers to an ordering relation "$<$", should not be expected to hold in an arbitrary field.)

Since the field operations behave much like operations over the rationals (or integers, or reals), we need not normally distinguish exactly which operations we mean when we use symbols like 1, $+$, etc.—even when we use several different domains at the same time. When we want to emphasize that the symbols refer to the ones of a specific domain F, we will write them with a subscript: "$+_F$", "$\cdot_F$", "0_F" and "1_F". But normally context suffices.

2. Exercise ("just for fun"). The element 0 is special in that it does not have a multiplicative inverse. What happens if 0^{-1} exists?

3. Definition. It may happen that a field F contains a smaller field G: that is, $G\subseteq F$ and G is closed under addition, multiplication and inverses. In this case, we call G a *subfield* of F and F an *extension field* of G. For example, the field $\mathbb{R}$ of real numbers is an extension field of the field $\mathbb{Q}$ of rational numbers.

4. Definition. We will use natural numbers as exponents for repeated multiplication. For $a\in\mathbb{F}$ and $n\in\mathbb{N}$, the notation a^n means the value of $e(n,a)$, where $e:\mathbb{N}\times\mathbb{F}\to\mathbb{F}$ is defined by induction (on $\mathbb{N}$) as $e(0_{\mathbb{N}},a)=1_{\mathbb{F}}$ and $e(i+_{\mathbb{N}}1_{\mathbb{N}},a)=a\cdot_{\mathbb{F}}e(i,a)$.

Some basic finite fields

The most familiar fields are infinite: the rationals, the reals, etc. But finite fields (that is, fields with only finitely many elements) also exist, as we now explore. For this section, let F denote an arbitrary finite field.

F must contain 0 and 1, with $0 \neq 1$ (why?). It must also contain a value of $1 +_F 1$.

5. Exercise.

1. Define $1 + 1 = 0$. Show that with this definition, the set $\{0, 1\}$ is a field.
2. Within the set $\{0, 1\}$, the only other possibility is $1 + 1 = 1$. Show that defining "+" this way does not produce a field. (Hint: what is missing?)

Thus there is exactly one field of size two. You can easily recognize it as the set of integers mod 2.

If F has more than two elements, $1 + 1$ can be some other element, which we may denote by '2'. And then we must have elements $2 + 1$ (i.e., 3), 4, 5, 6, etc. In a finite field like F, of course, this infinite sequence must have only finitely many distinct elements—in other words, some element(s) must appear repeatedly.

6. Exercise. Consider the sequence $f_0 = 0_F$, $f_1 = 1_F$, $f_2 = 2_F$, ... (That is, $f_{i+1} = f_i +_F 1_F$.) Fix $k \in \mathbb{N}$ and $d \in \mathbb{N}$ such that $f_k = f_{k+d}$, with $d > 0$.

1. Show that $f_{n+d} = f_n$ for every $n \in \mathbb{N}$. (In other words, the sequence is a simple cycle, repeated infinitely often.)
2. In particular, we have $0 = f_d$ for some $d > 0$. Show that the smallest such d is a prime number. (Hint: $\mathbb{F}$ has no zero-divisors.)

Thus any finite field $\mathbb{F}$ contains (a subset isomorphic to) the set of integers modulo p, for some prime p. We shall denote this set by GF_p.[2]

7. Definition. The *characteristic* of a field $\mathbb{F}$ is the least positive integer such that the value of the sum $1 + 1 + \cdots + 1$ (p copies of 1) is 0.

(If no such p exists in $\mathbb{F}$—which requires an infinite $\mathbb{F}$—then we say that the field has characteristic 0.[3])

We now show that GF_p is always a field, whatever the prime p.

8. Exercise.

1. Convince yourself that the integers modulo p form a ring; that is, that they satisfy the properties required of a field except that some elements may lack multiplicative inverses.

[2]The notation "GF" stands for "Galois field", in honour of Evariste Galois. The standard notation is actually "$GF(p)$", but the parentheses can get confusing in complicated constructions.

