Chapter 1

Natural numbers and primes

Throughout this book, we study natural numbers and the problem of distinguishing between those that are prime and those that are composite. Hence we begin by recalling and proving some of the fundamental properties of these numbers.

We will first learn about the principle of **mathematical induction** and define divisibility. We prove the key result that every natural number can be written uniquely as a product of prime numbers (the "Fundamental Theorem of Arithmetic") and derive and use the **Euclidean algorithm**, which allows us to tell whether two numbers have a common factor *without* having to compute their prime factor decomposition! To round off the chapter, we study the oldest known primality test – the **Sieve of Eratosthenes** – and also show that there are infinitely many different prime numbers.

1.1. The natural numbers

Let us think of the natural numbers quite naively as those quantities used to count things – whenever we count (finitely many) objects, their number should be an element of $\mathbb{N}$. We take the view that counting makes sense only if there actually is something to count:

the number 0 does *not* belong to $\mathbb{N}$. Hence the first and smallest natural number is 1.

From this intuitive point of view, let us discuss some important properties that distinguish $\mathbb{N}$ from other number systems. At first, we notice many things that *cannot* be done with natural numbers. They cannot always be subtracted from one another without leaving the set of natural numbers (as is possible for integers) nor can we divide them without restriction (as with fractions). Taking arbitrary square roots is certainly not possible either, in contrast to the positive real numbers. But $\mathbb{N}$ does have a striking property that has many useful consequences: the **well-ordering principle**.

1.1.1. Well-Ordering Principle.
Every non-empty subset of $\mathbb{N}$ *possesses a smallest element.*

We see at once that the well-ordering principle is false for $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$. Indeed, if a is an arbitrary integer, then $a-1$ is also an integer, and $a-1$ is smaller than a. So $\mathbb{Z}$ does not contain a smallest element. (However, the well-ordering principle does hold for subsets of $\mathbb{Z}$ that are *bounded from below*; see Exercise 1.1.11. We will frequently use this fact.)

On the other hand, we can justify the principle for the natural numbers as follows. If A is a non-empty subset of $\mathbb{N}$, then A contains some element $n_0 \in A$. If n_0 is not the smallest element of A, then there is some other element $n_1 \in A$ such that $n_1 < n_0$. If n_1 again is not the smallest element of A, then there is an even smaller one, and so on. But there are only $n_0 - 1$ natural numbers that are smaller than n_0, so the procedure must necessarily come to an end. Thus at some point, we will have found the smallest element of A.

Strictly speaking, this is not a proof, but merely an argument that the well-ordering principle is plausible. Indeed, we cannot give a formal proof because we did not *define* the natural numbers with the necessary accuracy. Instead, we accept the well-ordering principle as an **axiom**, i.e. a proposition whose truth we take to be evident without proof. (However, see Exercise 1.1.24.)

The method of infinite descent. The well-ordering principle provides us with a useful tool for proving statements about all natural numbers. The idea is best illustrated by an example. (Regarding the name of the following theorem, see Comment 1.1.21.)

1.1.2. Theorem (Irrationality of $\sqrt{2}$)**.**
Let n be a natural number. Then there is no natural number m with $2m^2 = n^2$.

Proof. We assume, by way of contradiction, that the claim is wrong; so suppose that there are natural numbers n and m with $2m^2 = n^2$. To see what this would entail, it might make things simpler to choose n and m as small as possible, using the well-ordering principle. Indeed, our assumption means that the set

$$A := \{n \in \mathbb{N} : \text{ there is some } m \in \mathbb{N} \text{ such that } 2m^2 = n^2\}$$

is not empty, so it has a smallest element n_0. Let us try to find out some more things about this number n_0. By definition, there exists $m_0 \in \mathbb{N}$ with

$$2m_0^2 = n_0^2. \tag{1.1.3}$$

Since $1 < 2$, we have $m_0^2 < n_0^2$, and thus m_0 is smaller than n_0.

We also see from (1.1.3) that n_0^2 is an even number. So n_0 must itself be even, since the square of an odd number is odd (Exercise 1.1.10). In other words, there is a number $\tilde{n} \in \mathbb{N}$ with $n_0 = 2\tilde{n}$. Substituting for n_0 in (1.1.3), we see that

$$2m_0^2 = n_0^2 = (2\tilde{n})^2 = 4\tilde{n}^2$$

and therefore $m_0^2 = 2\tilde{n}^2$. We conclude that m_0 also belongs to the set A. But this is impossible because $m_0 < n_0$ and n_0 was chosen to be the smallest element of A! So we have derived a contradiction – it follows that our original premise was false and that the theorem is true. ■

The principle underlying the preceding proof is called **the method of infinite descent** or **the method of the smallest counterexample**. The idea is to assume that there is a natural number

for which the claim (i.e. a statement that we wish to prove for all $n \in \mathbb{N}$) is false. By the well-ordering principle, there exists a "smallest counterexample": a smallest natural number that violates our statement. If, by studying the properties of this number, we can deduce that there would have to be an even smaller counterexample, then we obtain a contradiction. Throughout the book, the reader will encounter many applications of this valuable proof principle for the natural numbers.

Mathematical induction. The method of infinite descent is closely related to another method of proof, called **(mathematical) induction**. The following theorem formulates the concept as a general principle; Example 1.1.5 below illustrates how one applies the procedure to a specific problem.

