Elementary Number Theory and Algebra

Jan Nekovář

December 1, 2019

This is an expanded version of notes of a second-year course taught at the university formerly known as Paris 6 (at the time of writing in September 2019, it is a part of Sorbonne Université).

The arithmetic part of the course corresponds to chapters 1–4, which cover basic properties of prime numbers and divisibility, uniqueness of factorisation for integers, Euclid's algorithm and Bézout's theorem, congruences, Euler's theorem and its variants, and basic applications to cryptography.

The algebraic part of the course corresponds to chapters 6–8, which introduce groups, rings and basic constructions involving them, and to certain parts of chapters 9–10 treating polynomial rings in one variable and construction of fields, in particular of finite fields.

Contents

1 Integers, divisibility, prime numbers

1.1 Divisibility, prime numbers

In this section we introduce the main protagonists of the whole story.

1.1.1 Notation We use the standard notation

$$\mathbf{N} := \{0, 1, 2, 3, \ldots\}, \qquad \mathbf{N}_+ := \{1, 2, 3, \ldots\}, \qquad \mathbf{Z} := \{0, \pm 1, \pm 2, \pm 3, \ldots\}$$
$$\mathbf{Q} := \left\{ \tfrac{a}{b} \mid a, b \in \mathbf{Z},\, b \neq 0 \right\}$$

for the set of natural integers, positive integers, integers and rational numbers, respectively.

For complex numbers $a, b \in \mathbf{C}$ and subsets $X, Y \in \mathbf{C}$ we let

$$X + Y := \{x + y \mid x \in X,\, y \in Y\}, \qquad aX = Xa := \{aX \mid x \in X\}.$$

With this notation, $aX + bY = \{ax + by \mid x \in X,\, y \in Y\}$.

1.1.2 Factorisation into primes Experience shows that positive integers admit unique factorisation as products of prime numbers:

$1 = 1$	$7 = \boxed{7}$	$13 = \boxed{13}$
$2 = \boxed{2}$	$8 = 2 \cdot 2 \cdot 2 = 2^3$	$14 = 2 \cdot 7$
$3 = \boxed{3}$	$9 = 3 \cdot 3 = 3^2$	$15 = 3 \cdot 5$
$4 = 2 \cdot 2 = 2^2$	$10 = 2 \cdot 5$	$16 = 2 \cdot 2 \cdot 2 \cdot 2 = 2^4$
$5 = \boxed{5}$	$11 = \boxed{11}$	$17 = \boxed{17}$
$6 = 2 \cdot 3$	$12 = 2 \cdot 2 \cdot 3 = 2^2 \cdot 3$	$18 = 2 \cdot 3 \cdot 3 = 2 \cdot 3^2$

In this table, the boxed numbers $2, 3, 5, 7, 11, 13, 17 \ldots$ are **prime numbers** — they cannot be written as products in a non-trivial way.

1.1.3 Definition. An integer $a > 1$ is a **prime number** if $a \neq bc$ for any $b, c \in \mathbf{N}$ such that $b, c \neq 1$. Denote by $\mathcal{P} = \{2, 3, 5, 7, 11, 13, 17 \ldots\}$ the set of all prime numbers.

1.1.4 Definition. Let $a, b \in \mathbf{Z}$. We say that b **divides** a (equivalently, that a **is a multiple of** b, or that b **is a divisor of** a) if there exists $c \in \mathbf{Z}$ such that $a = bc$. **Notation:** $b \mid a$ (if b does not divide a, we write $b \nmid a$).

1.1.5 Basic properties of divisibility The following properties are immediate consequences of the definition.

- $b \mid a \iff b\mathbf{Z} \supseteq a\mathbf{Z}$ ("b divides a if and only if every multiple of a is a multiple of b")
- $\forall b \in \mathbf{Z} \quad b \mid 0 \quad$ (since $b \cdot 0 = 0$), $\qquad \forall a \in \mathbf{Z} \quad 1 \mid a \quad$ (since $1 \cdot a = a$)
- $b \mid \pm 1 \iff b = \pm 1$
- $b \mid a \iff \pm b \mid \pm a$
- $b \mid a_1, a_2 \Longrightarrow b \mid (a_1 \pm a_2)$
- $c \mid b$, $b \mid a \Longrightarrow c \mid a$
- $b_1 \mid a_1$, $b_2 \mid a_2 \Longrightarrow b_1 b_2 \mid a_1 a_2 \qquad$ (in particular, $b \mid a \Longrightarrow b^2 \mid a^2$)
- if $c \neq 0$, then it is equivalent: $b \mid a \iff bc \mid ac \quad$ (since $ac = bcd \iff a = bd$).

Moreover,

- if $a, b \in \mathbf{Z} \smallsetminus \{0\}$ satisfy $b \mid a$ and $a \mid b$, then $b = \pm a$.

Indeed, we have $b = au$ and $a = bv$ for some $u, v \in \mathbf{Z}$, hence $a = bv = auv$ and $a(1 - uv) = 0$, which implies that $uv = 1$ (since $a \neq 0$) and therefore $u, v = \pm 1$.

1.1.6 Exercise. (1) Find all solutions $x, y \in \mathbf{N}$ of $x^2 - y^2 = n$ for $n = 20, 21$ and 22.
(2) If $x \in \mathbf{Z}$ is even (resp. odd), then $x^2 = 4k$ (resp. $x^2 = 4k + 1$) for some $k \in \mathbf{Z}$.
(3) $\{x^2 - y^2 \mid x, y \in \mathbf{Z}\} \subseteq (2\mathbf{Z} + 1) \cup 4\mathbf{Z}$.
(4) $\{x^2 - y^2 \mid x, y \in \mathbf{Z}\} = (2\mathbf{Z} + 1) \cup 4\mathbf{Z}$.
(5) Does $x^2 + y^2 = 1000003$ have a solution $x, y \in \mathbf{Z}$?

1.1.7 Exercise. Assume that $a \in \mathbf{Z}$ and $2 \nmid a$. Show that:
(1) $2 \mid (a \pm 1)$ and either $4 \mid (a - 1)$ or $4 \mid (a + 1)$.
(2) $a^2 - 1 = (a - 1)(a + 1)$ is divisible by 8, $a^4 - 1 = (a^2 - 1)(a^2 + 1)$ is divisible by 16,..., $a^{2^k} - 1$ is divisible by 2^{k+2} (for any $k \geq 1$).

1.1.8 Exercise. If p is a prime and $p \neq 2$ (resp. $p \neq 2, 3$), then $p = 4k \pm 1$ (resp. $p = 6k \pm 1$) for some $k \in \mathbf{N}_+$.
[Hint: write $p = 4k + a$ (resp. $p = 6k + a$).]

1.1.9 Proposition (Characterisation of prime numbers)**.** *An integer $a > 1$ is a prime number if and only if a has precisely two positive divisors (namely, 1 and a).*

Proof. Indeed, the existence of a positive divisor $b \mid a$ such that $b \neq 1$, a is equivalent to a factorisation $a = bc$, where $b, c \in \mathbf{N}_+$ and $b, c \neq 1$. □

1.1.10 Proposition (Primality criterion)**.** *For an integer $a > 1$, the following properties are equivalent.*
(1) a is not a prime number.
(2) There exists an integer b such that $1 < b \leq \sqrt{a}$ and $b \mid a$.
(3) There exists a prime number p such that $p \leq \sqrt{a}$ and $p \mid a$.

Proof. The implications (3) $\Longrightarrow$ (2) $\Longrightarrow$ (1) are automatic. If (1) holds, then $a = bc$ for some integers $1 < b \leq c$; then $b \mid a$ and $b^2 \leq bc = a$, hence $b \leq \sqrt{a}$, as required in (2). If (2) holds, then there exists a prime number p dividing b, by Proposition 1.2.1 below; then $p \mid a$ and $p \leq \sqrt{a}$. □

1.1.11 Examples of primes (1) Is $a = 89$ a prime number? Yes, it is, since $89 < 10^2$ and $2, 3, 5, 7 \nmid 89$ (here $\{2, 3, 5, 7\} = \{p \in \mathcal{P} \mid p < 10\}$).
(2) Is $a = 91$ a prime number? Again $91 < 10^2$ and $2, 3, 5 \nmid 91$, but $7 \mid 91 = 7 \cdot 13$; therefore 91 is not a prime.

1.1.12 Exercise. (1) If (a_n) and (b_n) $(n \geq 0)$ are two sequences of complex numbers such that $a_n = b_n - b_{n-1}$ holds for all $n \geq 1$, then

$$\forall n \geq 0 \quad \sum_{k=1}^{n} a_k = b_n - b_0.$$

(2) Compute $b_n - b_{n-1}$ for $b_n = n$, $b_n = n(n+1)$, $b_n = n(n+1)(n+2)$ etc.
(3) Compute

$$\sum_{k=1}^{n} k, \qquad \sum_{k=1}^{n} k(k+1), \qquad \sum_{k=1}^{n} k(k+1)(k+2), \qquad \sum_{k=1}^{n} k^2, \qquad \sum_{k=1}^{n} k^3.$$

1.2 Existence of factorisation

1.2.1 Proposition (Existence of factorisation)**.** *Every integer $n \geq 1$ is a product of (not necessarily distinct) prime numbers $n = p_1 \cdots p_r$ $(r \geq 0)$.*
[Note that $n = 1 \iff r = 0$. Here we use the convention that an empty product $x_1 \cdots x_r$ for $r = 0$ is equal to 1, in the same way that an empty sum $x_1 + \cdots + x_r$ for $r = 0$ is defined to be equal to 0.]

Proof. We argue by induction. If $n = 1$, then we can take $r = 0$. Assume that $n > 1$ and that the statement holds for all positive integers $m < n$. There are two possibilities:

Case 1. n is a prime. In this case $n = p_1$.

Case 2. n is not a prime. In this case $n = ab$ for some positive integers $a, b > 1$ (which implies that $a = n/b < n$ and $b = n/a < n$). By induction hypothesis, both $a = p_1 \cdots p_r$ and $b = q_1 \cdots q_s$ are products of primes, hence $n = p_1 \cdots p_r q_1 \cdots q_s$ is a product of primes, too. □

1.2.2 Remark on proofs by induction The proof of Proposition 1.2.1 used mathematical induction. In general, proofs by induction can be reformulated in terms of the following Minimality Principle for subsets of $\mathbf{N}$.

1.2.3 Theorem (Minimality Principle)**.** *Every non-empty subset $S \subset \mathbf{N}$ contains a minimal element (i.e., an element $a \in S$ such that there is no $b \in S$ satisfying $b < a$).*

1.2.4 Using the Minimality Principle Let us rephrase the above proof of Proposition 1.2.1 in the language of the Minimality Principle 1.2.3. Let $S \subset \mathbf{N}_+$ be the set of all positive integers for which 1.2.1 does not hold: $S := \{a \in \mathbf{N}_+ \mid a \neq p_1 \cdots p_r \ (r \geq 0)\}$. We want to show that $S = \emptyset$. Arguing by contradiction, assume that S is not empty. Therefore there exists a minimal element $a \in S$. It satisfies $a > 1$ (since $1 \notin S$) and $a \notin \mathcal{P}$ (since $S \cap \mathcal{P} = \emptyset$), hence there exist integers $b, c > 1$ such that $a = bc$, by Proposition 1.1.9. As $b, c < a$, the minimality of $a \in S$ implies that $b, c \notin S$. Therefore $a = p_1 \cdots p_r$ and $b = q_1 \cdots q_s$ are products of primes, and so is $a = p_1 \cdots p_r q_1 \cdots q_s$.

1.2.5 There are infinitely many primes A proof of the following fundamental result is given already in Euclid's Elements.

1.2.6 Theorem. *There are infinitely many prime numbers.*

Proof. It is enough to show that, for any given finite set of primes $A = \{p_1, \dots, p_r\}$ ($r \geq 0$), there exists a prime $p \notin A$. Consider the positive integer $N := 1 + p_1 \cdots p_r \geq 1 + 1 = 2$ (where $N = 2 \iff r = 0$). According to Proposition 1.2.1, there exists a prime number $p \mid N$. We claim that $p \notin A$. Indeed, if $p \in A$, then $p = p_i$ for some i, hence $p \mid (N-1)$. Therefore $p \mid N - (N-1)$, which is impossible. This contradiction shows that $p \notin A$, as claimed. □

1.2.7 Primes modulo 10 An inspection of tables of primes

$p = 10k+1$	11	31	41	61	71	101	131	151
$p = 10k+3$	3	13	23	43	53	73	83	103
$p = 10k+7$	7	17	37	47	67	97	107	127
$p = 10k+9$	19	29	59	79	89	109	139	149

seems to suggest that there are infinitely many primes in each of the four arithmetic progressions $10k+1$, $10k+3$, $10k+7$ and $10k+9$. This is, indeed, true, as a consequence of a general result due to Dirichlet (see Theorem 5.2.12). In Propositions 5.2.2 and 5.2.4 we show that there are infinitely many primes of the form $p = 4k+3$ and $p = 4k+1$, respectively.

A second look at the above table suggests that, in a given range $p \leq X$, primes of the form $10k+3$ and $10k+7$ seem to appear more frequently than those of the form $10k+1$ and $10k+9$. This kind of a phenomenon was first noticed by Čebyšev and is still being actively investigated under the technical term "Prime number races".

1.3 Factorisation of numbers $a^n \pm 1$

It is natural to ask whether there are any hidden rules governing factorisation of integers of special form. For example, one can consider integers $a^n \pm 1$, for fixed $a > 1$ and variable n.

1.3.1 Useful formulas Recall that

$$X^4 - 1 = (X-1)(X^3 + X^2 + X + 1), \qquad Y^3 + 1 = (Y+1)(Y^2 - Y + 1).$$

More generally,

$$X^m - 1 = (X-1)(X^{m-1} + X^{m-2} + \cdots + X + 1), \tag{1.3.1.1}$$

$$Y^{2k+1} + 1 = -((-Y)^{2k+1} - 1) = (Y+1)(Y^{2k} - Y^{2k-1} + \cdots + Y^2 - Y + 1). \tag{1.3.1.2}$$

1.3.2 Factorisation of Mersenne numbers $2^n - 1$ The following table contains factorisations of the **Mersenne numbers** $M_n = 2^n - 1$ for small values of n.

n	M_n
1	1
2	$\boxed{3}$
3	$\boxed{7}$
4	$15 = 3 \cdot 5$
5	$\boxed{31}$
6	$63 = 3 \cdot 3 \cdot 7$
7	$\boxed{127}$
8	$255 = 3 \cdot 5 \cdot 17$
9	$511 = 7 \cdot 73$
10	$1023 = 3 \cdot 11 \cdot 31$

We see that, for $1 \leq n \leq 10$, the number M_n is a prime if and only $n = 2, 3, 5, 7$ is a prime. Does this behaviour continue for $n > 10$?

For the next prime number $n = 11$ we have $M_{11} = 2047 < 46^2$. In order to test its primality, we need to check its (non)divisibility by primes $p < 46$. Indeed, $2, 3, 5, 7, 11, 13, 17, 19 \nmid M_{11}$, but $23 \mid M_{11} = 23 \cdot 89$. So M_{11} is not a prime!

Nevertheless, the converse holds:

1.3.3 Proposition. *If $M_n = 2^n - 1$ is a prime, so is n.*

Proof. We must show that M_n is not a prime if n is not a prime. In this case either $n = 1$ (when $M_n = 1$ is not a prime), or $n = ab$ for some integers a, $b > 1$. The formula (1.3.1.1) for $X = 2^a$ and $m = b$ shows that

$$M_n = 2^{ab} - 1 = (2^a - 1)(2^{a(b-1)} + 2^{a(b-2)} + \cdots + 2^a + 1)$$

is a non-trivial product (since $1 < 2^a - 1 < 2^{ab} - 1$), hence M_n is not a prime. □

1.3.4 Mersenne primes If p is a prime for which M_p is also a prime, we say that M_p is a **Mersenne prime**. This property can be checked using the Lucas–Lehmer criterion in Theorem (1.3.5) below or its variants. In 1876, Lucas used his criterion to show that M_{127} is a prime. At the time of writing (August 2019), the biggest known explicit prime is the prime M_p for $p = 82589933$ (M_p has 24 862 048 decimal digits). See `https://www.mersenne.org/`

1.3.5 Theorem (Lucas, Lehmer)**.** *Let $(a_n)_{n \geq 0}$ be the sequence of integers defined by $a_0 = 2$, $a_1 = 1$ and $a_{n+2} = a_{n+1} + a_n$.*

n	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
a_n	*2*	*1*	*3*	*4*	*7*	*11*	*18*	*29*	*47*	*76*	*123*	*199*	*322*	*521*	*843*	*1364*	*2207*

If $p = 4k + 3$ ($k \in \mathbf{Z}$) is a prime, then M_p is a prime if and only M_p divides $a_{2^{p-1}}$. [This can be reformulated in terms of the recursive formulae $a_{2^{n+1}} = a_{2^n}^2 - 2$ ($n \geq 1$).] A similar result holds for the sequence $(b_n)_{n \geq 0}$ given by $b_0 = 2$, $b_1 = 4$ and $b_{n+2} = 4b_{n+1} - b_n$ (for which $b_{2^{n+1}} = b_{2^n}^2 - 2$ ($n \geq 0$)), without the restriction $p = 4k + 3$.

1.3.6 Factorisation of $2^n + 1$ Let us now turn to factorisations of the numbers $A_n = 2^n + 1$.

n	A_n
1	$\boxed{3}$
2	$\boxed{5}$
3	$9 = 3 \cdot 3$
4	$\boxed{17}$
5	$33 = 3 \cdot 11$
6	$65 = 5 \cdot 13$
7	$129 = 3 \cdot 43$
8	$\boxed{257}$
9	$513 = 3^3 \cdot 19$
10	$1025 = 5^2 \cdot 41$

We see that, for $1 \leq n \leq 10$, the number A_n is a prime if and only $n = 1, 2, 4, 8$ is a power of 2. Does this behaviour continue for $n > 10$? Firstly, there is an analogue of Proposition 1.3.3.

1.3.7 Proposition. *If A_n is a prime, then $n = 2^k$ ($k \in \mathbf{N}$).*

Proof. If $n \in \mathbf{N}_+$ is not of the form $n = 2^k$ ($k \in \mathbf{N}$), then it can be written as $n = ab$ for some integers a, $b > 1$ with $2 \nmid b$. The formula (1.3.1.2) for $Y = 2^a$ and $m = b$ shows that

$$A_n = 2^{ab} + 1 = (2^a + 1)(2^{a(b-1)} - 2^{a(b-2)} + \cdots - 2^a + 1)$$

is a non-trivial product (since $1 < 2^a + 1 < 2^{ab} + 1$), hence A_n is not a prime. □

1.3.8 Fermat numbers Secondly, we need to examine the Fermat numbers $F_k = A_{2^k} = 2^{2^k} + 1$. If F_k is a prime, it is called a **Fermat prime**. The first five Fermat numbers in the following table are all Fermat primes.

k	0	1	2	3	4
F_k	3	5	17	257	65537

However, Euler showed that the next one $F_5 = 2^{32} + 1$ is divisible by 641, and therefore is not a prime. In fact, at the of time of writing (August 2019), it is not known whether any Fermat number F_k for $k > 4$ is a prime. What is known is that none of the F_k for $5 \leq k \leq 32$ is a prime.

1.3.9 Fermat primes and geometry One of the earliest major discoveries of Gauss (and perhaps his most treasured one) was a geometric construction of a regular polygon with 17 sides. It is no coincidence that the address of the Mathematical Research Institute in Berkeley is 17 Gauss Way!

More precisely, Gauss showed that, for any Fermat prime $p = F_k$ (in particular, for $p = 17$), one can construct a regular p-gon inscribed to a unit circle by a geometric construction involving only iterated intersections of lines and circles beginning with two points at a unit distance.

In general, it is known that such a geometric construction exists for a regular n-gon if and only if $n = 2^a p_1 \cdots p_r$, where $p_j = 2^{2^{k_j}} + 1$ are distinct Fermat primes.

1.3.10 Factorisation of $3^n \pm 1$ What is going on for numbers of the form $3^n \pm 1$? Here are a few numerical examples.

n	1	3	5	7
$(3^n - 1)/2$	1	13	$121 = 11^2$	1093
$(3^n + 1)/4$	1	7	61	547

n	2	4	6	8
$(3^n-1)/8$	1	$10 = 2 \cdot 5$	$91 = 7 \cdot 13$	$820 = 2^2 \cdot 5 \cdot 41$
$(3^n+1)/2$	5	41	$365 = 5 \cdot 73$	3281

1.3.11 Exercise. Let $n \geq 1$ be an integer.
(1) If $2 \nmid n$, then $2 \nmid (3^n - 1)/2$.
(2) If $2 \nmid n$ and if $(3^n - 1)/2$ is a prime, then $n = p$ is a prime.
(3) If $2 \mid n$, then $2 \nmid (3^n + 1)/2$.
(4) If $2 \mid n$ and if $(3^n + 1)/2$ is a prime, then $n = 2^k$.

1.4 Uniqueness of factorisation, Euclid's Lemma, p-adic valuations

1.4.1 Theorem (Uniqueness of factorisation)**.** *Every non-zero integer $a \in \mathbf{Z} \smallsetminus \{0\}$ can be written in the form $a = \pm p_1 \cdots p_r$, where $r \geq 0$ and $p_1, \ldots, p_r$ are (not necessarily distinct) prime numbers (and $\pm 1 = \operatorname{sgn}(a)$). If $a \operatorname{sgn}(a) = p_1 \cdots p_r = q_1 \cdots q_s$, where $s \geq 0$ and $q_1, \ldots, q_s$ are prime numbers (again, not necessarily distinct), then $r = s$ and, after a suitable renumbering of the indices $j = 1, \ldots, s$, we have $p_i = q_i$ for all $i = 1, \ldots, r = s$.*

Proof. We need to prove the uniqueness statement. If $r = 0$, then $a \operatorname{sgn}(a) = 1$, hence $s = 0$. Assume that $r > 0$ and that the statement holds for $r - 1$. The product $q_1 \cdots q_s$ is divisible by p_r. Euclid's Lemma 1.4.3 below implies that $p_r \mid q_j$ for some $1 \leq j \leq s$. After renumbering, we can assume that $p_r \mid q_s$. The only positive divisors of q_s are 1 and q_s; thus $p_r = q_s$. We can divide the equality $p_1 \cdots p_r = q_1 \cdots q_s$ by p_r, obtaining $p_1 \cdots p_{r-1} = q_1 \cdots q_{s-1}$. By induction hypothesis, we have (after renumbering the indices), $r - 1 = s - 1$ and $p_i = q_i$ for all $i = 1, \ldots, r - 1$. □

1.4.2 Theorem (Uniqueness of factorisation (equivalent formulation))**.** *Every non-zero integer $a \in \mathbf{Z} \smallsetminus \{0\}$ can be written uniquelly in the form $a = \pm p_1^{k_1} \cdots p_t^{k_t}$, where $p_1 < \cdots < p_t$ are prime numbers, $t \geq 0$ and $k_1, \ldots k_t \geq 1$.*

Proof. One rewrites the product in Theorem 1.4.1 by putting the primes p occurring there in an increasing order and collecting together the terms with the same value of p. This defines a canonical numbering of the primes in the product. □

1.4.3 Lemma (Euclid's Lemma)**.** *Let $a, b \in \mathbf{Z} \smallsetminus \{0\}$. If a prime number p satisfies $p \mid ab$ and $p \nmid b$, then $p \mid a$.*

Proof. We know that p divides ab and ap; we want to show that it also divides $a \cdot 1 = a$. It is natural, therefore, to investigate the set

$$X := \{x \in \mathbf{Z} \mid \ p \mid ax\}.$$

Note that $x \pm y \in X$ whenever $x, y \in X$. As $p, b \in X$, it follows that

$$X \supset p\mathbf{Z} + b\mathbf{Z} = \{pu + bv \mid u, v \in \mathbf{Z}\}.$$

A weak form of Bézout's theorem proved in Theorems 2.3.3 and Theorem 2.4.2 below states that $p\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ for some $d \in \mathbf{N}_+$. In particular, both $p, b \in p\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ are multiples of d. However, p has only two positive divisors, 1 and itself; thus $d = p$ or $d = 1$. The case $d = p$ is impossible, since b is not a multiple of p. Therefore $d = 1$, which implies that $1 = d \in X$ and $p \mid a \cdot 1$. □

1.5 p-adic valuations and their applications

1.5.1 Exponents in the prime factorisation The factorisation of $a \in \mathbf{Z} \smallsetminus \{0\}$ in Theorem 1.4.2 can be written as

$$a = \pm p_1^{k_1} \cdots p_t^{k_t} = \pm \prod_{p \in \mathcal{P}} p^{k(p)}, \qquad k(p) = \begin{cases} k_i & \text{if } p = p_i \\ 0 & \text{if } p \notin \{p_1, \ldots, p_t\}. \end{cases}$$

The exponents $k(p) \in \mathbf{N}$ are non-zero only for finitely many $p \in \mathcal{P}$, and they are uniquely determined by a, thanks to Theorem 1.4.2. As we shall see, all divisibility relations between non-zero integers can be stated in terms of these exponents. It is important, therefore, to give the exponents $k(p)$ a name.

1.5.2 Definition. Let $a \in \mathbf{Z} \smallsetminus \{0\}$. For each prime number $p \in \mathcal{P}$ the **p-adic valuation** $v_p(a)$ of a is defined as the exponent with which p occurs in the (unique) factorisation of a as a product of prime powers:

$$a = \operatorname{sgn}(a) \prod_{p \in \mathcal{P}} p^{v_p(a)} \tag{1.5.2.1}$$

($v_p(a) \in \mathbf{N}$, and $v_p(a) = 0$ for all but finitely many $p \in \mathcal{P}$). It is also useful to define $v_p(0) := +\infty$.
[For example, $\pm 56 = \pm 2^3 \cdot 7$, $v_2(\pm 56) = 3$, $v_7(\pm 56) = 1$, $v_p(\pm 56) = 0$ for all $p \neq 2, 7$.]

1.5.3 Warning Note that the definition of $v_p(a)$ makes sense only if we admit the truth of Theorem 1.4.2, which has not yet been fully proved in the previous sections. Its proof will be completed only when we establish Bézout's theorem (at least in its weak form).

1.5.4 Proposition (Properties of the p-adic valuation)**.** *Let $a, b, c \in \mathbf{Z} \smallsetminus \{0\}$. Then:*
(1) $a = \pm b \iff \forall p \in \mathcal{P} \quad v_p(a) = v_p(b)$.
(2) $\forall p \in \mathcal{P} \quad v_p(bc) = v_p(b) + v_p(c)$.
(3) $b \mid a \iff \forall p \in \mathcal{P} \quad v_p(b) \leq v_p(a)$.
(4) For each $p \in \mathcal{P}$ and $k \geq 0$, $p^k \mid a \iff k \leq v_p(a)$.
(5) $\forall p \in \mathcal{P} \quad v_p(a+b) \geq \min(v_p(a), v_p(b))$.
(6) If $v_p(a) \neq v_p(b)$, then $\forall p \in \mathcal{P} \quad v_p(a+b) = \min(v_p(a), v_p(b))$.

Proof. The parts (1) and (2) follow immediately from the definition (1.5.2.1). In part (3), if $b \mid a$, then $a = bc$ for some $c \in \mathbf{Z} \smallsetminus \{0\}$, hence $v_p(a) = v_p(b) + v_p(c) \geq v_p(b)$ holds for all $p \in \mathcal{P}$, by (2). Conversely, if $\forall p \in \mathcal{P} \quad v_p(b) \leq v_p(a)$, let

$$c := \operatorname{sgn}(b)^{-1} \operatorname{sgn}(a) \prod_{p \in \mathcal{P}} p^{v_p(a) - v_p(b)}.$$

All but finitely many exponents $v_p(a) - v_p(b) \geq 0$ are zero, which implies that $c \in \mathbf{Z} \smallsetminus \{0\}$ is well defined and satisfies $bc = \operatorname{sgn}(a) \prod_{p \in \mathcal{P}} p^{v_p(a)} = a$. Therefore $b \mid a$. Part (4) is a special case $b = p^k$ of (3). In (5) and (6) we can assume that $k := \min(v_p(a), v_p(b)) = v_p(a) \leq v_p(b)$. According to (4), $p^k \mid a$ and $p^k \mid b$, hence $p^k \mid (a+b)$, which is equivalent to $k \leq v_p(a+b)$. This proves (5). In (6) we have, in addition $k < v_p(b)$, and therefore $p^{k+1} \mid b$. If it were true that $v_p(a+b) > k$, then we would have $p^{k+1} \mid (a+b)$, hence also $p^{k+1} \mid (a+b) - b = a$, which is false. Therefore $v_p(a+b) \leq k$, which proves (6).

□

1.5.5 Proposition. *For any $a, b \in \mathbf{Z} \smallsetminus \{0\}$ we have:* $b \mid a \iff b^2 \mid a^2$.

Proof. The implication '$\Longrightarrow$' is automatic: if $a = bc$ for some $c \in \mathbf{Z}$, then $a^2 = b^2c^2$ and $c^2 \in \mathbf{Z}$, hence $b^2 \mid a^2$. However, the converse '$\Longleftarrow$' requires unique factorisation and some of its consequences proved in Proposition 1.5.4 (cf. the discussion around Proposition 1.5.20 below): if $b^2 \mid a^2$, then

$$\forall p \in \mathcal{P} \quad \underbrace{v_p(b^2)}_{2v_p(b)} \leq \underbrace{v_p(a^2)}_{2v_p(a)}, \text{ hence } \forall p \in \mathcal{P} \quad v_p(b) \leq v_p(a), \text{ and therefore } b \mid a.$$

□

1.5.6 Divisors of n If $n = p_1^{k_1} \cdots p_t^{k_t}$ is a prime factorisation of a given positive integer n, then the set of positive divisors of n is equal to

$$\{p_1^{l_1} \cdots p_t^{l_t} \mid \forall i = 1, \ldots, t \quad 0 \leq l_i \leq k_i\}.$$

There are $k_i + 1$ possible values of the exponent l_i, which means that the number of positive divisors of n is equal to

$$(k_1 + 1) \cdots (k_t + 1) = \prod_{p \mid n} (v_p(n) + 1)$$

(with the convention that p always denotes a prime number).

Example: $n = 12 = 2^2 \cdot 3^1$. The set of all positive divisors of 12 is equal to

$$\{2^i \cdot 3^j \mid 0 \leq i \leq 2,\ 0 \leq j \leq 1\} = \{1, 2, 2^2, 3, 2 \cdot 3, 2^2 \cdot 3\} = \{1, 2, 4, 3, 6, 12\}.$$

The sum of all positive divisors of 12 is equal to

$$1 + 2 + 2^2 + 3 + 2 \cdot 3 + 2^2 \cdot 3 = (1 + 2 + 2^2)(1 + 3) = 7 \cdot 4 = 28.$$

1.5.7 Exercise. For $n = p_1^{k_1} \cdots p_t^{k_t}$ as above, the sum of all positive divisors of n is equal to

$$\sigma_1(n) := \sum_{d \mid n} d = \sigma_1(p_1^{k_1}) \cdots \sigma_1(p_t^{k_t}) = (1 + p_1 + \cdots + p_1^{k_1}) \cdots (1 + p_t + \cdots + p_t^{k_t}) =$$

$$= \prod_{i=1}^{t} \left(\frac{p_i^{k_i+1} - 1}{p_i - 1} \right) = \prod_{p \mid n} \left(\frac{p^{v_p(n)+1} - 1}{p - 1} \right).$$

More generally, for any $s \in \mathbf{C}$, the sum of the s-th powers of all positive divisors of n is equal to

$$\sigma_s(n) := \sum_{d \mid n} d^s = \prod_{i=1}^{t} \sigma_s(p_i^{k_i}), \qquad \sigma_s(p^k) = 1 + p^s + \cdots + p^{ks} = \frac{p^{(k+1)s} - 1}{p - 1}.$$

Explicitly, determine the number and the sum of all positive divisors of $n = 2160$.

1.5.8 Perfect numbers The sum of all **proper** positive divisors (i.e., those less than n) of $n = 6$ is equal to $1 + 2 + 3 = 6 = n$. Similarly, for $n = 28$ we have $1 + 2 + 4 + 7 + 14 = 28 = n$.

Positive integers satisfying this property

$$\sum_{\substack{d \mid n \\ d \neq n}} d = n$$

(which is equivalent to $\sigma_1(n) = 2n$) are called **perfect numbers**. Conjecturally, no odd perfect numbers exist. Even perfect numbers are classified as follows.

1.5.9 Exercise. If $M_p = 2^p - 1$ is a Mersenne prime, then the number $n = 2^{p-1}(2^p - 1)$ is perfect. Conversely, any even perfect number is of this form.

1.5.10 Irrationality of $\sqrt[n]{a}$ (1) $\sqrt{3}$ is irrational: if $\sqrt{3} \in \mathbf{Q}$, then $\sqrt{3} = \frac{a}{b}$ for some $a, b \in \mathbf{N}_+$, hence $3b^2 = a^2$. It follows that, for every prime $p \in \mathcal{P}$,

$$v_p(3b^2) = v_p(a^2) \implies v_p(3) + 2v_p(b) = 2v_p(a).$$

For $p = 3$ we obtain $1 + 2v_3(b) = 2v_3(a)$ and $1 = 2(v_3(a) - v_3(b))$, which is impossible, since $2 \nmid 1$.
(2) $\alpha = \frac{5+\sqrt[3]{20}}{7}$ is irrational: if $\alpha \in \mathbf{Q}$, then $\sqrt[3]{20} = 7\alpha - 5 \in \mathbf{Q}$. Writing $\sqrt[3]{20} = \frac{a}{b}$ with $a, b \in \mathbf{N}_+$, we obtain $20b^3 = a^3$ and

$$\forall p \in \mathcal{P} \quad v_p(20b^3) = v_p(a^3),$$

hence, as above, $v_p(20) = 3(v_p(a) - v_p(b))$ must divisible by 3. However, this is false for $p = 2, 5$, since $20 = 2^2 \cdot 5$: $v_2(20) = 2$ and $v_5(20) = 1$. So, taking $p = 2$ or $p = 5$, we obtain a contradiction.

The general result is the following.

1.5.11 Theorem (Irrationality of $\sqrt[n]{a}$)**.** *Let $a, n \in \mathbf{N}_+$. The following properties are equivalent:*
(1) $\sqrt[n]{a} \in \mathbf{Q}$.
(2) $\forall p \in \mathcal{P} \quad n \mid v_p(a)$.
(3) There exists $b \in \mathbf{N}_+$ such that $a = b^n$.
(4) $\sqrt[n]{a} \in \mathbf{Z}$.

Proof. The implications (3) $\iff$ (4) $\implies$ (1) are automatic, whereas (3) $\implies$ (2) follows from $v_p(a) = v_p(b^n) = nv_p(b)$. Conversely, if (2) holds, then $b := \prod_{p \in \mathcal{P}} p^{v_p(a)/n} \in \mathbf{N}_+$ and $b^n = a$, proving (3). Finally, if (1) holds, then $\sqrt[n]{a} = \frac{b}{c}$ for some $b, c \in \mathbf{N}_+$ such that $ac^n = b^n$. This implies, for each $p \in \mathcal{P}$, that $v_p(ac^n) = v_p(a) + nv_p(c)$ is equal to $v_p(b^n) = nv_p(b)$, hence $v_p(a) = n(v_p(b) - v_p(c))$ is divisible by n. □

1.5.12 Corollary. *If $a, n \in \mathbf{N}_+$ and if there exists $k \in \mathbf{N}_+$ such that $k^n < a < (k+1)^n$, then $\sqrt[n]{a} \notin \mathbf{Z}$, hence $\sqrt[n]{a} \notin \mathbf{Q}$.*

1.5.13 Theorem (Uniqueness of factorisation in $\mathbf{Q}$)**.** *Every non-zero rational number $a \in \mathbf{Q} \smallsetminus \{0\}$ can be written in a unique way as a product*

$$a = \operatorname{sgn}(a) \prod_{p \in \mathcal{P}} p^{k(p)},$$

where the exponents $k(p) \in \mathbf{Z}$ are non-zero only for finitely many $p \in \mathcal{P}$. For each $p \in \mathcal{P}$ we define the p-adic valuation of a to be the integer $v_p(a) := k(p) \in \mathbf{Z}$. [For example, $a = \frac{15}{100} = \frac{3^1 \cdot 5^1}{2^2 \cdot 5^2} = \frac{3}{20} = \frac{3}{2^2 \cdot 5^1} = 2^{-2} \cdot 3^1 \cdot 5^{-1}$ and $v_2(a) = -2$, $v_3(a) = 1$, $v_5(a) = -1$ and $v_p(a) = 0$ for $p \neq 2, 3, 5$.]

Proof. **Existence:** choose $b, c \in \mathbf{Z} \smallsetminus \{0\}$ such that $a = b/c$; then

$$a = \operatorname{sgn}(b) \prod_{p \in \mathcal{P}} p^{v_p(b)} / \operatorname{sgn}(c) \prod_{p \in \mathcal{P}} p^{v_p(c)} = \operatorname{sgn}(a) \prod_{p \in \mathcal{P}} p^{v_p(b) - v_p(c)}.$$

Uniqueness: assume that

$$\prod_{p \in \mathcal{P}} p^{k(p)} = \prod_{p \in \mathcal{P}} p^{l(p)}, \tag{1.5.13.1}$$

where the integers $k(p), l(p) \in \mathbf{Z}$ are non-zero only for p lying in a finite subset $S \subset \mathcal{P}$. After multiplying (1.5.13.1) by $\prod_{p \in S} p^n$, where $n \geq \max_{p \in S}(|k(p)| + |l(p)|$, we obtain

$$\prod_{p\in\mathcal{P}} p^{k(p)+n} = \prod_{p\in\mathcal{P}} p^{l(p)+n},$$

where $k(p)+n, l(p)+n \geq 0$ for all $p \in S$. The uniqueness statement in Theorem 1.4.2 implies that $k(p)+n = l(p)+n$ (hence $k(p) = l(p)$) for all $p \in S$. □

1.5.14 Exercise. The statements (1), (2), (5) and (6) in Proposition 1.5.4 hold for $a, b \in \mathbf{Q} \smallsetminus \{0\}$.

1.5.15 Exercise. The number $(4/7)^{4/7}$ is irrational.

1.5.16 Exercise. Let $a \in \mathbf{Q}$, $a > 0$ and $n \in \mathbf{N}_+$. It is equivalent: $\sqrt[n]{a} \in \mathbf{Q} \iff \forall p \in \mathcal{P} \quad n \mid v_p(a)$.

1.5.17 Digression on subrings of C Is it possible to prove the implication $b^2 \mid a^2 \Longrightarrow b \mid a$ in Proposition 1.5.5 without appealing to the uniqueness of factorisation? The answer is "NO". The point is that one can define divisibility in more general domains that in $\mathbf{Z}$, and in some of them the above implication fails to hold, as we are going to see.

1.5.18 Definition. A **subring** of $\mathbf{C}$ is a subset $A \subset \mathbf{C}$ such that

$$0, 1 \in A; \qquad a, b \in A \Longrightarrow a \pm b,\ ab \in A.$$

For $a, b \in A$ we define $b \mid a$ (b divides a) if there exists $c \in A$ such that $a = bc$.
A **subfield** of $\mathbf{C}$ is a subring $A \subset \mathbf{C}$ such that

$$a \in A \smallsetminus \{0\} \Longrightarrow a^{-1} \in A.$$

An **additive subgroup** of $\mathbf{C}$ is a subset $X \subset \mathbf{C}$ such that

$$0 \in X; \qquad x, y \in X \Longrightarrow x \pm y \in X.$$

In particular, any subring of $\mathbf{C}$ is an additive subgroup of $\mathbf{C}$.

1.5.19 Examples of subrings and additive subgroups of C (1) $X = \frac{1}{2}\mathbf{Z}$ is an additive subgroup of $\mathbf{C}$, but it is not a subring of $\mathbf{C}$ (since $\frac{1}{2} \cdot \frac{1}{2} \notin \frac{1}{2}\mathbf{Z}$).
(2) $A = \mathbf{Z}$ is a subring of $\mathbf{C}$.
(3) $A = \mathbf{Z} + i\mathbf{Z} = \{u + iv \mid u, v \in \mathbf{Z}\}$ is a subring of $\mathbf{C}$.
(4) $A = \mathbf{Z} + 2i\mathbf{Z} = \{u + 2iv \mid u, v \in \mathbf{Z}\}$ is a subring of $\mathbf{C}$.
(5) $A = \mathbf{Q} + i\mathbf{Q} = \{u + iv \mid u, v \in \mathbf{Q}\}$ is a subfield of $\mathbf{C}$ (since $(u+iv)^{-1} = (u-iv)/(u^2+v^2)$).

1.5.20 Proposition. *(1) If $A \subset \mathbf{C}$ is a subring, $a, b \in A$ and $b \mid a$, then $b^2 \mid a^2$.*
(2) In the subring $A = \mathbf{Z} + 2i\mathbf{Z}$, consider $b = 2$ and $a = 2i$. Then $b^2 \mid a^2$, but $b \nmid a$ in A.

Proof. (1) is proved as in Proposition 1.5.5. In part (2), $2^2 = 4$ divides $(2i)^2 = -4$ (since $-1 \in A$), but 2 does not divide $2i$ (since $i \notin A$). □

1.5.21 Metaremark In general, "easy" implications involving divisibility, such as $b \mid a \Longrightarrow b^2 \mid a^2$, hold in any subring of $\mathbf{C}$. This means that if we find a subring of $\mathbf{C}$ in which a certain implication involving divisibility (such as $b^2 \mid a^2 \Longrightarrow b \mid a$) does not hold, as in Proposition 1.5.20(2) above, then even in the case when such an implication happens to hold in $\mathbf{Z}$, its proof cannot follow from basic properties of divisibility, such as those listed in Section 1.1.5.

1.5.22 Exercise. Let $a, b \geq 1$ be positive integers. Which of the following statements are true, and why:
(1) if $b^2 \mid a^3$, then $b \mid a$.
(2) if $b^3 \mid a^2$, then $b \mid a$.
(3) if $b^3 \mid a^3$, then $b \mid a$.

1.5.23 Exercise. (1) What is the smallest additive subgroup of $\mathbf{C}$ containing $\frac{1}{2}$ (resp. $\frac{i}{2}$)?
(2) What is the smallest subring of $\mathbf{C}$ containing $\frac{1}{2}$ (resp. $\frac{i}{2}$)?

1.5.24 Binomial coefficients In Exercise 2.6.6 we are going to discuss divisibility of binomial coefficients by various primes. Here we consider the simplest possible case.

1.5.25 Proposition. *If p is a prime number, then $\binom{p}{j}$ is divisible by p, for all integers $0 < j < p$.*

Proof. The formula

$$\binom{p}{j} = \frac{p(p-1)\cdots(p-j+1)}{1 \cdot 2 \cdots j}$$

implies that

$$v_p(\binom{p}{j}) = v_p(p) + \sum_{i=1}^{j-1} v_p(p-i) - \sum_{i=1}^{j} v_p(i) = 1 + 0 - 0 = 1.$$

□

1.5.26 Theorem (Fermat's little theorem)**.** *If $p \in \mathcal{P}$ and $a \in \mathbf{Z}$, then $p \mid (a^p - a)$.*

Proof. The binomial formula implies that

$$(a+1)^p - (a+1) = (a^p - a) + R, \qquad R = \sum_{0<j<p} \binom{p}{j} a^j.$$

According to Proposition 1.5.25, the integer R is divisible by p. Consequently, the statement of the theorem holds for $a+1$ if and only if it holds for a. As it is trivially true for $a = 0$, it is true for all $a \in \mathbf{Z}$, by induction. □

1.5.27 Exercise. Prove Fermat's little theorem for $p = 2$ and $p = 3$ directly, using the formulas $a^2 - a = a(a-1)$ and $a^3 - a = (a-1)a(a+1)$. What about the case $p = 5$, when $a^5 - a = (a-1)a(a+1)(a^2+1)$?

1.6 The greatest common divisor, the least common multiple

1.6.1 Common divisors (example) For $a = 44 = 2^2 \cdot 11$ and $b = 16 = 2^4$ we have

$$\{\text{divisors of } 44\} = \{\pm 1, \pm 2, \pm 4, \pm 11, \pm 22, \pm 44\}, \quad \{\text{divisors of } 16\} = \{\pm 1, \pm 2, \pm 4, \pm 8, \pm 16\}$$
$$\{\text{common divisors of 44 and 16}\} := \{\text{divisors of } 44\} \cap \{\text{divisors of } 16\} =$$
$$= \{\pm 1, \pm 2, \pm 4\} = \{\text{divisors of } 4\}.$$

We see that 4 is the **greatest common divisor** of 44 and 16 **with respect to divisibility**: it is a common divisor of 44 and 16, and it is divisible by all common divisors of these two numbers. It is denoted by

$$\gcd(44, 16) = 4.$$

1.6.2 Theorem (Existence and uniqueness of the gcd)**.** *Let $a, b \in \mathbf{Z} \smallsetminus \{0\}$.*
(1) There exists a unique positive integer $d \in \mathbf{N}_+$ such that
(a) $d \mid a$ and $d \mid b$;
(b) if $c \in \mathbf{Z} \smallsetminus \{0\}$, $c \mid a$ and $c \mid b$, then $c \mid d$.
We denote it by $\gcd(a, b)$ *(the greatest common divisor of a and b).*
(2) If $a = \pm \prod_{p \in \mathcal{P}} p^{v_p(a)}$ and $b = \pm \prod_{p \in \mathcal{P}} p^{v_p(b)}$, then $d = \gcd(a, b) = \prod_{p \in \mathcal{P}} p^{\min(v_p(a), v_p(b))}$.
(3) If $\gcd(a, b) = 1$, we say that the integers a and b are ***relatively prime****.*

Proof. (1) The uniqueness is easy: if $d, d' \in \mathbf{N}_+$ satisfy (a) and (b), then $d \mid d'$ and $d' \mid d$, hence $d' = \pm d$. As both of them are positive, $d' = d$. We are going to prove the existence of $\gcd(a, b)$ as a consequence of the formula (2), which in turn depends on the unique factorisation property from Theorem 1.4.2 (which has not yet been fully proved in the previous sections). In Section 2.3 we give another proof of the existence of $\gcd(a, b)$ which uses (a weak form of) Bézout's theorem $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$, but does not rely on the uniqueness of factorisation.
(2) According to Proposition 1.5.4(3),

$$\{\text{divisors of } a\} = \Big\{ \pm \prod_{p \in \mathcal{P}} p^{c_p} \mid \forall p \in \mathcal{P} \quad 0 \le c_p \le v_p(a) \Big\}$$
$$\{\text{divisors of } b\} = \Big\{ \pm \prod_{p \in \mathcal{P}} p^{c_p} \mid \forall p \in \mathcal{P} \quad 0 \le c_p \le v_p(b) \Big\},$$

which implies that

$$\{\text{divisors of } a\} \cap \{\text{divisors of } b\} = \Big\{ \pm \prod_{p \in \mathcal{P}} p^{c_p} \mid \forall p \in \mathcal{P} \quad 0 \le c_p \le \min(v_p(a), v_p(b)) \Big\} =$$
$$= \{\text{divisors of } d\},$$

where $d = \prod_{p \in \mathcal{P}} p^{\min(v_p(a), v_p(b))}$. □

1.6.3 Theorem (Existence and uniqueness of the lcm)**.** *Let $a, b \in \mathbf{Z} \smallsetminus \{0\}$.*
(1) There exists a unique positive integer $m \in \mathbf{N}_+$ such that
(a) $a \mid m$ and $b \mid m$;
(b) if $c \in \mathbf{Z} \smallsetminus \{0\}$, $a \mid c$ and $b \mid c$, then $m \mid c$.
We denote it by $\operatorname{lcm}(a, b)$ *(the least common multiple of a and b).*
(2) If $a = \pm \prod_{p \in \mathcal{P}} p^{v_p(a)}$ and $b = \pm \prod_{p \in \mathcal{P}} p^{v_p(b)}$, then $m = \operatorname{lcm}(a, b) = \prod_{p \in \mathcal{P}} p^{\max(v_p(a), v_p(b))}$.

Proof. (1) Again, if $m' \in \mathbf{N}_+$ also satisfies (a) and (b), then $m \mid m'$ and $m' \mid m$, hence $m' = \pm m$ and $m' = m$, by positivity. The existence follows from (2), which is proved (using the uniqueness of factorisation) as in Theorem 1.6.2, except that the inequalities go in the opposite direction:

$$\{\text{multiples of } a\} \cap \{\text{multiples of } b\} = \Big\{ \pm \prod_{p \in \mathcal{P}} p^{c_p} \mid \forall p \in \mathcal{P} \quad c_p \ge \max(v_p(a), v_p(b)) \Big\} =$$
$$= \{\text{multiples of } m\}.$$

□

1.6.4 Characterisation of gcd and lcm To sum up, $d = \gcd(a,b), m = \mathrm{lcm}(a,b) \in \mathbf{N}_+$ are characterised by the following properties:

$$\{\text{divisors of } a\} \cap \{\text{divisors of } b\} = \{\text{divisors of } d\},$$
$$\{\text{multiples of } a\} \cap \{\text{multiples of } b\} = \{\text{multiples of } m\}.$$

1.6.5 Exercise. (1) $\forall x, y \in \mathbf{R} \quad \min(x,y) + \max(x,y) = x + y$.
(2) $\forall a, b \in \mathbf{Z} \smallsetminus \{0\} \quad \gcd(a,b)\mathrm{lcm}(a,b) = |ab|$.

1.6.6 Example of $\gcd(a,b)$ and $\mathrm{lcm}(a,b)$ For $a = 50 = 2^1 \cdot 5^2$ and $b = 15 = 3^1 \cdot 5^1$ we have $\gcd(a,b) = 5^1 = 5$ and $\mathrm{lcm}(a,b) = 2^1 \cdot 3^1 \cdot 5^2 = 150$.

1.6.7 Exercise. If $a_1, \ldots, a_r \in \mathbf{Z} \smallsetminus \{0\}$ $(r \geq 2)$, then

$$\{\text{common divisors of } a_1, \ldots, a_r\} = \{\text{divisors of } d = \gcd(a_1, \ldots, a_r)\},$$
$$\{\text{common multiples of } a_1, \ldots, a_r\} = \{\text{multiples of } m = \mathrm{lcm}(a_1, \ldots, a_r)\},$$

where

$$d = \prod_{p \in \mathcal{P}} p^{\min(v_p(a_1), \ldots, v_p(a_r))}, \qquad m = \prod_{p \in \mathcal{P}} p^{\max(v_p(a_1), \ldots, v_p(a_r))}.$$

Warning: for $r > 2$ there is no general relation between d and m.

1.6.8 Example of $\gcd(a,b,c)$ and $\mathrm{lcm}(a,b,c)$ The greatest common divisor $\gcd(6,10,15)$ is equal to $\gcd(\gcd(6,10),15) = \gcd(2,15) = 1$ (alternatively, it is equal to $\gcd(6, \gcd(10,15)) = \gcd(6,5) = 1$). We can also use the factorisations $6 = 2^1 \cdot 3^1$, $10 = 2^1 \cdot 5^1$ and $15 = 3^1 \cdot 5^1$ to compute $\gcd(6,10,15) = 1$ and $\mathrm{lcm}(6,10,15) = 2^1 \cdot 3^1 \cdot 5^1 = 30$.

1.6.9 Exercise. Let $a, b \in \mathbf{N}_+$. Show that there exist positive divisors $a' \mid a$ and $b' \mid b$ such that $\gcd(a', b') = 1$ and $a'b' = \mathrm{lcm}(a,b)$.
[For example, for $a = 15$ and $b = 20$ we can take either $a' = 15$ and $b' = 4$, or $a' = 3$ and $b' = 20$.]

2 Euclid's algorithm, Bézout's theorem

2.1 A preview

2.1.1 General description Euclid's algorithm is of fundamental importance not only in arithmetic, but in other contexts, too, as we shall see in Section 10. It consists of a repeated application of division with remainder.

Its input is a pair of non-zero integers $a, b \in \mathbf{Z} \smallsetminus \{0\}$. Its output is a non-zero integer d satisfying

$$a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z} = |d|\mathbf{Z} \tag{2.1.1.1}$$

(a **weak form of Bézout's theorem**). In fact, the algorithm also produces an explicit pair of integers $u, v \in \mathbf{Z}$ such that

$$au + bv = d \tag{2.1.1.2}$$

(a **Bézout relation**). It is easy to deduce from (2.1.1.1) that

$$|d| = \gcd(a, b)$$

in the sense that $|d| \in \mathbf{N}_+$ is a positive common divisor of a and b that is divisible by all common divisors of a and b.

2.1.2 Bézout's theorem This gives a direct proof of the existence of $\gcd(a, b)$, as well as a more precise version of (2.1.1.1):

$$a\mathbf{Z} + b\mathbf{Z} = \gcd(a, b)\mathbf{Z} \tag{2.1.2.1}$$

(**Bézout's theorem**; not to be confused with Bézout's theorem about intersection of two planar algebraic curves).

All of this is important not only from a theoretical point of view (as we saw in the proof of Lemma 1.4.3, the equality (2.1.1.1) implies Euclid's Lemma), but also for explicit calculations of $\gcd(a, b)$ and u, v in (2.1.1.2).

2.1.3 Subgroups of Z By induction, one obtains from (2.1.2.1) that

$$a_1\mathbf{Z} + \cdots + a_r\mathbf{Z} = \gcd(a_1, \ldots, a_r)\mathbf{Z}, \tag{2.1.3.1}$$

for any $r \geq 2$ and $a_i \in \mathbf{Z} \smallsetminus \{0\}$.

Note that the term on the left hand side of (2.1.3.1) is an additive subgroup of $\mathbf{Z}$ (it contains 0 and, together with any pair of elements x and y, also $x \pm y$).

It turns out that a more abstract version of (2.1.3.1) holds: **any additive subgroup** X of $\mathbf{Z}$ is of the form $X = d\mathbf{Z}$, for unique $d \in \mathbf{N}$ ("all subgroups of $\mathbf{Z}$ are cyclic", in the abstract language of Definition 7.2.7 below). In Theorem 2.3.2 we are going to prove this abstract statement directly, using division with remainder only once (not repeatedly, as in Euclid's algorithm).

2.2 Division with remainder, Euclid's algorithm (example)

2.2.1 Example: division of 44 **by** 16 Even though 16 does not divide 44, we can write

$$44 = 2 \cdot 16 + 12, \qquad 0 \leq 12 < 16. \tag{2.2.1.1}$$

In this equality, 2 is the **quotient** and 12 is the **remainder** of the division of 44 by 16. The formula (2.2.1.1) is equivalent to

$$\frac{44}{16} = 2 + \frac{12}{16}. \tag{2.2.1.2}$$

The general case is as follows.

2.2.2 Proposition (Division with remainder)**.** *Let $a, b \in \mathbf{Z} \smallsetminus \{0\}$. Then there exists a unique pair of integers $q, r \in \mathbf{Z}$ such that*

$$a = qb + r, \qquad 0 \leq r < |b|$$

*(q is the **quotient** and r is the **remainder** of the division of a by b). As in* (2.2.1.2)*, we obtain*

$$\frac{a}{b} = q + \frac{r}{b}.$$

Proof. **Uniqueness:** if $a = bq + r = bq' + r'$ and $0 \leq r, r' < |b|$, then $r' - r = b(q - q')$ and

$$-|b| = 0 - |b| \leq r' - |b| < r' - r = b(q - q') < |b| - r \leq |b| - 0 = |b|.$$

After dividing by $|b| > 0$, we obtain $-1 < q - q' < 1$. The only integer satisfying this inequality is $q - q' = 0$, which implies that $q = q'$ and $r = r'$.
Existence: apply the Minimality Principle 1.2.3 to the subset

$$S := \mathbf{N} \cap \{a - qb \mid q \in \mathbf{Z}\} = \mathbf{N} \cap (a + b\mathbf{Z}) \subset \mathbf{N}$$

(S is non-empty, since $0 \leq a + |a| \leq a + |a||b| \in S$). Let $r = a - qb \in S$ be a minimal element of S; then $r \geq 0$. If $r \geq |b|$, then $0 \leq r - |b| = a - (q + \operatorname{sgn}(b))b \in S$, which contradicts the minimality of r. Therefore $0 \leq r < |b|$ and $a = qb + r$, as required. □

2.2.3 Euclid's algorithm (example) Euclid's algorithm for a pair of non-zero integers (a, b) first computes the the quotient q and the remainder r of the division of a by b, and the replaces the pair (a, b) by (b, r). The procedure is repeated until we obtain a pair $(d, 0)$.

For example, for $a = 44$ and $b = 16$ we obtain

$$\begin{aligned} 44 &= 2 \cdot 16 + 12 \\ 16 &= 1 \cdot 12 + 4 \\ 12 &= 3 \cdot 4 + 0 \\ 4 &= d \end{aligned}$$

Going through these equalities from the bottom to the top, we see inductively that $d = 4$ divides all the numbers appearing on the left hand side, namely,

$$12 = 3 \cdot 4, \qquad 16 = 1 \cdot 12 + 4, \qquad 44 = 2 \cdot 16 + 12.$$

In particular, 4 is a common divisor of 44 and 16. Conversely, going from the top to the bottom, we see that we can write all numbers on the left hand side as integral linear combinations of 44 and 16:

$$44 = 1 \cdot 44 + 0 \cdot 16 \tag{2.2.3.1}$$
$$16 = 0 \cdot 44 + 1 \cdot 16 \tag{2.2.3.2}$$
$$12 = 1 \cdot 44 + (-2) \cdot 16 \tag{2.2.3.3}$$
$$4 = 16 - 1 \cdot 12 = 16 - (44 - 2 \cdot 16) = (-1) \cdot 44 + 3 \cdot 16 \tag{2.2.3.4}$$

In more scientific terms

$$4 \mid 44, \quad 4 \mid 16 \tag{2.2.3.5}$$

$$4 \in 44\mathbf{Z} + 16\mathbf{Z} \tag{2.2.3.6}$$

We claim that the conjunction of the two statements (2.2.3.5) and (2.2.3.6) is equivalent to

$$44\mathbf{Z} + 16\mathbf{Z} = 4\mathbf{Z} \tag{2.2.3.7}$$

Indeed, (2.2.3.5) is equivalent to

$$44\mathbf{Z} \subseteq 4\mathbf{Z}, \quad 16\mathbf{Z} \subseteq 4\mathbf{Z},$$

hence it implies that $44\mathbf{Z} + 16\mathbf{Z} \subseteq 4\mathbf{Z}$, whereas (2.2.3.6) implies the reverse inclusion $4\mathbf{Z} \subseteq 44\mathbf{Z} + 16\mathbf{Z}$.

Conversely, the Bézout relation (2.2.3.7) implies, on one hand, (2.2.3.6), and on the other hand that

$$4\mathbf{Z} = 44\mathbf{Z} + 16\mathbf{Z} \supseteq 44\mathbf{Z}, \quad 4\mathbf{Z} = 44\mathbf{Z} + 16\mathbf{Z} \supseteq 16\mathbf{Z},$$

which is equivalent to (2.2.3.5).

Moreover, if $c \in \mathbf{Z} \smallsetminus \{0\}$ is any common divisor of 44 and 16, then it divides all elements of $44\mathbf{Z} + 16\mathbf{Z} = 4\mathbf{Z}$; in particular, c divides 4.

Therefore (2.2.3.7) directly implies that 4 has the properties (a) and (b) in Theorem 1.6.2(1) (for $a = 44$ and $b = 16$), hence $4 = \gcd(44, 16)$ and

$$44\mathbf{Z} + 16\mathbf{Z} = \gcd(44, 16)\mathbf{Z}. \tag{2.2.3.8}$$

2.2.4 Euclid's algorithm and continued fractions (example) In the above example one can iterate the formula (2.2.1.2) to obtain

$$\frac{44}{16} = 2 + \frac{12}{16}, \qquad \frac{16}{12} = 1 + \frac{4}{12}, \qquad \frac{12}{4} = 3,$$

hence

$$\frac{44}{16} = 2 + \frac{1}{1 + \dfrac{1}{3}} = 2 + \frac{3}{4} = \frac{11}{4}. \tag{2.2.4.1}$$

This is a basic example of a (finite) continued fraction. It turns out that one can determine the coefficients in the linear combination that was obtained from Euclid's algorithm

$$4 = 3 \cdot 16 - 1 \cdot 44 \tag{2.2.4.2}$$

by taking the continued fraction (2.2.4.1) and deleting from it the last term:

$$2 + \frac{1}{1} = \frac{3}{1}. \tag{2.2.4.3}$$

Similarly, the coefficients in the linear combination

$$12 = (-2) \cdot 16 + 1 \cdot 44 \tag{2.2.4.4}$$

are obtained by deleting from (2.2.4.1) the last two terms:

$$2 = \frac{2}{1}. \tag{2.2.4.5}$$

In matrix form, we can rewrite the above formulas as follows:

$$\begin{pmatrix}-16 & 44\end{pmatrix}\begin{pmatrix}0 & 1 & 2 & 3 & 11\\ 1 & 0 & 1 & 1 & 4\end{pmatrix} = \begin{pmatrix}44 & -16 & 12 & -4 & 0\end{pmatrix}. \tag{2.2.4.6}$$

All of the above makes sense if we replace 44 and 16 by an arbitrary pair of non-zero integers. This will be explained in detail in Section 2.4 below.

2.3 Subgroups of Z, Bézout's theorem

2.3.1 Subgroups of Z Recall that a subset $X \subset \mathbf{Z}$ is an (additive) subgroup of $\mathbf{Z}$ if $0 \in X$ and if $x \pm y \in X$ for all $x, y \in X$. Note that:

- If $a_i \in \mathbf{Z}$, then $X = a_1\mathbf{Z} + \cdots + a_r\mathbf{Z} = \{a_1x_1 + \cdots + a_rx_r \mid x_i \in \mathbf{Z}\}$ is a subgroup of $\mathbf{Z}$.
- In particular, $d\mathbf{Z} = (-d)\mathbf{Z}$ is a subgroup of $\mathbf{Z}$, for any $d \in \mathbf{Z}$.
- If $X \subset \mathbf{Z}$ is a subgroup and $x_1, \ldots, x_r \in X$, then $a_1x_1 + \cdots + a_rx_r \in X$, for any $a_i \in \mathbf{Z}$ (in other words, $\mathbf{Z}x_1 + \cdots + \mathbf{Z}x_r \subset X$).