Computer scientists sometimes refer to the set of integers mod p as "$\mathbb{Z}_p$". Algebraists, however, reserve that notation for the ring of p-adic integers, which is uncountably infinite, not a field, and has characteristic 0. They use "$\mathbb{Z}/p\mathbb{Z}$" as a generic notation for the integers mod p. More uses of this "/" notation appear later, along with an explanation.

[3]Not "infinity"—the infinite sum may exist, but it can't have value 0. The non-standard reals provide an example.

2. To show GF_p is a field, we must show that every non-zero element a has an inverse a^{-1} such that $aa^{-1} = 1$. Consider the set of elements that we can obtain by multiplying by a; that is, let S_a be the set

$$S_a = \left\{ ab \mid b \in GF_p \right\} \quad .$$

Show that the set S_a has p *distinct* elements. That is, if $b_1 \neq b_2$ then $ab_1 \neq ab_2$.

3. Conclude that a^{-1} exists in GF_p, satisfying $aa^{-1} = 1$.

It's high time for an example.

9. Example. Fix $p = 5$, and consider the field GF_5 of integers mod 5. What are the values of $1/2$ and $1/3$ in this field?

Over the rational numbers $\mathbf{Q}$, we have $\frac{1}{2} - \frac{1}{3} = \frac{1}{6}$. Does this equation hold in GF_5? Why or why not?

Having a non-zero characteristic yields relations that have no direct analogues in the familiar cases with characteristic 0.

10. Exercise. Recall the "binomial theorem": for any natural number n and any a and b,

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i} \quad .$$

This equation holds in any field. (Its proof is simple algebra, combined with the combinatorial definition of $\binom{n}{i}$. You could have selected it in Exercise 4.)

Fix an arbitrary field $\mathbb{K}$ with characteristic $p > 0$.

1. Show that $(a+b)^p = a^p + b^p$ for every a and b in the field $\mathbb{K}$. (Hint: what is the value of $\binom{p}{i}$ mod p?)

2. Show that $\left(\sum_{i=0}^{k} a_i\right)^p = \sum_{i=0}^{k} a_i^p$, for all $k > 0$ and $a_i \in \mathbb{K}$ $(1 \leq i \leq k)$.

3. Show that $a^p = a$ for each $a \in GF_p$. In other words, every element of GF_p is a root of the polynomial $x^p - x$.
(Hint: write a as a sum.)

We shall see below that $x^p - x$ cannot have any other roots, whatever the field.

Are there other finite fields?

Thus far, we have identified infinitely many finite fields: GF_p for each prime p. If some other field exists, it must contain GF_p; that is, it must be an *extension* of GF_p.

11. Example. We look first at an example—choosing the smallest possible one. We start with GF_2 as the "base field", and presume that we can add another element, while retaining the property of being a field. Call the new element α; we assume $\alpha \neq 0$ and $\alpha \neq 1$, and naturally require $0 + \alpha = \alpha$ and $1 \cdot \alpha = \alpha$.

1. Show that $\alpha + \alpha = 0$. (Use the fact that $1 + 1 = 0$, together with the distributive law.)
Thus α is its own additive inverse: $-\alpha = \alpha$.

2. To satisfy closure, an element $\alpha + 1$ must exist. Show that if $\alpha + 1 \in \{0, 1, \alpha\}$, then the resulting structure is not a field. (Each of the three cases yields a contradiction.)
   Thus $\alpha + 1$ must be a fourth element of the field.

3. The value of $\alpha \cdot \alpha$ (a.k.a. α^2) must also exist. If $\alpha^2 = 0$, then α is a zero-divisor—no good.

   Show that $\alpha^2 = 1$ implies that $(\alpha + 1) \cdot (\alpha + 1) = 0$. ("Multiply it out" using the distributive law.)

   Show that $\alpha^2 = \alpha$ and $\alpha \neq 0$ imply $\alpha = 1$, violating our assumption that $\alpha \neq 1$. (Hint: multiply both sides by α^{-1}, which exists since $\alpha \neq 0$.)

4. Verify that defining $\alpha^2 = \alpha + 1$ determines a field. You may want to write out the full addition and multiplication tables. (Four elements and two operations yields 32 entries. Not so many, really.)

