Notes on Number Theory

Óscar Pereira[1]

6th February 2020

[1] oscar@randomwalk.eu.

Contents

Chapter 1

Modular arithmetic

1.1 Division

Good old division theorem tells us that given integers a, b, with $b > 0$, there exist integers q, r such that $a = bq + r$, with $r \in [0, b[$. What happens when $a < 0$? Well, if $r = 0$, everything stays the same, i.e. if $a = bq$, then $-a = b(-q)$. If $r \neq 0$ however, then $-a = b(-q) - r$, and to again place the remainder in $[0, b[$, we can do $-a = b(-q) - r + b - b = b(-q-1) + (b-r)$. As we have assumed that $r \in]0, b[$ (due to $r \neq 0$), we conclude that $b - r \in]0, b[$. Also, this shows that, when the remainder of both divisions is in $]0, b[$, we have: $(a - bq) + (-a - b[-q-1]) = b$.

The *floor* and *ceiling* functions are defined as usual: $\lfloor x \rfloor \stackrel{\text{def}}{=} x - \varepsilon$ and $\lceil x \rceil \stackrel{\text{def}}{=} x + \varepsilon$, with $\varepsilon \in [0, 1[$ and $x \in \mathbb{R}$ (and of course, $\lfloor x \rfloor, \lceil x \rceil \in \mathbb{Z}$). For each x, the corresponding ε is unique. Hence, if we manage to write x as $x = x' + \alpha$, with x' integer and $\alpha \in [0, 1[$, we can conclude that $x' = \lfloor x \rfloor$; and similarly for the ceiling function (with subtracting α instead of adding).

If $a = bq + r$ as above, then dividing everything by b yields $a/b = q + r/b$, and as $r/b \in [0, 1[$, it follows that $q = \lfloor a/b \rfloor$.

Computing the remainder of a by b is denoted $a \bmod b$. As integer division above has only been defined for a positive b, we generalise the remainder operation to any nonzero modulus by setting $a \bmod b \stackrel{\text{def}}{=} a - b\lfloor a/b \rfloor$. This coincides with the remainder of normal division when $b > 0$, but for $b < 0$, the modulus is in the interval $]-b, 0]$. This can be seen as as follows $a \bmod b \stackrel{\text{def}}{=} a - b\lfloor a/b \rfloor = a - b(a/b - \varepsilon) = b\varepsilon$—and in particular when $b < 0$, that expression ranges in the interval $]-b, 0]$.

Theorem 1.1.1. *Let n be an integer such that $n \geqslant 2$, and $x \in \mathbb{R}$. Then $n\lfloor x \rfloor \leqslant \lfloor nx \rfloor \leqslant n\lfloor x \rfloor + n - 1$ holds.*

Proof idea: $n\lfloor x \rfloor$ is smaller than $\lfloor nx \rfloor$ because the decimal factor of x, ε, is not multiplied by n (as is the case in $\lfloor nx \rfloor$). The difference between the two, $\lfloor n\varepsilon \rfloor$, is at most $n - 1$, which explains the last inequality.

Proof. We have that $n\lfloor x \rfloor = nx - n\varepsilon$. Thus $n\lfloor x \rfloor = nx - n\varepsilon = nx - \lfloor n\varepsilon \rfloor - \varepsilon'$, with $\varepsilon' \in [0, 1[$. As the lhs is an integer, so is the rhs, and as $\lfloor n\varepsilon \rfloor$ is also integer, ε' is the decimal part of nx—which means $\lfloor nx \rfloor = nx - \varepsilon'$, which is clearly greater or equal

to $nx - \lfloor n\varepsilon \rfloor - \varepsilon'$. This shows that $n\lfloor x \rfloor \leqslant \lfloor nx \rfloor$. And now $\lfloor nx \rfloor = nx - \varepsilon' = (nx - \lfloor n\varepsilon \rfloor - \varepsilon') + \lfloor n\varepsilon \rfloor \leqslant n\lfloor x \rfloor + n - 1$, as $\lfloor n\varepsilon \rfloor \leqslant n-1$. This shows $\lfloor nx \rfloor \leqslant n\lfloor x \rfloor + n - 1$. ■

