

§6 Integers Modulo n

In Example 2.3(e), we have defined the congruence of two integers a, b with respect to a modulus $n \in \mathbb{N}$. Let us recall that $a \equiv b \pmod{n}$ means $n \mid a - b$. We have proved that congruence is an equivalence relation on \mathbb{Z} . The equivalence classes are called the *congruence classes* or *residue classes* (modulo n). The congruence class of $a \in \mathbb{Z}$ will be denoted by \bar{a} . Notice that there is ambiguity in this notation, for there is no reference to the modulus. Thus $\bar{1}$ represents the residue class of 1 with respect to the modulus 1, also with respect to the modulus 2, also with respect to the modulus 3, in fact with respect to any modulus. However, the modulus will be usually fixed throughout a particular discussion and \bar{a} will represent the residue class of a with respect to that fixed modulus. The ambiguity is therefore harmless.

By the division algorithm (Theorem 5.3), any integer k can be written as $k = qn + r$, with $q, r \in \mathbb{Z}$, $0 \leq r < n$. So any integer k is congruent (mod n) to one of the numbers $0, 1, 2, \dots, n-1$. Furthermore, no two distinct of the numbers $0, 1, 2, \dots, n-1$ are congruent (mod n), for if $r_1, r_2 \in \{0, 1, 2, \dots, n-1\}$ and $r_1 \equiv r_2 \pmod{n}$, then $n \mid r_1 - r_2$, so $n \leq |r_1 - r_2|$ by Lemma 5.2(11), and so $n \leq (n-1) - 0$, which is impossible. Thus any integer is congruent to one of the numbers $0, 1, 2, \dots, n-1$, and these numbers are pairwise incongruent. This means that $0, 1, 2, \dots, n-1$ are the representatives of all the residue classes. Hence there are exactly n residue classes (mod n), namely

$$\begin{aligned} \bar{0} &= \{x \in \mathbb{Z} : x \equiv 0 \pmod{n}\} &= \{nz \in \mathbb{Z} : z \in \mathbb{Z}\} &=: n\mathbb{Z} \\ \bar{1} &= \{x \in \mathbb{Z} : x \equiv 1 \pmod{n}\} &= \{nz + 1 \in \mathbb{Z} : z \in \mathbb{Z}\} &=: n\mathbb{Z} + 1 \\ \bar{2} &= \{x \in \mathbb{Z} : x \equiv 2 \pmod{n}\} &= \{nz + 2 \in \mathbb{Z} : z \in \mathbb{Z}\} &=: n\mathbb{Z} + 2 \end{aligned}$$

.....

$$\overline{n-1} = \{x \in \mathbb{Z} : x \equiv n-1 \pmod{n}\} = \{nz + (n-1) \in \mathbb{Z} : z \in \mathbb{Z}\} =: n\mathbb{Z} + (n-1).$$

The set $\{\bar{0}, \bar{1}, \bar{2}, \dots, \overline{n-1}\}$ of residue classes (mod n) will be denoted by \mathbb{Z}_n . An element of \mathbb{Z}_n , that is, a residue class (mod n) is called an *integer modulo n* , or an *integer mod n* . An integer mod n is not an integer, not

an element of \mathbb{Z} ; it is a subset of \mathbb{Z} . An integer mod n is not an integer with a property "mod n ". It is an object whose name consists of the three words "integer", "mod(ulo)", " n ".

6.1 Lemma: *Let $n \in \mathbb{N}$, $a, a_1, b, b_1 \in \mathbb{Z}$. If $a \equiv a_1 \pmod{n}$ and $b \equiv b_1 \pmod{n}$, then $a + b \equiv a_1 + b_1 \pmod{n}$ and $ab \equiv a_1 b_1 \pmod{n}$.*

Proof: If $a \equiv a_1 \pmod{n}$ and $b \equiv b_1 \pmod{n}$, then $n \mid a - a_1$ and $n \mid b - b_1$. Hence $n \mid (a - a_1) + (b - b_1)$ by Lemma 5.2(5), which gives $n \mid (a + b) - (a_1 + b_1)$, so $a + b \equiv a_1 + b_1 \pmod{n}$. Also, $n \mid b(a - a_1) + a_1(b - b_1)$ by Lemma 5.2(7), which gives $n \mid ba - a_1 b_1$, so $ab \equiv a_1 b_1 \pmod{n}$. \square

We want to define a kind of addition \oplus and a kind of multiplication \otimes on \mathbb{Z}_n . We put

$$\overline{a} \oplus \overline{b} = \overline{a + b} \quad (*)$$

$$\overline{a} \otimes \overline{b} = \overline{ab} \quad (**)$$

for all $\overline{a}, \overline{b} \in \mathbb{Z}_n$ (for all $a, b \in \mathbb{Z}$). This is a very natural way of introducing addition and multiplication on \mathbb{Z}_n .

