1 Numbers

This chapter serves as an introduction to the modern theory of algebra through the natural numbers $0, 1, 2, \ldots$. The list of natural numbers never ends and most of them are far beyond everyday use. Gigantic numbers of more than 100 digits are used to protect information transmitted over the internet.

Suppose Alice has to send a message to Bob over the internet and it must be kept secret. Alice and Bob live far apart and many intermediate computers will see the message on its way. Alice will have to scramble (encrypt) the message and send it, but at the same time Bob will have to know how to unscramble (decrypt) it. How does Alice get this information through to him? She could call and tell him. But then again someone could be listening in on their phone call. Is there a way out of this problem?

The answer is an amazing "yes" and it builds on a current paradox of mathematics: the existence of so-called one-way functions $f(X)$. These are functions easy to compute given the input X. Once they are computed and only $f(X)$ is known, it appears to be exceedingly difficult to recover X unless some secret information is known.

Here is an example of a one-way function. Fix a natural number N and let $f(X) = [X^3]$, where $[Y]$ denotes the remainder of Y after division by N. This is a function $f : M \to M$, where $M = \{0, 1, 2, \ldots, N-1\}$. When $N = 15$, f can be tabulated as

X	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$f(X)$	0	1	8	12	4	5	6	13	2	9	10	11	3	7	14

Of course we can easily find X given $f(X)$ by using the above table. But in general, as N grows the difficulty of finding X given $f(X)$ seems insurmountable unless you know some secret information. In the above example the secret information is that $f(f(X)) = X$ (you can see this using the table). In a sense we are raising a number to the third power and then scrambling things up by

taking the remainder. So far nobody has found effective methods for finding cube roots in this setting. In the above example Alice sends the encrypted message $f(X)$ to Bob and Bob decrypts it using f. This is the basic principle behind the RSA cryptosystem [22], which was the first cryptosystem based on the groundbreaking idea [8] of using one-way functions (with a trapdoor).

On a more detailed level Bob computes two gigantic prime numbers (usually 100 digits or more) p and q and forms $N = pq$. He then uses p and q to compute a number e (for encryption) and a number d (for decryption). He makes the numbers N and e public so that people wishing to write secret messages to him can use the function $f(X) = [X^e]$ for encryption, where $[Y]$ denotes the remainder of Y after division by N. He keeps the function $g(X) = [X^d]$ secret (the point being that $g(f(X)) = X$). In the example above we have $p = 3, q = 5, N = 15, e = 3, d = 3$. One way of systematically finding the secret decryption function g in the RSA system is to find the prime factors p and q of N (N being available to the general public). The straightforward method of trial division (dividing with successive primes $2, 3, 5, \dots$) is much too slow. Mathematicians have tried at least since Gauss's time (1777–1855) to find faster methods for factoring numbers. In fact Gauss writes in ([11], Art. 329)

> The problem of distinguishing prime numbers from composite numbers and of resolving the latter into their prime factors is known to be one of the most important and useful in arithmetic. It has engaged the industry and wisdom of ancient and modern geometers to such an extent that it would be superfluous to discuss the problem at length. Nevertheless we must confess that all methods that have been proposed thus far are either restricted to very special cases or are so laborious and prolix that even for numbers that do not exceed the limits of tables constructed by estimable men, i.e., for numbers that do not yield to artificial methods, they try the patience of even the practiced calculator. And these methods do not apply at all to larger numbers.