2.3.2 Theorem (Structure of subgroups of $\mathbf{Z}$)**.** *If $X \subset \mathbf{Z}$ is a subgroup, then there exists $d \in \mathbf{N}$ (unique) such that $X = d\mathbf{Z}$.*

Proof. If $X = \{0\}$, then $d = 0$. Assume $X \neq \{0\}$. We want to show that $X = d\mathbf{Z}$ for some $d \in \mathbf{N}_+$. In this case $d = \min\{|a| \mid 0 \neq a \in d\mathbf{Z}\}$, so we are going to define d in the same way, replacing $d\mathbf{Z}$ by X. As the subset $S := \{|a| \mid 0 \neq a \in X\} \subset \mathbf{N}_+$ is non-empty, the Minimality Principle 1.2.3 tells us that there exists a minimal element $d \in S$. We claim that $X = d\mathbf{Z}$. Firstly, X contains d or $-d$ (or both), and hence $d\mathbf{Z} = (-d)\mathbf{Z} \subseteq X$, by the last point in 2.3.1. If there were $a \in X$ such that $a \notin d\mathbf{Z}$, then we could subtract from a a suitable multiple of d to obtain a non-zero element of X with smaller absolute value than d: indeed, if we perform the division with remainder of a by d, we obtain $a = qd + r$, where $q, r \in \mathbf{Z}$, $0 \leq r < d$ and $r \neq 0$ (since $a \notin d\mathbf{Z}$). We know that $a \in A$ and $qd \in d\mathbf{Z} \subset X$; thus $0 \neq r = a - qd \in X$ and $|r| = r < d$. This is a contradiction with the minimality of d. Therefore no such a exists and $X = d\mathbf{Z}$. □

2.3.3 Theorem (Bézout's theorem)**.** *If $a, b \in \mathbf{Z} \smallsetminus \{0\}$, then there exists a unique positive integer $d \in \mathbf{N}_+$ such that $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$. It satisfies (a) and (b) from Theorem 1.6.2(1):*
(a) $d \mid a$ and $d \mid b$;
(b) if $c \in \mathbf{Z} \smallsetminus \{0\}$, $c \mid a$ and $c \mid b$, then $c \mid d$.
In other words, $d = \gcd(a, b)$ and $a\mathbf{Z} + b\mathbf{Z} = \gcd(a, b)\mathbf{Z}$.

Proof. The existence and uniqueness of $d \in \mathbf{N}_+$ such that $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ is a special case of Theorem 2.3.2, for $X = a\mathbf{Z} + b\mathbf{Z}$. Properties (a) and (b) are checked as in 2.2.3: firstly, $a\mathbf{Z} \subseteq a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$, hence $d \mid a$ (and similarly for $d \mid b$). Secondly, if $c \in \mathbf{Z} \smallsetminus \{0\}$ divides both a and b, then $c\mathbf{Z} \subset \mathbf{Z}$ is a subgroup containing a and b, hence $c\mathbf{Z} \supseteq a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ (by the last point in 2.3.1), which means that $c \mid d$. □

2.3.4 Remarks The above proof of the existence of $\gcd(a, b)$ is unconditional. It does not rely on the uniqueness of factorisation. In addition, it characterises $\gcd(a, b)$ as the unique positive integer $d > 0$ such that $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$.

Note that the relation $a\mathbf{Z} + b\mathbf{Z} = \gcd(a, b)\mathbf{Z}$ is equivalent to

$$\frac{1}{a}\mathbf{Z} + \frac{1}{b}\mathbf{Z} = \frac{1}{ab}(b\mathbf{Z} + a\mathbf{Z}) = \frac{\gcd(a, b)}{ab}\mathbf{Z} = \frac{1}{\operatorname{lcm}(a, b)}\mathbf{Z}.$$

By induction, Theorem 2.3.3 implies that, for any $a_i \in \mathbf{Z} \smallsetminus \{0\}$,

$$a_1\mathbf{Z} + \cdots + a_r\mathbf{Z} = d\mathbf{Z},$$

where $d \in \mathbf{N}_+$ satisfies

$$\{\text{common divisors of } a_1, \ldots, a_r\} = \{\text{divisors of } d\}.$$

The latter property characterises $\gcd(a_1, \ldots, a_r)$; therefore $d = \gcd(a_1, \ldots, a_r)$ and

$$a_1\mathbf{Z} + \cdots + a_r\mathbf{Z} = \gcd(a_1, \ldots, a_r)\mathbf{Z}.$$

2.3.5 Exercise. (1) Determine $\frac{2}{5}\mathbf{Z} + \frac{3}{7}\mathbf{Z}$.
(2) Show that, for any $a_i \in \mathbf{Z} \smallsetminus \{0\}$,

$$\frac{1}{a_1}\mathbf{Z} + \cdots + \frac{1}{a_r}\mathbf{Z} = \frac{1}{\operatorname{lcm}(a_1, \ldots, a_r)}\mathbf{Z}.$$

2.3.6 Corollary. *Let $a, b, c \in \mathbf{Z} \smallsetminus \{0\}$. If $c \mid a$ and $c \mid b$, then $\gcd(a/c, b/c) = \gcd(a, b)/|c|$. In particular, if $d = \gcd(a, b)$, then $\gcd(a/d, b/d) = 1$. As a result, $a/b = a'/b'$, where $a' = a/d$, $b' = b/d$ and $\gcd(a', b') = 1$.*

Proof. We know that $c \mid d$; we can divide, therefore, the equality $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ by c, obtaining $(a/c)\mathbf{Z} + (b/c)\mathbf{Z} = (d/c)\mathbf{Z} = (d/|c|)\mathbf{Z}$, with $d/|c| > 0$; therefore $\gcd(a/c, b/c) = \gcd(a, b)/|c|$. □

2.3.7 Exercise. (1) Give an alternative proof of Corollary 2.3.6 using Theorem 1.6.2(2).
(2) Show: if $\gcd(a, b) = 1$, then $\gcd(a^m, b^n) = 1$, for all $m, n \in \mathbf{N}_+$.

2.3.8 Lemma. *Let $a, b, c \in \mathbf{Z} \smallsetminus \{0\}$. If $c \mid ab$ and $\gcd(c, b) = 1$, then $c \mid a$.*

Proof. According to Bézout's theorem, $c\mathbf{Z} + b\mathbf{Z} = \gcd(c, b)\mathbf{Z} = \mathbf{Z}$. After multiplying this equality by a, we obtain (since $c\mathbf{Z} \supseteq ab\mathbf{Z}$) that

$$a\mathbf{Z} = ac\mathbf{Z} + ab\mathbf{Z} \subseteq ac\mathbf{Z} + c\mathbf{Z} \subseteq c\mathbf{Z} + c\mathbf{Z} = c\mathbf{Z},$$

hence $c \mid a$. □

2.3.9 Lemma (Euclid's Lemma (bis))**.** *Let $a, b \in \mathbf{Z} \smallsetminus \{0\}$. If a prime number p satisfies $p \mid ab$ and $p \nmid b$, then $p \mid a$.*

Proof. The positive integer $d := \gcd(p, b)$ divides p, hence $d = 1$ or $d = p$. It also divides b, but $p \nmid b$; thus $d = 1$. Lemma 2.3.8 then applies with $c = p$. □

2.3.10 Remark This completes the proof of the uniqueness of factorisation (Theorem 1.4.2). A less abstract proof of Bézout's theorem will be given in Theorem 2.4.2 below, as a consequence of Euclid's algorithm.

2.3.11 Exercise. Let p be a prime and $x \in \mathbf{Z}$. Show that:
(1) If $p \mid (x^2 - 1)$, then $p \mid (x - 1)$ or $p \mid (x + 1)$.
(2) If $p^k \mid (x^2 - 1)$, $k > 1$ and $p \neq 2$, then $p^k \mid (x - 1)$ or $p^k \mid (x + 1)$.
(3) What happens if $p = 2$ in (2)?

2.3.12 Theorem (Yet another proof of irrationality of $\sqrt[n]{a}$)**.** *If $a, n \in \mathbf{N}_+$, then: $\sqrt[n]{a} \in \mathbf{Q} \iff \sqrt[n]{a} \in \mathbf{Z}$.*

Proof. If $\sqrt[n]{a} \in \mathbf{Q}$, then $\sqrt[n]{a} = c/b$ for some $b, c \in \mathbf{N}_+$ satisfying $\gcd(c, b) = 1$. As $ab^n = c^n$ is divisible by c^n and $\gcd(c^n, b^n) = \gcd(c, b)^n = 1$, Lemma 2.3.8 applied to the triple a, b^n, c^n tells us that $c^n = ab^n$ divides a. This is possible only if $|b^n| = 1$, which implies that $b = 1$ and $\sqrt[n]{a} = c/b = c \in \mathbf{N}_+$. □

2.3.13 Rational roots of polynomials Theorem 2.3.12 can be reformulated as follows: "if $\alpha \in \mathbf{Q}$ is a root of the polynomial $X^n - a$ (where $a, n \in \mathbf{N}_+$), then $\alpha \in \mathbf{Z}$." What happens for more general polynomials?

2.3.14 Theorem (Rational roots of polynomials)**.** *Let $a_0, \dots, a_n \in \mathbf{Z}$, $a_0 \neq 0$, $n \geq 1$. If $\alpha = a/b \in \mathbf{Q}$ (where $a, b \in \mathbf{Z} \smallsetminus \{0\}$ and $\gcd(a, b) = 1$) is a root of the polynomial equation $f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n = 0$, then $a \mid a_n$ and $b \mid a_0$. In particular, if $a_0 = 1$, then $\alpha = a/b \in \mathbf{Z}$.*

Proof. After multiplying the equality

$$a_0 \left(\frac{a}{b}\right)^n + a_1 \left(\frac{a}{b}\right)^{n-1} + \cdots + a_{n-1}\left(\frac{a}{b}\right) + a_n = 0$$

by b^n we obtain

$$a_0 a^n + a_1 a^{n-1} b + \cdots + a_{n-1} a b^{n-1} + a_n b^n = 0.$$

The sum of all the terms is equal to zero and all terms except the last one $a_n b^n$ are divisible by a, hence $a \mid a_n b^n$. As $\gcd(a, b^n) = 1$, Lemma 2.3.8 implies that $a \mid a_n$. Similarly, all terms except the first one $a_0 a^n$ are divisible by b, hence $b \mid a_0 a^n$. The same argument based on $\gcd(b, a^n) = 1$ then implies that $b \mid a_0$. □

2.3.15 Exercise. Find all rational solutions $\alpha \in \mathbf{Q}$ of the equation $f(x) = x^3 + x^2 - 5x + 3 = 0$.

2.3.16 Terminology A complex number $\alpha \in \mathbf{C}$ satisfying a polynomial equation

$$a_0 \alpha^n + a_1 \alpha^{n-1} + \cdots + a_n = 0$$

for some $a_j \in \mathbf{Z}$ (with $n \geq 1$ and $a_0 \neq 0$) is called an **algebraic number**. If, in addition, $a_0 = 1$, then we say that α is an **algebraic integer**. The second part of Theorem 2.3.14 says that an algebraic integer contained in $\mathbf{Q}$ is a usual integer.

2.4 Euclid's algorithm

2.4.1 The general case of Euclid's algorithm Let $a, b \in \mathbf{Z} \smallsetminus \{0\}$. As in 2.2.3, we are going to perform the division with remainder of a by b, then replace the pair (a, b) by (b, r), and repeat the procedure until we obtain a pair $(d, 0)$. All the successive remainders will be explicit linear combinations of a and b; they will also be divisible by d. This will give a weak Bézout's theorem $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ (hence also $\gcd(a, b) = |d|$), as well as an explicit pair of integers $u, v \in \mathbf{Z}$ such that $au + bv = d$.

We write, recursively, $r_{-2} = a$, $r_{-1} = b$ and

$$\begin{aligned} a &= a_0 b + r_0 & 0 \leq r_0 < |b| \\ b &= a_1 r_0 + r_1 & 0 \leq r_1 < r_0 \\ r_0 &= a_2 r_1 + r_2 & 0 \leq r_2 < r_1 \\ &\vdots & \vdots \\ r_{k-2} &= a_k r_{k-1} + r_k & 0 = r_k < r_{k-1} \end{aligned} \tag{2.4.1.1}$$

(where $k \geq 0$). We let $d := r_{k-1}$. Note that the decreasing sequence of positive integers

$$|b| = |r_{-1}| > r_0 > r_1 > \cdots > r_{k-1} = d > r_k = 0$$

must, indeed, terminate at some point.

This is a general version of what was done in 2.2.3 for $a = 44$ and $b = 16$. All the phenomena that were observed there in this special case hold in general. For example, the identities

$$\frac{r_i}{r_{i+1}} = a_{i+2} + \frac{1}{r_{i+1}/r_{i+2}} \qquad (0 \le i+2 \le k)$$

give a continues fraction

$$\frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_k}}} \tag{2.4.1.2}$$

2.4.2 Theorem (Properties of Euclid's algorithm)**.** *(1) $d \mid r_i \quad (-2 \le i \le k)$*
(2) $r_i \in a\mathbf{Z} + b\mathbf{Z} \quad (-2 \le i \le k)$
(3) $d \mid a$, $d \mid b$ and there exist $u, v \in \mathbf{Z}$ such that $au + bv = d$.
(4) $a\mathbf{Z} + b\mathbf{Z} = d\mathbf{Z}$ (weak form of Bézout's theorem).
(5) $|d| = \gcd(a, b)$ and $a\mathbf{Z} + b\mathbf{Z} = \gcd(a, b)\mathbf{Z}$ (strong form of Bézout's theorem).

Proof. (1) Decreasing induction on i starting from $i = k$ and $i = k-1$: $r_k = 0$, $r_{k-1} = d$ and $r_i = a_{i+2}r_{i+1} + r_{i+2}$.
(2) Increasing induction on i starting from $i = -2$ and $i = -1$: $r_{-2} = a$, $r_{-1} = b$ and $r_{i+2} = r_i - a_{i+2}r_{i+1}$.
(3) This is a special case of the statement (1) for $i = -2$, $i = -1$ and of (2) for $i = k-1$.
(4) As in 2.2.3, (3) implies, on one hand, that $d\mathbf{Z} \supseteq a\mathbf{Z}$ and $d\mathbf{Z} \supseteq b\mathbf{Z}$, hence $d\mathbf{Z} \supseteq a\mathbf{Z} + b\mathbf{Z}$, and on the other hand $d \in a\mathbf{Z} + b\mathbf{Z}$ implies that $d\mathbf{Z} \subseteq a\mathbf{Z} + b\mathbf{Z}$.
(5) This follows from (4), as in the proof of Theorem 2.3.3. □

2.4.3 Explicit Bézout relations It is very important that Euclid's algorithm gives, for all $-2 \le i \le k$, explicit integers $u_i, v_i \in \mathbf{Z}$ such that $u_i a + v_i b = r_i$. In particular, for $i = k-1$ we obtain $u = u_{k-1}$ and $v = v_{k-1}$ such that $ua + vb = d = \operatorname{sgn}(d)\gcd(a, b)$.

For small values of a and b we can compute u_i and v_i by hand, as in (2.2.3.1)–(2.2.3.4). In general, the coefficients u_i, v_i are related to the continued fraction (2.4.1.2), as in 2.2.4. The general formulas are as follows.

For each $j \ge 0$, we need to simplify the continued fraction

$$[a_0, \dots, a_j] := a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_j}}} \tag{2.4.3.1}$$

and write it as a usual fraction p_j/q_j. For small values of j we have

$$\frac{p_0}{q_0} = \frac{a_0}{1}, \qquad \frac{p_1}{q_1} = a_0 + \frac{1}{a_1} = \frac{a_0a_1 + 1}{a_1}, \qquad \frac{p_2}{q_2} = a_0 + \frac{a_2}{a_1a_2 + 1} = \frac{a_2(a_0a_1 + 1) + a_0}{a_1a_2 + 1}.$$

j	-2	-1	0	1	2
a_j			a_0	a_1	a_2
p_j	0	1	a_0	$a_0a_1 + 1$	$a_2(a_0a_1 + 1) + a_0$
q_j	1	0	1	a_1	$a_1a_2 + 1$

This table seems to suggest that (if we let $p_{-2} = q_{-1} = 0$ and $p_{-1} = q_{-2} = 1$), for all $j \geq 0$,

$$\begin{aligned} p_j &= a_j p_{j-1} + p_{j-2} \\ q_j &= a_j q_{j-1} + q_{j-2}, \end{aligned}$$

which is equivalent to matrix equations

$$\begin{pmatrix} p_j & p_{j-1} \\ q_j & q_{j-1} \end{pmatrix} = \begin{pmatrix} p_{j-1} & p_{j-2} \\ q_{j-1} & q_{j-2} \end{pmatrix} M_j, \qquad M_j = \begin{pmatrix} a_j & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.4.3.2}$$

This is, indeed, the case (see Section 2.4.6 below for more details). Inductively, these formulas give

$$\forall j \geq -1 \qquad M_0 \cdots M_j = \begin{pmatrix} p_j & p_{j-1} \\ q_j & q_{j-1} \end{pmatrix}, \tag{2.4.3.3}$$

which implies that

$$\forall j \geq -1 \quad \begin{vmatrix} p_j & p_{j-1} \\ q_j & q_{j-1} \end{vmatrix} = (-1)^{j+1}, \quad \begin{pmatrix} p_j & p_{j-1} \\ q_j & q_{j-1} \end{pmatrix}^{-1} = (-1)^{j+1} \begin{pmatrix} q_{j-1} & -p_{j-1} \\ -q_j & p_j \end{pmatrix}.$$

Similarly, the relations $r_{j-2} = a_j r_{j-1} + r_j$ can be written in a matrix form

$$\forall j = 0, \ldots, k \qquad M_j \begin{pmatrix} r_{j-1} \\ r_j \end{pmatrix} = \begin{pmatrix} r_{j-2} \\ r_{j-1} \end{pmatrix}.$$

Inductively, we obtain

$$\forall j = -1, \ldots, k \qquad M_0 \cdots M_j \begin{pmatrix} r_{j-1} \\ r_j \end{pmatrix} = \begin{pmatrix} r_{-2} \\ r_{-1} \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix},$$

hence

$$\forall j = -1, \ldots, k \qquad \begin{pmatrix} r_{j-1} \\ r_j \end{pmatrix} = (-1)^{j+1} \begin{pmatrix} q_{j-1} & -p_{j-1} \\ -q_j & p_j \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

The last formula can be rewritten as

$$\forall j = -2, \ldots, k \qquad (-1)^j r_j = q_j a - p_j b = \begin{pmatrix} -b & a \end{pmatrix} \begin{pmatrix} p_j \\ q_j \end{pmatrix}. \tag{2.4.3.4}$$

This system of equations can be reformulated as a single matrix equation

$$\begin{pmatrix} -b & a \end{pmatrix} \begin{pmatrix} p_{-2} & p_{-1} & p_0 & \cdots & p_{k-1} & p_k \\ q_{-2} & q_{-1} & q_0 & \cdots & q_{k-1} & q_k \end{pmatrix} = \begin{pmatrix} a & -b & r_0 & \cdots & (-1)^j r_j & \cdots & (-1)^{k-1} d & 0 \end{pmatrix}. \tag{2.4.3.5}$$

We have already seen a special case of this matrix formula in (2.2.4.6).

In particular, the case $j = k - 1$ of (2.4.3.4) gives an explicit Bézout relation

$$q_{k-1} a - p_{k-1} b = (-1)^{k-1} d, \qquad |d| = \gcd(a, b).$$

2.4.4 Sample computations To sum up the previous discussion: Euclid's algorithm applied to $a, b \in \mathbf{Z} \smallsetminus \{0\}$ produces integers a_j and r_j $(0 \le j \le k)$ given by (2.4.1.1). We let $r_{-2} = a$ and $r_{-1} = b$. After that we compute integers p_j and q_j $(-2 \le j \le k)$ by the recursive formulas (2.4.3.2) and put them into a table

j	-2	-1	0	1	$\cdots$	$k-1$	k
a_j			a_0	a_1	$\cdots$	a_{k-1}	a_k
p_j	0	1	a_0	a_0a_1+1	$\cdots$	p_{k-1}	p_k
q_j	1	0	1	a_1	$\cdots$	q_{k-1}	q_k

In the notation of (2.4.3.1) we have

$$\forall j = 0, \dots, k \qquad [a_0, \dots, a_j] = \frac{p_j}{q_j}.$$

Moreover, $\gcd(p_j, q_j) = 1$, since $p_j q_{j-1} - q_j p_{j-1} = \pm 1$. On the other hand, $\gcd(a, b) = |d|$, where $d = r_{k-1}$. In particular, for $j = k$ we obtain that

$$\frac{a}{b} = [a_0, \dots, a_k] = \frac{p_k}{q_k}, \qquad |a/d| = |p_k|, \qquad |b/d| = |q_k|.$$

Furthermore, the matrix formula (2.4.3.5) expresses each r_j (in particular, also $d = r_{k-1}$ and $\gcd(a, b) = |d|$) as an explicit linear combination of a and b (an **explicit Bézout relation** for a and b).

Example 1: $a = 18$, $b = 11$. In this case

$$18 = 1 \cdot 11 + 7, \qquad 11 = 1 \cdot 7 + 4, \qquad 7 = 1 \cdot 4 + 3, \qquad 4 = 1 \cdot 3 + 1, \qquad 3 = 3 \cdot 1 + 0$$

j	-2	-1	0	1	2	3	4
a_j			1	1	1	1	3
p_j	0	1	1	2	3	5	18
q_j	1	0	1	1	2	3	11

Therefore

$$\gcd(18, 11) = 1, \qquad 5 \cdot 11 + (-3) \cdot 18 = 1$$

and, more generally,

$$\begin{pmatrix} -11 & 18 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 2 & 3 & 5 \\ 1 & 0 & 1 & 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 18 & -11 & 7 & -4 & 3 & -1 \end{pmatrix}.$$

One can also compute directly

$$7 = 18 - 11, \qquad 4 = 11 - (18 - 11) = 2 \cdot 11 - 18, \qquad 3 = (18 - 11) - (2 \cdot 11 - 18) =$$
$$= 2 \cdot 18 - 3 \cdot 11, \qquad 1 = (2 \cdot 11 - 18) - (2 \cdot 18 - 3 \cdot 11) = 5 \cdot 11 - 3 \cdot 18$$

or

$$1 + \cfrac{1}{1 + \cfrac{1}{1 + \frac{1}{1}}} = 1 + \cfrac{1}{1 + \frac{1}{2}} = 1 + \frac{2}{3} = \frac{5}{3}.$$

Example 2: $a = 31$, $b = 13$. In this case

$$31 = 2 \cdot 13 + 5, \qquad 13 = 2 \cdot 5 + 3, \qquad 5 = 1 \cdot 3 + 2, \qquad 3 = 1 \cdot 2 + 1, \qquad 2 = 2 \cdot 1 + 0$$

j	-2	-1	0	1	2	3	4
a_j			2	2	1	1	2
p_j	0	1	2	5	7	12	31
q_j	1	0	1	2	3	5	13

Therefore

$$\gcd(31,13)=1, \qquad 12\cdot 13+(-5)\cdot 31=1$$

and

$$\begin{pmatrix}-13 & 31\end{pmatrix}\begin{pmatrix}0 & 1 & 2 & 5 & 7 & 12\\ 1 & 0 & 1 & 2 & 3 & 5\end{pmatrix}=\begin{pmatrix}31 & -13 & 5 & -3 & 2 & -1\end{pmatrix}.$$

Again, one can compute directly

$$\begin{gathered}5=31-2\cdot 13, \qquad 3=13-2\cdot(31-2\cdot 13)=5\cdot 13-2\cdot 31,\\ 2=(31-2\cdot 13)-(5\cdot 13-2\cdot 31)=3\cdot 31-7\cdot 13,\\ 1=(5\cdot 13-2\cdot 31)-(3\cdot 31-7\cdot 13)=12\cdot 13-5\cdot 31\end{gathered}$$

or

$$2+\cfrac{1}{2+\cfrac{1}{1+\frac{1}{1}}}=2+\frac{1}{2+\frac{1}{2}}=2+\frac{2}{5}=\frac{12}{5}.$$

2.4.5 Modified Euclid's algorithm It is often useful to modify the division with remainder from Proposition 2.2.2 by allowing a **negative remainder**, but minimising its absolute value. The conditions

$$a=qb+r, \qquad 0\le r<|b|$$

are then replaced by

$$a=q'b+r', \qquad |r'|\le |b|/2.$$

For example, we can write

$$\begin{aligned}a&=5q+3=5(q+1)+(-2)\\ a&=4q+3=4(q+1)+(-1)\\ a&=4q+2=4(q+1)+(-2).\end{aligned}$$

If b is odd, then the pair (q',r') is unique. The same is true if b is even and $r<|b|/2$. On the other hand, if b is even and $r=|b|/2$, then there is a choice between $a=qb+|b|/2$ and $a=(q+\operatorname{sgn}(b))b+(-|b|/2)$.

In the **modified Euclid's algorithm** one replaces the usual division with remainder by its modification discussed above. The algorithm then becomes more efficient. For example, for $a=44$ and $b=16$ from 2.2.3 one had to perform three divisions with remainder. The modified version needs only two divisions:

$$\begin{aligned}44&=3\cdot 16+(-4)\\ 16&=(-4)\cdot(-4)+0\\ -4&=d'\end{aligned}$$

The corresponding continued fraction

$$\frac{44}{16} = 3 + \frac{1}{-4}$$

and its table of coefficients

j	-2	-1	0	1
a_j			3	-4
p_j	0	1	3	-11
q_j	1	0	1	-4

then give explicit Bézout relations

$$1 \cdot 44 + (-3) \cdot 16 = -4, \qquad \gcd(44, 16) = 4$$

and

$$\begin{pmatrix} -16 & 44 \end{pmatrix} \begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 44 & -16 & -4 \end{pmatrix}.$$

The two examples from Section 2.4.4 can be treated in the same way.
Example 1: $a = 18$, $b = 11$. In this case

$$\begin{gathered} 18 = 2 \cdot 11 + (-4), \qquad 11 = (-3) \cdot (-4) + (-1), \qquad -4 = 4 \cdot (-1) + 0, \qquad -1 = d', \\ -4 = 18 - 2 \cdot 11, \qquad -1 = 11 + 3(18 - 2 \cdot 11) = 3 \cdot 18 - 5 \cdot 11, \qquad 1 = 5 \cdot 11 - 3 \cdot 18. \end{gathered}$$

Alternatively, we compute

j	-2	-1	0	1	2
a_j			2	-3	4
p_j	0	1	2	-5	-18
q_j	1	0	1	-3	-11

This gives, as before (but faster), the following relations:

$$\gcd(18, 11) = 1, \qquad 5 \cdot 11 + (-3) \cdot 18 = 1$$

and

$$\begin{pmatrix} -11 & 18 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 & -5 \\ 1 & 0 & 1 & -3 \end{pmatrix} = \begin{pmatrix} 18 & -11 & -4 & 1 \end{pmatrix}.$$

Example 2: $a = 31$, $b = 13$. In this case

$$\begin{gathered} 31 = 2 \cdot 13 + 5, \qquad 13 = 3 \cdot 5 + (-2), \qquad 5 = (-2) \cdot (-2) + 1, \qquad -2 = (-2) \cdot 1 + 0, \qquad 1 = d', \\ 5 = 31 - 2 \cdot 13, \qquad -2 = 13 - 3(31 - 2 \cdot 13) = 7 \cdot 13 - 3 \cdot 31, \\ 1 = (31 - 2 \cdot 13) + 2(7 \cdot 13 - 3 \cdot 31) = 12 \cdot 13 - 5 \cdot 31 \end{gathered}$$

j	-2	-1	0	1	2	3
a_j			2	3	-2	-2
p_j	0	1	2	7	-12	31
q_j	1	0	1	3	-5	13

This gives, as before,

$$\gcd(31,13)=1, \qquad 12\cdot 13+(-5)\cdot 31=1$$

and

$$\begin{pmatrix}-13 & 31\end{pmatrix}\begin{pmatrix}0 & 1 & 2 & 7 & -12\\ 1 & 0 & 1 & 3 & -5\end{pmatrix}=\begin{pmatrix}31 & -13 & 5 & 2 & 1\end{pmatrix}.$$

2.4.6 Continued fractions and matrices Where do the matrix formulas (2.4.3.2) and (2.4.3.3) come from?

The point is that if we let a complex matrix $M=\begin{pmatrix}A & B\\ C & D\end{pmatrix}\in M_2(\mathbf{C})$ with $\det(M)\neq 0$ act on $\mathbf{C}\cup\{\infty\}$ by the Möbius transformation

$$M(z):=\frac{Az+B}{Cz+D}$$

(with $M(\infty)=A/C$ and $M(-D/C)=\infty$), then

$$M_1(M_2(z))=(M_1M_2)(z), \qquad \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}(z)=z. \tag{2.4.6.1}$$

The formula (2.4.6.1) is a consequence of a natural identification of the complex projective line $\mathbf{C}\cup\{\infty\}=\mathbf{P}^1(\mathbf{C})$ with the set of lines passing through the origin (i.e., one-dimensional vector subspaces) in $\mathbf{C}^2$. The standard linear action of $M\in M_2(\mathbf{C})$ on $\mathbf{C}^2$ sends the subspace $\mathbf{C}\begin{pmatrix}z\\1\end{pmatrix}$ onto

$$M\big(\mathbf{C}\begin{pmatrix}z\\1\end{pmatrix}\big)=\mathbf{C}\cdot M\begin{pmatrix}z\\1\end{pmatrix}=\mathbf{C}\cdot\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}z\\1\end{pmatrix}=\mathbf{C}\cdot\begin{pmatrix}Az+B\\ Cz+D\end{pmatrix}=\mathbf{C}\cdot\begin{pmatrix}\frac{Az+B}{Cz+D}\\ 1\end{pmatrix},$$

and the subspace $\mathbf{C}\begin{pmatrix}1\\0\end{pmatrix}$ onto

$$M\big(\mathbf{C}\begin{pmatrix}1\\0\end{pmatrix}\big)=\mathbf{C}\cdot M\begin{pmatrix}1\\0\end{pmatrix}=\mathbf{C}\cdot\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\mathbf{C}\cdot\begin{pmatrix}A\\ C\end{pmatrix}=\mathbf{C}\cdot\begin{pmatrix}\frac{A}{C}\\ 1\end{pmatrix}.$$

This is relevant to continued fractions, thanks to the formula

$$a+\frac{1}{z}=\frac{a\cdot z+1}{1\cdot z+0}=\begin{pmatrix}a & 1\\ 1 & 0\end{pmatrix}(z),$$

which implies, in the notation of (2.4.3.1), that

$$[a_0,\dots,a_j,z]=M_0(M_1(\cdots(M_j(z))\cdots)=(M_0\cdots M_j)(z), \qquad M_i=\begin{pmatrix}a_j & 1\\ 1 & 0\end{pmatrix} \tag{2.4.6.2}$$

(for any $a_1,\dots,a_j,z\in\mathbf{C}$). If we define $p_j,q_j\in\mathbf{C}$ by

$$M_0\cdots M_j=\begin{pmatrix}p_j & *\\ q_j & *\end{pmatrix},$$

then

$$\begin{pmatrix}p_{j-1} & *\\ q_{j-1} & *\end{pmatrix}\begin{pmatrix}a_j & 1\\ 1 & 0\end{pmatrix}=\begin{pmatrix}* & p_{j-1}\\ * & q_{j-1}\end{pmatrix},$$

which implies that

$$M_0 \cdots M_j = \begin{pmatrix} p_j & p_{j-1} \\ q_j & q_{j-1} \end{pmatrix}.$$

Finally, substituting $z = \infty$ into (2.4.6.2), one obtains

$$[a_0, \ldots, a_j] = (M_0 \cdots M_j)(\infty) = \frac{p_j}{q_j}.$$

This finishes the proof of the formula (2.4.3.3).

2.5 Equations $ax + by = c$ $\quad (x, y \in \mathbf{Z})$

2.5.1 Example $44x + 16y = 6 \qquad (x, y \in \mathbf{Z})$.

The left hand side $44x + 16y = 4(11x + 4y)$ is always divisible by $4 = \gcd(44, 16)$, but the right hand side is not: $4 \nmid 6$. It follows that there is no solution.

2.5.2 Example $44x + 16y = 40 \qquad (x, y \in \mathbf{Z})$.

In this case the right hand side is divisible by $4 = \gcd(44, 16)$. After dividing by 4, we obtain an equivalent equation

$$11x + 4y = 10 \qquad (x, y \in \mathbf{Z}), \tag{2.5.2.1}$$

this time with $\gcd(11, 4) = 1$.

Assume that we know a particular solution $x_1, y_1 \in \mathbf{Z}$ of (2.5.2.1):

$$11x_1 + 4y_1 = 10. \tag{2.5.2.2}$$

Subtracting (2.5.2.2) from (2.5.2.1), we obtain

$$11(x - x_1) = 4(y_1 - y), \tag{2.5.2.3}$$

hence $11 \mid 4(y_1 - y)$. As $\gcd(11, 4) = 1$, Lemma 2.3.8 implies that $11 \mid (y_1 - y)$. Therefore $y_1 - y = 11t$ for some $t \in \mathbf{Z}$; substitution into (2.5.2.3) yields $x - x_1 = 4t$. In other words,

$$x = x_1 + 4t, \qquad y = y_1 - 11t \qquad (t \in \mathbf{Z}). \tag{2.5.2.4}$$

Conversely, if (2.5.2.4) and (2.5.2.2) hold true, then

$$11x + 4y = 11(x_1 + 4t) + 4(y_1 - 11t) = 11x_1 + 4y_1 + (11 \cdot 4 - 4 \cdot 11)t = 10.$$

It remains to find a particular solution $x_1, y_1 \in \mathbf{Z}$. We use the Bézout relation

$$(-1) \cdot 44 + 3 \cdot 16 = 4$$

from (2.2.3.4); it implies that

$$(-1) \cdot 11 + 3 \cdot 4 = 1. \tag{2.5.2.5}$$

After multiplying the last equation by 10, we obtain a particular solution

$$(-10) \cdot 11 + 30 \cdot 4 = 10$$

$x_1 = -10$, $y_1 = 30$ of (2.5.2.1), hence a general solution

$$x = 4t - 10, \qquad y = 30 - 11t \qquad (t \in \mathbf{Z}). \tag{2.5.2.6}$$

2.5.3 Making the numbers smaller The constants appearing in (2.5.2.6) can be made smaller, by performing a (modified) division with remainder. For example, if we write $30 = 3 \cdot 11 + (-3)$, we obtain $y = -3 - 11(t-3)$. A change of variables $t - 3 = s$ then yields $t = s + 3$ and $x = 4(s+3) - 10 = 4s + 2$, hence

$$x = 4s + 2, \qquad y = -3 - 11s \qquad (s \in \mathbf{Z}). \tag{2.5.3.1}$$

We can make the numbers smaller already during the calculations. For example, if we write $10 = 2 \cdot 4 + 2$, then we can replace (2.5.2.1) by

$$11x + 4(y-2) = 2. \tag{2.5.3.2}$$

A particular solution $x_2, y_2 \in \mathbf{Z}$ of this equation can be found by multiplying (2.5.2.5) by 2: $x_2 = -2$ and $y_2 - 2 = 6$ (hence $y_2 = 8$). As above, this yields

$$x = 4t - 2, \qquad y = 8 - 11t \qquad (t \in \mathbf{Z}). \tag{2.5.3.3}$$

If we are not satisfied with this form of the solution, we can write $8 = 1 \cdot 11 + (-3)$ and $y = -3 - 11(t-1)$ and let $t - 1 = s$. This will give again the formulas in (2.5.3.1).

The general case is as follows.

2.5.4 Theorem. *Let $a, b \in \mathbf{Z} \setminus \{0\}$ and $c \in \mathbf{Z}$. Consider the equation*

$$ax + by = c \qquad (x, y \in \mathbf{Z}). \tag{2.5.4.1}$$

(1) If $d := \gcd(a, b)$ does not divide c, then (2.5.4.1) *has no solution.*
(2) If d divides c, then (2.5.4.1) *does have a solution. For any such a solution $x_1, y_1 \in \mathbf{Z}$, the general solution of* (2.5.4.1) *is given by*

$$x = x_1 + (b/d)t, \qquad y = y_1 - (a/d)t \qquad (t \in \mathbf{Z}).$$

Proof. We proceed as in the above examples.
(1) The left hand side $ax + by$ is divisible by d, for all $x, y \in \mathbf{Z}$, but c is not.
(2) If d divides c, then we can divide (2.5.4.1) by d and obtain an equivalent equation

$$a'x + b'y = c' \qquad (x, y \in \mathbf{Z}), \tag{2.5.4.2}$$

where $a' = a/d$, $b' = b/d$ and $c' = c/d$. Note that $\gcd(a', b') = d/d = 1$.

If $x_1, y_1 \in \mathbf{Z}$ is a particular solution of (2.5.4.2), then we can subtract $a'x_1 + b'y_1 = c'$ from (2.5.4.2) and obtain

$$a'(x - x_1) = b'(y_1 - y), \tag{2.5.4.3}$$

hence $a' \mid b'(y_1 - y)$. As $\gcd(a', b') = 1$, it follows that $a' \mid (y_1 - y)$, hence $y_1 - y = a't$ for some $t \in \mathbf{Z}$. Substitution of this relation into (2.5.4.3) yields

$$x = x_1 + b't, \qquad y = y_1 - a't \qquad (t \in \mathbf{Z}).$$

Conversely, this formula produces solutions of (2.5.4.2), since

$$a'x + b'y = a'x_1 + b'y_1 + (a'b' - b'a')t = c'.$$

It remains to show that (2.5.4.2) always has a solution. This follows immediately by multiplying a Bézout relation

$$a'u + b'v = \gcd(a', b') = 1 \qquad (u, v \in \mathbf{Z})$$

by c'; the pair $x_1 := c'u$, $y_1 := c'v$ will be a solution of (2.5.4.2).

However, this solution can be fairly large, so it would make sense to apply the methods of 2.5.3 in order to obtain a smaller particular solution. □

2.5.5 Exercise. Find all solutions $x, y \in \mathbf{Z}$ of $10x + 16y = 18$.

2.6 Expansions of integers in base b

2.6.1 Base 10 The decimal expansion of $n = 468$ is given by

$$\begin{aligned} 468 &= 46 \cdot 10 + 8 \\ 46 &= 4 \cdot 10 + 6 \\ 4 &= 0 \cdot 10 + 4 \end{aligned} \qquad\qquad 468 = 4 \cdot 10^2 + 6 \cdot 10^1 + 8 \cdot 10^0$$

2.6.2 Base 7 The expansion of $n = 468$ in base 7 is given by

$$\begin{aligned} 468 &= 66 \cdot 7 + 6 \\ 66 &= 9 \cdot 7 + 3 \\ 9 &= 1 \cdot 7 + 2 \\ 1 &= 0 \cdot 7 + 1 \end{aligned} \qquad\qquad 468 = 1 \cdot 7^3 + 2 \cdot 7^2 + 3 \cdot 7^1 + 6 \cdot 7^0 = (1236)_7$$

2.6.3 General base Given an integer $b > 1$, one can expand any $n \in \mathbf{N}$ in base b using digits a_i representing the values $0, 1, \ldots, b-1$ (for example, in the hexadecimal base $b = 16$ one uses digits $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F$). One writes

$$n = a_k b^k + a_{k-1} b^{k-1} + \cdots + a_1 b + a_0 = (a_k \cdots a_1 a_0)_b \qquad (a_i \in \{0, 1, \ldots, b-1\}),$$

and the digits a_i of n in base b can be computed recursively as follows:

- $a_0 :=$ the remainder of the division of n by b,
- $n_1 := (n - a_0)/b$ the quotient of the division with remainder of n by b,
- $a_1 :=$ the remainder of the division of n_1 by b,
- $n_2 := (n_1 - a_1)/b$ the quotient of the division with remainder of n_1 by b, etc.

2.6.4 First application: computing high powers One can compute a^m for big $m \in \mathbf{N}$ by successive squarings. The point is to write $m = \sum 2^{k_i}$ in base 2 and then compute successively

$$a^2, \qquad (a^2)^2 = a^{2^2} = a^4, \qquad \left((a^2)^2\right)^2 = a^{2^3} = a^8, \qquad \cdots$$

For example,

$$m = 105 = 64 + 32 + 8 + 1 = 2^6 + 2^5 + 2^3 + 2^0 = (1101001)_2, \qquad a^{105} = a^{64} \cdot a^{32} \cdot a^8 \cdot a,$$

which means that, in order to compute a^{105}, it is sufficient to compute the following 9 products:

$a \cdot a = a^2$	
$a^2 \cdot a^2 = a^4$	
$a^4 \cdot a^4 = a^8$	$a^8 \cdot a$
$a^8 \cdot a^8 = a^{16}$	
$a^{16} \cdot a^{16} = a^{32}$	$a^{32} \cdot (a^8 \cdot a)$
$a^{32} \cdot a^{32} = a^{64}$	$a^{64} \cdot (a^{32} \cdot (a^8 \cdot a)) = a^{105}$

2.6.5 Second application: Legendre's formula for $v_p(n!)$ **Example:** let us compute $v_3(11!)$ (the exponent with which 3 occurs in the prime factorisation of 11!). In the product

$$11! = 1 \cdot 2 \cdot \boxed{3} \cdot 4 \cdot 5 \cdot \boxed{6} \cdot 7 \cdot 8 \cdot \boxed{9} \cdot 10 \cdot 11$$

only the boxed terms

$$3 = 3 \cdot 1, \qquad 6 = 3 \cdot 2, \qquad 9 = 3 \cdot 3$$

are divisible by 3. As a result,

$$v_3(11!) = v_3(3 \cdot 6 \cdot 9) = v_3(3^3) + v_3(1 \cdot 2 \cdot 3) = 3 + v_3(3!).$$

Similarly,

$$3! = 1 \cdot 2 \cdot \boxed{3}, \qquad v_3(3!) = v_3(3) = 1 + v_3(1!) = 1,$$

which implies that $v_3(11!) = 3 + 1 = 4$.

The general case: let us compute $v_p(n!)$ for $n \in \mathbf{N}_+$ and $p \in \mathcal{P}$ by the same method. Writing $n = pn_1 + a_0$ with $0 \leq a_0 < p$, we see that the terms divisible by p in the product $n! = 1 \cdots n$ are those equal to $p \cdot 1, p \cdot 2, \ldots p \cdot n_1$. Therefore

$$v_p(n!) = v_p(p^{n_1}(n_1)!) = n_1 + v_p((n_1)!), \qquad n_1 = \left\lfloor \frac{n}{p} \right\rfloor.$$

We use the notation $\lfloor x \rfloor$ for the integral part of a real number $x \in \mathbf{R}$: $\lfloor x \rfloor \in \mathbf{Z}$ and $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

The same procedure can be applied to n_1. Recursively, we obtain (for $n_0 = n$)

$$\begin{aligned} n_0 &= pn_1 + a_0 & 0 \leq a_0 < p \\ n_1 &= pn_2 + a_1 & 0 \leq a_1 < p \\ &\vdots & \vdots \\ n_{k-1} &= pn_k + a_{k-1} & 0 \leq a_{k-1} < p \\ n_k &= a_k & 0 \leq a_k < p \end{aligned}$$

with

$$n_i = \left\lfloor \frac{n_{i-1}}{p} \right\rfloor = \left\lfloor \frac{n}{p^i} \right\rfloor, \qquad v_p((n_i)!) = n_{i+1} + v_p((n_{i+1})!),$$

which implies that

$$v_p(n!) = n_1 + \cdots + n_k = \sum_{i=1}^{k} \left\lfloor \frac{n}{p^i} \right\rfloor = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor . \tag{2.6.5.1}$$

This formula can also be rewritten in terms of the expansion of n in base p

$$n = a_k p^k + \cdots + a_1 p + a_0 = (a_k \cdots a_0)_p \qquad (a_i = 0, 1, \ldots, p-1)$$

as follows:

$$\begin{aligned} v_p(n!) &= (a_k p^{k-1} + \cdots + a_2 p + a_1) + (a_k p^{k-2} + \cdots + a_3 p + a_2) + \cdots + a_k = \\ &= a_k(1 + p + \cdots + p^{k-1}) + a_{k-1}(1 + p + \cdots + p^{k-2}) + \cdots + a_2(1+p) + a_1 = \\ &= \frac{a_k(p^k - 1) + a_{k-1}(p^{k-1} - 1) \cdots + a_1(p-1)}{p-1} = \frac{n - s_p(n)}{p-1}, \end{aligned} \tag{2.6.5.2}$$

where

$$s_p(n) = a_k + a_{k-1} + \cdots + a_0$$

is the sum of the digits in the expansion of n in base p. Note that Legendre's formula (2.6.5.2) holds also for $n = 0$, hence for all $n \in \mathbf{N}$.

2.6.6 Exercise. (1) For $r \geq 1$ and $0 < a \leq p^r$, the p-adic valuation of $\binom{p^r}{a}$ is equal to $r - v_p(a)$.
(2) For any $m, n \in \mathbf{N}$ and $p \in \mathcal{P}$, the p-adic valuation of $\binom{m+n}{m} = \frac{(m+n)!}{m!n!}$ is equal to the number of times we need to carry over a digit when performing the addition of m and n in base p.

3 Congruences, arithmetic in $\mathbf{Z}/n\mathbf{Z}$

3.1 Basic concepts

3.1.1 A preview Congruences modulo a positive integer n arise when we **retain only the last digit** of the expansion in base n (equivalently, when we retain only the remainder of the division by n) of the various integers involved.

The set of integers $\mathbf{Z}$ then naturally decomposes into n classes, according to this remainder. The usual operations (addition, subtraction, multiplication) can be performed with these classes. In other words, the remainders under the division by n of $a \pm b$ and ab depend only on the remainders of a and b (the division is more subtle).

For example, for $n = 2$,

$$\mathbf{Z} = (2\mathbf{Z}) \cup (2\mathbf{Z}+1), \qquad \begin{array}{c} 2\mathbf{Z} = \{2k \mid k \in \mathbf{Z}\} = \{\text{even numbers}\} \\ 2\mathbf{Z}+1 = \{2k+1 \mid k \in \mathbf{Z}\} = \{\text{odd numbers}\} \end{array}$$

$$\begin{array}{rl} (2k)+(2l) = 2(k+l) & \text{even} + \text{even} = \text{even} \\ (2k)+(2l+1) = 2(k+l)+1 & \text{even} + \text{odd} = \text{odd} \\ (2k+1)+(2l+1) = 2(k+l+1) & \text{odd} + \text{odd} = \text{even} \\ (2k)\cdot(2l) = 2(2kl) & \text{even} \cdot \text{even} = \text{even} \\ (2k)\cdot(2l+1) = 2(2kl+k) & \text{even} \cdot \text{odd} = \text{even} \\ (2k+1)\cdot(2l+1) = 2(2kl+k+l)+1 & \text{odd} \cdot \text{odd} = \text{odd} \end{array}$$

Similarly, we can see immediately that $276 + 389 \neq 667$ by looking at the remainders under the division by 10, since $\ldots 6 + \ldots 9 = \ldots 5$.

From now on until the end of Section 3 we let $m, n \geq 1$ be positive integers.

3.1.2 Definition. Let $a, b \in \mathbf{Z}$. We say that a and b are **congruent modulo** n if they have the same remainder when divided by n (equivalently, when $n|(a-b)$). **Notation:** $a \equiv b \pmod{n}$, or $a \equiv b\ (n)$, or $a \equiv b\ [n]$. The **residue class modulo** n of an integer a is the set

$$a \pmod{n} := \{x \in \mathbf{Z} \mid x \equiv a \pmod{n}\} = \{a + ny \mid y \in \mathbf{Z}\} = a + n\mathbf{Z}.$$

The set of all residue classes modulo n will be denoted by $\mathbf{Z}/n\mathbf{Z}$.

3.1.3 Examples We have $17 \equiv 7 \equiv -13 \pmod{10}$, hence

$$17 \pmod{10} = 7 \pmod{10} = -13 \pmod{10} = \{7, 17, 27, \ldots, -3, -13, -23, \ldots\} = 10\mathbf{Z} + 7 = 10\mathbf{Z} - 3.$$

There are two residue classes modulo 2, namely

$$\begin{array}{l} 0 \pmod{2} = \quad 2 \pmod{2} = -4 \pmod{2} = \{0, 2, 4, 6, \ldots, -2, -4, -6, \ldots\} = 2\mathbf{Z} \\ 1 \pmod{2} = -1 \pmod{2} = -7 \pmod{2} = \{1, 3, 5, 7, \ldots, -1, -3, -5, \ldots\} = 2\mathbf{Z}+1 \end{array}$$

(hence $\mathbf{Z}/2\mathbf{Z} = \{0 \pmod{2},\ 1 \pmod{2}\}$) and three residue classes modulo 3, namely

$$0 \pmod 3 = 3 \pmod 3 = -3 \pmod 3 = \{0, 3, 6, 9, \ldots, -3, -6, -9, \ldots\} = 3\mathbf{Z}$$
$$1 \pmod 3 = 4 \pmod 3 = -2 \pmod 3 = \{1, 4, 7, 10, \ldots, -2, -5, -8, \ldots\} = 3\mathbf{Z} + 1$$
$$2 \pmod 3 = 5 \pmod 3 = -1 \pmod 3 = \{2, 5, 8, 10, \ldots, -1, -4, -7, \ldots\} = 3\mathbf{Z} + 2 = 3\mathbf{Z} - 1$$

(hence $\mathbf{Z}/3\mathbf{Z} = \{0 \pmod 3, \pm 1 \pmod 3\}$). In general, $\mathbf{Z}/n\mathbf{Z}$ contains the following n residue classes

$$\mathbf{Z}/n\mathbf{Z} = \{1 \pmod n, 2 \pmod n, \ldots, n \pmod n\}$$

(with $n \pmod n = 0 \pmod n$). If $n = 2k + 1$ is odd, then we can also write

$$\mathbf{Z}/(2k+1)\mathbf{Z} = \{0 \pmod n, \pm 1 \pmod n, \pm 2 \pmod n, \ldots, \pm k \pmod n\}.$$

Similarly, if $n = 2k$ is even, then

$$\mathbf{Z}/2k\mathbf{Z} = \{0 \pmod n, \pm 1 \pmod n, \pm 2 \pmod n, \ldots, \pm(k-1) \pmod n, k \pmod n\}.$$

3.1.4 Proposition. *If $a \equiv a' \pmod n$ and $b \equiv b' \pmod n$, then*

$$a \pm b \equiv a' \pm b' \pmod n, \qquad ab \equiv a'b' \pmod n.$$

Proof. There exist $x, y \in \mathbf{Z}$ such that $a = a' + nx$ and $b = b' + ny$, which implies that

$$a \pm b = a' \pm b' + n(x \pm y), \qquad ab = (a' + nx)(b' + ny) = a'b' + n(a'y + b'x + nxy).$$

□

3.1.5 Congruences modulo m and mn It is important to understand relations between congruences modulo different integers.

The first observation is that

$$a \equiv b \pmod{mn} \implies a \equiv b \pmod m \tag{3.1.5.1}$$

(for example, if $a \equiv b \pmod{18}$, then $a \equiv b \pmod 6$). Indeed, if $mn \mid (a-b)$, then $m \mid (a-b)$. Equivalently, $mn\mathbf{Z} \subseteq m\mathbf{Z}$, hence $a + mn\mathbf{Z} \subseteq a + m\mathbf{Z}$.

In the opposite direction, each residue class modulo m is a disjoint union of n residue classes modulo mn. For example, let us express the residue class $6\mathbf{Z}+1$ as a disjoint union of three residue classes modulo 18. Firstly,

$$x \equiv 1 \pmod 6 \iff \exists\, y \in \mathbf{Z} \quad x = 6y + 1.$$

Secondly, y belongs to one of the three residue classes modulo 3:

$$y = 3z, \quad \text{or} \quad y = 3z + 1, \quad \text{or} \quad y = 3z + 2,$$

for some $z \in \mathbf{Z}$. Putting these two conditions together, we obtain:

$$x \equiv 1 \pmod 6 \iff \exists\, z \in \mathbf{Z} \quad \begin{cases} x = 6(3z) + 1 = 18z + 1, \text{ or} \\ x = 6(3z+1) + 1 = 18z + 7, \text{ or} \\ x = 6(3z+2) + 1 = 18z + 13, \end{cases}$$

hence

$$x \equiv 1 \pmod 6 \iff \begin{cases} x \equiv 1 \pmod{18}, \text{ or} \\ x \equiv 7 \pmod{18}, \text{ or} \\ x \equiv 13 \pmod{18}. \end{cases}$$

In general,

$$\mathbf{Z} = \coprod_{b=0}^{n-1} (b + n\mathbf{Z}), \qquad a + m\mathbf{Z} = \coprod_{b=0}^{n-1} (a + m(b + n\mathbf{Z})) = \coprod_{b=0}^{n-1} ((a + bm) + mn\mathbf{Z})$$

(where the symbol $\coprod$ denotes a disjoint union), which implies that

$$x \equiv a \pmod m \iff x \equiv a, a + m, a + 2m, \ldots, a + (n-1)m \pmod{mn}. \tag{3.1.5.2}$$

3.1.6 Proposition. *For any $m, n \geq 1$ and $a, b \in \mathbf{Z}$, it is equivalent:*

$$a \equiv b \pmod m \iff na \equiv nb \pmod{mn}.$$

Proof. The statement follows from the equivalences

$$a \equiv b \pmod m \iff m \mid (a - b) \iff mn \mid n(a - b) \iff na \equiv nb \pmod{mn}.$$

□

3.1.7 Proposition (A very useful property of congruences)**.** *It is equivalent:*

$$\left\{ \begin{matrix} a \equiv b \pmod m \\ a \equiv b \pmod n \end{matrix} \right\} \iff a \equiv b \pmod{\operatorname{lcm}(m, n)}$$

Proof. We need to show that

$$\left\{ \begin{matrix} m \mid (a - b) \\ n \mid (a - b) \end{matrix} \right\} \iff \operatorname{lcm}(m, n) \mid (a - b),$$

but this is true by the defining property of $\operatorname{lcm}(m, n)$:

$$\{\text{multiples of } m\} \cap \{\text{multiples of } n\} = \{\text{multiples of } \operatorname{lcm}(m, n)\}.$$

□

3.1.8 Corollary. *If $a, b \in \mathbf{Z}$ and if $n = p_1^{k_1} \cdots p_r^{k_r}$ is the prime factorisation of an integer $n \geq 1$, then it is equivalent:*

$$a \equiv b \pmod n \iff \forall i = 1, \ldots, r \quad a \equiv b \pmod{p_i^{k_i}}.$$

3.2 The values of $a^k \pmod n$

3.2.1 Fermat's little theorem revisited In the language of congruences, Fermat's little theorem 1.5.26 states that

$$\forall p \in \mathcal{P} \quad \forall a \in \mathbf{Z} \qquad a^p \equiv a \pmod p.$$

We are going to investigate, first numerically, then theoretically, the behaviour of the sequence of residue classes $1, a, a^2, a^3, \ldots \pmod n$.

3.2.2 Example: $a^k \pmod 3$ The table

$a \pmod 3$	$a^2 \pmod 3$	$a^3 \pmod 3$
0	$0^2 \equiv 0$	$0 \cdot 0 \equiv 0$
± 1	$(\pm 1)^2 \equiv 1$	$(\pm 1) \cdot 1 \equiv \pm 1$

shows that

$$\forall a \in \mathbf{Z} \quad a^3 \equiv a \pmod 3, \qquad a^2 \equiv \begin{cases} 0 \pmod 3, & 3 \mid a \\ 1 \pmod 3, & 3 \nmid a, \end{cases} \tag{3.2.2.1}$$

which implies that

$$\forall a \in \mathbf{Z}\ \forall k \in \mathbf{N}_+ \quad a^{2k-1} \equiv a \pmod 3, \qquad a^{2k} \equiv \begin{cases} 0 \pmod 3, & 3 \mid a \\ 1 \pmod 3, & 3 \nmid a. \end{cases} \tag{3.2.2.2}$$

Note that we have proved Fermat's little theorem for $p = 3$ by this method.
Application: for any fixed $k \in \mathbf{Z}$, the equation

$$x^2 - 3y^2 = 3k - 1 \qquad (x, y \in \mathbf{Z})$$

has no solution, since

$$x^2 - 3y^2 \equiv x^2 \equiv 0, 1 \pmod 3, \qquad 3k - 1 \equiv -1 \not\equiv 0, 1 \pmod 3.$$

3.2.3 Example: $a^k \pmod 4$ The table

$a \pmod 4$	$a^2 \pmod 4$
0	$0^2 \equiv 0$
2	$2^2 \equiv 0$
± 1	$(\pm 1)^2 \equiv 1$

shows that

$$\forall a \in \mathbf{Z} \quad a^2 \equiv \begin{cases} 0 \pmod 4, & 2 \mid a \\ 1 \pmod 4, & 2 \nmid a. \end{cases} \tag{3.2.3.1}$$

Again, this implies that

$$\forall a \in \mathbf{Z}\ \forall k \in \mathbf{N}_+ \quad a^{2k} \equiv \begin{cases} 0 \pmod 4, & 2 \mid a \\ 1 \pmod 4, & 2 \nmid a. \end{cases} \tag{3.2.3.2}$$

Application: for any fixed $k \in \mathbf{Z}$, the equation

$$x^2 + y^2 = 4k + 3 \qquad (x, y \in \mathbf{Z})$$

has no solution, since

$$4k + 3 \equiv 3 \pmod 4, \qquad x^2, y^2 \equiv 0, 1 \pmod 4, \qquad x^2 + y^2 \equiv 0, 1, 2 \pmod 4 \not\equiv 3 \pmod 4.$$

3.2.4 Example: $a^2 \pmod 8$ The table

$a \pmod 8$	$a^2 \pmod 8$
0	$0^2 \equiv 0$
± 2	$(\pm 2)^2 \equiv 4$
4	$4^2 \equiv 0$
± 1	$(\pm 1)^2 \equiv 1$
± 3	$(\pm 3)^2 \equiv 1$

shows that

$$\forall a \in \mathbf{Z} \quad a^2 \equiv \begin{cases} 0, 4 \pmod 8, & 2 \mid a \\ 1 \pmod 8, & 2 \nmid a. \end{cases} \tag{3.2.4.1}$$

Application: for any fixed $k \in \mathbf{Z}$, the equation

$$x^2 + y^2 + z^2 = 8k + 7 \qquad (x, y, z \in \mathbf{Z})$$

has no solution, since

$$8k+7 \equiv 7 \pmod 8, \qquad x^2, y^2, z^2 \equiv 0, 1, 4 \pmod 8, \qquad x^2+y^2+z^2 \equiv 0, 1, 2, 3, 4, 5, 6 \pmod 8 \not\equiv 7 \pmod 8.$$

3.2.5 Example: $a^k \pmod 5$ The table

$a \pmod 5$	$a^2 \pmod 5$	$a^4 \pmod 5$	$a^5 \pmod 5$
0	$0^2 \equiv 0$	$0^2 \equiv 0$	$0 \cdot 0 \equiv 0$
± 1	$(\pm 1)^2 \equiv 1$	$1^2 \equiv 1$	$(\pm 1) \cdot 1 \equiv \pm 1$
± 2	$(\pm 2)^2 \equiv 4 \equiv -1$	$(-1)^2 \equiv 1$	$(\pm 2) \cdot 1 \equiv \pm 2$

shows that

$$\forall a \in \mathbf{Z} \quad a^5 \equiv a \pmod 5, \qquad a^4 \equiv \begin{cases} 0 \pmod 5, & 5 \mid a \\ 1 \pmod 5, & 5 \nmid a, \end{cases} \tag{3.2.5.1}$$

which proves Fermat's little theorem for $p = 5$.

3.2.6 Example: $a^k \pmod{3^2}$ **for** $3 \nmid a$ The table

$a \pmod{3^2}$	$a^3 \pmod{3^2}$	$a^6 \pmod{3^2}$
± 1	$(\pm 1)^3 \equiv \pm 1$	$(\pm 1)^2 \equiv 1$
± 2	$(\pm 2)^3 \equiv \pm 8 \equiv \mp 1$	$(\mp 1)^2 \equiv 1$
± 4	$(\pm 4)^3 \equiv \pm 64 \equiv \pm 1$	$(\pm 1)^2 \equiv 1$

shows that

$$\forall a \in \mathbf{Z} \quad \left[3 \nmid a \implies a^3 \equiv \pm 1 \pmod{3^2}, \qquad a^6 \equiv 1 \pmod{3^2}\right] \tag{3.2.6.1}$$

$$\forall a \in \mathbf{Z} \qquad a^3 \equiv 0, \pm 1 \pmod{3^2}. \tag{3.2.6.2}$$

Application: if $x, y, z \in \mathbf{Z}$ and $3 \nmid xyz$, then $x^3, y^3, z^3 \equiv \pm 1 \pmod{3^2}$, which implies that $x^3 + y^3 \equiv 0, \pm 2 \not\equiv z^3 \pmod{3^2}$.

3.2.7 Example: $a^k \pmod{3^3}$ for $3 \nmid a$ According to 3.2.6, $a^6 \equiv 1 \pmod{3^2}$, which is equivalent to

$$a^6 \equiv 1, 1+3^2, 1+2\cdot 3^2 \equiv 1, 10, -8 \pmod{3^3}.$$

The table

$a^6 \pmod{3^3}$	$a^{12} \pmod{3^3}$	$a^{18} \pmod{3^3}$
1	$1^2 \equiv 1$	$1 \cdot 1 \equiv 1$
10	$10^2 \equiv 100 \equiv -8$	$10 \cdot (-8) \equiv 1$
-8	$(-8)^2 \equiv \pm 64 \equiv 10$	$(-8) \cdot 10 \equiv 1$

shows that

$$\forall a \in \mathbf{Z} \quad \left[3 \nmid a \implies a^{18} \equiv 1 \pmod{3^3}\right]. \tag{3.2.7.1}$$

3.2.8 Example: $a^k \pmod{5^2}$ for $5 \nmid a$ We have shown in (3.2.5.1) that if $a \in \mathbf{Z}$ is not divisible by 5, then $a^4 \equiv 1 \pmod 5$, which is, in turn, equivalent to

$$a^4 \equiv 1,\ 1+5,\ 1+2\cdot 5,\ 1+3\cdot 5,\ 1+4\cdot 5 \equiv 1, 6, 11, -9, -4 \pmod{5^2}. \tag{3.2.8.1}$$

The table

$a^4 \pmod{5^2}$	$a^8 \pmod{5^2}$	$a^{16} \pmod{5^2}$	$a^{20} \pmod{5^2}$
1	$1^2 \equiv 1$	$1^2 \equiv 1$	$1 \cdot 1 \equiv 1$
-4	$(-4)^2 \equiv 16 \equiv -9$	$(-9)^2 \equiv 81 \equiv 6$	$(-4) \cdot 6 \equiv 1$
6	$6^2 \equiv 36 \equiv 11$	$11^2 \equiv 121 \equiv -4$	$6 \cdot (-4) \equiv 1$
-9	$(-9)^2 \equiv 6$	$6^2 \equiv 11$	$(-9) \cdot 11 \equiv 1$
11	$11^2 \equiv -4$	$(-4)^2 \equiv -9$	$11 \cdot (-9) \equiv 1$

then implies that

$$\forall a \in \mathbf{Z} \quad \left[5 \nmid a \implies a^{20} \equiv 1 \pmod{5^2}\right]. \tag{3.2.8.2}$$

3.2.9 Fermat's Last Theorem (FLT) Fermat claimed that he was able to show that, for any $n \geq 3$, the equation

$$x^n + y^n = z^n \qquad (x, y, z \in \mathbf{Z}) \tag{3.2.9.1}$$

has no solution with $xyz \neq 0$. He gave a proof for $n = 4$ using his method of infinite descent, but the general case was solved only in 1995 by Wiles (in the 19th century, very important contributions to this question were made by Kummer).

