Columbia University
E6291: Topics in Cryptography
Professor Luca Trevisan

Handout 2
January 29, 1999

Notes on Algebra

This section contains as little theory as possible, and most results are stated without proof. Any introductory book on algebra will contain proofs and put the results in a more general, and more beautiful framework.

For example, a book by Childs [C95] covers all the required material without getting too abstract. It also points out the cryptographic applications.

1 Prime Numbers

By *integer*, we mean a positive or negative integer. We denote by $\mathbf{Z}$ the set $\mathbf{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. A *natural* number is a non-negative integer. We denote by $\mathbf{N}$ the set $\mathbf{N} = \{0, 1, 2, 3 \ldots\}$. We also denote by $\mathbf{Z}^+$ the set $\mathbf{Z}^+ = \{1, 2, 3, \ldots\}$ of positive integers.

For an integer n, we denote by $||n||$ the *length* of n, i.e. the number of bits needed to represent it, i.e. $||n|| = \lceil \log_2 n \rceil$. Logarithms will always be to the base 2, so we will omit the base hereafter. We will denote by $\ln n$ the natural logarithm of n, i.e. the logarithm taken to the base $e = 2.71828\ldots$

For integers k, n, we say that k *divides* n (or that k is *divisor* of n) if n is a multiple of k. For example 5 divides 35. We write $k|n$ when k divides n.

A *prime number* is a positive integer $p \geq 2$ whose only divisors are 1 and p. Notice that 2 is the only even prime number.

The first few prime numbers are $2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots$.

When a number is not prime, it is called *composite*. A composite can always be written (in a unique way) as a product of primes, possibly with repetitions. E.g. $300 = 2 \times 2 \times 5 \times 5$.

There are infinitely many prime numbers (which is very easy to prove), and in fact there are quite a lot of them (which is harder to prove). Specifically, if we define $\pi(n)$ to be the number of prime numbers p such that $2 \leq p \leq n$, then $\pi(n)$ is about $n/\ln n$. Formally

Theorem 1.1 (Prime Numbers Theorem) $\lim_{n\to\infty} \frac{\pi(n)}{n/\ln n} = 1$

The following bounds are also known

$$\pi(n) \geq \frac{n}{\ln n}$$

and, for $n \geq 17$,

$$\pi(n) \leq 1.10555 \frac{n}{\ln n}$$

There is an efficient randomized algorithm that on input an integer tests whether it is prime or not. Therefore if we want to generate a large prime (in the interval from 1 to n, where n can be thought of as a number around 10^{200}) we can just pick a random number in the set $\{1, \ldots, n = 10^{200}\}$ and then test whether it is prime. If it is not, we try again. Each time we have a probability $\approx 1/\ln n \approx 1/460$ of succeeding, so we expect to succeed after less than 500 attempts. Big prime numbers are very important in applied cryptography, and the Prime Number Theorem is a very useful tools to analyze certain cryptographic protocols.

The Prime Number Theorem has exceedingly difficult proofs, but it is easy to prove at least that there are infinitely many primes. Suppose, by contradiction, that there are only finitely many primes, and let them be $p_1, \ldots, p_n$. Consider the number $m = p_1 \cdot p_2 \cdots p_n + 1$. Since m is bigger than any prime, it must be composite, and hence it must be divisible by some prime. We note that m is not divisible by p_1, as when we divide m by p_1 we get the quotient $p_2 \cdots p_n$ and the remainder 1. Similarly, m is not divisible by p_2, neither by p_3, $\ldots$, neither by p_n. So we get a contradiction.

2 Modular Arithmetic

Let a, n be integers ($n \geq 2$). If we try to divide a by n using the grade-school algorithm we end up with two numbers q and r (the quotient and the remainder) such that $aq + r = n$ and $0 \leq r \leq n - 1$. For example, if we divide 15 by 7 we get a quotient 2 and a remainder 1 and the equation $2 \cdot 7 + 1 = 15$. Such numbers q and r are unique. For integers a, b, n we write

$$a = b \pmod n$$

if a and b have the same remainder when divided by n (equivalently, if $a - b$ is a multiple of n). For example $15 = 8 \pmod 7$.