RSA Labs has put forward several factoring challenges. The hardest unsolved challenge is called RSA-2048. This is the 2048-bit number (617 digits) N on the cover of this book. It is known to be the product of two prime numbers p and q. A computer was instructed to forget p and q after forming $N = pq$. Given two candidates p' and q', it is easy to multiply them to see if their product equals N. This can be done in a small fraction of a second on any modern computer. Nevertheless, finding p and q knowing only N seems to be a painstakingly slow process not within the limits of modern computers and algorithms. If you can find p and q you will be able to claim the $200 000 prize by submitting your factorization via http://www.rsasecurity.com/go/factorization.html. Alternatively, you could settle for the less ambitious RSA factoring challenges presented at

http://www.rsasecurity.com/rsalabs/challenges/factoring/numbers.html. It has not been proved mathematically that factoring a number is a difficult problem in a precise sense, so a fast algorithm may exist waiting to be discovered. In a sense this would disrupt the pillars of the modern information age. The algebraic reasoning behind the RSA cryptosystem is founded on basic results (more than 300 years old) about the natural numbers.

1.1 The natural numbers and the integers

The natural numbers $1, 2, 3, \ldots$ were handed over to mankind by God (in the words of Kronecker (1823–91)). Mankind later added the important natural number 0. We will reserve the symbol $\mathbb{N}$ for the natural numbers $\{0, 1, 2, 3, \ldots\}$. The need for negative numbers leads us to introduce the set of integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ containing the natural numbers $\mathbb{N}$. We have deliberately cut through the red tape of formally defining $\mathbb{N}$ and $\mathbb{Z}$ here. We will also take the addition (and subtraction) and multiplication of integers for granted. This will be the starting point of our study of numbers.

1.1.1 Well ordering and mathematical induction

For $X, Y \in \mathbb{Z}$ we define $X \leq Y$ if $Y - X \in \mathbb{N}$ and $X < Y$ if $X \neq Y$ and $X \leq Y$. This leads to the usual way of ordering the integers,

$$\cdots < -3 < -2 < -1 < 0 < 1 < 2 < 3 < \cdots .$$

An element s in a subset $S \subseteq \mathbb{Z}$ is said to be a first element in S if $s \leq x$ for every $x \in S$. There are many subsets of $\mathbb{Z}$ that do not have a first element. If a subset of $\mathbb{Z}$ has a first element then the latter has to be unique (see Exercise 1.1 at the end of the chapter). The basic axiom for starting our investigation of numbers says that *every non-empty subset of* $\mathbb{N}$ *has a first element.* We also say that the set of natural numbers is *well ordered*.

The property that $\mathbb{N}$ is well ordered is equivalent to mathematical induction. Recall that mathematical induction says that if we are given statements $P(n)$ for every integer $n \geq 1$ such that

(i) $P(1)$ is true and
(ii) $P(n)$ is true implies that $P(n + 1)$ is true

then $P(n)$ is true for every $n \geq 1$.

Example 1.1.1 Let us prove the formula

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2} \tag{1.1}$$

for $n \in \mathbb{N}$ using mathematical induction. This means that we consider (1.1) as a statement $P(n)$. Clearly $P(1)$ is true, since $1 \cdot (1+1) = 2$. Suppose now that $P(n)$ is true. Then

$$1 + 2 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1).$$

The right hand side can be rewritten as

$$\begin{aligned}\frac{n(n+1)}{2} + (n+1) &= \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}.\end{aligned}$$

This is the formula for $n+1$. So we have proved that $P(n)$ implies $P(n+1)$. By mathematical induction we have proved $P(n)$ for every $n \geq 1$.

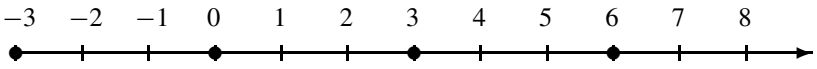
Of course, having the formal machinery for constructing a proof like this does not necessarily provide the beauty of a really ingenious mathematical argument. When Gauss was in school (at the age of seven) his mathematics teacher asked the class to sum up all numbers from 1 to 100. The students worked furiously with their small slates. Gauss was the first to give his slate with the number 5050 to the teacher. The teacher replied "Oh, I see, you probably knew the answer." "No, no! I just realized that

$$\begin{aligned}1 + 100 &= 101, \\ 2 + 99 &= 101, \\ 3 + 98 &= 101, \\ &\vdots \\ 100 + 1 &= 101.\end{aligned}$$

Therefore $1 + 2 + \cdots + 100 = (100 \cdot 101)/2 = 5050$," Gauss replied.