1.1.4. Theorem (Induction principle)**.**
Suppose that $M \subseteq \mathbb{N}$ is a set of natural numbers with the following properties:

(a) *the number 1 is an element of M and*

(b) *if n is a natural number in M, then the next number $n+1$ is also an element of M.*

Then it follows that $M = \mathbb{N}$, i.e. every natural number belongs to M.

Proof. We assume that $M \neq \mathbb{N}$ and deduce a contradiction, using the method of the smallest counterexample. By our assumption, the set $A := \{n \in \mathbb{N} : n \notin M\}$, i.e. the collection of natural numbers that do not belong to M, is a non-empty subset of $\mathbb{N}$. By the well-ordering principle, this set has a smallest element n_0. Our hypothesis (a) implies that 1 is not in A, and in particular $n_0 \neq 1$. Therefore $m := n_0 - 1$ is also a natural number. Since n_0 is the smallest element of A, the number m cannot belong to A and it follows that $m \in M$. But hypothesis (b) shows that $n_0 = m + 1$ must also be an element of M, and this is a contradiction. ■

We can visualize the principle of mathematical induction by imagining an (infinite) row of dominoes, placed in such a way that each

domino will, in falling, cause the next one to topple as well. The principle of mathematical induction tells us that, if we push over the first one, every domino will eventually fall. This is certainly supported by intuition!

In order to prove some property for all natural numbers using mathematical induction, we show:

- the number 1 has the desired property (**basis of the induction**) and
- if n has the desired property, then $n+1$ also does (**induction step**).

Then Theorem 1.1.4 tells us that the desired statement does indeed hold for *all* natural numbers.

1.1.5. Example. We demonstrate that, for all natural numbers n, the number $n^3 - n$ is a multiple of 3. (That is, there is an integer m such that $n^3 - n = 3m$.)

Proof. If $n = 1$, then

$$n^3 - n = 1^3 - 1 = 1 - 1 = 0 = 3 \cdot 0,$$

so the claim is true in this case. That is the basis of the induction.

Now let us take a look at an arbitrary natural number n for which the claim is true. Then there is an integer m such that $n^3 - n = 3m$. This is called the **induction hypothesis**.

We need to show that the claim is also true for $n + 1$. That is, we must check that $(n+1)^3 - (n+1)$ is a multiple of 3. To do so, we expand the power $(n + 1)^3$ (see also Theorem 1.1.9) and see that

$$(n + 1)^3 - (n + 1) = n^3 + 3n^2 + 3n + 1 - n - 1 = n^3 + 3n^2 + 3n - n.$$

Note that $n^3 - n$ appears on the right-hand side, and our induction hypothesis tells us that $n^3 - n = 3m$. So

$$(n + 1)^3 - (n + 1) = 3m + 3n^2 + 3n = 3(m + n^2 + n).$$

Since $m + n^2 + n$ is certainly an integer, we have shown that $(n + 1)^3 - (n + 1)$ is a multiple of 3, as desired.

This completes the induction step, and the claim is proved for all natural numbers n. ■

There are a few variants of the induction principle that will be used on occasion:

(a) Sometimes it is convenient to pass from $n-1$ to n in the induction step, rather than from n to $n+1$.

(b) Statements that are true for all integers greater than or equal to some number n_0 can also be proved using mathematical induction. In this case, in the basis of the induction, we simply check the statement for $n = n_0$ (instead of $n = 1$); everything else is exactly as above. In particular, if we would like to have a formula available for all numbers in $\mathbb{N}_0$, rather than all numbers in $\mathbb{N}$, we start the induction at $n = 0$.

(c) In some cases, we derive the desired property for $n+1$ using not only the induction hypothesis for n, but also for $n-1$, or even for all $m \leq n$. That is, the principle of mathematical induction still applies when the induction hypothesis takes the stronger form "We suppose that the claim is true for all smaller numbers"; see Exercise 1.1.25.

Since we deduced the principle of mathematical induction from the well-ordering principle, every inductive proof can also be formulated as a proof by infinite descent (and vice versa; see Exercise 1.1.24). However, induction yields a **direct** proof (i.e. one that does not rely on finding contradictions), which is often more elegant. We shall decide on a case-by-case basis which method to use, depending on the specific application and also on personal taste.

Recursive definitions. Using mathematical induction, we can define certain sequences $a_1, a_2, a_3, \ldots$ of numbers without having to explicitly write down a formula. Indeed, suppose that we specify the value of the first number a_1 and know how to obtain a_{k+1} from the numbers $a_1, \ldots, a_k$. Then there is enough information to determine a_k uniquely for all $k \in \mathbb{N}$. This is called a **recursive definition**.

For example, we define a sequence $a_1, a_2, ...$ via

$$a_1 := 1 \quad \text{and} \quad a_{k+1} := \frac{1}{1+a_k} \text{ for all } k \geq 1. \tag{1.1.6}$$

From this information we can compute all numbers in the sequence, starting with a_1:

$$a_1 = 1, \qquad a_2 = \frac{1}{1+a_1} = \frac{1}{2}, \qquad a_3 = \frac{1}{1+a_2} = \frac{2}{3},$$
$$a_4 = \frac{1}{1+a_3} = \frac{3}{5}, \qquad a_5 = \frac{1}{1+a_4} = \frac{5}{8}, \qquad a_6 = \frac{1}{1+a_5} = \frac{8}{13},$$

and so on. It is often possible to use mathematical induction to prove statements about a recursively defined sequence of number, even if we do not know an explicit formula for the k-th element.

