M3032 - Number Theory and Cryptography

Academic Year 2021/2022

Christian Lomp

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Departamento de Matemática

February 14, 2023

Contents

1 Numbers **2**

2 Some arithmetic functions **19**

3 Primitive roots of unity modulo n **30**

4 Euclidean domains **43**

5 Law of quadratic residues **57**

6 Cryptography **62**

7 Primality tests **63**

1

Numbers

Throughout this text, we will assume the axioms of Zermelo-Fraenkel[1][2] set theory (see [6]). Some of these axioms are for example the following: the Axiom of extensionality says that two sets are equal if they have the same elements. In particular a set B is called a subset of A, denoted by $B \subseteq A$ if any element of B is also an element of A. The Axiom schema of specification says that if A is a set, and ϕ is any property which might characterise the elements a of A, then there exists a subset B of A whose elements are precisely those elements $a \in A$ that satisfy property ϕ, i.e. the set

$$B = \{a \in A \mid \phi(a) \text{ is true } \} \tag{1.1}$$

exists. In particular this also implies that the empty set exists, namely if A is any set and ϕ is a property such that $\phi(a)$ is always false, then B formed as above does not contain any elements and will be denoted by $\emptyset$. The Axiom of union: says that for any set (of sets) $\mathcal{F}$ there exists a set A containing every set that is a member of some member of $\mathcal{F}$, i.e. $\bigcup_{B\in\mathcal{F}} B$ exists. The Axiom of power set says that for any set A there exists a set $\mathcal{P}(A)$ that consists of all subsets of A. [3]

Formally, the cartesian product $A \times B$ of two sets is defined as the subset of $\mathcal{P}(\mathcal{P}(A \cup B))$ consisting of the elements of the form $\{\{a\}, \{a, b\}\}$ with $a \in A$ and $b \in B$, i.e.

$$A \times B = \{\{\{a\}, \{a, b\}\} \mid a \in A, b \in B\}. \tag{1.2}$$

[1] Ernst Zermelo (1871-1953), Biography: MacTutor

[2] Abraham Fraenkel (1891-1965), Biography: MacTutor

[3] These axioms are necessary to avoid certain paradoxes, like *Russel's paradox*, which says that there does not exist the set of all sets, because if such a set would exist, call it Ω, then it would be an element of itself. Now this self-reference causes a paradox, because if we consider the subset $\mathcal{R} = \{A \in \Omega : A \notin A\}$, then this would be also a set and we would have two options $\mathcal{R} \in \mathcal{R}$ or $\mathcal{R} \notin \mathcal{R}$. In the first case, i.e. $\mathcal{R} \in \mathcal{R}$, the defining proposition of $\mathcal{R}$ says that $\mathcal{R} \notin \mathcal{R}$ which is a contradiction. On the other hand the latter condition, $\mathcal{R} \notin \mathcal{R}$ just means that $\mathcal{R}$ satisfies the defining proposition of $\mathcal{R}$, i.e. $\mathcal{R} \in \mathcal{R}$, which is a contradiction again. Thus either of these options leads to a contradiction and cannot be fulfilled.

This construction exists thanks to our axioms. Of course we will write elements of $A \times B$ in the familiar way (a, b) instead of $\{\{a\}, \{a, b\}\}$.

A relation R between two sets A and B is a subset of their cartesian product $A \times B$, i.e. $R \subseteq A \times B$. The interpretation is that two elements $a \in A$ and $b \in B$ are in relation if and only if the pair (a, b) belongs to R, i.e. $(a, b) \in R$. Usually we will use a symbol to denote that two elements are in relation. For instance considering the identity relation on a set $\mathrm{Id}_X = \{(x, x) : x \in X\}$, we would usually use the symbol "$=$" to indicate this occurrence, i.e. $a = b$ if and only if $(a, b) \in \mathrm{Id}_X$. The inverse relation of a relation $R \subseteq A \times B$ is the relation $R^{-1} = \{(b, a) \in B \times A \mid (a, b) \in R\}$.

Relations can be composed, i.e. if $R \subseteq A \times B$ is a relation and $S \subseteq B \times C$ is a relation between sets, then $S \circ R \subseteq A \times C$ is a relation defined as

$$S \circ R = \{(a, c) \in A \times C \mid \exists b \in B : (a, b) \in R,\ (b, c) \in S.\}.$$

In particular $R^{-1} \circ R = \mathrm{Id}_A$ and $R \circ R^{-1} = \mathrm{Id}_B$.

A function $f : A \to B$ between two sets A and B is defined as a relation $R \subseteq A \times B$ such that for any $a \in A$ there exists exactly one element $(a, b) \in R$ with $b \in B$. The familiar notation is that one writes $f(a) = b$ if $(a, b) \in R$. Moreover R is called the graph of f. A function f is called injective if whenever $f(a) = f(a')$, then $a = a'$. Note that the defining condition of a function, i.e. for any $a \in A$ there exists exactly one $(a, b) \in R$ does not say that f is injective. It only says that there is no ambiguity in defining the value of $f(a)$. A function f is surjective if for every $b \in B$ there exists $a \in A$ with $f(a) = b$. In general, there are functions that are injective, but not surjective and there are functions that are surjective, but not injective. A bijective function (or bijective map or bijection) is a function that is injective and surjective. In this case for every $b \in B$ there exists one and only one $a \in A$ with $f(a) = b$. Hence we can define the inverse function of f by $f^{-1}(b) = a$ if and only if $f(a) = b$. This means, that if R is the graph of f, then R^{-1} is the graph of f^{-1}.

Let $R \subseteq X \times X$ be a relation on a set X. Then R is called

- reflexive if $(a, a) \in R$ for all $a \in X$;
- symmetric if $(a, b) \in R$ implies $(b, a) \in R$;
- anti-symmetric if $(a, b) \in R$ and $(b, a) \in R$ imply $a = b$;
- transitive if $(a, b) \in R$ and $(b, c) \in R$ imply $(a, c) \in R$.

For example the identity relation Id_X is a reflexive, transitive and symmetric relation. A reflexive and transitive relation on a set X is called an equivalence relations if it is symmetric and it is called a *partial order (or partial ordering) if it is anti-symmetric. Such relations are usually denoted by symbols. We will recall equivalence relations later, but look first to partial orders.*

Let $R \subseteq X \times X$ be a partial ordering. We usually use a symbol to denote that two elements are in relation, for example $x \leq y$ if and only if $(x, y) \in R$. In this case we say that x an y are comparable. The partial ordering $\leq$ is said to be a total ordering if every two elements are comparable, that means we have $x \leq y$ or $y \leq x$ and using the set R this means $(x, y) \in R$ or $(y, x) \in R$. A minimum (with respect to $\leq$) of a subset $U \subseteq X$ is an element $x_0 \in U$ such

that $x_0 \leq y$ for any $y \in U$. A total ordering $\leq$ on a set X is a well-ordering if and only if every non-empty subset of X has a least element.

The usual order of the natural numbers is a well-ordering. This is one of the key-features of the natural numbers and we will revise the construction of the natural numbers in this section. The reader is of course well aware of the basic properties of the natural numbers and might skip this section. However, we want to illustrate the foundations of algebra given by set theory and logic. The following quote is attributed to the mathematician Leopold Kronecker[4]:

"Die natürlichen Zahlen hat Gott gemacht, alles übrige ist Menschenwerk."

which one could translated freely as "God made the natural numbers. Everything else is the work of man". The role of "god" will be played by the axioms of set theory in order to define the natural numbers $\mathbb{N}$: The Axiom of infinity says that there exists an inductive set, i.e. there exists a set X such that if $x \in X$, then X contains also the "successor" $S(x) := x \cup \{x\} \in X$. This axiom says basically that the natural numbers $\mathbb{N}$ exist as there exists a set that contains all the sets obtained by taking their successors, starting with $A = \emptyset = \{\}$:

$$\begin{aligned} \text{“0”} &= \emptyset \\ \text{“1”} &= S(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\} \\ \text{“2”} &= S(S(\emptyset)) = \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} \\ \text{“3”} &= S(S(S(\emptyset))) = \{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\} = \{\emptyset, \{\emptyset\}, \{\{\emptyset, \{\emptyset\}\}\}\} \\ \text{“4”} &= \cdots \end{aligned}$$

Another axiomatic (but equivalent) way to define the natural numbers is through the

Definition 1.1 (Peano axioms) *There exist a set $\mathbb{N}$, an injective function $S : \mathbb{N} \to \mathbb{N}$ and an element $0 \in \mathbb{N}$, such that*

1. *$0 \notin \mathrm{Im}(S)$ and*
2. *if $M \subseteq \mathbb{N}$ is a subset containing 0 and $[n \in M \Rightarrow S(n) \in M]$, then $M = \mathbb{N}$.*

Let us denote $\mathbb{N}^+ = \mathrm{Im}(S)$. Then the second property tells us that $\mathbb{N} = \{0\} \cup \mathbb{N}^+$. Hence, each natural number is either 0 or the successor of another (previous) natural number. With these axioms we define

$$1 := S(0), \qquad 2 := S(1), \qquad 3 := S(2), \qquad \ldots \qquad n+1 := S(n), \qquad \forall n \in \mathbb{N}.$$

The second property implies the usual induction principle: suppose $P(n)$ is a logical proposition for any $n \in \mathbb{N}$, then $P(n)$ is true for all $n \in \mathbb{N}$ if it is true for the first proposition $P(0)$ and whenever $P(n)$ is true for some $n \in \mathbb{N}$, then $P(n+1)$ is true. Using the second property of the Peano axioms we can form the set $M = \{n \mid P(n) \text{ is true}\}$. Thus if $P(0)$ is true and whenever $P(n)$ is true then $P(n+1)$ is true means that $0 \in M$ and for all $n \in M$ also $S(n) = n + 1 \in M$.

As we know, the natural numbers come with an addition, a multiplication and a well-ordering. The addition can be defined recursively, namely if n and m are numbers and $m =$

[4]Leopold Kronecker (1823-1891), Biography: MacTutor. The citation is taken from a memorial tribute to Kronecker by H.Weber published in the *Mathematische Annalen* from 1893; see [13].

$m'+1$, then $n+m=(n+m')+1$, meaning that we only need to add 1 recursively. Taking the successor of a number should mean "adding 1" to that number. Hence intuitively it is clear how to proceed, but in order to define the recursion properly, we will prove first the so-called Recursion Theorem in order to define addition as a function from $\mathbb{N}$ to $\mathbb{N}$ (see [6, Section 12]*):*

Theorem 1.2 (Recursion Theorem) *Let X be any set, $f: X \to X$ a function and $a \in X$ an element. Then there exists a function $f_a: \mathbb{N} \to X$ that satisfies $f_a(0) = a$ and $f_a(S(n)) = f(f_a(n))$, for all $n \in \mathbb{N}^+$.*

Proof: Recall that a function $f_a: \mathbb{N} \to X$ is a subset $R \subseteq \mathbb{N} \times X$ such that for any $n \in \mathbb{N}$ there exists exactly one $x \in X$ with $(n,x) \in R$ and $f_a(n) = x$. Thus consider the following set of subsets of $\mathbb{N} \times X$:

$$\Omega = \{U \subseteq \mathbb{N} \times X \mid (0,a) \in U \text{ and if } (n,x) \in U \text{ then } (S(n), f(x)) \in U\}.$$

Since $\mathbb{N} \times X \in \Omega$, Ω is not empty. Consider the intersection of all subsets of Ω, i.e. $R = \bigcap \{U \mid U \in \Omega\}$. It is clear that $R \in \Omega$ and that R is the minimum of Ω. Our aim will be to prove that R actually defines the function we are looking for. To do so we will use the induction principle of $\mathbb{N}$ and define

$$M = \{n \in \mathbb{N} \mid \text{ there exists precisely one } x \in X \text{ such that } (n,x) \in R\}.$$

By definition, $(0,a) \in R$. Suppose there exists $x \in X$ with $x \neq a$ and $(0,x) \in R$. We claim that $R' := R \setminus \{(0,x)\} \in \Omega$, which would contradict the minimality of R. Clearly $(0,a) \in R'$ and if $(n,y) \in R' \subseteq R$, then $(S(n), f(y)) \in R$. Since $S(n) \neq 0$, $(S(n), f(y)) \in R'$. Hence $R' \in \Omega$ would be properly contained in R, which is a contradiction. Hence $a \in X$ is the only element such that $(0,a) \in R$ and therefore, $0 \in M$.

Let $n \in M$. Then there exists a unique $x \in X$ with $(n,x) \in R$. By definition, $(S(n), f(x)) \in R$. Suppose that there exists $y \in X$ with $(S(n), y) \in R$ and $y \neq f(x)$. Consider $R' = R \setminus \{(S(n), y)\}$. Clearly $(0,a) \in R'$. If $(m,z) \in R'$, then $(S(m), f(z)) \in R$. Suppose $(S(m), f(z)) = (S(n), y)$, then $m = n$ as S is injective and $f(z) = y$. But then by the uniqueness of $(n,x) \in R$ also $(m,z) = (n,x)$, $z = x$ and finally $y = f(z) = f(x)$, which is a contradiction. Thus $(S(m), f(z)) \in R'$ and $R' \in \Omega$, which contradicts the minimality of R. Hence $(S(n), f(x)) \in R$ is unique and $S(n) \in M$. By the induction principle, $M = \mathbb{N}$.

Now we can define $f_a: \mathbb{N} \to X$ by $f_a(0) = a$ and $f_a(n) = x$ if and only if $(n,x) \in R$. By construction, if $f_a(n) = x$, then $(S(n), f(x)) \in R$ and hence $f_a(S(n)) = f(x) = f(f_a(n))$.

□

The constructed function f_a has the property that

$$f_a(1) = f(f_a(0)) = f(a), \qquad f_a(2) = f(f_a(1)) = f(f(a)), \qquad \text{and so on.}$$

Applied to $X = \mathbb{N}$ and the function $f = S$, we obtain for all $m \in \mathbb{N}$ a function $f_m: \mathbb{N} \to \mathbb{N}$ defined by $f_m(0) = m$ and $f_m(S(n)) = S(f_m(n))$, for all $n \in \mathbb{N}$. Note that $f_0 = id_{\mathbb{N}}$, because if

$$M_e = \{n \in \mathbb{N} \mid f_0(n) = n\},$$

then $0 \in M_e$ and if $n \in M_e$, then $f_0(S(n)) = S(f_0(n)) = S(n)$. Hence $S(n) \in M_e$. By the induction principle, $M_e = \mathbb{N}$ and $f_0 = id_{\mathbb{N}}$. In the same spirit we can show that the set

$$M_1 = \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : f_{S(m)}(n) = f_m(S(n))\}$$

is $\mathbb{N}$, because for any $m \in \mathbb{N}$: $f_m(S(0)) = S(f_m(0)) = S(m) = f_{S(m)}(0)$, i.e. $0 \in M_1$. If $n \in M_1$ and $m \in \mathbb{N}$, then $f_m(S(S(n))) = S(f_m(S(n))) = S(f_{S(m)}(n)) = f_{S(m)}(S(n))$. Similarly, $f_m(n) = f_n(m)$, for all $n, m \in \mathbb{N}$, by considering

$$M_c = \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : \ f_n(m) = f_m(n)\}.$$

We have $f_0(m) = m = f_m(0)$ and hence $0 \in M_c$. If $n \in M_c$ then $f_{S(n)}(m) = f_n(S(m)) = f_{S(m)}(n) = f_m(S(n))$, for all $m \in \mathbb{N}$. Hence $S(n) \in M_c$ and we have shown that $M_c = \mathbb{N}$, i.e. $f_n(m) = f_m(n)$ for all $n, m \in \mathbb{N}$.

This allows us now to define a binary operation on $\mathbb{N}$ as

$$+ : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}, \qquad (n, m) \mapsto f_m(n) =: n + m.$$

In particular, $+$ is commutative, since $f_n(m) = f_m(n)$ and 0 is the identity element as $f_0 = id_{\mathbb{N}}$, i.e. for all $n, m \in \mathbb{N}$:

$$\begin{aligned} 0 + m &= m \\ S(n) + m &:= S(n + m) \end{aligned}$$

Note that $S(n) = S(n+0) = n+S(0) = n+1$ shows that the successor of n is precisely $n+1$. The associativity follows again by an application of the induction principle (see exercise 5).

Moreover, for all $n, m, k \in \mathbb{N}$: if $n + m = n + k$, then $m = k$. Again this can be proven by the induction principal, just consider the set

$$M = \{n \in \mathbb{N} \mid \forall m, k \in \mathbb{N} : \ \text{if } n + m = n + k \text{ then } m = k\}$$

and note that $0 \in M$ as 0 is the neutral element of $+$ and if $n \in M$, then $S(n)+m = S(n)+k$ implies $S(n + m) = S(n + k)$ and by injectivity of S, $n + m = n + k$. Since $n \in M$ we must have $m = k$. Thus $S(n) \in M$. By the induction principle $\mathbb{N} = M$. This means that $(\mathbb{N}, +, 0)$ is what we will call a cancelative commuative monoid.

Definition 1.3 *Define an ordering on $\mathbb{N}$ as follows:*

$$\forall n, m \in \mathbb{N} : \ n \leq m \text{ if and only if there exists } k \in \mathbb{N} \text{ such that } n + k = m. \tag{1.3}$$

It is clear that $\leq$ is reflexive, since $n = n + 0$ for all $n \in \mathbb{N}$ and also transitive, since $+$ is associative. For the anti-symmetry, let $n \leq m$ and $m \leq n$. Then there are $k, l \in \mathbb{N}$ such that $m = n + k$ and $n = m + l$. But then $n + (k + l) = n = n + 0$ and by cancellation, $k + l = 0$. If $k \neq 0$, then $k = S(k')$, for some $k' \in \mathbb{N}$ and hence $S(k' + l) = S(k') + l = k + l = 0$, contradicting $0 \notin \mathrm{Im}(S)$. Thus $k = 0$ and $m = n$.

To prove that $\leq$ is a total order we apply again the induction principle to the set $M = \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : n \leq m \text{ or } m \leq n\}$, where we have to use the fact that if $n \leq m \leq S(n)$,

then $n = m$ or $m = S(n)$. Finally, to prove the well-ordering, let $U \subseteq \mathbb{N}$ be a subset of $\mathbb{N}$. Suppose U does not have a minimum. Then consider the set of all lower bounds of U, i.e.

$$M = \{n \in \mathbb{N} \mid \forall x \in U : n < x\}.$$

Since by hypothesis U does not have a minimum, $0 \in M$. Let $n \in M$. Then for all $x \in U$, if $x \leq S(n)$, then $n < x \leq S(n)$ and hence $x = S(n)$. Thus if $S(n) \notin M$, then $S(n)$ would be a minimum of U, which is a contradction. Hence $S(n) \in M$. By the induction principle, $M = \mathbb{N}$. But then $U = \emptyset$ and we conclude that any non-empty subset of $\mathbb{N}$ must have a minimum.

Note also that $\leq$ is monotone with respect to $+$, i.e. if $n \leq m$, then for all $k \in \mathbb{N}$ also $n + k \leq m + k$. This means that $(\mathbb{N}, +)$ is an ordered monoid. We just proved the following Theorem:

Theorem 1.4 *The natural numbers $(\mathbb{N}, +)$ with the total ordering $\leq$ is a commutative cancellative well-ordered monoid.*

Let $P(n)$ be a logical proposition, for all $n \in \mathbb{N}$. In order to prove that all propositions $P(n)$ are true, the complete induction requires that $P(0)$ is true and that whenever $n \geq 0$ and $P(m)$ is true for all $m \leq n$, then also $P(n+1)$ is true. The complete induction is also a consequence of the second property of the Peano Axioms, by setting $M = \{n \mid P(m)$ is true for all $m \leq n\}$, because under the hypothesis of complete induction, $0 \in M$ and $n \in M$ implies $S(n) \in M$, shows $M = \mathbb{N}$.

The idea of a "finite" and "infinite" set can be made precise through the total ordering of the natural numbers. Intuitively it is clear what a finite set is, namely a set whose elements can be "enumerated" in "finitely many" steps. More precisely, two sets are called equivalent (or isomorphic) if there exists a bijection between them. For a number $n \in \mathbb{N}$ we say that a set X is a finite set of cardinality n if it is equivalent to the interval

$$[0, n[= \{m \in \mathbb{N} : m < n\}.$$

A set X that is not equivalent to a set of the form $[0, n[$ is called an infinite set. Note that for $n = 0$, the subset $[0, 0[= \emptyset$ is empty and hence a set X has cardinality 0 if and only if it is empty. We will denote the cardinality of a finite set X by $|X| = n$. If X is infinite we will write $|X| = \infty$. Let $n \in \mathbb{N}$. Every proper subset X of $[0, n[$ is a finite set of cardinality $|X| < n$. We can prove this by induction. If $n = 0$, then $[0, 0[$ is empty and has no proper subset, so assume $n = 1$. Then $[0, 1[= \{0\}$ and X can only be the empty set, which has cardinality $0 \leq 1$. Let $n \geq 0$ and suppose we have shown that any proper subset of $[0, n[$ is finite. If $X \subset [0, n + 1[$, then either $X \subseteq [0, n[$ or $n + 1 \in X$. In the first case, by induction, X is finite of cardinality at most n. [5] In the latter case, there must exists $k \leq n$ such that $k \notin X$. Define a map $f : X \to [0, n[$ by $f(x) = x$ for all $x \neq n + 1$ and$f(n + 1) = k$. Then X and the image $f(X)$ are equivalent. Since $f(X) \subseteq [0, n[$ it is finite of cardinality at most n.

The usual multiplication is again defined by the recursion theorem. For a given $n \in \mathbb{N}$ let $f_n : \mathbb{N} \to \mathbb{N}$ as above, defined by $f_n(m) = n + m$. For $a = 0$, the recursion theorem says that

[5] if X is not a proper subset of $[0, n[$, then it is equal to $[0, n[$ and finite by definition.

there exists a map $g_n := (f_n)_0 : \mathbb{N} \to \mathbb{N}$ *such that* $g_n(0) = 0$ *and* $g_n(S(m)) = f_n(g_n(m)) = n + g_n(m)$*, for all* $m \in \mathbb{N}$*. This allows us now to define a multiplication on* $\mathbb{N}$ *as*

$$\cdot : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{N}, \qquad (n, m) \mapsto n \cdot m := g_n(m).$$

In particular, for all $n, m \in \mathbb{N}$*:*

$$\begin{aligned} n \cdot 0 &= g_n(0) = 0 \\ n \cdot S(m) &:= g_n(S(m)) = n + g_n(m) = n + n \cdot m \end{aligned}$$

As an exercise, one checks that $\cdot$ *is commutative, associative and distributes over* $+$*. Then it follows that* $1 \cdot n = n \cdot 1 = n + n \cdot 0 = n$ *holds for all* $n \in \mathbb{N}$*.*

Lemma 1.5 *Let* $n, m \in \mathbb{N}$ *and* $k \in \mathbb{N}^+$*. Then*

(i) $n \cdot m = 0$ *if and only if* $n = 0$ *or* $m = 0$*.*

(ii) $n = m$ *if and only if* $n \cdot k = m \cdot k$*.*

(iii) $n < m$ *if and only if* $n \cdot k < m \cdot k$*.*

(iv) $n \cdot m = 1$ *if and only if* $n = m = 1$*.*

P*roof:* *(i) Let* $n > 0$ *and* $m > 0$*, then* $n = S(n')$ *and* $m = S(m')$ *for* $n', m' \in \mathbb{N}$*. Thus*

$$n \cdot m = n \cdot S(m') = n + S(n') \cdot m' = n + m' + n' \cdot m' > 0,$$

because $n > 0$ *and addition preserves the order. If* $n = 0$ *o or* $m = 0$*, then* $n \cdot m = 0$*.*

(ii) Suppose $n \cdot k = n \cdot m$*. Without loss of generality we can assume* $n \leq m$*. Then* $m = n + l$ *for some* $l \in \mathbb{N}$*. Hence, by distributivity,* $n \cdot k = m \cdot k = n \cdot k + l \cdot k$ *and by the cancellation property of the addition,* $l \cdot k = 0$*. Since* $k \neq 0$*, by (i),* $l = 0$ *and* $n = m$*.*

(iii) If $n < m$*, then there exists* $l > 0$ *with* $m = n + l$*. Thus,* $m \cdot k = n \cdot k + l \cdot k$*. By* (i)*,* $l \cdot k \neq 0$*. Thus* $n \cdot k < m \cdot k$*. Conversely, if* $n \cdot k < m \cdot k$ *and* $n \geq m$ *then* $n = m + l$ *for some* $l > 0$ *and*

$$m \cdot k + l \cdot k = n \cdot k < m \cdot k,$$

which implies $l \cdot k = 0$ *and hence by* (i)*,* $l = 0$ *as* $k > 0$*. But then* $n = m$ *and* $n \cdot k = m \cdot k$*, a contradiction. Thus* $n < m$*.*

(iv) Clearly $1 \cdot 1 = 1 + 1 \cdot 0 = 1$*. Conversely, if* $n \cdot m = 1$*, then* $n \neq 0$ *by* (i)*. Hence* $n \geq 1$ *by* (iii)*,* $1 = n \cdot m \geq m$*, showint* $m = 1$*.* □

Division with Rest in $\mathbb{N}$

One of the key features of the numbers is division with rest.

Proposition 1.6 (division with rest in $\mathbb{N}$***)*** *For any* $n, m \in \mathbb{N}$ *with* $n > 0$ *there exist* $q, r \in \mathbb{N}$ *such that*

$$m = q \cdot n + r \qquad \text{and} \qquad r < n. \tag{1.4}$$

The number q *is called the quotient of* m *divided by* n *and* r *is called the rest of the division of* m *by* n*, denoted by* $r := m \pmod n$*.*

P*roof: Let* $U = \{r \in \mathbb{N} : \exists q \in \mathbb{N} : m = q \cdot n + r\}$. *Then* $U \neq \emptyset$ *since* $m \in U$. *Hence by the total ordering,* U *has a minimum, say* $r \in U$ *and there exists* $q \in \mathbb{N}$ *such that* $m = q \cdot n + r$. *If* $r = 0$, *then* $r < n$. *Assume* $r > 0$. *If* $r \not< n$, *then* $n < r$ *ou* $n = r$. *If* $n = r$, *then* $m = q \cdot n + n = S(q) \cdot n + 0$ *and* $0 \in U$. *But as* r *was the minimum of* U, $r = 0$ *which contradicts our assumption. If* $n < r$, *then there exists* $r' > 0$ *with* $r = n + r'$ *and* $r' \in U$ *as* $m = q \cdot n + (n + r') = S(q) \cdot n + r'$. *This is impossible as* $r' < r$ *and* r *is the minimum of* U. □

Divisibility in $\mathbb{N}$

Let $n \in \mathbb{N}$. *A divisor of* n *is a number* $d \in \mathbb{N}$ *such that* $n = d \cdot e$ *for some* $e \in \mathbb{N}$. *We write* $d \mid n$ *in this case. If* $n = 0$, *any number* d *is a divisor of* 0, *because* $0 = d \cdot 0$. *Thus* $d \mid 0$ *for any* d. *On the other hand Lemma 1.5(i) says that if* $0 \mid n$ *then* $n = 0$.