1.2 Division with remainder

Suppose that you mark all multiples of 3 on the axis of the integers:



An integer is uniquely given by the closest multiple of 3 to its left and the remainder you have to walk to the right. Examples are $5 = 3 + 2 = 1 \cdot 3 + 2, 7 = 6 + 1 = 2 \cdot 3 + 1, -2 = -3 + 1 = -1 \cdot 3 + 1$ and $6 = 6 + 0 = 2 \cdot 3 + 0$. Division with remainder is the generalization of this simple fact.

Theorem 1.2.1 *Let $d \in \mathbb{Z}$, where $d > 0$. For every $x \in \mathbb{Z}$ there is a unique* remainder $r \in \mathbb{N}$ *such that*

$$x = qd + r,$$

where $q \in \mathbb{Z}$ and $0 \leq r < d$.

Proof. To prove the uniqueness of r assume that $x = q_1 d + r_1$ and $n = q_2 d + r_2$, where $q_1, q_2, r_1, r_2 \in \mathbb{Z}$ and $0 \leq r_1, r_2 < d$. Then

$$(q_1 - q_2)d = r_2 - r_1.$$

If $r_1 \neq r_2$ we may assume that $r_2 > r_1$. This implies that $r_2 - r_1 = md$, where $m \geq 1$. But this contradicts the fact that $r_2 - r_1 \leq r_2 < d$. To prove the existence of r, let $M = \{x - qd \mid q \in \mathbb{Z}\}$. Then $M \cap \mathbb{N} \neq \emptyset$ (see Exercise 1.2) and we let r be the first element in the subset $M \cap \mathbb{N}$ of $\mathbb{N}$. Now $r = x - qd$ for some q and we claim that $0 \leq r < d$. If $r \geq d$ then $r > r - d \geq 0$ and $r - d = x - (q + 1)d \in M \cap \mathbb{N}$. This contradicts that r is the first element in $M \cap \mathbb{N}$. □

Definition 1.2.2 Suppose that $a = bc$ where $a, b, c \in \mathbb{Z}$. Then we say that c is a *divisor* of a (*it divides* a). We write this as $c \mid a$.

Notice that 1 and -1 divide every integer and that 0 only divides 0.

Definition 1.2.3 If $x, d \in \mathbb{Z}$, where $d > 0$, we let $[x]_d$ denote the unique remainder r in Theorem 1.2.1. Sometimes we use the notation $[x]$ when it is clear which d we are using.

1.3 Congruences

Gauss published his monumental work [11] on numbers when he was 24 years old. He had begun his deep studies in the theory of numbers at age 18. At the

start of [11] he introduced the theory of congruences, which turned out to be of fundamental importance. Congruences form an elegant way of organizing the integers according to their remainders with respect to a fixed number.

Definition 1.3.1 Let $a, b, c \in \mathbb{Z}$. Then a and b are called *congruent modulo* c if c divides $b - a$. This is denoted

$$a \equiv b \pmod{c}.$$

This may seem strange at first, but using remainders the definition (for $c > 0$) just states that a and b are congruent modulo c if and only if a and b have the same remainder when divided by c. This is the content of the following:

Proposition 1.3.2 *Let $c \in \mathbb{Z}$, where $c > 0$. Then*

(i) $a \equiv [a]_c \pmod{c}$,
(ii) $a \equiv b \pmod{c}$ *if and only if* $[a]_c = [b]_c$,

for $a, b \in \mathbb{Z}$.

