18.200 Lecture notes March 28, 2024

Modular Arithmetic and Elementary Algebra

Lecturer: Ankur Moitra

These notes cover basic notions in algebra which will be needed for several topics later on. In particular, we will need them to describe the RSA cryptosystem, primality testing algorithms, and error-correcting codes, which we will be covering later in this course.

1 Modular Arithmetic

We start by introducing some simple algebraic structures, beginning with the important example of *modular arithmetic* (over the integers). This is the example we will need for the RSA cryptosystem. Modular arithmetic uses the notion of *equivalence* and *equivalence classes.*

First an easy definition: for integers a, b, m we say that

$$a \equiv b \pmod{m}$$

if $a - b$ is a multiple of m. That is, if

$$a - b = km$$

for some integer $k \in \mathbb{Z}$. Here, "mod" is short for "modulo". For example, $-2 \equiv 8 \pmod{5}$. This congruence relation $\equiv$ modulo m is an equivalence relation (i.e. is reflexive, symmetric and transitive) and partitions the set of all integers into m equivalent classes.

For any integer $n \in \mathbb{Z}$ there is a unique integer r in $\{0, 1, \ldots, m-1\}$ such that $n \equiv r \bmod m$. This follows from the division algorithm, and r is called the *residue* of n modulo m, and by slight overload of notation we refer to it as $n \bmod m$. There is a slight difference between $a = n \bmod m$ and $a \equiv n \pmod{m}$; in the former case, a is the residue and thus between 0 and $m - 1$, while in the latter case, there is no such restriction on a or b. In later notes, however, we often use these notations interchangeably, but the interpretation should be clear from context.

Every residue modulo m lies in the set $\mathbb{Z}_m := \{0, 1, \ldots, m-1\}$, and every element r of $\mathbb{Z}_m$ can be viewed as a representative of the equivalence class $\{n \in \mathbb{Z} : n \equiv r \bmod m\}$ composed of all integers with residue r modulo m. Addition and multiplication over the integers carry over to similar operations over $\mathbb{Z}_m$. More precisely, for $a, b \in \mathbb{Z}_m$ we define $a \oplus b$ to be the residue of $(a+b)$ modulo m. Similarly, we define $a \otimes b$ to be the residue of $(a \times b)$ modulo m. (Later on, we'll be using just + and × but at this point, it is useful to be able to emphasize the fact these operations are different (and apply to different sets) than the usual + and ×.)

Example. In $\mathbb{Z}_5$, one has $3 \oplus 4 = 2$ and $3 \otimes 4 = 2$.

With this definition of $\oplus$ over $\mathbb{Z}_n$, the reader is invited to check the following facts.

(i) For any $a, b \in \mathbb{Z}_n$, $a \oplus b$ also belongs to $\mathbb{Z}_n$.

(ii) For any $a, b, c \in \mathbb{Z}_n$, we have $a \oplus (b \oplus c) = (a \oplus b) \oplus c$.

(iii) The element $0 \in \mathbb{Z}_n$ is the *identity* for $\oplus$ over $\mathbb{Z}_n$. This means that for all $a \in Z_n$, we have $a \oplus 0 = a = 0 \oplus a$.

(iv) For every element $a \in \mathbb{Z}_n$, there exists an element b such that $a \oplus b = 0 = b \oplus a$.

In the next section we will see that the above relations are precisely the conditions showing that $(\mathbb{Z}_m, \oplus)$, like $(\mathbb{Z}, +)$ is an algebraic structure called a *group*. If we consider instead $\mathbb{Z}_n$ with the operation $\otimes$, do we still obtain a group? We'll consider this question in the next section after we define the notion of *group* precisely, and look at several examples. Considering the general notion of group allows one to prove theorems that are valid for *all* groups, instead of doing a proof for each individual example. For instance we will prove Fermat's Little Theorem using general results about groups.

2 Groups

The following treatment of groups is presented in more depth than we really need at this point. Given a set G and a binary operation $*$, if each element in the set obeys the following four properties, then the set and its operation $(G, *)$ is called a *group*.

(i) *Closure.* If $a, b \in G$, then $a * b \in G$.

(ii) *Associativity.* $(a * b) * c = a * (b * c)$ for all $a, b, c \in G$.