Fermat's result for $n = 4$ implies that it is sufficient to consider only the case when n is a prime number $n = p > 2$. It turns out that in this situation the case $p \nmid xyz$ (the **first case**) is much easier than the case $p \mid xyz$ (the **second case**).

For example, we were able to verify the first case of FLT for $p = 3$ in Section 3.2.6 by considering congruences modulo 3^2. A more sophisticated method due to Sophie Germain (and then further developed by Legendre) proves the first case of FLT for all primes $p < 100$ (among others).

3.2.10 Exercise. (1) Determine all possible values of $a^5 \pmod{5^2}$ if $5 \nmid a$. [Hint: It may be helpful to consult first Proposition 4.1.5 below.]
(2) Show that $x^5 + y^5 \not\equiv z^5 \pmod{5^2}$ if $x, y, z \in \mathbf{Z}$ and $5 \nmid xyz$.
(3) What happens if one replaces 5 by 7?

3.2.11 Exercise. Let $x, y, z \in \mathbf{Z}$.
(1) If $3 \mid (x^2 + y^2)$, then $3 \mid x$ and $3 \mid y$.
(2) If $x^2 + y^2 = 3z^2$, then $3 \mid x$, $3 \mid y$ and $3 \mid z$.
(3) If $x^2 + y^2 = 3z^2$, then $x = y = z = 0$.
(4) What happens if one replaces 3 by 5 (resp. by 7)?

3.2.12 Exercise. Let $x, y, z \in \mathbf{Z}$.
(1) If $4 \mid (x^2 + y^2 + z^2)$, then $2 \mid x$, $2 \mid y$ and $2 \mid z$.
(2) If $x^2 + y^2 + z^2 \equiv 3 \pmod{4}$, then $2 \nmid x$, $2 \nmid y$, $2 \nmid z$ and $x^2 + y^2 + z^2 \equiv 3 \pmod{8}$.
(3) $x^2 + y^2 + z^2 \neq 4^k(8l + 7)$ for any $k, l \in \mathbf{N}$.
[The **three square theorem (Legendre, Gauss)** states that a positive integer $n \neq 4^k(8l + 7)$ can always be written as $n = x^2 + y^2 + z^2$ for suitable $x, y, z \in \mathbf{Z}$.]

3.3 The Chinese Remainder Theorem (CRT)

3.3.1 Example: $\mathbf{Z}/6\mathbf{Z}$ and $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$ We know from (3.1.5.1) that a residue class $x \pmod 6$ modulo 6 determines a residue class $x \pmod 2$ modulo 2, as well as a residue class $x \pmod 3$ modulo 3. In concrete terms, this correspondence is given by the following table.

$x \pmod 6$	$x \pmod 2$	$x \pmod 3$
0	0	0
1	1	1
2	0	2
3	1	0
4	0	1
5	1	2
$3a + 4b$	$a = a \cdot 1 + b \cdot 0$	$b = a \cdot 0 + b \cdot 1$

Note that every possible combination

$$(0,0), (0,1), (0,2), (1,0), (1,1), (1,2)$$

appears in the last two columns of the table precisely once. In other words, each residue class modulo 6 is uniquely determined by the corresponding pair of residue classes modulo 2 and 3, respectively.

In scientific terms, this means that the map

$$\begin{array}{ccc} \mathbf{Z}/6\mathbf{Z} & \longrightarrow & \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z} \\ x \pmod 6 & \mapsto & (x \pmod 2, x \pmod 3) \end{array} \tag{3.3.1.1}$$

is bijective: every element on each side corresponds to exactly one element on the other side. We have used the standard notation

$$X \times Y = \{(x, y) \mid x \in X,\, y \in Y\} \tag{3.3.1.2}$$

for the cartesian product of two sets X and Y.

In fact, the last line of the table gives a formula for the inverse of the bijective map (3.3.1.1), namely

$$\begin{array}{rcl} \mathbf{Z}/6\mathbf{Z} & \longrightarrow & \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z} \\ 3a + 4b \pmod 6 & \mapsto & (a \pmod 2, b \pmod 3). \end{array} \tag{3.3.1.3}$$

In concrete terms, this means that, for any $a, b \in \mathbf{Z}$, the system

$$\left\{ \begin{array}{l} x \equiv a \pmod 2 \\ x \equiv b \pmod 3 \end{array} \right\} \tag{3.3.1.4}$$

has a unique solution modulo 6, namely

$$x \equiv 3a + 4b \pmod 6. \tag{3.3.1.5}$$

The formula (3.3.1.3) is analogous to the formula

$$\begin{array}{rcl} & & \mathbf{R}^2 \longrightarrow \quad \mathbf{R} \times \mathbf{R} \\ \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \mapsto & \qquad (a, b) \end{array} \tag{3.3.1.6}$$

expressing an arbitrary vector in the plane $\mathbf{R}^2$ in terms of the vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

of the standard basis of $\mathbf{R}^2$. The residue classes $3 \pmod 6$ and $4 \pmod 6$, which are the respective solutions of

$$\left\{ \begin{array}{l} 3 \equiv 1 \pmod 2 \\ 3 \equiv 0 \pmod 3 \end{array} \right\}, \qquad \left\{ \begin{array}{l} 4 \equiv 0 \pmod 2 \\ 4 \equiv 1 \pmod 3 \end{array} \right\} \tag{3.3.1.7}$$

are analogues of e_1 and e_2. The following formulas are also analogous to each other.

$$\left\{ \begin{array}{l} 3a + 4b \equiv a \pmod 2 \\ 3a + 4b \equiv b \pmod 3 \end{array} \right\}, \qquad \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.3.1.8}$$

The general case is as follows.

3.3.2 Theorem (The Chinese Remainder Theorem (CRT))**.** *Assume that $m, n \geq 1$ and $\gcd(m, n) = 1$. For all $a, b \in \mathbf{Z}$, the system*

$$\left\{ \begin{array}{l} x \equiv a \pmod m \\ x \equiv b \pmod n \end{array} \right\} \tag{3.3.2.1}$$

has a unique solution modulo mn. In other words, the map

$$\begin{array}{rcl} \mathbf{Z}/mn\mathbf{Z} & \longrightarrow & \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z} \\ x \pmod{mn} & \mapsto & (x \pmod m, x \pmod n) \end{array}$$

is bijective.

Proof. **Uniqueness:** if $x \equiv y \equiv a \pmod m$ and $x \equiv y \equiv b \pmod n$, then $m \mid (x - y)$ and $n \mid (x - y)$, which implies that $x - y$ is divisible by $\operatorname{lcm}(m, n) = mn/\gcd(m, n) = mn$.
Construction: as in (3.3.1.7), we first find solutions when $(a, b) = (1, 0)$ and $(a, b) = (0, 1)$, respectively, and we then combine them, as in (3.3.1.8), to obtain a general solution.

Euclid's algorithm applied to (m,n) produces $u,v \in \mathbf{Z}$ satisfying the Bézout relation $mu+nv = \gcd(m,n) = 1$. This implies that

$$\left\{\begin{matrix} nv \equiv 1 \pmod m \\ nv \equiv 0 \pmod n \end{matrix}\right\}, \qquad \left\{\begin{matrix} mu \equiv 0 \pmod m \\ mu \equiv 1 \pmod n \end{matrix}\right\} \tag{3.3.2.2}$$

as in (3.3.1.7), hence

$$x \equiv a(nv)+b(mu) = a+(b-a)mu = b+(a-b)nv \equiv \left\{\begin{matrix} a\cdot 1 + b\cdot 0 \equiv a \pmod m \\ a\cdot 0 + b\cdot 1 \equiv b \pmod n \end{matrix}\right\}. \tag{3.3.2.3}$$

□

3.3.3 Remarks (1) The numbers appearing in the formula $x \equiv a(nv)+b(mu) \pmod{mn}$ can be quite large. However, if $av \equiv A \pmod m$ and $bu \equiv B \pmod n$ with $|A|, |B|$ not too big, then $x \equiv nA+mB \pmod{mn}$ (this congruence follows from Proposition 3.1.6) and $nA+mB$ is not too big, either. (2) The sets $\mathbf{Z}/mn\mathbf{Z}$ and $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ have the same cardinality. This means that, once we know that the map in the theorem is injective (which is equivalent to the uniqueness statement), it is automatically bijective. Unlike the proof above, this argument does not give an explicit formula for the inverse map.

3.3.4 Solving a system of congruences (example) We are going to solve the system of congruences

$$\left\{\begin{matrix} x \equiv 5 \pmod{27} \\ x \equiv 7 \pmod{37} \end{matrix}\right\}. \tag{3.3.4.1}$$

We apply modified Euclid's algorithm to the pair $37, 27$:

$$37 = 1\cdot 27 + 10, \qquad 27 = 3\cdot 10 + (-3), \qquad 10 = (-3)\cdot(-3)+1, \qquad -3 = (-3)\cdot 1 + 0,$$

which gives linear combinations

$$10 = 37-27, \qquad -3 = 27-3(37-27) = 4\cdot 27 - 3\cdot 37, \qquad 1 = (37-27)+3(4\cdot 27 - 3\cdot 37) = 11\cdot 27 - 8\cdot 37.$$

Equivalently, we can use the continued fraction

$$1+\cfrac{1}{3+\cfrac{1}{-3+\frac{1}{-3}}} = 1+\cfrac{1}{3+\frac{-3}{10}} = 1+\frac{10}{27} = \frac{37}{27}, \qquad 1+\cfrac{1}{3+\frac{1}{-3}} = 1+\frac{3}{8} = \frac{11}{8}$$

or its more formal form

j	-2	-1	0	1	2	3
a_j			1	3	-3	-3
p_j	0	1	1	4	-11	37
q_j	1	0	1	3	-8	27

to obtain the same Bézout relation

$$11\cdot 27 + (-8)\cdot 37 = 1.$$

Therefore

$$x \equiv 5 \cdot (-8) \cdot 37 + 7 \cdot 11 \cdot 27 \equiv 7 + (7-5) \cdot 8 \cdot 37 \pmod{27 \cdot 37}.$$

Note that $(7-5) \cdot 8 \equiv 16 \equiv -11 \pmod{27}$; it follows that $(7-5) \cdot 8 \cdot 37 \equiv -11 \cdot 37 \pmod{27 \cdot 37}$ (again, by Proposition 3.1.6), hence

$$x \equiv 7 - 11 \cdot 37 \equiv -400 \equiv 599 \pmod{27 \cdot 37}$$

(since $27 \cdot 37 = 999$).

3.3.5 Another approach to the CRT A system of congruences

$$\left\{\begin{matrix} x \equiv a \pmod{m} \\ x \equiv b \pmod{n} \end{matrix}\right\} \tag{3.3.5.1}$$

(without any conditions on m and n) can be reformulated in an equivalent way as follows: the first congruence is equivalent to $x = a + my$, for some $y \in \mathbf{Z}$. Substituting this equality to the second congruence yields

$$my \equiv b - a \pmod{n}. \tag{3.3.5.2}$$

Congruences of this type will be studied in Section 3.4.

Another reformulation uses the fact that the second congruence is equivalent to the existence of $z \in \mathbf{Z}$ such that $x = b + nz$. Putting the two conditions together, we obtain the equation $a + my = b + nz$, which can be written in the form

$$my - nz = b - a \qquad (y, z \in \mathbf{Z}) \tag{3.3.5.3}$$

(which is equivalent to (3.3.5.2)). The equation 3.3.5.3 can be solved by the method described in the proof of Theorem 2.5.4; this will then give a solution of the system (3.3.5.1) (provided a solution exists).

3.3.6 Exercise. Assume that $m_1, \ldots, m_r \geq 1$ and $\gcd(m_i, m_j) = 1$ for all $1 \leq i < j \leq r$. Then, for any $a_1, \ldots, a_r \in \mathbf{Z}$, the system of congruences

$$\left\{\begin{matrix} x \equiv a_1 \pmod{m_1} \\ \vdots \\ x \equiv a_r \pmod{m_r} \end{matrix}\right\}$$

has a unique solution $x \pmod{m_1 \cdots m_r}$.

3.3.7 Exercise. (1) Let $a, b \in \mathbf{Z}$. Show that the system

$$\left\{\begin{matrix} x \equiv a \pmod{4} \\ x \equiv b \pmod{6} \end{matrix}\right\}$$

has a solution if and only if $a \equiv b \pmod{2}$. If this is the case, then the solution is unique modulo 12.
(2) What happens if one replaces the numbers $4, 6$ by arbitrary $m, n \in \mathbf{N}_+$?
(3) Is there a common generalisation of (2) and Exercise 3.3.6?

3.4 Invertible elements in $\mathbf{Z}/n\mathbf{Z}$, congruences $ax \equiv b \pmod n$

3.4.1 Inverse of a residue class Throughout Section 3.4, $n \geq 1$ is a positive integer. We are interested in the congruence

$$ax \equiv 1 \pmod n. \tag{3.4.1.1}$$

A solution of this congruence — if it exists — should be thought of as an inverse of $a \pmod n$. For example, the congruence $3 \cdot 5 \equiv 1 \pmod 7$ tells us that $5 \pmod 7$ is an inverse of $3 \pmod 7$.

3.4.2 Theorem (Invertible elements in $\mathbf{Z}/n\mathbf{Z}$)**.** *If $a \in \mathbf{Z}$ and $n \geq 1$, then the congruence $ax \equiv 1 \pmod n$ has a solution if and only if $\gcd(a, n) = 1$. If this is the case, then there is a unique solution $x \pmod n$; we say that the residue class $a \pmod n$ is **invertible** and that $x \pmod n$ is its **inverse**. **Notation:** we denote $x \pmod n$ by $\frac{1}{a} \pmod n$ or by $a^{-1} \pmod n$. The set of all invertible elements of $\mathbf{Z}/n\mathbf{Z}$ will be denoted by $(\mathbf{Z}/n\mathbf{Z})^*$ or by $(\mathbf{Z}/n\mathbf{Z})^\times$.*

Proof. **Uniqueness:** if $ax \equiv ay \equiv 1 \pmod n$, then $y \equiv 1 \cdot y \equiv (ax)y \equiv x(ay) \equiv x \cdot 1 \equiv x \pmod n$.
Existence: this follows from the equivalences

$$\exists x \in \mathbf{Z} \;\; ax \equiv 1 \pmod n \iff \exists x, y \in \mathbf{Z} \;\; ax + ny = 1 \iff 1 \in a\mathbf{Z} + n\mathbf{Z} = \gcd(a, n)\mathbf{Z} \iff \gcd(a, n) = 1.$$

□

3.4.3 Computing the inverse of $a \pmod n$ If n is small, then one can usually determine the inverse directly. In general, one applies Euclid's algorithm (in its modified form) to find $d = \gcd(a, n)$ and an explicit Bézout relation $au + nv = d$. The residue class $a \pmod n$ is invertible if and only if $d = 1$. If this is the case, then $au \equiv 1 \pmod n$, which means that the inverse of $a \pmod n$ is equal to $u \pmod n$.

For example, the calculations from Section 2.4.5 give

$$5 \cdot 11 \equiv 1 \pmod{18}, \quad 11^{-1} \pmod{18} = 5 \pmod{18}, \quad 12 \cdot 13 \equiv 1 \pmod{31}, \quad 13^{-1} \pmod{31} = 12 \pmod{31}.$$

3.4.4 Exercise. (1) Determine all invertible elements in $\mathbf{Z}/12\mathbf{Z}$. Compute the inverse of each of them.
(2) Idem for $\mathbf{Z}/18\mathbf{Z}$.
(3) If both $a \pmod n$ and $b \pmod n$ are invertible, so is their product $ab \pmod n$, and its inverse is given by $(ab)^{-1} \equiv a^{-1}b^{-1} \pmod n$.

3.4.5 Powers of $a \pmod n$ If $a \pmod n$ is an invertible residue class, we define, for any $m \in \mathbf{N}_+$,

$$a^{-m} \pmod n := (a^{-1})^m \pmod n,$$

which is also equal to $(a^m)^{-1} \pmod n$, by Exercise 3.4.4(3).

3.4.6 Exercise. If $a \pmod n$ is invertible, show that

$$\forall l, m \in \mathbf{Z} \quad a^l \cdot a^m \equiv a^{l+m} \pmod n, \qquad (a^l)^m \equiv a^{lm} \pmod n.$$

3.4.7 Dividing congruences Assume that $a, b, c \in \mathbf{Z}$ satisfy $a \neq 0$ and

$$ac \equiv b \pmod n. \tag{3.4.7.1}$$

We would like to divide the congruence (3.4.7.1) by a, if possible.

Case 1: a **divides** n**.** This implies that a also divides b. The congruence (3.4.7.1) is equivalent to the existence of $y \in \mathbf{Z}$ such that

$$ac + ny = b. \tag{3.4.7.2}$$

As both terms n and b are divisible by a, the equality (3.4.7.2) is equivalent to

$$c + (n/a)y = b/a. \tag{3.4.7.3}$$

Consequently, the congruence (3.4.7.1) is equivalent to

$$c \equiv (b/a) \pmod{n/a} \tag{3.4.7.4}$$

(see also Proposition 3.1.6).

Case 2: $\gcd(a, n) = 1$**.** Theorem 3.4.2 tells us that there exists a unique inverse $a^{-1} \pmod{n}$ of $a \pmod{n}$. After multiplying (3.4.7.1) by $a^{-1} \pmod{n}$, we obtain

$$c \equiv ba^{-1} \pmod{n}. \tag{3.4.7.5}$$

Conversely, if we multiply (3.4.7.5) by $a \pmod{n}$, we obtain the original congruence (3.4.7.1). Therefore (3.4.7.5) is equivalent to (3.4.7.1).

Case 3: the general case. Let $d := \gcd(a, n)$; then d divides b. We first apply Case 1 with d instead of a, obtaining

$$(a/d)c \equiv (b/d) \pmod{n/d}. \tag{3.4.7.6}$$

As $\gcd(a/d, n/d) = 1$, we can then apply Case 2 to (3.4.7.6), obtaining

$$c \equiv (b/d)(a/d)^{-1} \pmod{n/d} \tag{3.4.7.7}$$

(which is equivalent to the original congruence (3.4.7.1)).

Example: we want to simplify the congruence

$$12x \equiv 15 \pmod{21}. \tag{3.4.7.8}$$

It is equivalent to the existence of $y \in \mathbf{Z}$ such that

$$12x + 21y = 15;$$

this equality can be divided by 3, which yields

$$4x + 7y = 5,$$

hence an equivalent congruence

$$4x \equiv 5 \pmod{7}. \tag{3.4.7.9}$$

As $4 \cdot 2 \equiv 1 \pmod{7}$, we then multiply (3.4.7.9) by $4^{-1} \pmod{7} = 2 \pmod{7}$, obtaining

$$x \equiv 2 \cdot 4x \equiv 2 \cdot 5 \equiv 3 \pmod{7}, \tag{3.4.7.10}$$

which is equivalent to the original congruence (3.4.7.8), by the previous discussion.

The general congruence

$$ax \equiv b \pmod{n}$$

can be treated in the same way. The final result is as follows.

3.4.8 Theorem. *Let $a, b, n \in \mathbf{Z}$ be integers such that $a \neq 0$ and $n \geq 1$; let $d := \gcd(a, n)$.*
(1) If $d \nmid b$, then the congruence $ax \equiv b \pmod{n}$ has no solution.
(2) If $d \mid b$, then the congruence $ax \equiv b \pmod{n}$ is equivalent to $(a/d)x \equiv (b/d) \pmod{n/d}$, which has a unique solution modulo n/d, namely $x \equiv (b/d)(a/d)^{-1} \pmod{n/d}$.

Proof. Everything was already done in Section 3.4.7 (for $c = x$). If we are only interested in the existence and uniqueness of a solution (but not in an explicit formula), then we can argue directly as follows.
Uniqueness: if $ax \equiv ay \equiv b \pmod{n}$, then there exists $z \in \mathbf{Z}$ such that $a(x-y) = nz$, hence $(a/d)(x-y) = (n/d)z$ is divisible by n/d. As $\gcd(a/d, n/d) = 1$, it follows from Lemma 2.3.8 that n/d divides $x - y$, hence $x \equiv y \pmod{n/d}$.
Existence: this follows from the equivalences

$$\exists x \in \mathbf{Z} \quad ax \equiv b \pmod{n} \iff \exists x, y \in \mathbf{Z} \quad ax + ny = b \iff b \in a\mathbf{Z} + n\mathbf{Z} = \gcd(a, n)\mathbf{Z} \iff \gcd(a, n) \mid b.$$

□

3.4.9 Example Let us solve the system (3.3.4.1)

$$\left.\begin{matrix} x \equiv 5 \pmod{27} \\ x \equiv 7 \pmod{37} \end{matrix}\right\}$$

by the methods developed in the previous sections. The second congruence is equivalent to $x = 7 + 37y$ for some $y \in \mathbf{Z}$. Substitution to the first congruence gives

$$5 \equiv x \equiv 7 + 37y \equiv 7 + 10y \pmod{27},$$

which is equivalent to $10y \equiv -2 \equiv 25 \pmod{27}$, hence to $2y \equiv 5 \pmod{27}$ (since $\gcd(5, 27) = 1$) and $y \equiv 14 \cdot 2y \equiv 14 \cdot 5 \equiv -11 \pmod{27}$. Therefore $y = -11 + 27z$ for some $z \in \mathbf{Z}$ and the original system is equivalent to

$$x = 7 - 11 \cdot 37 + 27 \cdot 37z \equiv 7 - 11 \cdot 37 \equiv -400 \pmod{27 \cdot 37}.$$

3.4.10 Exercise. (1) Solve the system of congruences $21x \equiv 33 \pmod{45}$, $15x \equiv 6 \pmod{66}$.
(2) Solve the system of congruences $x \equiv 9 \pmod{15}$, $x \equiv 3 \pmod{16}$, $x \equiv 13 \pmod{17}$.

4 Euler's function φ, Euler's theorem

4.1 Consequences of Fermat's little theorem

4.1.1 Examples and a preview In Section 3.2 we proved the following congruences for $a \in \mathbf{Z}$.

$$3 \nmid a \implies a^2 \equiv 1 \pmod 3, \qquad a^6 \equiv 1 \pmod{3^2}, \qquad a^{18} \equiv 1 \pmod{3^3}, \tag{4.1.1.1}$$

$$5 \nmid a \implies a^4 \equiv 1 \pmod 5, \qquad a^{20} \equiv 1 \pmod{5^2}. \tag{4.1.1.2}$$

One is tempted to speculate that this pattern holds in general, namely, that

$$\forall p \in \mathcal{P} \quad p \nmid a \implies a^{p-1} \equiv 1 \pmod p, \tag{4.1.1.3}$$

$$\implies a^{(p-1)p^{k-1}} \equiv 1 \pmod{p^k} \qquad (k \geq 1). \tag{4.1.1.4}$$

This is, indeed, the case, as we are now going to show. We deduce (4.1.1.3) from Fermat's little theorem

$$\forall p \in \mathcal{P} \quad \forall a \in \mathbf{Z} \qquad a^p \equiv a \pmod p \tag{4.1.1.5}$$

and then apply Proposition 4.1.5 below on improvement of congruences to prove (4.1.1.4). After that we use Corollary 3.1.8 to obtain analogous congruences modulo any $n = p_1^{k_1} \cdots p_r^{k_r} \geq 1$. In particular, we prove Euler's theorem

$$\gcd(a,n) = 1 \implies a^{\varphi(n)} \equiv 1 \pmod n, \qquad \varphi(n) = \varphi(p_1^{k_1}) \cdots \varphi(p_r^{k_r}), \qquad \varphi(p^k) = (p-1)p^{k-1} \tag{4.1.1.6}$$

and its improvement (Theorem 4.1.11).

In Section 4.2 we give an alternative treatement of Euler's function φ and Euler's theorem from a more conceptual point of view.

4.1.2 Proposition (Euler's theorem for $n = p$)**.** *If $p \in \mathcal{P}$ and if $a \in \mathbf{Z}$, $p \nmid a$, then $a^{p-1} \equiv 1 \pmod p$.*

Proof. If we multiply the congruence given by Fermat's little theorem (4.1.1.5) by the inverse $a^{-1} \pmod p$ of $a \pmod p$ (it exists, since $\gcd(a,p) = 1$, by assumption), then we obtain $a^p a^{-1} \equiv aa^{-1} \pmod p$, hence the desired congruence

$$a^{p-1} \equiv a^{p-1} a a^{-1} \equiv a^p a^{-1} \equiv a a^{-1} \equiv 1 \pmod p.$$

□

4.1.3 Exercise. Show that, for any $a \in \mathbf{Z}$ and $m, n \in \mathbf{N}_+$, $a^{m+10n} \equiv a^m \pmod{11}$. Determine $b \equiv 2019^{9102} \pmod{11}$, $0 \leq b \leq 10$.

4.1.4 Improvement of congruences by $x \mapsto x^p$ The general principle is very simple: for each prime number p, the **Frobenius map** sending x to x^p improves congruences modulo powers of p.

For example, in the simplest case when $a \equiv b \pmod 2$, we can write $a = b + 2c$ for some $c \in \mathbf{Z}$, which implies that

$$a^2 = (b+2c)^2 = b^2 + 4bc + 4c^2 \equiv b^2 \pmod{2^2}.$$

In general, we use the divisibility of the binomial coefficients $\binom{p}{j}$ $(0 < j < p)$ by p, proved in Proposition 1.5.25.

4.1.5 Proposition. *If p is a prime and $a, b \in \mathbf{Z}$ satisfy $a \equiv b \pmod{p^k}$ $(k \geq 1)$, then $a^p \equiv b^p \pmod{p^{k+1}}$.*

Proof. There exists $c \in \mathbf{Z}$ such that $a = b + p^k c$, hence

$$a^p = (b + p^k c)^p = b^p + \binom{p}{1} b^{p-1}(p^k c) + \binom{p}{2} b^{p-2}(p^k c)^2 + \cdots + \binom{p}{p-1} b(p^k c)^{p-1} + p^{pk} c^p.$$

Each term $\binom{p}{j} b^{p-j}(p^k c)^j$ for $0 < j < p$ is divisible by $p \cdot p^k = p^{k+1}$, thanks to Proposition 1.5.25. The last term $p^{pk} c^p$ is also divisible by p^{k+1}, since $pk \geq 2k \geq k + 1$. □

4.1.6 Proposition (Euler's theorem for $n = p^k$)**.** *If $p \in \mathcal{P}$ and if $a \in \mathbf{Z}$, $p \nmid a$, then $a^{(p-1)p^{k-1}} \equiv 1 \pmod{p^k}$ holds for all $k \geq 1$.*

Proof. For $k = 1$ this was proved in Proposition 4.1.2. The general case follows by induction from Proposition 4.1.5: if we know that $a^{(p-1)p^{k-1}} \equiv 1 \pmod{p^k}$, then

$$a^{(p-1)p^k} \equiv (a^{(p-1)p^{k-1}})^p \equiv 1^p \equiv 1 \pmod{p^{k+1}}.$$

□

4.1.7 Improvement for $p = 2$ For $p = 2$ the previous proposition says that

$$2 \nmid a \implies a^2 \equiv 1 \pmod 4, \qquad a^4 \equiv 1 \pmod 8, \qquad a^8 \equiv 1 \pmod{16}, \qquad a^{16} \equiv 1 \pmod{32}, \qquad \cdots$$

However, we showed in Section 3.2.4 that

$$2 \nmid a \implies a^2 \equiv 1 \pmod 8,$$

which implies, by a repeated application of Proposition 4.1.5, that

$$2 \nmid a \implies a^2 \equiv 1 \pmod 8, \qquad a^4 \equiv 1 \pmod{16}, \qquad a^8 \equiv 1 \pmod{32}, \qquad \cdots \tag{4.1.7.1}$$

We record this result (see also Exercise 1.1.7) as a formal proposition.

4.1.8 Proposition (Improved Euler's theorem for $n = 2^k$)**.** *If $k \geq 3$ and $2 \nmid a$, then $a^{2^{k-2}} \equiv 1 \pmod{2^k}$.*

4.1.9 Examples modulo 15, 35 and 504 What happens for moduli that are divisible by more than one prime? The answer is simple: we combine the previous results for several primes using the general principle proved in Proposition 3.1.7 and its Corollary 3.1.8.

For example, $n = 15 = 3 \cdot 5 = \operatorname{lcm}(3, 5)$. If $\gcd(a, 15) = 1$ (which is equivalent to $3 \nmid a$ and $5 \nmid a$), then

$$\left.\begin{array}{l} 3 \nmid a \implies a^{3-1} = a^2 \equiv 1 \pmod 3 \implies a^4 \equiv 1 \pmod 3 \\ 5 \nmid a \implies a^{5-1} = a^4 \equiv 1 \pmod 5 \end{array}\right\} \implies a^4 \equiv 1 \pmod{\operatorname{lcm}(3,5)} \equiv 1 \pmod{15}.$$

Similarly, $n = 35 = 5 \cdot 7$. If $\gcd(a, 35) = 1$, then

$$\left.\begin{array}{l} 5 \nmid a \implies a^{5-1} = a^4 \equiv 1 \pmod 5 \implies a^{12} \equiv (a^4)^3 \equiv 1 \pmod 5 \\ 7 \nmid a \implies a^{7-1} = a^6 \equiv 1 \pmod 7 \implies a^{12} \equiv (a^6)^2 \equiv 1 \pmod 7 \end{array}\right\} \implies a^{12} \equiv 1 \pmod{\underbrace{\operatorname{lcm}(5,7)}_{35}}.$$

Finally, $n = 504 = 7 \cdot 8 \cdot 9 = 2^3 \cdot 3^2 \cdot 7$. If $\gcd(a, 504) = 1$, then

$$\left.\begin{array}{l} 2 \nmid a \implies a^2 \equiv 1 \pmod{2^3} \implies a^6 \equiv (a^2)^3 \equiv 1 \pmod{2^3} \\ 3 \nmid a \implies a^{(3-1)\cdot 3} \equiv a^6 \equiv 1 \pmod{3^2} \\ 7 \nmid a \implies a^{7-1} \equiv a^6 \equiv 1 \pmod 7 \end{array}\right\} \implies a^6 \equiv 1 \pmod{\underbrace{\operatorname{lcm}(2^3, 3^2, 7)}_{504}}.$$

4.1.10 Notation: Euler's function φ Euler's function will be defined and investigated in detail in Section 4.2 below. Here we use its explicit values (proved in Theorem 4.2.5 below) simply as a convenient notation.

For a prime number p and $k \geq 1$, we let

$$\varphi(p^k) := (p-1)p^{k-1} = p^k - p^{k-1} = p^k(1 - \tfrac{1}{p}).$$

Using this notation, Proposition 4.1.6 states that $a^{\varphi(p^k)} \equiv 1 \pmod{p^k}$ if $p \nmid a$.

For an integer $n \geq 1$ with prime factorisation $n = p_1^{k_1} \cdots p_r^{k_r}$ ($r \geq 0$, $k_i \geq 1$) we let

$$\varphi(n) := \varphi(p_1^{k_1}) \cdots \varphi(p_r^{k_r}) = (p_1^{k_1} - p_1^{k_1-1}) \cdots (p_r^{k_r} - p_r^{k_r-1}) = n(1 - \tfrac{1}{p_1}) \cdots (1 - \tfrac{1}{p_r}).$$

4.1.11 Theorem (Improved Euler's theorem)**.** *Let $n = p_1^{k_1} \cdots p_r^{k_r} \geq 1$ be an integer (here p_i are distinct primes, $r \geq 0$, $k_i \geq 1$). If $\gcd(a, n) = 1$ and if $u \geq 1$ is divisible by $\varphi(p_i^{k_i}) = p_i^{k_i} - p_i^{k_i-1}$ for all $i = 1, \ldots, r$, then*

$$a^u \equiv 1 \pmod{n}.$$

Moreover, if $p_1 = 2$ and $k_1 \geq 3$, then one can replace $\varphi(p_1^{k_1}) = 2^{k_1-1}$ above by $\varphi(p_1^{k_1})/2 = 2^{k_1-2}$.

Proof. For each $i = 1, \ldots, r$, write $u = \varphi(p_i^{k_i})v_i$. According to Proposition 4.1.6,

$$a^{\varphi(p_i^{k_i})} \equiv 1 \pmod{p_i^{k_i}} \implies a^u \equiv \left(a^{\varphi(p_i^{k_i})}\right)^{v_i} \equiv 1 \pmod{p_i^{k_i}},$$

which implies the same congruence modulo $\operatorname{lcm}(p_1^{k_1}, \ldots, p_r^{k_r}) = n$, by Proposition 3.1.7.

In the case $p_1 = 2$ and $k_1 \geq 3$ we appeal to Proposition 4.1.8 instead of Proposition 4.1.6. □

4.1.12 Improvement of Euler's theorem (optimal version) The smallest value of u satisfying the assumptions of the previous theorem is equal to

$$u = \operatorname{lcm}(\varphi(p_1^{k_1}), \ldots, \varphi(p_r^{k_r})) \tag{4.1.12.1}$$

(again, with $\varphi(p_1^{k_1})$ replaced by $\varphi(p_1^{k_1})/2$ if $p_1 = 2$ and $k_1 \geq 3$).

4.1.13 Theorem (Euler's theorem)**.** *Let a, n be integers such that $n \geq 1$ and $\gcd(a, n) = 1$. Then $a^{\varphi(n)} \equiv 1 \pmod{n}$.*

Proof. This is a special case of Theorem 4.1.11 for $u = \varphi(n)$. Another proof will be given in 4.2.9 below. □

4.1.14 Comparison of Euler's theorem and its improvement Let us compare the two versions of Euler's theorem for the three values $n = 15, 35$ and 504 considered in Section 4.1.9. As

$$\varphi(15) = \varphi(3)\varphi(5) = (3-1)(5-1) = 8, \qquad \varphi(35) = \varphi(5)\varphi(7) = (5-1)(7-1) = 24,$$
$$\varphi(504) = \varphi(2^3)\varphi(3^2)\varphi(7) = (2^3 - 2^2)(3^2 - 3^1)(7-1) = 144,$$

Euler's theorem tells us that

$$\begin{aligned} \gcd(a, 15) = 1 &\implies a^8 \equiv 1 \pmod{15} \\ \gcd(a, 35) = 1 &\implies a^{24} \equiv 1 \pmod{35} \\ \gcd(a, 504) = 1 &\implies a^{144} \equiv 1 \pmod{504}. \end{aligned}$$

Its improved version that was made explicit in 4.1.9 is much more precise:

$$\begin{aligned} \gcd(a,15)=1 &\implies a^4 \equiv 1 \pmod{15} \\ \gcd(a,35)=1 &\implies a^{12} \equiv 1 \pmod{35} \\ \gcd(a,504)=1 &\implies a^6 \equiv 1 \pmod{504}. \end{aligned}$$

4.1.15 Exercise. (1) Compute $2^7 \pmod{5^3}$, $3^{15} \pmod{5^3}$ and $3^{15} \pmod{2^3}$.
(2) Using the Chinese Remainder Theorem, compute $3^{15} \pmod{1000}$.

4.1.16 Exercise. Let $a \in \mathbf{Z}$.
(1) If $2 \nmid a$ and $3 \nmid a$, then $a^2 \equiv 1 \pmod{24}$.
(2) If $2 \nmid a$, $3 \nmid a$ and $5 \nmid a$, then $a^4 \equiv 1 \pmod{240}$.
(3) If $2 \nmid a$ and $5 \nmid a$, then $a^{100} \equiv 1 \pmod{1000}$.
(4) If $b \in \mathbf{Z}$, then $b^{100} \equiv 0, 1, 376, 625 \pmod{1000}$.

4.1.17 Exercise. Let $a \in \mathbf{Z}$.
(1) Determine all the possible values of $a^{12} \pmod 7$, $a^{12} \pmod{13}$ and $a^{12} \pmod{91}$, for $a \in \mathbf{Z}$.
(2) Idem for a^6 instead of a^{12}.
(3) If $n \geq 1$ is an integer such that $n \equiv 1 \pmod{12}$, then $a^n \equiv a \pmod{91}$ holds for all $a \in \mathbf{Z}$.

4.2 Euler's function φ

4.2.1 Notation We use the notation $|X|$ for the number of elements of a set X. Note that $|X \times Y| = |X| \cdot |Y|$. Recall that, for any $n \in \mathbf{N}_+$,

$$\begin{aligned} \mathbf{Z}/n\mathbf{Z} &= \{\text{all residue classes } (\bmod n)\} = \{1 \pmod n,\ 2 \pmod n, \ldots,\ n \pmod n\} \\ (\mathbf{Z}/n\mathbf{Z})^* &= \{\text{invertible residue classes } (\bmod n)\} = \{a \pmod n \mid 1 \leq a \leq n,\ \gcd(a,n)=1\} \end{aligned}$$

The **Euler function** $\varphi : \mathbf{N}_+ \longrightarrow \mathbf{N}_+$ is defined by the formula

$$\varphi(n) := |(\mathbf{Z}/n\mathbf{Z})^*|\,. \tag{4.2.1.1}$$

For example,

$$\begin{gathered} (\mathbf{Z}/1\mathbf{Z})^* = \{1 \pmod 1\}, \quad (\mathbf{Z}/2\mathbf{Z})^* = \{1 \pmod 2\}, \quad (\mathbf{Z}/3\mathbf{Z})^* = \{1 \pmod 3,\ 2 \pmod 3\}, \\ (\mathbf{Z}/4\mathbf{Z})^* = \{1 \pmod 4,\ 3 \pmod 4\}, \quad (\mathbf{Z}/5\mathbf{Z})^* = \{1 \pmod 5,\ 2 \pmod 5,\ 3 \pmod 5,\ 4 \pmod 5\}, \\ (\mathbf{Z}/6\mathbf{Z})^* = \{1 \pmod 6,\ 5 \pmod 6\}, \end{gathered}$$

which implies that

$$\varphi(1)=1, \quad \varphi(2)=1, \quad \varphi(3)=2, \quad \varphi(4)=2, \quad \varphi(5)=4, \quad \varphi(6)=2.$$

4.2.2 Example: $(\mathbf{Z}/6\mathbf{Z})^*$ and $(\mathbf{Z}/2\mathbf{Z})^* \times (\mathbf{Z}/3\mathbf{Z})^*$ Note that $\varphi(6) = \varphi(2)\varphi(3)$, which suggests a possible link to the Chinese Remainder Theorem.

We reproduce below the table from Section 3.3.1 which makes explicit the correspondence between $\mathbf{Z}/6\mathbf{Z}$ and $\mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z}$. We mark the invertible elements in each column by putting them into a box.

$x \pmod 6$	$x \pmod 2$	$x \pmod 3$
0	0	0
$\boxed{1}$	$\boxed{1}$	$\boxed{1}$
2	0	$\boxed{2}$
3	$\boxed{1}$	0
4	0	$\boxed{1}$
$\boxed{5}$	$\boxed{1}$	$\boxed{2}$

We see that a residue class modulo 6 is invertible if an only if both the correponding residue class modulo 2 and the residue class modulo 3 are invertible.

In other works, under the bijective map (3.3.1.1)

$$\begin{array}{rcl} \mathbf{Z}/6\mathbf{Z} & \longrightarrow & \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/3\mathbf{Z} \\ x \pmod 6 & \mapsto & (x \pmod 2, x \pmod 3) \end{array}$$

the subset $(\mathbf{Z}/6\mathbf{Z})^*$ corresponds to $(\mathbf{Z}/2\mathbf{Z})^* \times (\mathbf{Z}/3\mathbf{Z})^*$. In particular, both subsets have the same number of elements, which explains the equality $\varphi(6) = \varphi(2)\varphi(3)$ that we observed earlier.

This phenomenon is completely general, as we are now going to show.

4.2.3 Proposition. *If $a \in \mathbf{Z}$, $m, n \geq 1$ and $\gcd(m,n) = 1$, then $a \pmod{mn}$ is invertible in $\mathbf{Z}/mn\mathbf{Z}$ if and only if both $a \pmod m$ and $a \pmod n$ are invertible in $\mathbf{Z}/m\mathbf{Z}$ and $\mathbf{Z}/n\mathbf{Z}$, respectively.*
In other words, under the bijective map $\mathbf{Z}/mn\mathbf{Z} \longrightarrow \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$ in the Chinese Remainder Theorem 3.3.2, the subset $(\mathbf{Z}/mn\mathbf{Z})^*$ corresponds to $(\mathbf{Z}/m\mathbf{Z})^* \times (\mathbf{Z}/n\mathbf{Z})^*$.

Proof. If $a \pmod{mn}$ is invertible in $\mathbf{Z}/mn\mathbf{Z}$, then there exists $b \in \mathbf{Z}$ such that $ab \equiv 1 \pmod{mn}$, which implies that $ab \equiv 1 \pmod m$ and $ab \equiv 1 \pmod n$.

Conversely, if $a \pmod m$ is invertible in $\mathbf{Z}/m\mathbf{Z}$ and $a \pmod n$ is invertible in $\mathbf{Z}/n\mathbf{Z}$, then there exist $b, c \in \mathbf{Z}$ such that $ab \equiv 1 \pmod m$ and $ac \equiv 1 \pmod n$. As $\gcd(m,n) = 1$, the Chinese Remainder Theorem implies that there exists $x \in \mathbf{Z}$ such that $x \equiv b \pmod m$ and $x \equiv c \pmod n$. It follows that $ax \equiv ab \equiv 1 \pmod m$ and $ax \equiv ac \equiv 1 \pmod n$, hence $ax \equiv 1 \pmod{\operatorname{lcm}(m,n)} \equiv 1 \pmod{mn}$. □

4.2.4 Corollary. *If $m, n \geq 1$ and $\gcd(m,n) = 1$, then $\varphi(mn) = \varphi(m)\varphi(n)$.*

Proof. By Proposition 4.2.3, both subsets $(\mathbf{Z}/mn\mathbf{Z})^*$ and and $(\mathbf{Z}/m\mathbf{Z})^* \times (\mathbf{Z}/n\mathbf{Z})^*$ have the same number of elements. □

4.2.5 Theorem (Properties of $\varphi(n)$)**.** *(1) If p is a prime, then $\varphi(p) = p - 1$.*
(2) If p is a prime and $k \geq 1$, then $\varphi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1) = p^k(1 - \frac{1}{p})$.
(3) If $m, n \geq 1$ and $\gcd(m,n) = 1$, then $\varphi(mn) = \varphi(m)\varphi(n)$.
(4) If $n = p_1^{k_1} \cdots p_r^{k_r}$ (where $r \geq 0$, p_i are distinct primes, $k_i \geq 1$), then

$$\varphi(n) = \varphi(p_1^{k_1}) \cdots \varphi(p_r^{k_r}) = (p_1^{k_1} - p_1^{k_1 - 1}) \cdots (p_r^{k_r} - p_r^{k_r - 1}) = n(1 - \tfrac{1}{p_1}) \cdots (1 - \tfrac{1}{p_r}) = n \prod_{p \mid n} (1 - \tfrac{1}{p}). \quad (4.2.5.1)$$

Proof. (1) $(\mathbf{Z}/p\mathbf{Z})^* = \{1 \pmod p, 2 \pmod p, \ldots, p-1 \pmod p\}$.
(2) In this case $(\mathbf{Z}/p^k\mathbf{Z})^* = \{a \pmod{p^k} \mid 1 \leq a \leq p^k, p \nmid a\} = \{a \pmod{p^k} \mid 1 \leq a \leq p^k\} \smallsetminus \{pb \pmod{p^k} \mid 1 \leq b \leq p^k/p\}$, which implies that this set has $p^k - p^k/p$ elements.
Part (3) was proved in Corollary 4.2.4, and Part (4) is a straightforward combination of (2) and (3). □

4.2.6 Remarks and examples (1) If p is a prime, then $\varphi(p^2) = p^2 - p = p(p-1) \neq (p-1)^2 = \varphi(p)^2$. (2) $\varphi(120) = \varphi(2^3 \cdot 3 \cdot 5) = (2^3 - 2^2)(3-1)(5-1) = 4 \cdot 2 \cdot 4 = 32 = 2^5$.

4.2.7 Exercise. Show that $\varphi(2n)/\varphi(n)$ is equal to 1 (resp. to 2) if $2 \nmid n$ (resp. if $2 \mid n$). What happens if one replaces $2n$ by $3n$ (resp. by $6n$)?

4.2.8 The inclusion-exclusion principle The formula (4.2.5.1) for $\varphi(n)$ can be proved directly using the inclusion-exclusion principle, as we are going to explain.

Example 1: $n = 12 = 2^2 \cdot 3$ In this case $\gcd(x, 12) = 1 \iff 2 \nmid x$ and $3 \nmid x$.
The subsets $A, B \subset \mathbf{Z}/12\mathbf{Z}$ defined as

$$A = \{1 \leq x \leq 12;\ 2 \mid x\} = \{2a \mid 1 \leq a \leq 6\} = \{2, 4, 6, 8, 10, 12\}$$
$$B = \{1 \leq x \leq 12;\ 3 \mid x\} = \{3b \mid 1 \leq b \leq 4\} = \{3, 6, 9, 12\}$$

satisfy

$$A \cap B = \{1 \leq x \leq 12;\ 6 \mid x\} = \{6c \mid 1 \leq c \leq 2\} = \{6, 12\}$$
$$A \cup B = \{1 \leq x \leq 12;\ x \pmod{12} \text{ not invertible}\} = \{2, 4, 6, 8, 10, 12, 3, 9\}$$
$$(\mathbf{Z}/12\mathbf{Z})^* = \{1 \leq x \leq 12;\ x \pmod{12} \text{ invertible}\} = \mathbf{Z}/12\mathbf{Z} \smallsetminus (A \cup B) = \{1, 5, 7, 11\}.$$

As

$$|A \cup B| = |A| + |B| - |A \cap B|, \qquad |A| = 12/2, \qquad |B| = 12/3, \qquad |A \cap B| = 12/6,$$

it follows that

$$\varphi(12) = 12 - |A \cup B| = 12 - |A| - |B| + |A \cap B| = 12\left(1 - \tfrac{1}{2} - \tfrac{1}{3} + \tfrac{1}{6}\right) = 12(1 - \tfrac{1}{2})(1 - \tfrac{1}{3}). \quad (4.2.8.1)$$

Example 2: $n = p_1^{k_1} p_2^{k_2} p_3^{k_3}$ In this case $\gcd(x, n) = 1 \iff p_1 \nmid x$, $p_2 \nmid x$ and $p_3 \nmid x$. If we define subsets $A, B, C \subset \mathbf{Z}/n\mathbf{Z}$ by the formulas

$$A = \{1 \leq x \leq n;\ p_1 \mid x\} = \{p_1 a \mid 1 \leq a \leq n/p_1\}$$
$$B = \{1 \leq x \leq n;\ p_2 \mid x\} = \{p_2 b \mid 1 \leq b \leq n/p_2\}$$
$$C = \{1 \leq x \leq n;\ p_3 \mid x\} = \{p_3 c \mid 1 \leq c \leq n/p_3\},$$

then

$$A \cap B = \{1 \leq x \leq n;\ p_1 p_2 \mid x\}, \quad A \cap C = \{1 \leq x \leq n;\ p_1 p_3 \mid x\}, \quad B \cap C = \{1 \leq x \leq n;\ p_2 p_3 \mid x\},$$
$$A \cap B \cap C = \{1 \leq x \leq n;\ p_1 p_2 p_3 \mid x\}, \quad A \cup B \cup C = \{1 \leq x \leq n;\ x \pmod{n} \text{ not invertible}\},$$

which implies that $\varphi(n) = n - |A \cup B \cup C|$ and

$$|A| = n/p_1,\ |B| = n/p_2,\ |C| = n/p_3,\ |A \cap B| = n/p_1 p_2,$$
$$|A \cap C| = n/p_1 p_3,\ |B \cap C| = n/p_2 p_3,\ |A \cap B \cap C| = n/p_1 p_2 p_3. \quad (4.2.8.2)$$

On the other hand,

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|, \tag{4.2.8.3}$$

hence

$$\varphi(n) = n\left(1 - \tfrac{1}{p_1} - \tfrac{1}{p_2} - \tfrac{1}{p_3} + \tfrac{1}{p_1p_2} + \tfrac{1}{p_1p_3} + \tfrac{1}{p_2p_3} - \tfrac{1}{p_1p_2p_3}\right) = n\left(1 - \tfrac{1}{p_1}\right)\left(1 - \tfrac{1}{p_2}\right)\left(1 - \tfrac{1}{p_3}\right). \tag{4.2.8.4}$$

Example 3: $n = p_1^{k_1} \cdots p_r^{k_r}$ We leave this general case as an exercise.

4.2.9 Theorem (Euler's Theorem (bis))**.** *Let a, n be integers such that $n \geq 1$ and $\gcd(a, n) = 1$. Then $a^{\varphi(n)} \equiv 1 \pmod{n}$.*

Proof. We give an algebraic proof which will work in the general abstract context of finite abelian groups (see Corollary 7.5.9 below). Denote the invertible elements in $\mathbf{Z}/n\mathbf{Z}$ by $x_1 \pmod{n}, \ldots, x_{\varphi(n)} \pmod{n}$. The idea is to multiply each of them by $a \pmod{n}$ and consider the elements $ax_1 \pmod{n}, \ldots, ax_{\varphi(n)} \pmod{n}$. Each residue class $ax_j \pmod{n}$ is again invertible in $\mathbf{Z}/n\mathbf{Z}$, and these classes are distinct (indeed, if $ax \equiv ay \pmod{n}$, then $x \equiv a^{-1}ax \equiv a^{-1}ay \equiv y \pmod{n}$). This implies that

$$\{x_1 \pmod{n}, \ldots, x_{\varphi(n)} \pmod{n}\} = (\mathbf{Z}/n\mathbf{Z})^* = \{ax_1 \pmod{n}, \ldots, ax_{\varphi(n)} \pmod{n}\}. \tag{4.2.9.1}$$

For example, if $n = 12$ and $a = 5$, then

$$\{1, 5, 7, 11 \pmod{12}\} = (\mathbf{Z}/12\mathbf{Z})^* = \{5, 1, 11, 7 \pmod{12}\}$$

(since $5 \cdot 1 \equiv 5$, $5 \cdot 5 \equiv 1$, $5 \cdot 7 \equiv 11$, $5 \cdot 11 \equiv 7 \pmod{12}$).

The equality (4.2.9.1) implies that the product of all invertible residue classes is equal both to

$$x := x_1 \cdots x_{\varphi(n)} \pmod{n} \tag{4.2.9.2}$$

and

$$(ax_1) \cdots (ax_{\varphi(n)}) \pmod{n} \equiv a^{\varphi(n)}x \pmod{n}, \tag{4.2.9.3}$$

hence

$$a^{\varphi(n)}x \equiv x \pmod{n} \tag{4.2.9.4}$$

(with the residue classes on both sides being invertible, being products of invertible classes). We can, therefore, multiply (4.2.9.4) by $x^{-1} \pmod{n}$, obtaining the desired congruence $a^{\varphi(n)} \equiv 1 \pmod{n}$. □

4.2.10 Consequences of Euler's theorem (1) If $n = p$ is a prime, then $\varphi(p) = p - 1$ and we recover the fact that $p \nmid a$ implies $a^{p-1} \equiv 1 \pmod{p}$ (Proposition 4.1.2).
(2) More generally, if p is a prime and $k \geq 1$, then $\varphi(p^k) = (p-1)p^{k-1}$ and we obtain that $p \nmid a$ implies $a^{(p-1)p^{k-1}} \equiv 1 \pmod{p^k}$ (Proposition 4.1.6).
(3) Note that Fermat's little theorem follows immediately from (1): if $p \mid a$, then $a^p \equiv 0 \equiv a \pmod{p}$, whereas if $p \nmid a$, then (1) implies that $a^p \equiv a \cdot a^{p-1} \equiv a \cdot 1 \equiv a \pmod{p}$. As we saw in the proof of Proposition 4.1.2, this argument can be reversed.
(4) Is there an analogue of Fermat's little theorem for congruences $\pmod{n}$ if n is not a prime? In other words, is there an integer $m > 1$ such that $a^m \equiv a \pmod{n}$ holds for all $a \in \mathbf{Z}$? The answer is given in Proposition 4.2.11 below. This result will be used in Section 4.4.2 on cryptographic applications and also in Section 5.3.

4.2.11 Proposition. *Let $n > 1$ be an integer.*
(1) If there exists a prime p such that $p^2 \mid n$, then there is no $m > 1$ satisfying

$$\forall a \in \mathbf{Z} \quad a^m \equiv a \pmod{n}.$$

(2) If $n = p_1 \cdots p_r$ is a product of $r \geq 1$ distinct primes and if $m \geq 1$ satisfies $(p_i - 1) \mid (m-1)$ for all $i = 1, \ldots, r$ (equivalently, if $\operatorname{lcm}(p_1 - 1, \ldots, p_r - 1) \mid (m-1)$), then

$$\forall a \in \mathbf{Z} \quad a^m \equiv a \pmod{n}.$$

Proof. (1) For $a = p$ and $m > 1$, $p^2 \mid a^m$ and $p^2 \nmid a$, which implies that $p^2 \nmid (a^m - a)$. Therefore $n \nmid (a^m - a)$.
(2) According to Proposition 3.1.7 and its Corollary 3.1.8, it is enough to show that

$$\forall a \in \mathbf{Z} \quad a^m \equiv a \pmod{p_i}$$

holds for each $i = 1, \ldots, r$. Write $(m-1) = (p_i - 1)t_i$. If $p_i \mid a$, then $a^m \equiv 0 \equiv a \pmod{p_i}$. If $p_i \nmid a$, then $a^{p_i - 1} \equiv 1 \pmod{p_i}$, hence

$$a^m \equiv a \cdot (a^{p_i - 1})^{t_i} \equiv a \cdot 1 \equiv a \pmod{p_i}.$$

□

4.2.12 Examples (1) $\forall a \in \mathbf{Z} \quad a^{21} \equiv a \pmod{55}$. In this case $n = 55 = 5 \cdot 11$ and $m - 1 = 20$.
(2) $\forall a \in \mathbf{Z} \quad a^{561} \equiv a \pmod{561}$. In this case $m = n = 561 = 3 \cdot 11 \cdot 17$ and $m - 1 = 560 = 2^4 \cdot 5 \cdot 7$. In the terminology of Definition 5.3.4, 561 is a **Carmichael number**.

4.3 Structure of $(\mathbf{Z}/n\mathbf{Z})^*$

4.3.1 Motivation We know that if two residue classes $a \pmod{n}$ and $b \pmod{n}$ are invertible, so is their product $ab \pmod{n}$, as well as the inverse $a^{-1} \pmod{n}$ (in the abstract language of Section 7 below, $(\mathbf{Z}/n\mathbf{Z})^*$ is an abelian group with respect to multiplication). However, multiplication in $(\mathbf{Z}/n\mathbf{Z})^*$ is much more complicated than addition in $\mathbf{Z}/n\mathbf{Z}$.

This happens already for real numbers: multiplication of (positive) real numbers is complicated, but it can be reduced to addition using the logarithm map with respect to a fixed base $a > 1$

$$\log_a : (\mathbf{R}_{>0}, \cdot) \longrightarrow (\mathbf{R}, +), \qquad \log_a(a^t) = t, \qquad \log_a(xy) = \log_a(x) + \log_a(y). \tag{4.3.1.1}$$

This works because each positive real number can be written as a suitable power a^t ($t \in \mathbf{R}$) of a.

Can something similar be done for $(\mathbf{Z}/n\mathbf{Z})^*$ instead of $\mathbf{R}_{>0}$? Only in the case if each element of $(\mathbf{Z}/n\mathbf{Z})^*$ can be written as a power of a fixed invertible residue class $a \pmod{n}$ (which is then called a **generator of** $(\mathbf{Z}/n\mathbf{Z})^*$, or a **primitive root modulo** n).

Let us investigate whether such a generator exists for small values of n.

4.3.2 Looking for generators of $(\mathbf{Z}/n\mathbf{Z})^*$ For $n = 3, 4, 6$ we have $(\mathbf{Z}/n\mathbf{Z})^* = \{\pm 1 \pmod{n}\}$, with $-1 \pmod{n}$ being a generator of $(\mathbf{Z}/n\mathbf{Z})^*$.

Example 1: $n = 5$**.**

$a \pmod 5$	$a^2 \pmod 5$	$a^3 \pmod 5$	$a^4 \pmod 5$
1	1	1	1
2	4	3	1
3	4	2	1
4	1	4	1

We see that both $2 \pmod 5$ and $3 \pmod 5 = 2^{-1} \pmod 5$ are generators of $(\mathbf{Z}/5\mathbf{Z})^*$.

Example 2: $n = 7$**.**

$a \pmod 7$	$a^2 \pmod 7$	$a^3 \pmod 7$	$a^4 \pmod 7$	$a^5 \pmod 7$	$a^6 \pmod 7$
1	1	1	1	1	1
2	4	1	2	4	1
3	2	6	4	5	1
4	2	1	4	2	1
5	4	6	2	3	1
6	1	6	1	6	1

Both $3 \pmod 7$ and $5 \pmod 7 = 3^{-1} \pmod 7$ are generators of $(\mathbf{Z}/7\mathbf{Z})^*$.

Example 3: $n = 8$**.**

$a \pmod 8$	$a^2 \pmod 8$	$a^3 \pmod 8$	$a^4 \pmod 8$
1	1	1	1
3	1	3	1
5	1	5	1
7	1	7	1

In this case there are no generators of $(\mathbf{Z}/8\mathbf{Z})^*$.

Example 4: $n = 9$**.**

$a \pmod 9$	$a^2 \pmod 9$	$a^3 \pmod 9$	$a^4 \pmod 9$	$a^5 \pmod 9$	$a^6 \pmod 9$
1	1	1	1	1	1
2	4	8	7	5	1
4	7	1	4	7	1
5	7	8	4	2	1
7	4	1	7	4	1
8	1	8	1	8	1

Both $2 \pmod 9$ and $5 \pmod 9 = 2^{-1} \pmod 9$ are generators of $(\mathbf{Z}/9\mathbf{Z})^*$.

Example 5: $n = 15$**.** In this case $(\mathbf{Z}/15\mathbf{Z})^*$ has $\varphi(15) = (3-1)(5-1) = 8$ elements, but we know from Section 4.1.9 that $a^4 \equiv 1 \pmod{15}$ if $\gcd(a, 15) = 1$. This means that $a^{4k+l} \equiv (a^4)^k a^l \equiv a^l \pmod{15}$ can take at most 4 values (corresponding to $l = 0, 1, 2, 3$), and therefore $a \pmod{15}$ is never a generator of $(\mathbf{Z}/15\mathbf{Z})^*$.

As we are going to see, this argument applies in a very general context. It tells us that a generator of $(\mathbf{Z}/n\mathbf{Z})^*$ can exist only in the case when the exponent $\varphi(n)$ in Euler's theorem cannot be improved (in other words, when $u = \varphi(n)$, in the notation of Theorem 4.1.11).

4.3.3 Definition. Assume that $a, n \in \mathbf{Z}$, $n \geq 1$ and $\gcd(a, n) = 1$. The **order** of the residue class $a \pmod n$ in $(\mathbf{Z}/n\mathbf{Z})^*$ is the smallest integer $d \geq 1$ such that $a^d \equiv 1 \pmod n$ (Euler's theorem implies that $d \leq \varphi(n)$). Note that $a^{kd+l} \equiv (a^d)^k a^l \equiv 1^k \cdot a^l \equiv a^l \pmod n$, for all $k, l \in \mathbf{Z}$.

4.3.4 Order of $a \pmod n$ in $(\mathbf{Z}/n\mathbf{Z})^*$ (examples) It is easy to extract from the tables in Section 4.3.2 the orders of all invertible residue classes $\pmod n$.

$a \pmod 5$	$1 \pmod 5$	$2 \pmod 5$	$3 \pmod 5$	$4 \pmod 5$
d	1	4	4	2

$a \pmod 7$	$1 \pmod 7$	$2 \pmod 7$	$3 \pmod 7$	$4 \pmod 7$	$5 \pmod 7$	$6 \pmod 7$
d	1	3	6	3	6	2

$a \pmod 8$	$1 \pmod 8$	$3 \pmod 8$	$5 \pmod 8$	$7 \pmod 8$
d	1	2	2	2

$a \pmod 9$	$1 \pmod 9$	$2 \pmod 9$	$4 \pmod 9$	$5 \pmod 9$	$7 \pmod 9$	$8 \pmod 9$
d	1	6	3	6	3	2

In these examples, the order of any element of $(\mathbf{Z}/n\mathbf{Z})^*$ divides $\varphi(n)$, and it is equal to $\varphi(n)$ if and only if the element is a generator of $(\mathbf{Z}/n\mathbf{Z})^*$. This happens in general.

4.3.5 Exercise. For $n \in \mathbf{N}$, let $a_n = 3^n$, $b_n = 4^n$ and $c_n = 1018 \cdot 2018^n + 1026 \cdot 2019^n$.
(1) What can one say about $a_n \pmod{13}$ and $b_n \pmod{13}$?
(2) What can one say about $c_n \pmod{13}$? When is $c_n \equiv 0 \pmod{13}$ (resp. $c_n \equiv 3 \pmod{13}$)?

4.3.6 Proposition. *Assume that $a, n \in \mathbf{Z}$, $n \geq 1$ and $\gcd(a, n) = 1$. Let $d \geq 1$ be the order of $a \pmod n$ in $(\mathbf{Z}/n\mathbf{Z})^*$.*
(1) For $l \in \mathbf{Z}$, it is equivalent: $a^l \equiv 1 \pmod n \iff d \mid l$.
(2) For $l, m \in \mathbf{Z}$, it is equivalent: $a^l \equiv a^m \pmod n \iff d \mid (l-m) \iff l \equiv m \pmod d$.
(3) The set $\{a^m \mid m \in \mathbf{Z}\}$ is equal to $\{a, a^2, \ldots, a^d \equiv 1 \pmod n\}$ and has d elements.