Theorem 1.1.2. *For $x \in \mathbb{R}$ and n a positive integer, we have: $\lfloor \lfloor x \rfloor / n \rfloor = \lfloor x/n \rfloor$.*

Proof. Do integer division for $\lfloor x \rfloor$ and n, to get $\lfloor x \rfloor = nq + r$, from where we get:

$$\frac{\lfloor x \rfloor}{n} = q + \frac{r}{n} \tag{1.1}$$

As q is an integer, and $r/n < 1$, we see that $q = \lfloor \lfloor x \rfloor / n \rfloor$. And thus:

$$\frac{\lfloor x \rfloor}{n} = \left\lfloor \frac{\lfloor x \rfloor}{n} \right\rfloor + \frac{r}{n} \tag{1.2}$$

On the other hand,

$$\frac{x}{n} = \frac{\lfloor x \rfloor + \varepsilon}{n} = \frac{\lfloor x \rfloor}{n} + \frac{\varepsilon}{n} = \left(\left\lfloor \frac{\lfloor x \rfloor}{n} \right\rfloor + \frac{r}{n} \right) + \frac{\varepsilon}{n} \tag{1.3}$$

Now r/n is at most $(n-1)/n$, and $\varepsilon/n < 1/n$, which means their sum is always strictly less than 1. As $\lfloor \lfloor x \rfloor / n \rfloor$ is an integer, then $r/n + \varepsilon/n$ is precisely the decimal part of x/n, entailing that $\lfloor \lfloor x \rfloor / n \rfloor = \lfloor x/n \rfloor$. ■

1.2 gcd and lcm

One of the ways of the defining the gcd is straightforward: given two integers a and b, it is just the greatest of their non-negative common divisors. Given that 1 is a common divisor of every number, the set of nonneg common divisors is always nonempty, and it's also finite, *almost* always. The rub lies precisely in what happens when both numbers are 0: for then every integer is a common divisor, and thus there is no "greatest" common divisor. But we can work around that case.

Definition 1.2.1. *Given two integers, not simultaneously zero, a and b, their* **greatest common divisor** *(gcd) is the greatest non-negative integer d such that it divides both a and b. If $a = b = 0$, then we define $gcd(0,0) = 0$.*

From this way of defining the gcd we get that $gcd(0,a) = |a|$ holds for any integer a.

Lemma 1.2.2. *Let a and b be two integers, and let $d = gcd(a,b)$. Then $d = xa + nb$, for some $x, y \in \mathbb{Z}$. Furthermore, every other common divisor of both a and b, also divides d.*

Proof. From the way we have defined the gcd, the result is obvious when either a or b, or both, are 0. So let us assume that neither is 0.

Consider the set $S = \{ r,s \in \mathbb{Z} \mid ar + bs \geqslant 1 \}$. This set is not empty (e.g. make $r = a$ and $s = b$); thus, by the WOP it contains a smallest element.[1] Let $d = xa + by$ be that element. Dividing a by d, we get $a = dq + r \Leftrightarrow a = (xa + by)q + r \Leftrightarrow r = a(1 - xq) - byq$. Thus the remainder is also a linear combination of a and b—which

[1] For the Well-Ordering Principle (WOP), cf. Abstract Algebra.

means that $r \in S$. But r must be smaller than d, and d is the smallest element of S, which means the remainder is zero (i.e. $d \mid a$). With b a similar reasoning shows that $d \mid b$—and thus d is a common divisor of both a and b.