(iii) *Existence of an identity element.* Suppose $e \in G$ is the identity element, then $a * e = e * a = a$ for all $a \in G$.

(iv) *Inverse.* For every $a \in G$, there exists an (inverse element) $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$.

If, in addition, each pair of elements $a, b \in G$ satisfies the commutative property, $a * b = b * a$, then the group $(G, *)$ is called an *Abelian group*.

Examples:

- The set of integers forms an (Abelian) group under addition as the rule of composition; and so do the rational, real, or complex numbers. The identity element e in these cases is the number 0, and the inverse of a is $-a$.

- The integers under multiplication, $(\mathbb{Z}, *)$, is not a group since an integer a doesn't have an inverse unless $a = 1$ or $a = -1$. So property (iv) is not satisfied.

- This is an important observation: if one leaves out zero (the additive identity element) the rational, real, and complex numbers each form an (Abelian) group under the operation of multiplication. We need to leave out zero since this element does not have a multiplicative inverse.

- Let Gl_n be the set of invertible $n \times n$ matrices with the usual matrix product as operation. Then $(Gl_n, \times)$ is a group; its identity is the identity matrix. This group is not Abelian (as matrix multiplication is not commutative in general). It is important to restrict to *invertible* matrices to make sure that every element of the group has an inverse.

- Remainders formed by dividing by a polynomial $q(x)$ also form a group with the operation being addition modulo this polynomial $q(x)$. This is a generalization of modular arithmetic (over the integers). Here two polynomials are equivalent if their difference is a polynomial multiple of $q(x)$, i.e. can be written as $k(x)q(x)$ for some polynomial $k(x)$. The possible residues/remainders will be all polynomials of degree smaller than the degree of $q(x)$. For example, if we take the remainder after dividing by say $q(x) = x^3 + 2x^2 + 1$, we can get all polynomials of degree 2 as remainders. The identity is the 0 polynomial ($p(x) = 0$ everywhere) and the inverse of $p(x)$ is the polynomial $-p(x)$.

In Section 1, we saw that, like $(\mathbb{Z}, +)$, the pair $(\mathbb{Z}_m, \oplus)$ forms a group (and this is true for every m), with identity element 0. Similarly, like $(\mathbb{Z}, \times)$ and $(\mathbb{R}, \times)$, the pair $(\mathbb{Z}_m, \otimes)$ is *not* a group because 0 has no multiplicative inverse. We were able to salvage a group for the reals by omitting 0. Can we do the same for $(\mathbb{Z}_m \setminus \{0\}, \otimes)$?

In general, the answer is negative since $a \neq 0$ may not have a multiplicative inverse modulo m. Let $\gcd(a, m)$ be the greatest common divisor of a and b. Then the connection between divisibility and multiplicative inverses in modular arithmetic is:

Lemma 1. *If* $\gcd(a, m) \neq 1$ *then a does not have a multiplicative inverse over $\mathbb{Z}_m$.*

Proof. Let $g = \gcd(a, m)$. Suppose for the purpose of contradiction there is an element b of $\mathbb{Z}_m$ so that $b \otimes a = 1$.Then, unpacking this statement, if we multiply b and a over $\mathbb{Z}$ we should have $ba = km + 1$ for some integer k. However this is impossible because the left hand side is divisible by g but the right hand side is not! □

As a concrete example, you can check that 2 does not have a multiplicative inverse over $\mathbb{Z}_6$. But what if we exclude from $\mathbb{Z}_m$ all elements that are not relatively prime with m? Then we are left with $\mathbb{Z}_m^* = \{a \in \mathbb{Z}_m : \gcd(a, m) = 1\}$. For instance if m is prime then $\mathbb{Z}_m^* = \mathbb{Z}_m - \{0\}$ while for $m = 15$ one gets $\mathbb{Z}_{15}^* = \{1, 2, 4, 7, 8, 11, 13, 14\}$. Our refined question is now: does $(\mathbb{Z}_m^*, \otimes)$ form a group?

Certainly $(\mathbb{Z}_m^*, \otimes)$ satisfies the closure property since the product of two integers relatively prime with m is also relatively prime with m. Associativity is also satisfied since it is inherited from $(\mathbb{Z}, \times)$. Similarly for the identity since $1 \in \mathbb{Z}_m^*$. So the key question is: Does every element of $\mathbb{Z}_m^*$ have a (multiplicative) inverse? The following lemma will help us answer this question in the affirmative.