What happens in general, if we start with a field and "adjoin" another element? Can we always create a larger field? Perhaps several different fields?

Extrapolating the previous example, adjoining a new element α requires other elements as well: sums ($\alpha + 1$, $2\alpha = \alpha + \alpha$, ...) and powers ($\alpha \cdot \alpha = \alpha^2$, α^3, ...) in all combinations (e.g., $1 - 2\alpha + 9\alpha^{12}$). In order to analyze this situation, we briefly review polynomials.

12. Definition. A *polynomial* in a variable x over a field $\mathbb{K}$ is an expression of the form

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

where $n \in \mathbb{N}$ and $a_i \in \mathbb{K}$ for $0 \leq i \leq n$, with the condition that either $a_n \neq 0$ or $n = 0$. The case $n = 0$ and $a_0 = 0$ is called the "zero polynomial".

- We denote the set of all polynomials in x over field $\mathbb{K}$ by $\mathbb{K}[x]$.

  In the case $n = 0$, we consider the polynomial ax^0 equivalent to the field element a. Thus we essentially have $\mathbb{K} \subseteq \mathbb{K}[x]$.

- We extend addition and multiplication in $\mathbb{K}$ to polynomials by supposing that the variable x obeys the associative, commutative, and distributive laws. (E.g., $ax + bx = (a + b)x$, $xa = ax$, etc.)

- The *degree* of a non-zero polynomial is the value of n; the degree of the zero polynomial is -1. We denote the degree of a polynomial q by $\deg(q)$.

- A polynomial of degree $n \geq 1$ is *monic* if its leading coefficient a_n is 1. If q has $a_n = c$, then $q = c \cdot q'$ where $q' = c^{-1} \cdot q$ is a monic polynomial of degree n.

13. Exercise.

1. Convince yourself that requiring the variable x to satisfy the associative, commutative, and distributive laws implies that all polynomials satisfies these laws.

   If $\deg(q_0) = n_0$ and $\deg(q_1) = n_1$, what can you say about the degrees of $q_0 + q_1$ and of $q_0 \cdot q_1$, in terms of n_0 and n_1?

2. $\mathbb{K}[x]$ has additive inverses for each element, but not multiplicative inverses (e.g., x has no inverse). Show that $\mathbb{K}[x]$ has no zero-divisors—if $s \cdot t = 0$, then either $s = 0$ or $t = 0$.

14. Definition. Let β be any object, in any domain, to which the operations $+$ and $\cdot$ can reasonably be extended, including satisfying associativity, commutativity, distributivity and the cancellation laws.

- For a polynomial $q(x)$, the object $q(\beta)$ is the value of the expression obtained from $q(x)$ by substituting β for x; that is, the value
$$\sum_{i=0}^{n} a_i \beta^i \quad .$$
(Where this value lies depends on β. If $\beta \in \mathbb{K}$ then also $q(\beta) \in \mathbb{K}$; if $\beta \in \mathbb{K}[x]$ then also $q(\beta) \in \mathbb{K}[x]$; we shall consider other possibilites below.)

- The set of all objects of the form $q(\beta)$ for some $q \in \mathbb{K}[x]$ is denoted $\mathbb{K}(\beta)$; i.e.,
$$\mathbb{K}(\beta) = \{\, q(\beta) \mid q \in \mathbb{K}[x] \,\} \quad .$$
We shall sometimes abuse nomenclature somewhat and refer to an element of $\mathbb{K}(\beta)$ as "a polynomial in β."

- The object β is a *root* of q if $q(\beta) = 0$.

These definitions allow a succinct summary of adjoining an element to a field: if $\alpha \in \mathbb{F}$ and $\alpha \notin GF_p$, then $GF_p(\alpha) \subseteq \mathbb{F}$. In fact, if $\mathbb{K} \subseteq \mathbb{F}$ is any sub-field of $\mathbb{F}$ and $\alpha \in \mathbb{F} - \mathbb{K}$, then $\mathbb{K}(\alpha) \subseteq \mathbb{F}$. We examine this subset of $\mathbb{F}$.