For a fixed integer n, the relation $\cdot = \cdot \pmod n$ has several properties of ordinary equality. For example,

- For every $a \in \mathbf{Z}$, $a = a \pmod n$;
- For every $a, b, c \in \mathbf{Z}$, if $a = b \pmod n$ and $b = c \pmod n$, then $a = c \pmod n$;
- For every $a, a', b, b' \in \mathbf{Z}$, if $a = a' \pmod n$ and $b = b' \pmod n$ then $a + b = a' + b'$;
- For every $a, a', k, k' \in \mathbf{Z}$, if $a = a' \pmod n$ and $k = k' \pmod n$, then $ak = a'k' \pmod n$.

The last two properties imply that when we do arithmetic operations modulo n, then we obtain the same result if we replace one term by another one that is equal modulo n. In particular, it is the same if every term a is replaced by the remainder of its division by n. So, when doing operations modulo n, we can restrict ourselves to use only the integers $0, \dots, n-1$.

We denote by $\mathbf{Z}_n = \{0, 1, \dots, n-1\}$, and we define on this set operations of addition and multiplication modulo n. Therefore, for every two elements $a, b \in \mathbf{Z}_n$ we define the element $a+b \pmod{n}$, which is defined as the (unique) element c of $\mathbf{Z}_n$ such that $c = a+b \pmod{n}$.

For example, $5 + 4 = 2 \pmod{7}$ and $2 + 1 = 0 \pmod{3}$.

Similarly, we define a product operation in $\mathbf{Z}_n$. For example, $3 \cdot 4 = 3 \pmod{9}$.

Addition in $\mathbf{Z}_n$ is "invertible." Specifically, for each element $a \in \mathbf{Z}_n$ there is an element $a' \in \mathbf{Z}_n$ such that $a + a' = 0 \pmod{n}$ (one can take $a' = n - a$). This gives an analog in $\mathbf{Z}_n$ of the subtraction operation.

Multiplication, alas, is not necessarily invertible. That is, it is not necessarily true that for an element $a \in \mathbf{Z}_n$ there is an element $a' \in \mathbf{Z}_n$ such that $a \cdot a' = 1 \pmod{n}$.

Consider for example $\mathbf{Z}_6$ and the element 2. If there was an element $a \in \mathbf{Z}_6$ such that $2 \cdot a = 1$, then we would have $3 \cdot 2 \cdot a = 3 \cdot 1 = 3 \pmod{6}$. But we also have $3 \cdot 2 \cdot a = 6 \cdot a = 0 \cdot a = 0 \pmod{6}$, and so we have a contradiction. It is possible to characterize precisely the cases where an element a has an inverse with respect to multiplication in $\mathbf{Z}_n$. To this aim, we need a result that is also useful for its algorithmic aspect. Recall that the greatest common divisor (abbreviated gcd) of two numbers n and m is the largest integer that is both a divisor of n and a divisor of m.

Theorem 2.1 (Euclid's algorithm) *There exists an algorithm that on input two positive integers m and n returns $k = \gcd(m, n)$ and two integers α, β such that $\alpha n + \beta m = k$. The algorithm runs in time polynomial in the number of digits of n and m.*

Example 1 On input 14 and 10, Euclid's algorithm returns $2 = \gcd(10, 14)$ and the coefficients $\alpha = 3$ and $\beta = -2$. Indeed, $3 \times 10 - 2 \times 14 = 2$. ■

Example 2 On input 60 and 17, Euclid's algorithm returns $1 = \gcd(60, 17)$ and the coefficients $\alpha = 2$ and $\beta = -7$. Indeed, $2 \times 60 - 7 \times 17 = 1$. ■

Let us return to the issue of inverses in $\mathbf{Z}_n$. Suppose a and n are such that $\gcd(a, n) = k > 1$, we will show that a cannot have an inverse. Let us call $b = n/k$. Note that $b \in \mathbf{Z}_n$, $b \neq 0$. Assume by contradiction that there exists $a' \in \mathbf{Z}_n$ such that $a \cdot a' = 1 \pmod{n}$, then $b \cdot (a \cdot a') = b \pmod{n}$, but also (doing operations over the integers now) $b \cdot a \cdot a' = (n/k) \cdot a \cdot a' = n \cdot (a/k) \cdot a'$ which is a multiple of n (since a/k is an integer), and therefore we have $(b \cdot a) \cdot a' = 0 \pmod{n}$.

Consider now the case $\gcd(a, n) = 1$. Then, using Euclid's algorithm, we can find coefficients α, β such that $\alpha a + \beta n = 1$, that is $\alpha a = 1 \pmod{n}$, that is α is an inverse of a. In this case not only does a have an inverse, but we can also find it efficiently using Euclid's algorithm.