The **Fibonacci numbers** are an important example of a recursively defined sequence. They are given by

$$f_1 := 1; \quad f_2 := 1; \quad f_k := f_{k-1} + f_{k-2} \text{ for all } k \geq 3. \tag{1.1.7}$$

The first few terms are 1, 1, 2, 3, 5, 8, 13, ...

Finally, we mention the **binomial coefficients**, from the mathematical field of **combinatorics**, which play an important role in this book on a number of occasions. If $n, k \in \mathbb{N}$, then the binomial coefficient $\binom{n}{k}$ is defined to be the number of ways of choosing k out of n different-colored balls, without repetitions and disregarding the order in which the balls are picked. (In other words, $\binom{n}{k}$ is *the number of subsets of* $\{1, 2, \ldots, n\}$ that contain exactly k elements.) For example, in the United Kingdom's National Lottery there are exactly $\binom{49}{6}$ different possibilities of picking 6 different numbers from the range 1–49. Binomial coefficients satisfy the following recursive formula:

$$\binom{n}{0} = 1, \quad \binom{0}{k} = 0, \quad \text{and} \quad \binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \tag{1.1.8}$$

for all $n \in \mathbb{N}_0$ and all $k \in \mathbb{N}$. (See Exercise 1.1.15.) In the following, we use this as a formal (recursive) definition of binomial coefficients. **Pascal's triangle** (Figure 1.1) is a succinct way of illustrating the formula (1.1.8) visually.

Binomial coefficients appear, in particular, when multiplying out powers of a sum of two elements. This is the content of the **binomial theorem**.

$$
\begin{array}{ccccccccccc}
 & & & & 1 & & & & & & \\
 & & & 1 & & 1 & & & & & \\
 & & 1 & & 2 & & 1 & & & & \\
 & 1 & & 3 & & 3 & & 1 & & & \\
1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots & & \ldots & & \ldots & & \ldots
\end{array}
$$

Figure 1.1. The first rows of Pascal's triangle. The $(n+1)$-th row contains the binomial coefficients $\binom{n}{0}, \ldots, \binom{n}{n}$. By (1.1.8), each entry is obtained by adding the two entries diagonally above it.

1.1.9. Theorem (Binomial theorem)**.**
Let a and b be arbitrary real numbers and let $n \geq 0$ be an integer. Then

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Remark. For $n = 2$, we obtain $(a+b)^2 = a^2 + 2ab + b^2$, the well-known formula for the square of a sum. In Example 1.1.5, multiplying out $(n+1)^3 = n^3 + 3n^2 + 3n + 1$ corresponds to the case $n = 3$.

Proof of the binomial theorem. It is a useful exercise to justify the claim using the combinatorial definition of binomial coefficients. (How many times does the term $a^k b^{n-k}$ appear when multiplying out $(a+b)^n$?) Instead, we use the opportunity to practice mathematical induction with a slightly more complicated example than before.

Let us keep the numbers a and b fixed. First suppose that $n = 0$; then $(a+b)^n = 1$ and

$$\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \binom{0}{0} a^0 b^0 = 1,$$

so the theorem is true in this case. That is the basis of the induction.

Now we suppose that the claim is true for n and we must show that it also holds for $n+1$. Using our induction hypothesis and then

multiplying out, we see that

$$\begin{aligned}(a+b)^{n+1} &= (a+b)\cdot(a+b)^n \\ &= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \\ &= \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n+1-k}.\end{aligned}$$

Substituting $j = k+1$ in the first sum and $j = k$ in the second, we obtain

$$(a+b)^{n+1} = \sum_{j=1}^{n+1}\binom{n}{j-1}a^j b^{n+1-j} + \sum_{j=0}^{n}\binom{n}{j}a^j b^{n+1-j}.$$

(We remind the reader that the Σ-notation for sums is merely an abbreviation, so it does not matter what name we give to the summation variable!)

Now separate the last term of the first sum and the first term of the second sum; then we can combine the remaining terms of both. By definition, $\binom{n}{0} = \binom{n}{n} = 1$, so

$$\begin{aligned}(a+b)^{n+1} &= \binom{n}{n}a^{n+1} + \binom{n}{0}b^{n+1} \\ &\quad + \sum_{j=1}^{n}\binom{n}{j-1}a^j b^{n+1-j} + \sum_{j=1}^{n}\binom{n}{j}a^j b^{n+1-j} \\ &= a^{n+1} + b^{n+1} + \sum_{j=1}^{n}\left(\binom{n}{j-1} + \binom{n}{j}\right)a^j b^{n+1-j}.\end{aligned}$$

Now we can use the recursive formula (1.1.8):

$$\begin{aligned}(a+b)^{n+1} &= a^{n+1} + b^{n+1} + \sum_{j=1}^{n}\binom{n+1}{j}a^j b^{n+1-j} \\ &= \sum_{j=0}^{n+1}\binom{n+1}{j}a^j b^{n+1-j}.\end{aligned}$$

This completes the induction step. By mathematical induction, the binomial theorem holds for all n. ■

Exercises.

1.1.10. Exercise (!). Recall that a natural number n is called **even** if there is a natural number m with $n = 2m$. The number n is called **odd** if there is a natural number m such that $n = 2m - 1$.

(a) Show that every natural number is either even or odd, but not both at the same time. (*Hint:* By the well-ordering principle, there is a smallest number m satisfying $2m \geq n$.)

(b) Show that the product of two even numbers is even and that the product of two odd numbers is odd. What can you say about the product of an odd number and an even number?