In case $n \neq 0$ *and* $n = de$, *then* $d, e \geq 1$, *hence by the monotony of the multiplication* $e \geq 1 \Rightarrow n = de \geq d$. *This shows that* $d \mid n$ *implies* $1 \leq d \leq n$.

Lemma 1.7 *The relation* $\mid$ *is a partial order relation on* $\mathbb{N}$.

P*roof:* $\mid$ *is reflexive as* $n = n \cdot 1$, *i.e* $n \mid n$. *The relation* $\mid$ *is transitive, since if* $d \mid n$ *and* $n \mid m$, *then* $n = de$ *and* $m = nf$ *for some* e *and* f. *Hence (by the associativity),* $m = (de)f = d(ef)$, *i.e.* $d \mid m$. *Moreover,* $\mid$ *is anti-symmetric, as if* $d \mid n$ *and* $n \mid d$, *then* $n = de$ *and* $d = nf$ *for some* e, f *and thus* $n = nfe$. *By Lemma 1.5(ii),* $1 = fe$ *and by 1.5(iv)* $f = e = 1$. *Thus* $n = d$. □

Later we will see that the anti-symmetry fails over the integers $\mathbb{Z}$, *because in the last step we might only conclude* $d = \pm 1$.

The relation $\mid$ *has still another important property, which follows from distributivity.*

Lemma 1.8 *Let* $d, n, m \in \mathbb{N}$.

1. *If* $d \mid n$, *then* $d \mid nm$ *if* $m \neq 0$.
2. *If* $d \mid n$ *and* $d \mid m$, *then* $d \mid n + m$.
3. *If* $d \mid n$ *and* $d \mid n + m$, *then* $d \mid m$.
4. *If* $d \mid 1$ *then* $d = 1$.

P*roof: (1) is clear by the associativity of the multiplication.*

(2) I If $d \mid n$ *and* $d \mid m$, *then* $n = de$ *and* $m = df$. *Thus* $n + m = d(e + f)$, *i.e.* $d \mid n + m$.

(3) If $d \mid n + m$ *and* $d \mid n$ *then there are* $e, f \in \mathbb{N}$ *such that* $de = n + m$ *and* $n = df$. *Hence,* $de = n + m = df + m$ *implies* $de \geq df$ *and therefore* $e \geq f$ *by Lemma 1.5(iii). Thus* $e = f + h$ *for some* h *and* $n + m = de = df + dh = n + dh$ *implies* $m = dh$, *i.e.* $d \mid m$.

(4) follows from Lemma 1.5.

□

Note that of course $d \mid n + m$ *does not need to imply* $d \mid n$ *or* $d \mid m$. *For instance* $2 \mid 2 = 1 + 1$.

Greatest common divisor in $\mathbb{N}$

A common divisor of two numbers a and b is an elements $d \in \mathbb{N}$ such that $d \mid a$ and $d \mid b$. Note that if $b = 0$, then any divisor of a is also a common divisor of a and 0. The greatest common divisor of two numbers a and b that are not both zero is a common divisor d such that for all common divisors e of a and b one also has $e \mid d$ and in particular $e \leq d$. Hence the greatest common divisor is in deed the largest common divisor with respect to the total ordering of $\mathbb{N}$. Does the greatest common divisor exist? Clearly if $b = 0$, then the a is the greatest common divisor of a and 0 (provided $a \neq 0$). In general if $b \neq 0$, then we can recursively calculate the greatest common divisor using the division algorithm.

Proposition 1.9 *Let a, b non-zero numbers. Then the set of common divisors of a and b is equal to the set of common divisors of b and r, where r is the rest of the division of a by b.*

P*roof: Let $a = qb + r$ with $0 \leq r < b$. If d is a common divisor of a and b, then $d \mid a = qb + r$ and $d \mid qb$. By Lemma 1.8, $d \mid r$. Hence d is a common divisor of b and r. Conversely, if d is a common divisor of b and r, then d divides $a = qb + r$, again by Lemma 1.8.* □

Theorem 1.10 *Every two non-zero numbers have a greatest common divisor.*

P*roof: We use complete induction to prove the claim. Let $a, b \in \mathbb{N}$ non-zero numbers. Set $M = \{b \in \mathbb{N}^+ \mid \forall a \in \mathbb{N} : \text{ the greatest common divisor of } a \text{ and } b \text{ exists}\}$. Clearly $1 \in M$, because for any $a \in \mathbb{N}$, $1 \mid a$ and hence 1 is the greatest common divisor of 1 and a. Let $b \in \mathbb{N}^+$ and suppose $r \in M$, for any $r < b$. Write $a = qb + r$ for $q, r \in \mathbb{N}$ and $0 \leq r < b$. If $r = 0$, then $b \mid a$ and b is the greatest common divisor of a and b. If $r \neq 0$, then $r \in M$ by hypothesis and the greatest common divisor of b and r exists, say d. By Proposition 1.9 The set of common divisors of a and b is the same as the set of common divisors of b and r. Hence d is also the greatest common divisor of a and b. Therefore, $b \in M$ and by complete induction we conclude $M = \mathbb{N}^+$.* □

After having introduced the integers we will describe the extended Euclidean Algorithm which let us easily calculate the greatest common divisor of two numbers.

Prime numbers

A number $p \in \mathbb{N}$ is called a prime number if $p \geq 2$ and 1 and p are the only divisors of p. We denote the set of prime numbers by $\mathbb{P}$. Clearly $2 \in \mathbb{P}$, because the only numbers less than or equal to 2 are $0, 1, 2$ and the only divisors are 1 and 2.

Theorem 1.11 *Any number greater than 1 is a product of prime numbers.*

P*roof: We use again complete induction to prove our claim. Let*

$$M = \{n \in \mathbb{N}_{>1} \mid n \text{ is a product of prime numbers}\}.$$

Then $2 \in M$. Let $n \in \mathbb{N}_{>1}$ and suppose that $a \in M$, for all numbers $a < n$. If n is not itself a prime number, then $n = ab$, for $a, b < n$. By hypothesis, a and b are products of prime numbers and so is n. □

The last Theorem and Lemma 1.8 allows us to conclude that the set of prime numbers is not finite.

Theorem 1.12 *The set $\mathbb{P}$ of prime numbers is not finite.*

P*roof: Suppse $\mathbb{P}$ is a finite set. Then we can enumerate $\mathbb{P} = \{p_1 = 2, p_2, \ldots, p_s\}$ for some $s \geq 1$. Let $m = p_1 \cdots p_s$ and $n = m + 1$. Then $n > p_i$ for all i, since multipilication is monotone and therefore $n \notin \mathbb{P}$. By Theorem 1.11 , there exists a prime number $q \mid n$. Thus $q = p_i \in \mathbb{P}$, for some i. Since $p_i \mid m$ and $p_i \mid n = m + 1$, we have by Lemma 1.8, $p_i \mid 1$, which implies $p_i = 1$ by the same Lemma. Since prime numbers are larger than 1, this is a contradiction. Hence $\mathbb{P}$ cannot be finite.* $\square$

The Integers

The integers are build from the natural numbers. Intuitively one could just take two copies of $\mathbb{N}$, "glue" them together at 0 and call one half the positive and the other half the negative integers. A more formal, but efficient way to define the integers is by introducing them as equivalent classes of pairs of natural numbers. On the set of pairs of natural numbers, $\mathbb{N}^2$, define the equivalence relation

$$(a, b) \sim (c, d) \Leftrightarrow a + d = b + c, \qquad \forall (a, b), (c, d) \in \mathbb{N}^2. \tag{1.5}$$

The reader should verify that $\sim$ is reflexive, symmetric and transitive (using the commutativity and associativity of the addition in $\mathbb{N}$). The integers are then defined as the set of equivalence classes: $\mathbb{Z} = \{[(a, b)]_\sim : (a, b) \in \mathbb{N}^2\}$. An equivalence class $[(a, b)]$ should be thought of as the difference $a - b$. The relation $(a, b) \sim (c, d)$ then simply says that the difference $a - b$ and $c - d$ are the same, because $a + d = b + c$.

A total order on $\mathbb{Z}$ is defined by

$$[(a, b)] < [(c, d)] \qquad \text{if and only if} \qquad a + d < b + c. \tag{1.6}$$

The arithmetic of $\mathbb{Z}$ is induced by the arithmetic of $\mathbb{N}$ by setting for $(a, b), (c, d) \in \mathbb{N}^2$:

$$\begin{aligned} [(a, b)] + [(c, d)] &:= [(a + c, b + d)] \\ [(a, b)] \cdot [(c, d)] &:= [(ac + bd, ad + bc)] \end{aligned}$$

The reader should verify that these definitions are well-defined and independent of the representative of the equivalence classes and that the ordinary laws of associativity, commutativity and distributivity hold. The additive inverse of an integer is obtained by reversing the order of the pair, i.e. $-[(a, b)] := [(b, a)]$, because $[(b, a)] + [(a, b)] = [(a + b, a + b)] = [(0, 0)]$. Subtraction is the set to

$$[(a, b)] - [(c, d)] := [(a, b)] + [(d, c)] = [(a + d, b + c)]. \tag{1.7}$$

It is easily verified that these definitions are independent of the choice of representatives of the equivalence classes $[(a, b)]$.

We actually have

$$\mathbb{Z} = \{[(k, 0)] : k > 0\} \cup \{[(0, 0)]\} \cup \{[(0, k)] : k > 0\}, \tag{1.8}$$

because if $(a,b) \in \mathbb{N}^2$*, then either* $a < b$*,* $a = b$ *or* $a > b$*. If* $a > b$*, then there exist* $k \in \mathbb{N}$ *such that* $a = b + k$*. Thus* $(a,b) \sim (k,0)$ *and hence* $[(a,b)] = [(k,0)]$*. Analogously, if* $a < b$*, then* $[(a,b)] = [(0,k)]$*, for some* $k \in \mathbb{N}$*. If* $a = b$*, then* $(a,b) \sim (0,0)$*, i.e.* $[(a,b)] = [(0,0)]$*. Note that* $(k,0) \sim (k',0)$ *if and only if* $k = k'$*. A natural number* $a \in \mathbb{N}$ *can be identified with the classes* $[(a,0)]$ *and the negative numbers are identified with the classes* $[(0,a)]$*. Elements of the form* $[(k,0)]$ *are simply written as* k*, while elements of the form* $[(0,k)]$ *as* $-k$*. Note that in particular* $-1 = [(0,1)]$*,* $(-1)(-1) = 1$*,* $(-1) \cdot k = [(0,1)] \cdot [(k,0)] = [(0,k)] = -k$ *and* $(-1) \cdot (-k) = k$ *for all* $k \in \mathbb{N}^+$*. We set*

$$\mathbb{Z}^+ = \{k \mid k \in \mathbb{N}^+\}, \qquad \mathbb{Z}^- = \{-k \mid k \in \mathbb{N}^+\}.$$

Thus $a < 0 < b$ *for any* $a \in \mathbb{Z}^-$ *and* $b \in \mathbb{Z}^+$*. The sign of* $a \in \mathbb{Z} \setminus \{0\}$ *is then defined as* $\operatorname{sign}(a) = 1$ *if* $a \in \mathbb{Z}^+$ *and* $\operatorname{sign}(a) = -1$ *if* $a \in \mathbb{Z}^-$*. The absolute value or norm* $|n|$ *of an integer* n *is defined as* $|n| = n$ *if* $\operatorname{sign}(n) = 1$ *and* $|n| = -n$ *if* n *is represented as* $[(0,k)]$*. We have,* $a = \operatorname{sign}(a)|a|$ *for any* $a \in \mathbb{Z} \setminus \{0\}$*.*

Division with Rest in $\mathbb{Z}$

We saw in Proposition 1.6 that division with rest holds in $\mathbb{N}$*. For* $a, b \in \mathbb{N}$ *with* $b \neq 0$*, there exist* $q, r \in \mathbb{N}$ *with* $a = qb + r$ *and* $0 \leq r < b$*. Hence also* $-a = (-q)b - r$ *holds and either* $r = 0$*, i.e.* $-a = (-q)b$*, or* $r \neq 0$ *and then* $0 < b - r < b$ *and* $-a = (-q-1)b + (b-r)$*. This shows that division with rest holds also for negative dividends. If the divisor is negative, then from* $a = qb + r$ *we get* $a = (-q)(-b) + r$ *and* $0 \leq r < b = |-b|$*.*

Lemma 1.13 (Division with rest in $\mathbb{Z}$***)*** *Let* $a, b \in \mathbb{Z}$ *with* $b \neq 0$*. Then there exist* $q, r \in \mathbb{Z}$ *such that*

$$a = qb + r \qquad \text{with } 0 \leq r < |b|. \tag{1.9}$$

The rest r *of* a *divided by* b *shall also be denoted by* $r = a \bmod b$*. The quotient* q *will be denoted by* a/b*.*

Divisibility in $\mathbb{Z}$

As in $\mathbb{N}$ *we say that an integer* d *divides an integer* $n \in \mathbb{Z}$ *if there exists* $e \in \mathbb{Z}$ *with* $de = n$*. We write* $d \mid n$*. As before, if* $d \mid 0$*, for any* d *and* $0 \mid n$ *if and only if* $n = 0$*. One of the differences between* $\mathbb{N}$ *and* $\mathbb{Z}$ *is that we cannot conclude that a divisor is necessarily smaller than a non-zero number it divides. For instance* $d = 2$ *divides* $n = -6$*, because* $-6 = 2 \cdot (-3)$*, but* $d > n$*. However, the divisor is always less than the norm of the non-zero number it divides.*

Lemma 1.14 *Let* $d, n \in \mathbb{Z} \setminus \{0\}$*.*

1. *If* $d \mid n$ *then* $-|n| \leq d \leq |n|$*.*
2. *If* $d \mid 1$ *then* $d = 1$ *or* $d = -1$*.*
3. *If* $d \mid n$ *and* $n \mid d$ *then* $d = n$ *or* $d = -n$*.*

Proof: (1) If $d \mid n$, $n < 0$ and $d > 0$, then $n = de$ and $e < 0$. Otherwise, if $e > 0$, then $n = de > 0$. Hence $(-n) = d(-e)$ and we conclude $-|n| < 0 < d \leq |-n| = |n|$. The other cases are easy.

(2) By (1), $-1 \leq d \leq 1$. Since $d \neq 0$, we are left with $d = 1$ or $d = -1$.

(3) If $d \mid n$ and $n \mid d$, then also $|d| \mid |n|$ and $|d| \mid |n|$. By Lemma 1.5, $|d| = |n|$. Hence, if $\text{sign}(d) = \text{sign}(n)$, then $d = n$ and otherwise $d = -n$. □

While $\mid$ is an equivalence relation on $\mathbb{N}$, we see from Lemma 1.14(3) that $\mid$ on $\mathbb{Z}$ is not anti-symmetric and that for each non-zero integer $n \in \mathbb{Z}$ the equivalence class of n contains exactly two elements $[n]_\sim = \{n, -n\}$.

The extended Euclidean Algorithm

A common divisor of two integers a and b is as defined as in $\mathbb{N}$ as an elements $d \in \mathbb{Z}$ such that $d \mid a$ and $d \mid b$. We can then define a common divisor d of two integers a and b (not both zero) to be a greatest common divisor of two integers if for any common divisor e of a and b one has $d \mid e$. Since $\mid$ is not anymore an equivalence relation, there are precisely two greatest common divisors, namely d and $-d$. By definition, we define the greatest common divisor of a and b to be the positive greatest common divisor and denote it by $\gcd(a, b)$ The greatest common divisor of two integers (not both zero) exists and the so-called (extended) Euclidean Algorithm calculates it.

Proposition 1.15 *Let $a, b \in \mathbb{Z}$ not both zero. Then $\gcd(a, b)$ exists and there are $r, s \in \mathbb{Z}$ such that $\gcd(a, b) = ra + sb$.*

Proof: Let $U = \{d \in \mathbb{N}^+ : \exists r, s \in \mathbb{Z} : \ d = ra + sb\}$. Then $U \subset \mathbb{N}^+$ is non-zero, since $a = 1 \cdot a + 0 \cdot b$ belongs to U if $a \neq 0$. By the well-ordering of $\mathbb{N}^+$, U has a minimum, say $d \in U$ with $r, s \in \mathbb{Z}$ such that $d = ra + sb$. By the division algorithm, there exist $q \in \mathbb{Z}$ and $t \in \mathbb{N}$ such that $a = qd + t$ and $0 \leq t < d$. If $t \neq 0$, then $t = a - qd = a - qra - qsb = (1 - qr)a + (-q)sb \in U$ contradicts the minimality of d. Thus $t = 0$ and $d \mid a$. Similarly one proves $d \mid b$. Suppose e is a common divisor of a and b with $a = ea'$ and $b = eb'$. Then $d = rea' + seb' = (ra' + sb')e$ shows $e \mid d$. Hence $d = \gcd(a, b)$. □

From the proof of Proposition 1.15 we have

Theorem 1.16 *Let $a, b \in \mathbb{Z}$ not both zero. Then $d \in \mathbb{N}^+$ is the greatest common divisor of $a$and b if and only if d is the least positive integer such that there are integers r and s with $d = ra + sb$.*

The extended Euclidean Algorithm calculates those elements r, s. We present here a recursive version for the integers, which at its heart uses Proposition 1.9, namely that $\gcd(a, b) = \gcd(b, a \bmod b)$ and the following reasoning: suppose

$$\gcd(b, a \bmod b) = x \cdot b + y \cdot (a \bmod b)$$

for some $x, y \in \mathbb{Z}$. Since $a = (a/b) \cdot b + (a \bmod b)$ it follows that $a \bmod b = a - (a/b) \cdot b$. Substituing this expression for $a \bmod b$ in the formula for $\gcd(b, a \bmod b)$ yields:

$$\gcd(a, b) = \gcd(b, a \bmod b) = x \cdot b + y \cdot (a - (a/b) \cdot b) = y \cdot a + (x - (a/b) \cdot y) \cdot b. \qquad (1.10)$$

Hence on the "way back" from our recursion, we can adjust the coefficients x and y by replacing x with y and y with $x - (a/b) \cdot y$.

Data: $a, b \in \mathbb{Z}$
Result: d, x, y *where* $d = \gcd(a, b)$ *and* $d = x \cdot a + y \cdot b$.
if $b \neq 0$ **then**
  d,x,y = gcd(b, a mod b);
  return (d, y, x-(a/b)y);
else
  return (|a|,1,0);
end

Algorithm 1: Extended Euclidean Algorithm

Example 1.17 *For instance if* $a = 144$ *and* $b = 80$, *then*

$$\begin{aligned} 144 &= 1 \cdot 80 + 64 && \Rightarrow 80 = 80 - (144 - 80) = (-1) \cdot 144 + 2 \cdot 80 \\ 80 &= 1 \cdot 64 + 16 && \Rightarrow 16 = 80 - 64 \\ 64 &= 4 \cdot 16 + 0 \end{aligned}$$

The following algorithm does not use recursion:

Data: $a, b \in \mathbb{Z}$ *not both zero.*
Result: d, x, y *where* $d = \gcd(a, b)$ *and* $d = x \cdot a + y \cdot b$.
$x = 1; v = 1; y = 0; u = 0;$
while $b \neq 0$ **do**
  $h = x;\ x = u;\ u = h - (a/b) \cdot u;$
  $h = y;\ y = v;\ v = h - (a/b) \cdot v;$
  $h = a;\ a = b;\ b = h \mod b;$
end
return x *e* y

As an example, take again the numbers $a = 144$ *e* $b = 81$. *The non-recursive algorithm would then work as follows:*

a	b	x	u	y	u	
144	*81*	*1*	*0*	*0*	*1*	*144 = 1*81 + 63*
81	*63*	*0*	*1*	*1*	*-1*	*81 = 1*63 + 18*
63	*18*	*-1*	*1*	*-1*	*2*	*63 = 3*18 + 9*
18	*9*	*-1*	*4*	*2*	*-7*	*18 = 2*9 + 0*
9	*0*	*4*	*-9*	*-7*	*16*	*algorithm stops*

Two integers a, b *are called relatively prime if* $\gcd(a, b) = 1$. *As a Corollary from Theorem 1.16 we conclude:*

Corollary 1.18 *Let* $a, b \in \mathbb{Z}$ *not both zero.*

1. *If there exist* $r, s \in \mathbb{Z}$ *such that* $1 = ra + sb$, *then* a *and* b *are relatively prime.*

2. *If* $a = a' \gcd(a,b)$ *and* $b = b' \gcd(a,b)$, *then* a' *and* b' *are relatively prime.*

3. *If* a *and* b *are relatively prime and* $c \in \mathbb{Z}$ *such that* $a \mid bc$ *then* $a \mid c$.

P*roof: (1) follows directly from Theorem 1.16 as* 1 *is the minimum of* $\mathbb{N}^+$.

(2) Since there are $r, s \in \mathbb{Z}$ *such that* $\gcd(a,b) = ra + sb = \gcd(a,b)(ra' + sb')$, *we have that* $1 = ra' + sb'$ *and therefore* a' *and* b' *relatively prime.*

(3) By Proposition 1.15 there exist $r, s \in \mathbb{Z}$ *such that* $1 = \gcd(a,b) = ra + sb$. *Hence* $c = rac + sbc$. *If* $a \mid bc$, *then* $a \mid sbc$ *and hence* $a \mid sbc + rac = c$. □

Prime integers

An integer $p \in \mathbb{Z}$ *is called a prime integer if* $|p| > 1$ *and whenever* $p \mid ab$, *then* $p \mid a$ *or* $p \mid b$, *for any* $a, b \in \mathbb{Z}$.

Lemma 1.19 *An integer* $p \in \mathbb{Z}$ *is prime if and only if* $p \notin \{-1, 0, 1\}$ *and the only divisors of* p *are* $1, -1, p, -p$. *In particular any prime number is also a prime integer.*

P*roof: Let* p *be a prime integer. Then* $|p| > 1$ *means that* $p \notin \{-1, 0, 1\}$. *Moreover, if* $p = ab$ *for some divisors* $a, b \in \mathbb{Z}$. *Then* $p \mid a$ *or* $p \mid b$. *As* $a \mid p$ *and* $b \mid p$, *we have that if* $p \mid a$ *then* $a = p$ *or* $a = -p$ *by Lemma 1.14. As* $p = ab = \pm pb$, *we conclude* $b = \pm 1$. *The case* $p \mid b$ *is treated analogously.*

On the other hand, suppose that $p \notin \{-1, 0, 1\}$ *and that* ± 1 *and* $\pm p$ *are the only divisors of* p. *Suppose* $p \mid ab$. *Then either* $\gcd(a,p) = 1$ *or* $\gcd(a,p) = p$. *In latter case* $p \mid a$. *Hence suppose* $\gcd(a,p) = 1$. *Then* a *and* p *are relatively prime and by Corollary 1.18* $p \mid b$.

If p *is a prime number, then* $p > 1$ *and the only divisors (in* $\mathbb{N}$*) are* 1 *and* p. *If* d *is any other integer divisor of* p, *then* $|d| \in \{1, p\}$ *and hence* $d \in \{\pm 1, \pm p\}$. *Thus* p *is an integer prime.* □

Theorem 1.20 (Unique factorization) *Let* $n \in \mathbb{Z} \setminus \{0, 1, -1\}$. *Then there exist prime integers* $p_1, \ldots, p_s$ *such that* $n = p_1 \cdot \cdots \cdot p_s$ *and for any other decomposition* $n = q_1 \cdots q_t$ *with* q_i *prime integer one has* $s = t$ *and there exists a permutation* σ *such that* $q_i = \pm p_{\sigma(i)}$, *for all* $1 \leq i \leq s$.

P*roof: Let* $n > 1$. *Then by Theorem 1.11 there exist prime numbers* $p_1, \ldots, p_s$ *such that* $n = p_1 \cdot \cdots \cdot p_s$. *We will prove the uniqueness of* s *and the factors* p_i *(up to a sign) by induction on* s. *Set* $M = \{n \in \mathbb{N}_{>1} \mid$ *the statement of the Theorem is true for* $n\}$.

Any prime number p *belongs to* M, *because* p *is also a prime integer and if* $p = q_1 \cdots q_t$ *for some prime integers* q_i, *then* $q_i \in \{\pm p\}$. *But then* $t = 1$, *since otherwise if* $t > 1$ *we had* $p = \pm p^t$ *and hence* $1 = \pm p^{t-1}$, *which would mean* $p = \pm 1$ *by Lemma 1.14. Thus* $p \in M$. *In particular* $2 \in M$.

Now suppose $n > 1$ *and* $m \in M$, *for all* $m < n$. *If* n *is a prime number, then* $n \in M$. *If* n *is not a prime number, then* $n = p_1 \cdot \cdots \cdot p_s$. *Suppose* $n = q_1 \ldots q_t$ *for some prime integers* q_i. *Since* p_s *is also a prime integer and* $p\_smidn = q_1 \cdots q_t$, *we must have* $p_s \mid q_i$ *for some* i. *Without loss of generality we can assume* $i = t$. *Since the only divisors of* q_i *are* ± 1 *and* $\pm q_i$ *and since* $p_s > 1$, *we have* $q_t = \pm p_s$. *Thus* $m = p_1 \cdots p_{s-1} = n/p_s = (\pm q_1) q_2 \cdots q_{t-1}$. *By complete induction hypothesis,* $s - 1 = t - 1$, *i.e.* $s = t$ *and up to a permutation* σ, $q_i = p_{\sigma(i)}$,

for all $1 \leq i \leq s$, i.e. $n \in M$. By complete induction, $M = \mathbb{N}_{>1}$, i.e. the Theorem is valid for all $n > 1$.