Lemma 2. *For any positive integers a, b, there exist integers s and t such that* $\gcd(a, b) = sa + tb$.

We will prove this lemma in the next section and will give an ancient algorithm for computing the integers s, t. But for now let's return to our question of whether $(\mathbb{Z}_m^*, \otimes)$ is a group. Consider $a \in \mathbb{Z}_m^*$, i.e., $\gcd(a, m) = 1$. By the above lemma, we have that $1 = sa + tm$ for some integers s, t. But then $sa \equiv 1 \pmod{m}$, so that $s \bmod m$ is the multiplicative inverse of a over $\mathbb{Z}_m$. Now that we know it exists, we can denote it by $a^{-1} \pmod{m}$. Observe that this multiplicative inverse must belong to $\mathbb{Z}_m^*$ (and not just $\mathbb{Z}_m$) since otherwise we couldn't get 1 as $aa^{-1} \bmod m$.

This completes the proof that $(Z_m^*, \otimes)$ forms a group, where $\mathbb{Z}_m^* = \{a \in \mathbb{Z}_m : \gcd(a, m) = 1\}$. In particular for m prime, we have $\mathbb{Z}_m^* = \mathbb{Z}_m \setminus \{0\}$. Summarizing, we can add the following examples to our collection of groups.

- $(\mathbb{Z}_m, \oplus)$ is a group with 0 as the additive identity.

- $(\mathbb{Z}_m^*, \oplus)$ is group for any positive integer m, with 1 as the multiplicative identity.
- In particular, $(\mathbb{Z}_m \setminus \{0\})$ is a group when m is prime.

3 Euclid's Algorithm

Euclid's algorithm (or the Euclidean algorithm) is a very efficient and ancient algorithm to find the $\gcd(a,b)$ of two integers a and b. It is based on the following observations. First, $\gcd(a,b) = \gcd(b,a)$, and so we can assume that $a \geq b$. Secondly $\gcd(a,0) = a$ by definition. Thirdly and most importantly, if

$$a = zb + c$$

where z is an integer then $\gcd(a,b) = gcd(b,c)$. Indeed any divisor of a and b will divide c, and conversely any divisor of b and c will divide a. We can compute c by taking the remainder after dividing a by b, i.e. c is $a \bmod b$. But $c < b < a$ and thus we have made progress by reducing the numbers we have to compute their gcd of. And therefore, we can proceed and express b as:

$$b = yc + d,$$

(thus $d = b \bmod c$) and thus $\gcd(b,c) = \gcd(c,d)$. We continue until we express $\gcd(a,b)$ as $\gcd(g,0) = g$, and at that point, we have found the gcd.

Example. Let $a = 365$ and $b = 211$. Then $c = 154$ and we have that $\gcd(365,211) = \gcd(211,154)$. Continuing, we get:

$$\begin{aligned}
\gcd(365,211) &= \gcd(211,154) \\
&= \gcd(154,57) \\
&= \gcd(57,40) \\
&= \gcd(40,17) \\
&= \gcd(17,6) \\
&= \gcd(6,5) \\
&= \gcd(5,1) \\
&= \gcd(1,0) \\
&= 1.
\end{aligned}$$

For example, 40 was found by taking 154 mod 57. The gcd of 365 and 211 is 1, which means that they are *relatively prime*.

Lemma 2 is an easy consequence of Euclid's algorithm. Indeed, Euclid's algorithm allows to find the integers s and t such that $as + bt = gcd(a,b)$. This clearly proves that no common divisor to a and b is greater than $\gcd(a,b)$ since any common divisor to a and b is also a divisor to $sa + tb$. To find s and t, we proceed bottom up. Suppose we have found u and v such that

$$\gcd(b,c) = ub + vc.$$

Then, knowing that $a = zb + c$ allows us to replace c by $a - zb$ and therefore get:

$$\gcd(a,b) = \gcd(b,c) = ub + v(a - zb) = va + (u - vz)b.$$

Thus, we have expressed the gcd as an integer combination of a and b, knowing it as an integer combination of b and c. Thus bottom up we can find s and t such that