Proof. (1) If $l = dk$, then $a^l \equiv (a^d)^k \equiv 1^k \equiv 1 \pmod n$. Conversely, assume that $a^l \equiv 1 \pmod n$. Apply division with remainder by d to l: write $l = dq + r$, where $q, r \in \mathbf{Z}$, $0 \leq r < d$. As $1 \equiv a^l \equiv (a^d)^q a^r \equiv 1 \cdot a^r \equiv a^r$, minimality of d implies that $r = 0$, hence $l = dq$ is divisible by d.
(2) The equivalences

$$a^l \equiv a^m \pmod n \iff (a^{-1})^m a^l \equiv (a^{-1})^m a^m \pmod n \iff a^{l-m} \equiv 1 \pmod n$$

are automatic, and the last condition is equivalent to $d \mid (l-m)$, by (1).
(3) Any $m \in \mathbf{Z}$ can be written as $m = kd + l$, where $k, l \in \mathbf{Z}$ and $1 \leq l \leq d$; then $a^m \equiv (a^d)^k a^l \equiv a^l \pmod n$. Moreover, if $1 \leq i < j \leq d$, then $0 < j - i < d$, hence $d \nmid (j-i)$ and $a^i \not\equiv a^j \pmod n$, by (2).
□

4.3.7 Corollary. *The order of any element $a \pmod n$ of $(\mathbf{Z}/n\mathbf{Z})^*$ divides $\varphi(n)$ (since $a^{\varphi(n)} \equiv 1 \pmod n$, by Euler's theorem). More precisely, this order divides $u := \operatorname{lcm}(\varphi(p_1^{k_1}), \ldots, \varphi(p_r^{k_r}))$ if $n = p_1^{k_1} \cdots p_r^{k_r}$ (since $a^u \equiv 1 \pmod n$, by an improved version of Euler's theorem (Theorem 4.1.11)). If $p_1 = 2$ and $k_1 \geq 3$, then we can replace $\varphi(p_1^{k_1})$ by $\varphi(p_1^{k_1})/2$.*

4.3.8 Definition. An invertible residue class $a \pmod n \in (\mathbf{Z}/n\mathbf{Z})^*$ is a **generator of** $(\mathbf{Z}/n\mathbf{Z})^*$ (also called a **primitive root modulo** n) if $\{a^m \mid m \in \mathbf{Z}\} = (\mathbf{Z}/n\mathbf{Z})^*$ (i.e., if each invertible residue class $\pmod n$ is a power of $a \pmod n$).

4.3.9 Proposition. *An invertible residue class $a \pmod n$ is a generator of $(\mathbf{Z}/n\mathbf{Z})^*$ if and only if its order in $(\mathbf{Z}/n\mathbf{Z})^*$ is equal to $\varphi(n)$.*

Proof. This is a special case of Proposition 4.3.6(3) (for $d = \varphi(n)$). □

4.3.10 Computing the order of $a \pmod n$ in $(\mathbf{Z}/n\mathbf{Z})^*$ Example: $2 \pmod{67}$. In this case $n = 67$ is a prime, hence $\varphi(67) = 66 = 2 \cdot 3 \cdot 11$. Let d be the order of $2 \pmod{67}$ in $(\mathbf{Z}/67\mathbf{Z})^*$. We know that $d \mid 66$, by Corollary 4.3.7.

We are going to show that $d = 66$, arguing by contradiction. If $d \neq 66$, then there exists a prime $p \mid 66$ such that $d \mid (66/p)$. According to the three cases $p = 2, 3, 11$ we have

$$\begin{aligned} d \mid 3 \cdot 11 &\Longrightarrow 2^{33} \equiv 1 \pmod{67}, \text{ or} \\ d \mid 2 \cdot 11 &\Longrightarrow 2^{22} \equiv 1 \pmod{67}, \text{ or} \\ d \mid 2 \cdot 3 &\Longrightarrow 2^{6} \equiv 1 \pmod{67}. \end{aligned}$$

However,

$$\begin{gathered} 2^6 \equiv 64 \equiv -3 \not\equiv 1 \pmod{67}, \quad 2^{12} \equiv (-3)^2 \equiv 9 \pmod{67}, \quad 2^{24} \equiv 9^2 \equiv 81 \equiv 14 \pmod{67}, \\ 2^{23} \equiv 2^{-1} \cdot 14 \equiv 7 \pmod{67}, \quad 2^{22} \equiv 2^{-1} \cdot 7 \equiv 2^{-1} \cdot (-60) \equiv -30 \not\equiv 1 \pmod{67}, \\ 2^{35} \equiv 2^{12} \cdot 2^{23} \equiv 9 \cdot 7 \equiv -4 \pmod{67}, \quad 2^{33} \equiv 2^{-2} \cdot (-4) \equiv -1 \not\equiv 1 \pmod{67}. \end{gathered}$$

This contradiction shows that $d = 66$, as claimed.

The same argument proves the following general statement.

4.3.11 Proposition. *If $m \geq 1$ and $a^m \equiv 1 \pmod n$, then it is equivalent:*
m is equal to the order of $a \pmod n$ in $(\mathbf{Z}/n\mathbf{Z})^* \iff$ for each prime $p \mid m$ we have $a^{m/p} \not\equiv 1 \pmod n$.

4.3.12 Exercise. Let $a \in \mathbf{Z}$.
(1) If $17 \nmid a$, then it is equivalent: $a \pmod{17}$ is a generator of $(\mathbf{Z}/17\mathbf{Z})^*$ (a primitive root modulo 17) $\iff a^8 \neq 1 \pmod{17}$.
(2) Find such a generator (try small values of a).
(3) If $3 \nmid a$, then it is equivalent: $a \pmod{27}$ is a generator of $(\mathbf{Z}/27\mathbf{Z})^*$ (a primitive root modulo 27) $\iff$ $a^6, a^9 \neq 1 \pmod{27}$.
(4) Find such a generator (try small values of a).

4.3.13 Proposition (Order of $a^k \pmod n$)**.** *Let d be the order of $a \pmod n$ in $(\mathbf{Z}/n\mathbf{Z})^*$. For an integer $k \neq 0$, the order of $a^k \pmod n$ in $(\mathbf{Z}/n\mathbf{Z})^*$ is equal to $d/\gcd(d, |k|)$. In particular, it is also equal to d if and only if $\gcd(d, |k|) = 1$.*

Proof. If we denote the order of $a^k \pmod n$ by $e \geq 1$, then $|k|e \geq 1$ is the smallest positive multiple of $|k|$ for which $a^{|k|e} \equiv 1 \pmod n$. However, this congruence is equivalent to $|k|e$ being divisible by d, thanks to Proposition 4.3.6(1). This means that $|k|e$ is the smallest common multiple of $|k|$ and d, hence

$$e = \operatorname{lcm}(d, |k|)/|k| = d/\gcd(d, |k|).$$

□

4.3.14 Corollary (Number of generators)**.** *(1) If $a \pmod n$ is a generator of $(\mathbf{Z}/n\mathbf{Z})^*$, then the set of generators of $(\mathbf{Z}/n\mathbf{Z})^*$ is equal to $\{a^k \pmod n \mid 1 \leq k \leq \varphi(n), \gcd(\varphi(n), k) = 1\}$.*
(2) The number of generators of $(\mathbf{Z}/n\mathbf{Z})^*$ is equal either to zero, or to $\varphi(\varphi(n))$.

Proof. (1) This is a special case of the last sentence of Proposition 4.3.13. Part (2) is an immediate consequence of (1). □

4.3.15 Proposition. *Assume that $n = n_1 n_2$, where $n_1, n_2 > 2$ and $\gcd(n_1, n_2) = 1$. If $\gcd(a, n) = 1$, then $a^{\varphi(n)/2} \equiv 1 \pmod{n}$. In particular, the order of any $a \pmod{n}$ in $(\mathbf{Z}/n\mathbf{Z})^*$ divides $\varphi(n)/2$, hence $(\mathbf{Z}/n\mathbf{Z})^*$ contains no generator.*

Proof. The assumptions imply that $2 \mid \varphi(n_i)$ and $\varphi(n) = \varphi(n_1)\varphi(n_2)$. Therefore

$$\left.\begin{aligned} a^{\varphi(n)/2} &\equiv (a^{\varphi(n_1)})^{\varphi(n_2)/2} \equiv 1 \pmod{n_1} \\ a^{\varphi(n)/2} &\equiv (a^{\varphi(n_2)})^{\varphi(n_1)/2} \equiv 1 \pmod{n_2} \end{aligned}\right\} \implies a^{\varphi(n)/2} \equiv 1 \pmod{n}.$$

□

4.3.16 Theorem (Existence of a generator in $(\mathbf{Z}/n\mathbf{Z})^*$)**.** *A generator exists in $(\mathbf{Z}/n\mathbf{Z})^* \iff n = 1, 2, 4, p^k, 2p^k$, where $p \neq 2$ is a prime and $k \geq 1$.*

Proof. We are going to prove here only the easy implication "$\Longrightarrow$". As regards the converse implication "$\Longleftarrow$", the key point is to show that $(\mathbf{Z}/p\mathbf{Z})^*$ contains a generator, for every prime p. This result was proved by Gauss (see Theorems 5.5.2 and 5.5.4 below for more details).

Assume that $(\mathbf{Z}/n\mathbf{Z})^*$ contains a generator. If we write $n = p_1^{k_1} \cdots p_r^{p_r}$ for primes $p_1 < \cdots < p_r$, Proposition 4.3.15 implies that n must be of the form $n = 2^k, p^k$ or $2p^k$ (where $k \geq 0$ and $p \neq 2$ is a prime). But if $n = 2^k$ and $k \geq 3$, Proposition 4.1.8 tells us that the order of each element of $(\mathbf{Z}/2^k\mathbf{Z})^*$ is at most equal to $\varphi(2^k)/2$, hence there is no generator in this case. □

4.3.17 Discrete logarithm Assume that $a \pmod{n}$ is a generator of $(\mathbf{Z}/n\mathbf{Z})^*$ (hence n is as in Theorem 4.3.16). For any invertible residue class $x \pmod{n}$ there exists a unique integer $m \in \{0, 1, \ldots, \varphi(n)-1\}$ such that $x \equiv a^m \pmod{n}$. The integer m is called the **discrete logarithm** of $x \pmod{n}$ with respect to $a \pmod{n}$.

If we denote m by $l_a(x \pmod{n})$, then we have

$$l_a(xy \pmod{n}) \equiv l_a(x \pmod{n}) + l_a(y \pmod{n}) \pmod{\varphi(n)}. \tag{4.3.17.1}$$

4.3.18 Decimal expansion of rational numbers Such expansions are always (ultimately) periodic:

$$\frac{2}{3} = 0,666\ldots = 0,\overline{6} \qquad \frac{1}{15} = 0,0666\ldots = 0,0\overline{6} \qquad \frac{3}{7} = 0,428571428571\ldots = 0,\overline{428571}$$
$$\frac{16}{37} = 0,432432\ldots = 0,\overline{432}$$

What can be said about the length of the period? For example,

$$\frac{3}{7} = 0,\overline{428571} \implies 10^6 \cdot \frac{3}{7} = 428571,\overline{428571} \implies (10^6 - 1) \cdot \frac{3}{7} = 428571 \implies$$
$$\implies 7 \mid 3 \cdot (10^6 - 1) \implies 7 \mid (10^6 - 1) \implies 10^6 \equiv 1 \pmod{7}.$$

The order of $10 \pmod{7} = 3 \pmod{7}$ in $(\mathbf{Z}/7\mathbf{Z})^*$ is equal to 6, which is also the length of the period of the decimal expansion of $\frac{3}{7}$.

In general, for each $d \geq 1$, the real number $x := 0,\underbrace{\overline{000\cdots1}}_{d}$ satisfies

$$10^d x = 1,\underbrace{\overline{000\cdots1}}_{d} = 1 + x \implies x = \frac{1}{10^d - 1}.$$

For example,

$$\frac{1}{9} = 0,1111\ldots = 0,\overline{1} \qquad \frac{1}{99} = 0,010101\ldots = 0,\overline{01} \qquad \frac{1}{999} = 0,001001001\ldots = 0,\overline{001}.$$

This implies that, for any decimal digits $a_1, \ldots, a_d \in \{0, 1, \ldots, 9\}$,

$$0,\overline{a_1 a_2 \cdots a_d} = (a_1 \cdots a_d)_{10} \cdot x = \frac{(a_1 \cdots a_d)_{10}}{10^d - 1}. \tag{4.3.18.1}$$

For example,

$$0,\overline{432} = 432 \cdot 0,\overline{001} = \frac{432}{999} = \frac{27 \cdot 16}{27 \cdot 37} = \frac{16}{37}.$$

The formula (4.3.18.1) is equivalent to the following assertion.

4.3.19 Proposition (Decimal expansion of fractions)**.** *Assume that* $0 < \frac{a}{b} < 1$*, where* $a, b \in \mathbf{N}_+$ *and* $\gcd(b, 10) = 1$*; let* $d \geq 1$ *be the order of* $10 \pmod b$ *in* $(\mathbf{Z}/b\mathbf{Z})^*$*. Then* $10^d - 1 = bq$*, where* $q \in \mathbf{N}_+$*, and*

$$\frac{a}{b} = \frac{aq}{bq} = \frac{aq}{10^d - 1} = 0,\overline{a_1 a_2 \cdots a_d}, \qquad (a_1 \cdots a_d)_{10} = aq, \qquad 0 < aq < bq = 10^d - 1.$$

4.3.20 Exercise. (1) What happens if $\gcd(b, 10) > 1$?
(2) Write $0,15\overline{27} = 0,15272727\ldots$ in the form $\frac{a}{b}$.

4.4 Applications to cryptography

4.4.1 Creation of a common secret (Diffie–Hellman) The goal is to create a common secret between two people (Alice and Bob) who send messages to each other through a non-secure communication channel.

Public data: (i) a big integer $n \geq 1$; (ii) an invertible residue class $g \pmod n$ whose order in $(\mathbf{Z}/n\mathbf{Z})^*$ is big.

In practice, $n = p$ is a large prime and $g \pmod p$ is a generator of $(\mathbf{Z}/p\mathbf{Z})^*$.

Step 1: Alice chooses a secret integer $a \in \mathbf{Z}$ and sends to Bob $g^a \pmod n$. Bob chooses a secret integer $b \in \mathbf{Z}$ and sends to Alice $g^b \pmod n$.

Step 2: Alice computes $(g^b)^a \equiv g^{ab} \pmod n$, and Bob computes $(g^a)^b \equiv g^{ab} \pmod n$. The residue class $g^{ab} \pmod n$ is their common secret.

What if there was a spy intercepting their messages? The spy would know the values of $g \pmod n$, $g^a \pmod n$ and $g^b \pmod n$.

In general, it is difficult to compute the value of a (modulo the order of $g \pmod n$) from the knowledge of $g \pmod n$ and $g^a \pmod n$. However, experts keep improving the corresponding algorithms, which means that one should use at present (August 2019) primes $n = p$ of the size at least equal to 2^{1000} (and such that $p - 1$ is divisible by at least one large prime).

4.4.2 Public key cryptography (Rivest–Shamir–Adleman: RSA) In its abstract form, public key cryptography encrypts messages by a publicly known algorithm, whereas the decryption algorithm remains secret. It is based on **one-way functions**.

Such a function is a bijective map $f : X \longrightarrow X$ between a suitable large finite set X and itself, which has the following property: f is easy to compute, but the inverse map $g = f^{-1} : X \longrightarrow X$ is difficult to compute. One uses f (which can be made public) for encryption, and g (which remains secret) for decryption.

The RSA protocol was originally discovered in 1973 by C. Cocks, a mathematician working for the UK intelligence service. It was independently discovered and published by R. Rivest, A. Shamir and L. Adleman in 1977. It uses $X = \mathbf{Z}/n\mathbf{Z}$, where $n = pq$ is a product of two large primes, and maps

$$f(x) \equiv x^e \pmod{n}, \qquad g(y) \equiv y^d \pmod{n}, \tag{4.4.2.1}$$

for suitable integers $d, e \geq 1$. These two maps are inverse to each other if and only if

$$\forall a \in \mathbf{Z} \quad a^{de} \equiv a \pmod{n}. \tag{4.4.2.2}$$

Congruences of this type were discussed in Proposition 4.2.11(2), but we repeat the argument here.

4.4.3 Proposition (Congruences behind RSA)**.** *Let $n = pq$, where $p \neq q$ are prime numbers. If $d, e \geq 1$ are integers such that $(p-1) \mid (de-1)$ and $(q-1) \mid (de-1)$ (which is equivalent to $de \equiv 1 \pmod{\operatorname{lcm}(p-1, q-1)}$), then the maps $f, g : \mathbf{Z}/pq\mathbf{Z} \longrightarrow \mathbf{Z}/pq\mathbf{Z}$ given by*

$$f(x) \equiv x^e \pmod{pq}, \qquad g(y) \equiv y^d \pmod{pq}$$

are inverse to each other:

$$\forall x, y \in \mathbf{Z}/pq\mathbf{Z} \qquad g(f(x)) \equiv (x^e)^d \equiv x^{de} \equiv x \pmod{pq}, \qquad f(g(y)) \equiv (y^d)^e \equiv y^{de} \equiv y \pmod{pq}.$$

Proof. Let $x \in \mathbf{Z}$. If $p \mid x$, then $x^{de} \equiv 0 \equiv x \pmod{p}$. If $p \nmid x$, then $x^{p-1} \equiv 1 \pmod{p}$. As $de = 1 + (p-1)a$ for some $a \in \mathbf{N}$, we have $x^{de} \equiv x \cdot (x^{p-1})^a \equiv x \pmod{p}$. Therefore $x^{de} \equiv x \pmod{p}$ always holds. Similarly, $x^{de} \equiv x \pmod{q}$ always holds. Combining the two congruences using Proposition 3.1.7 and its Corollary 3.1.8, we obtain $x^{de} \equiv x \pmod{pq}$. □

4.4.4 RSA communication Alice wants to receive messages encrypted by a publicly known algorithm, while keeping the decryption algorithm secret.

Step 1: Alice chooses large prime numbers $p \neq q$ (this can be done very fast on a computer, since there are efficient algorithms for deciding whether a given integer is a prime or not).

Step 2: Alice chooses an integer $e > 1$ (which will be used for encryption), and then computes an integer $d > 1$ (which will be used for decryption) such that

$$de \equiv 1 \pmod{\operatorname{lcm}(p-1, q-1)}. \tag{4.4.4.1}$$

Step 3: Alice makes the pair (pq, e) (the **public key**) public, but keeps d (the **secret key**) secret.

Step 4: Messages to Alice will consist of pieces, where a piece is an element of $\mathbf{Z}/pq\mathbf{Z}$. Each piece will be encrypted by the map $x \pmod{pq} \mapsto y \equiv x^e \pmod{pq}$ (note that both e and pq are publicly known!), and then sent to Alice through a non-secure communication channel.

Step 5: Alice will decrypt the message received by $y \pmod{pq} \mapsto y^d \pmod{pq} \equiv x^{de} \equiv x \pmod{pq}$.

4.4.5 Remarks Alice can also sign (and therefore authenticate) messages using the secret key d (by sending to Bob a message consisting of a pair (or of several pairs) $(y_1 \pmod{pq}, y_2 \pmod{pq})$, where $y_2 \equiv y_1^d \pmod{pq}$. Bob will compute $y_2^e \equiv (y_1^d)^e \pmod{pq}$ and check that it is congruent to $y_1 \pmod{pq}$).

In practice, one does not use RSA to encrypt the text of the message, but to encrypt a key to a more conventional encryption algorithm, which is then used for encrypting the message.

The point of the RSA protocol is that, in general, the knowledge of pq is insufficient to determine easily p and q (factorisation of integers is hard, unless one has a quantum computer). If one knew p, q and e, then it would be easy to compute the secret key d from the congruence (4.4.4.1).

However, one needs to make certain precautions. For example, the primes p and q should be generated in a sufficiently random way, and they should not be too close to each other. In addition, $p-1$ (and also $q-1$) should be divisible by a large prime.

5 More advanced topics for enthusiasts

5.1 Congruences $f(x) \equiv 0 \pmod{n}$

5.1.1 Proposition (Congruence $x^2 \equiv 1 \pmod{p^k}$)**.** *Let p be a prime, let $k \geq 1$. The solutions of $x^2 \equiv 1 \pmod{p^k}$ are given by*

$$\begin{cases} x \equiv \pm 1 \pmod{p^k}, & p \neq 2 \\ x \equiv \pm 1 \pmod{2^{k-1}}, & p = 2,\ k > 1. \end{cases}$$

Proof. This is the content of Exercise 2.3.11. □

5.1.2 Corollary. *If p is a prime and if $p \nmid a$, then it is equivalent:*

$$a \equiv a^{-1} \pmod{p} \iff a \equiv \pm 1 \pmod{p}.$$

Proof. The condition on the left hand side is equivalent to $a \cdot a \equiv a \cdot a^{-1} \pmod{p}$, hence to $a^2 \equiv 1 \pmod{p}$. □

5.1.3 Application of the CRT (example) Let us solve the congruence

$$x^2 \equiv 1 \pmod{15}, \tag{5.1.3.1}$$

which is equivalent to the system

$$\left\{ \begin{matrix} x^2 \equiv 1 \pmod{3} \\ x^2 \equiv 1 \pmod{5} \end{matrix} \right\}$$

Each of the two solutions $x \equiv \pm 1 \pmod{3}$ of the first congruence can be combined with each of the two solutions $x \equiv \pm 1 \pmod{5}$ of the second congruence to obtain a solution of (5.1.3.1). Altogether, we obtain the following $2 \cdot 2 = 4$ solutions:

$x \pmod{3}$	$x \pmod{5}$	$x \pmod{15}$
1	1	1
-1	-1	-1
1	-1	4
-1	1	-4

To sum up,

$$x^2 \equiv 1 \pmod{15} \iff x \equiv \pm 1, \pm 4 \pmod{15}.$$

5.1.4 Application of the CRT (general principle) Assume that we are given a polynomial with integer coefficients $f(x) = a_0 + a_1 x + \cdots + a_d x^d$ ($d \geq 0$, $a_i \in \mathbf{Z}$). For each $n \geq 1$, denote by

$$N(f; n) := |\{x \pmod{n} \mid f(x) \equiv 0 \pmod{n}\}|$$

the number of solutions $\pmod{n}$ of the congruence $f(x) \equiv 0 \pmod{n}$.

If $m, n \geq 1$ and $\gcd(m, n) = 1$, the Chinese Remainder Theorem implies that the canonical bijective map

$$\mathbf{Z}/mn\mathbf{Z} \xrightarrow{\sim} \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}$$

gives a bijection between the subsets

$$\{x \pmod{mn} \mid f(x) \equiv 0 \pmod{mn}\} \subset \mathbf{Z}/mn\mathbf{Z}$$

and

$$\{x \pmod{m} \mid f(x) \equiv 0 \pmod{m}\} \times \{x \pmod{n} \mid f(x) \equiv 0 \pmod{n}\} \subset \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}.$$

In particular,

$$\gcd(m,n) = 1 \implies N(f;mn) = N(f;m)N(f;n). \tag{5.1.4.1}$$

In Section 5.1.3 we considered the case $m = 3$, $n = 5$, $f(x) = x^2 - 1$.

5.1.5 Exercise. (1) If p is a prime and $k \geq 1$, then the solutions of $x^2 \equiv x \pmod{p^k}$ are $x \equiv 0, 1 \pmod{p^k}$. [Hint: $x^2 - x = x(x-1)$.]
(2) The solutions of $x^2 \equiv x \pmod{10^4}$ are $x \equiv 0,\ 1,\ 625,\ 9376 \pmod{10^4}$.
(3) For each $k \geq 1$, the solutions of $x^2 \equiv x \pmod{10^k}$ are equal to $x \equiv 0, 1, e_k, 1 - e_k \pmod{10^k}$, where $e_k \equiv \cdots 0625 \pmod{10^k}$ and $1 - e_k \equiv \cdots 9376 \pmod{10^k}$.

5.1.6 10-adic numbers In Part (3) of 5.1.5 one can pass to the limit $k \longrightarrow +\infty$, obtaining a funny number $e = \cdots 0625$ whose decimal expansion has **infinitely many digits on the left** and which satisfies $e^2 = e$ (this implies that $1 - e = \cdots 9376$ also satisfies $(1-e)^2 = 1 - e$). Both e and $1 - e$ are examples of 10**-adic integers**. Another example is given by $a := \cdots 11111$, which satisfies $9a = \cdots 99999$, hence $9a + 1 = \cdots 00000 = 0$ and $a = -\frac{1}{9}$.

5.1.7 Congruences $x^2 \equiv a \pmod{n}$ (examples) Let us compute all possible values of squares $x^2 \pmod{p}$ for small primes $p \neq 2$ and $p \nmid x$.

$x \pmod 3$	± 1
$x^2 \pmod 3$	1

$x \pmod 5$	± 1	± 2
$x^2 \pmod 5$	1	$4 \equiv -1$

$x \pmod 7$	± 1	± 2	± 3
$x^2 \pmod 7$	1	$4 \equiv -3$	$9 \equiv 2$

$x \pmod{11}$	± 1	± 2	± 3	± 4	± 5
$x^2 \pmod{11}$	1	4	$9 \equiv -2$	$16 \equiv 5$	$25 \equiv 3$

$x \pmod{13}$	± 1	± 2	± 3	± 4	± 5	± 6
$x^2 \pmod{13}$	1	4	$9 \equiv -4$	$16 \equiv 3$	$25 \equiv -1$	$36 \equiv -3$

$x \pmod{17}$	± 1	± 2	± 3	± 4	± 5	± 6	± 7	± 8
$x^2 \pmod{17}$	1	4	$9 \equiv -8$	$16 \equiv -1$	$25 \equiv 8$	$36 \equiv 2$	$49 \equiv -2$	$64 \equiv -4$

$x \pmod{19}$	± 1	± 2	± 3	± 4	± 5	± 6	± 7	± 8	± 9
$x^2 \pmod{19}$	1	4	9	$16 \equiv -3$	$25 \equiv 6$	$36 \equiv -2$	$49 \equiv -8$	$64 \equiv 7$	$81 \equiv 5$

These tables contain the following information about solvability of $x^2 \equiv a \pmod{p}$ for $a = -1, \pm 3, 5$ (and variable $p \nmid a$):

	YES	NO
$x^2 \equiv -1 \pmod{p}$	$5, 13, 17$	$3, 7, 11, 19$
$x^2 \equiv -3 \pmod{p}$	$7, 13, 19$	$5, 11, 17$
$x^2 \equiv 3 \pmod{p}$	$11, 13$	$5, 7, 17, 19$
$x^2 \equiv 5 \pmod{p}$	$11, 19$	$3, 7, 13, 17$

This would seem to suggest that

- $x^2 \equiv -1 \pmod p$ has a solution $\stackrel{?}{\iff} p \equiv 1 \pmod 4$
- $x^2 \equiv -3 \pmod p$ has a solution $\stackrel{?}{\iff} p \equiv 1 \pmod 6$
- $x^2 \equiv 3 \pmod p$ has a solution $\stackrel{?}{\iff} p \equiv \pm 1 \pmod{12}$
- $x^2 \equiv 5 \pmod p$ has a solution $\stackrel{?}{\iff} p \equiv \pm 1 \pmod 5$

All of this is true, as a consequence of the **quadratic reciprocity law** (first proved in full generality by Gauss), according to which

- $x^2 \equiv -1 \pmod p$ has a solution $\iff p \equiv 1 \pmod 4$
- $x^2 \equiv 2 \pmod p$ has a solution $\iff p \equiv \pm 1 \pmod 8$
- $x^2 \equiv (-1)^{(q-1)/2} q \pmod p$ has a solution $\iff x^2 \equiv p \pmod q$ has a solution

(above, $p, q > 2$ are distinct odd primes).

5.1.8 Exercise. (1) Find all solutions of $x^2 \equiv -1 \pmod{35}$.
(2) Find all solutions of $x^2 \equiv -1 \pmod{85}$.

5.1.9 Theorem. *Let $p \neq 2$ be a prime. The congruence $x^2 \equiv -1 \pmod p$ has a solution $\iff p \equiv 1 \pmod 4$.*

Proof. Let us begin with the easier implication "$\Longrightarrow$": if $x \in \mathbf{Z}$ satisfies $x^2 \equiv -1 \pmod p$, then

$$\underbrace{(-1)^{\frac{p-1}{2}}}_{=\pm 1} \equiv (x^2)^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \pmod p.$$

As $-1 \not\equiv 1 \pmod p$, we must have $(-1)^{\frac{p-1}{2}} = 1$, hence $\frac{p-1}{2} = 2k$ and $p = 4k+1$ (for some $k \in \mathbf{Z}$).
In order to prove the more difficult implication "$\Longleftarrow$" we use Wilson's theorem 5.1.10 below. If $p = 4k+1$, then

$$-1 \equiv (p-1)! \equiv (4k)! \equiv 1 \cdot 2 \cdots (2k) \cdot \underbrace{(2k+1)}_{\equiv -2k} \cdots \underbrace{(4k)}_{\equiv -1} \equiv (2k)!\,(-1)^{2k} \cdot (2k)! = ((2k)!)^2 \pmod p.$$

□

5.1.10 Theorem (Wilson's theorem)**.** *If p is a prime, then $(p-1)! \equiv -1 \pmod p$.*

Proof. We can assume that $p \neq 2$. By definition, $(p-1)! \equiv 1 \cdot 2 \cdots (p-1) \pmod p$ is congruent to the product of all invertible residue classes $\pmod p$. The idea is to put together each residue class with its inverse, provided they are distinct. For example, for $p = 7$, we can rewrite the product defining 6! as follows:

$$6! = \underbrace{(2 \cdot 4)}_{\equiv 1} \cdot \underbrace{(3 \cdot 5)}_{\equiv 1} \cdot \underbrace{(1 \cdot 6)}_{\equiv -1} \equiv -1 \pmod 7.$$

In general, Corollary 5.1.2 tells us that invertible residue classes $\pmod p$ different from $\pm 1 \pmod p$ come in pairs $a \pmod p$, $a^{-1} \pmod p$. Consequently,

$$(p-1)! \equiv (a \cdot a^{-1}) \cdot (b \cdot b^{-1}) \cdots (t \cdot t^{-1}) \cdot (1 \cdot (-1)) \equiv 1 \cdot (-1) \equiv -1 \pmod{p}.$$

□

5.1.11 Exercise. Let $p \neq 2, 3$ be a prime. If $x \in \mathbf{Z}$ satisfies $x^2 \equiv -3 \pmod{p}$, show that:
(1) The residue class $y := 2^{-1} \cdot (-1 + x) \pmod{p}$ behaves like $\frac{1}{2}(-1 + i\sqrt{3}) = e^{2\pi i/3} \in \mathbf{C}$: $y^2 + y + 1 \equiv 0 \pmod{p}$, $y^3 \equiv 1 \not\equiv y \pmod{p}$.
(2) Write $p = 3k + a$, where $k \in \mathbf{Z}$ and $a \in \{1, 2\}$. Deduce from $y^{p-1} \equiv 1 \pmod{p}$ that $a = 1$, hence $p \equiv 1 \pmod{3}$.

5.1.12 Exercise. If $x, y \in \mathbf{Z}$ and if $p \equiv 3 \pmod{4}$ is a prime such that $p \mid (x^2 + y^2)$, then $p \mid x$ and $p \mid y$.

5.1.13 Polynomial congruences with many solutions A polynomial of degree $d \geq 1$ with complex coefficients has at most d complex roots. What happens for polynomial congruences? We know that the quadratic congruence

$$x^2 - 1 \equiv 0 \pmod{8}$$

has 4 solutions $x \equiv \pm 1, \pm 5 \pmod{8}$. Similarly, if $p_1, \ldots, p_r \neq 2$ are distinct odd primes, then the quadratic congruence

$$x^2 - 1 \equiv 0 \pmod{p_1 \cdots p_r}$$

has $2^r > 2$ solutions, by (5.1.4.1).

The following result, which is a special case of a general abstract result proved in Theorem 9.2.7, shows that congruences modulo primes behave in a more reasonable way.

5.1.14 Theorem. *Let p be a prime, let $a_0, \ldots, a_d \in \mathbf{Z}$ and $p \nmid a_d$ $(d \geq 0)$. The congruence*

$$f(x) = a_0 + a_1 x + \cdots + a_d x^d \equiv 0 \pmod{p}$$

has at most d solutions $\pmod{p}$.

Proof. Induction on d. The case $d = 0$ is straightforward. Assume that $d > 0$ and that the result holds for polynomials of degree $\deg < d$. Assume that $a \in \mathbf{Z}$ satisfies $f(a) \equiv 0 \pmod{p}$. The formulas $x^k - a^k = (x - a)(x^{k-1} + ax^{k-2} + \cdots + a^{k-1})$ imply that there is a polynomial identity

$$f(x) - f(a) = (x - a)g(x), \qquad g(x) = b_0 + b_1 x + \cdots + b_{d-1} x^{d-1}, \qquad b_i \in \mathbf{Z}, \qquad p \nmid b_{d-1} = a_d.$$

If $b \not\equiv a \pmod{p}$ is a solution of $f(b) \equiv 0 \pmod{p}$, then

$$(b - a)g(b) \equiv f(b) - f(a) \equiv 0 \pmod{p}.$$

As $b - a \not\equiv 0 \pmod{p}$ is invertible $\pmod{p}$, it follows that $g(b) \equiv 0 \pmod{p}$. By induction, this congruence has at most $d - 1$ possible solutions $b \pmod{p}$. Together with $a \pmod{p}$, this gives at most d solutions of $f(x) \equiv 0 \pmod{p}$. □

5.1.15 Corollary. *Let p be a prime. If $a_0, \ldots, a_d \in \mathbf{Z}$ $(d \geq 0)$ and if the congruence $f(x) = a_0 + a_1 x + \cdots + a_d x^d \equiv 0 \pmod{p}$ has at least $d + 1$ solutions $\pmod{p}$, then $p \mid a_i$ for all i, hence $f(a) \equiv 0 \pmod{p}$ holds for all $a \in \mathbf{Z}$.*

5.1.16 Proposition (Another proof of Wilson's theorem)**.** *If p is a prime, then $(p - 1)! \equiv -1 \pmod{p}$.*

Proof. The polynomial

$$f(x) := (x^{p-1} - 1) - \prod_{j=1}^{p-1}(x - j)$$

has coefficients in $\mathbf{Z}$, its degree is $\deg(f) < p - 1$, and the congruence $f(x) \equiv 0 \pmod p$ is satisfied for $x \equiv 1, \dots, p-1 \pmod p$. By Corollary 5.1.15, all coefficients of $f(x)$ are divisible by p, which implies that

$$f(0) \equiv 0 \pmod p, \qquad f(0) = -1 + (-1)^p(p-1)!$$

□

5.2 Primes in arithmetic progressions

5.2.1 Primes modulo 4 and 6 Recall from Exercise 1.1.8 that a prime $p \neq 2$ (resp. $p \neq 2, 3$) satisfies $p \equiv \pm 1 \pmod 4$ (resp. $p \equiv \pm 1 \pmod 6$).

5.2.2 Proposition. *There are infinitely many primes* $p \equiv -1 \pmod 4$.

Proof. We need to show the following: if $p_1, \dots, p_r \equiv -1 \pmod 4$ are primes ($r \geq 0$), then there exists a prime $p \equiv -1 \pmod 4$ such that $p \neq p_1, \dots, p_r$. Let $N := 4p_1 \cdots p_r - 1 \geq 4 \cdot 1 - 1 = 3$. Note that $N \equiv -1 \pmod 4$. Write $N = q_1 \cdots q_s$, where the q_j are primes (not necessarily distinct). As $2 \nmid N$, $q_j \neq 2$, hence $q_j \equiv \pm 1 \pmod 4$. If $q_j \equiv 1 \pmod 4$ for all $j = 1, \dots, r$, then $N \equiv 1 \pmod 4$, which is not true. Therefore there exists $p = q_j$ such that $p \not\equiv 1 \pmod 4$, which implies that $p \equiv -1 \pmod 4$. Note that $p \mid N$. If $p = p_i$ for some i, then $p \mid 4p_1 \cdots p_r = N + 1$, hence $p \mid (N+1) - N$, which is impossible. Therefore $p \neq p_1, \dots, p_r$. □

5.2.3 Exercise. There are infinitely many primes $p \equiv -1 \pmod 6$.

5.2.4 Proposition. *There are infinitely many primes* $p \equiv 1 \pmod 4$.

Proof. We need to show the following: if $p_1, \dots, p_r \equiv 1 \pmod 4$ are primes ($r \geq 0$), then there exists a prime $p \equiv 1 \pmod 4$ such that $p \neq p_1, \dots, p_r$. Let $N := (2p_1 \cdots p_r)^2 + 1 \geq 2^2 + 1 = 5$, let $p \mid N$ be any prime dividing N. Then $p \neq 2$ and $x := 2p_1 \cdots p_r \in \mathbf{Z}$ satisfies $x^2 \equiv -1 \pmod p$. The "easier part" of Theorem 5.1.9 implies that $p \equiv 1 \pmod 4$. If $p = p_i$ for some i, then $p \mid (2p_1 \cdots p_r)^2 = N - 1$, hence $p \mid N - (N-1)$, which is impossible. Therefore $p \neq p_1, \dots, p_r$. □

5.2.5 Exercise. There are infinitely many primes $p \equiv 1 \pmod 6$.
[Hint: modify the method of proof of Proposition 5.2.4 by appealing to Exercise 5.1.11 instead of Theorem 5.1.9.

5.2.6 Exercise. There are infinitely many primes $p \equiv 5 \pmod{12}$.
[Hint: combine the methods of Proposition 5.2.4 with those of Exercise 5.2.3.]

5.2.7 Exercise. There are infinitely many primes $p \equiv 7 \pmod{12}$.
[Hint: combine the methods of Proposition 5.2.2 with those of Exercise 5.2.5.]

5.2.8 Exercise. What would one need to know in order to prove (by the same method) that there are infinitely many primes $p \equiv 11 \pmod{12}$?

5.2.9 Exercise. (1) $(x^{12} - 1)/((x^6 - 1)(x^2 + 1)) = x^4 - x^2 + 1 = (x^2 - 1)^2 + x^2 = (x^2 - \frac{1}{2})^2 + \frac{3}{4}$.
(2) If p is a prime and if $x^4 - x^2 + 1 \equiv 0 \pmod p$ has a solution, then $p \equiv 1 \pmod{12}$.
(3) There are infinitely many primes $p \equiv 1 \pmod{12}$.

5.2.10 Exercise. (1) If $p \neq 2$ is a prime and if $x^4 \equiv -1 \pmod p$ has a solution, then $p \equiv 1 \pmod 8$.
(2) There are infinitely many primes $p \equiv 1 \pmod 8$.
(3) If $p \neq 2$ is a prime and if $x^{2^k} \equiv -1 \pmod p$ has a solution, then $p \equiv 1 \pmod{2^{k+1}}$.
(4) For each $n \geq 2$ there are infinitely many primes $p \equiv 1 \pmod{2^n}$.

5.2.11 More general results Elementary methods of this kind can be used to show that there are infinitely many primes of the form $p \equiv a \pmod n$, provided that $a^2 \equiv 1 \pmod n$. The following general result is due to Dirichlet.

5.2.12 Theorem (Dirichlet's theorem on primes in arithmetic progressions)**.** *If* $\gcd(a, n) = 1$*, then there are infinitely many primes of the form* $p \equiv a \pmod n$*. More precisely,*

$$\sum_{\substack{p \in \mathcal{P} \\ p \equiv a \pmod n}} \frac{1}{p} = +\infty. \tag{5.2.12.1}$$

5.2.13 Dirichlet's method A first result of this kind was proved by Euler, who showed that

$$\sum_{p \in \mathcal{P}} \frac{1}{p} = +\infty.$$

One can reformulate Euler's method as follows: one considers the identity

$$\prod_{p \in \mathcal{P}} \frac{1}{1 - p^{-s}} = \sum_{n=1}^{\infty} n^{-s} \qquad (s > 1)$$

(which is an analytic reformulation of the uniqueness of factorisation in $\mathbf{Z}$), then applies the expansion

$$-\log(1 - T) = \sum_{k=1}^{\infty} \frac{T^k}{k}$$

to each term of the product, and finally lets $s \longrightarrow 1+$.

Dirichlet's result (5.2.12.1) in the simplest case $n = 4$, $a = \pm 1$, can be proved in the same way, by considering also the product

$$\prod_{\substack{p \in \mathcal{P} \\ p \equiv 1 \pmod 4}} \frac{1}{1 - p^{-s}} \prod_{\substack{p \in \mathcal{P} \\ p \equiv -1 \pmod 4}} \frac{1}{1 + p^{-s}} = 1 - 3^{-s} + 5^{-s} - 7^{-s} + 9^{-s} \cdots \qquad (s > 1).$$

5.3 Pseudoprimes, Carmichael numbers

5.3.1 Question According to Fermat's little theorem, $a^p \equiv a \pmod p$ holds for all $a \in \mathbf{Z}$ if p is a prime. Does this property characterise primes?

5.3.2 Definition. An integer $n > 1$ is a **pseudoprime in base** $a \in \mathbf{Z}$ if n is not a prime number and $a^n \equiv a \pmod n$.

5.3.3 Example: $2^{341} \equiv 2 \pmod{341}$ If $a = 2$ and $n = 341 = 11 \cdot 31$, then $2^5 = 32$ and

$$\left.\begin{aligned} 2^5 \equiv 1 \pmod{31} &\Longrightarrow 2^{10} \equiv 1 \pmod{31} \\ 2^5 \equiv -1 \pmod{11} &\Longrightarrow 2^{10} \equiv 1 \pmod{11} \end{aligned}\right\} \Longrightarrow 2^{10} \equiv 1 \pmod{11 \cdot 31} \Longrightarrow$$
$$\Longrightarrow 2^{341-1} \equiv (2^{10})^{34} \equiv 1 \pmod{11 \cdot 31} \Longrightarrow 2^{341} \equiv 2 \pmod{11 \cdot 31}.$$

5.3.4 Definition. An integer $n > 1$ is a **Carmichael number** if it is a pseudoprime in every base, i.e., if n is not a prime number and $a^n \equiv a \pmod{n}$ holds for all $a \in \mathbf{Z}$.

5.3.5 Proposition. *An integer $n > 1$ is a Carmichael number $\iff$ $n = p_1 \cdots p_r$ is a product of $r \geq 2$ distinct prime numbers such that $\forall i = 1, \ldots, r \quad (p_i - 1) \mid (n - 1)$.*

Proof. The implication '$\Longleftarrow$' is a special case of Proposition 4.2.11(2). The implication '$\Longrightarrow$': Proposition 4.2.11(1) implies that a Carmichael number n must be square-free, hence $n = p_1 \cdots p_r$ ($r \geq 2$, since $n > 1$ is not a prime). In order to show that $p_i - 1$ divides $n - 1$ we need to use the fact that there exists $a_i \in \mathbf{Z}$ such that $a_i \pmod{p_i}$ is a generator of $(\mathbf{Z}/p_i\mathbf{Z})^*$ (see Theorem 5.5.2). This integer satisfies $p_i \nmid a_i$ and $a_i^n \equiv a_i \pmod{p_i}$, hence $a_i^{n-1} \equiv 1 \pmod{p_i}$, which implies that $n - 1$ is divisible by the order of $a_i \pmod{p_i} \in (\mathbf{Z}/p_i\mathbf{Z})^*$, which is equal to $p_i - 1$. □

5.3.6 Example: $n = 561$ The smallest Carmichael number is $n = 561 = 3 \cdot 11 \cdot 17$, since $3 - 1 = 2$, $11 - 1 = 2 \cdot 5$, $17 - 1 = 2^4$ and $561 - 1 = 2^4 \cdot 5 \cdot 7$.

5.3.7 Remark It is known that there exist infinitely many Carmichael numbers.

5.3.8 Exercise. Show that, if $n = p_1 \cdots p_r$ is a Carmichael number, then $r \geq 3$.

5.4 Möbius inversion formula

5.4.1 Fractions $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}$ We know that every rational number $\frac{a}{n}$ ($a, n \in \mathbf{Z}$, $n \geq 1$) can be simplified by dividing both the numerator and the denominator by their greatest common divisor $m = \gcd(a, n)$:

$$a = ma', \qquad n = mn', \qquad \frac{a}{n} = \frac{a'}{n'}, \qquad \gcd(a', n') = 1. \tag{5.4.1.1}$$

What happens if we apply this procedure to all rational numbers $\frac{a}{n}$ ($a \in \mathbf{Z}$) for fixed $n \geq 1$? It is enough to consider the numerators in the range $1 \leq a \leq n$, since $\frac{a+n}{n} = \frac{a}{n} + 1$ and $\gcd(a + n, n) = \gcd(a, n)$.

Example: For $n = 6$,

$$\frac{1}{6} = \frac{1}{6}, \qquad \frac{2}{6} = \frac{1}{3}, \qquad \frac{3}{6} = \frac{1}{2}, \qquad \frac{4}{6} = \frac{2}{3}, \qquad \frac{5}{6} = \frac{5}{6}, \qquad \frac{6}{6} = \frac{1}{1},$$
$$\left\{\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}, \frac{5}{6}, \frac{6}{6}\right\} = \left\{\frac{1}{1}\right\} \cup \left\{\frac{1}{2}\right\} \cup \left\{\frac{1}{3}, \frac{2}{3}\right\} \cup \left\{\frac{1}{6}, \frac{5}{6}\right\}.$$

In general: For arbitrary $n \geq 1$ we obtain, after simplifying each fraction $\frac{a}{n}$ ($1 \leq a \leq n$) as in (5.4.1.1), fractions of the form $\frac{b}{d}$, where $d \mid n$, $1 \leq b \leq d$ and $\gcd(b, d) = 1$ (and each fraction of this form arises from some $\frac{a}{n}$). Therefore

$$\left\{\frac{a}{n} \,\middle|\, 1 \leq a \leq n\right\} = \bigcup_{d \mid n} \left\{\frac{b}{d} \,\middle|\, 1 \leq b \leq d,\ \gcd(b, d) = 1\right\} \tag{5.4.1.2}$$

(a disjoint union). If we compare the number of elements on each side, we obtain

$$\forall n \geq 1 \qquad n = \sum_{d \mid n} \varphi(d). \tag{5.4.1.3}$$

5.4.2 Möbius inversion formula The equalities (5.4.1.3) ($n \geq 1$) give a system of linear relations for the values $\varphi(d)$ ($d \geq 1$).

More generally, for any function $f : \mathbf{N}_+ \longrightarrow \mathbf{C}$ we can consider the function $g : \mathbf{N}_+ \longrightarrow \mathbf{C}$ defined by

$$g(n) = \sum_{d|n} f(d). \tag{5.4.2.1}$$

Explicitly,

$$g(1) = f(1), \quad g(2) = f(2) + f(1), \quad g(3) = f(3) + f(1), \quad g(4) = f(4) + f(2) + f(1), \quad \cdots$$

These relations can be inverted and the values of f can be expressed as linear combinations of the values of g:

$$f(1) = g(1), \quad f(2) = g(2) - g(1), \quad f(3) = g(3) - g(1), \quad f(4) = g(4) - g(2), \quad \cdots$$

In general, $f(n)$ is given by the **Möbius inversion formula**

$$f(n) = \sum_{n=dm} \mu(d)\, g(m) = \sum_{d|n} \mu(d)\, g\big(\frac{n}{d}\big) = \sum_{d|n} \mu\big(\frac{n}{d}\big)\, g(d), \tag{5.4.2.2}$$

where $\mu : \mathbf{N}_+ \longrightarrow \{0, \pm 1\}$ is the **Möbius function**, defined by

$$\mu(n) := \begin{cases} 1, & n = 1 \\ (-1)^r, & n = p_1 \cdots p_r,\ p_i \text{ distinct primes} \\ 0, & \exists p \in \mathcal{P} \quad p^2 \mid n. \end{cases} \tag{5.4.2.3}$$

Indeed, the function μ satisfies

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & n = 1 \\ 0, & n > 1, \end{cases} \tag{5.4.2.4}$$

which implies that if we begin with a function $g : \mathbf{N}_+ \longrightarrow \mathbf{C}$ and define $f : \mathbf{N}_+ \longrightarrow \mathbf{C}$ by the formula (5.4.2.2), then

$$\sum_{d|n} f(d) = \sum_{d|n} \sum_{d=em} \mu(e)\, g(m) = \sum_{m|n} g(m) \sum_{e|\frac{n}{m}} \mu(e) = \sum_{m|n} g(m) \left\{ \begin{matrix} 1, & \frac{n}{m} = 1 \\ 0, & \frac{n}{m} > 1 \end{matrix} \right\} = g(n).$$

5.4.3 Exercise. Prove the formula (5.4.2.4). [Hint: write $n = p_1^{k_1} \cdots p_r^{k_r}$.]

5.4.4 Function $\varphi(n)$ Applying the Möbius inversion formula to (5.4.1.3), we obtain

$$\varphi(n) = \sum_{d|n} \mu(d)\, \frac{n}{d}\,.$$

If $n = p_1^{k_1} \cdots p_r^{k_r}$, then the only divisors $d \mid n$ with $\mu(d) \neq 0$ are $d = p_{i_1} \cdots p_{i_s}$, where $1 \leq i_1 < \cdots < i_s \leq r$, which gives

$$\frac{\varphi(n)}{n} = \sum_{d|n} \frac{\mu(d)}{d} = \sum_{s=0}^{r} (-1)^s \sum_{1 \leq i_1 < \cdots < i_s \leq r} \frac{1}{p_{i_1} \cdots p_{i_s}} = \left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right) = \prod_{p|n} \left(1 - \frac{1}{p}\right). \tag{5.4.4.1}$$

5.4.5 Exercise. For a given integer $M \geq 1$, show that the function $\varphi_M : \mathbf{N}_+ \longrightarrow \mathbf{C}$ defined by

$$\varphi_M(n) := |\{(a_1, \ldots, a_M) \mid 1 \leq a_i \leq n,\ \gcd(a_1, \ldots, a_M, n) = 1\}|$$

satisfies

$$\sum_{d|n} \varphi_M(n) = n^M.$$

Prove a formula for $\varphi_M(n)$ similar to (5.4.4.1).

5.5 Structure of $(\mathbf{Z}/p^k\mathbf{Z})^*$

5.5.1 $p \neq 2$ We are going to show that in this case $(\mathbf{Z}/p^k\mathbf{Z})^*$ always has a generator. The main point is to prove this result for $k = 1$.

5.5.2 Theorem (Gauss)**.** *For each prime p there exists a generator of $(\mathbf{Z}/p\mathbf{Z})^*$ (i.e., an invertible residue class $a \pmod p$ such that $(\mathbf{Z}/p\mathbf{Z})^* = \{a, a^2, \ldots, a^{p-1} \pmod p\}$).*

Proof. We know that $a \pmod p \in (\mathbf{Z}/p\mathbf{Z})^*$ is a generator of $(\mathbf{Z}/p\mathbf{Z})^*$ if and only if its order is equal to $p - 1$. Let us denote by $\psi(d)$ (for any $d \geq 1$) the number of elements of $(\mathbf{Z}/p\mathbf{Z})^*$ whose order is equal to d. We know that $\psi(d) = 0$ if $d \nmid (p-1)$, which implies that

$$\sum_{d|(p-1)} \psi(d) = |(\mathbf{Z}/p\mathbf{Z})^*| = p - 1 = \sum_{d|(p-1)} \varphi(d). \tag{5.5.2.1}$$

The key point is to prove the following implication:

$$\psi(d) \neq 0 \implies \psi(d) = \varphi(d). \tag{5.5.2.2}$$

When combined with (5.5.2.1), it implies that

$$\forall d \mid (p-1) \quad \psi(d) = \varphi(d).$$

In particular, the number of generators of $(\mathbf{Z}/p\mathbf{Z})^*$ is equal to $\psi(p-1) = \varphi(p-1) > 0$.

It remains to prove (5.5.2.2). Assume that $d \mid (p-1)$ and that there exists $a \pmod p$ of order d. There is an obvious inclusion

$$\{a, a^2, \ldots, a^d \pmod p\} \subseteq \{x \pmod p \mid x^d - 1 \equiv 0 \pmod p\}.$$

The set on the left has d elements, whereas the set on the right has at most $\deg(x^d - 1) = d$ elements, by Theorem 5.1.14 (this is the point where we are using the fact that p is a prime). Therefore the two sets are equal:

$$\{a, a^2, \ldots, a^d \pmod p\} = \{x \pmod p \mid x^d - 1 \equiv 0 \pmod p\}.$$

The rest is easy: any element $x \pmod p$ of order d must satisfy $x^d - 1 \equiv 0 \pmod p$, hence $x \equiv a^k \pmod p$ for some $k = 1, \ldots, d$. According to Proposition 4.3.13, the order of $a^k \pmod p$ is equal to d if and only if $\gcd(k, p) = 1$, which happens for $\varphi(d)$ values of k. The implication (5.5.2.2) is proved. □

5.5.3 Proposition (Improvement of congruences by $x \mapsto x^p$ is uniform)**.** *Assume that p is a prime, $k \geq 1$, $p^k > 2$ and that $a, b \in \mathbf{Z}$ satisfy $p \nmid a$, $a \equiv b \pmod{p^k}$ and $a \not\equiv b \pmod{p^{k+1}}$. Then $a^p \equiv b^p \pmod{p^{k+1}}$ and $a^p \not\equiv b^p \pmod{p^{k+2}}$.*
[Note that the assumption $p^k > 2$ is necessary: $1 \equiv 3 \pmod 2$ and $1 \not\equiv 3 \pmod{2^2}$, but $1^2 \equiv 3^2 \pmod{2^3}$.]

Proof. There exists $c \in \mathbf{Z}$ such that $a = b + p^k c$ and $p \nmid c$, hence

$$a^p - b^p = (b + p^k c)^p - b^p = \binom{p}{1} b^{p-1}(p^k c) + \binom{p}{2} b^{p-2}(p^k c)^2 + \cdots + \binom{p}{p-1} b(p^k c)^{p-1} + p^{pk} c^p.$$

As observed in the proof of Proposition 4.1.5, each term on the right hand side is divisible by p^{k+1}. The first term $\binom{p}{1} b^{p-1}(p^k c) = p^{k+1} b^{p-1} c$ is not divisible by p^{k+2}, since $p \nmid b$ and $p \nmid c$. The terms $\binom{p}{j} b^{p-j}(p^k c)^j$ for $1 < j < p$ are divisible by $p \cdot (p^k)^2 = p^{2k+1}$, hence by p^{k+2}. The last term $p^{pk} c^p$ is also divisible by p^{k+2}, since the assumption $p^k > 2$ implies that $pk - (k+2) = (p-1)k - 2 \geq 0$. Therefore $a^p - b^p \equiv p^{k+1} b^{p-1} c \not\equiv 0 \pmod{p^{k+2}}$. □

5.5.4 Theorem. *Let $p \neq 2$ be a prime, let $k > 1$.*
(1) If $a \in \mathbf{Z}$, $p \nmid a$, $a^{p-1} \not\equiv 1 \pmod{p^2}$ and if $a \pmod p$ is a generator of $(\mathbf{Z}/p\mathbf{Z})^*$, then $a \pmod{p^k}$ is a generator of $(\mathbf{Z}/p^k\mathbf{Z})^*$ (i.e., $(\mathbf{Z}/p^k\mathbf{Z})^* = \{a, a^2, \ldots, a^{(p-1)p^{k-1}} \pmod{p^k}\}$).
(2) A generator of $(\mathbf{Z}/p^k\mathbf{Z})^*$ always exists.
(3) A generator of $(\mathbf{Z}/2p^k\mathbf{Z})^*$ always exists.

Proof. (1) Let d be the order of $a \pmod{p^k}$. We know that $d \mid (p-1)p^{k-1}$. On the other hand $a^d \equiv 1 \pmod{p^k}$ implies that $a^d \equiv 1 \pmod p$, hence $(p-1) \mid d$, since $a \pmod p$ is a generator of $(\mathbf{Z}/p\mathbf{Z})^*$. This means that $d = (p-1)p^l$ for some $0 \leq l \leq k-1$. We must show that $l = k-1$, but this follows from the assumptions $p > 2$ and $a^{p-1} \not\equiv 1 \pmod{p^2}$ (and the fact that $a^{p-1} \equiv 1 \pmod p$), by applying successively Proposition 5.5.3: $a^{(p-1)p} \not\equiv 1 \pmod{p^3}$, …, $a^{(p-1)p^{k-2}} \not\equiv 1 \pmod{p^k}$.

(2) According to Theorem 5.5.2, there exists $a \in \mathbf{Z}$ such that $a \pmod p$ is a generator of $(\mathbf{Z}/p\mathbf{Z})^*$. If $a^{p-1} \not\equiv 1 \pmod{p^2}$, then $a \pmod{p^k}$ is a generator of $(\mathbf{Z}/p^k\mathbf{Z})^*$, by (1). If $a^{p-1} \equiv 1 \pmod{p^2}$, then $b := a(1+p) \equiv a \pmod p$ and $b^{p-1} \equiv a^{p-1}(1 + \binom{p-1}{1} p) \equiv 1 - p \not\equiv 1 \pmod{p^2}$, hence $b \pmod{p^k}$ is a generator of $(\mathbf{Z}/p^k\mathbf{Z})^*$.

(3) Let $a \pmod{p^k}$ be a generator of $(\mathbf{Z}/p^k\mathbf{Z})^*$. After possibly replacing a by $a + p^k$ we can assume that $2 \nmid a$; then $a \pmod{2p^k}$ will be a generator of $(\mathbf{Z}/2p^k\mathbf{Z})^* = (\mathbf{Z}/2\mathbf{Z})^* \times (\mathbf{Z}/p^k\mathbf{Z})^* = \{1\} \times (\mathbf{Z}/p^k\mathbf{Z})^*$. □

5.5.5 $p = 2$ It turns out that in this case $(\mathbf{Z}/2^k\mathbf{Z})^*$ (for $k \geq 3$) does not have a generator, but it does have a "generator up to a sign".

5.5.6 Theorem. *Assume that $k > 1$. For every $a \in \mathbf{Z}$ satisfying $a \equiv 1 \pmod{2^2}$ and $a \not\equiv 1 \pmod{2^3}$ (for example, for $a = 5$), $(\mathbf{Z}/2^k\mathbf{Z})^* = \{\pm a, \pm a^2, \ldots, \pm a^{2^{k-2}} \pmod{2^k}\}$.*

Proof. There is a decomposition

$$(\mathbf{Z}/2^k\mathbf{Z})^* = X_+ \cup X_-, \qquad X_\pm = \{x \pmod{2^k} \mid x \equiv \pm 1 \pmod{2^2}\}, \qquad |X_\pm| = \tfrac{1}{2}\varphi(2^k) = 2^{k-2}.$$

The assumption $a \in X_+$ implies that $a^k \in X_+$, for all $k \in \mathbf{Z}$. As in the proof of Theorem 5.5.4(2), the assumptions $a \equiv 1 \pmod{2^2}$ and $a \not\equiv 1 \pmod{2^3}$ can be used as an input into Proposition 5.5.3, which will then give $a^2 \equiv 1 \pmod{2^3}$ and $a^2 \not\equiv 1 \pmod{2^4}$, …, $a^{2^{k-3}} \equiv 1 \pmod{2^{k-1}}$ and $a^{2^{k-3}} \not\equiv 1 \pmod{2^k}$, and finally $a^{2^{k-2}} \equiv 1 \pmod{2^k}$. This means that the order of $a \pmod{2^k}$ is equal to 2^{k-2}, hence both sets

$$\{a, a^2, \ldots, a^{2^{k-2}} \pmod{2^k}\} \subseteq X_+$$

have the same cardinality $2^{k-2} = |X_+|$. Therefore they are equal, which implies that

$$X_+ = \{a, a^2, \ldots, a^{2^{k-2}} \pmod{2^k}\}, \qquad X_- = \{-a, -a^2, \ldots, -a^{2^{k-2}} \pmod{2^k}\}.$$

□

6 Algebra – motivation

6.1 A preview

6.1.1 Abstract theory The main goal of the abstract algebraic part of this course is to introduce basic objects of abstract algebra: groups (with special emphasis on abelian groups), rings (with special emphasis on commutative rings) and fields, and describe simple constructions involving them. We are going to illustrate the theory by a number of concrete examples, many of them related to number-theoretical constructions from the first part of the course.

We have already seen quite a few examples of groups, rings and fields:

- in an additive subgroup $X \subset \mathbf{C}$ one can perform operations "+" and "−" satisfying the usual rules (X is an abelian group with respect to addition);
- in a subring $A \subset \mathbf{C}$ one can perform operations "+", "−" and "$\cdot$" satisfying the usual rules (A is a commutative ring);
- in a subfield of $\mathbf{C}$ one can also perform division by non-zero elements;
- $\mathbf{Z}/n\mathbf{Z}$ is a (commutative) ring (it has operations "+", "−" and "$\cdot$"), but it is not a subring of $\mathbf{C}$;
- if p is a prime, then $\mathbf{Z}/p\mathbf{Z}$ is a field (it is a commutative ring in which one can make division by non-zero elements);
- the set $M_2(\mathbf{R})$ of 2×2 matrices with real coefficients is a non-commutative ring (it has operations "+", "−" and "$\cdot$", but matrix multiplication is not commutative: $M \cdot N \neq N \cdot M$, in general).

Many groups occur as transformation groups in geometry, but such groups will not be treated in detail in this course.

6.1.2 Important example: polynomial rings In the concrete algebraic part of the course we are going to study polynomials in one variable from an algebraic point of view. Perhaps the most important point of the whole course is the fact that the ring of integers $\mathbf{Z} = \{0, \pm 1, \pm 2, \ldots\}$ behaves, from a purely algebraic point of view, in the same way as the ring of polynomials $K[X] = \{f(X) = a_0 + a_1 X + \cdots + a_n X^n \mid a_j \in K,\ n \geq 0\}$ in one variable with coefficients in a field K (we can take, for example, $K = \mathbf{Q}, \mathbf{R}$ or $\mathbf{C}$).

In both cases there is a notion of divisibility, and it has the same properties: there is a division with remainder, Euclid's algorithm, Bézout's theorem, the gcd, Euclid's Lemma and uniqueness of factorisation.

As a result, theory of congruences (including the Chinese Remainder Theorem) works in the same way in both situations. This is not just an abstract theory; as we shall see, congruences between polynomials are related to various concrete objects appearing in algebra, analysis and linear algebra.

Arithmetic	Algebra
$\mathbf{Z}$	$K[X]$
$n \in \mathbf{Z}$	$f = f(X) \in K[X]$
$a \equiv b \pmod{n}$	$u(X) \equiv v(X) \pmod{f}$
$\mathbf{Z}/n\mathbf{Z}$	$K[X]fK[X] = K[X]/(f)$
$\mathbf{Z}^* = \{\pm 1\}$	$K[X]^* = K^* = K \setminus \{0\}$
prime numbers p	irreducible monic polynomials $g(X)$
$n = \pm \prod p^{v_p(n)}$	$f(X) = c \prod g(X)^{v_g(f)}$

6.2 Polynomials

6.2.1 Division with remainder **Example 1:** $\mathbf{R}[X]/(X) = \mathbf{R}$. Indeed, any polynomial $g(X) = b_0 + b_1X + \cdots + b_nX^n$ ($b_j \in \mathbf{R}$) can be written in a unique way as

$$g(X) = Xh(X) + r, \qquad h(X) \in \mathbf{R}[X], \qquad r \in \mathbf{R},$$
$$h(X) = b_1 + b_2X + \cdots + b_nX^{n-1}, \qquad r = b_0 = g(0).$$

This means that the set of residue classes modulo X is equal to

$$\mathbf{R}[X]/(X) = \{r \pmod X \mid r \in \mathbf{R}\}.$$

More precisely, the maps

$$\mathbf{R} \longrightarrow \mathbf{R}[X]/(X), \qquad r \mapsto r \pmod X,$$
$$\mathbf{R}[X]/(X) \longrightarrow \mathbf{R}, \qquad g(X) \pmod X \mapsto g(0)$$

are inverse to each other, and they are compatible with arithmetic operations (addition, subtraction, multiplication).