We say that two integers n and m are co-prime if $\gcd(n, m) = 1$. Putting everything together we have

Theorem 2.2 *For an element $a \in \mathbf{Z}_n$, there exists an element $a' \in \mathbf{Z}_n$ such that $a \cdot a' = 1 \pmod{n}$ if and only if a and n are co-prime.*

3 Groups

Definition 3.1 (Group) *A group is a set G endowed with an operation, that we denote, say, by $\otimes$, that given two elements of G returns an element of G (i.e. for every $a, b \in G, (a \otimes b) \in G$; the operation must satisfy the following properties:*

1. *For every $a, b \in G$, $a \otimes b = b \otimes a$;*
2. *For every $a, b, c \in G$, $(a \otimes b) \otimes c = a \otimes (b \otimes c)$;*
3. *There exists an element $u \in G$ such that for every $a \in G$, $a \otimes u = u \otimes a = a$;*
4. *For every element $a \in G$ there exists an element $a' \in G$ such that $a \otimes a' = u$.*

Remark. To be precise, what we have just defined is the notion of Abelian group, that is a special type of groups. In general, G can be a group even if Property 1 is not satisfied (in such a case, it will be called a non-Abelian group). In this course we will never consider non-Abelian group, so it is not necessary to insist on the difference. An example of a non-Abelian group is the set of $n \times n$ matrices, together with the matrix multiplication operation. ■

Our canonical example of a group is $\mathbf{Z}_n$ with the operation $\cdot + \cdot \pmod{n}$.

Here is another interesting example. For a prime p, define $\mathbf{Z}^*{}_p = \{1, 2, \ldots, p-1\}$.

Theorem 3.2 *If p is a prime, then $\mathbf{Z}^*{}_p$ together with multiplication $\pmod{p}$ is a group.*

To prove the theorem we have to check the required properties of a group. First of all, it is definitely true that $a \cdot b = b \cdot a \pmod{p}$ and that $a \cdot (b \cdot c) = (a \cdot b) \cdot c \pmod{p}$. Furthermore we have a special element u, namely 1, such that $a \cdot 1 = 1 \cdot a = a \pmod{p}$. Using the results of the previous section, and the fact that a prime number is co-prime with every other number small than itself, we also have that for every $a \in \mathbf{Z}^*{}_p$ there is an $a' \in \mathbf{Z}^*{}_p$ such that $a \cdot a' = 1 \pmod{p}$. In fact, we should also check one more property, namely that for every $a, b \in \mathbf{Z}^*{}_p$ it is true that $a \cdot b \pmod{p} \in \mathbf{Z}^*{}_p$, i.e. that is impossible that $a \cdot b = 0 \pmod{p}$. This follows by contradiction: if $a \cdot b = 0 \pmod{p}$, this means that $a \cdot b$ (taking the product over the integers) is a multiple of p. Since p is a prime number, it means that either a or b is a multiple of p. But both a and b are smaller than p, and so we have a contradiction.

It would be nice to have a similar result for arbitrary n, and say that $\mathbf{Z}_n - \{0\}$ is a group with respect to multiplication $\pmod{n}$. Unfortunately this is not true when n is a composite number (elements of $\mathbf{Z}_n$ having some common factor with n do not have an inverse, as seen in the previous section). Yet, it is still possible to define a group.

Define $\mathbf{Z}^*{}_n = \{a : 1 \leq a \leq n-1 \text{ and } \gcd(a,n) = 1\}$. For example, $\mathbf{Z}^*{}_6 = \{1,5\}$ and $\mathbf{Z}^*{}_{10} = \{1,3,7,9\}$. Note that the definition of $\mathbf{Z}^*{}_p$ for p prime is a special case of the previous definition.

Theorem 3.3 *For every positive integer n, $\mathbf{Z}^*{}_n$ is a group with respect to multiplication.*

We denote by $\phi(n)$ the number of elements of $\mathbf{Z}^*{}_n$, i.e. the number of elements of $\{1,2,\ldots,n-1\}$ that are co-prime with n. It is easy to compute $\phi(n)$ given a factorization of n (but is hard otherwise).

Theorem 3.4

1. *If p is prime and $k \geq 1$ then $\phi(p^k) = (p-1)p^{k-1}$.*
2. *If n and m are co-prime then $\phi(nm) = \phi(n)\phi(m)$.*
3. *If the factorization of n is $\prod_i q_i^{k_i}$ then $\phi(n) = \prod_i (q_i - 1)q_i^{k_i - 1}$.*

Note that the third item in the theorem follows from the first two.