$$\gcd(a, b) = sa + tb.$$

This procedure is often referred to as the *extended Euclidean algorithm.*

Example. Consider again the example with $a = 365$ and $b = 211$. We express their $\gcd(365, 211) = 1$ by going bottom up in the derivation above, and derive:

$$\begin{array}{rclcl}
1 & = & 6 - 5 & & \\
 & = & 6 - (17 - 2 \cdot 6) & = & -17 + 3 \cdot 6 \\
 & = & -17 + 3 \cdot (40 - 2 \cdot 17) & = & -7 \cdot 17 + 3 \cdot 40 \\
 & = & 3 \cdot 40 - 7 \cdot (57 - 40) & = & 10 \cdot 40 - 7 \cdot 57 \\
 & = & 10 \cdot (154 - 2 \cdot 57) - 7 \cdot 57 & = & 10 \cdot 154 - 27 \cdot 57 \\
 & = & 10 \cdot 154 - 27 \cdot (211 - 154) & = & 37 \cdot 154 - 27 \cdot 211 \\
 & = & 37 \cdot (365 - 211) - 27 \cdot 211 & = & 37 \cdot 365 - 64 \cdot 211
\end{array}$$

There is another way to implement the extended Euclidean algorithm, which is a little easier to do on a spreadsheet. If you are finding $\gcd(a, b) = \gcd(a_0, a_1)$ and, after i steps, you have reduced to it to the calculation of $\gcd(a_i, a_{i+1})$ with $a_i \geq a_{i+1}$, you can keep track of two numbers x_i and y_i such that $x_i a + y_i b$ equals a_i. Initially, $(x_0, y_0) = (1, 0)$ and $(x_1, y_1) = (0, 1)$, as $a_0 = 1a + 0b$ and $a_1 = 0a + 1b$. To find x_i and y_i, we know that

$$a_{i-2} = x_{i-2}a + y_{i-2}b$$

and

$$a_{i-1} = x_{i-1}a + y_{i-1}b.$$

As $a_i = a_{i-2} - qa_{i-1}$ where q is the integer part of the quotient between a_{i-2} and a_{i-1}, we simply let $x_i = x_{i-2} - qx_{i-1}$ and $y_i = y_{i-2} - qy_{i-1}$. Indeed, we have:

$$\begin{array}{rcl}
x_i a + y_i b & = & (x_{i-2} - qx_{i-1})a + (y_{i-2} - qy_{i-1})b \\
 & = & a_{i-2} - qa_{i-1} \\
 & = & a_i.
\end{array}$$

For example, we have

gcd	x	y
$\gcd(365, 211)$	1	0
$\gcd(211, 154)$	0	1
$\gcd(154, 57)$	1	-1
$\gcd(57, 40)$	-1	2
$\gcd(40, 17)$	3	-5
$\gcd(17, 6)$	-4	7
$\gcd(6, 5)$	11	-19
$\gcd(5, 1)$	-26	45
$\gcd(1, 0)$	37	-64

Here we found that $6 = 11 * 365 - 19 * 211$ by subtracting twice the equation $17 = -4 \cdot 365 + 7 * 211$ from $40 = 3 * 365 - 5 * 211$.

4 The Chinese Remainder Theorem

Let's start off with a simple but illustrative puzzle:

Example. *A treasure chest contains gold coins, and if the coins are evenly divided among* 5 *pirates there are* 4 *left over. And if they are divided among* 7 *pirates,* 1 *are left over. What is the smallest number of coins that could be in the treasure chest? For this problem, it is easy to work out the answer by hand and figure out that it is* 29*.*

Thinking about this riddle more abstractly, we are given a system of linear equations using modular arithmetic and we'd like to find a solution. Can we determine the remainder $\bmod ab$ given the pairs of remainders $\bmod a$ and $\bmod b$?

An important and powerful theorem was discovered by the Chinese mathematician Sun Tzu in the 4-th century AD and written in his book the Sun Tzu Suan Ching. It says that whenever a and b are relatively prime then there is a bijection between the possible remainders $\bmod ab$ and the pairs of possible remainders $\bmod a$ and $\bmod b$. In other words, the two numbers (the remainder of x upon dividing by a and the remainder of x upon dividing by b) uniquely determines the number x upon dividing by ab, and vice versa. Let's look at another example. Let $a = 7$ and $b = 13$, then $ab = 91$. Any arbitrary remainder, say $73 \bmod 91$, is equivalent to the pair $(3, 8) = (73 \bmod 7, 73 \bmod 13)$. No other remainder mod 91 leads to the pair $(3, 8)$.