For $n < -1$, we replace n with $-n$ and hence $-n = p_1 \dots p_s$ for some prime numbers p_i which are unique up to permutation. Hence $n = (-p_1)p_2 \cdots p_s$. □

Unique factorization is a powerful tool. In the language of algebra, $\mathbb{Z}$ is an ordered integral domain. We will extend most of this chapter to the general setting of an Euclidean domain, but the principal idea is always the same.

The Rational numbers

Rational numbers are defined as equivalence classes of pairs of integers. Define on $\mathbb{Z} \times \mathbb{N}^+$ the following equivalence relation:

$$(a, b) \sim (c, d) \Leftrightarrow ad = bc \qquad \forall (a, b), (c, d) \in \mathbb{Z} \times \mathbb{N}^+. \tag{1.11}$$

The equivalence class of an element $(a, b) \in \mathbb{Z} \times \mathbb{N}^+$ is denoted by $\frac{a}{b}$ and consists of the following subset of $\mathbb{Z} \times \mathbb{N}^+$:

$$\frac{a}{b} := [(a, b)]_\sim = \{(c, d) \in \mathbb{Z} \times \mathbb{N}^+ \mid ad = bc.\} \tag{1.12}$$

The set of equivalence classes $(\mathbb{Z} \times \mathbb{N}^+) / \sim$ is denoted by $\mathbb{Q}$ and called the rational numbers. Note that $a/a = 1/1$ for any $a > 0$ and $0/b = 0/1$ for any $b > 0$. Moreover, any non-zero fraction a/b, i.e. $a/b \neq 0/1$, can be represented by a reduced fraction $a/b = c/d$, with c and d relatively prime. Because for a/b with $a \neq 0 \neq b$ and $d = \gcd(a, b)$, we have $a = da'$ and $b = db'$, for some $a' \in \mathbb{Z}$ and $b' \in \mathbb{N}^+$. Then $\gcd(a', b') = 1$ by Corollary 1.18 shows that a' and b' are relatively prime. Thus $ab = d(ab') = d(a'b)$ implies $ab' = a'b$ and therefore $a/b = a'/b'$.

It is tedious, but straightforward that with the following operations $\mathbb{Q}$ becomes a field:

$$\frac{a}{b} + \frac{c}{d} = \frac{a \cdot d + b \cdot c}{c \cdot d} \qquad \text{and} \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d};$$

The integers $\mathbb{Z}$ embed into $\mathbb{Q}$ through the map $n \mapsto \frac{n}{1}$ and form a subring of $\mathbb{Q}$. Note that 0 maps to $\frac{0}{1} = \frac{0}{b}$ for any $b \in \mathbb{N}^+$. One defines an order relation on $\mathbb{Q}$ as follows: Let $\frac{a}{b} \in \mathbb{Q}$. Then

$$\frac{a}{b} \geq 0 \qquad \Leftrightarrow \qquad ab \geq 0. \tag{1.13}$$

For any $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$ one sets

$$\frac{a}{b} \leq \frac{c}{d} \qquad \Leftrightarrow \qquad 0 \leq \frac{c}{d} + \frac{-a}{b}.$$

The Real numbers

The real numbers are built upon the rational numbers. A sequence of rational numbers $\mathbb{Q}$ is any function $f : \mathbb{N} \to \mathbb{Q}$, i.e. $f \in \mathbb{N}^{\mathbb{Q}}$. We also use the notation $(a_n)_{n \in \mathbb{N}}$ for the function f if it is understood that $a_n = f(n)$ for all $n \in \mathbb{N}$. A sequence (a_n) of rational numbers is called a Cauchy sequence if for any $q \in \mathbb{Q}_{>0}$ there exists $N \in \mathbb{N}$ such that

$$\forall n, m \geq N : \qquad |a_n - a_m| < q. \tag{1.14}$$

A Cauchy sequence (a_n) is said to converge against 0 if for every $q \in \mathbb{Q}_{>0}$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$: $|a_n| < q$.

The set of Cauchy sequences R is a ring with addition and multiplication as follows for all $(a_n), (b_n) \in R$:

$$(a_n) + (b_n) = (a_n + b_n) \qquad (a_n) \cdot (b_n) = (a_n \cdot b_n) \tag{1.15}$$

Defining the equivalence relation $(a_n) \sim (b_n)$ if and only if $(a_n - b_n)$ converges to 0, we obtain the set of equivalence classes

$$\mathbb{R} = R/\sim \tag{1.16}$$

called the field of real numbers. Hence each real number is the equivalence class of a Cauchy sequence of rational numbers. It can be shown that the total order on $\mathbb{Q}$ yields a total order on $\mathbb{R}$.

The complex numbers

The complex numbers are build upon the real numbers as $\mathbb{C} = \mathbb{R} \times \mathbb{R}$. An element $(a, b) \in \mathbb{C}$ is denoted by $a + b\imath$ and while addition is the same as the componentwise addition on $\mathbb{R} \times \mathbb{R}$, its multiplication is defined as

$$(a + b\imath)(c + d\imath) := (ac - bd) + (ad + bc)\imath, \qquad a, b, c, d \in \mathbb{R}.$$

The conjugate of a complex number $\omega = a + b\imath$ is $\overline{\omega} = a - b\imath$. And $\omega\overline{\omega} = a^2 + b^2 =: ||\omega||^2 \in \mathbb{R}$ is called the square of the norm of ω. In particular, $\omega = 0$ if and only if $a = b = 0$ if and only if $||\omega||^2 = 0$. Thus, if $\omega \neq 0$, then $\omega^{-1} = \frac{1}{||\omega||^2}\overline{\omega}$ is the (multiplicative) inverse of ω in $\mathbb{C}$. Hence $\mathbb{C}$ is a field. However, in contrast to $\mathbb{R}$, $\mathbb{C}$ is not an ordered field.

Exercises

Ex. 1 — *Using the Peano axioms, show that if $a, b \in \mathbb{N}$ such that $ab = 1$, then $a = b = 1$.*

Ex. 2 — *One can also define a partial subtraction on $\mathbb{N}$ as follows. Let $n, m \in \mathbb{N}$ and define*

$$n \dot{-} m = \begin{cases} k & \text{if } n = m + k \\ 0 & \text{else} \end{cases} \tag{1.17}$$

That means that $n \dot{-} m = k$ if $m \leq n$ and $n + k = m$, while $n \dot{-} m = 0$ if $m \not\leq n$. Prove that for any $m \leq n$ one has $n = m + (n \dot{-} m)$.

Ex. 3 — *Recall the equivalence relation on $\mathbb{N}^2$ that defines $\mathbb{Z}$: $(a, b) \sim (c, d)$ if and only if $a + d = b + c$, for all $(a, b), (c, d) \in \mathbb{N}^2$. Show that addition and multiplication defined on $\mathbb{Z} = \mathbb{N}^2/\sim$ are well-defined:*

$$\begin{aligned} [(a, b)] + [(c, d)] &:= [(a + c, b + d)] \\ [(a, b)] \cdot [(c, d)] &:= [(ac + bd, ad + bc)] \end{aligned}$$

Ex. 4 — *Let $A = \mathbb{N}_{>0}$ be the set of natural numbers and set $a \mid b$ if and only if a divides b. Then $\mid$ is a partial ordering, but not a total one.*

Ex. 5 — *Prove the associativity and distributivity of the arithmetic operations of* $\mathbb{N}$.

Ex. 6 — *Prove the following statements for* $n, m, k \in \mathbb{N}$*:*

1.$n \neq S(n)$.

2.If $n + k = n + m$ *then* $k = m$.

3.If $n \leq m \leq S(n)$, *then* $n = m$ *or* $m = S(n)$

2

Some arithmetic functions

The Euler totient function is counting the positive numbers that are relatively prime to a given number:

$$\varphi : \mathbb{N}^+ \to \mathbb{N}^+ \qquad n \longmapsto \varphi(n) = |\{a \in \{1, \ldots, n\} \mid \gcd(a, n) = 1\}|.$$

The Euler function counts not just the numbers between 1 and n that are relatively prime with n but also how many generators there are in a cyclic group of order n. We recall that the order of an element x in a group G is the least positive integer n such that $x^n = 1$ is the identity element of G. In particular if $x^m = 1$ for some $m \geq 1$, then the division algorithm will tell us that $m = qn + r$, for $0 \leq r < n$. Hence $1 = x^m = (x^n)^q x^r = x^r$ and $r < n$ shows that $r = 0$. Hence $n \mid m$.

Lemma 2.1 *Let $n > 1$ and let $C_n = \langle \omega \rangle$ be a (multiplicative) cyclic group of order n. Then*

$$\varphi(n) = |\{y \in C_n : y \text{ is a generator for } C_n\}| \tag{2.1}$$

P*roof: Let $A = \{k : 0 < k < n, \gcd(k, n) = 1\}$ and $B = \{y \in C_n : y \text{ is a generator for } C_n\}$. We show that the mapping $f : A \to B$ with $f(k) = \omega^k$ is a bijection. Let $k \in A$ and consider $y = f(k) = \omega^k$. By the extended Euclidean Algorithm, there exist $s, t \in \mathbb{Z}$ such that $1 = sk + tn$. Thus*

$$y^s = \omega^{sk} = \omega^{1-tn} = \omega(\omega^n)^{-t} = \omega. \tag{2.2}$$

This shows that $\omega \in \langle y \rangle$ and therefore $C_n = \langle y \rangle$. Let $k_1, k_2 \in A$ such that $f(k_1) = f(k_2)$. Without loss of generality we can assume $k_1 \geq k_2$, then

$$\omega^{k_1} = f(k_1) = f(k_2) = \omega^{k_2} \Rightarrow \omega^{k_1 - k_2} = 1. \tag{2.3}$$

Since ω has order n, $n \mid k_1 - k_2$. However, $0 \leq k_1 - k_2 < n$. Thus $k_1 = k_2$, showing that f is injective. To prove surjectivity, let $y \in B$. Then $y = \omega^m$ for some $m \in \mathbb{Z}$. By the division

algorithm we can divide m by n and obtain $m = qn + k$, for $0 \leq k < n$. Since $\omega^n = 1$, $\omega^{qn} = 1$ and hence $y = \omega^m = \omega^{qn}\omega^k = \omega^k$. If $k = 0$, then $y = 1$ would be a generator, meaning $C_n = \langle 1 \rangle = \{1\}$, which is absurd as $n > 1$. Thus $0 < k < n$. Suppose $\gcd(k, n) = d$. Then there are $a, b \in \mathbb{N}$ such that $k = ad$ and $n = bd$. Hence

$$(\omega^k)^b = \omega^{adb} = \omega^{na} = 1 \tag{2.4}$$

and as the order of $y = \omega^k$ is n, we must have $n \mid b$. But then $n = bd$ and $n \mid b$ implies $1 \mid d$, i.e. $d = 1$. Therefore $k \in A$ and $y = f(k)$. □

How can be calculate $\varphi(n)$? In case $n = p$ is a prime number, then $\gcd(a, p) = 1$ for any non-zero element $1 \leq a < p-1$. Hence $\varphi(p) = p-1$. How to calculate $\varphi(n)$ for an arbitrary n? In order to answer this question we will first of all look at the set of all arithmetic functions.

Denote by $\mathrm{Fun}(\mathbb{N}^+, \mathbb{C})$ the set of all functions $f : \mathbb{N}^+ \to \mathbb{C}$. The Dirichlet convolution on $\mathrm{Fun}(\mathbb{N}^+, \mathbb{C})$ is define as follows. For any two functions $f, g : \mathbb{N}^+ \to \mathbb{C}$ let $f \bullet g$ be the function defined by

$$(f \bullet g)(n) := \sum_{d|n} f(d)g(n/d) = \sum_{n=de} f(d)g(e)$$

where the sum runs over all positive divisors d of n and where n/d denotes the quotient of n divided by n which is a number.

The product is associative, because for any $f, g, h \in \mathrm{Fun}(\mathbb{N}^+, \mathbb{C})$ and $n \in \mathbb{N}^+$:

$$\begin{aligned}
((f \bullet g) \bullet h)(n) &= \sum_{d|n} (f \bullet g)(d)\, h(n/d) \\
&= \sum_{d|n} \sum_{e|n/d} f(e)\, g(d/e)\, h(n/d) \\
&= \sum_{e_1 e_2 e_3 = n} f(e_1)\, g(e_2)\, h(e_3) \\
&= \sum_{e|n} \sum_{d|n/e} f(e)\, g(d)\, h(n/ed) = \sum_{e|n} f(e)(g \bullet h)(n/e) = (f \bullet (g \bullet h))(n)
\end{aligned}$$

Therefore, $(f \bullet g) \bullet h = f \bullet (g \bullet h)$. The operation is also commutative, because

$$(f \bullet g)(n) = \sum_{de=n} f(d)\, g(e) = \sum_{ed=n} g(e)\, f(d) = (g \bullet f)(n).$$

Furthermore, the Dirichlet product has also a neutral element, namely the function

$$\epsilon(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{else} \end{cases}$$

Then $(f \bullet \epsilon)(n) = \sum_{d|n} f(d)\epsilon(n/d) = f(n)\epsilon(1) = f(n)$, for all $n \in \mathbb{N}^+$. In other words, $(\mathrm{Fun}(\mathbb{N}^+, \mathbb{C}), \bullet, e)$ is a commutative monoid.

The function $\mathbf{1} : \mathbb{N}^+ \to \mathbb{C}$ with $\mathbf{1}(n) = 1$ for all $n \in \mathbb{N}^+$ is invertible in this monoid and its inverse function is the Möbius function μ, which we will define now. Let us call a number $n \in \mathbb{N}_{>1}$ square free if whenever a prime number p divides n, then p^2 does not divide n. This means that if $n = p_1 \cdots p_s$ is a decompositon of n into prime numbers, all pimes p_i are different. The Möbius function is defined as follows:

$$\mu(n) := \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n \text{ is not square free} \\ (-1)^s & \text{if } n = p_1 \cdots p_s \text{ with } p_i \neq p_j \text{ for } i \neq j \end{cases}$$

Lemma 2.2 *$\sum_{d|n} \mu(d) = 0$, for any $n > 1$.*

Proof: If $d \mid n$ is not square free, then $\mu(d) = 0$. Hence we only need to consider $d \mid n$ that are square free. Let $\{p_1, \ldots, p_s\}$ be all distinct prime divisors of n. Then a square free divisor d of n is of the form $d = p_1^{e_1} \cdots p_s^{e_s}$, for $e_i \in \{0, 1\}$.

There are precisely $\binom{s}{i}$ choices for square free divisors d of n that are products of i distinct primes, in which case $\mu(d) = (-1)^i$. Hence

$$\sum_{d|n} \mu(d) = \sum_{(e_1,\ldots,e_s)\in\{0,1\}^s} \mu(p_1^{e_1} \cdots p_s^{e_s}) = \sum_{i=0}^{s} \binom{s}{i} (-1)^i = (-1+1)^s = 0.$$

□

Corollary 2.3 *μ is the inverse of $\mathbf{1}$ in the monoid $(\mathrm{Fun}(\mathbb{N}^+, \mathbb{C}), \bullet, e)$.*

Proof: For $n = 1$ we have $(\mu \bullet \mathbf{1})(1) = \mu(1)\mathbf{1}(1) = 1$. For $n > 1$ we get

$$(\mu \bullet \mathbf{1})(n) = \sum_{d|n} \mu(d)\mathbf{1}(n/d) = \sum_{d|n} \mu(d) = 0.$$

Hence $\mu \bullet \mathbf{1} = \epsilon$. □

As a consequence we obtain the Möbius inversion formula:

Theorem 2.4 (Möbius inversion formula) *Let $f \in \mathrm{Fun}(\mathbb{N}^+, \mathbb{C})$ and define $F \in \mathrm{Fun}(\mathbb{N}^+, \mathbb{C})$ by $F(n) = \sum_{d|n} f(d)$, for all $n \in \mathbb{N}^+$. Then*

$$f(n) = \sum_{d|n} \mu(d)F(n/d), \qquad \forall n \in \mathbb{N}^+.$$

Proof: We have $F = f \bullet \mathbf{1}$. Thus $f = \epsilon \bullet f = \mu \bullet \mathbf{1} \bullet f = \mu \bullet f \bullet \mathbf{1} = \mu \bullet F$. □

Returning to the Euler function φ, we first prove that $\varphi \bullet \mathbf{1} = id$ is the identity function $id(n) = n$.

Lemma 2.5 *$\sum_{d|n} \varphi(d) = n$, for all $n \in \mathbb{N}^+$.*

Proof: For $d \mid n$, $\varphi(d)$ counts the numbers $1 \leq k \leq d$ with $\gcd(k, d) = 1$. Let

$$\begin{aligned} A &:= \left\{(k, d) \in \mathbb{N}^2 \mid d \mid n,\ 1 \leq k \leq d,\ \gcd(k, d) = 1\right\} \\ &= \bigcup_{d|n} \left\{(k, d) \in \mathbb{N}^2 : 1 \leq k \leq d, \gcd(k, d) = 1\right\}. \end{aligned}$$

Then $|A| = \sum_{d|n} \varphi(d)$. *We claim that the map* $f : A \to B := \left\{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n}{n}\right\} \subseteq \mathbb{Q}$ *with* $f(k,d) = \frac{k}{d}$ *is a bijection between* A *and* B *and therefore* $\sum_{d|n} \varphi(d) = |A| = |B| = n$. *First of all* f *is well-defined, because if* $d \mid n$, *then* $n = de$ *for some* $e \in \mathbb{N}$. *Hence* $\frac{k}{d} = \frac{ke}{n} \in B$, *because* $1 \le k \le d$ *implies* $1 \le ke \le n$. f *is surjective, because any fraction* $\frac{a}{n}$. *with* $1 \le a \le n$, *can be reduced to* $\frac{k}{d}$ *with* $\gcd(k,d) = 1$, $1 \le k \le d$ *and* $d \mid n$. *The function* f *is injective, because if* $f(k,d) = f(k',d')$ *for* $(k,d),(k',d') \in A$. *Then* $kd' = dk'$. *Since* k *and* d *are relatively prime,* $k \mid k'$ *and analogously* $k' \mid k$. *Thus* $k = k'$ *and* $d = d'$. *Thus* f *is a bijection as claimed.* □

Corollary 2.6 *Let* $n = p_1^{a_1} \cdots p_s^{a_s}$ *with* $p_i \neq p_j$ *for* $i \neq j$, $a_i \ge 1$ *and* $s \ge 1$. *Then*

$$\varphi(n) = \prod_{i=1}^{s} \left(p_i^{a_i} - p_i^{a_i - 1}\right) = n \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right).$$

P*roof: Lemma 2.5 says that* $\varphi \bullet \mathbf{1} = id$ *and the Möbius inversion formula says that*

$$\varphi(n) = \sum_{d|n} \mu(d) \frac{n}{d}, \qquad \forall n \in \mathbb{N}^+.$$

Let $n = p^a m$ *and* $p \nmid m$. *Note that if* $d \mid n$, *then either* $p \nmid d$, *in which case* $d \mid m$ *or* $p \mid d$. *In the later case,* $\mu(d) = 0$ *if* $p^2 \mid d$ *and if* $p^2 \nmid d$, *then* $d = pd'$ *with* $p \nmid d'$ *and* $d' \mid m$. *Thus* $\mu(d) = -\mu(d')$.

$$\varphi(n) = \sum_{d|p^a m} \mu(d) \frac{p^a m}{d} = p^a \sum_{d|m} \mu(d) \frac{m}{d} - p^{a-1} \sum_{d'|m} \mu(d') \frac{m}{d'} = \left(p^a - p^{a-1}\right) \varphi(m).$$

Hence by induction the result follows. □

Example: *How many positive numbers less than* 712 *are relatively prime* 712*? Since* $712 = 2^3 \cdot 89$, *we have* $\varphi(712) = (2^3 - 2^2) \cdot (89 - 1) = 352$.

Arithmetic functions from prime decomposition

Let $\mathbb{P}$ *denote again the set of positive prime numbers and define the following arithmetic functions:*

$$\nu_p : \mathbb{N}^+ \to \mathbb{N}. \qquad \nu_p(n) = \max\{k : p^k \mid n\}, \qquad \text{for } p \in \mathbb{P}$$

Clearly we have $\nu_p(ab) = \nu_p(a) + \nu_p(b)$ *and also* $\nu_p(a/b) = \nu_p(a) - \nu_p(b)$ *if* $b \mid a$.

Lemma 2.7 *For any* $p \in \mathbb{P}$ *and* $n \ge 1$ *one has* $\nu_p(n!) = \sum_{k=1}^{\nu_p(n)} \left\lfloor \frac{n}{p^k} \right\rfloor$.

P*roof: For numbers* $m, k \in \mathbb{N}^+$ *set*

$$\chi(m,k) := \begin{cases} 1 & \text{if } p^k \mid m \\ 0 & \text{else} \end{cases}$$

Then $\nu_p(m) = \sum_{k\geq 1} \chi(m,k)$. *Using this technical tool we obtain*

$$\nu_p(n!) = \sum_{m=1}^{n} \nu_p(m) = \sum_{m=1}^{n}\sum_{k\geq 1} \chi(m,k) = \sum_{k\geq 1}\sum_{m=1}^{n} \chi(m,k) = \sum_{k=1}^{\nu_p(n)} \left\lfloor \frac{n}{p^k} \right\rfloor,$$

because $\sum_{m=1}^{n} \chi(m,k)$ *counts how many numbers* $1 \leq d \leq n$ *are multiples of* p^k. *If* $q = \left\lfloor \frac{n}{p^k} \right\rfloor$, *then* $p^k, 2p^k, \ldots, qp^k$ *are precisely these numbers* d *and there are* $\left\lfloor \frac{n}{p^k} \right\rfloor$ *many of them.* □

This means for instance that

$$\begin{aligned} \nu_2(100!) &= \frac{100}{2} + \frac{100}{4} + \left\lfloor \frac{100}{8} \right\rfloor + \left\lfloor \frac{100}{16} \right\rfloor + \left\lfloor \frac{100}{32} \right\rfloor + \left\lfloor \frac{100}{64} \right\rfloor \\ &= 50 + 25 + 12 + 6 + 3 + 1 = 97 \end{aligned}$$

Hence $2^{96} \mid 100!$.

A number $n \in \mathbb{N}_{>1}$ *is called square-free if for all* $c > 1$ *such that* $c \mid n$, $c^2 \nmid n$. *In other words* n *is square-free if and only if* $\nu_p(n) \leq 1$ *for all* $p \in \mathbb{P}$.

Two more important arithmetic functions are the following:

$\eta : \mathbb{N}^+ \to \mathbb{N}^+$,	$\eta(n) = \lvert\{d \in \mathbb{N} : d \mid n\}\rvert$,	*número de divisores positivos*
$\sigma : \mathbb{N}^+ \to \mathbb{N}^+$,	$\sigma(n) = \sum_{d\mid n} d$,	*soma dos divisores positivos*

The divisor sum function σ *is multiplicative in the following sense. A function* $f : \mathbb{N}^+ \to \mathbb{C}$ *is a multiplicative function if* $f(1) = 1$ *and for all* $a, b \in \mathbb{N}^+$ *that are relatively prime* $f(ab) = f(a)f(b)$ *holds. For example the identity function* id, *the constant* $\mathbf{1}$ *function and the neutral element* e *of* $\mathrm{Fun}(\mathbb{N}^+, \mathbb{C})$ *are multiplicative. Moreover, given two multiplicative function, also their Dirichlet product is multiplicative and so is* σ, *since* $\sigma = id \bullet \mathbf{1}$. *To prove this claim, let* f, g *be multiplicative, then clearly* $f \bullet g(1) = f(1)g(1) = 1$ *and for* $a, b \in \mathbb{N}^+$ *relatively prime, we get*

$$\begin{aligned} (f \bullet g)(ab) &= \sum_{d\mid ab} f(d)g\left(\frac{ab}{d}\right) \\ &= \sum_{d_1\mid a, d_2\mid b} f(d_1d_2)g\left(\frac{ab}{d_1d_2}\right) \\ &= \sum_{d_1\mid a}\sum_{d_2\mid b} f(d_1)g\left(\frac{a}{d_1}\right) f(d_2)g\left(\frac{b}{d_2}\right) \\ &= \left(\sum_{d_1\mid a} f(d_1)g\left(\frac{a}{d_1}\right)\right)\left(\sum_{d_2\mid b} f(d_2)g\left(\frac{b}{d_2}\right)\right) = (f \bullet g)(a)(f \bullet g)(b) \end{aligned}$$

The set of multiplicative function is actually a group.

Proposition 2.8 *Let $n \in \mathbb{N}_{>1}$ such that $n = p_1^{a_1} \cdots p_t^{a_t}$, with $p_i \in \mathbb{P}$, $p_i \neq p_j$ and $a_i \geq 1$. Then the following hold:*

1. *There exists a square free number a and a number b such that $n = ab^2$.*
2. *$\eta(n) = (a_1 + 1)(a_2 + 1) \cdots (a_t + 1) = \prod_{i=1}^{t} (\nu_{p_i} + 1)$.*
3. *$\sigma(n) = \left(\frac{p_1^{a_1+1}-1}{p_1-1}\right)\left(\frac{p_2^{a_2+1}-1}{p_2-1}\right)\cdots\left(\frac{p_t^{a_t+1}-1}{p_t-1}\right) = \prod_{i=1}^{t}\left(\frac{p_i^{\nu_{p_i}+1}-1}{p_i-1}\right)$.*

P*roof: (1) We have $n = \prod_{i=1}^{t} p_i^{\nu_{p_i}}$. For any $1 \leq i \leq t$ we have $\nu_{p_i}(n) = 2r_i + s_i$, onde $s_i \in \{0, 1\}$. Hence, if $a = \prod_{i=1}^{t} p_i^{s_i}$ and $b = \prod_{i=1}^{t} p_i^{r_i}$, then $ab^2 = \prod_{i=1}^{t} p_i^{s_i+2r_i} = n$. Clearly a is square-free.*

(2) There is a bijection between the set of positive divisors d of n and the set

$$U = \{(b_1, \ldots, b_t) \in \mathbb{N}^t \mid 0 \leq b_i \leq a_i, \text{ for } 1 \leq i \leq t\}.$$

Any $(b_1, \ldots, b_t) \in U$ corresponds to a positive divisor $d = \prod_{i=1}^{t} p_i^{b_i}$ of n and any positive divisor of n can be uniquely written in that way by the fact that $\mathbb{Z}$ is a unique factorization domain. Hence $\eta(n) = |U| = \prod_{i=1}^{t} (a_i + 1)$.