15. Exercise. Let $\mathbb{F}$ be a finite field of characteristic p and $\mathbb{K}$ a sub-field of $\mathbb{F}$. Suppose that $\alpha \in \mathbb{F}$, $\alpha \notin \mathbb{K}$.

1. Show that if $\beta \in \mathbb{K}(\alpha)$ and $q \in \mathbb{K}[x]$, then $q(\beta) \in \mathbb{K}(\alpha) \subseteq \mathbb{F}$.
(That is, all of the values in this exercise are elements of $\mathbb{F}$. We can do arithmetic on them freely.)

2. Since $\mathbb{F}$ is finite, we must have many instances of $q_0(\alpha) = q_1(\alpha)$ for distinct polynomials $q_0(x)$ and $q_1(x)$. Letting $r = q_0 - q_1$, we have $r(\alpha) = 0$—i.e., α is a root of the polynomial r. Fix such a polynomial r of minimal degree n. Without loss of generality, we can take r to be monic. Why?

3. Show that for every polynomial $q \in \mathbb{K}[x]$, there is a polynomial $q' \in \mathbb{K}[x]$ of degree $\deg(q') < n$ such that $q(\alpha) = q'(\alpha)$.
(Hint: we know $r(\alpha) = 0$, where $\deg(r) = n$. This gives an equation involving α^n.)

4. Show that if $\deg(q) < n$, $\deg(q') < n$, and $q \neq q'$, then $q(\alpha) \neq q'(\alpha)$.
(Hint: we chose n to be minimal.)

5. Show that the minimial polynomial $r(x)$ must be irreducible—it has no nontrivial factors. In other words, if $r(x) = s(x) \cdot t(x)$ for monic polynomials s and t over $\mathbb{K}$, then either $s = 1$ or $t = 1$.
(Here, "irreducible" means irreducible in $\mathbb{K}[x]$. If we expand our set of polynomials to $\mathbb{K}(\alpha)[x]$, then $r(x)$ has the factor $x - \alpha$; see below.)

To summarize the first three parts of the exercise, even though the set of polynomials $\mathbb{K}[x]$ is infinite, the set $\mathbb{K}(\alpha)$ of values of these polynomials at α is finite:

$$\mathbb{K}(\alpha) = \left\{ \sum_{i=0}^{n-1} b_i \alpha^i \;\middle|\; b_i \in \mathbb{K} \right\} \quad .$$

By part 4, $\mathbb{K}(\alpha)$ has k^n distinct elements, where k is the size of $\mathbb{K}$.

The definition of the set $\mathbb{K}(\alpha)$ makes no mention of inverses. Nevertheless, it contains an inverse of each of its non-zero elements.

16. Exercise.

1. Consider the equation $\alpha \cdot \left(\sum_{i=0}^{n-1} b_i \alpha^i\right) = 1$, which asserts that the value of the sum is a multiplicative inverse of α. Show that this equation has a unique solution for values b_i in the field $\mathbb{F}$—and these values actually lie in $\mathbb{K}(\alpha)$.
   (Hint: manipulate the left-hand side into a polynomial of degree $n-1$ (or less). Then observe that "1" is actually a constant polynomial (of degree 0).)

2. Extend the above argument to show that each element of $\mathbb{K}(\alpha)$ has a multiplicative inverse in $\mathbb{K}(\alpha)$—in other words, $\mathbb{K}(\alpha)$ is a field.

(In the exercise, we used the finiteness of $\mathbb{F}$ only to justify the existence of the polynomial r. In general, the construction works in any field: if $\mathbb{K}$ is any field and $\alpha \notin \mathbb{K}$ is a root of an irreducible polynomial $r(x) \in \mathbb{K}[x]$, then $\mathbb{K}(\alpha)$ is a field.)

If the original field $\mathbb{F}$ contains yet more elements, not in $\mathbb{K}(\alpha)$, we can select another one and repeat the construction with $\mathbb{K}(\alpha)$ as the field of coefficients, in place of $\mathbb{K}$. This process can continue as often as necessary and yields the following theorem.