1.1.11. Exercise (!). Suppose that M is a non-empty set of integers. Then M is called **bounded from above** or **bounded from below** if there exists an integer K such that $x \leq K$ for all $x \in M$ or $x \geq K$ for all $x \in M$, respectively.

Prove the following: if M is bounded from below, then M contains a smallest element. Similarly, if M is bounded from above, then M has a largest element. Are these statements also true for subsets of $\mathbb{Q}$ or $\mathbb{R}$?

(*Hint:* For the first part apply the well-ordering principle to the set of all numbers of the form $1 + x - K$, with $x \in M$. For the second part, apply the first claim to the set $\{x \in \mathbb{Z} : -x \in M\}$.)

1.1.12. Exercise (!). Prove by mathematical induction for all $n \in \mathbb{N}$:

(a) $2^n \geq 2n$.

(b) $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ ("Gauss summation").

(c) $\displaystyle\sum_{k=0}^{n-1} x^k = \frac{1-x^n}{1-x}$ for all real numbers $x \neq 1$.

(d) For all real numbers $x \neq 1$,
$\displaystyle\sum_{k=0}^{n-1} (k+1) \cdot x^k = \frac{nx^{n+1} - (n+1)x^n + 1}{(1-x)^2}$.

(e) $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

(f) $\displaystyle\sum_{k=0}^{n} (k \cdot k!) = (n+1)! - 1$.

(g) $\displaystyle\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{n}{n+1}$.

1.1.13. Exercise.

(a) Prove that $n^5 - n$ is a multiple of 5 for all $n \in \mathbb{N}$.

(b) Is $n^4 - n$ a multiple of 4 for all $n \in \mathbb{N}$? If not, give a counterexample and explain why the proof from (a) fails here.

1.1.14. Exercise. Suppose that $g_1, \ldots, g_n$ are straight lines in the plane, no two of which are parallel to each other. (We say that the lines are **pairwise** not parallel.) Furthermore no more than two lines should intersect at any given point in the plane.

Into how many different pieces do these lines separate the plane? Develop an idea, write down a formula, and prove it using mathematical induction. What changes if some of the lines are allowed to be parallel?

1.1.15. Exercise (!).

(a) Using the intuitive definition of binomial coefficients, prove that the recursive formula (1.1.8) is true.

(b) Also explain why the equation

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{1.1.16}$$

holds for all $k, n \in \mathbb{N}_0$ with $k \leq n$.

(c) Deduce (1.1.16) from the recursive formula (1.1.8), using mathematical induction.

1.1.17. Exercise (!). Let $n, k, \ell \in \mathbb{N}_0$. Prove the following:

(a) $\displaystyle\binom{n+\ell}{k} \geq \binom{n}{k}$.

(b) $\displaystyle\binom{n+\ell}{k+\ell} \geq \binom{n}{k}$.

(c) The "middle" binomial coefficients $\binom{2n}{n}$ grow at least exponentially, i.e.

$$\binom{2n}{n} \geq 2^n.$$

(*Hint:* For the first two parts it is useful to take a look at Pascal's triangle. For the third part, use induction together with the recursive formula for binomial coefficients and the first two parts of the exercise.)

1.1.18. Exercise (!). Let $n, k \in \mathbb{N}_0$ and write $a(n, k)$ for the number of ways of choosing up to k (not necessarily different) numbers from 1

to n, disregarding the order in which they are picked. (Here we allow the possibility of not picking *any* numbers; for example $a(n,0) = 1$ and $a(n,1) = n+1$ for all n.)

(a) Justify that, for all natural numbers $n, m \geq 1$, the number $a(n,m)$ satisfies the recursive formula

$$a(n,m) = a(n-1,m) + a(n,m-1).$$

(b) Prove by induction that

$$a(n,m) = \binom{n+m}{m}.$$

1.1.19. Exercise. Prove that the rational numbers a_k defined by (1.1.6) satisfy $a_k = f_k/f_{k+1}$ for all k, where f_k is the k-th Fibonacci number as defined in (1.1.7).

1.1.20. Exercise. Show that the k-th Fibonacci number is given by

$$f_k = \frac{(1+\sqrt{5})^k - (1-\sqrt{5})^k}{2^k \cdot \sqrt{5}} \qquad \text{(for all } k \geq 1\text{)}.$$

(*Hint:* Use version (c) of the induction principle.)

Further Exercises and Comments.

1.1.21. Theorem 1.1.2 is equivalent to the well-known fact that $\sqrt{2}$ is an irrational number. Indeed, $2m^2 = n^2$ just means that

$$\left(\frac{n}{m}\right)^2 = 2.$$

There also is a geometric interpretation of this statement: there is no square for which both the sidelength a and the length d of the diagonal are natural numbers. (Otherwise, Pythagoras's Theorem implies that $2a^2 = d^2$.)

1.1.22. If n is a natural number, then $n+1$ is usually called the **successor** of n. The following are evident:

(I) 1 is a natural number;

(II) every natural number has exactly one successor;

(III) there is no natural number whose successor is 1, but every natural number $n \neq 1$ is itself the successor of some other natural number;

(IV) different natural numbers have different successors;

(V) if M is a subset of $\mathbb{N}$ that contains 1 and also contains the successor of each of its elements, then $M = \mathbb{N}$.

Here statement (V) is exactly the induction principle proved in Theorem 1.1.4. Properties (I) to (V) are known as the **Peano axioms**, after the Italian mathematician Giuseppe Peano.