(3) We calculate:

$$\begin{aligned}
\sigma(n) &= \sum_{d|n} d \\
&= \sum_{(b_1,\ldots,b_t)\in U} p_1^{b_1} \cdots p_t^{b_t} \\
&= \sum_{b_1=0}^{a_1} p^{b_1} \sum_{(b_1,\ldots,b_t)\in U} p_2^{b_2} \cdots p_t^{b_t} \\
&= \cdots \\
&= \left(\sum_{b_1=0}^{a_1} p^{b_1}\right) \cdots \left(\sum_{b_t=0}^{a_t} p^{b_t}\right) = \left(\frac{p_1^{a_1+1}-1}{p_1-1}\right) \cdots \left(\frac{p_t^{a_t+1}-1}{p_t-1}\right)
\end{aligned}$$

where we usethe property of telescopic sums $(p-1)\left(\sum_{i=0}^{a} p^i\right) = p^{a+1} - 1$. □

Example:

$$\begin{aligned}
\eta(3) &= \eta(3^1) = 1 + 1 = 2 \\
\eta(6) &= \eta(2^1 3^1) = (1 + 1)(1 + 1) = 4 \\
\eta(12) &= \eta(2^2 3^1) = (2 + 1)(1 + 1) = 6 \\
\eta(28) &= \eta(2^2 7^1) = (2 + 1)(1 + 1) = 6
\end{aligned}$$

$$\begin{aligned}
\sigma(3) &= \sigma(3^1) = \frac{3^{1+1}-1}{3-1} = \frac{8}{2} = 4 \\
\sigma(6) &= \sigma(2^1 3^1) = \left(\frac{2^{1+1}-1}{2-1}\right)\cdot\left(\frac{3^{1+1}-1}{3-1}\right) = \frac{3}{1}\cdot\frac{8}{2} = 12 \\
\sigma(12) &= \sigma(2^2 3^1) = \left(\frac{2^{2+1}-1}{2-1}\right)\cdot\left(\frac{3^{1+1}-1}{3-1}\right) = \frac{7}{1}\cdot\frac{8}{2} = 28 \\
\sigma(28) &= \sigma(2^2 7^1) = \left(\frac{2^{2+1}-1}{2-1}\right)\cdot\left(\frac{7^{1+1}-1}{7-1}\right) = \frac{7}{1}\cdot\frac{48}{6} = 56 = 2\cdot 28.
\end{aligned}$$

A positive number n is called perfect if $\sigma(n) = 2n$. There are two unsolved question, namely: Are there odd perfect numbers? Are there infinitely many perfect numbers?

A result by Pascal Ochem and Michael Rao from 2012 says that if there exists an odd perfect number n, then $n > 10^{1500}$.

Euler proved that any even number is perfect number if and only if it is of the form $2^m(2^{m+1}-1)$ with $2^{m+1}-1$ being a prime number. To see this, let n be an even perfect number. Then $n = 2^m x$ for $m = \nu_2(n) \geq 1$ and x odd. Since n is perfect, σ is multiplicative and 2^m and x are relatively prime, we obtain

$$2^{m+1}x = 2n = \sigma(n) = \sigma(2^m x) = \sigma(2^m)\sigma(x) = \left(2^{m+1}-1\right)\sigma(x).$$

Since $2^{m+1}-1$ is an odd integer greater than 1 dividing $2^{m+1}x$ we must have $(2^{m+1}-1) \mid x$. Hence there exists $y \in \mathbb{Z}$ such that $x = (2^{m+1}-1)y$. Canceling $(2^{m+1}-1)$ from the equation above leads to

$$2^{m+1}y = \sigma(x)$$

Let $u = \sum_{d\mid x, d\neq x, d\neq y} d$. Then

$$\sigma(x) = x + y + u = \left(2^{m+1}-1\right)y + y + u = 2^{m+1}y + u.$$

Comparing the last two equations, leads to $u = 0$. Hence y and x are the only two divisors of x (note that $x \neq 1$ since 2^m is not perfect). Thus $y = 1$ and $x = 2^{m+1}-1$ is an odd prime number. An odd prime number p is called a Mersenne prime if it is of the form $p = 2^{m+1}-1$, for some $m \geq 1$.

The converse of the argument above also holds: Let $p = 2^{m+1}-1$ be a Mersenne prime. Then $n = 2^m \cdot p = 2^m \cdot \left(2^{m+1}-1\right)$ is a perfect number, because

$$\sigma(n) = \left(\frac{2^{m+1}-1}{2-1}\right)\cdot\left(\frac{p^2-1}{p-1}\right) = (2^{m+1}-1)\frac{(2^{m+1}-1)^2-1}{2^{m+1}-2} = (2^{m+1}-1)2^{m+1} = 2n$$

Thus even perfect numbers and Mersenne prime numbers are in correspondence. The so-called Great Internet Mersenne Prime Search, see `https://www.mersenne.org`, gathers information about Mersenne primes. As of 2022, only 51 Mersenne prime numbers have been identified and the largest one is

$$2^{82589933}-1$$

which is a number with 24862048 digits.

Arithmetic function about the prime number distribution

At the end of this chapter we will have a short look at the prime number theorem, that says that the number of primes in the real interval $[0, x]$ is assymptotic to $x/\ln(x)$. For any $x \in \mathbb{R}_{\geq 1}$ we denote the set of positive prime numbers less than x by

$$\mathbb{P}_x := \mathbb{P} \cap [2, x] = \{p \in \mathbb{P} : 2 \leq p \leq x\}.$$

Associated to this set we have

$$\pi : \mathbb{R}_{\geq 1} \to \mathbb{N}, \qquad x \longmapsto \pi(x) := |\mathbb{P}_x|, \qquad \forall x \in \mathbb{R}^+$$

$$\theta : \mathbb{R}_{\geq 1} \to \mathbb{N}, \qquad x \longmapsto \theta(x) := \sum_{p \in \mathbb{P}_x} \ln(p) \qquad \text{and } \theta(1) = 0.$$

If $\lfloor x \rfloor$ denotes the largest number less or equal to x, then we clearly have $\mathbb{P}_x = \mathbb{P}_{\lfloor x \rfloor}$ and $\pi(x) = \pi(\lfloor x \rfloor)$. Hence we can assume that x is a natural number.

Moreover, since the sum in $\theta(x)$ has $\pi(x)$ summands and $p < x$ implies $\ln(p) < \ln(x)$ we have

$$\theta(x) < \pi(x) \ln(x), \qquad \forall x \geq 2.$$

Proposition 2.9 *$\theta(x) < 4 \ln(2) x$, For any $x \in \mathbb{R}^+$*

P*roof: Let $n \geq 2$ and $p \in \mathbb{P}_{2n} \setminus \mathbb{P}_n$ be a prime number between n and $2n$, i.e. $n < p < 2n$. Then any p divides $\binom{2n}{n}$, because $p \nmid k$ for any $k \leq n$. Hence*

$$p \mid \binom{2n}{n} = \frac{(2n)!}{(n!)^2} = \frac{(2n) \cdot (2n-1) \cdots (n+2) \cdot (n+1)}{n \cdot (n-1) \cdots 2 \cdot 1}$$

ans therefore $p < \binom{2n}{n}$ and since $2^{2n} = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{2n}{n}$ we have also:

$$\prod_{p \in \mathbb{P}_{2n} \setminus \mathbb{P}_n} p < \binom{2n}{n} < 2^{2n}.$$

Taking the logarithm, we conclude:

$$\theta(2n) - \theta(n) = \sum_{p \in \mathbb{P}_{2n} \setminus \mathbb{P}_n} \ln(p) \ < 2n \ln(2).$$

The property of the telescopic sum yields:

$$\theta(2^m) = \sum_{k=1}^{m-1} \Big(\theta(2^{k+1}) - \theta(2^k) \Big) < \sum_{k=1}^{m-1} 2 \cdot (2^k) \cdot \ln(2) = 2 \ln(2) \, (2^m - 1) < 2^{m+1} \ln(2)$$

Thus, if m is the least positive number such that $2^{m-1} < x \leq 2^m$, then

$$\theta(x) \leq \theta(2^m) < 2^{m+1} \ln(2) = 4 \ln(2) 2^{m-1} < 4 \ln(2) x.$$

□

Corollary 2.10 *There exist constants $C_1, C_2 > 0$ such that*

$$C_1 \ln(x) \leq \pi(x) < C_2 \frac{x}{\ln(x)},$$

for all $x \geq 2$, where $C_1 = (2\ln(2))^{-1}$ and $C_2 = 8\ln(2) + 1$.

P*roof: We shall first determine the upper bound. Clearly, $\theta(x) > \sum_{p \in \mathbb{P}_x \setminus \mathbb{P}_{\sqrt{x}}} \ln(p)$ and as this sum has $\pi(x) - \pi(\sqrt{x})$ summands and $ln(p) > \ln(\sqrt{x})$, for all $p \notin \mathbb{P}_{\sqrt{x}}$, we conclude:*

$$\theta(x)) \geq \sum_{p \in \mathbb{P}_x \setminus \mathbb{P}_{\sqrt{x}}} \ln(p) \geq \ln(\sqrt{x}) \left(\pi(x) - \pi(\sqrt{x})\right) \geq \ln(\sqrt{x})\pi(x) - \sqrt{x}\ln(\sqrt{x}).$$

Using Proposition 2.9 and the fact that $\sqrt{x} < \frac{2x}{\ln(x)}$,[1] *for $x \geq 2$, we have therefore*

$$\pi(x) \leq \frac{\theta(x)}{\ln(\sqrt{x})} + \sqrt{x} \leq \frac{8\ln(2)x}{\ln(x)} + \frac{2x}{\ln(x)} = C_2 \frac{x}{\ln(x)},$$

for $C_2 = 8\ln(2) + 1$.

Let us determine the lower bound. Let $p(n) \subseteq \mathbb{P}$ be the set of prime divisors of a number n. For any finite set of positive primes $S \subseteq \mathbb{P}$ define

$$f_S : \mathbb{R}^+ \to \mathbb{N}, \qquad f_S(x) = |\{n \in \mathbb{N}^+ : n \leq x \ \text{ and } \ p(n) \subseteq S\}|.$$

That means, $f_S(x)$ counts the numbers less than x whose prime divisors belong to S. Let $n \leq x$ be a number whose prime divisors belong to S and write $n = ab^2$ with a square-free. Then we have $b \leq \sqrt{(}x)$ and for a we have at most $|S|$ choices. Hence there are at most $|S|\sqrt{x}$ possible numbers $n \leq x$ such that all prime divisors of n belong to S, i.e.

$$f_S(x) \leq |S|\sqrt{x}.$$

Now let $S = \{p_1, \ldots, p_{\pi x}\}$ be all prime numbers less or equal to x, then $f_S(x) = x$, because the prime divisors of any number $n \leq x$ belong to S. Thus

$$x = f_S(x) \leq 2^{\pi(x)}\sqrt{x} \Rightarrow \sqrt{x} \leq 2^{\pi(x)} \Rightarrow \pi(x) \geq \frac{\ln(x)}{2\ln(2)}.$$

□

The last Corollary implies that $\frac{\pi(x)}{x} \to 0$ when $x \to +\infty$.

Proposition 2.11 $\lim_{x \to +\infty} \frac{\pi(x)}{\frac{x}{\ln(x)}} > \frac{ln(2)}{2}$.

P*roof: Let $n \geq 2$. Then by Lemma 2.7, for any $p \in \mathbb{P}$:*

$$\nu_p(\binom{2n}{n}) = \nu_p((2n)!) - \nu_p((n!)^2) = \sum_{k \geq 1}^{\nu_p(2n)} \left\lfloor \frac{2n}{p^k} \right\rfloor - 2 \sum_{k \geq 1}^{\nu_p(n)} \left\lfloor \frac{n}{p^k} \right\rfloor = \sum_{k \geq 1}^{\nu_p(2n)} \left(\left\lfloor \frac{2n}{p^k} \right\rfloor - 2 \left\lfloor \frac{n}{pk} \right\rfloor \right).$$

[1] since $\frac{\ln(x)}{2} = \ln(\sqrt{x}) < \sqrt{x}$ and hence $\sqrt{x} < \frac{2x}{\ln(x)}$

For any $x \in \mathbb{R}_{>1}$, we have $\lfloor 2x \rfloor - 2\lfloor x \rfloor \in \{0,1\}$, because if $x = m + \epsilon$ with $m \in \mathbb{N}$ and $\epsilon \in [0, 1/2[$, then $2x = 2m + 2\epsilon$, with $0 < 2\epsilon < 1$. Hence $\lfloor 2x \rfloor = 2m = 2\lfloor x \rfloor$. Otherwise, if $x = m + \epsilon$ and $\epsilon \in [1/2, 1[$, then $2x = 2m + 1 + (2\epsilon - 1)$ with $0 < 2\epsilon - 1 < 1$. Hence $\lfloor 2x \rfloor - 2\lfloor x \rfloor = 1$.

Hence

$$\nu_p\left(\binom{2n}{n}\right) = \sum_{k\geq 1}^{\nu_p(2n)} \left(\left\lfloor \frac{2n}{p^k} \right\rfloor - 2 \left\lfloor \frac{n}{pk} \right\rfloor \right) < \nu_p(2n).$$

Then

$$2^n < \left(\frac{n+1}{1}\right) \cdot \left(\frac{n+2}{2}\right) \cdots \left(\frac{n+n}{n}\right) = \frac{(2n)!}{(n!)^2} = \binom{2n}{n} = \prod_{p \in \mathbb{P}_{2n}} p^{\nu_p(\binom{2n}{n})} \leq \prod_{p \in \mathbb{P}_{2n}} p^{\nu_p(2n)}.$$

As $p^{\nu_p(2n)} \leq 2n$, we have $\nu_p(2n) \leq \lfloor \frac{\ln(2n)}{\ln(p)} \rfloor$ and

$$n \ln(2) \leq \sum_{p \in \mathbb{P}_{2n}} \nu_p(2n) \ln(p) \leq \sum_{p \in \mathbb{P}_{2n}} \left\lfloor \frac{\ln(2n)}{\ln(p)} \right\rfloor \ln(p)$$

If $2n \geq p > \sqrt{2n}$, then $1 \leq \frac{\ln(2n)}{\ln(p)} < 2$. Hence $\left\lfloor \frac{\ln(2n)}{\ln(p)} \right\rfloor = 1$ and

$$n \ln(2) \leq \left(\sum_{p \in \mathbb{P}_{\sqrt{2n}}} \left\lfloor \frac{\ln(2n)}{\ln(p)} \right\rfloor \ln(p) \right) + \left(\sum_{p \in \mathbb{P}_{2n} \setminus \mathbb{P}_{\sqrt{2n}}} \ln(p) \right) \leq \sqrt{2n} \ln(2n) + \theta(2n).$$

We conclude

$$\pi(2n) \ln(2n) > \theta(2n) \geq n \ln(2) - \sqrt{2n} \ln(2n)$$

Let n be such that $2n \leq x < 2n + 1$. Then

$$\pi(x) \ln(x) > n \ln(2) - \sqrt{2n} \ln(2n) > \frac{(x-1)\ln(2)}{2} - \sqrt{x} \ln(x)$$

$$\frac{\pi(x)}{x/\ln(x)} > \frac{ln(2)}{2} - \frac{\ln(2)}{2x} - \frac{\ln(x)}{\sqrt{x}} \qquad \overset{x\to\infty}{\longrightarrow} \qquad \frac{ln(2)}{2}$$

□

We have seen that there exist $C, C' > 0$ such that

$$C > \frac{\pi(x)}{x/\ln(x)} > C'$$

for large enough x. The prime number theorem says that the limit of $\frac{\pi(x)}{x/\ln(x)}$ is actually 1, this means that asymptotically $\pi(x)$ is as big as $\frac{x}{\ln(x)}$.

Theorem 2.12 (Prime Number Theorem) *$\pi(x) \sim \frac{x}{\ln(x)}$.*

x	$\lfloor x/\ln(x) \rfloor$	$\pi(x)$	$\pi(x)/(x/\ln(x))$
10^1	*4*	*4*	*0.9210340371976184*
10^2	*21*	*25*	*1.151292546497023*
10^3	*144*	*168*	*1.1605028868689988*
10^4	*1085*	*1229*	*1.131950831715873*
10^5	*8685*	*9592*	*1.104319810599443*
10^6	*72382*	*78498*	*1.0844899477790795*
10^7	*620420*	*664579*	*1.0711747889618228*
10^8	*5428681*	*5761455*	*1.0612992317564809*
10^9	*48254942*	*50847534*	*1.0537269642351712*
10^{10}	*434294481*	*455052511*	*1.047797128358109*

3

Primitive roots of unity modulo n

We have already seen that division with rest is possible in $\mathbb{Z}$ *(and in* $\mathbb{N}$*). For all* $n \leq 1$ *the relation* $\sim_n$ *on* $\mathbb{Z}$ *defined as*

$$a \sim_n b \qquad \Leftrightarrow \qquad n \mid a - b \tag{3.1}$$

for all $a, b \in \mathbb{Z}$ *is an equivalence relation. The set of equivalence classes* $\mathbb{Z}/\sim_n$ *is denoted by* $\mathbb{Z}_n$*. If* $a \sim_n b$ *then one usually writes* $a \equiv b \,(\bmod n)$*. The case* $n = 1$ *is the trivial case in which every number is related via* $\sim_1$*. In this case* $\mathbb{Z}_1 = \{[0]_\sim\}$*.*

For $n > 1$ *there are exactly* n *different equivalence classes in* $\mathbb{Z}_n$*, because of the division algorithm: for all* $a \in \mathbb{Z}$ *there exist* $q \in \mathbb{Z}$ *and* $0 \leq r < n$ *such that* $a = q \cdot n + r$*, i.e.* $n \mid a - r$ *and therefore* $a \sim_n r$*. Moreover if* $0 \leq r, s < n$ *are in the same equivalence class of* $\sim_n$*, then* $r = s$*, as* $-n < r - s < n$*. Thus* $\{0, 1, \ldots, n-1\}$ *is a set of representatives for* $\mathbb{Z}_n$ *and one can identify* $\mathbb{Z}_n$ *with this set.*

$\mathbb{Z}_n$ *has addition and multiplication as follows: for all* $a, b \in \{0, 1, \ldots, n-1\}$ *one sets*

$$a + b = (a + b)(\bmod n) \qquad a \cdot b = (a \cdot b)(\bmod n) \tag{3.2}$$

where $x(\bmod n)$ *denotes the rest of the division of* x *by* n*.*

The operations $+$ *and* $\cdot$ *turn* $\mathbb{Z}_n$ *into a commutative unital ring, which means that the addition and multiplication is associative and commutative and* $\cdot$ *distributes over* $+$*. Calculating* modulo n *is sometimes referred to as modular arithmetic.*

Calculating modulo a number simplifies sometimes certain arguments. We can for example easily prove the following

Lemma 3.1 *There are infinitely many prime numbers that are congruent with* 3 *modulo* 4

Proof: Suppose to the contrary that the set of prime numbers that are congruent with 3 *modulo* 4 *is finite, say* $\{p_1, \ldots, p_s\}$ *is this set with* $p_i \neq p_j$ *if* $i \neq j$. *We assume that* $p_1 = 3$. *Then form the number*

$$n = 4p_2 \cdots p_s + 3$$

Since $\mathbb{Z}$ *is a unique factorization domain,* $n = q_1 \cdots q_t$ *for some prime numbers* q_i. *Since* n *is odd, none of the primes* q_i *is equal to* 2.

If $q_i \equiv 1(\bmod 4)$ *would hold for all* i, *i.e.* $q_i = 1$ *in* $\mathbb{Z}_4$, *then also*

$$n = q_1 \cdots q_t \equiv 1(mod 4),$$

but $n \equiv 3(\bmod 4)$. *Hence there exists a prime* q_i *such that* $q_i \equiv 3(\bmod 4)$, *i.e.* $q_i = p_j$ *for some* $1 \leq j \leq s$. *However, then* $q_i \mid 4p_2 \cdots p_s$ *and as* $q_i \mid n$, *we must have* $q_i \mid 3$, *i.e.* $q_i = 3$. *But then* $3 \mid 4p_2 \cdots p_s$ *leads to a contradiction, since* $3 \nmid 2$ *and* $3 \nmid p_l$ *for* $l \neq 1$. *Therefore our initial assumption, that there are only finitely many of such prime numbers is wrong.* □

An element $a \in \mathbb{Z}_n$ *is said to have a multiplicative inverse if there exists* $b \in \mathbb{Z}_n$ *such that* $ab = 1$ *in* $\mathbb{Z}_n$. *This means, recalling that the elements of* $\mathbb{Z}_n$ *are equivalence classes, that the class of* $ab \mod n$ *is the same as the class of* $1 \mod n$, *i.e.* $n \mid ab - 1$.

Proposition 3.2 *Let* $n > 1$ *and* a *a non-zero element of* $\mathbb{Z}_n$. *Then* a *has a multiplicative inverse in* $\mathbb{Z}_n$ *if and only if* $\gcd(a, n) = 1$. *In this case the extended Euclidean algorithm can be used to calculate its inverse.*

Proof: By the Euclidean Algorithm there exist $r, s \in \mathbb{Z}$ *such that* $\gcd(a, n) = ra + sn$. *In other words* $\gcd(a, n) \equiv ra \bmod n$. *If* $\gcd(a, n) = 1$, *then* r *is the inverse of* a *in* $\mathbb{Z}_n$. *If* $\gcd(a, n) \neq 1$, *then* $b = \frac{n}{\gcd(a,n)}$ *is a positive number, satisfying*

$$ba = \frac{n}{\gcd(a, n)}a = n\frac{a}{\gcd(a, n)} \equiv 0 \mod n. \tag{3.3}$$

Thus, if a *had a multiplicative inverse, say* $c \in \mathbb{Z}_n$, *then* $b = bac = 0$, *contradicts* $b > 0$. □

We denote the set of (multiplicative) invertible elements of a ring R *by* $U(R)$, *which is always a group, since if* a *and* b *are invertible, then* ab *is invertible with inverse* $(ab)^{-1} = b^{-1}a^{-1}$. *For* $R = \mathbb{Z}$ *we saw that* $U(\mathbb{Z}) = \{1, -1\}$. *Hence the map* $f : U(\mathbb{Z}) \to \mathbb{Z}_2 = \{0, 1\}$ *with* $1 \mapsto 0$ *and* $-1 \mapsto 1$ *is an isomorphism of groups. Proposition 3.2 shows that the unit group of* $\mathbb{Z}_n$ *is given by the elements*

$$U(\mathbb{Z}_n) = \{a \in \mathbb{Z}_n \mid \gcd(a, n) = 1\}, \qquad \text{and} \qquad |U(\mathbb{Z}_n)| = \varphi(n).$$

Corollary 3.3 $\mathbb{Z}_n$ *is a field if and only if* n *is a prime number.*

Proof: Clearly if n *is prime, then* $|U(\mathbb{Z}_n)| = \varphi(n) = n - 1 = |\mathbb{Z}_n \setminus \{0\}|$. *Thus every non-zero element of* $\mathbb{Z}_n$ *is invertible. Conversely, if* n *is not prime, then there exist* $a, b \in \mathbb{Z}$ *such that* $ab = n$ *and* $1 < a, b < n$. *Hence in* $\mathbb{Z}_n$ *we have* $ab = 0(\bmod n)$. *Thus neither* a *nor* b *can be invertible. Therefore,* $\mathbb{Z}_n$ *is not a field.* □

What is the structure of $U(\mathbb{Z}_n)$? *When is it cyclic? A first step to describe* $U(\mathbb{Z}_n)$ *is the following Theorem of Euler, which is basically an application of Lagrange's Theorem.*

Theorem 3.4 (Euler) *Let $n \in \mathbb{N}_{>1}$ and $a \in \mathbb{Z}$. If $\gcd(a,n) = 1$, then $a^{\varphi(n)} = 1$ in $\mathbb{Z}_n$.*

P*roof: By hypothesis $a \in U(\mathbb{Z}_n)$. Hence Lagrange Theorem says that the order of a divides the order of $U(\mathbb{Z}_n)$ and therfore $a^{\varphi(n)} = 1$ in $\mathbb{Z}_n$.* $\square$

As a consequence we obtain

Theorem 3.5 (Fermat) *If p is a prime number and $a \in \mathbb{Z}$, then $a^p = a(\bmod p)$.*

P*roof: Since $\varphi(p) = p-1$ and $a(\bmod p) \in U(\mathbb{Z}_p)$, for all $a \in \mathbb{Z}$ with $\gcd(a,p) = 1$, we obtain by Euler's $a^{p-1} = 1(\bmod p)$ and therefore, multiplying by a, also $a^p = a(\bmod p)$. In case $\gcd(a,p) \neq 1$, then $p \mid a$ and $a = 0(\bmod p)$ and therefore also $a^p = 0 = a(\bmod p)$.* $\square$

When is $U(\mathbb{Z}_n)$ a cyclic group ? Using Lemma 2.5 we can also give a nice criteria to show that a finite group G is cyclic, namely precisely if for each divisor d of $|G$ ther exists at most one subgroup of order d. This will be a key step to prove that the multiplicative group of a finite field is cyclic.

Theorem 3.6 *The following statements are equivalent for a finite group G of order n.*

(a) G is cyclic;

(b) for every divisor d of n, there exists exactly one subgroup H of G of order d;

(c) for every divisor d of n, there exists at most one subgroup H of G of order d.