(*) and (**) seem quite innocent, but we must check that \oplus and \otimes are really binary operations on \mathbb{Z}_n . The reader might say at this point that \oplus and \otimes are clearly defined on \mathbb{Z}_n and that there is nothing to check. But yes, there is. Let us remember that a binary operation on \mathbb{Z}_n is a function from $\mathbb{Z}_n \times \mathbb{Z}_n$ into \mathbb{Z}_n (Definition 3.18). As such, to each pair $(\overline{a}, \overline{b})$ in $\mathbb{Z}_n \times \mathbb{Z}_n$, there must correspond a *single* element $\overline{a} \oplus \overline{b}$ and $\overline{a} \otimes \overline{b}$ if \oplus and \otimes are to be binary operations on \mathbb{Z}_n (Definition 3.1) We must check that the rules (*) and (**) produce elements of \mathbb{Z}_n that are uniquely determined by \overline{a} and \overline{b} .

The rules (*) and (**) above convey the wrong impression that $\overline{a} \oplus \overline{b}$ and $\overline{a} \otimes \overline{b}$ are uniquely determined by \overline{a} and \overline{b} . In order to penetrate into the matter, let us try to evaluate $X \oplus Y$, where $X, Y \in \mathbb{Z}_n$ are not given directly as the residue classes of integers $a, b \in \mathbb{Z}$. (We discuss \oplus ; the discussion applies equally well to \otimes .) How do we find $X \oplus Y$? Since $X, Y \in \mathbb{Z}_n$, there are integers $a, b \in \mathbb{Z}$ with $\overline{a} = X$, $\overline{b} = Y$. Now add a and b in \mathbb{Z} to get $a+b \in \mathbb{Z}$, then take the residue class of $a+b$. The result is $X \oplus Y$.

The result? The question is whether we have only *one* result to justify the article "the". We summarize telegraphically. To find $X \oplus Y$,

- 1) choose $a \in \mathbb{Z}$ from X ,
- 2) choose $b \in \mathbb{Z}$ from Y ,
- 3) find $a + b$ in \mathbb{Z} ,
- 4) take the residue class of $a + b$.

This sounds a perfectly good recipe for finding $X \oplus Y$, but notice that we use some auxiliary objects, namely a and b , to find $X \oplus Y$, which must be determined by X and Y alone. Indeed, the result $\overline{a + b}$ depends explicitly on the auxiliary objects a and b . We can use our recipe with different auxiliary objects. Let us do it. 1) I choose a from $X \subseteq \mathbb{Z}$ and you choose a_1 from X . 2) I choose b from $Y \subseteq \mathbb{Z}$ and you choose b_1 from Y . 3) I compute $a + b$ and you compute $a_1 + b_1$. In general, $a + b \neq a_1 + b_1$. Hence our recipe gives, generally speaking, distinct elements $a + b$ and $a_1 + b_1$. So far, both of us followed the same recipe. I cannot claim that my computation is correct and yours is false. Nor can you claim the contrary. Now we carry out the fourth step. I find the residue class of $a + b$ as $X \oplus Y$, and you find the residue class of $a_1 + b_1$ as $X \oplus Y$. Since $a + b \neq a_1 + b_1$ in \mathbb{Z} , it can very well happen that $\overline{a + b} \neq \overline{a_1 + b_1}$ in \mathbb{Z}_n . On the other hand, if \oplus is to be a binary operation on \mathbb{Z}_n , we must have $\overline{a + b} = \overline{a_1 + b_1}$. This is the central issue. In order that \oplus be a binary operation on \mathbb{Z}_n , there must work a mechanism which ensures $\overline{a + b} = \overline{a_1 + b_1}$ whenever $\overline{a} = \overline{a_1}$, $\overline{b} = \overline{b_1}$, even if $a + b \neq a_1 + b_1$. If there is such a mechanism, we say \oplus is a well defined operation on \mathbb{Z}_n . This means \oplus is really a genuine operation on \mathbb{Z}_n : $X \oplus Y$ is uniquely determined by X and Y alone. Any dependence of $X \oplus Y$ on auxiliary integers $a \in X$ and $b \in Y$ is only apparent. We will prove that \oplus and \otimes are well defined operations on \mathbb{Z}_n , but before that, we discuss more generally well definition of functions.