It turns out that the Peano axioms describe the natural numbers uniquely, i.e. there is (up to a relabeling of the elements) no other set with the same properties. For this reason, they can be used to *define* the natural numbers; this is the usual way of introducing them in a university-level mathematics course. We decided to begin with the well-ordering principle instead, feeling that this would appear more intuitive to readers who have not encountered either concept before.

1.1.23. Exercise. Note that the Peano axioms do not include any statements about elementary arithmetic; they require only that for every natural number n the successor $n + 1$ is defined.

Show that, using this successor function, the sums $n+m$, products $n \cdot m$, and powers n^m for all natural numbers n, m can be defined recursively.

1.1.24. Exercise. Show that the well-ordering principle follows from the principle of mathematical induction (and hence from the Peano axioms).

1.1.25. Exercise. Prove – using either the well-ordering principle or the Peano axioms – that an analogue of Theorem 1.1.4 holds for each of the alternative versions of the induction principle introduced in the text.

1.1.26. The proof principle of infinite descent was first described by Pierre de Fermat (1601–1665), a French lawyer who intensively studied mathematics in his spare time. He did not publish mathematical texts himself; we know about his work from the letters he exchanged with other mathematicians and from handwritten notes that he made in the margins of textbooks. Fermat is famous for his claim that, for all natural numbers $n > 2$, there are no natural numbers a, b, and c such that $a^n + b^n = c^n$.

Although it is believed today that Fermat did not know of a correct proof of this statement, it has become known as **Fermat's Last Theorem**. It took 300 years until the problem was finally solved by the English mathematician Andrew Wiles. (His proof appeared in the *Annals of Mathematics* in 1995.)

1.1.27. Both Example 1.1.5 and Exercise 1.1.13(a) are special cases of **Fermat's Little Theorem**, which we will encounter in Section 3.2.

1.1.28. Exercise. We claim that every natural number can be described by a sentence containing at most two hundred letters.

"Proof": Assume that the claim is false. Then there exists *the smallest natural number that cannot be described by a sentence containing at most two hundred letters.* This description contains less than two hundred letters – a contradiction!

However, the claim above is clearly wrong. Indeed, there are only finitely many sentences containing at most two hundred letters, but infinitely many natural numbers .... Where did we make a mistake?

1.2. Divisibility and primes

Having discussed the natural numbers, we are now ready to discuss the question of when we can divide one natural number by another, which naturally leads to the definition of prime numbers. We begin by formally introducing a few concepts with which the reader is likely to be familiar.

1.2.1. Definition (Divisors, multiples, and primes)**.**
Let $n, k \in \mathbb{Z}$. We call k a **divisor** of n (and conversely n a **multiple** of k) if there exists an integer m such that $k \cdot m = n$. In this case, we say that k **divides** n, or that n is **divisible** by k, and write $k \mid n$.

A **prime** is a natural number that is divisible by exactly two different natural numbers (namely by 1 and itself).

A natural number $n > 1$ that is not prime is called a **composite number**.

For example, $3 \mid 6$ because $3 \cdot 2 = 6$. We also have $-3 \mid 6$, since $(-3)\cdot(-2) = 6$. Every integer is divisible by both 1 and -1. Similarly, every integer is a divisor of zero. (This is a good reason to leave zero out of the natural numbers.) If a, b are natural numbers and a divides b, then we see immediately that $a \leq b$. Readers are invited to convince themselves that this is not true when a and b are allowed to be negative!

If $n \in \mathbb{Z}$, then 1, -1, n, and $-n$ are certainly divisors of n; they are called the **trivial divisors**. If there are any other divisors, then they are called **non-trivial divisors**. Thus prime numbers only have trivial divisors. Note that 1 is *not* prime, since there is only one

natural number that divides it! The next result provides us with some simple rules for working with divisors.

1.2.2. Theorem (Rules for divisibility)**.**
Let $a, b, c \in \mathbb{Z}$. If b and c are divisible by a, then so are $b + c$, $b - c$, and $b \cdot c$. Every divisor of b divides every multiple of b; that is, if $a \mid b$ and $b \mid c$, then also $a \mid c$.

Proof. Suppose that a divides both b and c. Then, by definition, there exist integers n and m such that $a \cdot n = b$ and $a \cdot m = c$. Hence

$$\begin{aligned} b + c &= a \cdot n + a \cdot m = a \cdot (n + m), \\ b - c &= a \cdot n - a \cdot m = a \cdot (n - m), \\ b \cdot c &= (an) \cdot (am) = a \cdot (n \cdot a \cdot m). \end{aligned}$$

Since $n + m$, $n - m$, and $n \cdot a \cdot m$ are integers, it follows that a is a divisor of $b + c$, of $b - c$, and of $b \cdot c$. The final claim follows similarly; we leave the details to the reader as an exercise. ■

1.2.3. Definition (Common divisors and multiples; gcd and lcm)**.**
Let $a, b \in \mathbb{Z}$ and let $k \in \mathbb{Z}$ be a number that divides both a and b. Then k is called a **common divisor** or a **common factor** of a and b. Similarly, a number $v \in \mathbb{Z}$ that is a multiple of both a and b is called a **common multiple** of a and b.

Now suppose that $a \neq 0$ or $b \neq 0$. Then the largest number $k \in \mathbb{N}$ that is a common divisor of a and b is referred to as the **greatest common divisor** (gcd) of a and b. We write $k = \gcd(a, b)$. (The gcd is sometimes also called the **"highest common factor"**.)