P*roof: Let G be a group of order n and denote by $\sim$ the equivalence relation $a \sim b$ if and only if $\langle a \rangle = \langle b \rangle$, for all $a, b \in G$. Then $\sim$ yields a partition of G into distinct equivalent classes, i.e.*

$$G = \dot{\bigcup}_{a \in \Lambda} [a]_\sim \tag{3.4}$$

for some set of representatives $\Lambda \subset G$. Note that if $a \in \Lambda$, then $[a]_\sim$ is precisely the set of generators of the cyclic subgroup $C = \langle a \rangle$. Moreover, $|C| = d \mid n$ and $|[a]_\sim| = \varphi(d)$. For any divisor $d \mid n$, let c_d be the number of different cyclic subgroups of order d of G. Then the partition 3.4 yields:

$$n = |G| = \sum_{a \in \Lambda} |[a]_\sim| = \sum_{d \mid n} c_d \varphi(d). \tag{3.5}$$

$(a) \Rightarrow (b)$: Suppose G is cyclic, then we can assume $G = \langle x \rangle$. Any subgroup of H is generated by a power of x. To see this, let H be a subgroup of G. If $H = \{1\}$, then $H = \langle x^0 \rangle$. If $H \neq \{1\}$, then choose $k > 0$ minimal such that $x^k \in H$. For any element $h \in H$, there exists $m \geq 1$ such that $h = x^m$. By the minimality of k, $k \leq m$ and by the division algorithm, there exist $q, r \in \mathbb{Z}$ such that $m = qk + r$ and $0 \leq r < k$. Hence $x^r = x^{m-qk} = h\left(x^k\right)^{-q} \in H$. However since $k > 0$ was the least positive integer exponent of x such that $x^k \in H$ and $x^r \in H$ with $0 \leq r < k$, we must have $r = 0$, i.e. $h = x^{kq}$, which shows $H = \langle x^k \rangle$. Therefore, any subgroup of G is cyclic.

Let $d \mid n$ be a divisor of n. Then there exist at least one subgroup H of order d, which is

$$H = \langle x^{n/d} \rangle = \{1, x^{n/d}, x^{2n/d}, \ldots, x^{(d-1)n/d}\}. \tag{3.6}$$

Hence $c_d \geq 1$, for all $d \mid n$. Supose there exists another subgroup of order d, say $H' = \langle x^k \rangle$, which is of course also cyclic, then $x^{kd} = 1 = x^n$. Hence $n \mid dk$ and $(n/d) \mid k$, i.e. $H' \subseteq H$. But since $|H'| = d = |H|$, both subgroups are equal. Hence $c_d = 1$, for all $d \mid n$ and (b) hold.

$(b) \Rightarrow (c)$ holds trivially.

$(c) \Rightarrow (a)$: Let G be a group of order n, then by (c), $c_d \leq 1$, for all $d \mid n$. Hence by (3.5)

$$n = |G| = \sum_{d|n} c_d \varphi(d) \leq \sum_{d|n} \varphi(d) = n. \tag{3.7}$$

This shows $c_d = 1$ for all $d \mid n$ and in particular also $c_n = 1$, i.e. G is cyclic. □

Let F be a field. Then we know that any polynomial $f \in F[x]$ of degree $n \geq 1$ has no more than n roots. Using this simple fact and Theorem 3.6 we obtain that

Corollary 3.7 *The multiplicative group $U(F) = F \setminus \{0\}$ of a finite field is a cyclic group.*

P*roof: Let $G = F \setminus \{0\}$ be the multiplicative group of a finite field F. For any divisor $d \mid |G|$, consider two subgroups H, K of G of order d, i.e. $|H| = d = |K|$. Then for any element $a \in H \cup K$, we obtain that the order of a in the H (or K) divides d, i.e. $a^d = 1$. Hence any $a \in H \cup K$ is a root of the polynomial $x^d - 1 \in F[x]$. Since there are at most d roots, we have $d = |H| = |K| \leq |H \cup K| \leq d$ and therefore $H = H \cup K = K$, i.e. $H = K$. By Theorem 3.6, G is a cyclic group.* □

Since for any prime number p, $\mathbb{Z}_p$ is a finite field we conclude:

Corollary 3.8 *For any prime number p, $U(\mathbb{Z}_p)$ is a cyclic group.*

We should note, that if $U(\mathbb{Z}_n)$ is a cyclic group, then n need not be a prime number. The smallest example is $U(\mathbb{Z}_4) = \{1, 3\} \simeq \mathbb{Z}_2$. We will show now that $U(\mathbb{Z}_{p^n})$ is always cyclic if $p \neq 2$ and $n \geq 1$. The case $n = 1$ has been just proven. Note that if p is a prime number, then

$$p \mid \binom{p}{k}, \qquad \text{for all } 1 < k < p, \tag{3.8}$$

because in this case $p \nmid k!$ and $p \nmid (p-k)!$ since all factors of $k!$ and $(p-k)!$ are less than p. Therefore $p \mid p\frac{(p-1)!}{k!(p-k)!} = \binom{p}{k}$.

Lemma 3.9 *Let p be a prime number, $n \geq 1$ and $a, b \in \mathbb{Z}$ then*

1. *if $a = b (\bmod p^n)$ then $a^p = b^p (\bmod p^{n+1})$.*
2. *$a = 1 (\bmod p^n)$ if and only if $a^p = 1 (\bmod p^{n+1})$.*

P*roof: (1) If $a = b (\bmod p^n)$, then $a = b + cp^n$, for some $c \in \mathbb{Z}$. Hence*

$$a^p = (b + cp^n)^p = \sum_{i=0}^{p} \binom{p}{i} b^i (cp^n)^{p-i} = b^p + (cp^n)^p = b^p (\bmod p^{n+1}),$$

since $p \mid \binom{p}{i}$ for $1 < i < p$, i.e. $p^{n+1} \mid \binom{p}{i} p^{n(p-i)}$.

(2) By (1) we have proven already the only if part. We will use induction on n. Let $n = 1$ and $a^p = 1 (\bmod p^2)$, then by Fermat's Theorem $a = a^p = 1 (\bmod p)$. Suppose $n \geq 1$ and

suppose that we have already shown that if $a^p = 1(\bmod p^{n+1})$, *then* $a = 1(\bmod p^n)$. *Suppose* $a^p = 1(\bmod p^{n+2})$, *then we also have* $a^p = 1(\bmod p^{n+1})$. *By assumption,* $a = 1(\bmod p^n)$ *and* $a = 1 + cp^n$ *for some* $c \in \mathbb{Z}$. *Thus*

$$a^p = (1+cp^n)^p = \sum_{i=0}^{p} \binom{p}{i} (cp^n)^{p-i} = 1 + p(cp^n)(\bmod p^{n+2})$$

since $\binom{p}{i}(cp^n)^i$ *will be divisible by* p^{n+2} *for* $i \geq 2$. *However, by assumption* $a^p = 1(\bmod p^{n+2})$, *hence* $1 = a^p = 1 + cp^{n+1}(\bmod p^{n+2})$ *shows that* $p^{n+2} \mid cp^{n+1}$ *and therefore* $p \mid c$. *This implies that* $a - 1 = cp^n$ *is divisible by* p^{n+1}, *i.e.* $a^p = 1(\bmod p^{n+1})$. *By induction the result follows.* □

Theorem 3.10 *Let* p *be an odd prime. Then* $U(\mathbb{Z}_{p^n})$ *is a cyclic group for any* $n \geq 1$.

P*roof: We know from Corollary 3.8 that* $U(\mathbb{Z}_p)$ *is a cyclic group, which shows that the Theorem holds for* $n = 1$.

Let $n = 2$ *and choose any generator* $g \in U(\mathbb{Z}_p)$. *Let* m *be the order of* g *in* $U(\mathbb{Z}_{p^2})$. *Then* $m \mid \varphi(p^2) = p(p-1)$. *On the other hand* $g^m = 1(\bmod p^2)$ *implies* $g^m = 1(\bmod p)$ *and therefore* $\varphi(p) = p-1 \mid m$. *Hence* $m \in \{p-1, p(p-1)\}$. *If* $m = p(p-1)$, *then* g *is a generator in* $U(\mathbb{Z}_{p^2})$. *Suppose* $m = p-1$, *i.e.* $g^{p-1} = 1(\bmod p^2)$. *Since* $p+g$ *is also a generator in* $U(\mathbb{Z}_p)$, *the same reasoning as above shows that the order of* $p+g$ *is either* $p-1$ *or* $p(p-1)$. *However,*

$$(p+g)^{p-1} = \sum_{i=0}^{p-1} \binom{p-1}{i} p^i g^{p-1-i} = g^{p-1} + (p-1)pg^{p-2} = 1 + (p-1)pg^{p-2} (\bmod p^2) \neq 1 (\bmod p^2),$$

because $p \nmid (p-1)g^{p-2}$ *as* g *has order* $p-1$ *and* $p \nmid p-1$. *This shows that the order of* $p+g$ *in* $U(\mathbb{Z}_{p^2})$ *cannot be* $p-1$ *and therefore* $p+g$ *must be a generator in* $U(\mathbb{Z}_p^2)$.

Let $n \geq 2$ *and suppose* g *is a generator of* $U(\mathbb{Z}_{p^n})$. *Then the order of* g *in* $U(\mathbb{Z}_{p^n})$ *is* $\varphi(p^n) = p^{n-1}(p-1)$. *Let* m *be the order of* g *in* $U(\mathbb{Z}_{p^{n+1}})$. *Then* $g^m = 1(\bmod p^{n+1})$ *implies* $m \mid \varphi(p^{n+1}) = p^n(p-1)$ *and also* $g^m = 1(\bmod p^n)$. *Since the order of* g *in* $U(\mathbb{Z}_{p^n})$ *is* $p^{n-1}(p-1)$, *we obtain* $p^{n-1}(p-1) \mid m$. *If* $m = p^{n-1}(p-1)$, *then* $\left(g^{p^{n-2}(p-1)}\right)^p = g^m = 1(\bmod p^{n+1})$ *would imply by Lemma 3.9,* $g^{p^{n-2}(p-1)} = 1(\bmod p^n)$, *contradicting that the order of* g *in* $U(\mathbb{Z}_{p^n})$ *is* $p^{n-1}(p-1)$. *Hence* $m \neq p^{n-1}(p-1)$ *and as* $p^{n-1}(p-1) \mid m \mid p^n(p-1)$, *we obtain* $m = p^n(p-1)$, *i.e.* g *is a generator of* $U(\mathbb{Z}_{p^{n+1}})$.

□

Generators of $U(\mathbb{Z}_n)$ *are called primitive roots of unity modulo* n. *The above Theorem says that primitive roots of unity modulo* p^n *always exist for* p *an odd prime and moreover, if* g *is a primitive root of unity modulo* p, *then either* g *or* $p+g$ *is a primitive root of unity modulo* p^n *for any* $n \geq 1$.

For example, if $p = 5$, *then* 2 *is a generator of* $U(\mathbb{Z}_5)$, *since* $2^2 = 4 \neq 1(\bmod 5)$ *and* $2^4 = 1(\bmod 5)$. *Note that since* $|U(\mathbb{Z}_{25})| = 5(5-1) = 20$, *the only orders of elements of* $U(\mathbb{Z}_{25})$ *are* $1, 2, 4, 10, 20$ *since* $2^2 = 4 \neq 0(\bmod 25)$, $2^4 = 16 \neq 0(\bmod 25)$ *and* $2^{10} = 1024 = 24(\bmod 25)$, 2 *must have order* 20 *and is a generator in* $U(\mathbb{Z}_{25})$, *hence a primitive root of unity modulo* 5^n *for any* $n \geq 1$.

If $p = 29$ then 14 is a primitive root of unity modulo 29, because $p - 1 = 28 = 2^2 \times 7$ and the only possible orders are $1, 2, 4, 7, 14, 28$. One checks that $14^k \neq 1 (\bmod 29)$ for $k \in \{1, 2, 4, 7, 14\}$. Hence 14 has order 28 in $U(\mathbb{Z}_{29})$. On the other hand 14 is not a primitive root of unity modulo $29^2 = 841$, because $14^{28} = 1 (\bmod 841)$, while $\varphi(841) = 29 \times 28 = 812$. By the proof of the last Theorem, $14 + 29 = 43$ is a primitive root of unity modulo 29^n, for any $n \geq 1$.

The case for $p = 2$ is different, while $U(\mathbb{Z}_2) = \{1\}$ is the trivial group and $U(\mathbb{Z}_4) = \{1, 3\} \simeq \mathbb{Z}_2$ is cyclic, $U(\mathbb{Z}_{2^n})$ is never cyclic for $n \geq 3$.

Theorem 3.11 *Let $n \geq 3$. Then $U(\mathbb{Z}_{2^n}) \simeq \mathbb{Z}_2 \times \mathbb{Z}_{2^{n-2}}$ is never cyclic. More precisely,*

$$U(\mathbb{Z}_{2^n}) = \{(-1)^a 5^b : a \in \{0, 1\}, 0 \leq b \leq 2^{n-2}\}.$$

P*roof: We will first prove that the order of 5 in $\mathbb{Z}_{2^n}$ is 2^{n-2}. To do so we first prove the following equation by induction*

$$5^{2^{n-3}} = 1 + 2^{n-1} (\bmod 2^n) \tag{3.9}$$

For $n = 3$ this is clear, since $5^{2^{3-3}} = 5 = 1 + 4 (\bmod 2^3)$. Suppose equation (3.9) has been proved for some $n \geq 3$. Then $5^{2^{n-3}} = 1 + 2^{n-1} (\bmod 2^n)$ implies

$$(5^{2^{n-3}})^2 = \left(1 + 2^{n-1}\right)^2 (\bmod 2^{n+1}).$$

by Lemma 3.9. Since $n \geq 3$, $2(n-1) \geq n+1$, we have $2^{2(n-1)} = 0 (\bmod 2^{n+1})$ and therefore, $5^{2^{n-2}} = 1 + 2^n (\bmod 2^{n+1})$. Thus by induction, equation (3.9) holds for all $n \geq 3$.

In particular, there exists $c \in \mathbb{Z}$ such that

$$5^{2^{n-2}} = 1 + 2^n + c2^{n+1} = 1 + 2^n(1 + 2c)$$

and therefore $5^{2^{n-2}} = 1 (\bmod 2^n)$, which means that the order of 5 divides 2^{n-2}. On the other hand

$$5^{2^{n-3}} = 1 + 2^{n-1} (\bmod 2^n) \neq 1 (\bmod 2^n)$$

shows that the order is larger than 2^{n-3} and therefore must be 2^{n-2}.

We claim that

$$U(\mathbb{Z}_{2^n}) = \{(-1)^a 5^b : a \in \{0, 1\}, b \in \{0, 1, \ldots, 2^{n-2} - 1\}\}. \tag{3.10}$$

To show this it is enough to verify that the cardinality of the right side is $2^{n-1} = \varphi(2^n)$. Let $a, a' \in \{0, 1\}$ and $b, b' \in \{0, 1, \ldots, 2^{n-2} - 1\}$, such that

$$(-1)^a 5^b = (-1)^{a'} 5^{b'} (\bmod 2^n)$$

Then in particular $(-1)^a 5^b = (-1)^{a'} 5^{b'} (\bmod 4)$ and as $5 = 1 (\bmod 4)$ we get $(-1)^a = (-1)^{a'} (\bmod 4)$ which means $a = a' (\bmod 2)$. Since $a, a' \in \{0, 1\}$ we get $a = a'$. Therefore, $5^b = 5^{b'} (\bmod 2^n)$ and hence $b = b' (\bmod 2^{n-2})$ as we have shown that 2^{n-2} is the order of 5 in $U(\mathbb{Z}_{2^n})$. Since $b, b' < 2^{n-2}$ we obtain equality, i.e. $b = b'$. Therefore the set on the right side of (3.10) has cardinality, $2^{n-1} = \varphi(2^n)$.

Consider the function

$$f : \mathbb{Z}_2 \times \mathbb{Z}_{2^{n-2}} \to U(\mathbb{Z}_{2^n}), \qquad (a, b) \mapsto (-1)^a 5^b$$

then it is easily seen that f is a surjective homomorphism of groups, as

$$f((a,b)+(a',b')) = f(a+a', b+b') = (-1)^{a+a'}5^{b+b'} = (-1)^a 5^b (-1)^{a'} 5^{b'} = f(a,b)f(a',b').$$

As $|\mathbb{Z}_2 \times \mathbb{Z}_{2^{n-2}}| = 2^{n-1} = |U(\mathbb{Z}_{2^n})|$, f is an isomorphism of groups. □

In order to understand the structure of $U(\mathbb{Z}_n)$ for any n we need to decompose this group using the prime decomposition of n. The so-called Chinese Remainder Theorem will be the important tool and we present version of it for arbitrary rings and ideals.

Recall that a (two-sided) ideal of a ring R is an additive subgroup $I \subseteq R$ such that $ax \in I$ and $xa \in I$ for all $x \in I$ and $a \in R$. The ideals of the ring $R = \mathbb{Z}$ are precisely the subgroups $n\mathbb{Z} = \{na : a \in \mathbb{Z}\}$ for any $n \in \mathbb{Z}$.

Given two ideals I and J of a ring R one can form their sum, i.e.

$$I + J = \{a + b : a \in I, b \in J\}$$

which is again an ideal of R. For the case of integers, if $I = n\mathbb{Z}$ and $J = m\mathbb{Z}$, then

$$n\mathbb{Z} + m\mathbb{Z} = \gcd(n,m)\mathbb{Z}, \tag{3.11}$$

because by the Euclidean algorithm, there exist $x, y \in \mathbb{Z}$ such that $\gcd(n,m) = xn + ym$; hence

$$\gcd(n,m)\mathbb{Z} = \{z\gcd(n,m) : z \in \mathbb{Z}\} = \{(zx)n + (zy)m : z \in \mathbb{Z}\} \subseteq n\mathbb{Z} + m\mathbb{Z}.$$

On the other hand, n and m are multiples of $\gcd(n,m)$ and therefore any sum of multiples of n and m is also a multiple of $\gcd(n,m)$, i.e. $n\mathbb{Z} + m\mathbb{Z} \subseteq \gcd(n,m)\mathbb{Z}$. This shows (3.11).

Two proper ideals I and J of a ring R are called comaximal ideals if $I + J = R$. For the integers, this means that $n\mathbb{Z}$ and $m\mathbb{Z}$ are comaximal if and only if $\gcd(n,m) = 1$, i.e. n and m are relatively prime.

The direct product of rings $R_1, \ldots, R_k$ is the cartesian product

$$\prod_{i=1}^{k} R_i := R_1 \times \cdots \times R_k := \{(a_1, \ldots, a_k) : a_i \in R_i\}$$

with operations defined by

$$\begin{aligned}(a_1,\ldots,a_k) + (b_1,\ldots,b_k) &:= (a_1+b_1,\ldots,a_k+b_k)\\ (a_1,\ldots,a_k)\cdot(b_1,\ldots,b_k) &:= (a_1b_1,\ldots,a_kb_k),\end{aligned}$$

for all $(a_1,\ldots,a_k),(b_1,\ldots,b_k) \in \prod_{i=1}^k R_i$. Its zero element is $(0_{R_1},\ldots,0_{R_k})$ and its identity is $(1_{R_1},\ldots,1_{R_k})$, where 0_{R_i} and 1_{R_i} are the zero element and identity of the ring R_i.

Theorem 3.12 (Chinese Remainder Theorem) *Let $\{M_1,\ldots,M_k\}$ be a family of proper ideals of a unital ring R and $I = M_1 \cap \ldots \cap M_k$. Then the canonical ring homomorphism*

$$\Phi : R/I \to R/M_1 \times \cdots \times R/M_k, \qquad \text{with} \qquad \Phi(a+I) := (a+M_1,\ldots,a+M_k) \tag{3.12}$$

is an isomorphism of rings if and only if the ideals $M_1,\ldots,M_k$ are pairwise comaximal.

Proof: Note that Φ is well-defined, because if $x+I = x'+I$, then $x-x' \in I = M_1 \cap \cdots \cap M_k$ and $x + M_i = x' + M_i$ for all $1 \leq i \leq k$. Hence $\Phi(x+I) = \Phi(x'+I)$. Furthermore, it is not difficult to see that Φ is a ring homomorphism, by the way how addition and multiplication is defined on R/I and $R/M_1 \times \cdots \times R/M_k$. Moreover, Φ is always injective, because if $\Phi(x+I) = (x+M_1, \ldots, x+M_k) = (0+M_1, \ldots, 0+M_k)$, then $x \in M_i$ for all $1 \leq i \leq k$ and hence $x \in M_1 \cap \cdots \cap M_k = I$ and $x + I = 0 + I$.

So the question is to prove the surjectivity of Φ. If Φ is surjective, then for any $1 \leq i \leq k$ and $a \in R \setminus M_i$ there exists $a' \in R \setminus I$ such that $\Phi(a'+I) = (0+M_1, \ldots, 0+M_{i-1}, a+M_i, 0+M_{i+1}, \ldots, 0+M_k)$. Hence $a'-a \in M_i$ and $a' \in M_j$ for all $j \neq i$. This shows $a = a' - (a'-a) \in M_j + M_i$ and thus $M_j + M_i = R$ for any $i \neq j$. Hence the ideals $M_1, \ldots, M_k$ are pairwise comaximal.

On the contrary, suppose that the ideals M_i are pairwise comaximal. Then for any $1 < i \leq k$ there exist $a_i \in M_1$ and $b_i \in M_i$ such that $1 = a_i + b_i \in M_1 + M_i$. Hence

$$1 = (a_2+b_2)(a_3+b_3)\cdots(a_k+b_k) = a + b_2 b_3 \cdots b_k \tag{3.13}$$

for some $a \in M_1$ and $b_2 b_3 \cdots b_k \in M_2 M_3 \cdots M_k \subseteq M_2 \cap \cdots \cap M_k$. This shows that M_1 and $\bigcap_{j \neq 1} M_j$ are comaximal. Analogously, one shows that M_i and $\bigcap_{j \neq i} M_j$ are comaximal, for any $1 \leq i \leq k$.

Given any element $\gamma = (r_1 + M_1, \ldots, r_k + M_k) \in \prod_{i=1}^k R/M_i$, there exist elements $b_i \in M_i$ and $c_i \in \bigcap_{j \neq i} M_j$, for each $1 \leq i \leq k$, such that $r_i = b_i + c_i$. Let $x = c_1 + \cdots + c_k$, then

$$\Phi(x+I) = (x+M_1, \ldots, x+M_k) = (c_1+M_1, \ldots, c_k+M_k) = (r_1+M_1, \ldots, r_k+M_k) = \gamma,$$

because $c_j \in M_l$ for any $l \neq j$ and since $c_i - r_i = b_i \in M_i$. This shows that Φ is surjective. □

In the case of $R = \mathbb{Z}$ we have that the intersection of $n\mathbb{Z}$ and $m\mathbb{Z}$ is given by the least common multiple of n and m, which is defined as $\operatorname{lcm}(n,m) = \frac{n}{\gcd(n,m)} m = n \frac{m}{\gcd(n,m)}$, i.e.

$$n\mathbb{Z} \cap m\mathbb{Z} = \operatorname{lcm}(n,m)\mathbb{Z}. \tag{3.14}$$

To see this, note first that $\operatorname{lcm}(n,m)$ is a multiple of n and m, hence an element of $n\mathbb{Z} \cap m\mathbb{Z}$. Therefore, $\operatorname{lcm}(n,m)\mathbb{Z} \subseteq n\mathbb{Z} \cap m\mathbb{Z}$. Conversely, if $x \in n\mathbb{Z} \cap m\mathbb{Z}$, then $x = na = mb$ for some $a, b \in \mathbb{Z}$. In particular,

$$\frac{n}{\gcd(n,m)} a = \frac{x}{\gcd(n,m)} = \frac{m}{\gcd(n,m)} b \qquad \Rightarrow \qquad \frac{n}{\gcd(n,m)} \mid \frac{m}{\gcd(n,m)} b.$$

Since $\frac{n}{\gcd(n,m)}$ and $\frac{m}{\gcd(n,m)}$ are relatively prime, we conclude $\frac{n}{\gcd(n,m)} \mid b$. Hence

$$\operatorname{lcd}(n,m) = m \frac{nm}{\gcd(n,m)} \mid mb = x,$$

i.e. $n\mathbb{Z} \cap m\mathbb{Z} \subseteq \operatorname{lcd}(n,m)\mathbb{Z}$, proving equation (3.14). This means for ideals $n\mathbb{Z}$ and $m\mathbb{Z}$ of $\mathbb{Z}$ the following are equivalent:

n and m are relatively prime;

$\Leftrightarrow$ *$n\mathbb{Z}$ and $m\mathbb{Z}$ are comaximal ideals*

$\Leftrightarrow$ *$n\mathbb{Z} + m\mathbb{Z} = \mathbb{Z}$*

$\Leftrightarrow\ n\mathbb{Z} \cap m\mathbb{Z} = (nm)\mathbb{Z}.$

Identifying $\mathbb{Z}_n$ with $\mathbb{Z}/n\mathbb{Z}$, the Chinese Remainder Theorem says now the following:

Theorem 3.13 (Chinese Remainder theorem for integers) *Let $m_1, \ldots, m_s$ be positive numbers that are pairwise relatively prime and $n = m_1 \cdots m_s$. Then*

$$\Phi : \mathbb{Z}_n \to \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_s} \qquad \text{with} \qquad \Phi(x \bmod n) = (x \bmod m_1, \ldots, x \bmod m_s) \quad (3.15)$$

is an isomorphism of rings. In particular, for any integers $b_1, \ldots, b_s$ the system

$$\begin{cases} x &= b_1 (\bmod m_1) \\ &\vdots \\ x &= b_s (\bmod m_s) \end{cases}$$

has a unique solution modulo n.

P*roof: Taking $R = \mathbb{Z}$ and $M_i = m_i\mathbb{Z}$, we first note that $\bigcap_{i=1}^s M_i = \bigcap_{i=1}^s m_i\mathbb{Z} = n\mathbb{Z}$ since the m_i are pairwise relatively prime. Applying the Chines Remainder Theorem 3.12 we obtain the desired isomorphism Φ between*
ZZ_n and $\mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_s}$. The surjectivity of Φ yields that the given system has a solution and the injectivity of Φ shows that the solution is unique. $\square$

Note that if there exists an isomorphism of rings $\Phi : R \to S$, then it restricts to an isomorphism of their unit group $\Phi : U(R) \to U(S)$, because if $x \in U(R)$ is invertible in R with inverse x^{-1}. Then

$$\Phi(x)\Phi(x^{-1}) = \Phi(xx^{-1}) = \Phi(1_R) = 1_S = \Phi(1_R) = \Phi(x^{-1}x) = \Phi(x^{-1})\Phi(x).$$

This show that $\Phi(x)^{-1} = \Phi(x^{-1})$ and therefore $\Phi(x) \in U(S)$.