Example 3 . In order to compute $\phi(45)$ we compute the factorization $45 = 3^2 \cdot 5$ and then we apply the formula $\phi(45) = 3 \cdot (3-1)^{2-1} \cdot (5-1) = 24$. Indeed, we can check that $\mathbf{Z}^*{}_{45} = \{1,2,4,7,8,11,13,14,16,17,19,22,23,26,28,29,31,32,34,37,38,41,43,44\}$ has 24 elements. ■

For every n, the group $\mathbf{Z}_n$ with respect to the sum has the following nice property: every element $k \in \mathbf{Z}_n$ can be obtained by summing 1 to itself k times. One can define a generalization of this property for arbitrary groups.

Definition 3.5 *let G be a group with n elements and operation $\otimes$. Suppose there exists an element $g \in G$ such that g, $g \otimes g$, $g \otimes g \otimes g, \ldots,$ $\underbrace{g \otimes g \otimes \cdots \otimes g}_{n \text{ times}}$ are all different (and so cover all the elements of G). Then G is said to be* cyclic *and g is said to be a generator of G.*

Therefore $\mathbf{Z}_n$ is always cyclic and 1 is a generator. Quite interestingly, there is a similar result for $\mathbf{Z}^*{}_p$, p prime.

Theorem 3.6 *For every prime p, $\mathbf{Z}^*{}_p$ is a cyclic wit respect to multiplication* $\pmod{p}$.

We denote by $a^k \pmod{n}$ the value $\underbrace{a \cdot a \cdots a}_{k \text{ times}} \pmod{n}$.

Example 4 The group $\mathbf{Z}^*{}_7$ has generators 3 and 5. Indeed, we have

$$3^1 = 3 \pmod 7 \ , \ 3^2 = 2 \pmod 7 \ , \ 3^3 = 6 \pmod 7 \ ,$$

$$3^4 = 4 \pmod 7 \ , \ 3^5 = 5 \pmod 7 \ , \ 3^6 = 1 \pmod 7$$

The sequence of powers of 5 is $5, 4, 6, 2, 3, 1$. ■

The following results are useful in the analysis of RSA.

Theorem 3.7 (Fermat's Little Theorem) *If p is a prime and $a \in \mathbf{Z}_p^*$, then $a^{p-1} = 1 \pmod p$.*

Fermat's theorem is a special case of the following result.

Theorem 3.8 (Euler's theorem) *If $n \geq 2$ and $a \in \mathbf{Z}_n^*$, then $a^{\phi(n)} = 1 \pmod n$.*

In fact the definition of the function $\phi()$ is due to Euler.

Some of our interest in the notions of cyclic groups and generators is related to the exponentiation function (a candidate one-way function). In that application, one has to be able to generate at random a generator of a group $\mathbf{Z}^*{}_p$ (p prime). For this random generation, it is important to be able to test whether a given element is a generator and to know how many elements are generators.

Given n, $a \in \mathbf{Z}^*{}_n$ *and the factorization of* $\phi(n)$, there is an efficient algorithm that checks whether a is a generator for $\mathbf{Z}^*{}_n$. (By "efficient" algorithm we mean an algorithm that runs in time polynomial in the number of digits of n and a — and in the number of bits needed to represent the factorization of $\phi(n)$, which is polynomial in the number of digits of n anyway.)

Furthermore, the following results guarantee that there are several generators.

Theorem 3.9 *For a prime p, $\mathbf{Z}^*{}_p$ has $\phi(p-1)$ generators.*

Theorem 3.10 *For every n, $\phi(n) \geq n/6 \ln \ln n$.*

Using the prime number theorem, it is a triviality to prove (for $n \geq 17$) the weaker bound $\phi(n) \geq \frac{(n-1)}{\ln(n-1)} - \log n$. Indeed, there are $\pi(n-1) \geq \frac{(n-1)}{\ln(n-1)}$ primes in the interval $2, \ldots, n-1$, and all of them are co-prime with n, except those that are factors of n. But n has at most $\log n$ factors.

4 Some more algorithmic tools

4.1 The Chinese Remainders Theorem

The following is a very useful algorithmic result.