Correspondingly, if neither a nor b is equal to zero, the smallest natural number v that is a common multiple of a and b is called the **least common multiple** (lcm) of a and b; we write $v = \operatorname{lcm}(a, b)$. For completeness, we also define $\gcd(0, 0) := 0$ and $\operatorname{lcm}(a, 0) := \operatorname{lcm}(0, a) := 0$ for all $a \in \mathbb{Z}$.

Two integers a and b are called **coprime** if $\gcd(a, b) = 1$, i.e. if a and b do not have any positive common divisors apart from 1.

Remarks.

(a) The least common multiple exists by the well-ordering principle. Furthermore, for every $c \neq 0$ the set of divisors of c is bounded from above by the absolute value $|c|$. Hence the gcd also always exists (see Exercise 1.1.11).

(b) From school, we know that every common divisor of a and b also divides $\gcd(a, b)$. This is not completely obvious from the definition and therefore requires a proof, which we shall give in the next section. A simple method of computing $\gcd(a, b)$ will be introduced in Section 1.4.

Examples.

(a) If $n \in \mathbb{Z}$ is arbitrary, then $\gcd(1, n) = 1$ and $\gcd(0, n) = n$.

(b) The positive common divisors of 12 and 18 are 1, 2, 3, and 6. Hence $\gcd(12, 18) = 6$.

(c) The numbers 3 and -6 are *not* coprime as $\gcd(3, -6) = 3$. In contrast, $\gcd(5, 12) = \gcd(5, 17) = \gcd(12, 17) = 1$, so the numbers 5, 12, and 17 are pairwise coprime.

We now turn to the concept of **division with remainder**. Since this is central for our book and for number theory in general, we shall prove formally that division with remainder is always possible. This is sometimes referred to as the **"division algorithm"**; however, it is a theorem, not an algorithm in the sense of our book.

1.2.4. Theorem (Division theorem)**.**
Let $a \in \mathbb{Z}$ and $b \in \mathbb{N}$. Then there exist integers q and r such that $0 \leq r < b$ and $a = qb + r$.

*We say that b **divides** a **with remainder** r. The numbers q and r are uniquely determined by a and b.*

Remark. In particular, b divides a with remainder 0 if and only if b is a divisor of a.

Proof. The idea is simple: we choose q as large as possible while requiring that the remainder $r = a - qb$ is not negative. Then q

and r have the desired property and are uniquely determined by this description.

To develop this outline into a formal proof, consider the set $Q := \{q \in \mathbb{Z} : a - qb \geq 0\}$. We have

$$a - (-|a| \cdot b) = a + |a| \cdot b \geq a + |a| \geq 0,$$

so $-|a| \in Q$ and in particular $Q \neq \varnothing$. On the other hand, we have $a - nb < 0$ for all $n > |a|$, so it follows that the set Q is bounded from above. By Exercise 1.1.11, Q has a largest element q. We set $r := a - qb \geq 0$. By maximality of q, we see that $q + 1 \notin Q$ whence

$$r - b = a - qb - b = a - (q + 1)b < 0.$$

Hence it follows that $0 \leq r < b$ and $a = qb + r$ as claimed.

To prove uniqueness, let us assume that q', r' are integers such that $a = q'b + r'$ with $0 \leq r' < b$. We need to show that $q' = q$ and $r' = r$. By definition of the set Q, we know that $q' \in Q$. Also, we see for all $n \geq q' + 1$ that

$$a - nb \leq a - (q' + 1)b = a - q'b - b = r' - b < 0.$$

Therefore $n \notin Q$. This means that q' is the largest element of Q, giving $q' = q$. But then we also have $r' = a - q'b = a - qb = r$, as desired. ■

Example. The number 5 divides 47 with remainder 2, as $47 = 9 \cdot 5 + 2$ and $0 \leq 2 < 5$. For larger examples, the numbers q and r can be found with the usual method of **long division**. For example, to divide 10 007 by 101:

$$\begin{array}{r} 99 \\ 101\,\overline{)\,10007} \\ \underline{9090} \\ 917 \\ \underline{909} \\ 8 \end{array}$$

So $10\,007 = 99 \cdot 101 + 8$, and 101 divides 10 007 with remainder 8.

In Section 3.1, we will discuss many further properties of division with remainder. However, the concept will already be important for developing the **Euclidean algorithm** in Section 1.4.

Exercises.

1.2.5. Exercise. Suppose that n and m are integers and let $k \in \mathbb{N}$. Prove or disprove the following statements!

(a) If k divides $m+n$, then it also divides n and m.

(b) If k divides $m \cdot n$, then k is also a divisor of n and m.

(c) If k is a divisor of $m \cdot n$, then k divides n or m.

(d) If k divides m but not n, then k does not divide $m+n$.

(e) If k divides m, but not n, then k does not divide $m \cdot n$.

(f) If k divides n and m with remainder 1, then the remainder of $n \cdot m$ after division by k is also 1.

(g) If k divides both n and m with remainder 1, then k also divides $n+m$ with remainder 1.

1.2.6. Exercise (!).

(a) Show that every natural number $n > 1$ has at least one **prime factor**. (This means that there exists a prime number p such that $p \mid n$.)

(b) Show that a composite number $n > 1$ has at least one non-trivial divisor k such that $k^2 \leq n$.