Since Φ is an isomorphism, there exists an isomorphism $\Phi^{-1} : S \to R$ that is the inverse function of Φ. Hence for any $y \in U(S)$, $\Phi^{-1}(y) \in U(R)$, with $y = \Phi(\Phi^{-1}(y))$. Showing that $\Phi : U(R) \to U(S)$ is actually bijective. As it is multiplicative, it is also a group homomorphism between $U(R)$ and $U(S)$.[1]

Now let $R_1 \times \cdots \times R_k$ be a product of rings, then it is not difficult to see (and left to the reader) that $U(R_1 \times R_2) = U(R_1) \times \cdots \times U(R_k)$.

This means that if a ring R is isomorphic to a direct product $R_1 \times \cdots \times R_k$, then the unit groups $U(R)$ is isomorphic to the direct product of unit groups $U(R_1) \times \cdots \times U(R_k)$. Applying the Chinese Remainder Theorem yields:

Corollary 3.14 *Let $n = p_1^{a_1} \cdots p_s^{a_s}$ a positive number with $s \geq 1$, primes $p_i \neq p_j$ for $i \neq j$ and $a_i \geq 1$. Then*

$$U(\mathbb{Z}_n) \simeq U(\mathbb{Z}_{p_1^{a_1}}) \times \cdots \times U(\mathbb{Z}_{p_s^{a_s}})$$

In particular $U(\mathbb{Z}_n)$ is cyclic if and only if

[1] Be aware that if Φ is not an isomorphism, the unit groups might be different. Consider for example the inclusion $\Phi : \mathbb{Z} \to \mathbb{Q}$. Then it is true that $\Phi(x)$ is invertible, for any invertible element $x \in \mathbb{Z}$, but there exist elements $x \in \mathbb{Z}$ such that $\Phi(x)$ is invertible in $\mathbb{Q}$, but x is not invertible in $\mathbb{Z}$ as $U(\mathbb{Z}) = \{\pm 1\}$ and $U(\mathbb{Q}) = \mathbb{Q} \setminus \{0\}$.

1. $n = p^a$, *for some odd prime* p *and* $a \geq 1$ *or*
2. $n = 2p^a$, *for some off prime* p *and* $a \geq 1$ *or*
3. $n = 2$ *or* $n = 4$.

P*roof: By the Chinese Remainder Theorem there exists an isomorphism of rings*

$$\mathbb{Z}_n \simeq \mathbb{Z}_{p_1^{a_1}} \times \cdots \times \mathbb{Z}_{p_s^{a_s}}$$

which induces an isomorphism of groups $U(\mathbb{Z}_n) \simeq U(\mathbb{Z}_{p_1^{a_1}}) \times \cdots \times U(\mathbb{Z}_{p_s^{a_s}})$.

Theorem 3.10 shows that $U(\mathbb{Z}_{p^a})$ *is cyclic for* p *an odd prime. Clearly* $U(\mathbb{Z}_2) = \{1\}$ *and* $U(\mathbb{Z}_4) = \{1, 3\} \simeq \mathbb{Z}_2$ *are cyclic groups and for* $n = 2p^a$, *with* p *an odd prime we have*

$$U(\mathbb{Z}_{2p^a}) \simeq U(\mathbb{Z}_2) \times U(\mathbb{Z}_{p^a}) = \{1\} \times U(\mathbb{Z}_{p^a}) \simeq U(\mathbb{Z}_{p^a})$$

is also a cyclic group.

Conversely, note that iany subgroup of a cyclic group is cyclic. Hence a group that contains a non-cyclic subgroup, like the Klein group $\mathbb{Z}_2 \times \mathbb{Z}_2$ *cannot be a cyclic group.*

Recall that if p *is an odd prime, then* $U(\mathbb{Z}_{p^a})$ *is a cyclic group by Theorem 3.10. Since the group is cyclic and* 2 *divides its even order* $p^{a-1}(p-1)$, *it has exactly one subgroup of order* 2 *by 3.6.*

Hence if n *has two odd prime divisors, say* p_i *and* p_j, $i \neq j$. *Then* $U(\mathbb{Z}_n)$ *has a subgroup isomorphic to* $U(\mathbb{Z}_{p_i^{a_i}}) \times U(\mathbb{Z}_{p_j^{a_j}})$ *which each containing a subgroup of order* 2 *and hence their direct product contains a subgroup isomorphic to the Klein group* $\mathbb{Z}_2 \times \mathbb{Z}_2$, *which is not cyclic. Therefore,* $U(\mathbb{Z}_n)$ *cannot be cyclic.*

If $n = 2^a$ *for* $a \geq 3$, *then* $U(\mathbb{Z}_{2^a}$ *is not cyclic by Theorem 3.11.*

Also if n *has an odd prime divisor* p^b *and is divisble by* 4, *then* $U(\mathbb{Z}_n)$ *contains a subgroup isomorphic to* $U(\mathbb{Z}_{2^a}) \times U(\mathbb{Z}_{p^b})$ *with* $a \geq 2$. *If* $a = 2$, $U(\mathbb{Z}_4) = \mathbb{Z}_2$ *and if* $a \geq 3$, *then* $U(\mathbb{Z}_{2^a})$ *also contains a subgroup of order* 2. *Hence in either case* $U(\mathbb{Z}_n)$ *contains again a copy of the Klein group and cannot be cyclic.*

This shows that if $U(\mathbb{Z}_n)$ *is cyclic, then the only possibilities for* n *are* $n = p^a$ *or* $n = 2p^a$, *for* p *an odd prime,* $n = 2$ *or* $n = 4$, *which are precisely the cases in the statement of the Theorem.*□

Whether a number a *is a primitive root of unity modulo a prime number* p, *i.e. whether* a *is a generator of* $U(\mathbb{Z}_p)$ *is not always easy to answer. The following example connects the order of the number* 10 *modulo a prime number* p *with the period of the decimal expression of* $1/p$. *Consider the fraction* $1/7$ *and express it in decimal form:*

$$\begin{aligned} \frac{1}{7} &= 0.\underbrace{142857}_{6}\underbrace{142857}_{6}\underbrace{142857}_{6}\cdots = 0.\overline{142857} \\ \frac{1}{11} &= 0.\underbrace{09}_{2}\underbrace{09}_{2}\underbrace{09}_{2}\cdots = 0.\overline{09} \end{aligned}$$

Why does $1/7$ *has periodicity* 6 *and* $1/11$ *periodicity* 2*? How is the periodicity related to the prime number and how large can the periodicity be? Let* p *be a prime different from* 2 *and* 5, *then* $1/p$ *has some periodicity* n*:*

$$\frac{1}{p} = \left(\frac{a_1}{10} + \cdots + \frac{a_n}{10^n}\right) + 10^{-n}\left(\frac{a_1}{10} + \cdots + \frac{a_n}{10^n}\right) + 10^{-2n}\left(\frac{a_1}{10} + \cdots + \frac{a_n}{10^n}\right) + \cdots$$

For $M = \frac{a_1}{10} + \cdots + \frac{a_n}{10^n}$ *we have, using the geometric series* $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$*, for* $0 < q < 1$*:*

$$\frac{1}{p} = \sum_{i=0}^{\infty} (10^{-n})^i M = \frac{1}{1-10^{-n}} M = \frac{10^n M}{10^n - 1}.$$

Equivalently this means $10^n - 1 = pM10^n$ *or in other words*

$$10^n = 1 (\bmod p)$$

Therefore, n *is divisible by the order of* 10 *in* $U(\mathbb{Z}_p)$*. The order of* 10 *modulo* p *is maximal (and equal to* $p-1$*) if and only if* 10 *is a primitive root of unity modulo* p*. For instance in the case* $p = 7$*,* 10 *has order* 6*, while* 10 *has order* 2 *modulo* 11*. What would be the periodicity of* $1/29$*?* [2]

Emil Artin asked in 1927 whether for a given square free number $a > 1$ *there are infinitely many primes* p *such that* a *is a primitive root of unity modulo* p*. This is called the Artin Conjecture which up to today has not been solved.*

How to calculate quickly powers modulo a number

In some cryptographic algorithm it is necessary to calculate rapidly expressions like $a^b (\bmod n)$*. Before addressing this problem we will first have a look at the representation of natural numbers with respect to some basis. Any integer can be written as a linear combination of powers of a given positive base number. We will show this using the Peano axioms.*

Proposition 3.15 *Let* $b \in \mathbb{N}$ *with* $b > 1$*. For any natural number* $a \in \mathbb{N}$ *there exists a number* $n \geq 0$ *and numbers* $a_1, \ldots, a_k \in \{0, 1, \ldots, b-1\}$ *such that*

$$a = a_0 + a_1 \cdot b + a_2 \cdot b^2 + \cdots + a_n \cdot b^n = \sum_{k=0}^{n} a_k \cdot b^k. \tag{3.16}$$

In this case we will write $a = (a_0, \ldots, a_n)_b$ *to indicate the b-ary representation of* a*.*

P*roof: We will show this statement by induction on* a*. Let* $b > 1$ *be fixed and consider the set*

$$M = \left\{ a \in \mathbb{N} \mid \exists n \geq 0, a_1, \ldots, a_k \in \{0, 1, \ldots, b-1\} : a = \sum_{k=0}^{n} a_k \cdot b^k \right\}. \tag{3.17}$$

Clearly $0 \in M$ *setting* $n = 0$ *and* $a_0 = 0$*. Hence suppose* $a \in M$*. We have to show that* $S(a) = a + 1 \in M$*. As* $a \in M$*, there exist* $n \geq 0, a_1, \ldots, a_k \in \{0, 1, \ldots, b-1\}$ *such that* $a = \sum_{k=0}^{n} a_k \cdot b^k$*. If* $a_0 < b - 1$*, then set* $a_0' = a_0 + 1$ *and* $a_k' = b_k$ *for all* $k > 0$*. Thus* $a + 1 = \sum_{k=0}^{n} a_k' \cdot b^k$*. If* $a_0 = b - 1$*, then choose the least index* $0 \leq m \leq n$ *such that* $a_k = b - 1$ *for all* $k \leq m$*. Note that*

$$\sum_{k=0}^{m} (b-1) b^k + 1 = (b-1) \left(\sum_{k=0}^{m} b^k \right) + 1 = b^{m+1} - 1 + 1 = b^{m+1}. \tag{3.18}$$

[2] answer 28

If $m = n$, *then* $a + 1 = b^{n+1} \in M$. *If* $m < n$, *then set* $a'_k = \begin{cases} 0 & \text{if } k \leq m \\ a_{m+1} + 1 & \text{if } k = m+1 \\ a_k & \text{if } k > m \end{cases}$ *and we have* $a + 1 = \sum_{k=0}^{n} a'_k \cdot b^k \in M$. *By the principle of induction* $M = \mathbb{N}$. □

Only a finitely many numbers can be represented by a computer, since its memory is finite. The smallest information unit is a bit which can have the values 0 *or* 1. *Using binary representation* ($b = 2$) *any number can be expressed as a list of bits. For example one* byte *consists of* 8 *bits and can therefore represent the (positive) numbers between* 0 *and* $255 = 2^8 - 1$. *One kilo-, mega-, giga- and terabyte consist of* 2^{10}, 2^{20}, 2^{30} *and* 2^{40} *bytes. For instance with one kilobyte consists of* $8 \cdot 2^{10} = 8192$ *bits. Hence one can represent the positive numbers from* 0 *to* $2^{8192} - 1$ *with the memory of one kilobyte. A gigabyte consists of* $8 \cdot 2^{30} = 2^{33} = 8589934592$ *bits and hence can represent the numbers from* 0 *to* $2^{8589934592} - 1$. *These numbers are extremely large. However they are still (finite) numbers.*

How to calculate $a^b \pmod n$ *quickly for large* b? *Multiplication by* 2, *division by* 2 *and calculating the rests mod* 2 *is easy in binary representation. Given* $a = (a_0, a_1, \ldots, a_n)_2$ *one has*

$$2a = (0, a_0, a_1, \ldots, a_n)_2 \qquad a/2 = (a_1, \ldots, a_n)_2 \qquad a \bmod 2 = a_0. \tag{3.19}$$

Notice that multiplication and division by 2 *correspond to a shift to the right resp. to the left of the string* $(a_0, \ldots, a_n)$. *There are hardware solutions that can do this operation quickly. The idea of the following algorithm to calculate* $a^b \pmod n$ *is that if the binary representation of* b *is* $(b_0, \ldots, b_n)_2$, *then* $b = \sum_{k=0}^{n} b_k 2^k$ *and*

$$a^b = a^{\sum_{k=0}^{n} b_k 2^k} = \prod_{k=0}^{n} a^{b_k 2^k} = \prod_{b_k \neq 0} a^{2^k}. \tag{3.20}$$

For $k < l$ *and* $b_k \neq 0 \neq b_l$ *we can use* a^{2^k} *in order to calculate* a^{2^l}, *which is going to be the advantage of the following algorithm whose runtime is* $n = \log_2(b)$.

Data: $a, b, n \in \mathbb{N}$ *with* $n, b > 1$.
Result: $a^b \pmod n$.
$r = 1$
while $b > 0$ **do**
  if $b\%2 == 1$ **then**
    $r = (r \cdot a)\%n$;
  end
  $a = (a \cdot a)\%n$;
  $b = b/2$;
end
return r;

Algorithm 2: Fast modular power algorithm

Example 3.16 *Let us try to calculate* $14^{1000} \pmod{71}$. *The algorithm will perform the following steps:*

$b\%2 == 1?$	r	a	b
	1	14	1000
no	1	$14^2 \equiv 54 \mod 71$	500
no	1	$54^2 \equiv 5 \mod 71$	250
no	1	$5^2 = 25$	125
yes	25	$25^2 \equiv 57 \mod 71$	62
no	25	$57^2 \equiv 54 \mod 71$	31
yes	$25 \cdot 54 \equiv 1 \mod 71$	$54^2 \equiv 5 \mod 71$	15
yes	$1 \cdot 5 = 5$	$5^2 = 25$	7
yes	$5 \cdot 25 \equiv 54 \mod 71$	$25^2 \equiv 57$	3
yes	$54 \cdot 57 \equiv 25 \mod 71$	$57^2 \equiv 54$	1
yes	$25 \cdot 54 \equiv 1 \mod 71$	$54^2 \equiv 5$	0

Hence $14^{1000} \bmod 71 = 1$.

Exercises

Ex. 7 — *Write a program that given a number $b > 1$ and a number $a \geq 0$ yields the list of coefficients $a_0, \ldots, a_n$ such that $a = \sum_{k=0}^{n} a_k b^k$.*

Ex. 8 — *Show that the representation of a positive number $a = \sum_{k=0}^{n} a_k b^k$ as a linear combination of powers of b is unique with respect to all such representations where the highest term a_n is non-zero.*

4

Euclidean domains

Let R be a **commutative** *unital non-trivial ring and $a, b \in R$ elements. Like for numbers, we say that a divides b, in symbol $a \mid b$, if and only if there exists $c \in R$ with $b = ac$.*

Note that $1 \mid a$ holds for any $a \in R$ and $a \mid 1$ holds if and only if $a \in U(R)$.

Moreover, any element $a \in R$ divides (trivially) 0, because $0 = a \cdot 0$, i.e. $a \mid 0$. However, we will call an element a a zero divisor in R, if there exists a non-zero $b \in R$ such that $ab = 0$. A (non-trivial)[1] commutative ring R is called an integral domain if 0 is the only zero divisor of R. This means that whenever $ab = 0$, then $a = 0$ or $b = 0$. Clearly any field is an integral domain, because if $ab = 0$ and $b \neq 0$, then $a = abb^{-1} = 0$, i.e. 0 is the only zero divisor. We have already seen that for $n \geq 2$, $R = \mathbb{Z}_n$ is a field if and only if n is a prime number. It is not difficult to check and left as an exercise that $\mathbb{Z}_n$ is an integral domain if and only if $\mathbb{Z}_n$ is a field. Clearly this is not true for an arbitrary integral domain like for example $\mathbb{Z}$.

The set of multiples of a, $aR = \{ac : c \in R\}$ is an ideal and called the ideal generated by a in R. By definition, if $a \mid b$, then $b = ac$ for some $c \in R$. Therefore, for any $r \in R$, $br = acr$, i.e. $bR \leq aR$. Conversely, if $bR \leq aR$, then $b \in aR$ and hence there exists $c \in R$ with $b = ac$, i.e. $a \mid b$. This shows that for any $a, b \in R$:

$$a \mid b \qquad \text{if and only if} \qquad bR \leq aR \tag{4.1}$$

Thus, the divisibility relation is reflexive and transitive, but not necessarily anti-symmetric and therefore not an equivalence relation. For instance $2 \mid (-2)$ and $(-2) \mid 2$, but $2 \neq -2$ or more generally, if $u \in U(R) \setminus \{1\}$ is an invertible element in R and $a \in R$ is a non-zero element, then $a \mid ua$ and $ua \mid a$, but $a \neq ua$. This motivates the following definition:

Definition 4.1 *Let R be a commutative unital non-trivial ring. Then two non-zero elements $a, b \in R$ are said to be associated in R if and only if $a \mid b$ and $b \mid a$.*

[1] $0 \neq 1$

From (4.1) we deduce immediately that two elements $a, b \in R \setminus \{0\}$ are associated if and only if they generate the same ideal, i.e. $aR = bR$. The relation of being associated is an equivalence relation and its equivalence classes $[a]$ form a partition of $R \setminus \{0\}$. The divisibility relation defines then a partial ordering on the set of equivalence classes.

It is also clear by the previous comments that a and ua are associated, for any unit $u \in R$ and element $a \in R$, i.e. $\{au : u \in U(R)\} \subseteq [a]$. The converse is true for non-zero associated elements in an integral domain R, because if a and b are associated and non-zero, then there exist $c, d \in R$ such that $b = ac$ and $a = bd$. Thus $b = ac = bdc$ implies $b(1 - dc) = 0$. Since b is non-zero and R is an integral domain, we must have $1 - dc = 0$, i.e. $dc = 1$. Hence $c, d \in U(R)$, which shows $[a] = \{au : u \in U(R)\}$.

For any ring R we can define the power series ring $R[[x]]$ with coefficient in R. Each series can be thought of as a sequence $(a_n)_{n\in\mathbb{N}}$ of elements of R and is usually represented as a power series:

$$f = \sum_{n=0}^{\infty} a_n x^n.$$

Given $f = \sum a_n x^n$ and $g = \sum b_n x^n$ in $R[[x]]$, addition and multiplication are defined as

$$f + g = \sum_{n=0}^{\infty} (a_n + b_n) x^n \qquad f \cdot g = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i} \right) x^n.$$

The zero element of $R[[x]]$ is the polynomial with all coefficients 0, while the identity of $R[[x]]$ is defined as the polynomial $\sum_{n=0}^{\infty} a_n x^n$ with $a_0 = 1$ and $a_n = 0$, for all $n \geq 1$. The support of $f \in R[[x]]$ is defined as

$$\sup(f) = \{n \in \mathbb{N} : a_n \neq 0\}.$$

Note that $\sup(f) = \emptyset$ if and only if $f = 0$ is the zero element. In general power series might have infinite support. The polynomial ring $R[x]$ is defined as the set of series that have finite support, i.e.

$$R[x] = \{f \in R[[x]] : \sup(f) \text{ is finite } \}.$$

It is not difficult to check that $R[x]$ is closed under addition and multiplication and forms a ring. Moreover, $f \sum_{n=0}^{\infty} a_n x^n \in R[x]$ if and only if there exists some number $n \in \mathbb{N}$ such that $a_m = 0$ for all $m \geq n$. The degree $\deg(f)$ of a non-zero polynomial $f = \sum a_i x^i \in R[x]$ is set to be the maximum of the supremum, i.e. $\deg(f) = \max(\sup(f)) = N$ if and only if $a_N \neq 0$ and $a_n = 0$ for all $n > N$. The leading coefficient of f is then a_N. The degree of the zero polynomial is set to be $\deg(0) = -\infty$, which is merely a symbol to extend the natural numbers by an element that satisfies $(-\infty) + n = n + (-\infty) = -\infty$ for any $n \neq 0$.[2] The leading coefficient of the zero polynomial is not defined.

Lemma 4.2 *Let R be an integral domain. Then the polynomial ring $R[x]$ is an integral domain.*

Proof: Let $f = \sum a_n x^n$ and $g = \sum b_n x^n$ be non-zero polynomials with $\deg(f) = N$ and $\deg(g) = M$ and leading coefficients a_N and b_M. Then

$$fg = \left(\sum_{n=0}^{\infty} a_n x^n \right) \left(\sum_{m=0}^{\infty} b_m x^m \right) = \sum_{n=0} \left(\sum_{i=0}^{n} a_i b_{n-i} \right) x^n.$$

[2] In semigroup theory, the symbol $-\infty$ would be called an *absorbing* element in $(\mathbb{N} \cup \{-\infty\}, +)$

Examining the coefficient of x^n *with* $n > N + M$ *yields:*

$$\sum_{i=0}^{n} a_i b_{n-i} = \sum_{i=0}^{N} a_i b_{n-i} = 0,$$

because $a_i = 0$ *for* $i > N$ *and* $b_{n-i} = 0$ *for* $i \leq N$*, because then* $n - i > N + M - i \geq M$*. This shows that* $\deg(fg) \leq N + M = \deg(f) + \deg(g)$*. Similarly if we examine the coefficient of* x^{N+M} *yields:*

$$\sum_{i=0}^{N+M} a_i b_{N+M-i} = \sum_{i=0}^{N} a_i b_{N+M-i} = a_N b_M.$$

Therefore, $\deg(fg) = N + M = \deg(f) + \deg(g)$ *if and only if* $a_N b_M \neq 0$*. In particular, if* R *is an integral domain, then for any non-zero polynomials* f, g *we obtain* $\deg(fg) = \deg(f) + \deg(g) \geq 0$*, which shows that* $fg \neq 0$*.* □

Like the integers, the ring of polynomials $K[x]$ *over a field* K *has a division algorithm. More concretely, for any* $f, g \in K[x]$ *with* $g \neq 0$*, there exist* $q, r \in K[x]$ *such that*

$$f = qg + r \qquad \text{and} \qquad \deg(r) < \deg(g).$$

Here we are making use of our convention that $\deg(r) = -\infty$ *if* $r = 0$*.*

Definition 4.3 *An integral domain* R *is called an Euclidean domain if there exists a function* $\lambda : R \setminus \{0\} \to \mathbb{N}$ *such that for any* $a, b \in R$ *with* $b \neq 0$ *there exist* $q, r \in R$ *with* $a = qb + r$ *and* $r = 0$ *or* $\lambda(r) < \lambda(b)$*.*

The importance of Euclidean Domains is that the Euclidean algorithm holds for them. Define the following degree function on $\mathbb{Z}$*: for all* $a \in \mathbb{Z} \setminus \{0\}$ *set* $\deg(a) =\mid a \mid$*. With this function* $\mathbb{Z}$ *becomes an Euclidean domain with the ordinary division algorithm for* $\mathbb{Z}$*.*

Another example stems from the ring of polynomials over a field, $R = K[x]$*. Here the ordinary polynomial division and the degree function* $\deg : K[x] \setminus \{0\} \to \mathbb{N}$ *are the ingredients to turn* $K[x]$ *into an Euclidean domain.*

A common divisor of two elements a *and* b *is an elements* $d \in D$ *such that* $d \mid a$ *and* $d \mid b$*. A greatest common divisor of* a *and* b *is a common divisor* d *of* a *and* b *such that if* d' *is another common divisor of* a *and* b*, then* $d' \mid d$*. The greatest common divisor of* a *and* b*, if exists, is denoted by* $\gcd(a, b)$*. Let* D *be an Euclidean domain with degree function* deg*. If* d *is a greatest common divisor of* a *and* b *in* D*, then for any other common divisor* d' *of* a *and* b *one has* $\deg(d') \leq \deg(d)$ *(see exercises). Note that for any greatest common divisor* d *of* a *and* b *and invertible element* u*, one also has that* ud *is a greatest common divisor. Thus a greatest common divisor is defined up to an invertible factor and in certain classes of rings one requires further conditions. For example in the ring of integers one defines* the *greatest common divisor to be a positive greatest common divisor, while* the *greatest common divisor of two polynomials over a field is usually defined as the unique monic greatest common divisor.*

Assume that D *is an Euclidean domain. If* $a, b \in D$ *with* $b \neq 0$ *there exist* $q, r \in D$ *with*

$$a = q \cdot b + r \qquad \text{and } r = 0 \text{ or } \deg(r) < \deg(b). \tag{4.2}$$

Call q *the* quotient of a by b*, denoted by* $q = a/b$ *and* r *the* remainder of a divided by b*, denoted by* $r = a(\bmod b)$*.*

Proposition 4.4 *Let D be an Euclidean domain and $a, b \in D$. The set of common divisors of a and b is equal to the set of common divisors of b and $a(\bmod b)$.*

P*roof: By the division algorithm, there exist $q, r \in D$ such that $a = q \cdot b + r$. If $d \mid a$ and $d \mid b$, then $d \mid a - q \cdot b = r = a(\bmod b)$. Hence d is a common divisor of b and $a(\bmod b)$. On the contrary, if $d \mid b$ and $d \mid a(\bmod b)$, then $d \mid q \cdot b + a(\bmod b) = a$.* □

Proposition 4.4 allows us to replace the pair (a, b) by $(b, a(\bmod b))$ if we want to calculate the greatest common divisor. Applying the Proposition again leads to $(a(\bmod b), b(\bmod a(\bmod b)))$. Each time we apply the Proposition the degree of the second component goes down, i.e.

$$\deg(b) > \deg(a(\bmod b)) > \deg(b(\bmod a(\bmod b))) > \cdots . \tag{4.3}$$

Thus, this procedure leads to the following Algorithm which terminates since deg *takes values into $\mathbb{N}$ which is a well-ordered set.*

Data: *$a, b \in D$ an Euclidean domain*
Result: *a greatest common divisor of a and b.*
if $b \neq 0$ **then**
 | *return* $\gcd(b, a(\bmod b))$
else
 | *return a;*
end

Algorithm 3: Euclidean Algorithm

Proposition 4.4 secures that a greatest common divisor of a and b is the same as a greatest common divisor of b and $a\%b$. It is clear that a greatest common divisor of a and 0 is a. The algorithm terminates because the set of degrees of remainders is a non-empty subset of $\mathbb{N}$ and must have a least element. In the exercises you are invited to write a computer program to find the greatest common divisor of two integers.