Example 2: $\mathbf{R}[Y]/(Y - a) = \mathbf{R}$. For any $a \in \mathbf{R}$, one can make a change of variables $X = Y - a$ in Example 1, by writing $g(X) = g(Y-a) = b_0+b_1(Y-a)+\cdots+b_n(Y-a)^n = c_0+c_1Y+\cdots+c_nY^n = h(Y)$. This gives

$$\mathbf{R}[Y]/(Y - a) = \{r \pmod{(Y - a)} \mid r \in \mathbf{R}\}$$

and mutually inverse maps (again, compatible with operations)

$$\mathbf{R} \longrightarrow \mathbf{R}[Y]/(Y - a), \qquad r \mapsto r \pmod{(Y - a)},$$
$$\mathbf{R}[Y]/(Y - a) \longrightarrow \mathbf{R}, \qquad h(Y) \pmod{(Y - a)} \mapsto h(a)$$

We say that the **evaluation map** at $Y = a$

$$\mathrm{ev}_a : \mathbf{R}[Y] \longrightarrow \mathbf{R}, \qquad h(Y) \mapsto h(a)$$

induces an isomorphism of rings

$$\overline{\mathrm{ev}}_a : \mathbf{R}[Y]/(Y - a) \xrightarrow{\sim} \mathbf{R}. \tag{6.2.1.1}$$

Example 3: Construction of C. We are going to construct $\mathbf{C}$ using polynomials with real coefficients. To do that, we need to answer the following fundamental question: **what is C**? Of course, $\mathbf{C} = \{a + bi \mid a, b \in \mathbf{R}\}$, but **what is** i? One cannot really answer this question, but one can say what i^2 is: $i^2 = -1$. In other words, whenever we see i^2, we replace it by -1.

Equivalently, whenever we see $i^2 + 1$ (or its multiple), we replace it by 0.

This is analogous to what we do when we work with congruences $\pmod n$: we perform the usual arithmetic operations with integers, and whenever we see a multiple of n, we replace it by 0.

This suggests that we should consider congruences modulo $X^2 + 1$ for real polynomials.

The first thing to understand is division with remainder by $X^2 + 1$. For example,

$$X^3 + 2X^2 + 5 = (X^2 + 1)X + (2X^2 - X + 5), \qquad 2X^2 - X + 5 = (X^2 + 1) \cdot 2 - X + 3,$$
$$X^3 + 2X^2 + 5 = (X^2 + 1)(X + 2) + (3 - X).$$

In general, each polynomial $g(X) \in \mathbf{R}[X]$ can be written in a unique way as

$$g(X) = (X^2+1)h(X) + (a+bX), \qquad h(X) \in \mathbf{R}[X], \qquad a, b \in \mathbf{R}, \tag{6.2.1.2}$$

which implies that

$$g(X) \pmod{(X^2+1)} = a + bX \pmod{(X^2+1)}$$

and

$$\mathbf{R}[X]/(X^2+1) = \{a + bX \pmod{(X^2+1)} \mid a, b \in \mathbf{R}\} = \{a + bI \mid a, b \in \mathbf{R}\}$$

($a + bI = a' + b'I \iff a = a'$ and $b = b'$), where

$$I = X \pmod{(X^2+1)} \in \mathbf{R}[X]/(X^2+1), \qquad I^2 + 1 = X^2 + 1 \pmod{(X^2+1)} = 0 \pmod{(X^2+1)}.$$

This means that we have, indeed,

$$\boxed{\mathbf{R}[X]/(X^2+1) = \mathbf{C}} \tag{6.2.1.3}$$

in the sense that the evaluation map at i induces a ring isomorphism (a bijective map compatible with arithmetic operations)

$$\overline{\mathrm{ev}}_i : \mathbf{R}[X]/(X^2+1) \xrightarrow{\sim} \mathbf{C}, \qquad g(X) \pmod{(X^2+1)} \mapsto g(i). \tag{6.2.1.4}$$

Note that $-i$ is also a root of X^2+1, which means that there is another ring isomorphism

$$\overline{\mathrm{ev}}_{-i} : \mathbf{R}[X]/(X^2+1) \xrightarrow{\sim} \mathbf{C}, \qquad g(X) \pmod{(X^2+1)} \mapsto g(-i) = \overline{g(i)}, \tag{6.2.1.5}$$

which is obtained from (6.2.1.4) by composing it with complex conjugation.

Example 4: Construction of fields. In general, if K is a field and if $f(X) \in K[X]$ is an **irreducible** polynomial of degree $\deg(f) = d \geq 1$, then the quotient ring $L = K[X]/(f)$ is a field containing K, and each element of L can be written in a unique way as

$$a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1}, \qquad a_j \in K, \qquad \alpha = X \pmod{f(X)} \in L, \qquad f(\alpha) = 0. \tag{6.2.1.6}$$

For example, $\mathbf{Q}[X]/(X^2+1) = \{a + bi \mid a, b \in \mathbf{Q}\}$.

Another important case is when $K = \mathbf{Z}/p\mathbf{Z}$ for a prime p. The field L is then finite, with p^d elements (and every finite field arises in this way).

Example 5: Interpolation. If $f(X) = (X - a_1)\cdots(X - a_n)$ for distinct $a_1, \ldots, a_n \in \mathbf{C}$, then the quotient ring $\mathbf{C}[X]/(f)$ can be described in terms of Lagrange interpolation (finding a polynomial of degree $\deg < n$ with prescribed values at $a_1, \ldots, a_n$).

7 Groups

7.1 Definition and examples

7.1.1 Transformation groups We are going to treat groups from an abstract algebraic point of view, but it is important to know that many groups occur in a very concrete way as transformation groups.

Such a "concrete" group is simply a set G of invertible maps $g : X \longrightarrow X$ (for a given set X) that is stable under composition (if $g, h \in G$, then $g \circ h \in G$), under inverse (if $g \in G$, then $g^{-1} \in G$) and contains the identity map $\mathrm{id} : X \longrightarrow X$. Example : $X = \mathbf{R}^2$ and $G = \{\text{rotations of } \mathbf{R}^2 \text{ around the origin}\}$.

However, it is important to separate G from X, since the same abstract group can act as a transformation group of many different sets (in the above example, G acts not only on the set of points of $\mathbf{R}^2$, but also on the set of lines). If we do that and forget X, what will remain will be the set G equipped with a binary operation "$\circ$" (for $g, h \in G$ there is another element $g \circ h \in G$) and a distinguished element $\mathrm{id} \in G$. Their properties are given, in an abstract form, in Definition 7.1.3 below.

7.1.2 Example: $G = \mathbf{Z}$ Another example of a group is given by the set of integers $\mathbf{Z} = \{0, \pm 1, \pm 2, \ldots\}$ equipped with the operation "+" (and the corresponding inverse operation "$-$").

These operations satisfy the following identities (for all $a, b, c \in \mathbf{Z}$).

$$\begin{aligned} &(1) \quad (a+b)+c = a+(b+c) \\ &(2) \quad a+0 = 0+a = a \\ &(3) \quad a+(-a) = (-a)+a = 0 \\ &(4) \quad a+b = b+a \end{aligned} \tag{7.1.2.1}$$

7.1.3 Definition (Definition of a group)**.** A **group** is a pair $(G, *)$, where G is a set and $* : G \times G \longrightarrow G$ is a binary operation (i.e., a rule that assigns to any $g, h \in G$ an element $g * h \in G$) satisfying the following three axioms.

$$\begin{aligned} &(1) \quad \textbf{(Associativity)} \quad \forall g, h, k \in G \quad (g*h)*k = g*(h*k) \\ &(2) \quad \textbf{(Neutral element)} \quad \exists e \in G \ \forall g \in G \quad g*e = e*g = g \\ &(3) \quad \textbf{(Inverse element)} \quad \forall g \in G \ \exists h \in G \quad g*h = h*g = e \end{aligned}$$

If, in addition, the following property holds

$$(4) \quad \textbf{(Commutativity)} \quad \forall g, h \in G \quad g*h = h*g \tag{7.1.3.1}$$

then we say that $(G, *)$ an **abelian group** (named after a celebrated Norwegian mathematician N.H. Abel (1802–1829)).

7.1.4 Uniqueness We are going to show in Proposition 7.1.6 below that $e \in G$ in Axiom (2) is unique (**the neutral element of** G), and that $h \in G$ in Axiom (3) (which depends on g) is also unique (**the inverse of** g).

7.1.5 Examples of groups (1) In Example $(G, *) = (\mathbf{Z}, +)$ from 7.1.2 we have $G = \mathbf{Z}$, $* = +$, the neutral element is $e = 0$ and the inverse of $a \in \mathbf{Z}$ is equal to $-a$.
(2) $(\mathbf{R}, +)$ (more generally, any additive subgroup of $\mathbf{C}$) is an abelian group in which $e = 0$ and the inverse of a is equal to $-a$.
(3) $(\mathbf{R} \smallsetminus \{0\}, \cdot)$ and $(\mathbf{C} \smallsetminus \{0\}, \cdot)$ are abelian groups in which $e = 1$ and the inverse of a is equal to a^{-1}.
(4) $(\mathbf{Z} \smallsetminus \{0\}, \cdot)$ is **not** a group: Axioms (1) and (2) are satisfied (with $e = 1$), but $g = a \in \mathbf{Z} \smallsetminus \{0\}$ satisfies Axiom (3) if and only if there exists $b \in \mathbf{Z} \smallsetminus \{0\}$ such that $ab = ba = 1$. Such an element b exists only for $a = \pm 1$. This means that only the subset $(\{\pm 1\}, \cdot)$ is a group (abelian).
(5) Denote by $GL_2(\mathbf{R}) := \{M \in M_2(\mathbf{R}) \mid \det(M) \neq 0\}$ the set of invertible real 2×2 matrices. Then $(GL_2(\mathbf{R}), \cdot)$ is a group, in which $e = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Indeed, matrix multiplication satisfies $(M \cdot N) \cdot P = M \cdot (N \cdot P)$ and $M \cdot I = I \cdot M = M$. Furthermore the inverse of $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $M^{-1} = (ad-bc)^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. This group is not abelian, since $M \cdot N \neq N \cdot M$ for $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $N = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.
(6) $(\mathbf{Z}/n\mathbf{Z}, +)$ is an abelian group, where $e = 0 \pmod n$ and the inverse of $a \pmod n$ is $(-a) \pmod n$.
(7) $((\mathbf{Z}/n\mathbf{Z})^*, \cdot)$ is an abelian group, where $e = 1 \pmod n$ and the inverse of an invertible residue class $a \pmod n$ is $a^{-1} \pmod n$.

7.1.6 Proposition. *Let $(G,*)$ be a group.*
(1) ***(Higher associativity)*** *For any $n \geq 3$ and $g_1, \ldots, g_n \in G$, the element $g_1 * \cdots g_n \in G$ does not depend on the order in which we insert parentheses (for example, $(g_1*(g_2*g_3))*g_4 = (g_1*g_2)*(g_3*g_4) = ((g_1*g_2)*g_3)*g_4 = \cdots$).*
(2) ***The neutral element*** *$e \in G$ in Axiom (2) in* (7.1.3.1) ***is unique.***
(3) For each $g \in G$, the element $h \in G$ in Axiom (3) in (7.1.3.1) *is unique. We denote it by g^{-1} and call it* ***the inverse of*** *g.*
(4) ***(Left inverse = right inverse)*** *If $g, h \in G$ satisfy $g*h = e$, then $h = g^{-1}$ and $g = h^{-1}$ (hence $h*g = e$).*
(5) For any $g, h \in G$ we have $(g*h)^{-1} = h^{-1}*g^{-1}$ and $(g^{-1})^{-1} = g$.
(6) If $g, h, k \in G$ satisfy $g*h = g*k$ or $h*g = k*g$, then $h = k$.

Proof. (1) Exercise. (2) If e, e' satisfy $g*e = e*g = g$ and $h*e' = e'*h = h$ for all $g, h \in G$, then $e' = e*e' = e$, by taking $g = e'$ and $h = e$. (3) Similarly, if $g*h = h*g = e = g*h' = h'*g$, then $h = h*e = h*(g*h') = (h*g)*h' = e*h' = h'$. (4) If $g*h = e$, then $g^{-1} = g^{-1}*e = g^{-1}*(g*h) = (g^{-1}*g)*h = e*h = h$. (5) The identity $(g*h)*(h^{-1}*g^{-1}) = (g*(h*h^{-1}))*g^{-1} = (g*e)*g^{-1} = g*g^{-1} = e$ implies, by (4), that $(g*h)^{-1}$ is equal to $h^{-1}*g^{-1}$. (6) If $g*h = g*k$, then $g^{-1}*(g*h) = g^{-1}*(g*k)$. The left hand side is equal to $(g^{-1}*g)*h = e*h = h$ and the right hand side to $(g^{-1}*g)*k = e*k = k$. The case $h*g = k*g$ is similar. □

7.1.7 Notation (1) In an nonabelian group one often uses multiplicative notation, where $e = 1$ and the symbol for the operation is omitted. The formulae in Proposition 7.1.6 and its proof then become

$$(g_1(g_2g_3))g_4 = (g_1g_2)(g_3g_4) = ((g_1g_2)g_3)g_4, \qquad (gh)^{-1} = h^{-1}g^{-1}, \qquad g^{-1}gh = eh = h. \tag{7.1.7.1}$$

(2) In an abelian group one can choose between a **multiplicative notation** as in (1) and an **additive notation**, where $e = 0$, the operation is denoted by "+", and the inverse of $x \in G$ is denoted by $-x$. The first two formulae (7.1.7.1) then become

$$(x_1+(x_2+x_3))+x_4 = (x_1+x_2)+(x_3+x_4) = ((x_1+x_2)+x_3)+x_4, \qquad -(x+y) = (-y)+(-x) = (-x)+(-y) \tag{7.1.7.2}$$

(the last equality holds, since $(G,+)$ is an abelian group).

7.1.8 Product of groups If $(G,*)$ and $(H,\square)$ are groups, their **product** is the group

$$(G \times H, \triangle), \qquad G \times H = \{(g,h) \mid g \in G,\, h \in H\}, \qquad (g,h)\triangle(g',h') = (g*g', h\square h'). \tag{7.1.8.1}$$

In this group, $e_{G\times H} = (e_G, e_H)$ and $(g,h)^{-1} = (g^{-1}, h^{-1})$.
Example: $(\mathbf{R},+) \times (\mathbf{R},+) = (\mathbf{R}^2,+)$.

7.2 Subgroups

7.2.1 Example: $\mathbf{Z} \subset \mathbf{R}$ The additive group $(\mathbf{Z},+)$ is a subgroup of $(\mathbf{R},+)$.

7.2.2 Definition. Let $(G,*)$ be a group. A subset $H \subset G$ is a **subgroup** of $(G,*)$ if $(H,*)$ (with the operation $*$ inherited from G) is a group.
[If this is the case, then the uniqueness of the neutral element and of the inverse implies that $e_H = e_G =: e$, and that the inverse h^{-1} of any $h \in H$ is the same in G and in H.]

7.2.3 Proposition. *Let $(G,*)$ be a group, let $H \subset G$ be a subset. The following are equivalent:*
(1) H is a subgroup of $(G,*)$.
(2) $e_G \in H$ and, for all $h, h' \in H$, $h * h' \in H$ and $h^{-1} \in H$.
(3) $H \neq \emptyset$ and, for all $h, h' \in H$, $h' * h^{-1} \in H$.

Proof. (1) and (2) are equivalent, by definition (and by the uniqueness of the neutral element and of the inverse).
(2) $\Longrightarrow$ (3): H is non-empty, since $e = e_G \in H$. If $h, h' \in H$, then $h^{-1} \in H$, hence $h' * h^{-1} \in H$.
(3) $\Longrightarrow$ (2): as H is non-empty, there exists $k \in H$; then $k * k^{-1} = e \in H$. If $h, h' \in H$, then $e * h^{-1} = h^{-1} \in H$, hence $h' * (h^{-1})^{-1} = h' * h \in H$. □

7.2.4 Examples of subgroups (1) Both $\{e\}$ and G are subgroups of $(G,*)$.

(2) According to Theorem 2.3.2, subgroups of $(\mathbf{Z},+)$ are the subsets $d\mathbf{Z}$ $(d \in \mathbf{N})$.

(3) For a non-empty set X, denote by $(S_X, \circ)$ the group of permutations of X (with operation given by composition):

$$S_X := \{\text{bijective maps } \alpha : X \longrightarrow X\}, \qquad (\beta \circ \alpha)(x) = \beta(\alpha(x)), \qquad \beta \circ \alpha : X \xrightarrow{\alpha} X \xrightarrow{\beta} X.$$

The neutral element is the identity map $e = \mathrm{id}_X$ $(\mathrm{id}(x) = x$, for all $x \in X)$.

In the special case $X = \{1, 2, \ldots, n\}$ $(n \geq 1)$, $S_X = S_n$ is the **symmetric group** on n elements. A permutation $\alpha \in S_n$ is then written as

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha(1) & \alpha(2) & \cdots & \alpha(n) \end{pmatrix}, \qquad e = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

For $n > 2$ the group S_n is nonabelian. For example, if $n = 3$, then

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \qquad \beta \circ \alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \neq \alpha \circ \beta = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

For any subset $Y \subset X$,

$$H := \{\alpha \in S_X \mid \alpha(Y) = Y\}$$

is a subgroup of S_X. For example, if $X = \{1, 2, \ldots, n\}$ and $Y = \{n\}$, then $S_X = S_n$ and $H = S_{n-1}$.

(4) If V is a vector space over a field K, then the **general linear group**

$$GL(V) := \{\alpha : V \longrightarrow V \mid \alpha \text{ is bijective and } K\text{-linear}\} \subset S_V$$

is a subgroup of S_V. In the special case when $V = K^n$ consists of column vectors with $n \geq 1$ entries, then each linear map $\alpha : K^n \longrightarrow K^n$ can be written in a matrix form $\alpha(x) = Ax$, for some $A \in M_n(K)$. The composition of two linear maps $\alpha : x \mapsto Ax$ and $\beta : y \mapsto By$ is then equal to $\beta \circ \alpha : x \mapsto B(Ax) = (BA)x$. Therefore

$$GL(V) = GL(K^n) = \{x \mapsto Ax \ (x \in K^n) \mid A \in M_n(K), \det(A) \neq 0\}$$

can be identified with the matrix group

$$GL_n(K) = \{A \in M_n(K) \mid \det(A) \neq 0\}$$

(with operation given by matrix multiplication). The **special linear group**

$$SL_n(K) := \{A \in M_n(K) \mid \det(A) = 1\}$$

is a subgroup of $GL_n(K)$.

(5) In the situation of (4), one can combine linear maps with translations and obtain the **general affine group**

$$GA(V) := \{x \mapsto \alpha(x) + a \ (x \in V) \mid \alpha \in GL(V),\, a \in V\} \subset S_V.$$

The group of **translations**

$$\{x \mapsto x + a \ (x \in V) \mid a \in V\} \subset GA(V)$$

is a subgroup of $GA(V)$.

For $V = K^n$, the general affine group

$$GA_n(K) = GA(K^n) = \{x \mapsto Ax + a \ (x \in K^n) \mid A \in GL_n(K),\, a \in K^n\}$$

can be expressed in a matrix form as

$$GA_n(K) = \left\{ \begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix} \mid A \in GL_n(K),\, a \in K^n \right\},$$

since the composition of the maps $x \mapsto Ax + a$ and $y \mapsto By + b$ is equal to $x \mapsto B(Ax + a) + b = (BA)x + (Ba + b)$ and

$$\begin{pmatrix} B & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} BA & Ba + b \\ 0 & 1 \end{pmatrix}.$$

(6) For $n \geq 1$, the **orthogonal group**

$$O(n) := \{A \in M_n(\mathbf{R}) \mid {}^tAA = I\}$$

(where ${}^tA$ denotes the transpose matrix) is a subgroup of $GL_n(\mathbf{R})$. The **special orthogonal group**

$$SO(n) := \{A \in O(n) \mid \det(A) = 1\} = O(n) \cap SL_n(\mathbf{R})$$

is a subgroup of both $O(n)$ and $SL_n(\mathbf{R})$.

For $n = 2$, $SO(2)$ is the group of matrices representing the rotations of $\mathbf{R}^2$ around the origin.

Similarly, the **unitary group**

$$U(n) := \{A \in M_n(\mathbf{C}) \mid {}^tA\overline{A} = I\}$$

is a subgroup of $GL_n(\mathbf{C})$. The **special unitary group**

$$SU(n) := \{A \in U(n) \mid \det(A) = 1\} = U(n) \cap SL_n(\mathbf{C})$$

is a subgroup of both $U(n)$ and $SL_n(\mathbf{C})$.

(7) For any integer $n \geq 1$, the set of **n-th roots of unity**

$$\mu_n := \{z \in \mathbf{C} \mid z^n = 1\}$$

is a subgroup of $(\mathbf{C} \smallsetminus \{0\}, \cdot)$. For example,

$$\mu_1 = \{1\}, \qquad \mu_2 = \{\pm 1\}, \qquad \mu_3 = \{1, \tfrac{-1 \pm i\sqrt{3}}{2}\}, \qquad \mu_4 = \{\pm 1, \pm i\}, \qquad \mu_6 = \{\pm 1, \tfrac{1 \pm i\sqrt{3}}{2}, \tfrac{-1 \pm i\sqrt{3}}{2}\}.$$

(8) For any non-zero complex number $a \in \mathbf{C} \smallsetminus \{0\}$, the set of all powers of a

$$\langle a\rangle := \{a^n \mid n \in \mathbf{Z}\} \tag{7.2.4.1}$$

is a subgroup of $(\mathbf{C}\smallsetminus\{0\},\cdot)$. For example,

$$\begin{gathered}\langle -1\rangle = \{-1,1\} = \mu_2, \qquad \langle i\rangle = \{i, i^2=-1, i^3=-i, i^4=1\} = \mu_4,\\ \langle -i\rangle = \{i, (-i)^2=-1, (-i)^3=i, (-i)^4=1\} = \mu_4.\end{gathered} \tag{7.2.4.2}$$

(9) The **centre** $Z(G) := \{z \in G \mid \forall g \in G \quad zg = gz\}$ of $(G,*)$ is a subgroup of $(G,*)$. Note that $Z(G) = G \iff (G,*)$ is abelian.

7.2.5 Exercise. Show that $O(n)$ (resp. $U(n)$) is, indeed, a subgroup of $GL_n(\mathbf{R})$ (resp. of $GL_n(\mathbf{C})$).

7.2.6 Proposition. *Let $(G,*)$ be a group. The intersection $H := \bigcap_{i\in I} H_i \subset G$ of any set of subgroups $H_i \subset G$ $(i \in I)$ is again a subgroup of $(G,*)$.*

Proof. Each H_i contains the neutral element e of G, which implies that $e \in H$. If $h, h' \in H = \bigcap H_i$, then $h' * h^{-1} \in H_i$ (since each H_i is a subgroup), and therefore $h' * h^{-1} \in \bigcap_{i\in I} H_i = H$. □

7.2.7 Definition (Subgroup generated by a subset)**.** Let $(G,*)$ be a group. For a non-empty subset $S \subset G$, the intersection

$$\langle S\rangle := \bigcap_{\substack{H\subset(G,*)\\ S\subset H}} H$$

of all subgroups $H \subset (G,*)$ containing S is the smallest subgroup of $(G,*)$ containing S. We say that $\langle S\rangle$ is **the subgroup of** $(G,*)$ **generated by** S. If $S = \{g\}$ consists of one element, then we say that $\langle g\rangle := \langle\{g\}\rangle \subset G$ is **the cyclic subgroup generated by** g.

7.2.8 Example: cyclic subgroups of $\mathbf{C}^*$ Any subgroup of $(\mathbf{C}\smallsetminus\{0\},\cdot)$ containing a given complex number $a \in \mathbf{C}\smallsetminus\{0\}$ must also contain the following elements (for all integers $n \geq 1$):

$$\begin{gathered}1, \quad a, \quad a^2 = a\cdot a, \quad a^3 = a\cdot a\cdot a, \quad \cdots \quad a^n = \underbrace{a\cdot a\cdots a}_{n \text{ times}}, \quad \cdots \quad a^{-1},\\ (a^{-1})^2 = a^{-2}, \cdots \quad a^{-n} = \underbrace{a^{-1}\cdot a^{-1}\cdots a^{-1}}_{n \text{ times}},\end{gathered} \tag{7.2.8.1}$$

which implies that

$$\langle a\rangle = \{a^n \mid n \in \mathbf{Z}\} \tag{7.2.8.2}$$

is, indeed, given by the formula (7.2.4.1).

7.2.9 Isometries of $\mathbf{R}^n$ One uses the standard scalar product on $\mathbf{R}^n$

$$(x \mid y) := {}^t\!xy = x_1y_1 + \cdots x_ny_n, \qquad x = \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix} \in \mathbf{R}^n, \qquad y = \begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix} \in \mathbf{R}^n$$

to define the **norm** $\|x\| := (x \mid x)^{1/2}$ and the **distance** $d(x,y) := \|x-y\|$ $(x, y \in \mathbf{R}^n)$ on $\mathbf{R}^n$.

An **isometry of $\mathbf{R}^n$** is a map $f : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ that preserves distance:

$$\forall x, y \in \mathbf{R}^n \quad d(f(x), f(y)) = d(x, y). \tag{7.2.9.1}$$

For example, any translation $f(x) = x + a$ ($a \in \mathbf{R}^n$) is an isometry.

One can show that any isometry satisfying $f(0) = 0$ is an $\mathbf{R}$-linear map $x \mapsto Ax$ ($A \in M_n(\mathbf{R})$) (exercise!). Condition (7.2.9.1) is then equivalent to

$$\forall x, y \in \mathbf{R}^n \quad (Ax \mid Ay) = (x \mid y) \iff \forall x, y \in \mathbf{R}^n \quad {}^t x({}^t AA)y = {}^t xy \iff {}^t AA = I_n \iff A \in O(n).$$

Consequently,

$$\{\text{Isometries of } \mathbf{R}^n\} = \{x \mapsto Ax + a \ (x \in \mathbf{R}^n) \mid A \in O(n),\, a \in \mathbf{R}^n\} \subset GA_n(\mathbf{R}).$$

7.3 Cyclic groups, cyclic subgroups

We begin by checking that the formula (7.2.8.2) holds in general.

7.3.1 Powers of $g \in G$ Let $(G, *)$ be a group, let $g \in G$. Define integral powers $g^n \in G$ ($n \in \mathbf{Z}$) as follows (cf. Section 3.4.5).

$$g^0 := e, \qquad g^m := \underbrace{g * g * \cdots * g}_{m \text{ times}}, \qquad g^{-m} := \underbrace{g^{-1} * g^{-1} * \cdots * g^{-1}}_{m \text{ times}} \qquad (m \in \mathbf{N}_+). \tag{7.3.1.1}$$

7.3.2 Proposition. *If $(G, *)$ is a group and $g \in G$, then*

$$\forall m, n \in \mathbf{Z} \qquad g^m * g^n = g^{m+n} = g^n * g^m, \qquad (g^m)^n = g^{mn}. \tag{7.3.2.1}$$

Proof. This can be checked by a somewhat tedious but straightforward argument. For example, $g^3 * g^{-5} = g*g*g*g^{-1}*g^{-1}*g^{-1}*g^{-1}*g^{-1} = g^{-1}*g^{-1} = g^{-2}$. The details are left to the reader as an exercise. □

7.3.3 Corollary. *The cyclic subgroup $\langle g \rangle \subset G$ is equal to $\{g^n \mid n \in \mathbf{Z}\}$ and is abelian. Moreover, $\langle g \rangle = \langle g^{-1} \rangle$.*

7.3.4 Definition. A group $(G, *)$ is **cyclic** and $g \in G$ is its **generator** if $G = \langle g \rangle = \{g^n \mid n \in \mathbf{Z}\}$. [According to Corollary 7.3.3, if g is a generator of G, so is g^{-1}.]

7.3.5 Multiplicative vs additive notation Let us compare the concepts that were introduced above using multiplicative notation (which makes sense in arbitrary groups) with additive notation (which makes sense only for abelian groups).

Multiplicative notation $(G, *)$ any group	Additive notation $(G, +)$ abelian group
$g * h = gh$	$g + h = h + g$
$e = 1$	$e = 0$
g^{-1}	$-g$
gh^{-1}	$g + (-h) = (-h) + g$
$(gh)^{-1} = h^{-1}g^{-1}$	$-(g + h) = (-h) + (-g) = (-g) + (-h)$
$g^m := \underbrace{g \cdots g}_{m \text{ times}} \quad (m > 0)$	$mg := \underbrace{g + \cdots + g}_{m \text{ times}}$
$g^{-m} := (g^{-1})^m = (g^m)^{-1}$	$(-m)g := m(-g) = -(mg)$
$\langle g \rangle = \{g^n \mid n \in \mathbf{Z}\} = \langle g^{-1} \rangle$	$\langle g \rangle = \{ng \mid n \in \mathbf{Z}\} = \langle -g \rangle$

7.3.6 Examples of cyclic groups (1) For any $d \in \mathbf{Z}$, $(d\mathbf{Z}, +) = \langle d \rangle = ((-d)\mathbf{Z}, +) = \langle -d \rangle \subset \mathbf{Z}$ is a cyclic subgroup of $(\mathbf{Z}, +)$. As shown in Theorem 2.3.2, all subgroups of $\mathbf{Z}$ are of this form.

(2) For any $z \in \mathbf{C}$, the cyclic subgroup of $(\mathbf{C}, +)$ generated by z is equal to $\langle z \rangle = \{0, \pm z, \pm 2z, \pm 3z, \ldots\} \subset \mathbf{C}$.

(3) For each $n \in \mathbf{N}_+$, $(\mathbf{Z}/n\mathbf{Z}, +)$ is a cyclic group, generated by $1 \pmod n$ (and also by $-1 \pmod n$). The set of all generators of this group will be described in Example 3 of Section 7.5.3 (see also 7.5.7).

(4) As we saw in Section 4.3.2, $((\mathbf{Z}/7\mathbf{Z})^*, \cdot)$ is a cyclic group, generated by $3 \pmod 7$ (and also by $5 \pmod 7$).

(5) For each $n \in \mathbf{N}_+$, the group of n-th roots of unity is cyclic:

$$\mu_n = \{z \in \mathbf{C} \mid z^n = 1\} = \{e^{2\pi ik/n} = \cos(\tfrac{2\pi ik}{n}) + i\sin(\tfrac{2\pi ik}{n}) = (e^{2\pi i/n})^k \mid 1 \leq k \leq n\} = \langle e^{2\pi i/n} \rangle.$$

(6) For each $n \in \mathbf{N}_+$, the group of rotations of $\mathbf{R}^2$ around the origin that preserve a regular polygon with n sides (with centre at the origin) form a cyclic group (a subgroup of $SO(2)$)

$$C_n = \{r, r^2, \ldots, r^n = \mathrm{id}\} = \langle r \rangle,$$

where r^k is a rotation with angle $\frac{2\pi k}{n}$.

7.3.7 Exercise. What is the relation between Examples (5) and (6) in Section 7.3.6?

7.4 Group homomorphisms

7.4.1 Exponential map It is important to understand not just individual groups, but also relations between them. For example, the exponential map

$$\exp : (\mathbf{R}, +) \longrightarrow (\mathbf{R} \smallsetminus \{0\}, \cdot), \qquad \exp(x) = e^x \tag{7.4.1.1}$$

relates the operations in the two groups involved (addition and multiplication):

$$\exp(x + y) = \exp(x) \cdot \exp(y). \tag{7.4.1.2}$$

This is a special case of the following general concept.

7.4.2 Definition. Let $(G, *), (H, \square)$ be groups. A **group homomorphism** $f : (G, *) \longrightarrow (H, \square)$ is a map $f : G \longrightarrow H$ such that $\forall g, g' \in G \quad f(g * g') = f(g) \square f(g')$.

7.4.3 Examples of group homomorphisms (1) $f : G \longrightarrow H$, $f(g) = e_H$ for all $g \in G$.

(2) $\mathrm{id} : G \longrightarrow G$, $\mathrm{id}(g) = g$ for all $g \in G$.

(3) $f = [\times 6] : (\mathbf{Z}, +) \longrightarrow (3\mathbf{Z}, +)$, $f(n) = 6n$. In this case $6(m + n) = 6m + 6n$.

(4) $\exp : (\mathbf{R}, +) \longrightarrow (\mathbf{R} \smallsetminus \{0\}, \cdot)$, $\exp(x) = e^x$. In this case $e^{x+y} = e^x e^y$.

(5) $\exp : (\mathbf{C}, +) \longrightarrow (\mathbf{C} \smallsetminus \{0\}, \cdot)$, $\exp(x + iy) = e^{x+iy} = e^x(\cos(y) + i\sin(y))$. In this case $e^{z+z'} = e^z e^{z'}$.

(6) $\det : (GL_n(\mathbf{R}), \cdot) \longrightarrow (\mathbf{R} \smallsetminus \{0\}, \cdot)$, $\det(MN) = \det(M)\det(N)$.

(7) If $(G, *)$ is a group and $g \in G$, then the map $f : \mathbf{Z} \longrightarrow G$, $f(n) = g^n$ is a group homomorphism, since $g^{m+n} = g^m g^n$.

(8) If $(G, *)$ is a group and $H \subset G$ a subgroup, then the inclusion $H \hookrightarrow G$ is a group homomorphism.

(9) Canonical projection $\mathrm{pr} : (\mathbf{Z}, +) \longrightarrow (\mathbf{Z}/n\mathbf{Z}, +)$, $\mathrm{pr}(a) = a \pmod n$. In this case $(a + b) \pmod n = (a \pmod n) + (b \pmod n)$.

(10) The vertical projection $f : (\mathbf{R}^2, +) \longrightarrow (\mathbf{R}, +)$, $f(\begin{pmatrix} x \\ y \end{pmatrix}) = x$.

(11) If $f_1 : (G, *) \longrightarrow (H, \square)$ and $f_2 : (H, \square) \longrightarrow (K, \triangle)$ are group homomorphisms, so is their composition $f_2 \circ f_1 : (G, *) \longrightarrow (H, \square) \longrightarrow (K, \triangle)$ ($(f_2 \circ f_1)(g) = f_2(f_1(g))$).

(12) If $f_i : G_i \longrightarrow H_i$ ($i = 1, 2$) are group homomorphisms, so is $f_1 \times f_2 : G_1 \times G_2 \longrightarrow H_1 \times H_2$.

7.4.4 Proposition. *If $f : (G, *) \longrightarrow (H, \square)$ is a group homomorphism, then:*
(1) ***(Compatibility with neutral elements)*** *$f(e_G) = e_H$.*
(2) ***(Compatibility with inverse)*** *$\forall g \in G \quad f(g^{-1}) = f(g)^{-1}$.*
(3) If the map $f : G \longrightarrow H$ is bijective, then its inverse $f^{-1} : (H, \square) \longrightarrow (G, *)$ is also a group homomorphism. We say that f is a ***group isomorphism*** *(which implies that f^{-1} is a group isomorphism, too). [In this case the two groups are "the same" from a purely algebraic point of view, but this "sameness" is given by the map f, which is a part of the data.]*
(4) If f is injective, then it defines a group isomorphism $f : (G, *) \xrightarrow{\sim} (\mathrm{Im}(f), \square)$.

Proof. (1) $\forall g \in G \quad f(g) \square e_H = f(g) = f(g * e_G) = f(g) \square f(e_G)$; thus $e_H = f(e_G)$. Part (2) follows from the fact that $f(g) \square f(g^{-1}) = f(g * g^{-1}) = f(e_G) = e_H$.
(3) If f is bijective and $h, h' \in H$, then there exist unique $g, g' \in G$ such that $h = f(g)$ and $h' = f(g')$ ($g = f^{-1}(h)$, $g' = f^{-1}(h')$). The identity $h \square h' = f(g) \square f(g') = f(g * g')$ then implies that $f^{-1}(h \square h') = g * g' = f^{-1}(h) * f^{-1}(h')$, as claimed in (3). Finally, (4) is an immediate consequence of the definitions, since the injectivity of f implies that $f : G \longrightarrow \mathrm{Im}(f)$ is bijective. □

7.4.5 Definition. The **kernel** and the **image** of a group homomorphism $f : (G, *) \longrightarrow (H, \square)$ are defined, respectively, as

$$\mathrm{Ker}(f) := \{g \in G \mid f(g) = e_H\} \subset G, \qquad \mathrm{Im}(f) := \{f(g) \mid g \in G\} \subset H.$$

7.4.6 Proposition. *If $f : (G, *) \longrightarrow (H, \square)$ is a group homomorphism, then:*
(1) $\mathrm{Ker}(f)$ *is a subgroup of G.*
(2) $\mathrm{Im}(f)$ *is a subgroup of H.*
(3) For $g, g' \in G$, the following are equivalent: $f(g) = f(g') \iff g^{-1} * g' \in \mathrm{Ker}(f) \iff g' * g^{-1} \in \mathrm{Ker}(f)$.

Proof. (1) We know that $f(e_G) = e_H$; thus $e_G \in \mathrm{Ker}(f)$. If $g, g' \in \mathrm{Ker}(f)$, then $f(g) = f(g') = e_H$. It follows that $f(g^{-1}) = f(g)^{-1} = e_H^{-1} = e_H$ and $f(g * g') = f(g) \square f(g') = e_H \square e_H = e_H$, hence $g^{-1}, g * g' \in \mathrm{Ker}(f)$.
(2) Similarly, $e_H = f(e_G) \in \mathrm{Im}(f)$. If $h, h' \in \mathrm{Im}(f)$, then $h = f(g)$ and $h' = f(g')$ for some $g, g' \in G$. Consequently, $h^{-1} = f(g)^{-1} = f(g^{-1}) \in \mathrm{Im}(f)$ and $h \square h' = f(g) \square f(g') = f(g * g') \in \mathrm{Im}(f)$.
(3) $f(g) = f(g') \iff e_H = f(g)^{-1} \square f(g') = f(g^{-1} * g') \iff g^{-1} * g' \in \mathrm{Ker}(f)$. The second equivalence is similar. □

7.4.7 Examples of $\mathrm{Ker}(f)$ **and** $\mathrm{Im}(f)$ Let us describe the kernel and the image of each group homomorphism from Section 7.4.3.

(1) The trivial homomorphism $f(g) = e_H$: $\mathrm{Ker}(f) = G$, $\mathrm{Im}(f) = \{e_H\}$.

(2) The identity $\mathrm{id} : G \longrightarrow G$: $\mathrm{Ker}(\mathrm{id}) = \{e_G\}$, $\mathrm{Im}(f) = G$.

(3) $f : (\mathbf{Z}, +) \longrightarrow (3\mathbf{Z}, +)$, $f(n) = 6n$: $\mathrm{Ker}(f) = \{0\}$, $\mathrm{Im}(f) = (6\mathbf{Z}, +)$.

(4) $f = \exp : (\mathbf{R}, +) \longrightarrow (\mathbf{R} \smallsetminus \{0\}, \cdot)$: $\mathrm{Ker}(f) = \{0\}$, $\mathrm{Im}(f) = (\mathbf{R}_{>0}, \cdot)$.

(5) $f = \exp : (\mathbf{C}, +) \longrightarrow (\mathbf{C} \smallsetminus \{0\}, \cdot)$: $\mathrm{Ker}(f) = \{x + iy \in \mathbf{C} \mid e^x(\cos(y) + i\sin(y)) = 1\} = 2\pi i\mathbf{Z}$, $\mathrm{Im}(f) = (\mathbf{C} \smallsetminus \{0\}, \cdot)$.

(6) $\det : (GL_n(\mathbf{R}), \cdot) \longrightarrow (\mathbf{R} \smallsetminus \{0\}, \cdot)$: $\mathrm{Ker}(\det) = SL_n(\mathbf{R})$, $\mathrm{Im}(\det) = (\mathbf{R} \smallsetminus \{0\}, \cdot)$.

(7) $f : (\mathbf{Z}, +) \longrightarrow (G, *)$, $f(n) = g^n$: $\mathrm{Im}(f) = \langle g \rangle$.

(8) $f : H \hookrightarrow (G, *)$: $\mathrm{Ker}(f) = \{e_H\} = \{e_G\}$, $\mathrm{Im}(f) = H$.
(9) $f = \mathrm{pr} : (\mathbf{Z}, +) \longrightarrow (\mathbf{Z}/n\mathbf{Z}, +)$: $\mathrm{Ker}(\mathrm{pr}) = (n\mathbf{Z}, +)$, $\mathrm{Im}(\mathrm{pr}) = \mathbf{Z}/n\mathbf{Z}$.
(10) The vertical projection $f : (\mathbf{R}^2, +) \longrightarrow (\mathbf{R}, +)$: $\mathrm{Ker}(f) = \{\begin{pmatrix} 0 \\ y \end{pmatrix} \mid y \in \mathbf{R}\}$, $\mathrm{Im}(f) = \mathbf{R}$.

7.4.8 Examples of isomorphisms (1) The identity $\mathrm{id} : G \longrightarrow G$.
(2) Conjugation by $g \in G$: $f : G \longrightarrow G$, $f(h) = ghg^{-1}$. In this case $g(hh')g^{-1} = (ghg^{-1})(gh'g^{-1})$. The inverse of f is given by $f^{-1}(g) = g^{-1}hg$, since $g^{-1}ghg^{-1}g = ehe = h$.
(3) If $m, n \in \mathbf{N}_+$ and $\gcd(m, n) = 1$, then the map in the Chinese Remainder Theorem

$$f : (\mathbf{Z}/mn\mathbf{Z}, +) \longrightarrow (\mathbf{Z}/m\mathbf{Z}, +) \times (\mathbf{Z}/n\mathbf{Z}, +), \qquad f(a \pmod{mn}) = (a \pmod{m}, a \pmod{n})$$

is a bijective group homomorphism, hence a group isomorphism.

7.4.9 Exercise. An **automorphism** of a group G is a group isomorphism $G \longrightarrow G$. Show that:
(1) $\mathrm{Aut}(G) := \{\text{automorphisms of } G\}$ is a subgroup of S_G.
(2) The map $G \longrightarrow \mathrm{Aut}(G)$ sending $g \in G$ to the conjugation by g (see Example 2 in Section 7.4.8) is a group homomorphism. Its image is called the group of **inner automorphisms** of G.
(3) Determine the kernel of $G \longrightarrow \mathrm{Aut}(G)$.

7.4.10 Proposition. *A group homomorphism $f : (G, *) \longrightarrow (H, \square)$ is injective $\iff$ $\mathrm{Ker}(f) = \{e_G\}$.*

Proof. By definition, f is injective if, for $g, g' \in G$, the equality $f(g) = f(g')$ is equivalent to $g = g'$. However,

$$\begin{aligned} g = g' &\iff g' * g^{-1} = e_G \\ f(g) = f(g') &\iff g' * g^{-1} \in \mathrm{Ker}(f), \end{aligned}$$

which means that the conditions $g = g'$ and $f(g) = f(g')$ are equivalent if and only if $\{e_G\} = \mathrm{Ker}(f)$. $\square$

7.4.11 Corollary. *A group homomorphism $f : (G, *) \longrightarrow (H, \square)$ with $\mathrm{Ker}(f) = \{e_G\}$ is injective, hence defines a group isomorphism $f : (G, *) \xrightarrow{\sim} (\mathrm{Im}(f), \square)$.*

7.4.12 Example: exp and log The group homomorphism (7.4.1.1) given by the exponential induces a group isomorphism

$$\exp : (\mathbf{R}, +) \xrightarrow{\sim} (\mathbf{R}_{>0}, \cdot).$$

Its inverse (which is also a group isomorphism) is given by the logarithm map

$$\ln : (\mathbf{R}_{>0}, \cdot) \xrightarrow{\sim} (\mathbf{R}, +).$$

7.4.13 Exercise. Let $f : (G, *) \longrightarrow (H, \square)$ be a group homomorphism.
(1) The image $f(G_1) = \{f(g) \mid g \in G_1\}$ of any subgroup G_1 of G is a subgroup of H.
(2) The inverse image $f^{-1}(H_1) = \{g \in G \mid f(g) \in H_1\}$ of any subgroup H_1 of H is a subgroup of G.
(3) In the notation of (1) and (2), $G_1 \subseteq f^{-1}(f(G_1))$ and $f(f^{-1}(H_1)) \subseteq H_1$.
(4) Give an example in which $G_1 \neq f^{-1}(f(G_1))$ and $f(f^{-1}(H_1)) \neq H_1$.

7.4.14 Embedding of G into S_G (Cayley) Let $(G,*)$ be a group. Left translations

$$L(g) : G \longrightarrow G, \qquad h \mapsto g * h$$

by elements $g \in G$ are bijective maps satisfying

$$L(e) = \mathrm{id}, \qquad L(g * g') = L(g) \circ L(g'), \qquad L(g)(e) = g,$$

which implies that the map

$$L : G \longrightarrow S_G, \qquad g \mapsto L(g)$$

is an injective group homomorphism. In particular, any finite group of order $|G| = n$ is isomorphic to a subgroup $\mathrm{Im}(L) \subset S_G \xrightarrow{\sim} S_n$ of S_n.

7.5 Order, cyclic (sub)groups, Lagrange's theorem

7.5.1 A preview In Section 4.3 we investigated various properties of powers of invertible residue classes $(\mathrm{mod}\, n)$. This can be done in the following general setting.

7.5.2 Definition (Order)**.** Let $(G,*)$ be a group, let $g \in G$. The **order of** G is the number of elements $|G|$ of G ($|G| \in \mathbf{N}_+ \cup \{\infty\}$). The **order of** g is defined as $\min\{k \in \mathbf{N}_+ \mid g^k = e\}$ (if no such k exists, then the order of g is defined to be ∞).

7.5.3 Examples (1) If $G = ((\mathbf{Z}/n\mathbf{Z})^*, \cdot)$, then we recover Definition 4.3.3.

(2) If $(G,+)$ is an abelian group written additively, then the order of $g \in G$ is $\min\{k \in \mathbf{N}_+ \mid kg = 0\}$ (or ∞, if no such k exists).

(3) If $G = (\mathbf{Z}/m\mathbf{Z}, +)$ and $g = a \ (\mathrm{mod}\, m)$ for some $a \in \mathbf{Z} \smallsetminus \{0\}$, then $kg = ka \ (\mathrm{mod}\, m)$, which means that the order of g is equal to

$$d = \min\{k \geq 1 \mid ka \equiv 0 \ (\mathrm{mod}\, m)\}.$$

In other words, $d|a| \geq 1$ is the smallest positive multiple of $|a|$ divisible by m, hence $d|a| = \mathrm{lcm}(|a|, m)$ and $d = \mathrm{lcm}(|a|, m)/|a| = m/\gcd(|a|, m)$. In particular, $d = m$ if and only if $\gcd(|a|, m) = 1$ (cf. Proposition 4.3.13 and its proof).

7.5.4 Cyclic groups We are now going to show that a cyclic group is determined, up to isomorphism, by its order. More precisely, it is isomorphic to $(\mathbf{Z}/m\mathbf{Z}, +)$ if it has finite order m, and to $(\mathbf{Z}, +)$ if its order is infinite. Furthermore, every subgroup of a cyclic group is cyclic.

7.5.5 Proposition. *Let $(G,*)$ be a group, let $g \in G$. Consider the group homomorphism $f : (\mathbf{Z}, +) \longrightarrow G$ given by $f(n) = g^n$.*
(1) If the order of $g \in G$ is equal to ∞, then $\mathrm{Ker}(f) = \{0\}$ and f induces a group isomorphism $(\mathbf{Z}, +) \xrightarrow{\sim} \langle g \rangle$. An element g^k ($k \in \mathbf{Z}$) is a generator of the cyclic group $\langle g \rangle$ if and only if $k = \pm 1$. The set of subgroups of $\langle g \rangle$ is equal to $\{\langle g^d \rangle \mid d \in \mathbf{N}\}$.
(2) If the order of $g \in G$ is equal to $m \in \mathbf{N}_+$, then $\mathrm{Ker}(f) = m\mathbf{Z}$.

Proof. By definition, $\mathrm{Ker}(f) = \{n \in \mathbf{Z} \mid g^n = e\}$. This is a subgroup of $(\mathbf{Z}, +)$, which means that $\mathrm{Ker}(f) = m\mathbf{Z}$ for some $m \in \mathbf{N}$. The order of g is equal to $\min\{k \in \mathrm{Ker}(f) = m\mathbf{Z} \mid k > 0\}$, hence to m if $m > 0$ (resp. to ∞ when no such k exists, which is equivalent to $m = 0$). This proves the description of $\mathrm{Ker}(f)$, both in (1) and (2).

If the order of g is infinite, then $\mathrm{Ker}(f) = \{0\}$ by the above, which implies that the group homomorphism f is injective. Therefore it induces a group isomorphism between $(\mathbf{Z}, +)$ and $(\mathrm{Im}(f), *) = \langle g \rangle$ (the cyclic

subgroup of G generated by g), hence a bijection between the set of subgroups $\{d\mathbf{Z} \mid d \in \mathbf{N}\}$ of $\mathbf{Z}$ and the set of subgroups $\{f(d\mathbf{Z}) = \langle g^d \rangle\}$ of $\langle g \rangle$. Finally, $f(k) = g^k$ is a generator of $\langle g \rangle$ if and only if k is a generator of $\mathbf{Z}$, which is equivalent to $k\mathbf{Z} = \mathbf{Z}$, hence to $k = \pm 1$. □

7.5.6 Theorem. *Let $(G, *)$ be a group. Assume that $g \in G$ has order $m \in \mathbf{N}_+$.*
(1) For $k \in \mathbf{Z}$, it is equivalent: $g^k = e \iff m \mid k$.
(2) For $k, l \in \mathbf{Z}$, it is equivalent: $g^k = g^l \iff m \mid (k-l) \iff k \equiv l \pmod m$.
(3) The subgroup $\{g^k \mid k \in \mathbf{Z}\} = \langle g \rangle$ is equal to $\{g, g^2, \dots, g^m = e\}$ and has m elements.
(4) The map $\overline{f} : \mathbf{Z}/m\mathbf{Z} \longrightarrow \langle g \rangle$ given by $\overline{f}(k \pmod m) = g^k$ is well-defined. It is a group isomorphism $\overline{f} : \mathbf{Z}/m\mathbf{Z} \xrightarrow{\sim} \langle g \rangle$.
(5) For $k \in \mathbf{Z} \smallsetminus \{0\}$, the order of g^k is equal to $m/\gcd(|k|, m)$.
(6) The cyclic group $\langle g \rangle$ (of order m) has $\varphi(m)$ generators.
(7) More generally, the number of elements of $\langle g \rangle$ of order $d \in \mathbf{N}_+$ is equal to zero (resp. to $\varphi(d)$) if d does not divide m (resp. if d divides m).
(8) The set of subgroups of $\langle g \rangle$ is equal to $\{\langle g^d \rangle \mid d \mid m\}$ (where $\langle g^d \rangle$ is a cyclic group of order m/d).

Proof. (1), (2) Let $f : (\mathbf{Z}, +) \longrightarrow G$ be the group homomorphism $f(k) = g^k$; then $g^k = e \iff k \in \operatorname{Ker}(f)$, but $\operatorname{Ker}(f) = m\mathbf{Z}$, as shown in Proposition 7.5.5(2). Consequently, $g^k = g^l \iff g^{k-l} = g^k(g^l)^{-1} = e \iff m \mid (k-l)$.
(3) Each $k \in \mathbf{Z}$ can be written as $k = ma + l$, where $a, l \in \mathbf{Z}$ and $1 \leq l \leq m$; then $g^k = (g^m)^a g^l = g^l$, and the m values of $g, g^2, \dots, g^m = e$ are distinct, by (2).
(4) If $k \equiv l \pmod m$, then $g^k = g^l$, which means that the map $\overline{f}$ is well-defined. It satisfies $\overline{f}((k \pmod m) + (l \pmod m)) = \overline{f}(k \pmod m)\overline{f}(l \pmod m))$, since the left hand side is equal to $\overline{f}((k+l) \pmod m) = g^{k+l}$ and the right hand side to $g^k g^l$. This means that $\overline{f}$ is a group homomorphism. Finally, $\overline{f}$ is bijective, by (3).
(5) The order of g^k is the smallest integer $d \geq 1$ such that $(g^k)^d = g^{dk} = e$, which is equivalent to $m \mid dk$. Therefore $d|k|$ is the smallest positive multiple of $|k|$ that is divisible by m, which implies that $d|k| = \operatorname{lcm}(|k|, m)$ and $d = \operatorname{lcm}(|k|, m)/|k| = m/\gcd(|k|, m)$.
(6) Among the m elements $g, g^2, \dots, g^m = e$ of the cyclic group $\langle g \rangle$, g^k is its generator if and only if the order of g^k is equal to m, which is equivalent to $\gcd(k, m) = 1$. There are $\varphi(m)$ possible values of such k in the range $1 \leq k \leq m$.
(7) According to (5), possible orders of the elements $g, g^2, \dots, g^m = e$ of $\langle g \rangle$ are all divisors of m. Given a divisor $d \mid m$, the order of g^k $(1 \leq k \leq m)$ is equal to d if and only if $m/\gcd(k, m) = d$, which is equivalent to $\gcd(k, m) = m/d$, hence to $k = (m/d)k'$, $1 \leq k' \leq d$, $m = (m/d)d$ and $\gcd(k', d) = 1$. There are $\varphi(d)$ possible values of such k'.
(8) For any divisor $d \mid m$, $\langle g^d \rangle = \{g^d, g^{2d}, \dots, (g^d)^{m/d} = g^m = e\}$ is a cyclic subgroup of $\langle g \rangle$ of order m/d. Conversely, if H is a subgroup of $\langle g \rangle$, then $A := \{k \in \mathbf{Z} \mid g^k \in H\}$ is a subgroup of $\mathbf{Z}$ containing $m\mathbf{Z}$; therefore $A = d\mathbf{Z}$ for some $d \in \mathbf{Z}$ satisfying $d \mid m$. □

7.5.7 Summary of properties of cyclic (sub)groups (1) The order of $g \in G$ (finite or infinite) is equal to the order $|\langle g \rangle|$ of the cyclic group generated by g.

(2) A cyclic group G of infinite order has two generators. A choice of a generator $g \in G$ defines a group isomorphism $f : (\mathbf{Z}, +) \xrightarrow{\sim} G$, $f(k) = g^k$. The other generator is g^{-1}.

(3) A cyclic group G of order $m \in \mathbf{N}_+$ has $\varphi(m)$ generators. A choice of a generator $g \in G$ defines a group isomorphism $\overline{f} : (\mathbf{Z}/m\mathbf{Z}, +) \xrightarrow{\sim} G$, $f(k \pmod m) = g^k$. The other generators are g^k for $k \pmod m \in (\mathbf{Z}/m\mathbf{Z})^*$ (i.e., for $\gcd(k, m) = 1$).

(4) For example, $G = \mu_m = (\{z \in \mathbf{C} \mid z^m = 1\}, \cdot)$ is generated by $g = e^{2\pi i/m}$, and the map $\overline{f} : (\mathbf{Z}/m\mathbf{Z}, +) \longrightarrow \mu_m$ given by $k \pmod m \mapsto e^{2\pi ki/m}$ is a group isomorphism.

(5) One can show that every finite abelian group is isomorphic to a product of cyclic groups $(\mathbf{Z}/n_1\mathbf{Z}, +) \times \cdots \times (\mathbf{Z}/n_r\mathbf{Z}, +)$, where $r \geq 0$, $n_1 > 1$ and $n_1 \mid n_2 \mid \cdots \mid n_r$. This can be proved, for example, by using

Euclid's algorithm and elementary operations on matrices with coefficients in $\mathbf{Z}$.

7.5.8 Theorem (Lagrange)**.** *If G is a finite group and $H \subset G$ is a subgroup, then $|H|$ divides $|G|$.*

Proof. See Theorem 7.6.15 and Section 7.6.19 below. □

7.5.9 Corollary (Lagrange)**.** *If G is a finite group and $g \in G$, then the order of g (which is equal to the order of the cyclic subgroup $\langle g \rangle \subset G$) divides $|G|$. Consequently, $g^{|G|} = e$.*

Proof. We are going to prove the corollary in the special case when the group G is **abelian**. Let $m = |G|$, $G = \{g_1, \ldots, g_m\}$. For fixed $g \in G$, the equalities $gg_i = gg_j \iff g_i = g_j$ are equivalent, by Proposition 7.1.6(6). Therefore the elements $gg_1, \ldots, gg_m$ are distinct, which means that $G = \{gg_1, \ldots, gg_m\}$. As the group G is abelian, if we compute the product of its elements in an arbitrary order, then we obtain the same result. Therefore

$$g_1 g_2 \cdots g_m = (gg_1)(gg_2)\cdots(gg_m) = g^m g_1 g_2 \cdots g_m \quad \Longrightarrow \quad e = g^m$$

(again, by Proposition 7.1.6(6)). □

7.5.10 Lagrange's theorem $\Longrightarrow$ Euler's theorem If $G = ((\mathbf{Z}/n\mathbf{Z})^*, \cdot)$, then $|G| = \varphi(n)$ and $g = a \pmod n$ for some $a \in \mathbf{Z}$ such that $\gcd(a, n) = 1$. The statement $g^{|G|} = e$ is then equivalent to $a^{\varphi(n)} \equiv 1 \pmod n$, and the proof given above is an abstract version of the proof of Euler's theorem 4.2.9.

7.6 The quotient group G/H (abelian case)

7.6.1 A preview The goal of this section is to give an abstract version of the construction of the additive group $(\mathbf{Z}/n\mathbf{Z}, +)$.

From now on until the end of Section 7.6 we assume that we are given the following data.

- An abelian group $(G, +)$ written additively (the neutral element is $e = 0$ and the inverse of $a \in G$ is written as $-a$). For $a, b \in G$, we use the notation $a - b := a + (-b) = (-b) + a$; then $-(a - b) = b + (-a) = b - a$ and $(a - b) + (b - c) = a - c$.
- A subgroup $H \subset G$.

We want to define and investigate the notion of congruences modulo H for elements of G. An example we should keep in mind is $G = \mathbf{Z}$ and $H = n\mathbf{Z}$.

7.6.2 Definition. For $a \in G$, let $a + H := \{a + h \mid h \in H\} \subset G$. Subsets of this form will be called **classes modulo H in G.**

7.6.3 Examples (1) If $G = \mathbf{Z}$, $H = n\mathbf{Z}$ and $a \in \mathbf{Z}$, then $a + n\mathbf{Z} = \{b \in \mathbf{Z} \mid a \equiv b \pmod n\}$, as in chapter 3.

(2) If $G = (\mathbf{R}^2, +)$, $0 \neq u \in \mathbf{R}^2$ and $H = (\mathbf{R}u, +)$, then the elements of G correspond to points in a plane, H is a line in the plane passing through the origin, and $a + H$ is the unique line passing through a that is parallel to H. In particular, for any $a, b \in \mathbf{R}^2$, the sets $a + H$ and $b + H$ are either disjoint, or equal. As we are going to show, this property holds in general.

7.6.4 Proposition. *If $a, b \in G$, then*

$$\begin{cases} a + H = b + H, & \text{if } a - b \in H \\ (a + H) \cap (b + H) = \emptyset, & \text{if } a - b \notin H. \end{cases}$$

In addition, the map $H \longrightarrow a + H$, $h \mapsto a + h$ is bijective, with inverse given by $a' \mapsto a' - a$. In particular, if H is finite, then $|a + H| = |H|$.

Proof. If $(a+H)\cap(b+H)\neq\emptyset$, then there exist $h_1,h_2\in H$ such that $a+h_1=b+h_2$, hence $a-b=h_2-h_1\in H$. Conversely, if $h_0:=a-b\in H$, then, for each $h\in H$, $a+h=b+(h+h_0)\in b+H$ and $b+h=a+(h-h_0)\in a+H$. Therefore $a+H\subseteq b+H$ and $b+H\subseteq a+H$, hence $a+H=b+H$. The second part of the proposition is clear, since $(a+h)-a=h$ and $a+(a'-a)=a'$. □

7.6.5 Corollary. *The set G is a disjoint union of various classes modulo H. Denote by G/H the set of these classes (each class taken only once).*

7.6.6 Definition. We say that $a,b\in G$ are **congruent modulo** H (and we write $a\equiv b\pmod H$) if $a-b\in H$. According to Proposition 7.6.4, this is equivalent to $a+H=b+H$. Note that $a+H=\{c\in G\mid c\equiv a\pmod H\}$. In order to simplify the notation, one often uses the notation $\overline{a}\in G/H$ for $a+H\in G/H$.

The **index of H in G**, denoted by $(G:H):=|G/H|$, is the number of classes modulo H in G (for example, $(\mathbf{Z}:n\mathbf{Z})=n$ and $(\mathbf{R}:\mathbf{Z})=\infty$).

7.6.7 Proposition. *Let $a,b,c\in G$.*
(1) $a\equiv a\pmod H$.
(2) If $a\equiv b\pmod H$, then $b\equiv a\pmod H$.
(3) If $a\equiv b\pmod H$ and $b\equiv c\pmod H$, then $a\equiv c\pmod H$.

Proof. This is a straightforward consequence of the fact that $a\equiv b\pmod H$ is equivalent to $a+H=b+H$. □

7.6.8 Proposition. *If $a,b,a',b'\in G$, then*

$$\left.\begin{matrix}a\equiv a'\pmod H\\ b\equiv b'\pmod H\end{matrix}\right\}\implies\left\{\begin{matrix}a+b\equiv a'+b'\pmod H\\ -a\equiv -a'\pmod H\end{matrix}\right.$$

Proof. If $a-a',b-b'\in H$, then $(-a)-(-a')=-(a-a')\in H$ and $(a+b)-(a'+b')=(a-a')+(b-b')\in H$, since H is a subgroup of G and the operation "+" satisfies $x+y=y+x$. □

7.6.9 Towards G/H We are now ready to show that the set G/H of all classes modulo H in G has a natural structure of an abelian group (this will be a generalisation of $(\mathbf{Z}/n\mathbf{Z},+)$).

There is a natural projection map

$$\mathrm{pr}:G\longrightarrow G/H,\qquad a\mapsto a+H=\overline{a}$$

that sends an element $a\in G$ to the unique class modulo H containing a.

7.6.10 Theorem. *Let H be a subgroup of an abelian group $(G,+)$.*
*(1) The set G/H of all classes modulo H in G has a natural structure of an abelian group (the **quotient group** of G by H) such that*

(Operation)	$(a+H)+(b+H)=(a+b)+H$
(Inverse)	$-(a+H)=(-a)+H$
(Neutral element)	$0_{G/H}=0_G+H$

(2) The projection map $\mathrm{pr}:G\longrightarrow G/H$ is a (surjective) group homomorphism.
(3) $\mathrm{Ker}(\mathrm{pr})=H$.