1.2.7. Exercise. Let $k \in \mathbb{N}$ and $a, b \in \mathbb{Z}$. Furthermore, suppose that k divides the number a with remainder r and the number b with remainder s. Develop rules for the remainders of $a+b$, $a-b$, and $a \cdot b$ when divided by k, and then prove the correctness of these rules!

1.2.8. Exercise. Let $n \in \mathbb{N}$. Show that exactly one of the numbers n, $n+1$, and $n+2$ is divisible by 3.

1.2.9. Exercise. Let $n \geq 3$ be an odd natural number. Show that exactly one of the numbers $n+1$ and $n-1$ is divisible by 4.

1.2.10. Exercise. Let a, b, c, and d be integers such that $a \mid b$ and $c \mid d$. Does this imply that $ac \mid bd$?

1.2.11. Exercise. Show that, for all integers a and b, the number $2a+b$ is divisible by 7 if and only if $100a+b$ is divisible by 7. Use this fact to decide whether 100002 is divisible by 7. Can you find and prove similar "division rules"?

1.2.12. Exercise (!). Let $n \in \mathbb{N}$ and let p be prime. Prove: if p does not divide n, then n and p are coprime.

1.3. Prime factor decomposition

The importance of prime numbers stems from the fact that they are in some sense the "building blocks" of natural numbers. Indeed, every natural number $n \geq 2$ has a **prime factor decomposition**[1]; i.e. it can be written as a product of prime numbers

$$n = p_1 \cdot p_2 \cdots p_k. \tag{1.3.1}$$

(These numbers $p_1, \ldots, p_k$ are called the **prime factors** of n.) Furthermore, the decomposition is unique – apart from the possibility of reordering the factors in (1.3.1). We now formally prove these facts, together called the **Fundamental Theorem of Arithmetic**.

1.3.2. Theorem (Fundamental Theorem of Arithmetic)**.**
Let $n \geq 2$ be a natural number. Then n is a product of prime numbers, and this decomposition into primes is unique up to a reordering of the factors.

(In particular, the number of distinct primes and their multiplicities in the prime factor decomposition are uniquely determined.)

Proof. We shall prove the theorem using variant (c) of the induction principle. That is, let $n_0 \geq 2$ be a natural number, and suppose that we know that all smaller numbers $n \in \{2, 3, \ldots, n_0 - 1\}$ have a unique prime factor decomposition. For convenient notation, let us denote this decomposition by $\Pi(n)$; we must show that n_0 also has a unique prime factor decomposition.

First observe that the claim is true when n_0 is prime: in this case the decomposition consists of a single prime number and is unique because no other prime divides n_0.

Now suppose that n_0 is not prime, and let p be the smallest prime divisor of n_0 (see Exercise 1.2.6). Then we can write $n_0 = p \cdot k$ for some k with $2 \leq k < n_0$. By the induction hypothesis, k has a unique

[1] Sometimes it is useful to agree by convention that $n = 1$ also has a prime factor decomposition, namely the *empty* decomposition into no prime factors. This makes some statements and proofs easier because it is not necessary to treat this case differently.

prime factor decomposition $\Pi(k)$. In particular, n_0 has the prime factor decomposition

$$n_0 = p \cdot \Pi(k),$$

and this is the unique decomposition (up to reordering) that contains the number p. Consider any decomposition

$$n_0 = p_1 \cdot p_2 \cdots p_m,$$

where the prime factors are written in non-decreasing order. Recall that p is the *smallest* prime divisor of n_0, so $p_1 \geq p$; we must show that $p_1 = p$.

Let us assume, by contradiction, that $p_1 > p$. Then we divide p_1 by p with remainder; hence we let $q, r \in \mathbb{Z}$ be such that

$$p_1 = q \cdot p + r$$

and $0 \leq r < p$. Since p and p_1 are distinct prime numbers by assumption, we must have $r \geq 1$. Now we can write

$$n_0 = (q \cdot p + r) \cdot p_2 \cdots p_m = p \cdot q \cdot \ell + r \cdot \ell, \tag{1.3.3}$$

where we abbreviate $\ell = p_2 \cdots p_m$. The number $r \cdot \ell$ has the prime factor decomposition

$$r \cdot \ell = \Pi(r) \cdot p_2 \cdots p_m. \tag{1.3.4}$$

But p divides both n_0 and $p \cdot q \cdot \ell$, so according to (1.3.3), $r \cdot \ell$ is divisible by p. By the induction hypothesis, this means that p must appear in the prime factor decomposition of $r \cdot \ell$. But p is not visible in (1.3.4)! (Recall that $r < p$.) This is a contradiction. The induction, and hence the proof of the theorem, is complete. ■

Looking a little more closely at the proof, we notice that it is *uniqueness* of the decomposition, rather than existence, that posed the greatest difficulty. More precisely, we needed to show that a prime divides the product of several integers only if it divides one of these integers themselves. As this observation is important in its own right, we shall record it here for further reference.

1.3.5. Corollary (Prime divisors of a product)**.**
Let a and b be integers and let p be a prime divisor of the product $a \cdot b$. Then p divides a or b.

More generally, suppose that $a, b \in \mathbb{Z}$ are both coprime to $k \in \mathbb{Z}$. Then the product $a \cdot b$ is also coprime to k.

Proof. To prove the first claim, we can assume that a and b are both natural numbers ≥ 2 (since the divisors of a are precisely the divisors of $|a|$, and the claim is trivial when one of a and b is zero or one). By hypothesis let $k \in \mathbb{N}$ be such that $p \cdot k = a \cdot b$.