Proposition 4.5 *Let D be an Euclidean domain and $a, b \in D$ not both zero. Then there exist $r, s \in D$ such that $\gcd(a, b) = ra + sb$.*

P*roof: Consider $I = \{\deg(ra+sb) \in \mathbb{N} : ra+sb \neq 0$ and $r, s \in D\} \subseteq \mathbb{N}$. As not both elements a, b are zero, $I \neq$. As $\mathbb{N}$ is well-ordered, I has a minimum, say $n = \deg(ra + sb)$ for some $r, s \in D$. We claim that $d = ra + sb$ is a greatest common divisor of a and b. By the division algorithm for Euclidean domains, there exist $q, t \in D$ such that $a = qd + t$ and either $t = 0$ or $t \neq 0$ and $\deg(t) < \deg(d) = n$. However, since $t = a - qd = (1 - r)a - sb$, we have that either $t = 0$ or $\deg(t) \in I$. The second option is impossible, as $n = \deg(d)$ is the minimum of I and $\deg(t)$ is smaller than n. Hence $t = 0$ and $a = qd$, i.e. $d \mid a$. Similarly one shows $d \mid b$. If d' is another common divisor of a and b, then $d' \mid d = ra + sb$. Thus d is a greatest common divisor of a and b and can be written as $d = ra + sb$.* □

As in the case of the integers, we have the extended Euclidean Algorithm to calculate $\gcd(a, b)$ of two elements $a, b \in D$ in an Euclidean domain as well as the elements $r, s \in D$ that satisfy $\gcd(a, b) = ra+sb$. This is crucial for many applications. As in the case of the integers, the algorithm uses the following reasoning: suppose $\gcd(b, a(\bmod b)) = x \cdot b + y \cdot (a\%b)$ for some $x, y \in D$, where $a\%b$ denotes the rest of the division of a by b. Since $a = (a/b)b + a(\bmod b)$

it follows that $a(\bmod b) = a - (a/b) \cdot b$, *where* a/b *denotes the quotient of the division of* a *by* b. *Substituting this expression for* $a(\bmod b)$ *in the formula for* $\gcd(b, a(\bmod b))$ *yields:*

$$\gcd(a,b) = \gcd(b, a(\bmod b)) = x \cdot b + y \cdot (a - (a/b) \cdot b) = y \cdot a + (x - (a/b) \cdot y) \cdot b. \quad (4.4)$$

Hence on the "way back" from our recursion, we can adjust the coefficients x *and* y *by replacing* x *with* y *and* y *with* $x - (a/b) \cdot y$.

Data: $a, b \in D$ *an Euclidean domain*
Result: d, x, y *where* $d = \gcd(a,b)$ *and* $d = x \cdot a + y \cdot b$.
if $b \neq 0$ **then**
  $d, x, y = gcd(b, a(\bmod b))$;
  return $(d, y, x - (a/b)y)$;
else
  return (a,1,0);
end

Algorithm 4: Extended Euclidean Algorithm

An integral domain all whose ideals are principal is called a principal ideal domain.

Lemma 4.6 *Any Euclidean domain is a principal ideal domain.*

P*roof: Given a non-zero ideal* I *of an Euclidean domain* R. *Consider* $\Lambda = \{\lambda(r) : r \in I\} \subseteq \mathbb{N}$. *Since* $\mathbb{N}$ *is well ordered, there exists a non-zero element* $b \in I$ *with* $\lambda(b)$ *minimal. Hence, for any* $a \in I$, *there are* $q, r \in R$ *with* $a = qb + r$ *and* $r = 0$ *or* $\lambda(r) < \lambda(b)$. *If* $r \neq 0$, *then* $r = a - qb \in I$ *has lower* λ*-value than* b *which contradicts its minimality. Therefore* $r = 0$ *and* $a = qb$, *i.e.* $I = \langle b \rangle$. □

Recall the notion of irreducible and prime elements in an integral domain R. *Let* p *be a non-zero, non-invertible element of* R. *Then* p *is called an irreducible element if and only if for any* $a, b \in R$ *if* $p = ab$, *then* $a \in U(R)$ *or* $b \in U(R)$.

The element p *is called a prime element if and only if for any* $a, b \in R$ *such that* $p \mid ab$, *then* $p \mid a$ *or* $p \mid b$. *In general, prime and irreducible elements might be different. For principal ideal domains however the notions coincide.*

Lemma 4.7 *Let* R *be a PID and* $p \in R \setminus U(R) \cup \{0\}$. *Then* p *is irreducible if and only if* p *is prime if and only if* R/pR *is a field.*

P*roof: Let* $P = pR$. *Suppose* p *is an irreducible element and let* $a \in R$ *such that* $p \nmid a$ *or equivalently* $a + P \neq 0 + P\rangle$ *in* R/P. *Consider the ideal* $I = pR + aR$ *generated by* p *and* a. *Since* R *is a PID,* $I = cR$ *for some* $c \in R$. *Hence* $p = cd$ *for some* $d \in R$. *As* p *is irreducible,* $d \in U(R)$ *or* $c \in U(R)$. *If* $d \in U(R)$, *then* $c = d^{-1}p$ *and as* $a \in I = cR$ *there exists* $e \in R$ *with* $a = ec = ed^{-1}p$, *i.e.* $p \mid a$, *which contradicts* $p \nmid a$. *Hence* $c \in U(R)$ *and* $R = I = pR + aR$. *Thus there exist* $x, y \in R$ *with* $1 = xa + yp$, *i.e.* $1 + P = (x + P)(a + P)$, *i.e.* $a + P \in U(R/P)$. *This shows that any non-zero element of* R/pR *is invertible, i.e.* R/pR *is a field.*

If R/P *is a field, then* p *is certainly prime, because if* $p \mid ab$ *for some* $a, b \in R$, *then* $(a + P)(b + P) = 0 + P$ *in* R/P *and as* R/P *is a field,* $a \in P$ *or* $b \in P$, *i.e.* $p \mid a$ *or* $p \mid b$.

If p *is a prime element, then it is always irreducible.* □

A (commutative) ring R is called Noetherian if any ascending chain of ideals

$$I_1 \subseteq I_2 \leq I_3 \subseteq \cdots$$

becomes stationary, i.e. if there exists $n \geq 1$ such that $I_n = I_m$, for all $m \geq n$.

Lemma 4.8 *A principal ideal domain is Noetherian.*

P*roof: If I_j are ideals of a principal ideal domain R that form a chain $I_1 \subseteq I_2 \subseteq \cdots$ then the union $I = \bigcup_{i\geq 1} I_i$ is an ideal of R (why?). Since R is a principal ideal domain, $I = aR$ for some $a \in I$. But then there exists an index $n \geq 1$ such that $a \in I_n$. As for all $m \geq n$,*

$$I_m \subseteq \bigcup_{i\geq 1} I_i = I = aR \subseteq I_n \subseteq I_m$$

we conclude $I_n = I_m$. □

Lemma 4.9 *Let R be a (commutative) Noetherian integral domain. Then any non-zero, non-invertible element is a product of irreducible elements.*

P*roof: Set*

$$\mathcal{P} = \{a \in R : \text{ } a \text{ is non-zero, non-invertible and not a product of irreducible elements}\}$$

Suppose $\mathcal{P} \neq \emptyset$. Then there exists an element $a_1 \in \mathcal{P}$ which is non-zero, non-invertible and not a product of irreducible, thus not irreducible itself. Hence there exist non-invertible elements $a_2, b_2 \in R$ with $a_1 = a_2 b_2$. Both a_2, b_2 are non-zero and cannot be both products of irreducibles, as otherwise a_1 would be a product of irreducibles. Thus one of them must belong to $\mathcal{P}$. Without loss of generality we can assume a_2 and in particular we have $a_1 R \subsetneq a_2 R$. Note that if $a_1 R = a_2 R$, then $a_2 = a_1 c = a_2 b_2 c$, for some $c \in R$, implies b_2 invertible, which is a contradiction. Hence the inclusion $a_1 R \subset a_2 R$ is proper. By the same argument, there must exist $a_3 \in \mathcal{P}$ such that $a_1 R \subsetneq a_2 R \subsetneq a_3 R$. Continuing in this way we obtain an infinite (properly) ascending chain of ideals which would not become stationary and would contradict R to be Noetherian. Hence $\mathcal{P}$ must be empty, i.e. any element is either 0, invertible or a product of irreducibles. □

Note that the proof only needs that the ascending chain of principal ideals holds. For a principal ideal domain (and in particualar for an Euclidean domain) we have just proven that any non-zero, non-invertible element is a product of irreducible (=prime) elements. However, in general a product of irreducible elements might not be unique and we make the following definition.

Definition 4.10 *An integral domain R is called a unique factorisation domain if every non-zero, non-invertible element a of R has a unique factorisation in irreducible elements, i.e.*

1. *every non-zero, non-invertible element is a product of irreducible elements.*
2. *for any two sets $\{p_1, \ldots, p_k\}$ and $\{q_1, \ldots, q_n\}$ of irreducible elements, such that $p_1 \cdots p_k = q_1 \cdots q_n$ one has $n = k$ and a permutation $\sigma \in S_n$ such that p_i and $q_{\sigma(i)}$ are associated, for any $1 \leq i \leq n$.*

Lemma 4.11 *Any Noetherian integral domain such that irreducible elements are prime is a unique factorisation domain. In particular any principal ideal domain and any Euclidean domain is a unique factorization domain.*

P*roof: By Lemma 4.9, any non-zero, non-invertible element is a product of irreducible elements, which are prime by hypothesis. The uniqueness follows now by induction on the number of prime factors: if* $p_1 \cdots p_n = q_1 \cdots q_m$*, with* p_i *and* q_i *prime, then* $p_1 \mid q_1 \dots q_m$ *implies* $p_1 \mid q_i$ *for some* i *and hence* p_1 *and* q_i *are associated. After rearranging we might assume* $i = 1$ *and can factor out* p_1 *to obtain a shorter product. Continuing this process yields* $n = m$ *and* $p_i \simeq q_{\sigma(i)}$*, for some permutation* $\sigma \in S_n$*.* $\square$

The Gaussian Integers

One particular example of Euclidean domains is the ring of Gaussian integers[3]*, this is the subring*

$$\mathbb{Z}[i] = \{a + bi \in \mathbb{C} : a, b \in \mathbb{Z}\}$$

of $\mathbb{C}$ *consisting the complex numbers with integer coefficients.*

Clearly the multiplication rule

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

shows that if $a, b, c, d \in \mathbb{Z}$*, then* $(a + bi)(c + di) \in \mathbb{Z}[i]$*. Since complex conjugation*

$$a + bi \mapsto \overline{a + bi} := a - bi$$

is a ring isomorphism of $\mathbb{C}$*, it is multiplicative, i.e.* $\overline{\alpha\beta} = \overline{\alpha}\overline{\beta}$*, for all* $\alpha, \beta \in \mathbb{C}$*. The norm of a complex number* $\alpha = a + bi \in \mathbb{C}$ *is defined to be* $N(\alpha) = \alpha\overline{\alpha} = a^2 + b^2 \in \mathbb{R}_{\geq 0}$*. In particular,* $N(\alpha) \in \mathbb{N}$ *for all* $\alpha \in \mathbb{Z}[i]$*. Moreover, the norm is multiplicative, i.e.*

$$N(\alpha\beta) = \alpha\beta\overline{\alpha\beta} = \alpha\beta\overline{\alpha}\overline{\beta} = \alpha\overline{\alpha}\beta\overline{\beta} = N(\alpha)N(\beta).$$

The multiplicative group of invertible elements of $\mathbb{Z}[i]$ *is* $U(\mathbb{Z}[i]) = \{1, -1, i, -i\}$*. It is easy to see that* $1, -1, i, -i$ *are invertible. Let* $\alpha \in U(\mathbb{Z}[i])$ *and* $\beta \in \mathbb{Z}[i]$ *such that* $\alpha\beta = 1$*, then*

$$1 = N(1) = N(\alpha\beta) = N(\alpha)N(\beta).$$

Hence $N(\alpha) = 1 = N(\beta)$ *as we have seen* $U(\mathbb{Z}) = \{1, -1\}$ *and* $N(\alpha) \geq 0$*. But if* $\alpha = a + bi$ *and* $a^2 + b^2 = N(\alpha) = 1$*, then* $a = \pm 1$ *and* $b = 0$ *or* $a = 0$ *and* $b = \pm 1$*, since otherwise* $N(\alpha) \geq 2$*. This shows*

$$U(\mathbb{Z}[i]) = \{1, -1, i, -i\}. \tag{4.5}$$

As a group $U(\mathbb{Z}[i]) \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$ *is the Klein group.*

In contrast to $\mathbb{Z}[i]$*, the subring* $\mathbb{Q}[i] = \{a + bi \in \mathbb{C} : a, b \in \mathbb{Q}\}$ *is a subfield of* $\mathbb{C}$*, because for any non-zero* $\alpha \in \mathbb{Q}[i]$*, also* $N(\alpha) \neq 0$ *and hence invertible in* $\mathbb{Q}$ *and hence* $\alpha^{-1} = \frac{1}{N(\alpha)}\overline{\alpha}$ *is the inverse of* α*.*

We will show that $\mathbb{Z}[i]$ *has a division algorithm and that* $\mathbb{Z}[i]$ *is an Euclidean domain and therefore a principal ideal domain and a unique factorization domain. But first wee need an elementary Lemma about* $\mathbb{Z}$ *and* $\mathbb{Q}$*.*

[3] Johann Carl Friederich Gauss (1777-1855), Biography: MacTutor

Lemma 4.12 *For any $x \in \mathbb{Q}$ there exists an integer $n \in \mathbb{Z}$ such that $|x - n| \leq \frac{1}{2}$.*

Note that n might not be uniquely determined, if for instance $x = \frac{1}{2}$, then $n = 0$ or $n = 1$ both satisfy $|x - n| = \frac{1}{2}$.
P*roof: If $x = 0$, then we can choose $n = 0$. Hence assume $x = \frac{a}{b} \neq 0$ such that a and b are relatively prime and $b > 0$. By the division algorithm in $\mathbb{Z}$, there exist $q, r \in \mathbb{Z}$ such that $a = qb + r$ and $0 \leq r < b$. If $2r \leq b$, then we set $n = q$ and obtain*

$$|x - n| = \frac{a - qb}{b} = \frac{r}{b} \leq \frac{1}{2}$$

If $b < 2r$, then $2(b - r) < b$ and we set $n = q + 1$ and obtain

$$|x - n| = \left|\frac{a - (q+1)b}{b}\right| = \left|\frac{r - b}{b}\right| = \frac{b - r}{b} < \frac{1}{2}.$$

□

We can prove now the division algorithm for $\mathbb{Z}[i]$.

Theorem 4.13 *Let $\alpha, \beta \in \mathbb{Z}[i]$ be two elements with $\beta \neq 0$. Then there exists $\gamma, r \in \mathbb{Z}[i]$ such that*

$$\alpha = \gamma\beta + r \qquad \text{with } N(r) < N(\beta).$$

In particular, $\mathbb{Z}[i]$ is an Euclidean domain, hence a principal ideal domain and a unique factorization domain.

P*roof: Lemma 4.12 shows that for any $\rho = x + yi \in \mathbb{Q}[i]$, there exist $a, b \in \mathbb{Z}$ such that $|x - a| < \frac{1}{2}$ and $|y - b| < \frac{1}{2}$. Hence for $\gamma = a + bi$ we obtain*

$$N(\rho - \gamma) = (x - a)^2 + (y - b)^2 = |x - a|^2 + |y - b|^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Let $\alpha = a + bi, \beta = c + di \in \mathbb{Z}[i]$ with $\beta \neq 0$ and consider

$$\frac{\alpha}{\beta} = \frac{1}{N(\beta)}\alpha\overline{\beta} \in \mathbb{Q}[i].$$

Then by the previous argument, there exists $\gamma \in \mathbb{Z}[i]$ such that $N\left(\frac{\alpha}{\beta} - \gamma\right) \leq \frac{1}{2}$. Therefore, for $r = \alpha - \gamma\beta \in \mathbb{Z}[i]$ we have $\alpha = \gamma\beta + r$ and

$$N(r) = N(\alpha - \gamma\beta) = N\left(\frac{\alpha}{\beta} - \gamma\right) N(\beta) \leq \frac{1}{2}N(\beta) < N(\beta).$$

As a subring of the field $\mathbb{C}$, $\mathbb{Z}[i]$ is an integral domain, which has division algorithm as we have seen with respect to the norm $N : \mathbb{Z}[i] \to \mathbb{N}$. Hence $\mathbb{Z}[i]$ is an Euclidean domain. □

For example if we want to divide $\alpha = 12 + 3i$ por $\beta = 2 + 2i$, then we can form the fraction in $\mathbb{Q}[i]$:

$$\frac{12 + 3i}{2 + 2i} = \frac{(12 + 3i)(2 - 2i)}{(2 + 2i)(2 - 2i)} = \frac{30}{8} - \frac{18}{8}i = 4 - 2i + (\frac{1}{4} - \frac{1}{4}i)$$

where $N(-\frac{1}{4}+\frac{1}{4}i)=\frac{1}{8}<\frac{1}{2}$. *Choosing* $\gamma=4-2i$ *and*

$$r=\alpha-\gamma\beta=12+3i-(2+2i)(4-2i)=12+3i-(12+4i)=-i$$

we obtain $\alpha=\gamma\beta+r$ *and* $1=N(r)<N(\beta)=8$. *Note that* $-i$ *is invertible, hence* α *and* β *are relatively prime in* $\mathbb{Z}[i]$. *A greatest common divisor of two elements of* $\mathbb{Z}[i]$ *is unique up to an invertible element, this is up to a multiple of* $1,-1,i$ *or* $-i$. *In particular we have*

$$1=i\alpha-i\gamma\beta=i\alpha-(2+4i)\beta.$$

Here is a second example: Let $\alpha=11+3i$ *and* $\beta=1+8i$. *Then*

$$\begin{aligned}11+3i &= (1-i)(1+8i)+2-4i\\ 1+8i &= (-1+i)(2-4i)-1+2i\\ 2-4i &= (-2)(-1+2i)+0\end{aligned}$$

Hence $-1+2i$ *is a greatest common divisor of* α *and* β *and we can write*

$$-1+2i=(1+8i)-(-1+i)(2-4i)=\beta+(1-i)(\alpha-(1-i)\beta)=(1-i)\alpha+(1+2i)\beta$$

as linear combination of α *and* β.

Applications

Since $\mathbb{Z}[i]$ *is a unique factorization domain, the notion of prime and irreducible elements coincide.*

Lemma 4.14 *Let* $\alpha\in\mathbb{Z}[i]$ *be an element such that* $N(\alpha)$ *is a prime number. Then* α *is a prime element in* $\mathbb{Z}[i]$.

P*roof: It is enough to show that* α *cannot be factorized into the product of two non-invertible elements. Hence suppose* $\alpha=\beta\gamma$. *Then* $N(\alpha)=N(\beta)N(\gamma)$ *since the norm is multiplicative. But since* $N(\alpha)$ *is a (positive) prime number,* $N(\beta)=1$ *or* $N(\gamma)=1$. *If* $N(\beta)=1$, *then* β *is invertible with inverse* $\overline{\beta}$, *i.e.* $\beta\in\{1,-1,i,-i\}$. *Analogously, if* $N(\gamma)=1$, *then* γ *is invertible. Thus* α *is an irreducible element and hence a prime element as* $\mathbb{Z}[i]$ *is a unique factorization domain.* □

Not all prime elements in $\mathbb{Z}[i]$ *have a norm that is a prime number (in* $\mathbb{Z}$*). For example, the prime number* 3 *is a prime element in* $\mathbb{Z}[i]$, *because if* $3=\alpha\beta$, *with* $\alpha,\beta\in\mathbb{Z}[i]$, *then* $9=N(3)=N(\alpha)N(\beta)$. *However, there does not exists* $\alpha\in\mathbb{Z}[i]$ *with* $N(\alpha)=3$, *because if* $\alpha=a+bi$, *then* $N(\alpha)=a^2+b^2=3$, *which has no solution on* $\mathbb{Z}$[4]

The prime number 2 *is however not a prime element in* $\mathbb{Z}[i]$ *since* $2=(1+i)(1-i)$ *has a factorisation in two non-invertible elements. Note that* $2=1^2+1^2$ *is the sum of two squares and we will prove that a prime number is precisely not a prime element in* $\mathbb{Z}[i]$ *if it is the sum of two squares. We will actually prove that a prime number* p *is not a prime element in* $\mathbb{Z}[i]$ *if and only if it can be written as the sum of two squares.*

We need the following technical result before we continue.

[4] if $a=0$, $b^2=3$ is impossible and if $a=\pm1$, then $b^2=2$ is impossible.

Theorem 4.15 *Let m be a positive integer such that $U(\mathbb{Z}_m)$ is cyclic. Let $a \in \mathbb{Z}$ be an integer that is relatively prime to m. Let $n \in \mathbb{N}^+$ and $d = \gcd(n, \varphi(m))$ Then the following are equivalent:*

(a) $x^n = a(\bmod m)$ *has a solution.*

(b) $a^{\frac{\varphi(m)}{d}} = 1(\bmod m)$.

A solution a of the equation $x^n = a(\bmod m)$ is called an nth power residue modulo m. In case $n = 2$, a is called a quadratic residue modulo m.

P*roof: Let g be a generator of $G = U(\mathbb{Z}_m)$. Then there exists $b \geq 0$ with $a = g^b$. Then*

$$\begin{aligned} \exists c \in \mathbb{N} : (g^c)^n = a(\bmod m) &\Leftrightarrow \exists c \in \mathbb{N} : g^{cn} = g^b(\bmod m) \\ &\Leftrightarrow \exists c \in \mathbb{N} : cn = b(\bmod \varphi(m)) \\ &\Leftrightarrow d = \gcd(n, \varphi(m)) \mid b \\ &\Leftrightarrow d \mid b \text{ and } a^{\frac{\varphi(m)}{d}} = g^{\frac{\varphi(m)b}{d}} = \left(g^{\varphi(m)}\right)^{\frac{b}{d}} = 1(\bmod m) \end{aligned}$$

□

By Theorem 4.15 we have in particular for an odd prime p, that a is a quadratic residue modulo p, i.e. $x^2 = a(\bmod p)$ has a solution[5]*, if and only if $a^{\frac{p-1}{2}} = 1(\bmod p)$. The special case $a = -1$ is important to us:*

Theorem 4.16 *Let p be a positive odd prime number, then the following statements are equivalent:*

(a) p can be written as the sum of two squares;

(b) p is not a prime element in $\mathbb{Z}[i]$;

(c) -1 is a quadratic residue modulo p;

(d) $(-1)^{\frac{p-1}{2}} = 1$;

(e) $p = 1(\bmod 4)$.

In the case $p = a^2 + b^2$, the numbers a and b are uniquely determined up to sign and order.

P*roof: $(a) \Leftrightarrow (b)$ if $p = a^2 + b^2$, then $p = (a + bi)(a - bi)$ and $N(a + bi) = a^2 + b^2 = p$ shows that $a \pm bi$ is a prime element in $\mathbb{Z}[i]$, in particular not invertible, hence p is not prime in $\mathbb{Z}[i]$. On the other hand, if $p = \alpha\beta$, with $\alpha, \beta \in \mathbb{Z}[i]$ not invertible, then $N(\alpha) \neq 1 \neq N(\beta)$ and $p^2 = N(a + bi)N(c + di)$ shows $N(\alpha) = p$. Thus if $\alpha = a + bi$, then $a^2 + b^2 = N(\alpha) = p$.*

$(a) \Rightarrow (c)$: If $p = a^2 + b^2$, then $p \nmid b$ otherwise if $p \mid b$, then also $p \mid a$ and hence $p^2 \mid p$ which is impossible. Thus $p = a^2 + b^2$ is equivalent to $a^2 = -b^2(\bmod p)$ and as b is invertible modulo p, $(ab^{-1})^2 = -1(\bmod p)$, i.e. -1 is a quadratic residue modulo p.

[5] or if you like a has a square root in $\mathbb{Z}_p$

(c) $\Rightarrow$ (b) if $-1 = a^2 (\bmod p)$ for some $a \in \mathbb{Z}$ means that $p \mid (a^2+1)$ in $\mathbb{Z}$. But then also $p \mid (a^2+1) = (a+i)(a-i)$ in $\mathbb{Z}[i]$. If p would be a prime element in $\mathbb{Z}[i]$, then since $\mathbb{Z}[i]$ is a unique factorization domain, $p \mid a+i$ or $p \mid a-i$. If $p \mid a+i$, then there exists $c+di \in \mathbb{Z}[i]$ such that $a+i = p(c+di) = pc+pdi$. Comparing the imaginary part, we see $1 = pd$ which is impossible, since the only invertible integers are 1 and -1. Hence p cannot be a prime element in $\mathbb{Z}[i]$, which shows (b).

(c) $\Leftrightarrow$ (d) $\Leftrightarrow$ (e) follows easily from Theorem 4.15 with $a = -1$ and $n = \varphi(2, p-1) = 2$. It is clear that $(-1)^{\frac{p-1}{2}} = 1$ holds if and only if $(p-1)/2$ is even if and only if $p = 1 (\bmod 4)$.

(Uniqueness): If $p = a^2+b^2$ then $p = (a+bi)(a-bi)$ in $\mathbb{Z}[i]$. For any other numbers $c, d \in \mathbb{Z}$ such that $p = c^2+d^2$ we would also have $p = (c+di)(c-di)$. Since $\mathbb{Z}[i]$ is a unique factorization domain, $a+bi$ must be associated to $c+di$ or $c-di$. Hence there exists $u \in U(\mathbb{Z}[i])$ such that $a+bi = u(c \pm di)$. Substituting the four choices $u \in \{\pm 1, \pm i\}$, we obtain that $a = \pm c$ and $b = \pm d$ or $a = \pm d$ and $b = \pm c$. $\square$

R*emarks: Note that the uniqueness of Theorem 4.16 does not say that for any number n that can be written as $n = a^2+b^2$, the numbers a and b are uniquely determined. For instance $50 = 5^2+5^2 = 1^2+7^2$.*

Fermat had thought that his so-called 5th Fermat number $n = 2^{(2^5)}+1 = 4294967297$ was a prime number. This number is a sum of two squares as

$$n = \left(2^{16}\right)^2 + 1^2 = 65536^2 + 1^2.$$

Euler found a different way to write this number as the sum of two squares, namely

$$n = 62264^2 + 20449^2.$$

Hence Proposition 4.16 shows that n cannot be a prime number, as the squares are different. We have proven that n is not a prime number without having to decompose it.