Given that any number that divides a and b must divide any linear combination of theirs, we conclude that any common divisor of a and b must also divide d. In particular this also shows that d must be the *greatest* common divisor—indeed if d' were a common divisor that was greater than d, then we would have $d' \mid d$, which is a contradiction.[2] ■

Corollary 1.2.3. *An integer r can be written as $r = as+bt$, if and only if $gcd(a,b) \mid r$.*

Proof. As $gcd(a,b) = as + bt$ for some s, t, it is obvious that any multiple of the gcd can also be written as a linear combination of a and b. Conversely, any linear combination of a and b is divisible by any common divisor of a and b, and in particular by the gcd. ■

Remark 1.2.4. Another is that when a, b are not simultaneously 0, the gcd can be seen as the smallest positive integer that can be written as a linear combination of a and b. But note that said decomposition is not unique. For example:

$$\begin{aligned} gcd(a,b) &= ax + by \\ &= ax - ab + by + ba \\ &= a(x-b) + b(y+a) \end{aligned}$$

Intuitively this can be understood by seeing $gcd(a,b) = ax + by$ as a straight line in $\mathbb{R}^2$ (on variables x and y), which has an infinite number of solutions which are integer pairs. △

Definition 1.2.5. *Two integers are said to be* **relatively prime** *if their gcd is* 1.

From the previous theorem we immediately get the following property:

Lemma 1.2.6. *If a and b are integers, then they are relatively prime if and only if $xa + yb = 1$, for some integers x and y.*

Proof. If $gcd(a,b) = 1$, then by lemma 1.2.2, $xa + yb = 1$, for some $x, y \in \mathbb{Z}$. Conversely if $xa+yb = 1$, then any common divisor of a and b must divide 1, which implies that the only common divisor of a and b is 1—and so $gcd(a,b) = 1$. ■

Corollary 1.2.7. *Two integers are relatively prime (gcd = 1) if and only if they have no common prime factors.*

Proof. (→) If $gcd = 1$, then we can write (call the integers a and b) $ax + by = 1$; thus, any common factor of a and b must also divide 1, and so 1 is the only common factor—and 1 has no prime factors. (←) If a and b have no common prime factors, then they cannot have any common factors (divisors) at all, other than 1, because otherwise such hypothetical common factors would have a nonempty prime factorisation, which would yield common prime factors. Thus, the gcd is 1. ■

[2] In my view this already shows the gcd to be *unique*, for no set can contain two distinct greatest elements. However, the gcd's uniqueness can also be shown explicitly: let d' now be another gcd. We would necessarily have $d \mid d'$ and $d' \mid d$, and as the gcd is always non-negative by definition, we conclude that $d = d'$.

As 0 has no nonzero factors, it follows that every nonzero integer must be relatively prime to zero—for otherwise there would exist common prime factors, which is absurd.

Another property, that is useful to understand the existence of inverses of residue classes, is that if $gcd(a,n) = 1$, then also $gcd(a+kn,n) = 1$, for $k \in \mathbb{Z}$. Indeed from lemma 1.2.2, let $ax + ny = 1$. Then we have $ax + ny + knx - nkx = 1 \Leftrightarrow (a+kn)x + n(y-kx) = 1 \Leftrightarrow gcd(a+kn,n) = 1$.

So if residue class $[a]_n$ has a multiplicative inverse, then we would expect that $gcd(a,n) = 1$—and the above result shows the same conclusion holds regardless of the representative of that class that is chosen.

The next theorem is a simple albeit not obvious result.

Theorem 1.2.8. *Let $n_1, \ldots, n_k$ be a family of integers, and let $n = \prod n_i$. For an integer a, $gcd(a,n) = 1$ if and only if $gcd(a,n_i) = 1$.*

Proof. ($\rightarrow$) If $gcd(a,n) = 1$, then $ax+ny = 1 \Leftrightarrow ax+(\prod n_i)\, y = 1$ from where we conclude that for each i we can write $ax+n_iy' = 1$, which entails that $gcd(a,n_i) = 1$.