A function $f: A \rightarrow B$ is essentially a rule by which each element a of A is associated with a unique element of $f(a) = b$ of B . The important point is that the rule produces an element $f(a)$ that depends only on a . Sometimes we consider rules having the following form. To find $f(a)$,

- 1) do this and that
- 2) take an x related to a in such and such manner
- 3) do this and that to x
- 4) the result is $f(a)$.

A rule of this type uses an auxiliary object x . The result then depends on a and x . At least, it seems so. This is due to the ambiguity in the second step. This step states that we choose an x with such and such property, but there may be many objects x, y, z, \dots related to a in the prescribed manner. The auxiliary objects x, y, z, \dots will, in general, produce different results, so we should perhaps that the result is $f(a, x)$ (or $f(a, y), f(a, z), \dots$). In order the above rule to be a function, it must produce the same result. Hence we must have $f(a, x) = f(a, y) = f(a, z) = \dots$. The rule must be so constructed that the same result will obtain even if we use different auxiliary objects. If this be the case, the function is said to be *well defined*.

This terminology is somewhat unfortunate. It sounds as though there are two types of functions, well defined functions and not well defined functions (or badly defined functions). This is definitely not the case. A well defined function is simply a function. Badly defined functions do not exist. Being well defined is not a property, such as continuity, boundedness, differentiability, integrability etc. that a function might or might not possess. That a function $f: A \rightarrow B$ is well defined means: 1) the rule of evaluating $f(a)$ for $a \in A$ makes use of auxiliary, foreign objects, 2) there are many choices of these foreign objects, hence 3) we have reason to suspect that applying the rule with different choices may produce different results, which would imply that our rule does not determine $f(a)$ uniquely and f is not a function in the sense of Definition 3.1, but 4) our suspicion is not justified, for there is a mechanism, hidden under the rule, which ensures that same result will obtain even if we apply the rule with different auxiliary objects. The question as to whether a "function" is well defined arises only if that "function" uses objects not uniquely determined by the element a in its "domain" in order to evaluate $f(a)$. We wrote "function" in quotation marks, for such a thing may not be a function in the sense of Definition 3.1. Given such a "function", which we want to be a function in the sense of Definition 3.1, we check whether $f(a)$ is uniquely determined by a , that is, we check whether $f(a)$ is independent of the auxiliary objects that we use for evaluating $f(a)$. If this be the case, our supposed "function" f is indeed a function in the sense of Definition 3.1. We say then that f is well defined, or f is a well defined function. This means f is a function. In fact, it is more accurate to say that a function is defined instead of saying that a function is well defined.

6.2 Examples: (a) Let L be the set of all straight lines in the Euclidean plane, on which we have a cartesian coordinate system. We consider the "function" $s: L \rightarrow \mathbb{R} \cup \{\infty\}$, which assigns the slope of the line l to l . How do we find $s(l)$? As follows: 1) choose a point, say (x_1, y_1) , on l ; 2) choose another point, say (x_2, y_2) , on l ; 3) evaluate $x_2 - x_1$ and $y_2 - y_1$; 4) put $s(l) = (y_2 - y_1)/(x_2 - x_1)$ if $x_1 \neq x_2$ and $s(l) = \infty$ if $x_1 = x_2$. Clearly we can choose the points in many ways. For example, we might choose $(x_1', y_1') \neq (x_1, y_1)$ as the first point, $(x_2', y_2') \neq (x_2, y_2)$ as the second point. Then we have, in general, $x_2' - x_1' \neq x_2 - x_1$ and $y_2' - y_1' \neq y_2 - y_1$, so we might suspect that $(y_2' - y_1')/(x_2' - x_1') \neq (y_2 - y_1)/(x_2 - x_1)$. It is known from analytic geometry that these two quotients are equal, hence $s(l)$ depends only on l , and not on the points we choose. Thus s is a well defined function. Ultimately, this is due to the fact that there passes one and only one straight line through two distinct points. The next example shows that well definition breaks down if we modify the domain a little.