Theorem 1 (Chinese Remainder Theorem)**.** *Let a and b be integers that are relatively prime. Each pair of remainders (r, s) mod a and b separately corresponds to exactly one remainder t mod ab such that $r = t \bmod a$ and $s = t \bmod b$.*

Moreover, if one adds or multiplies remainders with respect to ab, the corresponding remainders with respect to a and b separately add or multiply correspondingly.

Proof. In order to show this, first note that the number of possible remainders $\bmod ab$ is ab, while the number of pairs of possible remainders $\bmod a$ and $\bmod b$ is also ab. To any remainder $t \bmod ab$, there corresponds a pair $(t \bmod a, t \bmod b)$ of remainders $\bmod a$ and $\bmod b$. So we only need to show that there cannot exist two distinct remainders x and y upon dividing by ab, and that x and y have the same remainders upon dividing by a and by b. Suppose the contrary. In this case, both a and b divide the difference $x - y$. Since we assumed that a and b are relatively prime, it implies that ab divides $x - y$. This implies that x and y have the same remainder upon dividing by ab and are therefore equal. This is a contradiction. Thus, each remainder of ab corresponds to a unique pair of remainders for a and b separately. This proves the first statement of the theorem. The second statement is easy to check and is left to the reader. □

By the above theorem, we can now describe remainders with respect to ab by the corresponding pairs of remainders with respect to a and b separately: we let (s, t) represent the remainder that is s with respect to a and t with respect to b.

It is easy to find the remainders $\bmod a$ and $\bmod b$ that correspond to the remainder $\bmod ab$. But suppose we want to go in the other direction. That is, suppose that we want to find the remainder modulo ab that correspond to the pair (s, t). Let $x \in \{0, 1, \ldots, ab - 1\}$ be this remainder. We know that $s = x \bmod a$ and $t = x \bmod b$. In particular, $x = kb + t$ where $0 \leq k < a$. Moreover $s \equiv kb + t \pmod{a}$, hence $kb \equiv s - t \pmod{a}$. Now recall that b has a multiplicative inverse in the group $(\mathbb{Z}_a^*, \otimes)$, and that this inverse can be found using Euclid's algorithm. Using this inverse we can compute $k = b^{-1}(s-t) \bmod a$, and then compute $x = kb + t$.

Now let us play a little bit with the remainder pairs representations. Let us try for instance to find the solutions of the equation $x^2 = 1 \bmod ab$ where a and b are relatively prime. Since the remainder $1 \bmod ab$ is represented by the remainder pair $(1,1)$ (where the pair represents the values modulo a and b), it is easy to see that this equation has at least four solutions: the remainder pairs $(1,1),(-1,1),(1,-1),(-1,-1)$. (Here, $(-1,1)$ is a convenient notation for $(a-1,1)$.) Why are these all solutions to $x^2 = 1 \bmod ab$? For any $x \in \{(1,1),(-1,1),(1,-1),(1,1)\}$, we have that x^2 gets represented by $((\pm 1)^2,(\pm 1)^2) = (1,1)$ and, by the Chinese remainder theorem, the only remainder $\bmod(ab)$ that corresponds to this is 1.

For example, consider $a = 3$ and $b = 5$. The remainders $r = 1$, 4, 11, and 14 are relatively prime to $ab = 15$, and correspond to the remainder pairs $(1,1)$, $(1,-1) = (1,4)$, $(-1,1) = (2,1)$ and $(-1,-1) = (2,4)$ respectively. In each case, $r \equiv \pm 1 \pmod 3$ and $r \equiv \pm 1 \pmod 5$ so that $r^2 \equiv (\pm 1)^2 \equiv 1 \pmod 3$ and $r^2 \equiv (\pm 1)^2 \equiv 1 \pmod 5$, so that $r^2 \equiv 1 \pmod{15}$. And indeed $1^2 = 1$, $4^2 = 16$, $11^2 = 121$, and $14^2 = 196$ are all $\equiv 1 \bmod 15$.