Proof. (1) The operations are **well-defined:** we must check that, if $a+H=a'+H$ and $b+H=b'+H$, then $(a+b)+H=(a'+b')+H$ and $(-a)+H=(-a')+H$; this was proved in Proposition 7.6.8.

Associativity: we must check that, for any $a,b,c\in G$,

$$(a+H)+((b+H)+(c+H)) \stackrel{?}{=} ((a+H)+(b+H))+(c+H)$$

This follows from the fact that the left hand side (resp. the right hand side) is equal to $(a+(b+c))+H$ (resp. to $((a+b)+c)+H$).

Commutativity: we must check that, for any $a,b\in G$,

$$(a+H)+(b+H) \stackrel{?}{=} (b+H)+(a+H)$$

This follows from the fact that the left hand side (resp. the right hand side) is equal to $(a+b)+H$ (resp. to $(b+a)+H$).

Neutral element: we must check that, for any $a\in G$,

$$(a+H)+(0+H) \stackrel{?}{=} a+H \stackrel{?}{=} (0+H)+(a+H)$$

This follows from the fact that the term on the left (resp. on the right) is equal to $(a+0)+H=a+H$ (resp. to $(0+a)+H=a+H$).

Inverse: we must check that, for any $a\in G$,

$$(a+H)+((-a)+H) \stackrel{?}{=} 0+H \stackrel{?}{=} ((-a)+H)+(a+H)$$

This follows from the fact that the term on the left (resp. on the right) is equal to $(a+(-a))+H=0+H$ (resp. to $((-a)+a)+H=0+H$).

(2) For any $a,b\in H$,

$$\mathrm{pr}(a+b)=(a+b)+H=(a+H)+(b+H)=\mathrm{pr}(a)+\mathrm{pr}(b).$$

(3) $a\in\mathrm{Ker}(\mathrm{pr}) \iff a+H=0+H \iff a-0\in H \iff a\in H.$ □

7.6.11 Remark The proof shows that the formulae in (1) are uniquely detemined by (2).

7.6.12 Examples of quotient groups (abelian) (1) If $G=(\mathbf{Z},+)$ and $H=n\mathbf{Z}$, then $(G/H,+)=(\mathbf{Z}/n\mathbf{Z},+)$ is the set of residue classes modulo n.

(2) If $G=(\mathbf{R}^2,+)$ and $H=(\mathbf{R}u,+)$ for fixed $0\neq u\in\mathbf{R}^2$, then H is a line in the plane G and G/H is the set of all lines in G parallel to H. In this case G/H is not just an abelian group, but a vector space over $\mathbf{R}$ (since both G and H are real vector spaces). If one chooses $u'\in\mathbf{R}^2$ such that u and u' are linearly independent over $\mathbf{R}$, then $H'=\mathbf{R}u'$ is a line in G transversal to H, which implies that the composite map

$$H'\hookrightarrow G\xrightarrow{\mathrm{pr}} G/H \tag{7.6.12.1}$$

is bijective (its inverse sends a line $a+H$ parallel to H to the intersection $(a+H)\cap H'$). As (7.6.12.1) is a group morphism, it is a group isomorphism.

In general, if V is a vector space over a field K and $W\subset V$ a vector subspace, then the quotient abelian group $(V/W,+)$ has again a natural structure of a vector space over K (the **quotient space** of V by W) with respect to the scalar multiplication

$$t(v+W)=(tv)+W,\qquad v\in V,\qquad t\in K.$$

To show that, one needs to check the identities

$$(t+t')(v+W) \stackrel{?}{=} t(v+W)+t'(v+W)$$
$$t(t'(v+W)) \stackrel{?}{=} (tt')(v+W)$$
$$t((v+W)+(v'+W)) \stackrel{?}{=} t(v+W)+t(v'+W)$$

$(t,t' \in K,\ v,v' \in V)$, which is easy.

As above, if $W' \subset V$ is a subspace complementary to W (i.e., such that $V = W \oplus W'$), then the composite linear map

$$W' \hookrightarrow V \xrightarrow{\text{pr}} V/W \tag{7.6.12.2}$$

is bijective (its inverse sends $a+W$ to the intersection $(a+W)\cap W'$), hence it is an isomorphism of vector spaces.

(3) $G = (\mathbf{R},+)$, $H = 2\pi\mathbf{Z}$. The quotient $(\mathbf{R}/2\pi\mathbf{Z},+)$ has the following geometric interpretation.

For any $\alpha \in \mathbf{R}$, denote by $R(\alpha)$ the rotation around the origin in $\mathbf{R}^2$ with oriented angle α. These rotations form an abelian group (denoted by $SO(2)$) with respect to composition and satisfy

$$R(\alpha) = R(\beta) \iff \alpha \equiv \beta \pmod{2\pi\mathbf{Z}} \iff \alpha + 2\pi\mathbf{Z} = \beta + 2\pi\mathbf{Z} \tag{7.6.12.3}$$
$$R(\alpha)\circ R(\beta) = R(\alpha+\beta). \tag{7.6.12.4}$$

This is equivalent to the fact that the map

$$\mathbf{R}/2\pi\mathbf{Z} \xrightarrow{\sim} SO(2), \qquad \alpha + 2\pi\mathbf{Z} \mapsto R(\alpha) \tag{7.6.12.5}$$

is a group isomorphism. In other words, we should think of oriented angles in the plane $\mathbf{R}^2$ as classes $\alpha + 2\pi\mathbf{Z} \in \mathbf{R}/2\pi\mathbf{Z}$ of real numbers modulo integral multiples of 2π.

7.6.13 Exercise. Determine the order of $\frac{m}{n} + \mathbf{Z} \in (\mathbf{Q}/\mathbf{Z},+)$, where $m,n \in \mathbf{Z}$, $n \geq 1$ and $\gcd(m,n) = 1$.

7.6.14 Multiplicative notation If we use multiplicative notation for the abelian group $(G,\cdot)$ and its subgroup $H \subset G$, then the classes modulo H in G will be defined as

$$gH := \{gh \mid h \in H\}, \tag{7.6.14.1}$$

for $g \in G$. As before, two classes will be either disjoint, or equal. More precisely, $gH = g'H \iff g^{-1}g' \in H$.

The set G/H of all classes is an abelian group with operation $(gH)(g'H) = (gg')H$, neutral element $eH = H$ and inverse $(gH)^{-1} = g^{-1}H$.

7.6.15 Theorem (Lagrange, abelian case)**.** *If G is a finite abelian group and $H \subset G$ is a subgroup, then $|G| = |H| \cdot |G/H|$. In particular, $|H|$ divides $|G|$.*

Proof. We know that G is a disjoint union of $|G/H|$ sets of the form $a+H$, and that $|a+H| = |H|$ (for any $a \in G$). □

7.6.16 Theorem (Homomorphism theorem, abelian case)**.** *Assume that $(G,*)$ is an abelian group and that $f : (G,*) \longrightarrow (H,\square)$ is a group homomorphism. The map*

$$\begin{array}{ccc} \overline{f} : (G/\operatorname{Ker}(f),*) & \longrightarrow & (\operatorname{Im}(f),\square) \\ g\operatorname{Ker}(f) & \mapsto & f(g) \end{array}$$

is a group isomorphism. [We use here multiplicative notation for G and $G/\operatorname{Ker}(f)$, as in Section 7.6.14.]

Proof. The map $\overline{f}$ is **well-defined:** if $g\,\mathrm{Ker}(f) = g'\,\mathrm{Ker}(f)$, then $g^{-1}g' \in \mathrm{Ker}(f)$, which implies that $f(g') = f(gg^{-1}g') = f(g)f(g^{-1}g') = f(g)\square e_H = f(g)$.
The map $\overline{f}$ is a **group homomorphism:**

$$\overline{f}((g\,\mathrm{Ker}(f))(g'\,\mathrm{Ker}(f))) = \overline{f}(gg'\,\mathrm{Ker}(f)) = f(gg') = f(g)\square f(g') = \overline{f}(g\,\mathrm{Ker}(f))\square\overline{f}(g'\,\mathrm{Ker}(f)).$$

In order to finish the proof, it is enough to check that $\overline{f}$ has trivial kernel, and then apply Corollary 7.4.11.
If $g\,\mathrm{Ker}(f)$ lies in $\mathrm{Ker}(\overline{f})$, then $f(g) = e_H$, which implies that $g \in \mathrm{Ker}(f)$, hence $g\,\mathrm{Ker}(f) = \mathrm{Ker}(f)$ is the neutral element of $G/\mathrm{Ker}(f)$. □

7.6.17 Homomorphism theorem: examples (1) If $\mathrm{Ker}(f) = \{e_G\}$, then we recover Corollary 7.4.11 (for an abelian group G).
(2) The exponential map $\exp : (\mathbf{C}, +) \longrightarrow (\mathbf{C} \smallsetminus \{0\}, \cdot)$ is a group homomorphism satisfying $\mathrm{Ker}(\exp) = (2\pi i\mathbf{Z}, +)$ and $\mathrm{Im}(\exp) = \mathbf{C} \smallsetminus \{0\}$. It gives rise to a group isomorphism

$$(\mathbf{C}/2\pi i\mathbf{Z}, +) \xrightarrow{\sim} (\mathbf{C} \smallsetminus \{0\}, \cdot), \qquad z + 2\pi i\mathbf{Z} \mapsto e^z.$$

(3) Assume that $(G, *)$ is a group and $g \in G$ is an element of finite order $m \in \mathbf{N}_+$. The group homomorphism $f : (\mathbf{Z}, +) \longrightarrow (G, *)$, $f(n) = g^n$ from Section 7.4.3, Example 7 satisfies $\mathrm{Im}(f) = \langle g\rangle$ and $\mathrm{Ker}(f) = (m\mathbf{Z}, +)$. It defines, therefore, a group isomorphism

$$\overline{f} : (\mathbf{Z}/m\mathbf{Z}, +) \xrightarrow{\sim} \langle g\rangle, \qquad k + m\mathbf{Z} \mapsto g^k,$$

which we saw already in Theorem 7.5.6(4).
(4) A slightly modified exponential map $f : (\mathbf{C}, +) \longrightarrow (\mathbf{C} \smallsetminus \{0\}, \cdot)$, $f(z) = e^{2\pi iz}$ is a surjective group homomorphism satisfying $\mathrm{Ker}(f) = (\mathbf{Z}, +)$. It defines group isomorphisms

$$(\mathbf{C}/\mathbf{Z}, +) \xrightarrow{\sim} (\mathbf{C} \smallsetminus \{0\}, \cdot), \qquad (\mathbf{R}/\mathbf{Z}, +) \xrightarrow{\sim} (\{e^{2\pi i\alpha} \mid \alpha \in \mathbf{R}\}, \cdot) = (\{z \in \mathbf{C} \mid |z| = 1\}, \cdot) = U(1),$$
$$(\tfrac{1}{n}\mathbf{Z}/\mathbf{Z}, +) \xrightarrow{\sim} (\{z \in \mathbf{C} \mid z^n = 1\}, \cdot) = \mu_n, \qquad (\mathbf{Q}/\mathbf{Z}, +) \xrightarrow{\sim} \bigcup_{n\geq 1} \mu_n = \{\text{roots of unity in } \mathbf{C}\}.$$

7.6.18 Universal property of G/H Let H be a subgroup of an abelian group G (written multiplicatively). For every group homomorphism $f : G \longrightarrow G'$ such that $H \subset \mathrm{Ker}(f)$ there exists a unique group homomorphism $f' : G/H \longrightarrow G'$ satisfying

$$f = f' \circ \mathrm{pr} : G \longrightarrow G/H \longrightarrow G'.$$

This formula is equivalent to $f'(gH) = f(g)$ for all $g \in G$. The condition $H \subset \mathrm{Ker}(f)$ ensures that $f(g) = f(gh)$ for all $h \in H$, which means that f' is well-defined (it is a group homomorphism, since f is).

7.6.19 What happens if G is not abelian? If $(G, *)$ is an arbitrary group and $H \subset G$ its subgroup, then one must distinguish between the classes (called **left (resp. right) cosets of H in G**)

$$gH := \{gh \mid h \in H\} \subset G, \qquad Hg := \{hg \mid h \in H\} \subset G \qquad (g \in G).$$

As in Proposition 7.6.4, these classes have the following properties.

- $gH \cap g'H \neq \emptyset \iff g^{-1}g' \in H \iff gH = g'H$.
- The map $H \longrightarrow gH$, $h \mapsto gh$ is bijective (with inverse map $gH \longrightarrow H$ given by $g' \mapsto g^{-1}g'$).

Consequently, G is a disjoint union of the classes gH, and each class has the same number of elements (finite or infinite) as H.

If we denote by G/H the set of the classes gH, we obtain, as in the proof of Theorem 7.6.15, that $|G| = |H| \cdot |G/H|$.

Similarly, one denotes by $H\backslash G$ the set of the classes Hg. The map $G/H \longrightarrow H\backslash G$, $gH \mapsto Hg^{-1}$ is bijective. The index of H in G is defined as $(G : H) := |G/H| = |H\backslash G|$.

It is natural to ask whether the set G/H becomes a group when equipped with operation

$$(g_1H)(g_2H) \stackrel{?}{=} (g_1g_2)H. \tag{7.6.19.1}$$

In general, no. The point is that a non-abelian version of Proposition 7.6.8 does not hold, in general, which means that the formula (7.6.19.1) does not give a well-defined operation on G/H. What we need is a validity of the property

$$\left.\begin{matrix} g_1H = g_1'H \\ g_2H = g_2'H \end{matrix}\right\} \stackrel{?}{\implies} (g_1g_2)H = (g_1'g_2')H, \tag{7.6.19.2}$$

which is equivalent to $(g_1g_2)H = (g_1h_1g_2h_2)H$ for all $g_i \in G$ and $h_i \in H$, which is, in turn, equivalent to $g_2H = h_1g_2H$, hence to $H = g_2^{-1}h_1g_2H$, and to $g_2^{-1}h_1g_2 \in H$ (for all $g_2 \in G$ and $h_1 \in H$). In other words, we need

$$\forall g \in G \quad g^{-1}Hg \subseteq H.$$

Applying this condition with g and g^{-1}, we obtain an equivalent condition

$$\forall g \in G \quad g^{-1}Hg = H. \tag{7.6.19.3}$$

Subgroups satisfying (7.6.19.3) are called **normal subgroups of** G (notation: $H \lhd G$). They also have the property $gH = Hg$, which implies that

$$gH = g'H \implies Hg^{-1} = (gH)^{-1} = (g'H)^{-1} = Hg'^{-1} \implies g^{-1}H = g'^{-1}H. \tag{7.6.19.4}$$

It follows from (7.6.19.2) and (7.6.19.4) that, for a normal subgroup $H \lhd G$, G/H is a group with respect to the operation (7.6.19.1). As in Theorem 7.6.10, the projection $\mathrm{pr} : G \longrightarrow G/H$ sending g to gH is a surjective group homomorphism, and $\mathrm{Ker}(\mathrm{pr}) = H$.

Conversely, the kernel of any group homomorphism $f : (G, *) \longrightarrow (H, \square)$ is a normal subgroup of G. Indeed, if $g' \in \mathrm{Ker}(f)$, then $f(g') = e_H$ and $f(g^{-1}g'g) = f(g)^{-1}\square f(g')\square f(g) = f(g)^{-1}\square e_H \square f(g) = f(g)^{-1}\square f(g) = e_H$, hence $g^{-1}g'g \in \mathrm{Ker}(f)$.

Therefore normal subgroups of G are precisely the kernels of group homomorphisms $G \longrightarrow G'$.

The statement (and the proof) of the Homomorphism Theorem 7.6.16 then make sense for any group homomorphism.

The universal property of G/H in Section 7.6.18 holds for any group G and any normal subgroup $H \lhd G$.

8 Rings

8.1 Definition and examples

8.1.1 Example: $A = \mathbf{Z}$ A fundamental example of a (commutative) ring is the set of integers $\mathbf{Z} = \{0, \pm 1, \pm 2, \cdots\}$ equipped with the operations "+" and "$\cdot$". Other examples include $\mathbf{Z}/n\mathbf{Z}$ and matrix rings $M_k(\mathbf{R})$ (non-commutative if $k > 1$).

8.1.2 Definition. A **ring** (with a unit) is a triple $(A, +, \cdot)$, where A is a set equipped with two binary operations $a, b \mapsto a + b$ and $a, b \mapsto a \cdot b = ab$ satisfying the following axioms.

(1) **(Additive structure)** $(A, +)$ is an abelian group with neutral element $0 = 0_A$ and inverse $a \mapsto -a$

(2) **(Associativity)** $\forall a, b, c \in A \quad (a \cdot b) \cdot c = a \cdot (b \cdot c)$

(3) **(Unit)** $\exists 1 = 1_A \in A \; \forall a \in A \quad a \cdot 1 = 1 \cdot a = a$

(4) **(Distributivity)** $\forall a, b, c \in A \quad (a + b) \cdot c = (a \cdot c) + (b \cdot c), \quad a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ (8.1.2.1)

8.1.3 Definition. If, in addition, the following property holds

(5) **(Commutativity)** $\forall a, b \in A \quad a \cdot b = b \cdot a$

then we say that A is a **commutative ring**.

8.1.4 Basic properties

- $\forall a \in A \quad 0 \cdot a = a \cdot 0 = 0$ (since $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$).
- 1 **is unique:** if there exist $1, 1' \in A$ such that $\forall a, b \in A \quad a \cdot 1 = 1 \cdot a = a$ and $b \cdot 1' = 1' \cdot b = b$, then we obtain, by taking $a = 1'$ and $b = 1$, that $1' = 1 \cdot 1' = 1$.
- $0 = 1$ in $A \iff A = \{0\}$ (the **zero ring**). Indeed, if $0 \neq 1$ in A, then $A \neq \{0\}$. Conversely, if $0 = 1$ in A, then $a = a \cdot 1 = a \cdot 0 = 0$ holds, for any $a \in A$.

8.1.5 Examples of rings (1) Commutative rings $\mathbf{Z}$, $\mathbf{Z}/n\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$, $\mathbf{Z} + i\mathbf{Z}$, $\mathbf{Z} + 2i\mathbf{Z}$, $\mathbf{Q} + i\mathbf{Q}$.

(2) Non-commutative matrix rings $M_n(\mathbf{R})$, $M_n(\mathbf{C})$ $(n > 1)$. The unit is the identity matrix $I = I_n$.

(3) More generally, if V is a vector space over a field K, then the set of K-linear endomorphisms of V

$$\mathrm{End}_K(V) := \{\alpha : V \longrightarrow V \mid \alpha \text{ is } K- \text{ linear}\}$$

is a ring with respect to the operations $(\alpha + \beta)(v) = \alpha(v) + \beta(v)$ and $(\beta \circ \alpha)(v) = \beta(\alpha(v))$ $(v \in V)$. The unit is the identity map $\mathrm{id}_V : v \mapsto v$.

Of course, if $V = K^n$, then $\mathrm{End}_K(K^n) = M_n(K)$ (cf. Example (4) in Section 7.2.4).

(4) In fact, if A is any ring (not necessarily commutative), then the set $M_n(A)$ of $n \times n$ matrices with coefficients in A is a ring with respect to the standard matrix sum and product.

(5) If A is a commutative ring, so is the ring of polynomials $A[T] = \{a_0 + a_1T + \cdots + a_dT^d \mid d \geq 0,\ a_i \in A\}$ in one variable with coefficients in A. By induction, we obtain rings $A[T_1, \ldots, T_n]$ of polynomials in several variables.

(6) One sometimes encounters rings without a unit, but such rings usually have "approximate units". Example: $A := \{\alpha \in \mathrm{End}_K(V) \mid \dim_K(\mathrm{Im}(\alpha)) < \infty\}$, where V is a vector space of infinite dimension over a field K.

8.1.6 Definition. Let A be a ring. An element $a \in A$ is **invertible in** A if there exists $b \in A$ such that $ab = ba = 1$. Such an element $b \in A$ is then **unique**; we say that b is **the inverse of** a and we write $b = a^{-1}$. The set of invertible elements of A

$$A^* := \{a \text{ invertible in } A\}$$

is a group with respect to the product (its neutral element is 1). We say that A^* is the **multiplicative group of** A.

8.1.7 Remarks on the inverse (1) Uniqueness of the inverse: if $ab = ba = 1 = ac = ca$, then $b = b \cdot 1 = b(ac) = (ba)c = 1 \cdot c = c$.

(2) $(A^*, \cdot)$ is a group: firstly, $1 \in A^*$, since $1 \cdot 1 = 1$. Secondly, if $a, a' \in A^*$, then $ab = ba = 1 = a'b' = b'a'$ for $b = a^{-1}$, $b' = a'^{-1}$. It follows that $(aa')(b'b) = a(a'b')b = a \cdot 1 \cdot b = ab = 1$ and $(b'b)(aa') = b'(ba)a' = b' \cdot 1 \cdot a' = b'a' = 1$. Therefore $ab \in A^*$ and the inverse $(ab)^{-1}$ of ab is equal to $b'a' = b^{-1}a^{-1}$, as in Proposition 7.1.6(5). Finally, if $a \in A^*$, then $aa^{-1} = a^{-1}a = 1$, which implies that $a^{-1} \in A^*$ and $(a^{-1})^{-1} = a$ (again, as in Proposition 7.1.6(5)).

(3) In fact, the argument in (1) shows that a stronger statement holds: if $b, c \in A$ satisfy $ba = 1 = ac$, then $b = b \cdot 1 = b(ac) = (ba)c = 1 \cdot c = c$. In other words, if $a \in A$ admits both a **left inverse** $b \in A$ and a **right inverse** $c \in A$, then $b = c$, and this element is unique.

8.1.8 Invertible elements (examples) (1) $\mathbf{R}^* = (\mathbf{R} \smallsetminus \{0\}, \cdot)$, $\mathbf{C}^* = (\mathbf{C} \smallsetminus \{0\}, \cdot)$, $\mathbf{Q}^* = (\mathbf{Q} \smallsetminus \{0\}, \cdot)$, $(\mathbf{Q} + \mathbf{Q}i)^* = ((\mathbf{Q} + \mathbf{Q}i) \smallsetminus \{0\}, \cdot)$, $\mathbf{Z}^* = (\{\pm 1\}, \cdot)$, $(\mathbf{Z} + \mathbf{Z}i)^* = (\{\pm 1, \pm i\}, \cdot)$, $(\mathbf{Z} + \mathbf{Z}\sqrt{2})^* = (\{\pm(1 + \sqrt{2})^n \mid n \in \mathbf{Z}\}, \cdot)$, $(\mathbf{Z} + \mathbf{Z}\frac{1+\sqrt{5}}{2})^* = (\{\pm(\frac{1+\sqrt{5}}{2})^n \mid n \in \mathbf{Z}\}, \cdot)$.

(2) For any ring A denote the multiplicative group of the matrix ring $M_n(A)$ by

$$GL_n(A) := M_n(A)^* = \{M \in M_n(A) \mid \exists N \in M_n(A) \quad MN = NM = I_n\}.$$

8.1.9 Proposition. *If A is a commutative ring, then*

$$\begin{aligned} GL_n(A) &= \{M \in M_n(A) \mid \det(M) \in A^*\} \\ &= \{M \in M_n(A) \mid \exists N \in M_n(A) \quad MN = I_n\} \\ &= \{M \in M_n(A) \mid \exists N \in M_n(A) \quad NM = I_n\} \end{aligned}$$

Proof. If $M, N \in M_n(A)$ satisfy $MN = I_n$, then $\det(M)\det(N) = \det(I_n) = 1$, hence $\det(M), \det(N) \in A^*$.
Conversely, if we denote by $\mathrm{adj}(M) \in M_n(A)$ the **adjoint matrix** of M (recall that $(-1)^{i+j}\mathrm{adj}(M)_{ij}$ is defined to be the determinant of the $(n-1) \times (n-1)$ matrix obtained from M by deleting the i-th column and the j-th row), then

$$M \cdot \mathrm{adj}(M) = \mathrm{adj}(M) \cdot M = \det(M) I_n. \tag{8.1.9.1}$$

For example, for $n = 2$,

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \mathrm{adj}(M) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \qquad M \cdot \mathrm{adj}(M) = \mathrm{adj}(M) \cdot M = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}.$$

If $\det(M) \in A^*$, then the formulas (8.1.9.1) imply that the matrix $N := (\det(M))^{-1}\mathrm{adj}(M) \in M_n(A)$ satisfies $MN = NM = I_n$. □

8.1.10 Exercise. Give an example when an element a of a ring A has a left inverse, but not a right inverse (and vice versa).
[Hint: try $A = \mathrm{End}_K(V)$ for an infinite-dimensional vector space V.]

8.1.11 Product of rings Let A_1, A_2 be rings. Their product

$$A = A_1 \times A_2 = \{(a_1, a_2) \mid a_i \in A_i\}$$

is again a ring when equipped with the operations

$$(a_1,a_2)+(b_1,b_2)=(a_1+b_1,a_2+b_2),\qquad (a_1,a_2)\cdot(b_1,b_2)=(a_1\cdot b_1,a_2\cdot b_2),$$
$$0_A=(0_{A_1},0_{A_2}),\qquad 1_A=(1_{A_1},1_{A_2}).$$

Note that

$$A^*=A_1^*\times A_2^*=\{(a_1,a_2)\mid a_i\in A_i^*\},\qquad (a_1,a_2)^{-1}=(a_1^{-1},a_2^{-1}).$$

8.2 Subrings

8.2.1 Definition. A **subring** of a ring A is a subset $B\subset A$ which is a ring with respect to the operations "$+$" and "$\cdot$" in A (this implies that B has the same 0 and 1 as A, that $B^*\subset A^*\cap B$ and that the inverse of any $b\in B^*$ is the same when computed in B or in A).

8.2.2 Example: $\mathbf{Z}\subset\mathbf{C}$ $\mathbf{Z}$ is a subring of $\mathbf{C}$ and $\mathbf{Z}^*=\{\pm1\}\subsetneq\mathbf{Z}\cap\mathbf{C}^*=\mathbf{Z}\smallsetminus\{0\}$.

8.2.3 Proposition. *Let $B\subset A$ be a subset of a ring A. It is equivalent:*
(1) B is a subring of A;
(2) $0_A,1_A\in B$ and $\forall b,b'\in B\quad b-b',bb'\in B$.
[This shows that Definition 8.2.1 is, indeed, a generalisation of the definition of a subring of $\mathbf{C}$ from Definition 1.5.18.]

Proof. The implication (1) $\Longrightarrow$ (2) is automatic. Let us prove that (2) implies (1). We know, by Proposition 7.2.3(3), that (2) implies that $(B,+)$ is an abelian group. The rest is again automatic. □

8.2.4 Examples of subrings of $\mathbf{C}$ Proposition 8.2.3 implies that the following subsets of $\mathbf{C}$ are subrings of $\mathbf{C}$:

$$A_1=\mathbf{Z}+\mathbf{Z}\sqrt{6},\qquad A_2=\mathbf{Q}+\mathbf{Q}\sqrt{6},\qquad A_3=\mathbf{Z}+\mathbf{Z}\sqrt[3]{2}+\mathbf{Z}\sqrt[3]{4},\qquad A_4=\mathbf{Q}+\mathbf{Q}\sqrt[3]{2}+\mathbf{Q}\sqrt[3]{4}.\quad (8.2.4.1)$$

8.2.5 Exercise. (1) Every subring of $\mathbf{C}$ contains $\mathbf{Z}$.
(2) What is the smallest subring of $\mathbf{C}$ containing $2\sqrt{6}$ (resp. $\sqrt[4]{2}$, resp. $\sqrt{6}/2$)?

8.2.6 The centre of a ring The **centre**

$$Z(A):=\{z\in A\mid\forall a\in A\quad za=az\}$$

of a ring A is a subring of A. The ring A is commutative $\iff$ $Z(A)=A$.

8.2.7 Exercise. For any ring A and any $n\geq1$,

$$Z(M_n(A))=Z(A)\cdot I_n=\{a\cdot I_n\mid a\in Z(A)\}.$$

8.2.8 Example: $\mathbf{C}$ as a subring of $M_2(\mathbf{R})$ We can write elements $\mathbf{C}$ as complex numbers $z=x+iy$, or as pairs of real numbers $\begin{pmatrix}x\\y\end{pmatrix}$. This identification of real vector spaces

$$\mathbf{C}\xrightarrow{\sim}\mathbf{R}^2,\qquad z=x+iy\mapsto\begin{pmatrix}x\\y\end{pmatrix}$$

allows us to write every $\mathbf{C}$-linear map $\mathbf{C} \longrightarrow \mathbf{C}$ (i.e., a 1×1 complex matrix) as an $\mathbf{R}$-linear map $\mathbf{R}^2 \longrightarrow \mathbf{R}^2$ (i.e., as a 2×2 real matrix).

Explicitly, multiplication by a fixed complex number $w = a+bi$ of a variable complex number $z = x+yi$ defines an $\mathbf{R}$-linear map

$$z = x + yi \mapsto wz = (a + bi)(x + yi) = (ax - by) + (bx + ay)i \tag{8.2.8.1}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{8.2.8.2}$$

represented by the matrix $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

The correponding matrix representation of complex numbers

$$M : \mathbf{C} \longrightarrow M_2(\mathbf{R}), \qquad M(a + bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{8.2.8.3}$$

has the following properties.

- $\forall a \in \mathbf{R} \quad M(a) = a \cdot I_2$ (in particular, $M(1) = I_2$);
- $M(w + w') = M(w) + M(w')$ (since $(w + w')z = wz + w'z$);
- $M(ww') = M(w)M(w')$ (since $(ww')z = (ww')z$);
- $M(w) = 0 \iff w = 0$;
- more generally, $M(w) = M(w') \iff w = w'$.

These properties imply that we can consider $\mathbf{C}$ as a subring of $M_2(\mathbf{R})$ by identifying $w \in \mathbf{C}$ with the matrix $M(w)$. More precisely, in the language of Definition 8.4.1 below, M is an injective ring homomorphism.

In abstract terms, we have written $\mathbf{C} = M_1(\mathbf{C}) = \mathrm{End}_{\mathbf{C}}(\mathbf{C})$ as a subring of $\mathrm{End}_{\mathbf{R}}(\mathbf{C}) = \mathrm{End}_{\mathbf{R}}(\mathbf{R}^2) = M_2(\mathbf{R})$.

8.2.9 Example continued If we restrict M to the set of invertible elements of $\mathbf{C}$, we obtain an injective group homomorphism

$$M : \mathbf{C}^* \hookrightarrow GL_2(\mathbf{R}).$$

Note that, for each $\alpha \in \mathbf{R}$, the matrix

$$M(e^{i\alpha}) = r(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

represents the rotation of $\mathbf{R}^2$ around the origin with oriented angle α. We obtain, therefore, four different algebraic incarnations of the group of such rotations:

$$\mathbf{R}/2\pi\mathbf{Z}, \qquad U(1) = \{z \in \mathbf{C} \mid z\overline{z} = 1\}, \qquad SO(2) = \{a \in M_2(\mathbf{R}) \mid {}^tAA = I_2\}, \qquad \{r(\alpha) \mid \alpha \in \mathbf{R}\}.$$

They are related by the following explicit group isomorphisms.

$$\mathbf{R}/2\pi\mathbf{Z} \xrightarrow{\sim} U(1), \qquad \alpha + 2\pi\mathbf{Z} \mapsto e^{i\alpha}; \qquad M : U(1) \xrightarrow{\sim} \{r(\alpha) \mid \alpha \in \mathbf{R}\} = SO(2).$$

8.3 Integral domains, fields

8.3.1 Examples (1) A product of non-zero complex numbers is always non-zero.

(2) In the ring $A = \mathbf{Z}/6\mathbf{Z}$, the elements $a = 2 \pmod 6 \in \mathbf{Z}/6\mathbf{Z} \smallsetminus \{0 \pmod 6\}$ and $b = 3 \pmod 6 \in \mathbf{Z}/6\mathbf{Z} \smallsetminus \{0 \pmod 6\}$ are non-zero, but their product is equal to zero: $ab = 6 \pmod 6 = 0 \pmod 6$.

8.3.2 Definition. Let A be a **commutative** ring A.
(1) A is an **integral domain** if $A \neq \{0\}$ and if a product of non-zero elements is non-zero:

$$\forall a, b \in A \smallsetminus \{0\} \quad ab \neq 0.$$

(2) A is a **field** if $A \neq \{0\}$ and if every non-zero element is invertible:

$$A \smallsetminus \{0\} = A^*.$$

(3) A **subfield** of a field K is a subring of K which is itself a field.

8.3.3 Examples and Remarks (1) $\mathbf{Q}, \mathbf{R}, \mathbf{C}$ and $\mathbf{Q} + \mathbf{Q}i$ are fields (in fact, subfields of $\mathbf{C}$), whereas $\mathbf{Z}$ and $\mathbf{Z} + \mathbf{Z}i$ are not (but they are integral domains).

(2) There exist non-commutative rings $A \neq \{0\}$ satisfying $A \smallsetminus \{0\} = A^*$ (for example, the ring $\mathbf{H}$ of Hamilton quaternions). Such rings are called **skew-fields** or **division algebras**.

(3) If A is a ring and if $a, b \in A$ satisfy $ab = 0$ and $a \in A^*$, then $b = a^{-1}ab = a^{-1} \cdot 0 = 0$. This implies that every field is an integral domain.

(4) Every subring of an integral domain (in particular, every subring of a field) is again an integral domain.

(5) Conversely, one can show that every integral domain A is a subring of a suitably minimal field K constructed in terms of fractions whose numerators and denominators are contained in A (example: $\mathbf{Z} \subset \mathbf{Q}$). Informally,

$$K = \left\{ \frac{a}{b} \mid a, b \in A,\, b \neq 0 \right\}, \qquad \frac{a}{b} = \frac{c}{d} \iff ad - bc = 0 \in A, \quad \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd},$$
$$0_K = \frac{0}{1}, \qquad 1_K = \frac{1}{1}$$

(but one must show that these conditions do not lead to a contradiction, that the ring axioms (8.1.2.1) are satisfied by K, and that K is a field, with inverse $(\frac{a}{b})^{-1} = \frac{b}{a}$). We say that K is the **fraction field of** A (or the **quotient field of** A).

8.3.4 Exercise. Which among the rings A_j in (8.2.4.1) are fields?

8.3.5 Proposition. *Let $n \geq 1$ be an integer. The following conditions are equivalent:*
(1) $\mathbf{Z}/n\mathbf{Z}$ is a field.
(2) $\mathbf{Z}/n\mathbf{Z}$ is an integral domain.
(3) $n = p$ is a prime number.

Proof. (1) $\Longrightarrow$ (2): this is automatic.
(2) $\Longrightarrow$ (3): we follow Example 2 of Section 8.3.1. If $n \neq p$ is not a prime, then either $n = 1$ (when $\mathbf{Z}/n\mathbf{Z} = \{0\}$), or $n = kl$ for some integers $1 < k, l < n$. The residue classes $a = k \pmod n$ and $b = l \pmod n$ are non-zero elements of $\mathbf{Z}/n\mathbf{Z}$, but their product is zero: $ab = kl \pmod n = 0 \pmod n \in \mathbf{Z}/n\mathbf{Z}$.
(3) $\Longrightarrow$ (1): if $n = p$ is a prime, then $\varphi(p) = p - 1 > 0$, which implies that $(\mathbf{Z}/p\mathbf{Z})^* = \mathbf{Z}/p\mathbf{Z} \smallsetminus \{0\} \neq \emptyset$. $\square$

8.3.6 Finite integral domains are fields The non-trivial implication (2) $\Longrightarrow$ (1) in Proposition 8.3.5 can also be deduced from the following abstract statement.

8.3.7 Proposition. *If $A \neq \{0\}$ is a commutative ring with finitely many elements, then it is equivalent:*
(1) A is a field.
(2) A is an integral domain.

Proof. The implication (1) $\Longrightarrow$ (2) is automatic. Conversely, assume that (2) holds. This means that, for any $a \in A \smallsetminus \{0\}$, the multiplication map

$$m_a : A \longrightarrow A, \qquad b \mapsto ab$$

is injective. However, an injective map between two finite sets of the same cardinality must be bijective. Therefore there exists $b \in A$ such that $1 = m_a(b) = ab$, which means that a is invertible and $b = a^{-1}$. $\square$

8.3.8 Integral domains of finite dimension (over a field) are fields Proposition 8.3.7 has the following linear algebra analogue.

8.3.9 Exercise. Assume that A is a commutative ring containing a subring K such that K is a field. Show that:
(1) A is a K-vector space (with respect to the multiplication $A \times A \longrightarrow A$ restricted to $K \times A \longrightarrow A$).
(2) If the dimension of A as a K-vector space is finite, then it is equivalent: A is a field $\iff$ A is an integral domain.
(3) Solve Exercise 8.3.4 using (2).

8.3.10 Divisibility Let A be a commutative ring. Divisibility in A is defined in the usual way: if $a, b \in A$, we say that b **divides** a (notation: $b \mid a$) if there exists $c \in A$ such that $a = bc$. This is equivalent to $bA \supseteq aA$.

The properties of divisibility listed in Section 1.1.5 hold in this generality, with the following modifications:

- $b \mid 1 \iff b \in A^*$
- if $u \in A^*$, then it is equivalent: $b \mid a \iff b \mid au$
- if A is an integral domain and if $a, b \in A \smallsetminus \{0\}$ satisfy $b \mid a$ and $a \mid b$, then $b = au$ for some $u \in A^*$.

8.3.11 Irreducible elements Let A be an integral domain. An element $a \in A$ is **irreducible in** A if it has the following properties:

- $a \neq 0$
- $a \notin A^*$
- a is not a non-trivial product: if $a = bc$ for $b, c \in A$, then $b \in A^*$ or $c \in A^*$ (but not simultaneously, since $a \notin A^*$).

Note: if a is irreducible in A and $u \in A^*$, then au is irreducible, too.

Example: $\{\text{irreducible elements in } \mathbf{Z}\} = \{\pm p \mid p \text{ prime}\}$.

8.4 Ring homomorphisms

8.4.1 Definition. Let A, B be rings. A map $f : A \longrightarrow B$ is a **ring homomorphism** if it satisfies the following conditions.
(1) $\forall a, a' \in A \quad f(a + a') = f(a) + f(a')$.
(2) $\forall a, a' \in A \quad f(aa') = f(a)f(a')$.
(3) $f(1_A) = 1_B$.

8.4.2 Remarks and examples (1) Condition (1) in Definition 8.4.1 implies that f defines a homomorphism of additive groups $f : (A, +) \longrightarrow (B, +)$ (hence $f(0_A) = 0_B$ and $f(-a) = -f(a)$ for all $a \in A$). Similarly, conditions (2) and (3) imply that $f(A^*) \subset B^*$ and that f defines a homomorphism of multiplicative groups $f : (A^*, \cdot) \longrightarrow (B^*, \cdot)$ (in particular, if $a \in A^*$, then $f(a) \in B^*$ and $f(a)^{-1} = f(a^{-1})$).

(2) The projection $\mathrm{pr} : \mathbf{Z} \longrightarrow \mathbf{Z}/n\mathbf{Z}$, $\mathrm{pr}(a) = a \pmod n$ is a ring homomorphism.

(3) Similarly, the map $\mathbf{Z}/mn\mathbf{Z} \longrightarrow \mathbf{Z}/n\mathbf{Z}$ given by $a \pmod{mn} \mapsto a \pmod n$ is a ring homomorphism.

(4) Let A_1, A_2 be rings. The projections $A_1 \xleftarrow{\mathrm{pr}_1} A_1 \times A_2 \xrightarrow{\mathrm{pr}_2} A_2$, $\mathrm{pr}_j(a_1, a_2) = a_j$ are ring homomorphisms.

(5) The inclusion of a subring $B \hookrightarrow A$ is a ring homomorphism.

(6) **Warning:** conditions (1) and (2) in Definition 8.4.1 **do not imply** (3), in general. For example, the map $f : \mathbf{Z}/6\mathbf{Z} \longrightarrow \mathbf{Z}/6\mathbf{Z}$ given by $f(a \pmod 6) := 3a \pmod 6$ satisfies $3(a + a') \equiv 3a + 3a' \pmod 6$ and $3(aa') \equiv (3a)(3a') \pmod 6$ (since $3 \equiv 3^2 \pmod 6$), but $3 \cdot 1 \not\equiv 1 \pmod 6$.

(7) Similarly, the inclusions $A_1 \xrightarrow{i_1} A_1 \times A_2 \xleftarrow{i_2} A_2$ given by $i_1(a_1) = (a_1, 0)$ and $i_2(a_2) = (0, a_2)$ satisfy conditions (1) and (2) in Definition 8.4.1, but not (3) (if $A_1, A_2 \neq \{0\}$).

(8) For any ring A there is a unique ring homomorphism $f : \mathbf{Z} \longrightarrow A$. Indeed, such a homomorphism must satisfy $f(0) = 0_A$, $f(1) = 1_A$, $f(2) = f(1+1) = f(1)+f(1) = 1_A+1_A$, $f(3) = f(2)+f(1) = 1_A+1_A+1_A$, $f(-1) = -f(1) = -1_A$, $f(-2) = -(1_A + 1_A)$ etc. In other words, if we define, for $m \in \mathbf{N}_+$ and $a \in A$,

$$m \cdot a := \underbrace{a + \cdots + a}_{m \text{ times}}, \qquad (-m) \cdot a := -\underbrace{a + \cdots + a}_{m \text{ times}}, \qquad 0 \cdot a := 0_A,$$

then f must be given by the formula $f(n) = n \cdot 1_A$ ($n \in \mathbf{Z}$), which proves the uniqueness of f. Conversely, Proposition 7.3.2 for $G = (A, +)$ and $g = 1_A$ tells us that $f(m) + f(n) = f(m + n)$. The multiplicativity property $f(mn) = f(m)f(n)$ is a consequence of the distributivity rule 8.1.2.1(4). Finally, $f(1) = 1_A$.

Note that $f(\mathbf{Z}) = \{n \cdot 1_A \mid n \in \mathbf{Z}\} \subset Z(A)$ is contained in the centre $Z(A)$ of A.

(9) If $f : A \longrightarrow B$ and $g : B \longrightarrow C$ are ring homomorphisms, so is their composition $g \circ f : A \longrightarrow B \longrightarrow C$.

(10) If $f_i : A_i \longrightarrow B_i$ ($i = 1, 2$) are ring homomorphisms, so is $f_1 \times f_2 : A_1 \times A_2 \longrightarrow B_1 \times B_2$.

(11) If p is a prime number and if A is a commutative ring such that $p \cdot 1_A = 0_A$, then the map $\varphi : A \longrightarrow A$ given by $\varphi(a) = a^p$ is a ring homomorphism (called the **Frobenius morphism**). Indeed, $\varphi(1_A) = 1_A$, $\varphi(ab) = (ab)^p = a^p b^p = \varphi(a)\varphi(b)$ and

$$\varphi(a + b) = (a + b)^p = a^p + b^p + \sum_{j=1}^{p-1} \binom{p}{j} a^j b^{p-j} = a^p + b^p = \varphi(a) + \varphi(b),$$

since $\binom{p}{j} \in p\mathbf{Z}$ for $1 \leq j \leq p - 1$. For each $n \geq 1$, the n-fold composition $\varphi \circ \cdots \circ \varphi : A \longrightarrow A$ is given by the formula $a \mapsto \varphi_q(a) = a^q$, where $q = p^n$. It is also a ring homomorphism.

8.4.3 Proposition. *If a ring homomorphism $f : A \longrightarrow B$ is bijective, then its inverse $f^{-1} : B \longrightarrow A$ is also a ring homomorphism. We say that f is a **ring isomorphism** (which implies that f^{-1} is a ring isomorphism, too).*

Proof. Given $b, b' \in B$, let $a = f^{-1}(b), a' = f^{-1}(b') \in A$. The identities $f(a+a') = f(a)+f(a') = b+b'$ and $f(aa') = f(a)f(a') = bb'$ imply that $f^{-1}(b+b') = a+a' = f^{-1}(b)+f^{-1}(b')$ and $f^{-1}(bb') = aa' = f^{-1}(b)f^{-1}(b')$. Finally, $f^{-1}(1_B) = f^{-1}(f(1_A)) = 1_A$. Therefore $f^{-1} : B \longrightarrow A$ is a ring homomorphism. □

8.4.4 Example: the Chinese Remainder Theorem If $m, n \geq 1$ satisfy $\gcd(m,n) = 1$, then the bijective map in the Chinese Remainder Theorem

$$f : \mathbf{Z}/mn\mathbf{Z} \longrightarrow \mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z}, \qquad f(a \pmod{mn}) = (a \pmod{m}, a \pmod{n})$$

is a bijective ring homomorphism, hence a ring isomorphism.

8.4.5 Exercise. Explain, using the Chinese Remainder Theorem, why Example 6 in Section 8.4.2 is a special case of Example 7.

8.4.6 Definition. The **kernel** and the **image** of a ring homomorphism $f : A \longrightarrow B$ are defined, respectively, as

$$\mathrm{Ker}(f) := \{a \in A \mid f(a) = 0\} \subset A, \qquad \mathrm{Im}(f) := \{f(a) \mid a \in A\} \subset B.$$

8.4.7 Proposition. *If $f : A \longrightarrow B$ is a ring homomorphism, then:*
(1) $\mathrm{Im}(f)$ *is a subring of* B.
(2) If f is injective, then it defines a ring isomorphism $A \xrightarrow{\sim} \mathrm{Im}(f)$.
(3) For $a, a' \in A$, the following are equivalent: $f(a) = f(a') \iff a' - a \in \mathrm{Ker}(f)$.
(4) f is injective $\iff \mathrm{Ker}(f) = \{0\}$.

Proof. (1) We know that $f(1_A) = 1_B$ and $f(0_A) = 0_B$; thus $0_B, 1_B \in \mathrm{Im}(f)$. If $b, b' \in \mathrm{Im}(f)$, then $b = f(a)$ and $b' = f(a')$ for some $a, a' \in A$. Consequently, $b - b' = f(a-a') \in \mathrm{Im}(f)$ and $bb' = f(aa') \in \mathrm{Im}(f)$. This shows that $\mathrm{Im}(f)$ is a subring of B, by Proposition 8.2.3.
(2) This follows from the definitions.
(3) This is a special case of Proposition 7.4.6(3), for $G = (A, +)$ and $H = (B, +)$. We can also argue directly: $f(a) = f(a') \iff 0 = f(a') - f(a) = f(a'-a) \iff a' - a \in \mathrm{Ker}(f)$.
(4) This is a special case of Proposition 7.4.10 for $G = (A, +)$ and $H = (B, +)$ (or a direct consequence of (3)). □

8.4.8 Corollary. *If a ring homomorphism $f : A \longrightarrow B$ satisfies* $\mathrm{Ker}(f) = \{0\}$*, then it is injective, hence it defines a ring isomorphism* $f : A \xrightarrow{\sim} \mathrm{Im}(f)$.

8.4.9 $\mathrm{Ker}(f), \mathrm{Im}(f)$**: Examples** (1) Inclusion of a subring $B \subset A$: $\mathrm{Ker} = \{0\}$, $\mathrm{Im} = B$.
(2) Projection $\mathrm{pr} : \mathbf{Z} \longrightarrow \mathbf{Z}/n\mathbf{Z}$: $\mathrm{Ker}(\mathrm{pr}) = n\mathbf{Z}$, $\mathrm{Im}(\mathrm{pr}) = \mathbf{Z}/n\mathbf{Z}$.
(3) Projection $\mathbf{Z}/mn\mathbf{Z} \longrightarrow \mathbf{Z}/n\mathbf{Z}$: $\mathrm{Ker} = n\mathbf{Z}/mn\mathbf{Z}$, $\mathrm{Im}(\mathrm{pr}) = \mathbf{Z}/n\mathbf{Z}$.
(4) Projections $\mathrm{pr}_j : A_1 \times A_2 \longrightarrow A_j$ $(\mathrm{pr}(a_1, a_2) = a_j)$: $\mathrm{Im}(\mathrm{pr}_j) = A_j$, $\mathrm{Ker}(\mathrm{pr}_1) = \{0\} \times A_2$, $\mathrm{Ker}(\mathrm{pr}_2) = A_1 \times \{0\}$.

8.4.10 Exercise. (Decomposing a ring into a product of rings) (1) In the notation of Section 8.4.2, Example 4, the elements

$$e_1 := i_1(1_{A_1}) = (1_{A_1}, 0_{A_2}) = (1,0) \in A_1 \times A_2, \qquad e_2 := i_2(1_{A_2}) = (0_{A_1}, 1_{A_2}) = (0,1) \in A_1 \times A_2$$

have the following properties (they form a **full system of orthogonal central idempotents in** $A_1 \times A_2$):

(e_1, e_2 are central)	$e_1, e_2 \in Z(A_1 \times A_2)$
(e_1, e_2 are idempotents)	$e_1^2 = e_1,\ e_2^2 = e_2$
(e_1, e_2 are orthogonal)	$e_1 e_2 = e_2 e_1 = 0$
	$e_1 + e_2 = 1$

(2) Conversely, if A is a ring and $e_1, e_2 \in A$ have the properties listed in (1), then the subset $A_j := e_j A = Ae_j = \{e_j a = a e_j \mid a \in A\} \subset A$ equipped with the operations "+" and "·" from A is a ring with unit e_j $(j = 1, 2)$, and the map

$$A \longrightarrow A_1 \times A_2, \qquad a \mapsto (e_1 a, e_2 a)$$

is a ring isomorphism.

8.4.11 Exercise. Let A be a commutative ring. We know that there is a unique ring homomorphism $\mathbf{Z} \longrightarrow A$. What can one say about ring homomorphisms $\mathbf{Z} \times \mathbf{Z} \longrightarrow A$?

8.4.12 Exercise. Let $(G, +)$ and $(H, +)$ be abelian groups. Show that:
(1) The set $\mathrm{Hom}_{Ab}(G, H) := \{\text{group homomorphisms } f : G \longrightarrow H\}$ is an abelian group with respect to the operation $(f_1 + f_2)(g) := f_1(g) + f_2(g)$.
(2) The set $\mathrm{End}_{Ab}(G) := \mathrm{Hom}_{Ab}(G, G)$ is a ring, with product given by the composition of maps.
(3) $\mathrm{End}_{Ab}((\mathbf{Z}, +)) = \mathbf{Z}$ and $\mathrm{End}_{Ab}((\mathbf{Z}^2, +)) = M_2(\mathbf{Z})$.

8.5 The quotient ring A/I

8.5.1 A preview The goal of this section is to give an abstract version of the construction of the ring $(\mathbf{Z}/n\mathbf{Z}, +, \cdot)$.

8.5.2 Multiplication of congruences From now on until the end of Section 8.5 we assume that

- A is a ring;
- $I \subset (A, +)$ is an additive subgroup.

As in Section 7.6, we write, for $a, b \in A$, $a \equiv b \pmod I$ if and only if $a - b \in I$ (which is equivalent to $a + I = b + I$).

It was shown in Proposition 7.6.8 that congruences can be added and subtracted: for any $a, b, a', b' \in A$,

$$\left.\begin{matrix} a \equiv a' \pmod I \\ b \equiv b' \pmod I \end{matrix}\right\} \implies a \pm b \equiv a' \pm b' \pmod I.$$

Our goal is to understand under what conditions one can multiply congruences $\pmod I$. In other words, when does the following implication hold:

$$\left.\begin{matrix} a \equiv a' \pmod I \\ b \equiv b' \pmod I \end{matrix}\right\} \overset{?}{\implies} ab \equiv a'b' \pmod I. \tag{8.5.2.1}$$

As in the proof of Theorem 7.6.10, this property would imply that the set $A/I = \{a + I \mid a \in A\}$ of residue classes $\pmod I$ in A has a natural structure of a ring, for which the projection map $\mathrm{pr} : A \longrightarrow A/I$ $(\mathrm{pr}(a) = a + I)$ is a ring homomorphism.

8.5.3 Multiplication of congruences: examples (1) If $A = \mathbf{Z}$ and $I = n\mathbf{Z}$, then (8.5.2.1) holds.
(2) If $A = \mathbf{R}$ and $I = 2\pi\mathbf{Z}$, then (8.5.2.1) **does not hold**. Indeed, we have, for each $a \in \mathbf{R} \smallsetminus \mathbf{Z}$,

$$\begin{gathered} a \equiv a \pmod{2\pi\mathbf{Z}} \\ 0 \equiv 2\pi \pmod{2\pi\mathbf{Z}} \\ \underbrace{a \cdot 0}_{0} \not\equiv \underbrace{a \cdot 2\pi}_{2\pi a} \pmod{2\pi\mathbf{Z}} \end{gathered}$$

This means that **one cannot multiply two angles** $\alpha, \beta \in \mathbf{R}/2\pi\mathbf{Z}$.

8.5.4 Theorem. *An additive subgroup $I \subset (A, +)$ of a ring A satisfies* (8.5.2.1) *if and only if $AI \subset I$ and $IA \subset I$; in other words, if*

$$\forall a \in A \ \forall x \in I \quad ax \in I,\ xa \in I.$$

Such an additive subgroup is called a ***(bilateral) ideal*** *of A.*

Proof. Let $a \in A$ and $x \in I$. If (8.5.2.1) holds, then

$$\left.\begin{matrix} a \equiv a \pmod{I} \\ 0 \equiv x \pmod{I} \end{matrix}\right\} \implies a \cdot 0 \equiv a \cdot x \pmod{I} \implies ax \in I \tag{8.5.4.1}$$

and

$$\left.\begin{matrix} 0 \equiv x \pmod{I} \\ a \equiv a \pmod{I} \end{matrix}\right\} \implies 0 \cdot a \equiv x \cdot a \pmod{I} \implies xa \in I. \tag{8.5.4.2}$$

Conversely, if $ax \in I$ and $xa \in I$ whenever $a \in A$ and $x \in I$, then

$$\left.\begin{matrix} a \equiv a' \pmod{I} \\ b \equiv b' \pmod{I} \end{matrix}\right\} \implies a - a', b - b' \in I \implies ab - a'b' = (a - a')b + a'(b - b') \in IA + AI \subset I + I \subset I. \tag{8.5.4.3}$$

□

8.5.5 Examples of (bilateral) ideals (1) Both $\{0\}$ and A are (bilateral) ideals of A.
(2) A non-empty subset $I \subset A$ is a (bilateral) ideal if (and only if) it satisfies

$$\forall x, y \in I \ \forall a \in A \quad x + y \in I,\ ax \in I,\ xa \in I.$$

Indeed, taking $a = 0_A$ resp. $a = -1_A$ we obtain that $0_A \in I$ and $-x \in I$ if $x \in I$. Therefore I is a subgroup of $(A, +)$, by Proposition 7.2.3(3).
(3) If I, J are (bilateral) ideals of A, so are $I+J := \{x+y \mid x \in I,\ y \in J\}$, $I \cap J$ and $IJ := \{x_1y_1 + \cdots + x_ry_r \mid r \geq 0,\ x_i \in I,\ y_i \in J\}$.
(4) The kernel $\mathrm{Ker}(f) = \{a \in A \mid f(a) = 0_B\}$ of any ring homomorphism $f : A \longrightarrow B$ is a (bilateral) ideal of A. Indeed, $\mathrm{Ker}(f)$ is a subgroup of $(A, +)$, and if $a \in A$ and $x \in \mathrm{Ker}(f)$, then $f(x) = 0$ and $f(ax) = f(a)f(x) = f(a) \cdot 0 = 0$, which means that $ax \in \mathrm{Ker}(f)$ (similarly, $f(xa) = f(x)f(a) = 0 \cdot f(a) = 0$, hence $xa \in \mathrm{Ker}(f)$).
(5) Conversely, Theorem 8.5.9 below says that any (bilateral) ideal $I \subset A$ is the kernel of a (surjective) ring homomorphism $A \longrightarrow A/I$. Consequently,

$$\{\text{(bilateral) ideals of } A\} = \{\text{kernels of ring homomorphisms } A \longrightarrow B\} = \\ = \{\text{kernels of surjective ring homomorphisms } A \longrightarrow B\}$$

(6) If $f : A \longrightarrow B$ is a ring homomorphism and $J \subset B$ is a (bilateral) ideal of B, then $I := f^{-1}(J) = \{a \in A \mid f(a) \in J\}$ is a (bilateral) ideal of A.

(7) If the ring A is commutative, than the conditions $ax \in I$ and $xa \in I$ in Theorem 8.5.4 are equivalent. In this case we drop the adjective "bilateral" and simply say that I is an ideal of A.

(8) Assume that the ring A is commutative. The simplest example of an ideal $I \subset A$ is the **principal ideal generated by** $x \in A$:

$$(x) := xA = Ax = \{ax \mid a \in A\}$$

consisting of the multiples of x. Note that $(x) = (ux)$, for any invertible element $u \in A^*$.

More generally, for any $x_1, \ldots, x_r \in A$, the subset

$$(x_1, \ldots, x_r) := (x_1) + \cdots + (x_r) = Ax_1 + \cdots + Ax_r = \{a_1x_1 + \cdots + a_rx_r \mid a_i \in A\}$$

is an ideal of A, called **the ideal generated by** $x_1, \ldots, x_r$ (it is contained in any ideal $I \subset A$ containing $x_1, \ldots, x_r$).

(9) A subset $I \subset \mathbf{Z}$ is an ideal $\iff$ $I \subset (\mathbf{Z}, +)$ is an additive subgroup ($\iff$ $I = d\mathbf{Z} = (d) = (-d)$ for some $d \in \mathbf{N}$). In other words, every ideal of $\mathbf{Z}$ is principal.

(10) If A is commutative and if $I \subset A$ is an ideal containing an invertible element $u \in A^*$, then I contains $u \cdot (u^{-1}a) = a$, for all $a \in A$; therefore $I = A = (1)$.

(11) In particular, if $A = K$ is a field, then the only ideals of K are $(0) = \{0\}$ and $(1) = K$, since every non-zero element of K is invertible.

(12) A subset $I \subset \mathbf{Z}/n\mathbf{Z}$ is an ideal $\iff$ $I = d\mathbf{Z}/n\mathbf{Z}$ for some $d \mid n$.

(13) If A is not commutative, then there are two weaker notions of ideals in A. A subgroup $I \subset (A, +)$ of the additive group of A is a **left ideal** (resp. a **right ideal**) of A if $ax \in I$ (resp. $xa \in I$) holds, for all $a \in A$ and $x \in I$. This means that I is a bilateral ideal $\iff$ it is simultaneously a left and a right ideal.

Example: let $A = M_n(K)$, where K is a field. For any K-vector subspace $W \subset K^n$, let

$$I_W := \{M \in M_n(K) \mid \text{all columns of } M \text{ lie in } W\}, \quad {}^tI_W := \{M \in M_n(K) \mid \text{all columns of } {}^tM \text{ lie in } W\}.$$

One can show that

$$\{\text{right ideals of } M_n(K)\} = \{I_W \mid W \subset K^n\}, \qquad \{\text{left ideals of } M_n(K)\} = \{{}^tI_W \mid W \subset K^n\}, \\ \{\text{bilateral ideals of } M_n(K)\} = \{\{0\}, M_n(K)\}.$$

8.5.6 Exercise. For any $m, n \in \mathbf{Z} \smallsetminus \{0\}$,

$$(m) + (n) = (\gcd(m, n)), \qquad (m)(n) = (mn), \qquad (m) \cap (n) = (\operatorname{lcm}(m, n)).$$

8.5.7 Exercise. Assume that K is a field and $B \neq \{0\}$ is a ring. Show that any ring homomorphism $f : K \longrightarrow B$ is injective. What can one say about ring homomorphisms $f : M_n(K) \longrightarrow B$ for $n > 1$?

8.5.8 Exercise. Let A be a commutative ring.
(1) An element $a \in A$ is called **nilpotent** if there exists an integer $n \geq 1$ such that $a^n = 0$. Show that the **nilradical of** A $\mathrm{Nil}(A) := \{a \in A \mid a \text{ is nilpotent}\}$ is an ideal of A.
(2) If A is an integral domain, then $\mathrm{Nil}(A) = \{0\} = (0)$.
(3) Determine $\mathrm{Nil}(\mathbf{Z}/6\mathbf{Z})$, $\mathrm{Nil}(\mathbf{Z}/12\mathbf{Z})$ and $\mathrm{Nil}(\mathbf{Z}/n\mathbf{Z})$, for any $n \in \mathbf{N}_+$.
(4) If $f : A \longrightarrow B$ is a ring homomorphism (and both A and B are commutative), then $f(\mathrm{Nil}(A)) \subset \mathrm{Nil}(B)$. In particular, if B is an integral domain, then $\mathrm{Nil}(A) \subset \mathrm{Ker}(f)$.
(5) More generally, if $I \subset A$ is any ideal, show that $\sqrt{I} := \{a \in A \mid \exists n \geq 1 \; a^n \in I\}$ (the **radical of** I) is an ideal of A (note that $\sqrt{(0)} = \mathrm{Nil}(A)$).

8.5.9 Theorem. *If $I \subset A$ is a (bilateral) ideal of a ring A, then the set $A/I = \{a + I \mid a \in A\}$ of residue classes $(\mathrm{mod}\, I)$ in A is a ring with respect to the operations*

$$\begin{aligned}
&\textbf{\textit{(Sum)}} && (a+I)+(b+I) = (a+b)+I \\
&\textbf{\textit{(Product)}} && (a+I)\cdot(b+I) = ab+I \\
&\textbf{\textit{(Zero)}} && 0_{A/I} = 0_A + I \\
&\textbf{\textit{(Unit)}} && 1_{A/I} = 1_A + I
\end{aligned}$$

(2) The projection map $\mathrm{pr} : A \longrightarrow A/I$ *(*$\mathrm{pr}(a) = a + I$*) is a (surjective) ring homomorphism.*
(3) $\mathrm{Ker}(\mathrm{pr}) = I$.

Proof. (1) The operations "+" and "·" on A/I are well-defined, by Proposition 7.6.8 and Theorem 8.5.4, respectively. We need to check the ring axioms (8.1.2.1). Property (1) was proved in Theorem 7.6.10. Associativity follows from

$$\begin{aligned}
((a+I)\cdot(b+I))\cdot(c+I) &= (ab+I)\cdot(c+I) = (ab)c+I, \\
(a+I)\cdot((b+I)\cdot(c+I)) &= (a+I)\cdot(bc+I) = a(bc)+I,
\end{aligned}$$

the unit axiom from

$$\begin{aligned}
(a+I)\cdot(1+I) &= (a\cdot 1+I) = (a+I), \\
(1+I)\cdot(a+I) &= (1\cdot a+I) = (a+I),
\end{aligned}$$

and distributivity from

$$\begin{aligned}
((a+I)+(b+I))\cdot(c+I) &= ((a+b)+I)\cdot(c+I) = (a+b)c+I, \\
(a+I)\cdot(c+I)+(b+I)\cdot(c+I) &= (ac+I)+(bc+I) = (ac+bc)+I
\end{aligned}$$

(and a similar formula involving $a(b+c)$).
Part (2) is automatic (in fact, the definition of the operations in A/I was forced upon us by the the requirement (2)), and Part (3) was proved in Theorem 7.6.10. □

8.5.10 Remarks (1) If $I = \{0\}$, then $A/I = A$.
(2) If $I = A$, then $A/I = \{0\}$.
(3) If the ring A is commutative, so is A/I.