Let us write again $\Pi(n)$ for the prime factor decomposition of a natural number $n \geq 2$; then $a \cdot b$ has the decompositions

$$a \cdot b = p \cdot \Pi(k) \qquad \text{and} \qquad a \cdot b = \Pi(a) \cdot \Pi(b).$$

By the Fundamental Theorem of Arithmetic, the two decompositions agree up to reordering. So p must occur in $\Pi(a)$ or in $\Pi(b)$, as claimed.

The second claim follows from the first. Indeed, suppose that $a \cdot b$ is *not* coprime to k, and let p be a prime divisor of $\gcd(a \cdot b, k)$. Then p divides both k and $a \cdot b$. By the first part of the corollary, one of a and b is a multiple of p, and hence not coprime to k. ■

More generally, a natural number k divides another number n if and only if the prime factor decomposition of k is contained in the prime factor decomposition of n. So, if n and m are natural numbers, we can find the greatest common divisor and the lowest common multiple of n and m from their decompositions. For example, $n = 90$ and $m = 315$ can be written as $n = 2 \cdot 3 \cdot 3 \cdot 5$ and $m = 3 \cdot 3 \cdot 5 \cdot 7$, so

$$\gcd(n, m) = 3 \cdot 3 \cdot 5 = 45 \quad \text{and} \quad \operatorname{lcm}(n, m) = 2 \cdot 3 \cdot 3 \cdot 5 \cdot 7 = 630.$$

This representation of gcd and lcm using prime factors will be extremely helpful for us in several places. For example, it immediately implies an important property of the gcd:

1.3.6. Theorem (Divisors of the gcd)**.**
Let $a, b \in \mathbb{Z}$. Then every common divisor of a and b also divides $\gcd(a, b)$.

On the other hand, it is very difficult to find the prime factor decomposition of a large integer, so computing the gcd and lcm in this manner is not practical! In the next section, we develop a much more effective method. (Compare Exercises 1.3.8 and 1.4.5.)

Exercises.

1.3.7. Exercise. Find the prime factor decomposition of the numbers 600, 851, and 1449.

1.3.8. Exercise. Compute $\gcd(1961, 1591)$ by finding the prime factor decomposition of the two numbers.

1.3.9. Exercise (!). Let $a, b \in \mathbb{Z}$ and $d := \gcd(a, b)$. Prove:

(a) $\frac{a}{d}$ and $\frac{b}{d}$ are coprime.

(b) If v is a common multiple of a and b, then v is a multiple of $\operatorname{lcm}(a, b)$.

(c) $\operatorname{lcm}(a, b) \cdot d = |a \cdot b|$. In particular, if a and b are coprime, then $\operatorname{lcm}(a, b) = |a \cdot b|$.

1.3.10. Exercise. Let a, b, and c be integers such that c is a divisor of the product $a \cdot b$. Is it true that c divides $\gcd(a, c) \cdot \gcd(b, c)$?

Further Exercises and Comments.

1.3.11. The Fundamental Theorem of Arithmetic can be found implicitly already in Euclid's *Elements*. However, the theorem was not explicitly formulated and proved until 1801, when Gauss did so for the first time in his *Disquisitiones Arithmetica*. For a detailed overview over the history of this theorem we refer the reader to the article [**AÖ**].

1.3.12. Often the gcd is defined to be a positive common divisor that is divisible by every other common divisor (and thus is a "biggest" common factor in that sense). Theorem 1.3.6 is then needed to show that "the" gcd is unique.

1.3.13. Using the results of the next section, we can give an alternative proof of the Fundamental Theorem of Arithmetic that is perhaps more elegant, but less direct, than the one we gave here. (See Exercise 1.4.9.)

1.4. The Euclidean algorithm

So how do we find the greatest common divisor of two given integers a and b? As already mentioned, simply computing all divisors of a and b, or using their prime factor decompositions, is not workable for very large numbers. On the other hand, the **Euclidean algorithm**, which we shall now describe, can be performed very quickly even when a and b have thousands of digits. The development of this method will also give us some important theoretical insights. (Regarding the word "algorithm" and whether or not it is appropriate in this context, we refer the reader to Chapter 2.)

The basic idea is to use a and b to come up with new (and potentially smaller) numbers that have the same common divisors:

1.4.1. Lemma (Pairs of numbers with the same common divisors)**.** *Let a, b, and m be arbitrary integers. Then every common divisor of a and b is also a common divisor of a and $c := b + m \cdot a$; conversely each common divisor of a and c is a common divisor of a and b.*

In particular, $\gcd(a, b) = \gcd(a, b + m \cdot a)$.

Proof. If k is a common divisor of a and b, then, by Theorem 1.2.2, k divides $m \cdot a$, and hence also $c = b + m \cdot a$, as claimed.

On the other hand, let k be a common divisor of a and c. We can apply the fact we just proved to the numbers a, c, and $-m$. This shows that k is a common divisor of a and $c + (-m) \cdot a = (b + m \cdot a) - m \cdot a = b$, as desired. The final statement of the lemma follows from the definition of the greatest common divisor. ■

How do we use this observation? If $a, b \in \mathbb{N}$ are natural numbers with $a > b$, then we can divide a by b with remainder and hence find $q, r \in \mathbb{Z}$ such that $a = q \cdot b + r$ and $0 \leq r < b$. Lemma 1.4.1 tells us that $\gcd(a, b) = \gcd(b, r)$ – so we have reduced the problem to finding the greatest common divisor of another pair of numbers. If