17. Theorem. For every finite field, there are a prime $p \in \mathbb{N}$ and an integer $d \geq 1$ such that the field has characteristic p and the field has p^d distinct elements.

To illustrate the construction, we give some examples over familiar fields.

18. Example.

1. Over the field $\mathbf{Q}$ of rational numbers, the polynomial $x^2 - 2$ has no roots. The extension field $\mathbb{R}$, however, contains the root $\sqrt{2}$. Thus we can form $\mathbf{Q}(\sqrt{2})$ comprising all elements of the form $a + b\sqrt{2}$ where $a, b \in \mathbf{Q}$. Convince yourself that $\mathbf{Q}(\sqrt{2})$ is a field.
   (Hint: What is $\sqrt{2}^{-1}$, as an element of $\mathbf{Q}(\sqrt{2})$? What is $(a + b\sqrt{2})^{-1}$?)

2. Over field $\mathbb{R}$, $x^2 + 1$ has no root. If we imagine one—call it ι—then we get a field $\mathbb{R}(\iota)$ that contains a square root of -1. What is the usual name of $\mathbb{R}(\iota)$?

3. Over the field GF_2, the polynomial $x^2 + x + 1$ has no root. (Why?)

   Suppose that some $\alpha \notin GF_2$ satisfies $\alpha^2 + \alpha + 1 = 0$. The above show that $GF_2(\alpha)$ has $2^2 = 4$ elements, which must be 0, 1, α and one other. What is the other, in terms of the first three?
   Write out the multiplication table for $GF_2(\alpha)$, to confirm that it is a field. (Remember that it has characteristic 2.)

If you find yourself worrying that the α of the last example may not exist, compare it to the first two examples. Did you worry whether $\sqrt{2}$ exists? Or ι?

Polynomials over a field

The previous exercises showed the close relationship of finite fields to polynomials. In this part, we shall focus on polynomials. Armed with a few key results, we will then return to the question of classifying all finite fields. For this part, fix a field $\mathbb{K}$ (finite or not); we shall use the term "polynomial" to mean a polynomial with variable x and co-efficients in $\mathbb{K}$ (Definition 12).

As we have seen, polynomials have addition, subtraction and multiplication. What about division? The ratio of two polynomials need not be a polynomial. But we can nevertheless use a division algorithm to get a meaningful result.

19. Lemma (Uniqueness of quotient and remainder). Let s and t be polynomials, with t not zero. There are unique polynomials q (their "quotient") and r (their "remainder") such that

$$s(x) = q(x)t(x) + r(x) \quad \text{and} \quad \deg(r) < \deg(t) \ .$$

20. Exercise.

1. We will prove the lemma by induction on the degree of s. If $\deg(s) = 0$, then s is a constant; prove this case. (Consider two sub-cases separately: either $\deg(t) = 0$ or $\deg(t) > 0$.)

2. If $\deg(s) = n > 0$, and $\deg(t) = m \leq n$, let $s(x) = ax^n + s'(x)$ and $t(x) = bx^m + t'(x)$, where $\deg(s') < n$ and $\deg(t') < m$. Prove that any solution to the equation in the lemma must have $q(x) = ab^{-1}x^{n-m} + q'(x)$ where $\deg(q') < n - m$.

3. Complete the proof by induction on n.

Note that this implies an algorithm to compute q and r from s and t. (It uses $O(n \cdot m)$ operations on elements of $\mathbb{K}$. Asymptotically faster algorithms exist, just as over the integers. But we won't consider that here.)

A close analogy exists between this division of polynomials and division (with remainder) over the natural numbers. If we replace "polynomial" by "natural number" and "deg(s)" by "s" in the statement of the lemma, we get the following familiar statement.

> Let s and t be natural numbers, with t not zero. There are unique natural numbers q (their "quotient") and r (their "remainder") such that $s = qt + r$ and $r < t$.

For polynomials, as for integers, the most interesting cases of division are those that "come out even"—i.e., give a zero remainder.