8.5.11 Invertible elements of A/I (commutative case) Let $I \subset A$ be an ideal of a commutative ring A, let $a \in A$. As in the proof of Theorem 3.4.2, it is equivalent:

$$\begin{aligned} a \pmod I = a + I \in A/I \text{ is invertible in } A/I &\iff \exists u \in A \quad au \equiv 1 \pmod I \\ &\iff \exists u \in A \ \ \exists y \in I \quad au + y = 1 \\ &\iff 1 \in aA + I = (a) + I \\ &\iff (a) + I = A. \end{aligned} \tag{8.5.11.1}$$

In the special case when $I = (b) = bA$ is a principal ideal, then

$$a \pmod b = a + (b) \in A/(b) \text{ is invertible in } A/(b) \iff 1 \in aA + bA \iff aA + bA = A. \tag{8.5.11.2}$$

Explicitly, if $u, v \in A$ satisfy $au + bv = 1$, then $au \equiv 1 \pmod b$ and $u \pmod b$ is the inverse of $a \pmod b$ in $A/(b)$.

8.5.12 Theorem (Homomorphism theorem)**.** *Let $f : A \longrightarrow B$ be a ring homomorphism. The map*

$$\begin{array}{ccc} \overline{f} : A/\operatorname{Ker}(f) & \longrightarrow & \operatorname{Im}(f) \\ a + \operatorname{Ker}(f) & \mapsto & f(a) \end{array}$$

is a ring isomorphism.

Proof. According to Theorem 7.6.16, the map $\overline{f}$ is well-defined, bijective and satisfies $\overline{f}(x + y) = \overline{f}(x) + \overline{f}(y)$. It remains to show that it satifies the remaining properties (2) and (3) in Definition 8.4.1, but this is automatic: $\overline{f}(1_A + \operatorname{Ker}(f)) = f(1_A) = 1_B$ and

$$\forall a, b \in A \quad \overline{f}((a+\operatorname{Ker}(f))(b+\operatorname{Ker}(f))) = \overline{f}(ab+\operatorname{Ker}(f)) = f(ab) = f(a)f(b) = \overline{f}(a+\operatorname{Ker}(f))\overline{f}(b+\operatorname{Ker}(f)).$$

□

8.5.13 Reformulation Theorem 8.5.12 says that any ring homomorphism $f : A \longrightarrow B$ can be written in a natural way as a composition

$$f : A \xrightarrow{\text{pr}} A/\operatorname{Ker}(f) \xrightarrow{\overline{f}} \operatorname{Im}(f) \xrightarrow{i} B, \tag{8.5.13.1}$$

where pr is the projection on a quotient ring, $\overline{f}$ is an isomorphism and i is the inclusion of a subring.

There are two important special cases (the first of which we already know):

$$\begin{aligned} f \text{ is injective} &\iff \text{pr} = \text{id} \iff \overline{f} : A \xrightarrow{\sim} \operatorname{Im}(f), \\ f \text{ is surjective} &\iff i = \text{id} \iff \overline{f} : A/\operatorname{Ker}(f) \xrightarrow{\sim} B. \end{aligned}$$

There is a similar decomposition

$$f : G \xrightarrow{\text{pr}} G/\operatorname{Ker}(f) \xrightarrow{\overline{f}} \operatorname{Im}(f) \xrightarrow{i} H \tag{8.5.13.2}$$

of any group homomorphism $f : G \longrightarrow H$.

8.5.14 Exercise. Let I be a (bilateral) ideal of a ring A.
(1) If $\overline{J} \subset A/I$ is a (bilateral) ideal of A/I, then $J := \text{pr}^{-1}(\overline{J}) = \{a \in A \mid a \pmod I \in \overline{J}\}$ is a (bilateral) ideal of A containing I, and $\overline{J} = J/I$.
(2) Conversely, if $J \supset I$ is a (bilateral) ideal of A, then J/I is a (bilateral) ideal of A/I and $J = \text{pr}^{-1}(J/I)$.

8.5.15 Exercise. Let $I \subset A$ be an ideal of a commutative ring A. Show that $\mathrm{Nil}(A/I) = \sqrt{I}/I$, in the language of Exercise 8.5.8.

8.5.16 The characteristic of a ring Let A be a ring. We know that there is a unique ring homomorphism $f : \mathbf{Z} \longrightarrow A$, namely, $f(n) = n \cdot 1_A$. The image of f is contained in the centre $Z(A)$ of A.

Case 1. $\mathrm{Ker}(f) = 0$. We say that A has **characteristic zero**. In this case f is injective and we can consider $\mathbf{Z}$ as a subring of A if we identify $n \in \mathbf{Z}$ with its image $f(n) = n \cdot 1_A$.

Case 2. $\mathrm{Ker}(f) \neq 0$. In this case $\mathrm{Ker}(f)$ is a non-zero ideal of $\mathbf{Z}$, hence $\mathrm{Ker}(f) = m\mathbf{Z}$ for a (unique) integer $m \geq 1$. We say that A has **characteristic** m.

Note that $m = 1$ if and only if $f(1) = 0$, which is equivalent to $A = \{0\}$. Therefore $m > 1$ if $A \neq \{0\}$.

According to Theorem 8.5.12, f defines a ring isomorphism $\overline{f} : \mathbf{Z}/m\mathbf{Z} \xrightarrow{\sim} \mathrm{Im}(f)$, which means that we can consider $\mathbf{Z}/m\mathbf{Z}$ as a subring of A if we identify $n \pmod m = n + m\mathbf{Z} \in \mathbf{Z}/m\mathbf{Z}$ with $f(n) = n \cdot 1_A$.

8.5.17 The characteristic of a field If $A = K$ is a field, then one can say more.

If K has characteristic zero, then $\mathbf{Z}$ is a subring of K, via $n \mapsto n \cdot 1_K$ $(n \in \mathbf{Z})$. If $n \in \mathbf{Z} \smallsetminus \{0\}$, then $n \cdot 1_K \neq 0_K$, which means that $n \cdot 1_K$ is invertible in K. The injective ring homomorphism

$$\mathbf{Z} \hookrightarrow K, \qquad n \mapsto n \cdot 1_K$$

extends uniquely to a ring homomorphism (necessarily injective, by Exercise 8.5.7)

$$\mathbf{Q} \hookrightarrow K, \qquad \frac{m}{n} = mn^{-1} \mapsto (m \cdot 1_K)(n \cdot 1_K)^{-1}. \tag{8.5.17.1}$$

Therefore there is a unique ring homomorphism $\mathbf{Q} \longrightarrow K$, given by the formula (8.5.17.1). This homomorphism is injective and $\mathbf{Q}$ can be identified with its image, which is a subfield of K.

If K has characteristic $m > 0$, then $\mathbf{Z}/m\mathbf{Z}$ is a subring of K, via $n \pmod m \mapsto n \cdot 1_K$ $(n \in \mathbf{Z})$. The field K is an integral domain, which means that $\mathbf{Z}/m\mathbf{Z}$ is an integral domain, too. This implies that $m = p$ is a **prime number**, by Proposition 8.3.5. The ring $\mathbf{Z}/p\mathbf{Z}$ is then a field, often denoted by $\mathbf{F}_p$.

For future reference, let us summarise the previous discussion in a formal proposition.

8.5.18 Proposition. *Let K be a field.*
(1) If K has characteristic zero, then K contains $\mathbf{Q}$ as a subfield, via the map $\frac{m}{n} = mn^{-1} \mapsto (m \cdot 1_K)(n \cdot 1_K)^{-1}$.
(2) If K has non-zero characteristic $m \geq 1$, then $m = p$ is a prime number and K contains $\mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$ as a subfield, via the map $n \pmod p \mapsto n \cdot 1_K$.

9 Polynomial rings A[X]

Throughout Section 9, A is a commutative ring.

9.1 Definition and basic properties of $A[X]$

9.1.1 Informal definition of $A[X]$ Polynomials in one variable (say, X) with coefficients in A are formal expressions

$$a = a(X) = a_0 + a_1X + \cdots + a_mX^m = a_mX^m + \cdots + a_1X + a_0 \qquad (m \geq 0,\ a_i \in A).$$

By definition, a polynomial does not change if one adds terms with zero coefficients. For example:

$$a_0 + a_1X + \cdots + a_mX^m = a_0 + a_1X + \cdots + a_mX^m + 0 \cdot X^{m+1} + 0 \cdot X^{m+2}.$$

It is convenient to add all subsequent terms with zero coefficients $a_{m+1} = a_{m+2} = \cdots = 0$, and consider the polynomial $a = a(X)$ to be an expression

$$a = a(X) = \sum_{i=0}^{\infty} a_i X^i, \qquad a_i \in A, \qquad \text{only finitely many } a_i \text{ are non-zero.}$$

Such a polynomial can then be identified with its sequence of coefficients

$$(a_0, a_1, a_2, \ldots) = (a_i)_{i \in \mathbf{N}}, \qquad a_i \in A, \qquad \text{only finitely many } a_i \text{ are non-zero.}$$

Given another polynomial

$$b = b(X) = \sum_{i=0}^{\infty} b_i X^i, \qquad b_i \in A, \qquad \text{only finitely many } b_i \text{ are non-zero,}$$

one can compute the sum and the product of a and b by the usual formal computation:

$$\begin{aligned} a + b &= (a_0 + a_1 X + a_2 X^2 + \cdots) + (b_0 + b_1 X + b_2 X^2 + \cdots) = \\ &= (a_0 + b_0) + (a_1 + b_1) X + (a_2 + b_2) X^2 + \cdots \\ ab &= (a_0 + a_1 X + a_2 X^2 + \cdots)(b_0 + b_1 X + b_2 X^2 + \cdots) = \\ &= (a_0 b_0) + (a_0 b_1 + a_1 b_0) X + (a_0 b_2 + a_1 b_1 + a_2 b_0) X^2 + \cdots \end{aligned}$$

In terms of the coefficients,

$$\begin{aligned} (a_0, a_1, a_2, \ldots) + (b_0, b_1, b_2, \ldots) &= (a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots) \\ (a_0, a_1, a_2, \ldots) \cdot (b_0, b_1, b_2, \ldots) &= (c_0, c_1, c_2, \ldots), \qquad c_k = \sum_{i+j=k} a_i b_j. \end{aligned} \tag{9.1.1.1}$$

One uses these formulas to give a formal definition of the ring of polynomials.

9.1.2 Definition. The polynomial ring $A[X]$ in one variable X with coefficients in a commutative ring A is defined as follows. As a set,

$$A[X] = \{a = (a_0, a_1, a_2, \ldots) \mid a_i \in A, \text{ only finitely many } a_i \text{ are non-zero}\}.$$

Addition and multiplication in $A[X]$ are defined by the formulas (9.1.1.1). With these operations, $A[X]$ becomes a commutative ring, with zero $0 = (0, 0, 0, \ldots)$ and unit $1 = (1, 0, 0, \ldots)$.

9.1.3 Remarks on $A[X]$ (1) We leave it to the reader to check that the operations (9.1.1.1) on $A[X]$ are well-defined and satisfy the ring axioms 8.1.2.1.

(2) If one omits the condition "only finitely many a_i are non-zero" in the definition, one obtains the ring

$$A[[X]] = \{\sum_{i=0}^{\infty} a_i X^i \mid a_i \in A\}$$

of formal power series with coefficients on A, of which $A[X]$ is a subring.

(3) $A \subset A[X]$ is a subring of $A[X]$, with $\alpha \in A$ corresponding to the constant polynomial $(\alpha, 0, 0, \ldots) = \alpha + 0 \cdot X + 0 \cdot X^2 + \cdots$.

(4) For each polynomial $a \in A[X]$ there exists an integer $m \geq 0$ such that $a_i = 0$ for all $i > m$. One then writes

$$a = a(X) = a_0 + a_1X + \cdots + a_mX^m = a_mX^m + \cdots + a_1X + a_0.$$

If $a \neq 0$, then there is a unique such $m \geq 0$ satisfying $a_m \neq 0$. We define the **degree of** a to be $\deg(a) := m$. In particular, $a \in A \smallsetminus \{0\} \iff \deg(a) = 0$.

The degree of the zero polynomial is defined as $\deg(0) := -\infty$. With this convention, $\deg(ab) = \deg(a) + \deg(b) = -\infty$ if $a = 0$ or $b = 0$.

(5) A polynomial $a = a_mX^m + \cdots + a_1X + a_0$ of degree $m \geq 0$ is called **monic** if $a_m = 1$.

9.1.4 Example: $\deg(ab) \neq \deg(a) + \deg(b)$ Consider the polynomial ring $(\mathbf{Z}/4\mathbf{Z})[X]$. For any $m \in \mathbf{Z}$, denote the residue class $m \pmod 4 \in \mathbf{Z}/4\mathbf{Z}$ by $\overline{m}$. The polynomial $a = \overline{1} + \overline{2}X \in (\mathbf{Z}/4\mathbf{Z})[X]$ has the following properties (since $\overline{2} + \overline{2} = \overline{0} = \overline{2} \cdot \overline{2}$):

$$\deg(a) = 1, \qquad a^2 = \overline{1}^2 + (\overline{2} + \overline{2})X + \overline{2}^2X^2 = \overline{1}, \qquad \deg(a^2) = 0, \qquad a = a^{-1} \in (\mathbf{Z}/4\mathbf{Z})[X])^*.$$

9.1.5 Proposition. *Let $a, b \in A[X] \smallsetminus \{0\}$. Write $a = a_mX^m + \cdots + a_0$ and $b = b_nX^n + \cdots + b_0$, where $m = \deg(a) \geq 0$ and $n = \deg(b) \geq 0$.*
(1) $\deg(a + b) \leq \max(\deg(a), \deg(b))$.
(2) $\deg(ab) \leq \deg(a) + \deg(b)$.
(3) If $a_m \in A^*$, then $\deg(ab) = \deg(a) + \deg(b)$.
(4) If A is an integral domain, then $\deg(ab) = \deg(a) + \deg(b)$.

Proof. (1) If $i > m$ and $i > n$, then $a_i = b_i = 0$, hence $a_i + b_i = 0$.
(2) The equality $ab = a_mb_nX^{m+n} + \cdots + a_0b_0$ implies that $\deg(ab) \leq m + n$, with equality equivalent to $a_mb_n \neq 0$.
(3) The assumptions $a_m \in A^*$ and $b_n \neq 0$ imply that $a_mb_n \neq 0$ (see Section 8.3.3, Remark 3).
(4) The assumptions $a_m, b_n \neq 0$ imply that $a_mb_n \neq 0$, since A is an integral domain. □

9.1.6 Corollary. *If A is an integral domain, so is $A[X]$, and $A[X]^* = A^*$.*

Proof. If $a, b \in A[X] \smallsetminus \{0\}$, then $ab \neq 0$, by Proposition 9.1.5(4) (or its proof). Therefore $A[X]$ is an integral domain. The inclusion $A^* \subset A[X]^*$ is automatic. Conversely, if $a, b \in A[X]$ satisfy $ab = 1$, then $0 = \deg(1) = \deg(ab) = \deg(a) + \deg(b)$, hence $\deg(a) = \deg(b) = 0$, which implies that $a, b \in A \smallsetminus \{0\}$. As $ab = 1$, both a and b lie in A^*. □

9.1.7 Exercise. (1) $(A_1 \times A_2)[X] = A_1[X] \times A_2[X]$.
(2) If $A = A_1 \times \cdots \times A_r$ and if each $A_1, \ldots, A_r$ is an integral domain, then $A[X]^* = A^*$.
(3) If $n = p_1 \cdots p_r$ is a product of distinct primes, then $((\mathbf{Z}/n\mathbf{Z})[X])^* = (\mathbf{Z}/n\mathbf{Z})^*$.
(4) What can one say about $((\mathbf{Z}/n\mathbf{Z})[X])^*$ for general $n \geq 1$?

9.1.8 Exercise. Assume that $a = a_0 + a_1X + \cdots + a_mX^m \in A[X]$, $a_0 \in A^*$ and that each $a_1, \ldots, a_m \in A$ is nilpotent in A (see Exercise 8.5.8). Show that $a \in (A[X])^*$.
[Hint: consider $(a_0 - a)^n$ for a sufficiently large integer $n > 0$.]

9.2 Roots of polynomials

9.2.1 Evaluation morphisms Assume that $a = a_m X^m + \cdots + a_0 \in A[X]$, that A is a subring of a commutative ring B, and that $\beta \in B$. The **value of** a **at** β is defined as

$$a(\beta) := a_m \beta^m + \cdots + a_0 = \sum_{i=0}^{m} a_i \beta^i \in B.$$

If $a(\beta) = 0$, we say that β **is a root of** a.

It is useful to change the traditional point of view, according to which $a \in A[X]$ is fixed and $\beta \in B$ is variable, and consider instead $\beta \in B$ as being fixed and $a \in A[X]$ as being variable.

9.2.2 Proposition. *The map "evaluation at* β*"*

$$\mathrm{ev}_\beta : A[X] \longrightarrow B, \qquad a \mapsto a(\beta)$$

is a ring homomorphism satisfying $\mathrm{ev}_\beta(a) = a$*, for all* $a \in A$*.*

Proof. We need to check that

$$(a+b)(\beta) = a(\beta) + b(\beta), \qquad (ab)(\beta) = a(\beta)b(\beta), \qquad 1(\beta) = 1.$$

The only non-trivial statement is the one in the middle, which follows from

$$(ab)(\beta) = \sum_{k\geq 0} \Big(\sum_{i+j=k} a_i b_j \beta^k \Big) = \Big(\sum_{i\geq 0} a_i \beta^i \Big) \Big(\sum_{j\geq 0} b_j \beta^j \Big) = a(\beta)b(\beta).$$

□

9.2.3 Characterisation of roots When is $\beta \in A$ a root of a polynomial $a = \sum_{k=0}^{m} a_k X^k \in A[X]$? We are going to show that the answer is the same as in the case $A = \mathbf{C}$, namely

$$a(\beta) = 0 \iff (X-\beta) \mid a \iff a \equiv 0 \pmod{(X-\beta)}. \tag{9.2.3.1}$$

Indeed, if $a(X) = (X-\beta)b(X)$ for some $b \in A[X]$, then $a(\beta) = (\beta-\beta)b(\beta) = 0$, by Proposition 9.2.2 (in particular, $a \notin A \smallsetminus \{0\}$). In the opposite direction, the formulas

$$X^k - \beta^k = (X-\beta)(X^{k-1} + \beta X^{k-2} + \cdots + \beta^{k-1}) \qquad (k \geq 1)$$

imply that the difference

$$a - a(\beta) = a(X) - a(\beta) = (X-\beta) \sum_{k=1}^{m} a_k (X^{k-1} + \beta X^{k-2} + \cdots + \beta^{k-1}) \tag{9.2.3.2}$$

is divisible by $(X-\beta)$ in $A[X]$, hence

$$a(X) \equiv a(\beta) \pmod{(X-\beta)}. \tag{9.2.3.3}$$

In particular, if $a(\beta) = 0$, then $(X-\beta) \mid a$. The discussion above can be reformulated in more abstract terms as follows.

9.2.4 Proposition. *For any* $\beta \in A$*, the evaluation morphism* $\mathrm{ev}_\beta : A[X] \longrightarrow A$ *satisfies* $\mathrm{Ker}(\mathrm{ev}_\beta) = (X-\beta)$*, hence induces a ring isomorphism*

$$\overline{\mathrm{ev}}_\beta : A[X]/(X-\beta) \xrightarrow{\sim} A, \qquad a \pmod{(X-\beta)} \mapsto a(\beta).$$

Its inverse is given by the composition $A \hookrightarrow A[X] \xrightarrow{\mathrm{pr}} A[X]/(X-\beta)$*. In other words, the set of residue classes in* $A[X]$ *modulo* $(X-\beta)$ *is equal to* $\{\alpha \pmod{(X-\beta)} \mid \alpha \in A\}$*, with* $\alpha_1 \not\equiv \alpha_2 \pmod{(X-\beta)}$ *if* $\alpha_1 \neq \alpha_2$*.*

9.2.5 Taylor's expansion of a polynomial The congruence (9.2.3.2) can be refined to a congruence modulo $(X-\beta)^2$ as follows. The **derivative** of $a = a(X) = \sum_{k=0}^{m} a_k X^k$ is defined as

$$a' = a'(X) := \sum_{k=1}^{m} k a_k X^{k-1} \in A[X], \qquad k a_k := \underbrace{a_k + \cdots + a_k}_{k \text{ times}} \in A.$$

If we combine (9.2.3.2) with (9.2.3.3), we obtain that

$$a - a(\beta) - (X-\beta)a'(\beta) \tag{9.2.5.1}$$

is divisible by $(X-\beta)^2$ in $A[X]$, hence

$$a(X) \equiv a(\beta) + (X-\beta)a'(\beta) \pmod{(X-\beta)^2} \equiv a(\beta) + (X-\beta)a'(X) \pmod{(X-\beta)^2}. \tag{9.2.5.2}$$

What about congruences modulo higher powers of $(X-\beta)$?

9.2.6 Exercise. (Taylor's formula) For each $n \geq 0$,

$$a(X) \equiv \sum_{k=0}^{n} (D_k a)(\beta)\,(X-\beta)^k \pmod{(X-\beta)^{n+1}}, \tag{9.2.6.1}$$

where $D_k a = \sum_{i \geq k} \binom{i}{k} a_i X^{i-k} \in A[X]$.
[Note that, if $A = \mathbf{C}$, then $k!(D_k a)(X) = (d/dX)^k a(X)$ is the k-th derivative of $a(X)$.]

9.2.7 Theorem. *If A is an integral domain, then a non-zero polynomial $a \in A[X] \smallsetminus \{0\}$ has at most* $\deg(a)$ *distinct roots in A.*

Proof. We proceed as in the proof of Theorem 5.1.14, by induction on $d := \deg(a)$. If $d = 0$, then $a \in A \smallsetminus \{0\}$ and there is no root. Assume that $d > 0$ and that the result holds for polynomials of degree $\deg < d$. If $\alpha \in A$ is a root of a, then $a = (X-\alpha)b$ for some polynomial $b \in A[X] \smallsetminus \{0\}$ of degree $\deg(b) = d-1$. If $\beta \in A$, $\beta \neq \alpha$ and $a(\beta) = 0$, then $0 = a(\beta) = (\beta - \alpha)b(\beta)$. By assumption, A is an integral domain and $\beta - \alpha \neq 0$, which implies that $b(\beta) = 0$. There are at most $\deg(b) = d-1$ such values of β, by induction hypothesis. Together with α, they give at most $(d-1)+1 = d$ roots of a in A. □

9.2.8 Remark We saw in Section 5.1.13 examples of polynomials $a \in (\mathbf{Z}/n\mathbf{Z})[X]$ of degree $\deg(a) = 2$ with (at least) four roots in $\mathbf{Z}/n\mathbf{Z}$ (for $n = 8$ or $n = p_1 \cdots p_r$, where $r \geq 2$ and $p_i \neq 2$ are distinct primes).

9.2.9 Exercise. For each prime $p \neq 2$ give an example of a monic polynomial $a \in (\mathbf{Z}/p^2\mathbf{Z})[X]$ of degree $\deg(a) < p^2$ satisfying $\forall \alpha \in \mathbf{Z}/p^2\mathbf{Z} \quad a(\alpha) = 0$.

9.3 Division with remainder in $A[X]$

9.3.1 A preview Residue classes $\pmod n$ in $\mathbf{Z}$ correspond to remainders of division by $n \geq 1$ in $\mathbf{Z}$. In order to understand residue classes $\pmod b$ in $A[X]$ we need to understand remainders of division by $b \in A[X]$ in $A[X]$.

9.3.2 The quotient ring $A[X]/(b)$ For any $b \in A[X]$ one can consider congruences in $A[X]$ modulo b (more precisely, modulo the principal ideal $(b) = bA[X]$ generated by b), defined in the usual way:

$$a \equiv \widetilde{a} \pmod b \iff b \mid (a - \widetilde{a}) \iff \exists c \in A[X] \quad a - \widetilde{a} = bc.$$

The residue classes $(\mathrm{mod}\, b)$ form a commutative ring $A[X]/(b) = A[X]/bA[X]$ with respect to the usual operations: if we denote, for any polynomial $a \in A[X]$, by $\overline{a} = a + (b) = a \ (\mathrm{mod}\, b)$ its image in $A[X]/(b)$ (i.e., the residue class of a modulo b), then

$$a \ (\mathrm{mod}\, b) \pm \widetilde{a} \ (\mathrm{mod}\, b) = (a \pm \widetilde{a}) \ (\mathrm{mod}\, b), \qquad (a \ (\mathrm{mod}\, b)) \cdot (\widetilde{a} \ (\mathrm{mod}\, b)) = a\widetilde{a} \ (\mathrm{mod}\, b).$$

Furthermore,

$$a \ (\mathrm{mod}\, b) \text{ is invertible in } A[X]/(b) \iff 1 \in aA[X] + bA[X] \iff aA[X] + bA[X] = A[X], \quad (9.3.2.1)$$

by (8.5.11.2). If $u, v \in A[X]$ satisfy $au + bv = 1$, then $au \equiv 1 \ (\mathrm{mod}\, b)$ and $u \ (\mathrm{mod}\, b)$ is the inverse of $a \ (\mathrm{mod}\, b)$ in $A[X]/(b)$.

There is a **distinguished element** of $A[X]/(b)$, given by the residue class of the variable X. This class $\overline{X} = X \ (\mathrm{mod}\, b) \in A[X]/(b)$ satisfies the polynomial equation

$$b(\overline{X}) = 0 \in A[X]/(b), \quad (9.3.2.2)$$

since $b(\overline{X}) = \overline{b(X)}$ and $b(X) \equiv 0 \ (\mathrm{mod}\, b)$.

In the simplest non-trivial case $b = X - \beta$ ($\beta \in A$) considered in Proposition 9.2.4, the quotient ring $A[X]/(X - \beta)$ is identified with A via the evaluation morphism $\overline{\mathrm{ev}}_\beta$, and $\overline{X}$ corresponds to the value $\mathrm{ev}_\beta(X) = X(\beta) = \beta \in A$.

The main goal of Section 9.3 is to describe the quotient ring $A[X]/(b)$ for polynomials b whose leading coefficient is invertible in A (this condition is automatically satisfied if $A = K$ is a field). This will be done using division with remainder for polynomials.

9.3.3 Division with remainder (examples) (1) Division of $a = a(X) = 2X^3 + 2X^2 - X + 1 \in \mathbf{Q}[X]$ by $b = b(X) = 2X + 3$. We compute, consecutively,

$$\begin{aligned}
\boxed{2X^3} + 2X^2 - X + 1 &= (2X+3)X^2 + (-X^2 - X + 1),\\
\boxed{-X^2} - X + 1 &= (2X+3)(-\tfrac{1}{2}X) + (\tfrac{1}{2}X + 1),\\
\boxed{\tfrac{1}{2}X} + 1 &= (2X+3) \cdot \tfrac{1}{4} + \tfrac{1}{4},\\
2X^3 + 2X^2 - X + 1 &= (X^2 - \tfrac{1}{2}X + \tfrac{1}{4})(2X+3) + \tfrac{1}{4}.
\end{aligned}$$

These calculations can be performed purely mechanically, by manipulating the coefficients of the polynomials involved:

	X^3	X^2	X	1
b			2	3
a	2	2	-1	1
$X^2 b$	2	3		
$a - X^2 b$		-1	-1	1
$-\frac{1}{2}Xb$		-1	$-\frac{3}{2}$	
$a - (X^2 - \frac{1}{2}X)b$			$\frac{1}{2}$	1
$\frac{1}{4}b$			$\frac{1}{2}$	$\frac{3}{4}$
$a - (X^2 - \frac{1}{2}X + \frac{1}{4})b$				$\frac{1}{4}$

(2) Division of $a = a(X) = X^3 - 2X^2 - 7X + 3 \in \mathbf{Q}[X]$ by $b = b(X) = 2X^2 + 4X - 1$. In this case

$$\begin{aligned} \boxed{X^3} - 2X^2 - 7X + 3 &= (2X^2 + 4X - 1)(\tfrac{1}{2}X) + (-4X^2 - \tfrac{13}{2}X + 3), \\ \boxed{-4X^2} - \tfrac{13}{2}X + 3 &= (2X^2 + 4X - 1)(-2) + (\tfrac{3}{2}X + 1), \\ 2X^3 - 2X^2 - 7X + 3 &= (\tfrac{1}{2}X - 2)(2X^2 + 4X - 1) + (\tfrac{3}{2}X + 1). \end{aligned}$$

Alternatively,

	X^3	X^2	X	1
b		2	4	-1
a	1	-2	-7	3
$\frac{1}{2}Xb$	1	2	$-\frac{1}{2}$	
$a - \frac{1}{2}Xb$		-4	$-\frac{13}{2}$	3
$-2b$		-4	-8	2
$a - (\frac{1}{2}X - 2)b$			$\frac{3}{2}$	1

(3) Division of $a = \sum_{k=0}^{m} a_k X^k \in A[X]$ by $b = X - \beta$ ($\beta \in A$) is equivalent to the formula (9.2.3.2).

9.3.4 Proposition. *Assume that $b = b_n X^n + \cdots + b_0 \in A[X]$, $\deg(b) = n \geq 0$ and $b_n \in A^*$. Then, for each $a \in A[X]$, there exists a unique pair $q, r \in A[X]$ such that*

$$a = bq + r, \qquad \deg(r) < \deg(b).$$

Moreover, $b \mid a$ in $A[X]$ iff and only if $r = 0$.

Proof. If $n = 0$, then $b = b_0 \in A^*$ and $q = b_0^{-1}a$, $r = 0$. Assume that $n > 0$.

Uniqueness: if $a = bq + r = b\widetilde{q} + \widetilde{r}$ and $\deg(r), \deg(\widetilde{r}) < \deg(b)$, then $b(q - \widetilde{q}) = \widetilde{r} - r$, which implies that

$$\deg(b) > \deg(\widetilde{r} - r) = \deg(b(q - \widetilde{q})) = \deg(b) + \deg(q - \widetilde{q})$$

(we have used here the assumption $b_n \in A^*$ and Proposition 9.1.5(3)). Therefore $\deg(q - \widetilde{q}) < 0$, which means that $q - \widetilde{q} = 0$, hence $q = \widetilde{q}$ and $r = \widetilde{r}$.

Existence: induction on $m = \deg(a)$, $a = a_m X^m + \cdots + a_0$. If $m < n$, then we take $q = 0$, $r = a$. If $m \geq n$, then

$$a = (a_m b_n^{-1} X^{m-n})b + c, \qquad c \in A[X], \qquad \deg(c) < m.$$

By induction hypothesis, there exist $\widetilde{q}, \widetilde{r} \in A[X]$ such that $c = b\widetilde{q} + \widetilde{r}$ and $\deg(\widetilde{r}) < \deg(b)$, which implies that

$$a = b(a_m b_n^{-1} X^{m-n} + \widetilde{q}) + \widetilde{r}, \qquad \deg(\widetilde{r}) < \deg(b).$$

Divisibility: if $r = 0$, then $a = bq$ is divisible by b in $A[X]$. Conversely, if $a = bc$ for some $c \in A[X]$, then $a = cb + 0$ and $\deg(0) = -\infty < \deg(b)$; therefore $r = 0$ by uniqueness. □

9.3.5 Corollary. *Let $A \subset \widetilde{A}$ be a subring of a commutative ring $\widetilde{A}$, let $a \in A[X]$. Assume that $b = b_n X^n + \cdots + b_0 \in A[X]$, $\deg(b) = n \geq 0$ and $b_n \in A^*$. It is equivalent:*

$$b \mid a \text{ in } A[X] \iff b \mid a \text{ in } \widetilde{A}[X].$$

[Warning: the assumption $b_n \in A^*$ cannot be omitted (2 divides 1 in $\mathbf{C}[X]$, but not in $\mathbf{Z}[X]$)].

Proof. As $b_n \in A^* \subset \widetilde{A}^*$, there exist unique $q, r \in A[X]$ and $\widetilde{q}, \widetilde{r} \in \widetilde{A}[X]$ such that

$$a = bq + r, \qquad \deg(r) < \deg(b), \qquad a = b\widetilde{q} + \widetilde{r}, \qquad \deg(\widetilde{r}) < \deg(b).$$

Uniqueness implies that $\widetilde{q} = q$ and $\widetilde{r} = r$, hence

$$b \mid a \text{ in } A[X] \iff r = 0 \iff \widetilde{r} = 0 \iff b \mid a \text{ in } \widetilde{A}[X].$$

□

9.3.6 Consequences for $A[X]/(b)$ Assume that $b = b_n X^n + \cdots + b_0 \in A[X]$, $\deg(b) = n \geq 0$ and $b_n \in A^*$. Proposition 9.3.4 can be reformulated by saying that, for each polynomial $a \in A[X]$, there is a unique polynomial $r \in A[X]$ of degree $\deg(r) < n$ such that

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

In other words, if we write

$$A[X]_{\deg<n} := \{r \in A[X] \mid \deg(r) < n\} = \{r_0 + r_1 X + \cdots + r_{n-1}X^{n-1} \mid r_j \in A\}$$

(note that $A[X]_{\deg<0} = \{0\}$), then the composite map

$$A[X]_{\deg<n} \hookrightarrow A[X] \xrightarrow{\text{pr}} A[X]/(b) \tag{9.3.6.1}$$

is **bijective:**

$$\begin{aligned} A[X]/(b) &= \{r \pmod{b} \mid r \in A[X], \deg(r) < n\} \\ &= \{r_0 + r_1 X + \cdots + r_{n-1}X^{n-1} \pmod{b} \mid r_j \in A\} \\ &= \{r_0 + r_1 \overline{X} + \cdots + r_{n-1}\overline{X}^{n-1} \mid r_j \in A\}, \end{aligned} \tag{9.3.6.2}$$

where $\overline{X} = X \pmod{b} \in A[X]/(b)$ (the residue class $\pmod{b}$ of the variable X) satisfies $b(\overline{X}) = 0$, as observed in (9.3.2.2). If we denote this class by, say, $\alpha := X \pmod{b} \in A[X]/(b)$, then

$$\begin{aligned} A[X]/(b) = \{r(\alpha) \mid r \in A[X], \deg(r) < n\} = \{r_0 + r_1\alpha + \cdots + r_{n-1}\alpha^{n-1} \mid r_j \in A\}, \\ b(\alpha) = b_n\alpha^n + \cdots + b_0 = 0 \end{aligned} \tag{9.3.6.3}$$

and distinct n-tuples $(r_0, \ldots, r_{n-1})$ $(r_j \in A)$ correspond to distinct elements of $A[X]/(b)$.

The map (9.3.6.1) is A-linear: it is compatible with operations "+" and "multiplication by a constant $c \in A$" on both sides. Products $(a \pmod{b}) \cdot (\widetilde{a} \pmod{b})$ are computed by writing $a\widetilde{a} = qb + r$ with $\deg(r) < n$; then $(a \pmod{b}) \cdot (\widetilde{a} \pmod{b}) = r \pmod{b}$.

If we use the notation of (9.3.6.3) for elements of $A[X]/(b)$, then $a(\alpha)\widetilde{a}(\alpha) = r(\alpha)$.

10 Polynomial rings K[X]

Throughout Section 10, K is a field.

10.1 Basic properties of $K[X]$

10.1.1 Basic setup Recall that a field is a non-zero commutative ring in which every non-zero element is invertible. Basic examples are $K = \mathbf{Q}, \mathbf{R}, \mathbf{C}, \mathbf{Q} + \mathbf{Q}i$ or $\mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$ (for a prime p).

The property $K^* = K \smallsetminus \{0\}$ has the following consequences.

- $K[X]$ is an integral domain and $K[X]^* = K^* = K \smallsetminus \{0\} = \{a \in K[X] \mid \deg(a) = 0\}$.
- $\deg(ab) = \deg(a) + \deg(b)$ for any $a, b \in K[X]$.
- Any non-zero polynomial $a = a_n X^n + \cdots + a_0$ can be written in a unique way as $a = a_n(a_n^{-1}a)$, where $a_n \in K^*$ and $a_n^{-1}a = X^n + \cdots + a_n^{-1}a_0 \in K[X]$ is **monic** ($n = \deg(a) \geq 0$).
- A non-constant polynomial $a \in K[X] \smallsetminus K$ is **irreducible in** $K[X]$ in the sense of Section 8.3.11 if $a \neq bc$ for any non-constant polynomials $b, c \in K[X] \smallsetminus K$. For example, a is irreducible in $K[X]$ if $\deg(a) = 1$. Denote by $\mathcal{P}_K$ the set of all (non-constant) **monic irreducible** polynomials in $K[X]$. Note that $\{\pi \in \mathcal{P}_K \mid \deg(\pi) = 1\} = \{X - \alpha \mid \alpha \in K\}$.

As we are going to see, the polynomial ring $K[X]$ behaves very much like the ring of integers $\mathbf{Z}$; in particular, elements of $\mathcal{P}_K$ behave like prime numbers.

For example, much of the discussion in Section 1.1.5 applies with very minor modifications:

10.1.2 Proposition (Existence of factorisation)**.** *Every non-zero polynomial $a \in K[X] \smallsetminus \{0\}$ can be written as $a = \lambda\pi_1 \cdots \pi_r$, where $\lambda \in K^*$, $r \geq 0$ and $\pi_j \in \mathcal{P}_K$.*

Proof. Induction on $\deg(a)$, entirely analogous to induction on n in the proof of Proposition 1.2.1. □

10.1.3 Proposition (Irreducibility criterion)**.** *For a non-constant polynomial $a \in K[X] \smallsetminus K$, the following properties are equivalent.*
(1) a is not irreducible in $K[X]$.
(2) There exists $b \in K[X] \smallsetminus K$ such that $\deg(b) \leq \frac{1}{2}\deg(a)$ and $b \mid a$.
(3) There exists $\pi \in \mathcal{P}_K$ such that $\deg(\pi) \leq \frac{1}{2}\deg(a)$ and $\pi \mid a$.

Proof. If (1) holds, then $a = bc$ for some non-zero polynomials b, c such that $\deg(b) \leq \deg(c)$; then $b \mid a$ and $2\deg(b) \leq \deg(b) + \deg(c) = \deg(bc) = \deg(a)$, which proves (2). If (2) holds, then there exists $\pi \in \mathcal{P}_K$ dividing b; then $\pi \mid a$ and $\deg(\pi) \leq \deg(b) \leq \frac{1}{2}\deg(a)$. □

10.1.4 Corollary. *If $a \in K[X]$ is of degree $\deg(a) \in \{2, 3\}$, then it is equivalent:*

$$a \text{ is irreducible in } K[X] \iff a \text{ has no root in } K.$$

Proof. Indeed, $\pi \in \mathcal{P}_K$ satisfies Property (3) in Proposition 10.1.3 if and only if $\pi = X - \alpha$, where $\alpha \in K$ and $a(\alpha) = 0$. □

10.1.5 Examples (1) If $n \in \{2, 3\}$, $a \in \mathbf{N}_+$ and $\sqrt[n]{a} \notin \mathbf{N}_+$, then $X^n - a$ has no roots in $\mathbf{Q}$ (by Theorem 1.5.11), hence is irreducible in $\mathbf{Q}[X]$.
(2) This is no longer true for $n = 4$, since

$$X^4 + 4 = (X^2 - 2X + 2)(X^2 + 2X + 2).$$

10.2 Division with remainder in $K[X]$ and its consequences

10.2.1 Division with remainder in $K[X]$ The following proposition is entirely analogous to its arithmetic counterpart Proposition 2.2.2.

10.2.2 Proposition. *If $a, b \in K[X]$ and $b \neq 0$, then there exists a unique pair $q, r \in K[X]$ such that*

$$a = bq + r, \qquad \deg(r) < \deg(b).$$

Moreover, $b \mid a$ in $K[X]$ iff and only if $r = 0$.

Proof. This is a special case of Proposition 9.3.4 for $A = K$. The key assumption $b_n \in K^*$ ($n = \deg(b)$) in that proposition is an automatic consequence of the fact that $b \neq 0$. □

10.2.3 Consequences All the statements that were deduced from Proposition 2.2.2 in Sections 2 and 3 have immediate analogues in $K[X]$. Here is a brief list.

10.2.4 Euclid's algorithm, Bézout relations For any $a, b \in K[X] \smallsetminus \{0\}$, iterated division with remainder along the lines of Section 2.4 produces elements

$$d \in K[X] \smallsetminus \{0\}, \qquad u, v \in K[X]$$

such that

$$\left.\begin{array}{l} d \mid a, \;\; d \mid b \\ d = au + bv \end{array}\right\} \implies aK[X] + bK[X] = dK[X]. \tag{10.2.4.1}$$

In other words, the ideal $(a, b) = (a) + (b) = (d)$ of $K[X]$ is principal, generated by d. This is a weak form of Bézout's theorem for $K[X]$.

10.2.5 Greatest common divisor Relation (10.2.4.1) implies that

$$\begin{array}{l} d \mid a, \;\; d \mid b \\ \text{if } c \in K[X] \smallsetminus \{0\} \text{ and } c \mid a, \;\; c \mid b, \text{ then } c \mid d. \end{array} \tag{10.2.5.1}$$

The element d in (10.2.4.1) is unique up to multiplication by an element of $K[X]^* = K^* = K \smallsetminus \{0\}$. It will be unique if we require it to be **monic**.

We then say that d is the **greatest common divisor** of a and b, denoted by $\gcd(a, b) := d$. The relation

$$aK[X] + bK[X] = \gcd(a, b)K[X]. \tag{10.2.5.2}$$

is a strong form of Bézout's theorem for $K[X]$.

10.2.6 Examples Consider the first two examples from Section 9.3.3.

(1) $a = 2X^3 + 2X^2 - X + 1$, $b = 2X + 3 \in \mathbf{Q}[X]$. We know that

$$2X^3 + 2X^2 - X + 1 = (X^2 - \tfrac{1}{2}X + \tfrac{1}{4})(2X + 3) + \tfrac{1}{4},$$

which implies that $\gcd(a, b) = 1$, and gives directly

$$4(2X^3 + 2X^2 - X + 1) + (-4X^2 + 2X - 1)(2X + 3) = 1.$$

(2) $a = X^3 - 2X^2 - 7X + 3$, $b = 2X^2 + 4X - 1 \in \mathbf{Q}[X]$. We know that

$$2X^3 - 2X^2 - 7X + 3 = (\tfrac{1}{2}X - 2)(2X^2 + 4X - 1) + (\tfrac{3}{2}X + 1),$$

but we must perform one more division with remainder:

$$2X^2 + 4X - 1 = (\tfrac{4}{3}X + \tfrac{16}{9})(\tfrac{3}{2}X + 1) - \tfrac{25}{9},$$

obtaining $\gcd(a, b) = 1$ and

$$-\tfrac{25}{9} = b - (\tfrac{4}{3}X + \tfrac{16}{9})(a - (\tfrac{1}{2}X - 2)b) = -\tfrac{4}{9}(3X + 4)a + \tfrac{b}{9}(6X^2 - 16X - 23),$$

$$\tfrac{4}{25}(3X + 4)a + \tfrac{1}{25}(-6X^2 + 16X + 23)b = 1.$$

10.2.7 Ideals in $K[X]$ In fact, **every ideal of $K[X]$ is principal**. Indeed, if $I \subset K[X]$ is a non-zero ideal and if $b \in I \smallsetminus \{0\}$ is an element of minimal degree, the argument from the proof of Theorem 2.3.2 shows that $I = (b) = bK[X]$.

Again, b is unique up to multiplication by an element of K^*. It will be unique if we require it to be monic.

10.2.8 Euclid's Lemma As in Sections 1.4 and 2.3, Bézout's theorem in its weak form (10.2.4.1) implies Euclid's Lemma in $K[X]$, which implies, in turn, uniqueness of factorisation in $K[X]$.

10.2.9 Lemma (Euclid's Lemma in $K[X]$)**.** *If $\pi \in \mathcal{P}_K$, $a, b \in K[X] \smallsetminus \{0\}$, $\pi \mid ab$ and $\pi \nmid b$, then $\pi \mid a$.*

10.2.10 Theorem (Uniqueness of factorisation in $K[X]$)**.** *A non-zero polynomial $a \in K[X] \smallsetminus \{0\}$ has unique factorisation*

$$a = \lambda \prod_{\pi \in \mathcal{P}_K} \pi^{v_\pi(a)} \qquad (\lambda \in K^*)$$

Above, $v_\pi(a) \in \mathbf{N}$, and $v_\pi(a) = 0$ for all but finitely many $\pi \in \mathcal{P}_K$.

10.2.11 π-adic valuations The exponents v_π in Theorem 10.2.10 are analogues of p-adic valuations of integers. They have all the properties given in Proposition 1.5.4, with the following modification: $a = \lambda b$ for some $\lambda \in K^* \iff v_\pi(a) = v_\pi(b)$ holds for all $\pi \in \mathcal{P}_K$.

10.2.12 Least common multiple As in Theorem 1.6.2,

$$\forall a, b \in K[X] \smallsetminus \{0\} \quad \gcd(a, b) = \prod_{\pi \in \mathcal{P}_K} \pi^{\min(v_\pi(a), v_\pi(b))}.$$

Similarly, if we define the **least common multiple** of a and b by

$$m = \operatorname{lcm}(a, b) := \prod_{\pi \in \mathcal{P}_K} \pi^{\max(v_\pi(a), v_\pi(b))},$$

then m is monic, it is a common multiple of a and b, and every common multiple of a and b is a multiple of m, as in Theorem 1.6.3.

Note that $\gcd(a, b) \operatorname{lcm}(a, b) = \lambda ab$, for some $\lambda \in K^*$.

10.3 Algebraically closed fields

10.3.1 The Fundamental Theorem of Algebra This is the statement that any non-constant polynomial with complex coefficients has a complex root (in the terminology of Definition 10.3.2 below, the field of complex numbers is **algebraically closed**). In spite of its name, this theorem is of analytic nature.

10.3.2 Definition. A field K is **algebraically closed** if for each $f \in K[X] \smallsetminus K$ there exists $\alpha \in K$ such that $f(\alpha) = 0$.

10.3.3 Proposition. *If K is an algebraically closed field and if $f \in K[X]$, $\deg(f) = n \geq 0$, then there exist $\alpha_1, \ldots, \alpha_n \in K$ (not necessarily distinct) and $a_n \in K \smallsetminus \{0\}$ such that $f = a_n(X - \alpha_1) \cdots (X - \alpha_n)$.*

Proof. Induction on n. There is nothing to prove if $n = 0$. Assume that $n > 0$ and that the statement holds for polynomials of degree $\deg < n$. By assumption, there exists $\alpha_1 \in K$ such that $f(\alpha_1) = 0$, which implies that $f = (X - \alpha_1) f_1$ for some $f_1 \in K[X]$, by (9.2.3.1); then $\deg(f_1) = n - 1$. By induction hypothesis, $f_1 = a_n(X - \alpha_2) \cdots (X - \alpha_n)$ for some $\alpha_j \in K$ and $a_n \in K \smallsetminus \{0\}$; thus $f = a_n(X - \alpha_1) \cdots (X - \alpha_n)$. □

10.3.4 Theorem (The Fundamental Theorem of Algebra)**.** *For each $f \in \mathbf{C}[X] \smallsetminus \mathbf{C}$ there exists $\alpha \in \mathbf{C}$ such that $f(\alpha) = 0$.*

Proof. The proof below (due to Argand) uses the following analytic facts.

- **Compactness:** every continuous real-valued function $F : D_R \longrightarrow \mathbf{R}$ defined on a closed disc $D_R = \{z \in \mathbf{C} \mid |z| \leq R\}$ attains its infimum: there exists $z_0 \in D_R$ such that $F(z_0) = \inf_{z \in D_R} F(z)$.
- **Existence of m-th roots:** for any $z \in \mathbf{C}$ such that $|z| = 1$ and any integer $m \geq 1$ there exists $w \in \mathbf{C}$ satisfying $w^m = z$.

The polynomial $f(z) = a_n z^n + \cdots + a_0$ $(n = \deg(f) \geq 1)$ satisfies

$$\lim_{\substack{z \in \mathbf{C} \\ |z| \to +\infty}} f(z)/z^n = a_n \neq 0,$$

which implies that there exists $R > 0$ such that $|f(z)| > |f(0)|$ for all $z \in \mathbf{C}$ satisfying $|z| > R$. Therefore

$$\inf_{z \in \mathbf{C}} |f(z)| = \inf_{|z| \leq R} |f(z)| = |f(z_0)|$$

for some $z_0 \in \mathbf{C}$, $|z_0| \leq R$. This equality implies that $f(z_0) = 0$, by Lemma 10.3.5 below; thus z_0 is a root of f. □

10.3.5 Lemma (Argand)**.** *If $f \in \mathbf{C}[z] \smallsetminus \mathbf{C}$, $z_0 \in \mathbf{C}$ and $f(z_0) \neq 0$, then, for each $r > 0$, there exists $z \in \mathbf{C}$ such that $|z - z_0| < r$ and $|f(z)| < |f(z_0)|$.*

Proof. Consider the polynomial $g(z) := f(z_0 + z)/f(z_0)$. It is of the form

$$g(z) = 1 + b_m z^m + z^{m+1}(b_{m+1} + \cdots + b_n z^{n-m-1}) = 1 + b_m z^m + h(z), \qquad b_m \neq 0, \qquad (1 \leq m \leq n).$$

We must find $z \in \mathbf{C}$ such that $|z| < r$ and $|g(z)| < 1$. The idea is simple: if $|z| > 0$ is small enough, then the term $|h(z)|$ will be smaller than $|b_m z^m|$, so it will be sufficient to choose the argument of z in such a way that $b_m z^m \in \mathbf{R}$ and $b_m z^m < 0$.

More precisely, if $|z| < 1$ and $0 < |z| < r_1 := |b_m|/(|b_{m+1}| + \cdots + |b_n|)$, then

$$|h(z)| < |z^{m+1}|(|b_{m+1}| + \cdots + |b_n||z^{n-m-1}|) < |z^{m+1}|(|b_{m+1}| + \cdots + |b_n|) < |b_m z^m|.$$

Furthermore, if $|z| < r_2 := 1/|b_m|^{1/m}$, then $|b_m z^m| < 1$. Therefore $|h(z)| < |b_m z^m| < 1$ if $0 < |z| < r_0 := \min(1, r_1, r_2)$.

We need to find z such that $b_m z^m < 0$, which is equivalent to $(-b_m/|b_m|)z^m > 0$. We know that there exists $w \in \mathbf{C}$ such that $w^m = -|b_m|/b_m$; we let $z = tw$, for any $0 < t < \min(r, r_0)$. In this case $b_m z^m = -|b_m| t^m < 0$ and $|h(z)| < |b_m z^m| < 1$, hence

$$g(z) = 1 + b_m z^m + h(z) = 1 - |b_m z^m| + h(z), \qquad |g(z)| \leq |1 - |b_m z^m|| + |h(z)| = 1 - |b_m z^m| + |h(z)| < 1.$$

□

10.3.6 Another proof of Argand's Lemma Proof of the existence of m-th roots of complex numbers of modulus $|z| = 1$ requires some trigonometry: one needs to know that z can be written as $z = \cos(\alpha) + i\sin(\alpha)$ for some $\alpha \in \mathbf{R}$, and that $z = (\cos(\alpha/m) + i\sin(\alpha/m))^m$.

Here is another proof of Lemma 10.3.5 which uses only the existence of square roots of complex numbers:

- if $z = a + bi \in \mathbf{C}$, then $w := \sqrt{(a + \sqrt{a^2 + b^2})/2} + i\sqrt{(-a + \sqrt{a^2 + b^2})/2} \in \mathbf{C}$ satisfies $w^2 = z$.

If we write $g(z) = 1 + g_0(z)$, $g_0(z) = z^m(b_m + z(b_{m+1} + \cdots + b_n z^{n-m-1}))$ $(b_m \neq 0)$, then $|g(z)|^2 = (1 + g_0(z))(1 + \overline{g_0(z)}) = 1 + 2\,\mathrm{Re}(g_0(z)) + |g_0(z)|^2$. Fix $w \in \mathbf{C} \smallsetminus \{0\}$ and take $z = tw$, for $t > 0$. The fact that $\lim_{z\to 0} g_0(z)/z^m = b_m$ implies that

$$\lim_{t\to 0+} \frac{|g(tw)|^2 - 1}{t^m} = 2\,\mathrm{Re}(b_m w^m).$$

It is enough to show, therefore, that $\mathrm{Re}(b_m w^m) < 0$ for some $w \in \mathbf{C}$. Write $m = 2^k n$, where $2 \nmid n$.

- If $\mathrm{Re}(b_m) < 0$, take $w = 1$.
- If $\mathrm{Re}(b_m) > 0$, take w such that $w^{2^k} = -1$; then $w^m = -1$.
- If $\mathrm{Re}(b_m) = 0$ and $\mathrm{Re}(i^n b_m) < 0$, take w such that $w^{2^k} = i$; then $w^m = i^n$.
- If $\mathrm{Re}(b_m) = 0$ and $\mathrm{Re}(i^n b_m) > 0$, take w such that $w^{2^k} = -i$; then $w^m = -i^n$.

10.3.7 Exercise. Assume that $z_0 \in \mathbf{C}$, $R > 0$, $a_0, a_1, a_2, \ldots \in \mathbf{C}$, $a_m \neq 0$ for some $m > 0$ and $\sup_{n\in\mathbf{N}} |a_n| R^n < +\infty$.
(1) The series $f(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n$ is absolutely convergent if $z \in \mathbf{C}$ and $|z - z_0| < R$.
(2) There exists $z \in \mathbf{C}$ such that $|z - z_0| < R$ and $|f(z)| > |f(z_0)| = |a_0|$.
(3) If $f(z_0) = a_0 \neq 0$, then there exists $z \in \mathbf{C}$ such that $|z - z_0| < R$ and $|f(z)| < |f(z_0)| = |a_0|$.

10.3.8 Proposition. *(1) If K is an algebraically closed field (for example, $K = \mathbf{C}$), then the set $\mathcal{P}_K$ of (non-constant) monic irreducible polynomials in $K[X]$ is equal to $\{X - \alpha \mid \alpha \in K\}$.*
(2) The set $\mathcal{P}_{\mathbf{R}}$ is equal to $\{X - \alpha \mid \alpha \in \mathbf{R}\} \cup \{(X - \beta)(X - \overline{\beta}) = X^2 - 2\,\mathrm{Re}(\beta)X + |\beta|^2 \mid \beta \in \mathbf{C} \smallsetminus \mathbf{R}\}$.

Proof. (1) Let $f \in \mathcal{P}_K$. By assumption, there exists $\alpha \in K$ such that $f(\alpha) = 0$, which means that f is divisible by $X - \alpha$ in $K[X]$. As both f and $X - \alpha$ are irreducible (and non-constant), there exists $b \in K[X]^* = K^* = K \smallsetminus \{0\}$ such that $f = b(X - \alpha)$. However, both f and $X - \alpha$ are monic, and therefore $b = 1$.

(2) Let $f \in \mathcal{P}_{\mathbf{R}}$. By Theorem 10.3.4, there exists $\alpha \in \mathbf{C}$ such that $f(\alpha) = 0$. If $\alpha \in \mathbf{R}$, then the argument from (1) implies that $f = X - \alpha$. If $\alpha \notin \mathbf{R}$, then $0 = \overline{f(\alpha)} = \overline{f}(\overline{\alpha}) = f(\overline{\alpha})$. Consequently, f is divisible in $\mathbf{C}[X]$ by both $X - \alpha$ and $X - \overline{\alpha}$, hence also by $g := \mathrm{lcm}(X - \alpha, X - \overline{\alpha}) = (X - \alpha)(X - \overline{\alpha}) \in \mathbf{R}[X]$ (the equality follows from the fact that $\alpha \neq \overline{\alpha}$). Corollary 9.3.5 implies that $g \mid f$ in $\mathbf{R}[X]$, since $f = gh$ for some $h \in \mathbf{C}[X]$ and $f, g \in \mathbf{R}[X]$. The polynomial $g = X^2 - 2\,\mathrm{Re}(\alpha)X + |\alpha|^2 \in \mathbf{R}[X]$ is monic and irreducible in $\mathbf{R}[X]$, since $\deg(g) = 2$ and g has no roots in $\mathbf{R}$. The argument from (1) then shows that $f = g$. □

10.4 The quotient ring $K[X]/(b)$

10.4.1 Dimension of $K[X]/(b)$ If $b \in K[X] \smallsetminus \{0\}$ is a non-zero polynomial of degree $\deg(b) = n \geq 0$, then the map

$$K[X]_{\deg<n} \hookrightarrow K[X] \longrightarrow K[X]/(b) \tag{10.4.1.1}$$

is K-linear and bijective, as observed in 9.3.6.1. In particular, it is an isomorphism of K-vector spaces, hence the quotient ring $K[X]/(b)$ is a vector space of dimension n over K.

If $n = 1$, then $b = a_1(X - \alpha)$ for some $\alpha \in K$ and $a_1 \in K \smallsetminus \{0\}$. The map (10.4.1.1) is then a ring isomomorphism

$$K \xrightarrow{\sim} K[X]/(X - \alpha),$$

whose inverse is given by the evaluation isomorphism

$$\overline{\mathrm{ev}}_\alpha : K[X]/(X-\alpha) \xrightarrow{\sim} K, \qquad a \ (\mathrm{mod}\,(X-\alpha)) \mapsto a(\alpha). \tag{10.4.1.2}$$

10.4.2 The Chinese Remainder Theorem in $K[X]$ The arithmetic version of CRT given in Theorem 3.3.2 was deduced from the strong form of Bézout's theorem. One can use the relation (10.2.5.2) in the same way, or one can argue directly, in analogy with Remark 2 in Section 3.3.3.

10.4.3 Theorem (The Chinese Remainder Theorem in $K[X]$)**.** *If $a, b \in K[X] \smallsetminus \{0\}$ and $\gcd(a,b) = 1$, then the map*

$$\begin{aligned} f : K[X]/(ab) &\longrightarrow K[X]/(a) \times K[X]/(b) \\ c \ (\mathrm{mod}\, ab) &\mapsto (c \ (\mathrm{mod}\, a), c \ (\mathrm{mod}\, b)) \end{aligned}$$

is bijective (hence it is a ring isomorphism).

Proof. The map f is a ring homomorphism. If $c \ (\mathrm{mod}\, ab) \in \mathrm{Ker}(f)$, then c is divisible in $K[X]$ by both a and b, hence also by $\mathrm{lcm}(a,b) = \lambda ab / \gcd(a,b) = \lambda ab$ ($\lambda \in K^*$). Therefore $c \equiv 0 \ (\mathrm{mod}\, ab)$ and $\mathrm{Ker}(f) = \{0\}$. This implies that the map f is injective. It is then automatically bijective, since it is K-linear and both $K[X]/(ab)$ and $K[X]/(a) \times K[X]/(b)$ are K-vector spaces of dimension $\deg(ab) = \deg(a) + \deg(b)$. □

10.4.4 Example: $\mathbf{R}[X]/(X^2-1)$ If we combine the ring isomorphism

$$\begin{aligned} \mathbf{R}[X]/(X^2-1) &\xrightarrow{\sim} \mathbf{R}[X]/(X-1) \times \mathbf{R}[X]/(X+1) \\ f \ (\mathrm{mod}\,(X^2-1)) &\mapsto (f \ (\mathrm{mod}\,(X-1)), f \ (\mathrm{mod}\,(X+1))) \end{aligned}$$

given by the CRT with the evaluation isomorphisms (10.4.1.2)

$$\overline{\mathrm{ev}}_{\pm 1} : \mathbf{R}[X]/(X \mp 1) \xrightarrow{\sim} \mathbf{R}, \qquad f \ (\mathrm{mod}\,(X \mp 1)) \mapsto f(\pm 1),$$

we obtain a ring isomorphism

$$(\overline{\mathrm{ev}}_1, \overline{\mathrm{ev}}_{-1}) : \mathbf{R}[X]/(X^2-1) \xrightarrow{\sim} \mathbf{R} \times \mathbf{R}, \qquad f \ (\mathrm{mod}\,(X^2-1)) \mapsto (f(1), f(-1)). \tag{10.4.4.1}$$

Explicitly,

$$\mathbf{R}[X]/(X^2-1) = \{u+vX \ (\mathrm{mod}\,(X^2-1)) \mid u, v \in \mathbf{R}\}, \ \ (\overline{\mathrm{ev}}_1, \overline{\mathrm{ev}}_{-1}) : u+vX \ (\mathrm{mod}\,(X^2-1)) \mapsto (u+v, u-v).$$

The inverse to (10.4.4.1) is given, therefore, by the formula

$$\mathbf{R} \times \mathbf{R} \xrightarrow{\sim} \mathbf{R}[X]/(X^2-1), \qquad (s,t) \mapsto \tfrac{(s+t)+(s-t)X}{2} \ (\mathrm{mod}\,(X^2-1)). \tag{10.4.4.2}$$

In other words, $f(X) = \frac{(s+t)+(s-t)X}{2}$ is the unique polynomial in $\mathbf{R}[X]$ of degree $\deg < 2$ such that $f(1) = s$ and $f(-1) = t$.

10.4.5 Example: $\mathbf{R}[X]/(X^2+1)$ The polynomial X^2+1 is irreducible in $\mathbf{R}[X]$, but it factors in $\mathbf{C}[X]$ as $(X-i)(X+i)$; its roots are $\pm i$. The evaluation map

$$\mathrm{ev}_i : \mathbf{R}[X] \longrightarrow \mathbf{C}, \qquad f \mapsto f(i)$$

is a ring homomorphism such that $\mathrm{Im}(\mathrm{ev}_i) = \mathbf{C}$ (since $\mathrm{ev}_i(u+vX) = u+vi$) and $\mathrm{Ker}(\mathrm{ev}_i) = (X^2+1) = (X^2+1)\mathbf{R}[X]$.

Indeed, if $f \in \mathbf{R}[X]$ satisfies $f(i) = 0$, then $0 = \overline{f(i)} = \overline{f}(\bar{i}) = f(-i)$, which imples that f is divisible in $\mathbf{C}[X]$ by both $X-i$ and $X+i$, hence by their $\mathrm{lcm}(X-i, X+i) = X^2+1$. By Corollary 9.3.5, f is divisible by X^2+1 in $\mathbf{R}[X]$ (cf. the arguments in the proof of Proposition 10.3.8)).

The Homomorphism Theorem 8.5.12 then implies that the map

$$\overline{\mathrm{ev}}_i : \mathbf{R}[X]/(X^2+1) \xrightarrow{\sim} \mathbf{C}, \qquad f \pmod{(X^2+1)} \mapsto f(i)$$

is a ring isomorphism (the inverse of which sends $u+vi$ to the residue class $u+vX \pmod{(X^2+1)}$).