($\leftarrow$) Let $ax_1 + n_1y_1 = 1$ and $ax_2 + n_2y_2 = 1$. Multiply one by the other; we obtain:

$$a^2x_1x_2 + ax_1n_2y_2 + n_1y_1ax_2 + n_1y_1n_2y_2 = ax' + n_1n_2y' = 1^2 = 1 \qquad (1.4)$$

If we now have that $ax_3 + n_3y_3 = 1$, then multiplying member-wise by $ax' + n_1n_2y' = 1$ will yield $ax'' + n_1n_2n_3y'' = 1$. Now suppose that we have shown that $ar + (n_1n_2\ldots n_j)s = 1$, for numbers $n_1, n_2, \ldots,$ up to n_j (for some integers r,s). Now as $gcd(a,n_{j+1}) = 1$, there are r',s' such that $ar' + n_{j+1}s' = 1$. Doing sidewise multiplication, as in (1.4), will now yield $ar'' + (n_1n_2\ldots n_jn_{j+1})s'' = 1$, for some r'',s'', which, by induction, shows the result. ∎

1.2.1 Formulas for gcd and lcm

Suppose we have two positive integers a and b, defined as follows:[3]

$$a = p_1^{a_1}p_2^{a_2}\cdots = \prod p_i^{a_i} \qquad \text{and} \qquad b = p_1^{b_1}p_2^{b_2}\cdots = \prod p_i^{b_i} \qquad (1.5)$$

(the exponents are zero when not needed, so even though i ranges over all of $\mathbb{N}$, both products only involve a finite number of terms different than 1).

We make the following observation: $a \mid b$ if and only if $a_i \leqslant b_i$, for all i. We can then derive the following two formulas, where both products take place over the set of all primes (they also hold for more than two integers):

$$gcd(a,b) = \prod p_i^{g_i}\ , \text{where } g_i \stackrel{\text{def}}{=} min(a_i,b_i) \qquad (1.6)$$

$$lcm(a,b) = \prod p_i^{l_i}\ , \text{where } l_i \stackrel{\text{def}}{=} max(a_i,b_i) \qquad (1.7)$$

It is clear that $gcd(a,b)$ as defined above is a common divisor of both a and b. Consider how could we increase it: because of existence and uniqueness of prime

[3] We deal only with positive integers because the negative one is identical, and if at least one of the integers is zero, then their lcm is also zero. The gcd for nonpositive integers was explained in definiton 1.2.1.

factorisation, the only way would be to increase some exponents g_i in (1.6). But this would mean that at least one of the conditions $g_i \leqslant a_i$ or $g_i \leqslant b_i$ would be violated for some i—accordingly entailing that we would no longer be dealing with a common divisor of a and b.

As for the lcm, via a similar reasoning as above, the only way to decrease the lcm is to decrease some of the exponents l_i in (1.7). But this would mean that at least one of the conditions $a_i \leqslant l_i$ or $b_i \leqslant l_i$ would be violated for some i—accordingly entailing that we would no longer be dealing with a common multiple of a and b.

Remark 1.2.9. If two numbers a and b are relatively prime, then their lcm is their product, i.e. ab. Also, every other common multiple is also a multiple of the lcm, for if any of the prime factors of this common multiple had a exponent smaller than the maximum (cf. equation 1.7), it could not be a common multiple. A special case is if two numbers n_1 and n_2 are relatively prime; then any common multiple of both is also a multiple of $n_1 n_2$.[I] △

1.3 Congruences

Two integers a and b are said to be *congruent modulo* n if $n \mid (a-b)$. This is usually denoted as $a \equiv b \pmod{n}$—the indication of the module can be omitted when clear from context. If $a \equiv a'$ and $b \equiv b'$, simple algebraic manipulations show that $a+b \equiv a'+b'$ and $ab \equiv a'b'$. Furthermore, for all a we have that

$$\Big[a \equiv (a \bmod n)\Big] \pmod{n} \tag{1.8}$$

holds. This allows us to simplify computing the remainder of very large numbers. Indeed, we have that:

$$(a+b) \pmod{n} \equiv a+b \equiv (a \bmod n) + (b \bmod n) \equiv \\ [(a \bmod n) + (b \bmod n)] \pmod{n}$$

And similarly,

$$(ab) \pmod{n} \equiv ab \equiv (a \bmod n)(b \bmod n) \equiv \\ [(a \bmod n)(b \bmod n)] \pmod{n}$$