5 Lagrange's theorem and Fermat's little theorem

The number of elements in a group G is called the *order* of G, written $|G|$.

Example. The order of $(Z_N, \oplus)$ is N. The order of $(\mathbb{Z}_N^*, \otimes)$ is the number of remainders which are relatively prime to N. If N is a prime then $\mathbb{Z}_N^* = \mathbb{Z}_N - \{0\}$ and the order in $N-1$. Now consider the situation where $N = pq$, where p and q are primes. In this case the order of $(\mathbb{Z}_N^*, \otimes)$ is $(p-1)(q-1) = N - p - q + 1$. To see this, observe that the remainders in $\{1, \ldots, N-1\}$ that have a factor in common with N are multiples of either p or q but not of both. There are $q-1$ of the former type, $p-1$ of the latter. So the order of $\mathbb{Z}_N^*$ is $(N-1) - ((p-1)+(q-1)) = N - p - q + 1$. For example, for $N = 15$, we have that the order of $\mathbb{Z}_{15}^*$ is $2 \cdot 4 = 8$, and indeed, this is the number of elements we found.

A group G is said to have a *subgroup* H if H is a subset of G, and H is also a group (under the same operation $*$ as G). Check for yourself, that if we know G is a group, and we want to know if some subset H of G is a subgroup, the only group properties we really have to check are closure and inverses. We now are ready to state (and prove) one of the simplest and most fundamental facts about groups.

Theorem 2 (Lagrange's Theorem)**.** *Suppose $|G|$ is finite, and H is a subgroup of G. Then $|H|$ divides $|G|$.*

For example, take for G the multiplicative group $\mathbb{Z}_7^*$ with 6 elements, and consider H to be subgroup consisting of all the distinct powers of 2, that is $2, 2^1 = 2, 2^2 = 4, 2^3 = 1 \bmod 7$. H is a subgroup (since it is closed and every element has an inverse), and the order of H is 3 which divides the order of G (equal to 6).

Proof. Suppose H has h elements. We will partition the elements of G into disjoint 'copies' of H. Each such copy will be of the form xH, where x is an element of G and $xH = \{xy : y \in H\}$ denotes the h elements obtained by multiplying (i.e. performing the group operation) x by each element of

H. These copies are called *cosets.* (In the example above, $3H$ corresponds to $\{3\otimes 1, 3\otimes 2, 3\otimes 4\} = \{3,6,5\}$ where $\otimes$ corresponds to multiplication in $\mathbb{Z}_7^*$.)

Suppose we have already identified $x_1, \cdots, x_j$ such that all x_kH are disjoint for $k = 1, \cdots, j$. So far, these sets cover precisely jh distinct elements of G. To initialize this process, we set $x_1 = e$ (where e is the identity element) and $j = 1$. Now, either we have all elements of G or we don't. In the first case, we know that $|G| = j|H|$ and we are done. In the second case, let x_{j+1} be an element of G which is not of the form x_kq for $k \leq j$ and some q in H.

We now add to our list the h additional elements $x_{j+1}H$ of G. We prove that these elements are all new and distinct: if $x_{j+1}h_1 = x_{j+1}h_2$ holds then by multiplying on the left of both sides of this equation by x_{j+1}^{-1} we find that h_1 and h_2 are equal; if $x_kg = x_{j+1}h$ holds for some g and h in H and $k \leq j$, then upon postmultiplying both sides of this equation by h^{-1} we get

$$x_kgh^{-1} = x_{j+1}$$

and therefore $x_{j+1} \in x_kH$ (since $gh^{-1} \in H$ by closure and the existence of inverse), a contradiction. We then increase j by 1 and repeat. If G is finite, this argument must come to an end which can only happen when the order of G is jh for some j. □

Let x be an element of a finite group G. The powers of x form a subgroup of G called the group *generated by* x, and we define the *order* $o(x)$ of an element x to be the order of that subgroup. One can see that the order of x is the smallest positive power k such that $x^k = 1$ (indeed if there were two indices j, l with $j < l \leq k$ and $x^j = x^l$, then $x^{l-j} = 1$ contradicting the definition of k). Hence for all $x \in G$, we must have $x^{o(x)} = 1$, where $o(x)$ denote the order of x. If we apply Lagrange's theorem to G and $x \in G$, then we see that $o(x)$ divides the order $|G|$ of G, and therefore $x^{o(x)} = 1$ implies that $x^{|G|} = 1$ for all $x \in G$. In particular, if we take $G = \mathbb{Z}_p^*$ with p prime, we get Fermat's little theorem (since the order of $\mathbb{Z}_p^*$ is $p - 1$):