Proof. We may write $a = qc + [a]_c$ for some $q \in \mathbb{Z}$, by Theorem 1.2.1. Therefore $c \mid a - [a]_c = qc$. This proves (i). Now write $b = q'c + [b]_c$ for some $q' \in \mathbb{Z}$. Then $a - b = (q - q')c + [a]_c - [b]_c$. Therefore $c \mid a - b$ if and only if $c \mid [a]_c - [b]_c$. But $c \mid [a]_c - [b]_c$ if and only if $[a]_c = [b]_c$, since $0 \leq [a]_c, [b]_c < c$. This proves (ii). □

Example 1.3.3 The integers 24 and 14 can be written $24 = 4 \cdot 5 + 4$ and $14 = 2 \cdot 5 + 4$. So $[24]_5 = [14]_5 = 4$. This means that $24 \equiv 14 \pmod{5}$. Of course this could just as easily have been observed from the fact that $5 \mid 24 - 14$.

Proposition 1.3.4 *Suppose that $x_1 \equiv x_2 \pmod{d}$ and $y_1 \equiv y_2 \pmod{d}$. Then*

(i) $x_1 + y_1 \equiv x_2 + y_2 \pmod{d}$,
(ii) $x_1 y_1 \equiv x_2 y_2 \pmod{d}$

for $x_1, x_2, y_1, y_2, d \in \mathbb{Z}$.

Proof. If d divides $x_1 - x_2$ and $y_1 - y_2$ then it also divides $x_1 - x_2 + y_1 - y_2 = x_1 + y_1 - (x_2 + y_2)$. This proves (i). Rearranging, we also get that d divides $x_1 y_1 - x_2 y_2 = x_1(y_1 - y_2) + y_2(x_1 - x_2)$. This proves (ii). □

Proposition 1.3.4 may look innocuous at first. It is surprisingly useful. For one thing, when you combine it with Proposition 1.3.2, you get (see Exercise 1.3)

$$[xy] = [[x][y]]. \tag{1.2}$$

Using (1.2) you can tell in a flash that the remainder of 13^{2003} divided by 4 has to be 1 (how?). Take a look at the following example.

1.3.1 Repeated squaring – an example

How does one find the remainder of 12^{11} divided by 21 efficiently? This problem confronts a sender of a secret message in the RSA cryptosystem, where the encryption exponent is the number $e = 11$ and the possible messages are the natural numbers less than $N = 21$. As you may have guessed the trick is to avoid computing the integer 12^{11}, divide by 21 and find the remainder. First we write 11 in the binary expansion (11 can be expressed as 1011 in the binary positional system) as

$$2^3 + 2 + 1.$$

Then using (1.2) twice we see that

$$[12^{11}] = \left[12^{2^3} 12^2 12^1\right] = \left[\left[12^{2^3}\right][12^2][12^1]\right].$$

Again using (1.2) we build a table of remainders for use in the calculation

$$\begin{aligned}
[12^1] &= 12,\\
[12^2] &= 18,\\
\left[12^{2^2}\right] = [(12^2)^2] = [[12^2][12^2]] = [18 \cdot 18] &= 9,\\
\left[12^{2^3}\right] = \left[(12^{2^2})^2\right] = \left[\left[12^{2^2}\right]\left[12^{2^2}\right]\right] = [9 \cdot 9] &= 18.
\end{aligned}$$

Picking out the relevant numbers we get

$$\begin{aligned}
[12^{11}] &= [[18 \cdot 18] \cdot 12]\\
&= [9 \cdot 12]\\
&= 3.
\end{aligned}$$

We have reduced the horrendous procedure of computing the remainder of $12^{11} = 743008370688$ divided by 21 to computing the remainders of numbers less than $21^2 = 441$. The algorithm above is called *repeated squaring*, because we constantly use the following consequence of (1.2):

$$\left[a^{2^n}\right] = \left[\left(a^{2^{n-1}}\right)^2\right] = \left[\left[a^{2^{n-1}}\right]\left[a^{2^{n-1}}\right]\right],$$

for $a, n \in \mathbb{Z}$ where $n \geq 0$ (recall that $(a^b)^c = a^{bc}$, where $a, b, c \in \mathbb{Z}$ with $b, c \geq 0$).