These properties also show that when there are more than two factors, we can do things "piecewise". I.e., take the modulus as we go along the sum:

$$(a+b+c) \pmod{n} \equiv (a+b)+c \equiv \big[(a+b) \pmod{n}\big] + c \\ \equiv \Big\{\big[(a+b) \pmod{n}\big] + c\Big\} \pmod{n}$$

$$(abc) \pmod{n} \equiv (ab)c \equiv \big[(ab) \pmod{n}\big]c \\ \equiv \Big\{\big[(ab) \pmod{n}\big]c\Big\} \pmod{n}$$

Thus, whenever we have (to compute the remainder of) an expression that consists of sums of products, we can just compute the remainder of all parcels, and then the remainder of the full expression. A typical example is the rule to "cast out nines": as any integer can be written in the form $\sum d_i 10^i, 0 \leqslant d_i \leqslant 9$, and as $10 \equiv 1 \pmod 9$, to compute the remainder of the division of that integer by 9, we just sum the digits, casting out nines wherever possible—which is a lot simpler than remaindering over the whole integer.

Theorem 1.3.1. *$a \bmod n = b \bmod n$ if and only if $a \equiv b \pmod n$.*

Proof. ($\rightarrow$) $a \bmod n = b \bmod n \Leftrightarrow a - n\lfloor a/n \rfloor = b - n\lfloor b/n \rfloor \Leftrightarrow n \mid (a-b) \Leftrightarrow a \equiv b \pmod n$.

($\leftarrow$) $a \equiv b \pmod n \Leftrightarrow (a \bmod n + n\lfloor a/n \rfloor) - (b \bmod n + n\lfloor b/n \rfloor) = kn \Leftrightarrow a \bmod n - b \bmod n = n(k - \lfloor a/n \rfloor + \lfloor b/n \rfloor)$. This shows $a \bmod n \equiv b \bmod n \pmod n$. But as both $a \bmod n$ and $b \bmod n$ belong to $[0, n[$, their difference belongs to $]-n, n[$—and the only multiple of n in this range is 0. Hence $a \bmod n = b \bmod n$. ■

1.4 Number representation

The previous section mentioned the base-10 representation system, the well-known *decimal* system. Well, there is nothing special about the number 10. Let b be any positive integer; we can use it a basis for number representation, i.e. we can take any positive integer a and write it in the form:

$$a = r_n b^n + r_{n-1} b^{n-1} + \cdots + r_1 b + r_0, \text{ with } r_n \neq 0 \tag{1.9}$$

(the reason for labeling the coefficients r_i will become clear shortly). If $a < b$, the representation of a in the basis b is just a itself, so let's assume that $a \geqslant b$. Doing integer division, we get $a = q_0 b + r_0$. If $q_0 < b$ we are done. Otherwise we do integer division on q_0, to get $q_0 = q_1 b + r_1$, and plug this in the expression of the previous integer division, to get:

$$a = (q_1 b + r_1) b + r_0 = q_1 b^2 + r_1 b + r_0 \tag{1.10}$$

Note the general pattern is that q_i and r_i are produced in the $(i+1)$th division, which is $q_{i-1} = q_i b + r_i$. (According to this notation, $a = q_{-1}$.)

So suppose that after the k-th division (which yields q_{k-1} and r_{k-1}), the quotient obtained (q_{k-1}) finally drops below b. Then our representation of a would look something like this:

$$a = q_{k-1} b^k + r_{k-1} b^{k-1} + \cdots + r_1 b + r_0 \tag{1.11}$$

Observe that, while the coefficient obtained is greater than b, each new division introduces a new b factor; put another way, the number of divisions performed so far equals the exponent of the highest power of b. Also note that we cannot have $q_{k-1} = 0$, because as $q_{k-2} = q_{k-1} b + r_{k-1}$, this would mean that $q_{k-2} = r_{k-1}$, i.e. that the previous highest degree coefficient, q_{k-2}, had already dropped below b (and thus the algorithm would have stopped).