Theorem 4.1 (Chinese Remainders Theorem) *Consider a system of congruences of the form*

$$\begin{array}{rcll} x & = & a_1 & \pmod{n_1} \\ x & = & a_2 & \pmod{n_1} \\ & & \cdots & \\ x & = & a_k & \pmod{n_k} \end{array}$$

Where $n_1, \ldots, n_k$ are pairwise co-prime. Then there is always a solution x in the interval $1, \ldots, n_1 \cdot n_2 \cdots n_k - 1$, and this solution is the only one in the interval. Such a solution x is efficiently computable given $a_1, \ldots, a_k, n_1, \ldots, n_k$. Furthermore, the set of all solutions is precisely the set of integers y such that $y = x \pmod{n_1 \cdot n_2 \cdots n_k}$.

For example, consider the system

$$\begin{array}{rcll} x & = & 5 & \pmod{7} \\ x & = & 2 & \pmod{6} \\ x & = & 1 & \pmod{5} \end{array}$$

Then: $x = 26$ is a solution; 26 it is the only solution in the interval $1, \ldots, 209$; for every integer i we have that $26 + 210i$ is a solution; there is no other solution.

The algorithm to find the solution is simple. Let us call $N = n_1 \cdot n_2 \cdots n_k$ and $N_i = N/n_i$. Let us also call $y_i = (N_i)^{-1} \pmod{n_i}$, that is, y_i is such that $y_i \cdot N_i = 1 \pmod{n_i}$. By our assumptions on $n_1, \ldots, n_k$ it must be that $\gcd(N_i, n_i) = 1$, so y_i is well defined. Then a solution to the system is

$$\sum_{i=1}^{k} a_i N_i y_i \pmod{n_1 \cdot n_2 \cdots n_k}$$

4.2 Quadratic Residues

A (positive) integer x is said to be a *perfect square* if there is some integer $y \in \mathbf{Z}$ such that $x = y^2$; if so, y is said to be a *square root* of x. For example 25 is a perfect square, whose square roots are 5 and -5, while 20 is not a perfect square.

If an integer x is a perfect square then it has precisely two square roots, and they are efficiently computable given x. Unfortunately, things are not so easy with modular arithmetic.

Definition 4.2 *An element $x \in \mathbf{Z}_n$ is said to be a quadratic residue $\pmod{n}$ if there exists an element $y \in \mathbf{Z}_n$ such that $x = y \cdot y \pmod{n}$. If so, y is said to be a square root of $x \pmod{n}$.*

Let us first consider the case of $\mathbf{Z}_p$ with p prime. Then things are not too different from the case of the integers.

Theorem 4.3 *Let p be a prime number. If $x \in \mathbf{Z}_p$ $(x \neq 0)$ is a quadratic residue, then it has exactly two square roots.*

Theorem 4.4 *There exists an efficient randomized algorithm that on input p prime and $x \in \mathbf{Z}_p$ tests whether x is a quadratic residue and, if so, returns the two square roots of x.*

Among the integers, perfect squares are quite rare, and they get sparser and sparser: there are only about $\sqrt{n}$ perfect squares in the interval $1, \ldots, n$. The situation is quite different in the case of $\mathbf{Z}_p$.

Theorem 4.5 *Let $p \geq 3$ be a prime. Then $\mathbf{Z}_p$ has $(p-1)/2$ non-zero quadratic residues.*

This is easy to see: each one of the $p-1$ element $y \neq 0$ is the square root of $y \cdot y$, on the other hand every quadratic residue has two square roots, and so there must be $(p-1)/2$ quadratic residues.

When we are in $\mathbf{Z}_n$ the situation gets more involved, and it is conjectures that no efficient algorithm, on input x and n, can determine whether x is a quadratic residue, let alone find a root. The same conjecture holds when n is the product of two large primes.

On the other hand, if $n = pq$ is the product of two primes, and *the factorization of n is known*, then extracting square roots becomes feasible. The next two results are proved using the Chinese Remainders Theorem. The algorithm of Theorem 4.7 uses the algorithm of Theorem 4.4 as a subroutine.

Theorem 4.6 *Let n be the product of two primes. If x is a quadratic residue $\pmod{n}$, then x has precisely 4 square roots in $\mathbf{Z}_n$.*

Theorem 4.7 *There is an efficient randomized algorithm that on input (x, p, q) (where p and q are prime and $x \in \mathbf{Z}_{pq}$) tests whether x is a quadratic residue $\pmod{pq}$; if so, the algorithm finds all the four square roots of x.*

References

[C95] L.N. Childs. *A Concrete Introduction to Higher Algebra (second edition)*. Springer, 1995.