1.4 Greatest common divisor

Let

$$\operatorname{div}(n) = \{d \in \mathbb{N} \mid d \mid n\}$$

denote the set of natural divisors in $n \in \mathbb{Z}$. Notice that $\operatorname{div}(0) = \mathbb{N}$ and $\operatorname{div}(n) = \operatorname{div}(-n)$ for every $n \in \mathbb{Z}$.

Example 1.4.1 Let us list a few examples:

(i) $\operatorname{div}(18) = \{1, 2, 3, 6, 9, 18\}$,
(ii) $\operatorname{div}(24) = \{1, 2, 3, 4, 6, 8, 12, 24\}$,
(iii) $\operatorname{div}(36) = \{1, 2, 3, 4, 6, 9, 12, 18, 36\}$.

From this example we have

$$\operatorname{div}(24) \cap \operatorname{div}(18) = \{1, 2, 3, 6\} = \operatorname{div}(6)$$
$$\operatorname{div}(24) \cap \operatorname{div}(36) = \{1, 2, 3, 4, 6, 12\} = \operatorname{div}(12).$$

This indicates a striking fact. Given two integers m, n it seems that the common divisors $\operatorname{div}(m) \cap \operatorname{div}(n)$ of m and n are exactly the divisors $\operatorname{div}(d)$ of some third number. This is not a coincidence. It was discovered by the Greek mathematician Euclid of Alexandria (325–265BC) and is contained in book seven of his masterpiece, the *Elements*.

Lemma 1.4.2 (Euclid) *Let $m, n \in \mathbb{Z}$. There exists a unique natural number $d \in \mathbb{N}$ such that*

$$\operatorname{div}(m) \cap \operatorname{div}(n) = \operatorname{div}(d).$$

Proof. The uniqueness follows from the fact that $\operatorname{div}(d_1) = \operatorname{div}(d_2)$ if and only if $d_1 = d_2$ assuming that $d_1, d_2 \in \mathbb{N}$. When proving the existence of d we may assume that $m, n \in \mathbb{N}$, since $\operatorname{div}(x) = \operatorname{div}(-x)$ for $x \in \mathbb{Z}$. We proceed using induction on $\min(m, n)$, where $\min(m, n) = m$ if $m \leq n$ and $\min(m, n) = n$ if $m > n$. If $\min(m, n) = 0$ we may assume that $n = 0$. Therefore $\operatorname{div}(m) \cap \operatorname{div}(n) = \operatorname{div}(m)$. This settles the initial step $\min(m, n) = 0$ of the induction.

Now assume that we have proved $\mathrm{div}(m) \cap \mathrm{div}(n) = \mathrm{div}(d)$ for every $m, n \in \mathbb{N}$ with $\min(m, n) < N$, where $N > 0$. Suppose for the induction step that we are given $m, n \in \mathbb{N}$ with $\min(m, n) = N$ and that $m \geq n = N$. Then we may write $m = qn + r$, where $0 \leq r < n$ by Theorem 1.2.1. But (this is the clever step)

$$\mathrm{div}(m) \cap \mathrm{div}(n) = \mathrm{div}(m - qn) \cap \mathrm{div}(n) = \mathrm{div}(r) \cap \mathrm{div}(n),$$

since a number divides m and n if and only if it divides $m - qn$ and n. By induction we know that $\mathrm{div}(r) \cap \mathrm{div}(n) = \mathrm{div}(d)$ for some $d \in \mathbb{N}$, since $\min(r, n) = r < n = N$. This completes the proof. □

Definition 1.4.3 The unique number $d \in \mathbb{N}$ satisfying $\mathrm{div}(d) = \mathrm{div}(m) \cap \mathrm{div}(n)$ is called the *greatest common divisor* of m and n. It is denoted $\gcd(m, n)$.

If one of m and n is non-zero there is a finite number of common natural divisors. The greatest common divisor is really the greatest among these with respect to the usual ordering of $\mathbb{Z}$ (see Exercise 1.9). Notice that $\gcd(0, 0) = 0$.