If $0 < q_{k-1} < b$, then as $q_{k-1} = q_k b + r_k$, we conclude that $q_{k-1} = r_k$—but one does not need to actually compute this division (so it is not counted in the total number of divisions required). (1.11) then becomes:

$$a = r_k b^k + r_{k-1} b^{k-1} + \cdots + r_1 b + r_0 \tag{1.12}$$

This is representation of a in the basis b. It verifies the following bound (remember that $r_k (= q_{k-1})$ cannot be 0):

$$b^k \leqslant a < b^{k+1} \tag{1.13}$$

The first inequality comes from setting $r_k = 1$ and the remaining r_i to 0; as for second one, set *all* the r_i to $b-1$, and then add 1. We have:

$$\begin{aligned}
&(b-1)b^k + (b-1)b^{k-1} + \cdots + (b-1)b + (b-1) + 1 \\
&= (b-1)b^k + (b-1)b^{k-1} + \cdots + (b-1)b + b \\
&= (b-1)b^k + (b-1)b^{k-1} + \cdots + (b-1+1)b \\
&= (b-1)b^k + (b-1)b^{k-1} + \cdots + b^2 \\
&= \qquad\qquad \cdots \\
&= (b-1)b^k + (b-1)b^{k-1} + b^{k-1} \\
&= (b-1)b^k + (b-1+1)b^{k-1} \\
&= (b-1)b^k + b^k = bb^k = b^{k+1}
\end{aligned}$$

Equation (1.13) tells us that in the representation of a in basis b, the highest power of b with a nonzero coefficient is b^k, and as explained above, this requires k divisions. From the double inequality now comes $k = \lfloor \log_b a \rfloor$. Also the number of digits (i.e. length) of that representation is $\lfloor \log_b a \rfloor + 1$.

This result might be more easily memorised if thought of as follows: to get the representation of a number a in base b, we need 0 divisions if $b^0 \leqslant a < b^1$, 1 division if $b^1 \leqslant a < b^2$, ..., k divisions if $b^k \leqslant a < b^{k+1}$. And the number of digits is one plus the number of divisions.

1.5 The Chinese Remainder Theorem

Consider the following problem: given a family of integers $n_1, \ldots, n_k$, all pairwise relatively prime, and another family of integers $a_1, \ldots, a_k$, we want to find an integer a such that $a \equiv a_i \pmod{n_i}$, for all i. The way we do this is by finding a family of numbers $e_1, \ldots, e_k$, with the property that $e_i \equiv 1 \pmod{n_i}$, and $e_j \equiv 0 \pmod{n_i}$, for all $j \neq i$. Then it is straightforward to see that the number

$$a = \sum_i a_i e_i \tag{1.14}$$

solves the set of linear congruences. Indeed, modulo n_i we have:

$$a = a_i e_i + \sum_{j \neq i} a_j e_j \equiv a_i 1 + \sum_{j \neq i} a_j 0 = a_i \tag{1.15}$$

To construct the e_i, let $n = n_1 n_2 \dots n_k$. Then $e_i = (n/n_i)(n/n_i)^{-1}$, where $(n/n_i)^{-1}$ denotes the modular inverse of n/n_i, the modulus being n_i.[4] Note this inverse always exists, as $gcd(n/n_i, n_i) = 1$, due to the fact that all the n_i are pairwise prime.

Now let $a' \equiv a \pmod{n}$. This means $n \mid (a - a')$, and as $n_i \mid n$, for all i, then also $n_i \mid (a - a')$, and thus $a' \equiv a \pmod{n_i}$. And finally, as we also have that $a \equiv a_i \pmod{n_i}$, we conclude that also $a' \equiv a_i \pmod{n_i}$, i.e. that a' is also a solution to the set of congruences.

Conversely, suppose now that a' is a solution to the set of congruences, i.e. that $a' \equiv a_i \pmod{n_i}$ for all i. As we also have that $a \equiv a_i \pmod{n_i}$, we conclude that $a' \equiv a \pmod{n_i}$, again for all i. This equivalent to saying that $n_i \mid (a - a')$; and as this holds for all i, then we must also have that $lcm(n_1, \dots, n_k) \mid (a - a')$. As $lcm(n_1, \dots, n_k) = n$, due to the n_i being all pairwise relatively prime, we conclude that $n \mid (a - a')$, or equivalently, $a' \equiv a \pmod{n}$.