(b) Let C be the set of all curves in the Euclidean plane. We consider the "function" $s: C \rightarrow \mathbb{R} \cup \{\infty\}$, which assigns the "slope" of the curve c to c . How do we find $s(c)$? As follows: 1) choose a point, say (x_1, y_1) , on l ; 2) choose another point, say (x_2, y_2) , on l ; 3) evaluate $x_2 - x_1$ and $y_2 - y_1$; 4) put $s(l) = (y_2 - y_1)/(x_2 - x_1)$ if $x_1 \neq x_2$ and $s(l) = \infty$ if $x_1 = x_2$. This is the same rule as the rule in Example 6.2(a). Let us find the "slope" of the curve $y = x^2$. 1) Choose a point on this curve, for example $(0,0)$. If you prefer, you might choose $(-1,1)$. 2) Choose another point on this curve, for example $(1,1)$. If you prefer, you might choose $(2,4)$ of course. 3) Evaluate the differences of coordinates. We find $1 - 0$ and $1 - 0$. You find $2 - (-1)$ and $4 - 1$. Hence 4) the slope is $2/1$. You find it to be $3/3$. So $s(c) = 2$ and $s(c) = 1$. This is nonsense. We see that different choices of the points on the curve (different choices of the auxiliary objects) give rise to different results. So the above rule is not a function. We do not say " s is not a well defined function". s is simply not a function at all. s is not defined.

(c) Let F be the set of all continuous functions on a closed interval $[a,b]$. We want to "define" an integral "function" $I: F \rightarrow \mathbb{R}$, which assigns the real number $\int_a^b f(x)dx$ to $f \in F$. So $I(f) = \int_a^b f(x)dx$. I is a "function" whose "domain" is a set of functions. How do we find $I(f)$? As follows. 1) Choose an indefinite integral of f , that is, choose a function F on $[a,b]$ such that $F'(x) = f(x)$ for all $x \in [a,b]$ (we take one-sided derivatives at a

and b). 2) Evaluate $F(a)$ and $F(b)$. 3) Put $I(f) = F(b) - F(a)$. There are many functions F with $F'(x) = f(x)$ for all $x \in [a, b]$. For two different choices F_1 and F_2 , we have $F_1(b) \neq F_2(b)$ and $F_1(a) \neq F_2(a)$ in general. So we may suspect that $F_1(b) - F_1(a) \neq F_2(b) - F_2(a)$. In order to show that I is a well defined function, we must prove $F_1(b) - F_1(a) = F_2(b) - F_2(a)$ whenever F_1 and F_2 are functions on $[a, b]$ such that $F_1'(x) = f(x) = F_2'(x)$ for all $x \in [a, b]$. We know from the calculus that, when F_1 and F_2 have this property, there is a constant c such that $F_1(x) = F_2(x) + c$ for all $x \in [a, b]$. So $F_1(b) - F_1(a) = (F_2(b) + c) - (F_2(a) + c) = F_2(b) - F_2(a)$. Therefore, I is well defined.

After this lengthy digression, we return to the integers mod n and to the "operations" \oplus and \otimes .

6.3 Lemma: \oplus and \otimes are well defined operations on \mathbb{Z}_n .

Proof: We are to prove $\overline{a} \oplus \overline{b} = \overline{a'} \oplus \overline{b'}$ and $\overline{a} \otimes \overline{b} = \overline{a'} \otimes \overline{b'}$ whenever $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$ in \mathbb{Z}_n (different names for identical residue classes should not yield different results). This follows from Lemma 6.1. Indeed, if $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$, then $\overline{a} \equiv \overline{a'} \pmod{n}$ and $\overline{b} \equiv \overline{b'} \pmod{n}$ by definition, so we obtain $a + b \equiv a' + b' \pmod{n}$ and $ab \equiv a'b' \pmod{n}$ by Lemma 6.1, hence $\overline{a + b} = \overline{a' + b'}$ and $\overline{ab} = \overline{a'b'}$, which gives $\overline{a} \oplus \overline{b} = \overline{a + b} = \overline{a' + b'} = \overline{a'} \oplus \overline{b'}$ and $\overline{a} \otimes \overline{b} = \overline{ab} = \overline{a'b'} = \overline{a'} \otimes \overline{b'}$. \square

Having proved that \oplus and \otimes are well defined operations on \mathbb{Z}_n , we proceed to show that \oplus and \otimes possess many (but not all) properties of the usual addition and multiplication of integers. First we simplify our notation. From now on, we write $+$ and \cdot instead of \oplus and \otimes . In fact, we shall even drop \cdot and use simply juxtaposition to denote a product of two integers mod n . Thus we will have $\overline{a} + \overline{b} = \overline{a + b}$ and $\overline{a} \cdot \overline{b} = \overline{ab}$ or simply $\overline{a} \overline{b} = \overline{ab}$. The reader should note that the same sign "+" is used to denote two very distinct operations: \oplus in the old notation and the usual addition of integers. If anything, they are defined on distinct sets \mathbb{Z}_n and \mathbb{Z} . The same remarks apply to multiplication.