21. Definition. A polynomial t is a *factor*, or *divisor*, of a polynomial s if there is a polynomial q such that $s = q \cdot t$. If $\deg(t) = 0$ or $\deg(q) = 0$, then t is a *trivial* factor of s; otherwise t is a non-trivial factor of s.

If t is a factor of s (trivial or not), then we say that t *divides* s, and write $t \mid s$.

If all factors of s are trivial, and $\deg(s) > 0$, then s is *irreducible* over the field $\mathbb{K}$.

Note that irreducibility depends on the choice of field $\mathbb{K}$. A polynomial irreducible in $\mathbb{K}[x]$ may factor nontrivially in $\mathbb{K}'[x]$ (where $\mathbb{K} \subseteq \mathbb{K}'$).

Roots of polynomials have a close connection to factors.

22. Example. Suppose that $\alpha \in \mathbb{K}$ is a root of a polynomial s; i.e., $s(\alpha) = 0$. Show that $s(x)$ has $x - \alpha$ as a factor; i.e., there is a polynomial $t(x)$ such that $s(x) = (x - \alpha) \cdot t(x)$. The degree of t is $\deg(s) - 1$. (Hint: write $s(x) = q(x) \cdot (x - \alpha) + r(x)$. If $s(\alpha) = 0$, what does that imply about r?)

23. Definition. The *multiplicity* of a root α of a polynomial s is the largest integer k such that $(x - \alpha)^k$ is a factor of s.

Over the natural numbers, we can compute the greatest common divisor of s and t by repeated division, a.k.a. Euclid's algorithm. We can do the same with polynomials. We first recall the definition of a g.c.d.

24. Definition. Let s and t be polynomials, not both zero. A polynomial g is a *greatest common divisor* (g.c.d.) of s and t iff the following hold.

- $g \mid s$ and $g \mid t$.
- if $h \mid s$ and $h \mid t$, then $h \mid g$.

For polynomials (unlike the natural numbers), a g.c.d. is not unique.

25. Exercise. Let $s \in \mathbb{K}[x]$ and $t \in \mathbb{K}[x]$, not both 0. Show the following.

1. If g is a g.c.d. of s and t and $c \in \mathbb{K}$, $c \neq 0$, then cg is also a g.c.d. of s and t.
2. If g_1 and g_2 are each a g.c.d. of s and t, then there is a $c \in \mathbb{K}$ such that $g_2 = cg_1$.
   (Hint: consider g_i as the "h" in the definition of g.c.d.)

Thus any two g.c.d.'s are multiples of one another—almost as good as uniqueness.[4]

26. Definition. *Euclid's algorithm* takes two polynomials s and t as input, and produces a triple (g, a, b), as follows.

```
Euclid(s, t):
if t is zero, then return (s, 1, 0)
else
  Let q and r be the quotient and remainder of dividing s by t
        (i.e., solve s = q · t + r subject to deg(r) < deg(t))
  (g, a, b) ← Euclid(t, r)
  return (g, b, a − b · q)
```

27. Exercise. Let $s \in \mathbb{K}[x]$ and $t \in \mathbb{K}[x]$, not both 0. Suppose that a call to Euclid(s, t) returns a triple (g, a, b). Show the following.
(Use induction on the number of recursive invocations required—or, equivalently, on the degrees of s and/or t.)

1. $g \in \mathbb{K}[x]$, $a \in \mathbb{K}[x]$ and $b \in \mathbb{K}[x]$.

[4] A similar result holds over the integers: for example, both 2 and -2 are perfectly good g.c.d.'s of 4 and -6. But we ignore this, since their ratio $2/(-2) = -1$ exists in $\mathbb{Z}$. Similarly, the ratio c appearing in Exercise 25 exists in $\mathbb{K}[x]$.

2. Both $g \mid s$ and $g \mid t$.

3. For any h, if $h \mid s$ and $h \mid t$, then $h \mid g$.

4. $g = a \cdot s + b \cdot t$. (In other words, g is a linear combination of s and t.)

The natural numbers have a "unique factorization" property: any non-zero natural number has a unique factorization into primes. Over the integers, we agree to ignore factors of -1; we consider $2 \cdot 3$ and $(-2)(-3)$ as the same factorization of 6. Over $\mathbb{K}[x]$, we shall likewise ignore invertible elements (i.e., polynomials of degree 0). We may thus consider only monic polynomials.