So to sum up, given a solution a to the set of congruences, any other integer a' is also a solution if and only if $a \equiv a' \pmod{n}$ (where $n = n_1 \dots n_k$).

The CRT has a very neat interpretation in terms of residue classes [1, §2.5]. Indeed, suppose as above that a is a solution to the CRT congruences. Then we just saw that any other element of the residue class $[a]_n$ of $\mathbb{Z}_n$ is also a solution.[5] Furthermore, consider any one of the a_i; as $a \equiv a_i \pmod{n_i}$, we see that a belongs to the residue $[a_i]_{n_i}$, which thus coincides with the residue $[a]_{n_i}$. Hence the CRT be seen as a mapping from $\mathbb{Z}_n$ to $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \dots \times \mathbb{Z}_{n_k}$, as follows: $[a]_n \mapsto ([a]_{n_1} \times [a]_{n_2} \times \dots \times [a]_{n_k})$. As any number congruent modulo n with a solution to the CRT is also a solution, we see that the mapping is well-defined (i.e., it does not depend on the particular element of the residue class). The mapping is also a bijection, which we can show as follows: suppose there were two distinct residues, say $[a]_n$ and $[a']_n$ that got mapped to the same residue tuple. Then for any n_i, we must have that $[a]_{n_i}$ and $[a']_{n_i}$ are the same residue, which of course means that $a \equiv a' \pmod{n_i}$, for all n_i. As shown above, this implies that $a \equiv a' \pmod{n}$, which is the same saying that $[a]_n$ and $[a']_n$ are the same residue. This shows the mapping is injective, the CRT as formulated above shows that for any tuple of residues of the $\mathbb{Z}_{n_i}$, the exists a corresponding residue of $\mathbb{Z}_n$, which shows the mapping is surjective—and hence, bijective.

[4] **But note that you cannot reduce n/n_i modulo n_i!!** Otherwise it will no longer be a multiple of all the other n_j, with $j \neq i$!

[5] Recall that $[a]_n$ is composed of all integers b such that $b \equiv a \pmod{n}$, and that $\mathbb{Z}_n$ is the set of all such class, viz. $[0]_n, [1]_n, \dots, [n-1]_n$.

Chapter 2

Towards RSA

2.1 The theorems of Fermat and Euler

We begin with a combinatorial proof of Fermat's (so-called) little theorem.[1] As the proof (due to its combinatorial nature) applies only to positive integers, we begin by showing that if the theorem holds for positive integers, it holds for any integers.

Lemma 2.1.1. *Assume that for any positive integer a, and any prime p, we have that $a^p \equiv a \pmod{p}$. Then this holds for negative integers as well.*

Proof. So assume that a is nonzero, and in particular, that it is positive, and so that the modular equivalence above holds. This means there exists k such that $a^p - a = pk$. Multiplying by -1 and assuming p is odd:

$$\begin{aligned} & -a^p - (-a) = p(-k) && (2.1)\\ \Leftrightarrow\ & (-1)^p a^p - (-a) = p(-k) && (p \text{ is odd})\\ \Leftrightarrow\ & (-a)^p - (-a) = p(-k) && (2.2) \end{aligned}$$

This shows the modular equality holds for negative a's, if p is odd. If p is even, meaning $p = 2$, we get:

$$\begin{aligned} & a^2 - a = 2k && (\text{definition of congruence})\\ \Leftrightarrow\ & -a^2 - (-a) = 2(-k) && (\text{multiply by } -1)\\ \Leftrightarrow\ & 2a^2 - a^2 - (-a) = 2(-k) + 2a^2 && (\text{add } 2a^2 \text{ to both members})\\ \Leftrightarrow\ & a^2 - (-a) = 2(a^2 - n) && (\text{simplify})\\ \Leftrightarrow\ & (-a)^2 - (-a) = 2(a^2 - n) && (\text{as } (-1)^2 = 1)\\ \Leftrightarrow\ & a^2 \equiv a \pmod{2} && (\text{definition of congruence}) \end{aligned}$$

Thus the equality holds, for negative integers, also in the even exponent case, which concludes the proof. ■

Theorem 2.1.2. *For any integer a and prime p, $a^p \equiv a \pmod{p}$.*

[1] This proof is a modified version of the one that can be found in G. E. Andrews, *Number Theory*, Dover, New York, 1994, §3.2.