Theorem 3 (Fermat's little theorem)**.** *If p is prime and a is not divisible by p then $a^{p-1} = 1 \bmod p$.*

Also, if we take $G = \mathbb{Z}_N^*$ with $N = pq$, p and q being prime, we get that $|G| = (p-1)(q-1)$ and thus $x^{(p-1)(q-1)} = 1$ for all x relatively prime with pq.

Cosets of normal subgroups.
Let G be a group and H be a subgroup with h elements. For any group element $x \in G$, the h elements of G of the form xH form what is called a left *coset* of the subgroup H; a right coset is similarly defined as Hx. If the right and left cosets for each element are the same so that for all a in G we have $aH = Ha$, then H is said to be *normal.* In Abelian groups, all subgroups are normal.

If H is a normal subgroup of G, then its cosets form a group under the rule of composition $aHbH = abH$; this subgroup is called the *factor group* G/H of G with respect to H.

For example, if G is the group $\mathbb{Z}$ of integers under addition, and H is the subgroup consisting of those integers divisible by n (which we denote by $n\mathbb{Z}$), then the factor group has elements which correspond to the remainders upon dividing integers by n. This is called, as we remarked earlier, $\mathbb{Z}_n$, and is often referred to as the *integers mod n.* Thus $\mathbb{Z}_n$ can be seen as the factor group $\mathbb{Z}/n\mathbb{Z}$ of the group $(\mathbb{Z}, +)$.

6 Fields and algebraic equations

$(\mathbb{F}, +, *)$ is a *field* if

1. $(\mathbb{F}, +)$ is an Abelian group with identity element denoted 0.
2. $(\mathbb{F} - \{0\}, *)$ is an Abelian group with identity element denoted $1 \neq 0$.
3. *Distributive property.* For all $a, b, c \in \mathbb{F}$, $a * (b + c) = (a * b) + (a * c)$.

From our previous discussion on $\mathbb{Z}_m$ it is easy to see that if $(\mathbb{Z}_m, \oplus, \otimes)$ is a field if and only if m is a prime. The rationals, real or complex numbers (with usual addition and multiplication) are also fields.

Fields have the important property that the product of any two non-zero elements is not zero.

Lemma 3. *Let $\mathbb{F}$ be a field, and $a, b \in \mathbb{F}$. If $ab = 0$ then either $a = 0$ or $b = 0$.*

Proof. Suppose that neither a nor b is 0. Then a^{-1} and b^{-1} exist and $(b^{-1})(a^{-1})(ab) = 1$. This implies that ab has an inverse, but this cannot be true since $ab = 0$. A contradiction. □

Lemma 4. *If a is a solution of the polynomial equation $p(x) = 0$ with coefficients in a field, then $(x - a)$ divides $p(x)$.*

Proof. We can express $p(x) = q(x)(x - a) + r$ for some polynomial $q(x)$ and remainder r. Since $p(a) = 0$, this implies that $r = 0$. □

We can now prove the fundamental theorem of algebra.

Theorem 4. *A polynomial of degree $d \geq 1$ with coefficients in a field $\mathbb{F}$ can have at most d roots in $\mathbb{F}$.*

Proof. We will show this by induction on k. If $ax + b = 0$, then $x = -ba^{-1}$ is the unique solution, so the statement is true for $k = 1$. Let $p(x)$ be a polynomial of degree $d > 1$, and let $x = a$ be one of its roots. By lemma 4, we have $p(x) = (x - a)q(x)$, where $q(x)$ is a polynomial of degree $d - 1$, and by the induction hypothesis, $q(x)$ has at most $d - 1$ roots. Lemma 3 says that every root of $q(x)(x - a)$ is either a root of $q(x)$ or a root of $(x - a)$. Thus, this implies that $p(x) = q(x)(x - a)$ has at most $(d - 1) + 1 = d$ roots. □

18.200 Principles of Discrete Applied Mathematics
Spring 2024