All of the above holds if we replace everywhere i by $-i$. The corresponding ring isomorphism

$$\overline{\mathrm{ev}}_{-i} : \mathbf{R}[X]/(X^2+1) \xrightarrow{\sim} \mathbf{C}, \qquad f \pmod{(X^2+1)} \mapsto f(-i)$$

is obtained from $\overline{\mathrm{ev}}_i$ by composition with the complex conjugation on $\mathbf{C}$, since $\mathrm{ev}_{\pm i}(u+vX) = u \pm vi$.

10.4.6 Exercise. Describe the quotient ring $\mathbf{R}[X]/(X^2+X+1)$ along the same lines.

10.4.7 Theorem. *Let $a, b \in K[X]$, $b \neq 0$. The residue class $a \pmod b \in K[X]/(b)$ is invertible in $K[X]/(b)$ if and only if $\gcd(a,b) = 1$.*

Proof. As observed in (8.5.11.2) and (9.3.2.1), $a \pmod b$ is invertible in $K[X]/(b)$ if and only if $1 \in aK[X] + bK[X] = \gcd(a,b)K[X]$, which is equivalent to $\gcd(a,b) = 1$. □

10.4.8 Computing the inverse of $a \pmod b$ One computes $\gcd(a,b)$ using Euclid's algorithm. If $\gcd(a,b) \neq 1$, then $a \pmod b$ is not invertible in $K[X]/(b)$. If $\gcd(a,b) = 1$, Euclid's algorithm also gives an explicit Bézout relation $au + bv = 1$ for some $u, v \in K[X]$, which implies that $au \equiv 1 \pmod b$ and that $u \pmod b$ is the inverse of $a \pmod b$ in $K[X]/(b)$.

10.4.9 Exercise. (1) Compute the inverse of $X^2 - 2X + 3 \pmod{(X^3-2)}$ in $\mathbf{Q}[X]/(X^3-2)$.
(2) Write $(3 - 2\sqrt[3]{2} + \sqrt[3]{4})^{-1}$ in the form $a + b\sqrt[3]{2} + c\sqrt[3]{4}$ for suitable $a, b, c \in \mathbf{Q}$.
(3) What is the relation between (1) and (2)?

10.4.10 Theorem. *Let $b \in K[X]$, $b \neq 0$. The following properties are equivalent.*
(1) $K[X]/(b)$ is a field.
(2) $K[X]/(b)$ is an integral domain.
(3) The polynomial b is non-constant and irreducible in $K[X]$.

Proof. The implication (1) $\Longrightarrow$ (2) is automatic.
(2) $\Longrightarrow$ (3): If b is constant, then $(b) = (1) = K[X]$ and $K[X]/(b) = \{0\}$ is not an integral domain. If $b = fh$ is non-constant but reducible ($f, h \in K[X]$, f, h non-constant), then $\deg(f), \deg(h) < \deg(b)$, which implies that $b \nmid f$ and $b \nmid h$. This means that the residue classes $f \pmod b$ and $h \pmod b$ are non-zero elements of $K[X]/(b)$, but their product is equal to zero, since $fh \equiv 0 \pmod b$.
(3) $\Longrightarrow$ (1): we can assume that b is irreducible and monic. We must show that every non-zero residue class $a \pmod b \neq 0 \in K[X]/(b)$ is invertible in $K[X]/(b)$. The greatest common divisor $d := \gcd(a,b)$ divides b, hence is equal either to 1 or to b, by irreducibility of b. If $d = b$, then $b \mid a$, which implies that $a \pmod b = 0$. This contradiction shows that $d = 1$, hence $a \pmod b$ is invertible in $K[X]/(b)$, by Theorem 10.4.7. □

10.4.11 Remark The non-trivial implication (2) $\Longrightarrow$ (1) can also be proved in a more abstract way as in Proposition 8.3.7 (see Exercise 8.3.9).

10.5 Applications of $K[X]/(b)$ (examples)

10.5.1 A preview The examples below are intended to illustrate the general theory from Section 10.4. They are not used anywhere else in these notes and can be skipped by the reader.

10.5.2 Lagrange interpolation Assume that we are given the following data.

- A field K
- n distict elements $\alpha_1, \ldots, \alpha_n \in K$
- n elements $t_1, \ldots, t_n \in K$

The goal is to find a polynomial $g \in K[X]$ such that

$$\deg(g) < n, \qquad \forall i = 1, \ldots, n \quad g(\alpha_i) = t_i. \tag{10.5.2.1}$$

A very special case of this problem ($n = 2$, $\alpha_1 = 1$, $\alpha_2 = -1$) was solved in Section 10.4.4.

Uniqueness: if $g, h \in K[X]$ are solutions of (10.5.2.1), then $g - h \in K[X]$ is a polynomial of degree $\deg(g - h) < n$ with at least n distinct roots in K, namely $\alpha_1, \ldots, \alpha_n \in K$. Therefore $g - h = 0$, by Theorem 9.2.7.

Construction: there is a simple expression for g in terms of the following polynomials:

$$f(X) = \prod_{i=1}^{n} (X - \alpha_i), \qquad f_i(X) = \prod_{\substack{j=1 \\ j \neq i}}^{n} (X - \alpha_j) = f(X)/(X - \alpha_i).$$

Indeed,

$$(X - \alpha_i) f_i(X) = f(X), \qquad \deg(f_i) = n - 1, \qquad f_i(\alpha_j) = \begin{cases} f_i(\alpha_i) \neq 0, & j = i \\ 0, & j \neq i, \end{cases} \tag{10.5.2.2}$$

which implies that the polynomials $p_i(X) = f_i(X)/f_i(\alpha_i)$ satisfy $p_i(\alpha_j) = \delta_{ij}$ (Kronecker's symbol) and that

$$g = \sum_{i=1}^{n} t_i \, p_i(X)$$

is a solution of (10.5.2.1). Note that

$$(X - \alpha_i)(f_i(X) - f'(\alpha_i)) = f(X) - f(\alpha_i) - (X - \alpha_i) f'(\alpha_i) \equiv 0 \pmod{(X - \alpha_i)^2},$$

by (9.2.5.2), which implies that $f_i(\alpha_i) = f'(\alpha_i)$. Consequently,

$$p_i(X) = \frac{1}{f'(\alpha_i)} \frac{f(X)}{X - \alpha_i}, \qquad g = \sum_{i=1}^{n} t_i \, \frac{1}{f'(\alpha_i)} \frac{f(X)}{X - \alpha_i}$$

is a solution of (10.5.2.1).

The constant polynomial 1 is a solution of (10.5.2.1) for $t_1 = \cdots = t_n$, which implies that

$$\sum_{i=1}^{n} p_i(X) = 1.$$

10.5.3 Algebraic reformulation Conditions $g(\alpha_i) = t_i$ are equivalent to $g \equiv t_i \pmod{(X-\alpha_i)}$, which means that (10.5.2.1) is equivalent to inverting the ring isomorphism

$$\begin{aligned} K[X]/(f) &\xrightarrow{\sim} \prod_{i=1}^{n} K[X]/(X-\alpha_i) \xrightarrow{\sim} \prod_{i=1}^{n} K \\ g \pmod f &\mapsto (g \pmod{(X-\alpha_i)}) \mapsto (g(\alpha_1), \dots, g(\alpha_n)), \end{aligned} \tag{10.5.3.1}$$

since $K[X]/(f) = \{g \pmod f \mid \deg(g) < n\}$. Under this isomorphism, the residue classes $p_i(X) \pmod f$ correspond to the elements $e_i = (0, \dots, 1, \dots, 0) \in K \times \cdots \times K$.

10.5.4 Exercise. Define a ring isomorphism $\mathbf{R}[X]/(X^2+X-2) \xrightarrow{\sim} \mathbf{R} \times \mathbf{R}$ and give an explicit formula for its inverse.

10.5.5 Determinantal formulas For small values of n one can solve (10.5.2.1) by hand. For $n = 1$, $g(X) = t_1$.

For $n = 2$, $g(X) = u + vX$ and $u + v\alpha_i = t_i$ $(i = 1, 2)$, which implies that $v(\alpha_2 - \alpha_1) = t_2 - t_1$ and

$$v = \tfrac{t_2-t_1}{\alpha_2-\alpha_1}, \qquad u = t_1 - v\alpha_1 = \tfrac{t_1(\alpha_2-\alpha_1)-(t_2-t_1)\alpha_1}{\alpha_2-\alpha_1} = -\tfrac{\alpha_1 t_2-\alpha_2 t_2}{\alpha_2-\alpha_1}, \qquad g(X) = \tfrac{(t_2-t_1)X-(\alpha_1 t_2-\alpha_2 t_2)}{\alpha_2-\alpha_1}$$

These formulas can be rewritten as follows.

$$v = \frac{\begin{vmatrix} 1 & 1 \\ t_1 & t_2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \end{vmatrix}}, \qquad u = -\frac{\begin{vmatrix} \alpha_1 & \alpha_2 \\ t_1 & t_2 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \end{vmatrix}}, \qquad Y - g(X) = \frac{\begin{vmatrix} 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & X \\ t_1 & t_2 & Y \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \end{vmatrix}}$$

10.5.6 Exercise. Write an analogous formula for $Y - g(X)$ for arbitrary $n \geq 1$.

10.5.7 Diagonalisability of matrices Let $A \in M_n(K)$ be a square matrix with entries in a field K. Denote by $P_A(X) := \det(X \cdot I_n - A) \in K[X]$ its characteristic polynomial.

The matrix A defines a linear map

$$K^n \longrightarrow K^n, \qquad X \mapsto AX.$$

The image of this map is the subspace of K^n generated by the images Ae_i of the vectors $e_i = {}^t(0, \dots, 1, \dots 0)$ of the standard basis of K^n. Note that Ae_i is equal to the i-th column of A, which means that $\mathrm{Im}(A)$ is generated by the columns of A.

For any scalar $\alpha \in K$ denote by

$$V(\alpha) := \mathrm{Ker}(A - \alpha \cdot I_n) = \{v \in V = K^n \mid Av = \alpha v\} \subset V = K^n$$

the corresponding **eigenspace** of A. The subspace $V(\alpha)$ is non-zero if and only if $P_A(\alpha) = 0$. A non-zero element of some $V(\alpha)$ is called an **eigenvector** of A, and $\alpha \in K$ the corresponding **eigenvalue**.

Note that if $v \in V(\alpha)$, then $A^2v = A(\alpha v) = \alpha^2 v$, $A^3v = A(\alpha^2 v) = \alpha^3 v$, etc. Therefore $g(A)v = g(\alpha)v$ holds for all $g \in K[X]$.

The matrix A is called **diagonalisable over** K if there exists a basis of $V = K^n$ consisting of eigenvectors of A. If we put together elements of such a basis as columns of a matrix $P \in M_n(K)$, then $P \in GL_n(K)$ is invertible and $P^{-1}AP$ is a diagonal matrix, with diagonal entries equal to the corresponding eigenvalues.

We are going to show that eigenvectors and diagonalisability can be studied using Lagrange interpolation, i.e., using an explicit version of the Chinese Remainder Theorem

$$K[X]/\prod_{i=1}^{m}(X-\alpha_i) \xrightarrow{\sim} \prod_{i=1}^{m} K[X]/(X-\alpha_i) \xrightarrow{\sim} \prod_{i=1}^{m} K \tag{10.5.7.1}$$

(for $\alpha_1, \dots, \alpha_m \in K$ distinct).

10.5.8 Proposition. *Let $\alpha_1, \dots, \alpha_m \in K$ be distinct. Define polynomials*

$$f(X) = \prod_{i=1}^{m}(X-\alpha_i), \qquad f_i(X) = \prod_{\substack{j=1 \\ j\neq i}}^{m}(X-\alpha_j) = f(X)/(X-\alpha_i), \qquad p_i(X) = f_i(X)/f_i(\alpha_i)$$

as in (10.5.2.2).
(1) If $v = v_1 + \cdots v_m$ and $v_j \in V(\alpha_j)$ for all j, then $v_i = p_i(A)v$ holds for all i.
(2) If $0 \neq v_i \in V(\alpha_i)$, then $v_1, \dots, v_m$ are linearly independent in $V = K^n$.
*(3) The subspace $W := V(\alpha_1) + \cdots + V(\alpha_m) = \{v_1 + \cdots + v_m \mid v_i \in V(\alpha_i)\} \subset V$ is a **direct** sum of the subspaces $V(\alpha_1), \dots, V(\alpha_m)$.*
(4) The projection of $W = V(\alpha_1) \oplus \cdots \oplus V(\alpha_m)$ on its i-th factor is given by $v \mapsto p_i(A)v$. In particular, $p_i(A)W = V(\alpha_i)$.
(5) $W = \mathrm{Ker}(f(A)) := \{v \in V \mid f(A)v = 0\}$.

Proof. (1) This follows from the formulas

$$p_i(\alpha_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j, \end{cases} \qquad p_i(A)v_j = p_i(\alpha_j)v_j = \begin{cases} v_i, & i = j \\ 0, & i \neq j, \end{cases} \qquad p_i(A)\big(\sum_{j=1}^{m} v_j\big) = v_i.$$

(2) If $\sum t_j v_j = 0$ for some $t_j \in K$, then $0 = p_i(A)\sum t_j v_j = t_i v_i$, hence $t_i = 0$.
(3), (4) This is an abstract reformulation of (1).
(5) If $(A - \alpha_i \cdot I)v = 0$, then $f(A)v = f_i(A)(A - \alpha_i \cdot I)v = 0$. Therefore $\mathrm{Ker}(f)$ contains each $V(\alpha_i)$, hence also their sum W. Conversely, if $v \in \mathrm{Ker}(f(A))$, then $(A - \alpha_i \cdot I)p_i(A)v = \frac{f(A)}{f_i(\alpha_i)}\, v = 0$. Therefore $p_i(A)v \in V(\alpha_i)$ and $v = (\sum_i p_i(A))v \in \sum V(\alpha_i) = W$. □

10.5.9 Proposition. *Assume that the characteristic polynomial of A is of the form $P_A(X) = \prod_{i=1}^{n}(X - \alpha_i)$, where $\alpha_1, \dots, \alpha_n \in K$ lie in K and are distinct.*
(1) Each eigenspace $V(\alpha_i) = Kv_i$ is one-dimensional and the eigenvectors $v_1, \dots, v_n$ form a basis of $V = K^n$.
(2) The matrix A is diagonalisable over K.
(3) $P_A(A) = 0$ (the Cayley–Hamilton theorem for A).
(4) The projection of $V = Kv_1 \oplus \cdots \oplus Kv_n$ on its i-th factor Kv_i is given by $v \mapsto p_i(A)$, where $p_i(X) \in K[X]$ is defined as in Proposition 10.5.8 for $f(X) = P_A(X)$.

Proof. (1) According to Proposition 10.5.8, the subspace $W := V(\alpha_1) + \cdots + V(\alpha_n) \subset V = K^n$ is a direct sum $W = V(\alpha_1) \oplus \cdots \oplus V(\alpha_n)$ of n non-zero subspaces. If we count dimensions, we obtain

$$n = \dim(V) \geq \dim(W) = \sum_{i=1}^{n} \dim(V(\alpha_i)) \geq \sum_{i=1}^{n} 1 = n,$$

which gives equalities everywhere: $V = W$, $\dim(V(\alpha_i)) = 1$, $V(\alpha_i) = Kv_i$ and $K^n = V = Kv_1 \oplus \cdots \oplus Kv_n$. The last equality is equivalent to saying that the eigenvectors $v_1, \dots, v_n$ form a basis of K^n, which proves (2). Part (3) follows from the fact that $P_A(A)v_i = f_i(A)(A - \alpha_i \cdot I)v_i = 0$, for all i. Part (4) is a special case of Proposition 10.5.8(4). □

10.5.10 Examples (1) The matrix $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \in M_2(\mathbf{R})$ satisfies $\mathrm{Tr}(A) = 5$, $\det(A) = 6$, $P_A(X) = X^2 - 5X + 6 = (X-2)(X-3)$ and

$$A - 2I = \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix}, \qquad A - 3I = \begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix}, \qquad (A-2I)(A-3I) = 0,$$
$$\mathrm{Im}(A-2I) = \mathbf{R}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \mathrm{Ker}(A-3I), \qquad \mathrm{Im}(A-3I) = \mathbf{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathrm{Ker}(A-2I).$$

In this case $\alpha_1 = 2$, $\alpha_2 = 3$, $f_1 = X - 3$, $f_2 = X - 2$, $p_1 = 3 - X$, $p_2 = X - 2$.
(2) Consider the matrix

$$A = \begin{pmatrix} -7 & 18 & 27 \\ 0 & 5 & 6 \\ -2 & 2 & 4 \end{pmatrix} \in M_3(\mathbf{R}), \qquad P_A(X) = X^3 - 2X^2 - X + 2 = (X-1)(X-2)(X+1).$$

In this case $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = -1$, $f_1 = (X-2)(X+1)$, $f_2 = (X-1)(X+1)$, $f_3 = (X-1)(X-2)$, $p_1 = -f_1/2$, $p_2 = f_2/3$, $p_3 = f_3/6$ and

$$\mathrm{Ker}(A - \alpha_i I) = \mathbf{R}v_i, \qquad v_1 = \begin{pmatrix} 0 \\ -3 \\ 2 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix},$$
$$f_1(A) = \big(4v_1\big| - 10v_1\big| - 16v_1\big), \qquad f_2(A) = \big(-6v_2\big|18v_2\big|27v_2\big), \qquad f_3(A) = \big(6v_3\big| - 12v_3\big| - 18v_3\big).$$

10.5.11 Proposition (Characterisation of diagonalisable matrices)**.** *The following properties are equivalent:*
(1) $A \in M_n(K)$ is diagonalisable over K.
(2) There exists $f = \prod_{i=1}^m (X - \alpha_i) \in K[X]$ such that $\alpha_1, \ldots, \alpha_m \in K$ are distinct and $f(A) = 0$.
(3) There exists f as in (2) such that, in addition, $P_A(\alpha_i) = 0$ for all $i = 1, \ldots, m$.

Proof. (3) $\Longrightarrow$ (2) is automatic.
(2) $\Longrightarrow$ (3): Write

$$f = \prod_{P_A(\alpha_i)=0} (X - \alpha_i) \prod_{P_A(\alpha_i)\neq 0} (X - \alpha_i) = f_1 f_2.$$

If $P_A(\alpha_i) \neq 0$, then $A - \alpha_i \cdot I$ is invertible; thus $f_2(A)$ is invertible, too. Therefore $f(A) = 0$ implies that $f_1(A) = f(A)f_2(A)^{-1} = 0$.
(1) $\Longrightarrow$ (3): By (1), $V = Ku_1 \oplus \cdots \oplus Ku_n$, $Au_j = \lambda_j u_j$. Write $P_A(A) = \prod_{j=1}^n (X - \lambda_j) = \prod_{i=1}^m (X - \alpha_i)^{k_i}$, where $\alpha_1, \ldots, \alpha_m \in K$ are distinct and $k_i \geq 1$. For each $j = 1, \ldots, n$ there exists $i \in \{1, \ldots, m\}$ such that $\lambda_j = \alpha_i$, which implies that $f(A)u_j = \prod_{i=1}^m (A - \alpha_i \cdot I)u_j = 0$.
(2) $\Longrightarrow$ (1): According to Proposition 10.5.8, $\mathrm{Ker}(f(A)) = V(\alpha_1) \oplus \cdots \oplus V(\alpha_m)$, but $\mathrm{Ker}(f(A)) = V$, since $f(A) = 0$. Choosing an arbitrary basis of each $V(\alpha_i)$, we obtain a basis of V consisting of eigenvectors of A. □

10.5.12 Remarks (1) The set $J := \{g \in K[X] \mid g(A) = 0\} \subset K[X]$ is an ideal of $K[X]$. It is non-zero, since the $n^2 + 1 > n^2 = \dim_K(M_n(K))$ matrices $I, A, A^2, \ldots, A^{n^2} \in M_n(K)$ must satisfy a non-trivial linear relation with coefficients in K. Therefore $J = (f_A) = f_A K[X]$ for a unique monic polynomial $f_A \in K[X]$, called the **minimal polynomial** of A.

(2) Proposition 10.5.11 implies that $A \in M_n(K)$ is diagonalisable over K if and only if its minimal polynomial f_A is of the form $f_A(X) = \prod_{i=1}^{m}(X - \alpha_i)$, where $\alpha_1, \ldots, \alpha_m \in K$ lie in K and are distinct.

(3) According to the Cayley–Hamilton theorem, $P_A(A) = 0$. This implies that the minimal polynomial f_A divides the characteristic polynomial P_A in $K[X]$.

Examples:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad P_A(X) = f_A(X) = X^2; \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad P_B(X) = X^2, \quad f_B(X) = X.$$

(4) The evaluation map

$$\mathrm{ev}_A : K[X] \longrightarrow M_n(K) = \mathrm{End}_K(V), \qquad g(X) \mapsto g(A)$$

is a ring homomorphism with kernel $\mathrm{Ker}(\mathrm{ev}_A) = (f_A) = f_A K[X]$. It induces, therefore, a ring isomorphism

$$\overline{\mathrm{ev}}_A : K[X]/(f_A) \xrightarrow{\sim} \mathrm{Im}(\mathrm{ev}_A), \qquad g \pmod{f_A} \mapsto g(A).$$

In the special case when $f_A(X) = \prod_{i=1}^{m}(X - \alpha_i)$, where $\alpha_1, \ldots, \alpha_m \in K$ lie in K and are distinct, then there is another ring isomorphism (10.5.7.1)

$$(\overline{\mathrm{ev}}_{\alpha_1}, \ldots, \overline{\mathrm{ev}}_{\alpha_m}) : K[X]/(f_A) \xrightarrow{\sim} \prod_{i=1}^{m} K[X]/(X - \alpha_i) \xrightarrow{\sim} \prod_{i=1}^{m} K.$$

Under this isomorphism the residue class $p_i \pmod{f_A}$ corresponds to $e_i = (0, \ldots, 1, \ldots, 0) \in K \times \cdots \times K$. On the other hand, $\mathrm{ev}_A(p_i)$ is the matrix defining the projection of $V = K^n$ on $V(\alpha_i)$.

10.6 Construction of fields

10.6.1 Theorem. *If $f \in K[X]$ is an irreducible polynomial of degree $\deg(f) = n \geq 1$, then the ring $L := K[X]/(f)$ is a field containing K and a distinguished element $\alpha := \overline{X} = X \pmod{f}$ satisfying $f(\alpha) = 0$. Elements of L can be written in a unique way as*

$$\beta = r_0 + r_1\alpha + \cdots + r_{n-1}\alpha^{n-1} = r(\alpha) \qquad (r = r_0 + r_1 X + \cdots + r_{n-1}X^{n-1} \in K[X], \deg(r) < n).$$

Equivalently, L is a K-vector space of dimension n with basis $1, \alpha, \ldots, \alpha^{n-1}$.

Proof. According to Theorem 10.4.10, the ring L is a field. The rest was proved in Section 9.3.6. □

10.6.2 Corollary. *If $g \in K[X]$ is a polynomial of degree $\deg(g) = m \geq 1$, then there exists a field $L \supset K$ in which g has a root, and a field $M \supset K$ such that $g = a_m(X - \alpha_1)\cdots(X - \alpha_m)$ for some $a_m \in K^*$ and $\alpha_1, \ldots, \alpha_m \in M$ (not necessarily distinct).*

Proof. There exists an irreducible polynomial $f \in K[X]$ (non-constant) dividing g in $K[X]$. If we let $L := K[X]/(f)$ and $\alpha_1 := X \pmod{f} \in L$, then $L \supset K$ is a field and $f(\alpha_1) = 0$, hence $g = (X - \alpha_1)h$ for some $h \in L[X]$ of degree $\deg(h) = m - 1$. We conclude by induction on m. □

10.6.3 Exercise. In the situation of Corollary 10.6.2, it is equivalent: $\alpha_1, \ldots, \alpha_m$ are distinct $\iff$ $\gcd(g, g') = 1$ in $K[X]$.

10.6.4 Converse The reader can skip the rest of Section 10.6, as it will not be used anywhere else in these notes.

The construction in Theorem 10.6.1 can be reversed along the lines of what we did in Section 10.4.5.

Assume that $K \subset M$ are fields and that $\alpha \in M$ is a root of some non-constant polynomial with coefficients in K (we say that α is **algebraic over** K). A toy model is the case $K = \mathbf{R}$, $M = \mathbf{C}$ and $\alpha = i$ considered in Section 10.4.5.

- There exists a non-constant monic polynomial $f \in K[X]$ of minimal degree such that $f(\alpha) = 0$. It is unique, since the difference of two such polynomials has smaller degree and vanishes at α. We say that f is the **minimal polynomial of α over K**.
- The polynomial f is irreducible in $K[X]$ (if $f = gh$, then $g(\alpha) = 0$ or $h(\alpha) = 0$).
- If $g \in K[X]$ satisfies $g(\alpha) = 0$, division with remainder in $K[X]$ gives $g = qf + r$, $q, r \in K[X]$ and $\deg(r) < \deg(f)$. As $r(\alpha) = g(\alpha) - q(\alpha)f(\alpha) = 0$, minimality of $\deg(f)$ implies that $r = 0$, hence f divides g in $K[X]$.
- In particular, if $g \in K[X]$ is monic, irreducible and $g(\alpha) = 0$, then $g = f$.
- Conversely, if $g \in K[X]$ is divisible by f, then $g(\alpha) = 0$.
- In other words, the kernel of the evaluation ring homomorphism $\mathrm{ev}_\alpha : K[X] \longrightarrow M$ is equal to $\mathrm{Ker}(\mathrm{ev}_\alpha) = fK[X]$. According to the Homomorphism Theorem 8.5.12, ev_α induces a ring isomorphism (sending each element of K to itself) $\overline{\mathrm{ev}}_\alpha : K[X]/(f) \xrightarrow{\sim} \mathrm{Im}(\mathrm{ev}_\alpha)$ between the abstract ring $K[X]/(f)$ and the smallest subring $K[\alpha] \subset M$ containing both K and α.
- Theorem 10.6.1 tells us that $K[X]/(f)$ is a field (hence $K[\alpha]$ coincides with the smallest subfield $K(\alpha) \subset M$ containing both K and α). Moreover, $K(\alpha) = \mathrm{Im}(\mathrm{ev}_\alpha) = K + K\alpha + \cdots + K\alpha^{n-1} \subset M$ is a K-vector space of dimension $n = \deg(f)$, with basis $1, \alpha, \ldots, \alpha^{n-1}$.
- If $K = \mathbf{R}$, $M = \mathbf{C}$ and $\alpha = i$, then $f = X^2 + 1$ and $\mathbf{R}(i) = \mathbf{R}[i] = \mathbf{R} + \mathbf{R}i = \mathbf{C}$.
- An abstract form of the above argument goes as follows: by assumption, the kernel of the evaluation ring homomorphism $\mathrm{ev}_\alpha : K[X] \longrightarrow M$ ($\mathrm{ev}_\alpha(h) = h(\alpha)$) is non-zero. As it is an ideal in $K[X]$, it must be equal to $\mathrm{Ker}(\mathrm{ev}_\alpha) = fK[X]$ for some (non-constant) polynomial $f \in K[X]$. As $K[X]/(f)$ is isomorphic to $\mathrm{Im}(\mathrm{ev}_\alpha) \subset M$, which is an integral domain, f must be irreducible.
- If $K = \mathbf{Q}$, $M = \mathbf{C}$ and $\alpha = \sqrt{3}$, then $g = X^2 - 3$ is monic, irreducible in $\mathbf{Q}[X]$ (since it has no root in $\mathbf{Q}$ and $\deg(g) \leq 3$) and $g(\sqrt{3}) = 0$. Therefore $g = f$ is the minimal polynomial of $\sqrt{3}$ over $\mathbf{Q}$ and the map $\mathrm{ev}_{\sqrt{3}} : \mathbf{Q}[X]/(X^2 - 3) \xrightarrow{\sim} \mathbf{Q} + \mathbf{Q}\sqrt{3}$ sending $h \pmod{(X^2 - 3)}$ to $h(\sqrt{3})$ is an isomorphism of fields.
- If $K = \mathbf{Q}$, $M = \mathbf{C}$ and $\alpha = \sqrt[3]{2}$, then $g = X^3 - 2$ is monic, irreducible in $\mathbf{Q}[X]$ (since it has no root in $\mathbf{Q}$ and $\deg(g) \leq 3$) and $g(\sqrt[3]{2}) = 0$. Therefore $g = f$ is the minimal polynomial of $\sqrt[3]{2}$ over $\mathbf{Q}$ and the map $\mathrm{ev}_{\sqrt[3]{2}} : \mathbf{Q}[X]/(X^3 - 2) \xrightarrow{\sim} \mathbf{Q} + \mathbf{Q}\sqrt[3]{2} + \mathbf{Q}\sqrt[3]{4}$ sending $h \pmod{(X^3 - 2)}$ to $h(\sqrt[3]{2})$ is an isomorphism of fields. See also Exercise 10.4.9.

10.6.5 Irreducibility in $\mathbf{Q}[X]$ The above examples show that it is important to be able to decide whether a given polynomial $g \in \mathbf{Q}[X]$ is reducible or irreducible in $\mathbf{Q}[X]$. The following irreducibility criteria are very useful.

10.6.6 Theorem (Gauss)**.** *(1) If $g, h \in \mathbf{Z}[X]$ and if there exists a prime p such that $p \mid gh$ in $\mathbf{Z}[X]$, then $p \mid g$ or $p \mid h$ in $\mathbf{Z}[X]$.*
(2) (Gauss' Lemma) Define the ***content*** *$ct(g)$ of a non-zero polynomial $g \in \mathbf{Z}[X]$ to be the* gcd *of its*

coefficients. Then $ct(gh) = ct(g)ct(h)$.
(3) If $f \in \mathbf{Z}[X] \smallsetminus \mathbf{Z}$ *and if* $f = gh$ *for some* $g, h \in \mathbf{Q}[X] \smallsetminus \mathbf{Q}$, *then there exists* $u \in \mathbf{Q}^*$ *such that* $f = (ug)(u^{-1}h)$ *with* $ug, u^{-1}h \in \mathbf{Z}[X] \smallsetminus \mathbf{Z}$. *["If* f *is reducible in* $\mathbf{Q}[X]$, *it is reducible in* $\mathbf{Z}[X]$."]

Proof. (1) The quotient ring $\mathbf{Z}[X]/p\mathbf{Z}[X] = (\mathbf{Z}/p\mathbf{Z})[X]$ is an integral domain, since $\mathbf{Z}/p\mathbf{Z}$ is. By assumption, the product of the residue classes $g \pmod p, h \pmod p \in \mathbf{Z}[X]/p\mathbf{Z}[X]$ is equal to zero, since $gh \in p\mathbf{Z}[X]$. Therefore $g \pmod p$ or $h \pmod p$ must also be equal to zero in $\mathbf{Z}[X]/p\mathbf{Z}[X]$.
(2) After dividing g (resp. h) by its content, we can assume that $ct(g) = ct(h) = 1$. If $ct(gh) \neq 1$, then it is divisible by some prime p. Part (1) then implies that $p \mid ct(g)$ or $p \mid ct(h)$, which is a contradiction. Therefore $ct(gh) = 1$.
(3) There exist integers $c, d \geq 1$ such that $cg, dh \in \mathbf{Z}[X]$. The polynomials $G := cg/ct(cg) = ug$ ($u := c/ct(cg) \in \mathbf{Q}^*$) and $H := dh/ct(dh)$ then lie in $\mathbf{Z}[X]$. On the other hand, $cdf/(GH) = ct(cg)ct(dh) = ct((cg)(dh)) = ct(cdf) = cd\,ct(f)$, by (2). Therefore $GHct(f) = f$ and $u^{-1}h = Hct(f) \in \mathbf{Z}[X]$. □

10.6.7 Exercise. (1) Assume that $f = a_nX^n + \cdots + a_0 \in \mathbf{Z}[X]$, $n = \deg(f) \geq 1$, and that there exists a prime $p \nmid a_n$ such that $f \pmod p$ is irreducible in $(\mathbf{Z}/p\mathbf{Z})[X]$. Then f is irreducible in $\mathbf{Q}[X]$.
(2) The polynomial $a(X) = X^3 - 2X^2 - 7X + 3$ is irreducible in $\mathbf{Q}[X]$. [Hint: take $p = 2$.]

10.6.8 Theorem (Eisenstein's irreducibility criterion)**.** *If* $f = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in \mathbf{Z}[X]$ *is a polynomial for which there exists a prime* p *such that* $p \mid a_i$ *for all* $i = 0, \ldots, n-1$ *and* $p^2 \nmid a_0$, *then* f *is irreducible in* $\mathbf{Q}[X]$.

Proof. Exercise. □

10.6.9 Corollary. *For each* $n \geq 1$, *the polynomials* $X^n - 2$ *and* $X^n - 6$ *are irreducible in* $\mathbf{Q}[X]$.

10.7 Construction of finite fields

10.7.1 A preview We know that, for each prime p, the ring $\mathbf{Z}/p\mathbf{Z} = \mathbf{F}_p$ is a field with p elements. The goal of Section 10.7 is to construct more general finite fields F as quotient rings $F = \mathbf{F}_p[X]/(f)$, for irreducible polynomials $f \in \mathbf{F}_p[X]$.

Recall from Proposition 8.5.18 that a field F contains as a subfield in a canonical way either $\mathbf{Q}$ ("F is of characteristic zero"), or $\mathbf{F}_p$ ("F is of characteristic p") for some prime p, which is unique.

10.7.2 Proposition. *Let* F *be a field.*
(1) F *is finite* $\iff$ F *is of characteristic* p *for some prime* p *and* F *is of finite dimension as a vector space over its subfield* $\mathbf{F}_p$.
(2) If $\dim_{\mathbf{F}_p}(F) = n$, *then* $|F| = p^n = q$, $|F^*| = p^n - 1 = q - 1$ *and*

$$\forall a \in F^* \quad a^{q-1} = 1, \qquad \forall a \in F \quad a^q = a.$$

Proof. (1) Finiteness of F implies that F cannot contain $\mathbf{Q}$; thus $F \supset \mathbf{F}_p$ for some prime p. If $\dim_{\mathbf{F}_p}(F) = n < \infty$, then $F \xrightarrow{\sim} \mathbf{F}_p^n$ as a vector space, hence $|F| = |\mathbf{F}_p|^n = p^n$. If $\dim_{\mathbf{F}_p}(F) = \infty$, then F contains vector subspaces isomorphic to $\mathbf{F}_p^n$ for all $n \geq 1$, hence $|F| = \infty$.
(2) The multiplicative group $F^* = (F \smallsetminus \{0\}, \cdot)$ has $|F^*| = p^n - 1 = q - 1$ elements. Lagrange's theorem in the form of Corollary 7.5.9 tells us that $\forall a \in F^* \quad a^{q-1} = 1$ (which implies that $a^q = a$). If $a \in F \smallsetminus F^*$, then $a = 0$, hence $a^q = 0 = a$. □

10.7.3 Construction of finite fields We apply Theorem 10.6.1 in the special case $K = \mathbf{F}_p$.

For any monic irreducible polynomial $f \in \mathbf{F}_p[X]$ of degree $\deg(f) = n \geq 1$, the quotient ring $F = \mathbf{F}_p[X]/(f)$ is a field. It contains $\mathbf{F}_p$ and a distinguished element $\alpha = \overline{X} = X \pmod f \in F$ such that

$f(\alpha) = 0$. Its elements are written in a unique way as

$$r_0+r_1\alpha+\cdots+r_{n-1}\alpha^{n-1} = r(\alpha) = r(X) \pmod f \qquad (r = r_0+r_1X+\cdots+r_{n-1}X^{n-1} \in \mathbf{F}_p[X],\ \deg(r) < n).$$

Equivalently, F is an $\mathbf{F}_p$-vector space of dimension n with basis $1, \alpha, \ldots, \alpha^{n-1}$. In particular, $|F| = p^n$.

If f is reducible in $\mathbf{F}_p[X]$, then the quotient ring $\mathbf{F}_p[X]/(f)$ is no longer a field, but the remaining statements above still hold, by the discussion in Section 9.3.6.

10.7.4 Computations in finite fields If we write elements of the field $F = \mathbf{F}_p[X]/(f)$ as residue classes $r \pmod f = r(\alpha)$ of polynomials $r \in \mathbf{F}_p[X]$ of degree $\deg(r) < n$, then we can compute sums by addition of polynomials, and products by taking the remainder of a product of polynomials after division by f (see the discussion in Section 9.3.6).

If $r \neq 0$ and $\deg(r) < n$, then $\beta = r \pmod f = r(\alpha)$ is invertible in F. Euclid's algorithm applied to f and r gives an explicit Bézout relation $ru + vf = 1$ for some $u, v \in \mathbf{F}_p[X]$; then $ru \pmod f = 1 \pmod f$ and $u \pmod f = u(\alpha)$ is the inverse of β in F.

10.7.5 Examples (1) $p = 2$, $n = 2$. We write $\mathbf{F}_2 = \mathbf{Z}/2\mathbf{Z} = \{0, 1\}$, $1 + 1 = 0$ (hence $-1 = 1$ in $\mathbf{F}_2$). The polynomial $f = X^2 + X + 1 \in \mathbf{F}_2[X]$ is irreducible in $\mathbf{F}_2[X]$, since $\deg(f) \leq 3$ and f has no roots in $\mathbf{F}_2$ ($f(0) = f(1) = 1 \in \mathbf{F}_2$).

The quotient ring $F = \mathbf{F}_2[X]/(X^2 + X + 1)$ is therefore a field with $p^n = 2^2 = 4$ elements. Explicitly, $F = \{0, 1, \overline{X}, \overline{X} + 1\} = \{0, 1, \alpha, \alpha + 1\}$, where $\alpha = \overline{X} = X \pmod{(X^2 + X + 1)}$ satisfies $\alpha^2 + \alpha + 1 = 0$.

For example,

$$\alpha^2 = -\alpha - 1 = \alpha + 1, \qquad \alpha^3 = \alpha(\alpha + 1) = \alpha^2 + \alpha = -1 = 1, \qquad (\alpha + 1)^2 = \alpha^2 + 2\alpha + 1 = \alpha^2 + 1 = \alpha.$$

Note that

$$X^2 - X = X(X - 1), \qquad \frac{X^4 - X}{X^2 - X} = X^2 + X + 1 \in \mathbf{F}_2[X].$$

(2) $p = 3$, $n = 2$. In this case $\mathbf{F}_3 = \mathbf{Z}/3\mathbf{Z} = \{0, 1, 2\}$ with $2 = 1 + 1$ and $1 + 1 + 1 = 0$. In particular, $2 = -1 \in \mathbf{F}_3$. The polynomial $f = X^2 + 1 \in \mathbf{F}_3[X]$ is irreducible in $\mathbf{F}_3[X]$, since $\deg(f) \leq 3$ and f has no roots in $\mathbf{F}_3$ ($f(0) = 1$ and $f(\pm 1) = 2 = -1 \in \mathbf{F}_3$).

The quotient ring $F = \mathbf{F}_3[X]/(X^2 + 1)$ is therefore a field with $p^n = 3^2 = 9$ elements. Explicitly, $F = \{0, \pm 1, \pm\overline{X}, \overline{X} \pm 1, -\overline{X} \pm 1\} = \{0, \pm 1, \pm\alpha, \alpha \pm 1, -\alpha \pm 1\}$, where $\alpha = \overline{X} = X \pmod{(X^2 + 1)}$ satisfies $\alpha^2 + 1 = 0$.

For example,

$$0 = \alpha^2 + 1 = (\alpha + 1)(\alpha - 1) - 1, \qquad (\alpha + 1)^{-1} = \alpha - 1, \qquad (\alpha + 1)^2 = \alpha^2 + 2\alpha + 1 = 2\alpha = -\alpha,$$
$$(\alpha + 1)^4 = (-\alpha)^2 = \alpha^2 = -1, \qquad (\alpha + 1)^3 = -(\alpha + 1)^{-1} = 1 - \alpha,$$
$$\alpha^6 + \alpha^5 - \alpha^3 - \alpha + 1 = (\alpha^2 + 1)(\alpha^4 + \alpha^3 - \alpha^2 + \alpha + 1) + \alpha = \alpha.$$

One can construct other irreducible polynomials from f by a simple change of variables. Two such polynomials are

$$f_\pm(Y) := f(Y \pm 1) = (Y \pm 1)^2 + 1 = Y^2 \pm 2Y + 2 = Y^2 \mp X - 1 \in \mathbf{F}_3[X].$$

The same change of variables $X = Y \pm 1$ defines an isomorphism between F and the corresponding field $F_\pm := \mathbf{F}_3[Y]/(f_\pm(Y))$:

$$F_\pm = \mathbf{F}_3[Y]/(f_\pm(Y)) \xrightarrow{\sim} \mathbf{F}_3[X]/(f(X)) = F, \qquad r(Y \pm 1) \pmod{f_\pm(Y)} \mapsto r(X) \pmod{f(X)}.$$

Note that

$$\frac{X^9 - X}{X^3 - X} = X^6 + X^4 + X^2 + 1 = (X^2+1)(X^4+1) = f(X)f_+(X)f_-(X) \in \mathbf{F}_3[X].$$

10.7.6 General results on finite fields Here is a list of basic properties of finite fields. We give some indications as to how to prove them, without going into all the details.

(1) The multiplicative group F^* of a finite field F is cyclic (of order $q - 1 = |F| - 1$).

This was proved in Theorem 5.5.2 for $F = \mathbf{F}_p$. The same argument shows that any finite subgroup A of the multiplicative group K^* of any field K is cyclic (the statement above corresponds to $K = F$ and $A = K^* = F^*$).

(2) The construction from Section 10.7.3 gives all finite fields. More precisely, every finite field F is isomorphic to a field of the form $\mathbf{F}_p[X]/(f)$, for some monic irreducible polynomial $f \in \mathbf{F}_p[X]$.

This follows from the discussion in Section 10.6.4 applied to $K = \mathbf{F}_p$, $M = F$ and $\alpha \in F$ an arbitrary generator of the cyclic group F^* (this condition implies that $\mathbf{F}_p(\alpha) = F$).

(3) A field F with $|F| = q < \infty$ elements exists $\iff$ $q = p^n$ for some prime p and $n \geq 1$. Such a field is unique up to isomorphism; it is denoted by $\mathbf{F}_q$.

We know that $|F| = q = p^n$ if it exists. In such a case Proposition 10.7.2(2) implies that the elements of F must be precisely the roots of the polynomial $X^q - X \in \mathbf{F}_p[X]$. This polynomial has q roots $\alpha_1, \ldots, \alpha_q \in L$ contained in some field $L \supset \mathbf{F}_p$, by Corollary 10.6.2, and these roots are distinct, by Exercise 10.6.3 (since the derivative of $X^q - X \in \mathbf{F}_p[X]$ is equal to $-1 \in \mathbf{F}_p[X]$).

This makes the uniqueness of F intuitively obvious; the key words are "uniqueness of the splitting field". The existence follows from the fact that $F := \{\alpha \in L \mid \alpha^q = \alpha\}$ is a subfield of L, since the iterated Frobenius map $\varphi_q : \alpha \mapsto \alpha^q$ is a ring homomorphism $L \longrightarrow L$. The field F contains all q roots α_i of $X^q - X \in \mathbf{F}_p[X]$, but it has at most $\deg(X^q - X) = q$ elements; thus $|F| = q$.

(4) $\mathbf{F}_q$ is a subfield of $\mathbf{F}_{q'}$ $\iff$ $q' = q^m$ for some integer $m \geq 1$.

The implication "$\Longrightarrow$" is automatic, since $\mathbf{F}_{q'}$ is a vector space over its subfield $\mathbf{F}_q$. Conversely, if $q' = q^m$, then $X^{q'} - X$ is divisible by $X^q - X$ in $\mathbf{F}_p[X]$, and so the set of roots of $X^q - X \in \mathbf{F}_p[X]$ is contained in the set of roots of $X^{q'} - X \in \mathbf{F}_p[X]$.

(5) For $q = p^n$ and $m \geq 1$, the polynomial $X^{q^m} - X \in \mathbf{F}_q[X]$ factors as

$$X^{q^m} - X = \prod_{\substack{f \in \mathcal{P}_{\mathbf{F}_q} \\ \deg(f) \mid m}} f$$

(recall that $\mathcal{P}_K$ denotes the set of all monic irreducible polynomials in $K[X]$).

Indeed, for any $f \in \mathcal{P}_{\mathbf{F}_q}$ of $\deg(f) = d$, the field $F = \mathbf{F}_q[X]/(f)$ has q^d elements. As $a^{q^d} - a = 0$ for all $a \in F$, one must have $f \mid (X^{q^d} - X)$ in $\mathbf{F}_q[X]$. If $d \mid m$, then $(q^d - 1) \mid (q^m - 1)$, hence $(X^{q^d} - X) \mid (x^{q^m} - X)$ in $\mathbf{F}_q[X]$. Conversely, if $f \mid (x^{q^m} - X)$ in $\mathbf{F}_q[X]$, then $\alpha^{q^m} = \alpha$ for all $\alpha \in F$, hence $\alpha^{q^m - 1} = 1$ for all $\alpha \in F^*$. Taking α a generator of the cyclic group F^*, we see that $q^m - 1$ must be divisible by $q^d - 1$ (the order of α). Write $m = da + b$ with $a, b \in \mathbf{N}$, $0 \leq b < d$. Then $q^m = (q^d)^a q^b \equiv q^b \pmod{(q^d - 1)}$, hence $(q^d - 1) \mid (q^b - 1)$. This is possible only if $b = 0$, since $0 \leq q^b - 1 < q^d - 1$. Therefore $d \mid m$. Finally, $f^2 \nmid (X^{q^m} - X)$ in $\mathbf{F}_q[X]$, since the polynomial $X^{q^m} - X \in \mathbf{F}_q[X]$ has distinct roots, as observed in (3) above.

11 Appendix: Quotients

11.1 Abstract quotients

11.1.1 Quotients and partitions Quotients naturally arise from the following data:

- a set X;
- a partition $X = \coprod_{s\in S} X_s$ of X into a disjoint union of non-empty subsets ("classes") X_s. In other words, each element of X belongs to exactly one of the classes X_s.

Two elements $x, y \in X$ (not necessarily distinct) are called **equivalent** if they belong to the same class X_s (**notation:** $x \sim y$). The set $\{X_s\}_{s\in S}$ of equivalence classes (which can be identified with S) is called the **quotient of X by the equivalence relation** $\sim$ and is denoted by $X/\sim$.

There is a canonical surjective projection map $\mathrm{pr} : X \longrightarrow X/\sim$ assigning to each element $x \in X$ the unique class to which it belongs.

This construction can be reformulated in several equivalent ways.

11.1.2 Quotients and projections If $p : X \longrightarrow S$ is any surjective map between two sets, then its fibres

$$X_s := p^{-1}(s) = \{x \in X \mid p(x) = s\} \qquad (s \in S)$$

form a partition of X. Two elements $x, y \in X$ belong to the same fibre $p^{-1}(s)$ if and only if $p(x) = p(y) = s$. We recover the situation considered in Section 11.1.1 with $X/\sim = S$ and $\mathrm{pr} = p$.

11.1.3 Relations A **relation** on a set X is a subset $\mathcal{R} \subset X \times X$. If $x, y \in X$ satisfy $(x, y) \in \mathcal{R}$, we say that x and y are in relation $\mathcal{R}$, and we write $x\mathcal{R}y$.

Examples: $X = \mathbf{Z}$ and (1) $x\mathcal{R}_1 y$ if $x = y$;
(2) $x\mathcal{R}_2 y$ if $x \neq y$;
(3) $x\mathcal{R}_3 y$ if $x \leq y$;
(4) $x\mathcal{R}_4 y$ if $x < y$;
(5) $x\mathcal{R}_5 y$ if $x \equiv y \pmod 5$.

A relation $\mathcal{R}$ on X is called

- **reflexive** if $\quad \forall x \in X \quad x\mathcal{R}x$;
- **symmetric** if $\quad \forall x, y \in X \quad [x\mathcal{R}y \Longrightarrow y\mathcal{R}x]$;
- **transitive** if $\quad \forall x, y, z \in X \quad [x\mathcal{R}y,\ y\mathcal{R}z \Longrightarrow x\mathcal{R}z]$;
- **an equivalence relation** if it is reflexive, symmetric and transitive.

In the above examples, $\mathcal{R}_1, \mathcal{R}_3$ and $\mathcal{R}_5$ are reflexive, $\mathcal{R}_1, \mathcal{R}_2$ and $\mathcal{R}_5$ are symmetric, and $\mathcal{R}_1, \mathcal{R}_3, \mathcal{R}_4$ and $\mathcal{R}_5$ are transitive. In particular, $\mathcal{R}_1$ and $\mathcal{R}_5$ are equivalence relations (but $\mathcal{R}_2, \mathcal{R}_3, \mathcal{R}_4$ are not).

11.1.4 Quotients and equivalence relations If $X = \coprod_{s\in S} X_s$ is as in Section 11.1.1 and if we define $x \sim y$ as above, then $\sim$ is an equivalence relation on X.

Conversely, assume that $\mathcal{R}$ is an equivalence relation on a set X. For each $x \in X$ consider the subset

$$C_x := \{y \in X \mid x\mathcal{R}y\} \subset X$$

(in Example 5 of Section 11.1.3, $C_x = x \pmod 5$). We have $\bigcup_{x\in X} = X$ (since $x \in C_x$, by reflexivity). If $y \in C_x$ and $z \in C_y$, then $z \in C_x$ (by transitivity), hence $C_y \subset C_x$. However, symmetry implies that

$x \in C_y$, hence $C_x \subset C_y$. To sum up, $C_y = C_x$ whenever $y \in C_x$ (which implies that $C_x \cap C_{x'} = \emptyset$ if $x' \notin C_x$).

As a result, we obtain a partition $X = \coprod_{s \in S} X_s$ into classes $X_s = C_x$ (for any $x \in X_s$). The relation $x \sim y$ defined as in Section 11.1.1 then coincides with $x \mathcal{R} y$. As in Section 11.1.1, we denote the corresponding quotient by $X/\mathcal{R}$ and we say that $X/\mathcal{R}$ is the **quotient of X by the equivalence relation $\mathcal{R}$**.

11.1.5 Universal property of $X/\mathcal{R}$ Let $\mathcal{R}$ be an equivalence relation on a set X. The quotient $X/\mathcal{R}$ has the following universal property.

If Y is a set and $f : X \longrightarrow Y$ is a map such that $f(x) = f(x')$ whenever $x \mathcal{R} x'$, then there exists a unique map $f' : X/\mathcal{R} \longrightarrow Y$ satisfying

$$f = f' \circ \mathrm{pr} : X \longrightarrow X/\mathcal{R} \longrightarrow Y.$$

Indeed, f' is determined by this property: $f'(\mathrm{pr}(x)) = f(x)$, and this definition makes sense, since $f(x) = f(x')$ whenever $\mathrm{pr}(x) = \mathrm{pr}(x')$, by assumption.

11.1.6 Relation to G/H If H is a subgroup of an abelian group G, then the relation $x \equiv y \pmod{H}$ on G is an equivalence relation (see Proposition 7.6.7) and the corresponding abstract quotient $G/\equiv \pmod{H}$ coincides with the set G/H defined in Corollary 7.6.5.

12 Solutions to some of the exercises

1.1.7 (1) Every integer is of the form either $2b$ or $2b+1$, for some $b \in \mathbf{Z}$. As $2 \nmid a$, we have $a = 2b+1$, hence both $a-1 = 2b$ and $a+1 = 2(b+1)$ are divisible by 2. If $b = 2c$, then $a-1 = 4c$ is divisible by 4. If $b = 2c+1$, then $a+1 = 4(c+1)$ is divisible by 4.
(2) Both terms $(a-1)$ and $(a+1)$ are divisible by 2, and one of them is divisible by 4; their product is then divisible by $2 \cdot 4 = 8$. Similarly, (a^2-1) is divisible by 8 and (a^2+1) is divisible by 2 (since $a^2+1 = (2b+1)^2+1 = 2(2b^2+2b+1)$); thus their product is divisible by 16. The general case is proved in the same way by induction.

1.5.7 The solution is given, more all less, in the text of the exercise. As $2160 = 2^4 \cdot 3^3 \cdot 5^1$, the number of positive divisors of 2160 is equal to $(4+1)(3+1)(1+1) = 40$, and their sum is equal to $(1+2+2^2+2^3+2^4)(1+3+3^2+3^3)(1+5) = 31 \cdot 40 \cdot 6 = 2^4 \cdot 3 \cdot 5 \cdot 31 = 7440$.

1.5.9 If $M_p = 2^p - 1 = q$ is a prime, then the sum of all positive divisors of $n = 2^{p-1}q$ is equal to $\sigma_1(n) = \sigma_1(2^{p-1})\sigma_1(q) = (1+2+\cdots+2^{p-1})(1+q) = (2^p-1)2^p = 2n$.
Conversely, if $n = 2^{a-1}m$ is perfect and $a > 1$, $2 \nmid m$, then

$$2^a m = 2n = \sigma_1(n) = (1+2+\cdots+2^{a-1})\sigma_1(m) = (2^a-1)\sigma_1(m).$$

This implies that $(2^a-1) \mid m$, hence $m = (2^a-1)k$ and $2^a = \sigma_1((2^a-1)k) \geq \sigma_1(2^a-1) \geq 1+(2^a-1)$. Therefore both inequalities must be equalities: the first one implies that $k = 1$ and $m = 2^a-1$, the second one that 2^a-1 is a prime.

1.5.15 If $(4/7)^{4/7} = a/b$ for some $a, b \in \mathbf{N}_+$, then $2^8 b^7 = 7^4 a^7$, hence

$$\forall p \in \mathcal{P} \quad 8v_p(2) + 7v_p(b) = v_p(2^8 b^7) = v_p(7^4 a^7) = 4v_p(7) + 7v_p(a),$$

which implies that $8v_p(2) - 4v_p(7) = 7(v_p(a) - v_p(b))$ is divisible by 7, for each prime $p \in \mathcal{P}$. This is false both for $p = 2$ (when $8v_2(2) - 4v_2(7) = 8$) and for $p = 7$ (when $8v_7(2) - 4v_7(7) = -4$).

1.5.27 Among the two consecutive numbers $a-1, a$ there must be a multiple of 2: either $a = 2k$ (when $2 \mid a$), or $a = 2k+1$ (when $2 \mid a-1$). Similarly, among the three consecutive numbers $a-1, a, a+1$ there must be a multiple of 3: either $a = 3k$ (when $3 \mid a$), or $a = 3k \pm 1$ (when $3 \mid a \mp 1$). For the same reason, among the five consecutive numbers $a-2, a-1, a, a+1, a+2$ there must be a multiple of 5 (since $a = 5k, 5k \pm 1, 5k \pm 2$). However, the product $(a-2)(a-1)a(a+1)(a-2) = (a^3-a)(a^2-4)$ (which is divisible by 5, by the above) is not quite equal to $a^5 - a = (a^3-a)(a^2+1)$. Their difference $a^5 - a - (a-2)(a-1)a(a+1)(a-2) = 5(a^3-a)$ is divisible by 5, though.

1.6.5 (1) If $x \leq y$, then $\min(x,y) = x$ and $\max(x,y) = y$. The case $x \geq y$ is similar. (2) Use the formulas in Theorem 1.6.2(2) and Theorem 1.6.3(2) and apply (1) to $x = v_p(a)$ and $y = v_p(b)$.

1.6.7 Easy induction on r.

2.3.7 (1) We need to show that, for each $p \in \mathcal{P}$, the p-adic valuations $v_p(\gcd(a/c, b/c)) = \min(v_p(a) - v_p(c), v_p(b) - v_p(c))$ and $v_p(\gcd(a,b)/|c|) = \min(v_p(a), v_p(b)) - v_p(c)$ are equal to each other. This follows from the fact that $\min(x-z, y-z) = \min(x,y) - z$ holds for all $x, y, z \in \mathbf{R}$.
(2) For each $p \in \mathcal{P}$ we have $v_p(a) = 0$ (hence $v_p(a^m) = 0$) or $v_p(b) = 0$ (hence $v_p(b^n) = 0$). In either case, $\min(v_p(a^m), v_p(n)) = 0$. Alternatively, we can use Bézout's theorem: there exist $u, v \in \mathbf{Z}$ such that $au + bv = 1$. Raising this equality to the power $m+n-1$, we obtain $(au+bv)^{m+n-1} = 1$. Each term in the binomial expansion of the left hand side is divisible by a^m or by b^n. This means that the left hand

side is equal to $a^m U + b^n V$ for some $U, V \in \mathbf{Z}$. The equality $a^m U + b^n V = 1$ implies that $\gcd(a^m, b^n)$ divides 1.

2.3.11 (1) As $p \mid (x-1)(x+1)$, Euclid's Lemma implies that $p \mid (x-1)$ or $p \mid (x+1)$.
(2) By (1), p divides $x-1$ or $x+1$. However, it cannot divide both of them, since it does not divide their difference $(x+1)-(x-1)=2$. In other words, if $p \mid (x-a)$ for one of the two values $a \in \{1,-1\}$, then $\gcd(p, x+a) = 1$, hence $\gcd(p^k, x+a) = 1$. As $p^k \mid (x-a)(x+a)$, Lemma 2.3.8 implies that $p^k \mid (x-a)$.
(3) In this case 2 divides both $x \pm 1$, hence $x-1 = 2y$ and $x+1 = 2(y+1)$, for some $y \in \mathbf{Z}$. By assumption, 2^k divides $x^2 - 1 = 2^2 y(y+1)$, hence $2^{k-2} \mid y(y+1)$. If $k = 2$, then we can only conclude that $2 \nmid x$. Assume that $k > 2$. The greatest common divisor $\gcd(y, y+1)$ divides $(y+1) - y = 1$, and so is equal to 1. This means that 2 cannot divide both y and $y+1$. The same argument as in (2) then shows that either $2^{k-2} \mid y$, or $2^{k-2} \mid (y+1)$, which implies that $2^{k-1} \mid (x-1)$ or $2^{k-1} \mid (x+1)$. Conversely, if either of these conditions is satisfied, then 2^k divides $x^2 - 1$.

2.3.15 Write $\alpha = a/b$, where $a, b \in \mathbf{Z}$, $b \geq 1$ and $\gcd(a,b) = 1$. According to Theorem 2.3.14, a divides 3 and b divides 1; thus $b = 1$ and $\alpha = a \in \{\pm 1, \pm 3\}$. We compute $f(1) = 0$, $f(-1) = 8$, $f(3) = 24$ and $f(-3) = 0$. The rational roots of f are, therefore, $\alpha = 1, -3$. The polynomial $f(x)$ then factors as $f(x) = (x-1)(x+3)(x-1) = (x-1)^2(x+3)$.

2.6.6 (1) Use the formula

$$\binom{p^r}{a} = \frac{p^r}{a} \prod_{j=1}^{a-1} \frac{p^r - j}{j}.$$

As $1 \leq a \leq p^r$, each term in the product satisfies $v_p(j) < r = v_p(p^r)$; thus $v_p(p^r - j) = \min(v_p(j), v_p(p^r)) = v_p(j)$. As a result,

$$v_p\Big(\binom{p^r}{a}\Big) = v_p\Big(\frac{p^r}{a}\Big) = r - v_p(a).$$

3.2.10 (1), (2) According to Proposition 4.1.5, $a^5 \pmod{5^2}$ depends only on $a \pmod 5$. Taking $a = \pm 1, \pm 2$, we obtain $a^5 \equiv \pm 1, \pm 7 \pmod{5^2}$. Consequently, if $5 \nmid xyz$, then

$$x^5 + y^5 \equiv 0, \pm 2, \pm 6, \pm 8 \not\equiv \pm 1, \pm 7 \equiv z^5 \pmod{5^2}.$$

(3) Alas, $1^7 + 2^7 \equiv 3^7 \pmod{7^2}$. In fact, one can show that, for any prime $p \equiv 1 \pmod 3$ and any $r \geq 1$, there exist $x, y, z \in \mathbf{Z}$ such that $p \nmid xyz$ and $x^p + y^p \equiv z^p \pmod{p^r}$.

3.3.6 Induction on r (one needs to show that $\gcd(m_1 \cdots m_{r-1}, m_r) = 1$).

3.3.7 (1) The first (resp. the second) congruence is equivalent to the existence of $y \in \mathbf{Z}$ (resp. of $z \in \mathbf{Z}$) such that $x = a + 4y$ (resp. $x = b + 6z$). The system has a solution if and only if the equation $a + 4y = b + 6z$ (which is equivalent to $4y - 6z = b - a$) has a solution $y, z \in \mathbf{Z}$. This is true if and only if $2 = \gcd(4, -6)$ divides $b - a$, in which case the general solution is given by $y = y_1 + 3t$, $z = z_1 + 2t$ ($t \in \mathbf{Z}$), and therefore $x = x_1 + 12t$ is unique modulo 12. The same method works in (2), with 2 and 12 being replaced by $\gcd(m,n)$ and $\operatorname{lcm}(m,n)$, respectively.
(3) If one drops the assumption $\gcd(m_i, m_j) = 1$ for all $i < j$ in Exercise 3.3.6, then a necessary condition for the existence of a solution is that $a_i \equiv a_j \pmod{\gcd(m_i, m_j)}$ holds, for all $1 \leq i < j \leq r$. The solution is then unique modulo $\operatorname{lcm}(m_1, \ldots, m_r)$.

5.2.3 Replace everywhere in the proof of Proposition 5.2.2 the number 4 by 6.

5.2.5 Given primes $p_1, \ldots, p_r \equiv 1 \pmod 6$, consider $x := 2p_1 \cdots p_r \in \mathbf{Z}$ and $N := x^2 + 3 \geq 2^2 + 3 = 7$. Let $p \mid N$ be any prime dividing N. According to Exercise 5.1.11, $p \equiv 1 \pmod 3$, hence $p \equiv 1 \pmod 6$. If $p = p_i$ for some i, then $p \mid x$, hence $p \mid (N-3)$ and $p \mid N - (N-3)$, which is impossible. Therefore $p \neq p_1, \ldots, p_r$.

5.2.6 The integer N in the proof of Proposition 5.2.4 satisfies $N \equiv 1 \pmod 4$ and $N \equiv 1 + 1 \pmod 3$, hence $N \equiv 5 \pmod{12}$. We know that each prime $p \mid N$ satisfies $p \equiv 1 \pmod 4$, hence $p \equiv 1, 5 \pmod{12}$ ($p \not\equiv 9 \pmod{12}$, since $3 \mid 9$ and $3 \mid 12$). As in the proof of Proposition 5.2.2, there exists a prime $p \mid N$ such that $p \not\equiv 1 \pmod{12}$; therefore $p \equiv 5 \pmod{12}$. We also know that $p \neq p_1, \ldots, p_r$.

5.2.8 If we knew that the existence of a solution of $x^2 \equiv 3 \pmod p$ (for a prime $p \neq 2, 3$) implies that $p \equiv \pm 1 \pmod{12}$, then we could take $N := (4p_1 \cdots p_r)^2 - 3 \geq 13$ and argue as in the proof of Exercise 5.2.6.