1.5 The Euclidean algorithm

As already hinted in the inductive proof of Lemma 1.4.2, there is an algorithm for finding the greatest common divisor. The inductive step in the proof of Lemma 1.4.2 can be found in Euclid's *Elements* (around 300 BC) even though Euclid did not have the concept of induction and the rigor of a modern mathematical proof. The idea behind the modern version of Euclid's algorithm is the same.

Proposition 1.5.1 *Let $m, n \in \mathbb{Z}$. Then*

(i) $\gcd(m, 0) = m$ *if* $m \in \mathbb{N}$.
(ii) $\gcd(m, n) = \gcd(m - qn, n)$ *for every* $q \in \mathbb{Z}$.

Proof. Since $\mathrm{div}(0) = \mathbb{N}$, (i) follows. We get (ii) from the fact that

$$\mathrm{div}(m) \cap \mathrm{div}(n) = \mathrm{div}(m - qn) \cap \mathrm{div}(n).$$

This is a way of saying that a natural number d divides m and n if and only if d divides $m - qn$ and n, so that $\gcd(m, n) = \gcd(m - qn, n)$. □

Suppose that we wish to find the greatest common divisor of $m, n \in \mathbb{Z}$. We may assume that $m \geq n \geq 0$. If $n = 0$, we are done since $\gcd(m, 0) = m$ by Proposition 1.5.1(i). Assume that $n > 0$. The basic observation is that if we divide m by n and write $m = qn + r$ according to Theorem 1.2.1, then

$$\gcd(m, n) = \gcd(r, n) = \gcd(n, r)$$

and $n > r$. This follows from Proposition 1.5.1(ii). An example shows how this works.

Example 1.5.2 Let $m = 34$ and $n = 13$. Then

$$\begin{aligned}\gcd(34, 13) &= \gcd(13, 8) = \gcd(8, 5)\\ &= \gcd(5, 3) = \gcd(3, 2) = \gcd(2, 1)\\ &= \gcd(1, 0) = 1.\end{aligned}$$

This can also be illustrated as a sequence of divisions with remainders:

$$\begin{aligned}34 &= 2 \cdot 13 + 8,\\ 13 &= 1 \cdot 8 + 5,\\ 8 &= 1 \cdot 5 + 3,\\ 5 &= 1 \cdot 3 + 2,\\ 3 &= 1 \cdot 2 + 1,\\ 2 &= 2 \cdot 1 + 0.\end{aligned}$$

Now return to the general case $m \geq n \geq 0$. Put $r_{-1} = m$ and $r_0 = n$. If $r_0 = 0$ then $\gcd(r_{-1}, r_0) = r_{-1}$. Otherwise define r_1 to be the remainder of r_{-1} divided by r_0, so that $r_1 = r_{-1} - q_1 r_0$ for some integer q_1. Then we have

$$\gcd(r_{-1}, r_0) = \gcd(r_0, r_1)$$

and $r_{-1} > r_0 > r_1$. Proceeding in this way (if $r_1 \neq 0$) we let $r_2 = r_0 - q_2 r_1$ be the remainder of r_0 divided by r_1. Again we have

$$\gcd(r_0, r_1) = \gcd(r_1, r_2)$$

and $r_{-1} > r_0 > r_1 > r_2$. Eventually we are forced to the situation $r_N = 0$, for some step $N > 0$. This means that $\gcd(m, n) = \gcd(r_{N-1}, 0) = r_{N-1}$. The point is that the Euclidean algorithm gives rise to a strictly decreasing sequence of natural numbers $r_{-1} > r_0 > r_1 > \cdots$. If we consider the subset $R = \{r_{-1}, r_0, r_1, \dots\}$ as a subset of $\mathbb{N}$, it has a first element $r_N \in R$ since $\mathbb{N}$ is well ordered. If $r_N \neq 0$ we may continue division with remainder and get