CHAPTER 5

Modular Arithmetic: Calculating with Residue Classes

> Every fine story must leave in the mind of the sensitive reader an intangible residuum of pleasure . . .
>
> —Willa Cather, *Not Under Forty*, "Miss Jewett"

WE BEGIN THIS CHAPTER WITH a discussion of the principle of division with remainder. In relation to this we shall explain the significance of these remainders, their possible applications, and how one calculates with them. In order for the functions to be introduced later to be understandable, we begin with a bit of algebra.

We have seen that in division with remainder of an integer $a \in \mathbb{Z}$ by a natural number $0 < m \in \mathbb{N}$ one has the unique representation

$$a = qm + r, \qquad 0 \leq r < m.$$

Here r is called the *remainder after division of* a *by* m or the *residue of* a *modulo* m, and it holds that m divides $a - r$ without remainder, or in mathematical notation,

$$m \mid (a - r).$$

This statement about divisibility was given a new notation by Gauss, in analogy to the equal sign:[1]

$$a \equiv r \bmod m$$

(say "a is *congruent* to r modulo m").

Congruence modulo a natural number m is an *equivalence relation* on the set of natural numbers. This means that the set $R := \{\, (a, b) \mid a \equiv b \bmod m \,\}$

[1] Carl Friedrich Gauss, 1777–1855, is to be counted among the greatest mathematicians of all time. He made many significant discoveries in mathematics as well as in the natural sciences, and in particular, at the age of 24 he published his famous *Disquisitiones Arithmeticae*, which is the foundation upon which modern number theory has been built.

of integer pairs satisfying $m \mid (a - b)$ has the following properties, which result immediately from division with remainder:

(i) R is *reflexive:* For all integers a it holds that (a, a) is an element of R, that is, we have $a \equiv a \bmod m$.

(ii) R is *symmetric:* If (a, b) is in R, then so is (b, a); that is, $a \equiv b \bmod m$ implies $b \equiv a \bmod m$.

(iii) R is *transitive:* If (a, b) and (b, c) are in R, then so is (a, c); that is, $a \equiv b \bmod m$ and $b \equiv c \bmod m$ implies $a \equiv c \bmod m$.

The equivalence relation R partitions the set of integers into disjoint sets, called *equivalence classes:* Given a remainder r and a natural number $m > 0$ the set

$$\overline{r} := \{\, a \mid a \equiv r \bmod m \,\},$$

or, in other notation, $r + m\mathbb{Z}$, is called the *residue class* of r *modulo* m. This class contains all integers that upon division by m yield the remainder r.

Here is an example: Let $m = 7$, $r = 5$; then the set of integers that upon division by 7 yield the remainder 5 is the residue class

$$\overline{5} = 5 + 7 \cdot \mathbb{Z} = \{\ldots, -9, -2, 5, 12, 19, 26, 33, \ldots\}.$$

Two residue classes modulo a fixed number m are either the same or disjoint.[2] Therefore, a residue class can be uniquely identified by any of its elements. Thus the elements of a residue class are called *representatives*, and any element can serve as representative of the class. Equality of residue classes is thus equivalent to the congruence of their representatives with respect to the given modulus. Since upon division with remainder the remainder is always smaller than the divisor, for any integer m there can exist only finitely many residue classes modulo m.

Now we come to the reason for this extensive discussion: Residue classes are objects with which one can do arithmetic, and in fact, by employing their representatives. Calculating with residue classes has great significance for algebra and number theory and thus for coding theory and modern cryptography. In what follows we shall attempt to clarify the algebraic aspects of modular arithmetic.

Let a, b, and m be integers, $m > 0$. For residue classes $\overline{a}$ and $\overline{b}$ modulo m we define the relations "$+$" and "$\cdot$", which we call addition and multiplication (of residue classes), since they are based on the like-named operations on the integers:

$$\overline{a} + \overline{b} := \overline{a + b} \quad \text{(the sum of classes is equal to the class of the sum)};$$
$$\overline{a} \cdot \overline{b} := \overline{a \cdot b} \quad \text{(the product of classes is equal to the class of the product)}.$$

[2] Two sets are said to be *disjoint* if they have no elements in common, or put another way, if their intersection is the empty set.