Proof. Start by noting that if $a = 0$, the theorem is obvious. Let us now prove that the theorem holds for any positive a. We can think of a^p as the number of strings of length p that can be formed with an alphabet with a symbols: indeed we have a choices for the first position, a choices for the second, and so on, and finally a choices for the p-th position. In this lot, there are exactly a strings that are have the same symbol in all positions; removing them we are left with $a^p - a$ strings. On this reduced set we define the following equivalence relation: imagine the strings lay horizontally; we say two strings a and b are related if we can take string a, take off its rightmost element, re-place it as its leftmost element, and by repeating this operation a finite number of times, obtain string b. If we repeat the operation p times, we get the original string back, so the relation is reflexive. It is also symmetric, because if we can go from string a to string b, then as the operation loops around, we can also go from string b to string a. And it is transitive: if we can shift string from string a to string b, and from this to string c, then of course we can shift directly from string a to string c. So this right circular shift is indeed an equivalence relation.

The next step is to show that each of the equivalence classes induced by this relation have exactly p elements. As equivalence classes are pairwise disjoint, and form a partition of the original set of $a^p - a$ strings, this immediately yields that $a^p - a$ is a multiple of p, which proves the theorem for a positive a.

To show that each class has exactly p elements, first notice that as each string has at least two different symbols, there can be no repetitions until we get back the original element (i.e. until we "loop around"): starting with the original string, each subsequent shift produces a string that is different from all the predecessors—until we get the original one back. So saying that a class has p elements is tantamount to saying that the *minimum* number of shifts required to loop around is exactly p. Suppose this was *not* the case; i.e. suppose there existed $k < p$ such that k shifts yielded the original string back (as there are at least two symbols, it must be $k > 1$, for a shift of one never leaves the string unchanged[2]). Dividing p by k we get $p = kl + r$, with $r < k$. Now if k shifts get us back to the original string, so do kl shifts; and as p shifts also give leave back at the original, it must be the case that $r = p - kl$ shifts also do the same. But $r < k$, and k is the smallest number of shifts to loop around, so $r = 0$. So $p = kl$, but $k > 1$ and p is a prime, so it must be $l = 1$ and $p = k$. Thus the minimum number of shifts to loop around is indeed p—and hence each equivalence class has indeed p elements.

As discussed above, this shows the theorem holds for a positive a; and from lemma 2.1.1 we conclude the theorem also holds for any negative a, concluding the proof. ■

[2] This would happen for a string consisting of the repetition of just one symbol, which we excluded at the beginning of the proof.

Notes

Chapter 1 – Modular arithmetic

I. This generalises to products with more than two factors: if they are all relatively prime, then any common multiple is also a multiple of their product. To see this, let n_1, n_2 and n_3 be all (pairwise) relatively prime. Let n be a common multiple of the three; as n is, in particular, a common multiple of n_1 and n_2, from the above result we know that it will also be a multiple of $n_1 n_2$. But $gcd(n_1 n_2, n_3)$ must be 1, otherwise n_3 would have a common prime factor with either n_1 or n_2 (can't be with both, as they are relatively prime). Hence, as n is also a multiple of n_3, another application of the previous result shows that it is also a multiple of $n_1 n_2 n_3$. An inductive generalisation now suggests itself.

References

1. Victor Shoup. *A Computational Introduction to Number Theory and Algebra*. Cambridge University Press, New York, NY, USA, 2nd edition, 2009. Cited on page 9.