6.4 Lemma: For all $\bar{a}, \bar{b}, \bar{c} \in \mathbb{Z}_n$, the following hold.

- (1) $\bar{a} + \bar{b} \in \mathbb{Z}_n$;
- (2) $(\bar{a} + \bar{b}) + \bar{c} = \bar{a} + (\bar{b} + \bar{c})$;
- (3) $\bar{a} + \bar{0} = \bar{a}$;
- (4) $\bar{a} + \bar{-a} = \bar{0}$;
- (5) $\bar{a} + \bar{b} = \bar{b} + \bar{a}$;
- (6) $\bar{a} \cdot \bar{b} \in \mathbb{Z}_n$;
- (7) $(\bar{a} \cdot \bar{b}) \cdot \bar{c} = \bar{a} \cdot (\bar{b} \cdot \bar{c})$;
- (8) $\bar{a} \cdot \bar{1} = \bar{a}$;
- (9) if $(a, n) = 1$, then there is an $\bar{x} \in \mathbb{Z}_n$ such that $\bar{a} \cdot \bar{x} = \bar{1}$;
- (10) $\bar{a} \cdot \bar{b} = \bar{b} \cdot \bar{a}$;
- (11) $\bar{a} \cdot (\bar{b} + \bar{c}) = \bar{a} \cdot \bar{b} + \bar{a} \cdot \bar{c}$ and $(\bar{b} + \bar{c}) \cdot \bar{a} = \bar{b} \cdot \bar{a} + \bar{c} \cdot \bar{a}$;
- (12) $\bar{a} \cdot \bar{0} = \bar{0}$.

Proof: (1) is obvious. (2) follows from the corresponding property of addition in \mathbb{Z} . We indeed have

$$\begin{aligned} (\bar{a} + \bar{b}) + \bar{c} &= \overline{(a + b) + c} \\ &= \overline{a + (b + c)} \\ &= \bar{a} + \overline{(b + c)} \\ &= \bar{a} + (\bar{b} + \bar{c}). \end{aligned}$$

The remaining assertions are proved in the same way by drawing bars over integers in the corresponding equations in \mathbb{Z} . We prove only (9), which is not as straightforward as the other claims. If $(a, n) = 1$, Then there are integers x, y with $ax - ny = 1$ (Lemma 5.10). Using (3) and (12), we get $\bar{1} = \overline{ax - ny} = \overline{ax} - \overline{ny} = \overline{a \cdot x} - \overline{n \cdot y} = \overline{a \cdot x} - \overline{0 \cdot y} = \overline{a \cdot x} - \bar{0} = \bar{a} \cdot \bar{x}$. \square

Exercises

1. Determine whether the "function" $g: \mathbb{Z}_{13} \rightarrow \mathbb{N}$ is well defined, if g is defined as follows.

- (a) $g(\bar{a}) = (a, 13)$;
- (b) $g(\bar{a}) = (a, 26)$;
- (c) $g(\bar{a}) = (a, 169)$;
- (d) $g(\bar{a}) = (a^2, 13)$;

$$(e) g(\bar{a}) = (a^3, 169);$$

$$(f) g(\bar{a}) = (a, 6);$$

$$(g) g(\bar{a}) = (a^2, 65);$$

where $\bar{a} \in \mathbb{Z}_{13}$ and $a \in \mathbb{Z}$.

2) Let $f: \mathbb{Z}_{12} \times \mathbb{Z} \rightarrow \mathbb{Z}_{12}$ be such that $(\bar{a}, b) \mapsto \overline{a^2 + ab + b^2}$. Is f well defined?

3) For an integer a , we denote by \bar{a} the residue class of $a \pmod{12}$, by \tilde{a} the residue class of $a \pmod{6}$, and by \hat{a} the residue class of $a \pmod{5}$, so that $\bar{a} \in \mathbb{Z}_{12}$, $\tilde{a} \in \mathbb{Z}_6$ and $\hat{a} \in \mathbb{Z}_5$. Determine whether the following "functions" are well defined.

$$(a) \mathbb{Z}_{12} \rightarrow \mathbb{Z}_6, \bar{a} \rightarrow \tilde{a};$$

$$(b) \mathbb{Z}_6 \rightarrow \mathbb{Z}_{12}, \hat{a} \rightarrow \bar{a};$$

$$(c) \mathbb{Z}_{12} \rightarrow \mathbb{Z}_5, \bar{a} \rightarrow \hat{a};$$

$$(d) \mathbb{Z}_5 \rightarrow \mathbb{Z}_6, \hat{a} \rightarrow \tilde{a};$$

$$(e) \mathbb{Z}_5 \rightarrow \mathbb{Z}_6, \hat{a} \rightarrow a + \tilde{1}.$$