28. Theorem (Unique factorization). Let $s \neq 0$ be a polynomial over a field $\mathbb{K}$. Suppose that $s(x) = \alpha \prod_{i=1}^{k} t_i(x) = \beta \prod_{i=1}^{\ell} u_i(x)$, where $\alpha, \beta \in \mathbb{K}$ and each polynomial t_i and u_i is monic and irreducible. Then $\alpha = \beta$, $k = \ell$ and there is a permutation π such that $t_{\pi(i)} = u_i$ for each i. In other words, aside from the order of factors and the choice of leading coefficients,

every polynomial factors uniquely into irreducible factors.

29. Exercise. To prove the theorem, we shall make use of Euclid's algorithm. First, we note that a simple case implies the entire result.

1. Show that the following claim implies unique factorization.

   > Suppose that $t_1(x)t_2(x) = u_1(x)u_2(x)$ for polynomials t_1, t_2, u_1 and u_2, and that t_1 is irreducible (i.e., has only trivial factors). Then either $t_1 \mid u_1$ or $t_1 \mid u_2$.

   Use induction on the number of factors of $s(x)$.

2. Consider the computation $\text{Euclid}(t_1, u_1) = (g, a, b)$.

   - Why must g be a trivial factor of t_1?
   - Prove the claim for each of the two cases: (1) $g \in \mathbb{K}$ and (2) $\deg(g) = \deg(t_1)$.

   Easy, eh?

(Remark: although each polynomial has a unique factorization, and we have good algorithms to compute the division or g.c.d. of two polynomials, these do not imply a good algorithm to compute a factorization. Just as for integers. . . .)

30. Exercise. Show that a non-zero polynomial of degree n has at most n roots over any field, counting according to multiplicity.

Hint: if s has a root α of multiplicity $k \geq 1$ and thus $(x - \alpha)^k t(x) = s(x)$, what can you say about the degree of t? If $\beta \neq \alpha$ is another root of s, how does β relate to t?

31. Theorem. Let $\mathbb{F}$ have p^d elements, for some $d \geq 1$. Then $\mathbb{F}$ contains exactly the roots of the polynomial $x^{p^d} - x$.

We prove the theorem by induction on d. The base case $d = 1$ we have already proved, in Exercise 3.

32. Lemma. Let q be an irreducible polynomial of degree d over a finite field $\mathbb{F}$ of characteristic p. Let γ be a root of q (in any suitable domain). Then the field $\mathbb{F}(\gamma)$ contains all d roots of q. Further, if r is any irreducible polynomial over $\mathbb{F}$ that has a root $\beta \in \mathbb{F}(\gamma)$, then $\mathbb{F}(\gamma)$ contains $\deg(r)$ roots of r.

33. Exercise.

1. Suppose that an element α of $\mathbb{F}(\gamma)$ is a root of a polynomial r over $\mathbb{F}$. Show that α^p is also a root of r.
2. By the previous part, each of the elements $\gamma^{p^2}, \ldots, \gamma^{p^{d-1}}$ is a root of q.
3. Show that the elements $\gamma, \gamma^p, \gamma^{p^2}, \ldots, \gamma^{p^{d-1}}$ are d distinct roots of q.

34. Corollary. Let $\mathbb{F}$ have p^d elements, for p prime and $d \geq 1$. Then $\mathbb{F}$ contains all p^d roots of the polynomial $x^{p^d} - x$.

35. Exercise. The corollary is equivalent to the claim that every element α of $\mathbb{F}$ satisfies $\alpha^{p^d} = \alpha$.

1. Show that $\gamma^{p^d} = \gamma$.
2. Show that $\alpha^{p^d} = \alpha$ for every element α of $\mathbb{F}$.

36. Corollary. Let p be prime and $d \geq 1$. Then there is (up to isomorphism) exactly one field of size p^d: the set of p^d distinct roots of the polynomial $x^{p^d} - x$.