Condition 4.16(e) is easy to verify. Hence odd primes p that are not primes in $\mathbb{Z}[i]$ are precisely the primes p such that $p = 1+4n$ for some $n \in \mathbb{Z}$, like $5, 13, 17, \ldots$. The set of such primes is infinite:

Corollary 4.17 *There are infinitely many prime numbers that can be written as the sum of two squares and there are infinitely many prime numbers that cannot be written as the sum of two squares.*

P*roof: The proof goes by contradiction. Suppose there are only finitely many positive prime numbers of the form $1+4n$. Say, $\{p_1, \ldots, p_s\}$ are all distinct positive prime numbers $p_i \in \mathbb{N}$ such that $p_i = 1 (\bmod 4)$. Form the number*

$$n = (2p_1 \cdots p_s)^2 + 1.$$

Let p be any prime number that divides n. Then $p = 1 (\bmod 4)$ holds and there must exists $1 \le i \le s$ such that $p = p_i$. However, then $p \mid (2p_1 \cdots p_s)^2$ and as $p \mid n$, we conclude $p \mid 1$ and $p = 1$ a contradiction. Thus the set of prime numbers of the form $1+4n$ must be infinite. By Theorem 4.16 there are infinitely many prime numbers that can be written as the sum of two squares.

Suppose that the set of prime numbers that are of the form $3+4n$ *is finite. Say,* $\{p_1, \ldots, p_s\}$ *are all distinct positive prime numbers* $p_i \in \mathbb{N}$ *such that* $p_i = 3(\bmod 4)$*. Let* $p_1 = 3$ *and form the number*

$$n = 4p_2 \cdots p_s + 3.$$

Since $n = 3(\bmod 4)$*, there must exist a prime number* p *of the form* $3 + 4n$ *dividing* n *(otherwise if all prime divisors* p *of* n *satisfy* $p = 1(\bmod 4)$*, then also* $n = 1(\bmod 4)$ *- a contradiction). By assumption* $p = p_i$ *for some* $1 \leq i \leq s$*. If* $p = 3$*, then* $3 \mid n = 4p_2 \cdots p_s + 3$ *and therefore* $3 \mid 4p_2 \cdots p_s$ *which is a contradiction, since* $3 \nmid 4$ *and* $3 \nmid p_i$ *for* $i \geq 2$*. Hence* $p = p_i$ *for some* $2 \leq i \leq s$*. But then* $p_i \mid n = 4p_2 \cdots p_s + 3$ *implies* $p_i \mid 3$*, which is also a contradiction. Therefore the set of positive prime numbers of the form* $3+4n$ *must be infinite and by Theorem 4.16 there are infinitely many primes that are not the sum of two squares.* □

Theorem 4.16 allows us now to somehow characterise the prime elements of $\mathbb{Z}[i]$*. Recall that two elements* α *and* β *are called associated if* $\alpha = u\beta$ *for* u *an invertible element, i.e. in the case of* $\mathbb{Z}[i]$*,* $u \in \{\pm 1, \pm i\}$*.*

Corollary 4.18 *Any prime element in* $\mathbb{Z}[i]$ *is associated to one of the following primes:*

- $i + i$
- *a positive prime number* $p \in \mathbb{Z}$ *with* $p = 3(\bmod 4)$ *or*
- *a prime element* $\pi \in \mathbb{Z}[i]$ *such that* $p = N(\pi)$ *is an odd prime with* $p = 1(\bmod 4)$

P*roof: Suppose* $\alpha \in \mathbb{Z}[i]$ *is a prime element. Then* $\alpha\overline{\alpha} = N(\alpha) \in \mathbb{N}$*, i.e.* $\alpha \mid n$ *divides a number. Since* $n = p_1^{a_1} \cdots p_s^{a_s}$ *is a product of prime numbers in* $\mathbb{N}$ *and since* α *is a prime element,* $\alpha \mid p$*, for some prime number* p*.*

If $p = 2$*, then* $\alpha \mid 2 = (1 + i)(1 - i) = (-i)(1 + i)^2$*. Hence* α *is associated to the prime element* $1 + i$*.*

If $p = 3(\bmod 4)$*, then by Theorem 4.16* p *is a prime element in* $\mathbb{Z}[i]$ *and* $\alpha \mid p$ *shows that* α *is associated to* p*.*

If $p = 1(\bmod 4)$*, then* $p = a^2 + b^2$*, for some* $a, b \in \mathbb{Z}$*. Hence* $p = (a + bi)(a - bi)$ *with* $N(a \pm bi) = p$*. By Lemma 4.14,* $\pi = a+bi$ *and* $\overline{\pi}$ *are prime elements in* $\mathbb{Z}[i]$*. Since* $\alpha \mid p = \pi\overline{\pi}$*, either* $\alpha \mid \pi$ *or* $\alpha \mid \overline{\pi}$*.* □

Pythagorean triple

As a second application of the unique factorization of Gaussian integers we will classify Pythagorean triples. A triple $(a, b, c) \in \mathbb{Z}$ *of integers is called a Pythagorean triple if*

$$a^2 + b^2 = c^2 \tag{4.6}$$

holds. For instance $(3, 4, 5)$ *is such a triple, since* $9 + 16 = 25$ *or for instance* $(5, 12, 13)$*. A general formula to produce Pythagorean triples is given as follows: For any* $x, y \in \mathbb{Z}^2$ *we define*

$$(a, b, c) = (x^2 - y^2,\ 2xy,\ x^2 + y^2)$$

and verify that

$$a^2 + b^2 = (x^2 - y^2)^2 + (2xy)^2 = x^4 + 2x^2y^2 + y^4 = (x^2 + y^2)^2 = c^2.$$

However, some of these triples will be equal. In order to classify these triples we will restrict Pythagorean triples to primitive ones, because if (a,b,c) is a Pythagorean triple, then also $(\pm a, \pm b, \pm c)$ and (da, db, dc) are Pythagorean triples for all $d \in \mathbb{Z}$. Hence we call a Pythagorean triple (a,b,c) a primitive triple if a,b,c are positive and relatively prime, i.e.

$$a,b,c > 0 \qquad \text{and} \qquad \gcd(a,b,c) = 1.$$

Note that if (a,b,c) is a primitive Pythagorean triple, then $\gcd(a,b) = \gcd(b,c) = \gcd(a,c) = 1$, because whenever a number divides two of the three numbers a,b,c it will divide the third of them due to the equation $a^2 + b^2 = c^2$. In particular, at most one of the three numbers can be even.

Lemma 4.19 *For each $(x,y) \in \mathbb{N}^2$, such x and y are relatively prime, have different parity and $0 < y < x$, the Pythagorean triple $(x^2 - y^2, 2xy, x^2 + y^2)$ is primitive.*

P*roof: Since $0 < y < x$, each component of the constructed triple is positive.*

Since x and y have different parity, $x^2 + y^2$ is odd. Hence if there exists a positive prime number p that divides $x^2 - y^2$, $2xy$ and $x^2 + y^2$, then $p \neq 2$ and

$$p \mid x^2 - y^2 + (x^2 + y^2) = 2x^2$$

and

$$p \mid x^2 - y^2 - (x^2 + y^2) = -2y^2$$

shows that p is a common divisor of x and y, contradicting that x and y are relatively prime. Hence the components are relatively prime and the triple is primitive. □

Using factorzation in $\mathbb{Z}[i]$ we are going to show that any primitive Pythagorean triple (a,b,c) is constructed as in Lemma 4.19.

Let (a,b,c) be a primitive Pythagorean triple. If c would be even, then a and b have to be odd. But then $a^2 = 1 (\bmod 4)$ and $b^2 = 1 (\bmod 4)$ shows $c^2 = a^2 + b^2 = 2 (\bmod 4)$, which is impossible. Hence c must be odd and a and b have different parity. Without loss of generality we may assume that b is even and a is odd. Of course it is clear that if (a,b,c) is a Pythagorean triple, then so is (b,a,c).

Proposition 4.20 *Every primitive Pythagorean triple (a,b,c) with b even is of the form*

$$(a,b,c) = (x^2 - y^2, 2xy, x^2 + y^2)$$

with x and y relatively prime, x and y have different parity and $0 < y < x$.

P*roof: Let (a,b,c) be a primitive Pythagorean triple with b even. By the previous remark, we may assume that a and c are odd. We will write the condition $a^2 + b^2 = c^2$ as an equation in $\mathbb{Z}[i]$, namely as*

$$(a + bi)(a - bi) = c^2.$$

Let $\gamma = \gcd(a + bi, a - bi)$. Then $\gamma \mid 2a$ and $\gamma \mid 2b$ and therefore

$$\gamma \mid 2\gcd(a,b) = 2 = (1 + i)(1 - i).$$

By unique factorization in $\mathbb{Z}[i]$, if γ is not a unit, then $1 \pm i \mid \gamma$. But then $2 \mid N(\gamma) \mid c^2$, which contradicts c to be odd. Hence γ is a unit and $a+bi$ and $a-bi$ are relatively prime in $\mathbb{Z}[i]$.

By unique factorization, we have that $c = p_1^{a_1} \cdots p_s^{a_s}$ as a product of prime elements in $\mathbb{Z}[i]$. Since $a+bi$ and $a-bi$ are relatively prime and their product is $c^2 = p_1^{2a_1} \cdots p_s^{2a_s}$, we must have that there exists a subset $J \subseteq \{1, \dots s\}$ and a unit u such that

$$a + bi = u \prod_{j \in J} p_j^{2a_j} = u \left(\prod_{j \in J} p_j^{a_j} \right)^2 = u(x+yi)^2,$$

for some $x, y \in \mathbb{Z}$. If $u = -1$, then one can write $-1 = i^2$ and hence $a+bi = -(x+yi)^2 = (-y+xi)^2$. Hence we have actually only two options

$$a + bi = (x+yi)^2 \qquad or \qquad a + bi = i(x+yi)^2,$$

for some $x, y \in \mathbb{Z}$. Expanding the square yields

$$a + bi = (x^2 - y^2) + 2xyi \qquad or \qquad a + bi = -2xy + (x^2 - y^2)i.$$

Since we assume b to be even and a to be odd, we can only have the first option, i.e. $a = x^2 - y^2$ and $b = 2xy$. Hence

$$c^2 = (x^2 - y^2)^2 + 4x^2y^2 = (x^2 + y^2)^2$$

and as $c > 0$, we have $c = x^2 + y^2$. Moreover, as $b > 0$, x and y have the same sign and we can choose them to be positive. Furthermore $x^2 - y^2 = a > 0$ shows $x > y$. Any common factor of x and y would be a common factor of a, b, c. Hence x and y are relatively prime.$\square$

5

Law of quadratic residues

We return to our discussion on solving equations in $\mathbb{Z}_m$. *We have already seen when* $U(\mathbb{Z}_m)$ *is cyclic. In case it is, we can try to discuss the existence of* n*th roots of elements in* $U(\mathbb{Z}_m)$ *and Theorem 4.15 tells us that for a positive integer* m *such that* $U(\mathbb{Z}_m)$ *is cyclic,* $a \in \mathbb{Z}$ *an integer that is relatively prime to* m*;* $n \in \mathbb{N}^+$ *and* $d = \gcd(n, \varphi(m))$ *we have that* $x^n = a(\bmod m)$ *has a solution if and only if* $a^{\frac{\varphi(m)}{d}} = 1(\bmod m)$.

For example if we want to solve

$$x^4 = 2(\bmod 7), \tag{5.1}$$

here $m = 7$, $a = 2$ *and* $n = 4$. *In this case* $d = \gcd(n, \varphi(m)) = \gcd(4, 6) = 2$ *and* $a^{\varphi(m)/d} = 2^{6/2} = 8 = 1(\bmod 7)$. *Thus the equation* $x^4 = 2(\bmod 7)$ *has a solution. To obtain a solution, we need to find a generator of* $U(\mathbb{Z}_7)$. *Note that* 2 *is not a generator, since* $2^3 = 1(\bmod 7)$ *shows that* 2 *has order* 3 *in in* $U(\mathbb{Z}_7)$ *and not* $6 = \varphi(7)$. *However,* $g = 3$ *is a generator, since* $3^2 = 9 = 2(\bmod 7) \neq 1(\bmod 7)$ *and* $3^3 = 27 = -1(\bmod 7) \neq 1(\bmod 7)$. *Thus* $3^6 = 1(\bmod 7)$.

Write 2 *as a power of* $g = 3$ *modulo* 7. *We have already seen that* $3^2 = 2(\bmod 7)$. *For* $x = 3^k$ *for some* $k \in \mathbb{Z}$ *we calculate:*

$$\begin{aligned} x^4 = 2(\bmod 7) &\Leftrightarrow 3^{4k} = 3^2(\bmod 7) \\ &\Leftrightarrow 4k = 2(\bmod 6) \\ &\Leftrightarrow 2k = 1(\bmod 3) \\ &\Leftrightarrow k = 2(\bmod 3) \end{aligned}$$

Therefore $x = 3^{2+3l} = 3^2 3^{3^l} = 2 \times (-1)^l(\bmod 7)$, *for* $l \in \mathbb{Z}$ *are solutions for (5.1). Hence we have two different solutions,* $x \in \{2, -2\}$. *We verify that* $(\pm 2)^4 = 16 = 2(\bmod 7)$. *And by our reasoning,* 2 *and* -2 *are the only two solutions of the equation (5.1) modulo* 7.

How can we use the solutions we just found to obtain the solutions for

$$x^4 = 2(\bmod 7^n) \tag{5.2}$$

where $n \geq 2$? The idea is to start with one of the solutions, say 2 and try whether or when $x = 2 + 7y$ is a solution of (5.2), where y has to be determined. For example let $n = 2$ and we would like to obtain the solutions of $x^4 = 2(\bmod 49)$. Then we can calculate:

$$\begin{aligned}
& 2 = (2+7y)^4(\bmod 49) \\
\Leftrightarrow \quad & 2 = 2^4 + \binom{4}{1}2^3 \times (7y) + \binom{4}{2}2^2 \times (7y)^2 + \binom{4}{3}2 \times (7y)^3 + \binom{4}{4}(7y)^4(\bmod 49) \\
\Leftrightarrow \quad & 2 = 2^4 + 4 \times 2^3 \times 7y(\bmod 49) \\
\Leftrightarrow \quad & 2 - 2^4 = 4 \times 2^3 \times 7 \times y(\bmod 49) \qquad \textit{note that } 7 \mid 2 - 2^4 \\
\Leftrightarrow \quad & -2 = 4 \times 8 \times y(\bmod 7) \\
\Leftrightarrow \quad & 5 = 4y(\bmod 7) \\
\Leftrightarrow \quad & 3 = y(\bmod 7)
\end{aligned}$$

Hence $y = 3 + 7l$, for some $l \in \mathbb{Z}$ and

$$x = 2 + 7y = 2 + 7(3 + 7l) = 23 + 49l = 23(\bmod 49)$$

is a solution for (5.2). A similar calculation for the solution -2 and the attempt $x = -2+7y$ yields the solution $y = 4(\bmod 7)$ and therefore, $x = -2 + 7(4 + 7l) = 26(\bmod 49)$.

This procedure works in general and for each solution of (5.1) we obtain a solution of (5.2).

Proposition 5.1 *Let p be an odd prime number and $a \in \mathbb{Z}$, $n \in \mathbb{N}^+$ such that $p \nmid a$, $p \nmid n$. Then $x^n = a(\bmod p)$ has a solution if and only if $x^n = a(\bmod p^m)$ has a solution for any $m \geq 1$. Moreover, the solution sets of each equation have the same number of solutions.*

P*roof: Let $m \geq 1$ and let x_0 be a solution of $x^n = a(\bmod p^m)$ and let $u \in \mathbb{Z}$ be such that $a - x_0^n = up^m$. Since $p \nmid a$, we have $p \nmid x_0^{n-1}$, as otherwise if $p \mid x_0^{n-1}$ then also $p \mid a$. Recall that $p \nmid n$. Then the equation*

$$(nx_0^{n-1})y = u(\bmod p)$$

has a unique solution, say $y \in \mathbb{Z}$. Write $nx_0^n y = u + pl$ for some $l \in \mathbb{Z}$. We calculate:

$$\begin{aligned}
(x_0 + yp^m)^n &= \sum_{i=0}^{n} \binom{n}{i} x_0^{n-i}(yp^m)^i(\bmod p^{m+1}) \\
&= x_0^n + \binom{n}{1}x_0^{n-1}yp^m(\bmod p^{m+1}) \\
&= x_0^n + (u + pl)p^m(\bmod p^{m+1}) \\
&= x_0^n + up^m(\bmod p^{m+1}) = x_0^n + a - x_0^m(\bmod p^{m+1}) = a(\bmod p^{m+1})
\end{aligned}$$

Hence $x_0 + yp^m$ is a solution of the equation $x^n = a(\bmod p^{m+1})$. Since n and $\varphi(p^m)$ are relatively prime, the solution for y is unique and each solution of $x^n = a(\bmod p^m)$ corresponds

to a unique solution of $x^n = a(\bmod p^{m+1})$. *Note that any solution of* $x^n = a(\bmod p^{m+1})$ *is also a solution of* $x^n = a(\bmod p^m)$. *Hence the solution sets of both equations are in correspondence. By induction, we obtain solutions for any* $m \geq 1$, *starting from* $m = 1$. □

Let us specialize to the case of nth power residues module p with $n = 2$. *These residues are called quadratic residues, i.e. an integer* $a \in \mathbb{Z}$ *is called a quadratic residue modulo n if* $\gcd(a, n) = 1$ *and* $x^2 = a(\bmod n)$ *has a solution.*

From Theorem 4.15 we have seen for any odd prime p and $a \in \mathbb{Z}$ *with* $p \nmid a$:

$$x^2 = a(\bmod p) \text{ has a solution} \qquad \text{if and only if} \qquad a^{\frac{p-1}{2}} = 1(\bmod p).$$

What can we say about quadratic residues modulo some number n?

Proposition 5.2 *Let* $n = 2^b p_1^{a_1} \cdots p_s^{a_s}$ *be a positive number and* $p_i \neq p_j$, $i \neq j$ *odd prime numbers with* $b \geq 0$ *and* $a_i \geq 1$. *Let* $a \in \mathbb{Z}$ *with* $\gcd(a, n) = 1$. *then a is a quadratic residue modulo n, i.e.*

$$x^2 = a(\bmod n)$$

has a solution if and only if

(i) $a^{\frac{p_i-1}{=}2} 1(\bmod p_i)$, *for all* $1 \leq i \leq s$

(ii) if $b = 2$ *then* $a = 1(\bmod 4)$ *and if* $b \geq 3$ *then* $a = 1(\bmod 8)$.

P*roof: By the Chinese Remainder Theorem,* $x^2 = a(\bmod m)$ *has a solution if and only if the system*

$$\begin{cases} x^2 &= a(\bmod 2^b) \\ x^2 &= a(\bmod {p_i}^{a_i}) \\ &\vdots \\ x^2 &= a(\bmod {p_s}^{a_s}) \end{cases}$$

has a solution. If $b = 1$, $x^2 = a(\bmod 2)$ *has always a solution. If* $b = 2$, *then* $x^2 = a(\bmod 4)$ *has only a solution of* $a = 1(\bmod 4)$. *If* $b \geq 3$, *then* $x^2 = a(\bmod 2^b)$ *has a solution implies that* $x^2 = a(\bmod 8)$ *has a solution. But the only squares modulo* 8 *are* $0, 1, 4$ *and* $2 \mid 4$. *Hence* $a = 1(\bmod 8)$. □

As a technical tool we define the Legendre symbol of an integer a and a prime number p as

$$\left(\frac{a}{p}\right) := \begin{cases} 1 & a \text{ is a quadratic residue modulo } p \\ 0 & p \text{ divides } a \\ -1 & a \text{ is not a quadratic residue modulo } p \end{cases}$$

Proposition 5.3 *Let p be a positive odd prime number and* $a, b \in \mathbb{Z}$:

1. $a^{\frac{p-1}{2}} = \left(\frac{a}{p}\right) (\bmod p)$.
2. *If* $a = b(\bmod p)$ *then* $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$;
3. $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

P*roof: If* $p \mid a$ *or* $p \mid b$*, then all of the statements (1-3) are true. Hence assume* $p \nmid a, b$.

(1) By Fermat's Theorem **??** *we have that* $a^{p-1} = 1(\bmod p)$*. Hence*

$$\left(a^{\frac{p-1}{2}} + 1\right)\left(a^{\frac{p-1}{2}} - 1\right) = 0(\bmod p).$$

Thus $a^{\frac{p-1}{2}} = \pm 1(\bmod p)$*. By Theorem 4.15,* $\left(\frac{a}{p}\right) = 1$ *if and only if* $a^{\frac{p-1}{2}} = 1(\bmod p)$ *and by the same Theorem,* $\left(\frac{a}{p}\right) = -1$ *if and only if* $a^{\frac{p-1}{2}} \neq 1(\bmod p)$*. Hence* $\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}}(\bmod p)$.

(2) is clear.

(3) Using (1) we calculate:

$$\left(\frac{ab}{p}\right) = (ab)^{\frac{p-1}{2}}(\bmod p) = a^{\frac{p-1}{2}} b^{\frac{p-1}{2}}(\bmod p) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)(\bmod p)$$

Hence $p \mid \left(\frac{ab}{p}\right) - \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$*. But as the latter difference is between* -2 *and* 2 *and* $p \geq 3$ *we must have equality.* □

For $a \in U(\mathbb{Z}_p)$ *we have* $p \nmid a$*, hence* $\left(\frac{a}{p}\right) \in \{-1, 1\} = U(\mathbb{Z})$*. By Proposition 5.3,* $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$ *for all* $a, b \in \mathbb{Z}$*. Therefore this means that the Legendre symbol*

$$\left(\frac{?}{p}\right) : U(\mathbb{Z}_p) \to U(\mathbb{Z}) \qquad \textit{is a group homomorphism.}$$

Therefore we have that:

- *The product of two quadratic residues or two nonquadratic residues modulo* p *is a quadratic residue module* p*.*
- *The product of a quadratic residue and a nonquadratic residue modulo* p *is a nonquadratic residue modulo* p*.*
- *There are as many quadratic residues as nonquadratic residues.*

The last item is true, since $\left(\frac{?}{p}\right) : U(\mathbb{Z}_p) \to U(\mathbb{Z})$ *is a group homomorphism and the kernel* $Q = \mathrm{Ker}(\left(\frac{-}{p}\right))$ *is the set of all quadratic residues modulo* p*. By Lagrange's Theorem,*

$$p - 1 = |U(\mathbb{Z}_p)| = |Q| \times |U(\mathbb{Z})| = 2|Q|.$$

Hence $|Q| = \frac{p-1}{2}$ *and the set of nonquadratic residues module* p *must have the same cardinality.*

Theorem 5.4 *For any positive odd prime number* p *we have* $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.

P*roof: Apply Theorem 4.15 with* $a = -1$ *and* $n = 2$ *and* $m = p$ *as* $d = \gcd(2, p-1) = 2$. □

This means that -1 *is a quadratic residue modulo* p *if and only if* $p = 1(\bmod 4)$.

Lemma 5.5 (Gauss) *Let p be a positive odd prime number and $a \in \mathbb{Z}$ relatively prime with p. For any $1 \leq k \leq \frac{p-1}{2}$ let $m_k \in \{1, \ldots, p-1\}$ such that $m_k = ka(\mod p)$. Then*

$$\left(\frac{a}{p}\right) = (-1)^n (\mod p), \qquad \text{for} \qquad n = \left|\left\{m_k : m_k > \frac{p-1}{2}\right\}\right|.$$

P*roof:* . □

Theorem 5.6 *For any positive odd prime number p we have $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.*

P*roof: Consider the multiples of* 2*:*

$$S = \{2, 4, 6, \ldots, 2 \times \frac{p-1}{2}\}$$

which are all of them less than p. Consider m such that

$$2m \leq \frac{p-1}{2} < 2(m+1).$$

Then $n = |\{2k > \frac{p-1}{2} : k \in \{1, \ldots, \frac{p-1}{2}\}\}| = \frac{p-1}{2} - m$ and by Gauss Lemma, $\left(\frac{a}{p}\right) = (-1)^n$.

Consider $l \in \{1, 3, 5, 7\}$ such that $p = l + 8k$ for some $k \in \mathbb{Z}$. Then

$$2m \leq \frac{p-1}{2} = \frac{l-1+8k}{2} = 4k + \frac{l-1}{2} < 2(m+1)$$

If $l = 1$, then $m = 4k$, and $\left(\frac{a}{p}\right) = (-1)^{2k} = 1$.

If $l = 7$, then $m = 2k+1$ and $\left(\frac{a}{p}\right) = (-1)^{2k} = 1$.

□

This means that 2 *is a quadratic residue modulo p if and only if $p = 1(\mod 8)$ or $p = 7(\mod 8)$.*

6

Cryptography

7

Primality tests

Bibliography

[1] John A. Beachy, *Introductory lectures on rings and modules*, London Mathematical Society Student Texts, vol. 47, Cambridge University Press, Cambridge, 1999. MR1723048

[2] Matej Brešar, *Introduction to noncommutative algebra*, Universitext, Springer, Cham, 2014. MR3308118

[3] Benson Farb and R. Keith Dennis, *Noncommutative algebra*, Graduate Texts in Mathematics, vol. 144, Springer-Verlag, New York, 1993. MR1233388

[4] F.G. Frobenius, *Über lineare Substitutionen und bilineare Formen*, J. Reine Angew. Math. **84** (1878), 1–63, DOI 10.1515/crelle-1878-18788403. MR1581640

[5] K. R. Goodearl and R. B. Warfield Jr., *An introduction to noncommutative Noetherian rings*, 2nd ed., London Mathematical Society Student Texts, vol. 61, Cambridge University Press, Cambridge, 2004. MR2080008

[6] Paul R. Halmos, *Naive set theory*, Undergraduate Texts in Mathematics, Springer-Verlag, New York-Heidelberg, 1974. Reprint of the 1960 edition. MR0453532

[7] I. N. Herstein, *Topics in ring theory*, The University of Chicago Press, Chicago, Ill.-London, 1969. MR0271135

[8] T. Y. Lam, *A first course in noncommutative rings*, 2nd ed., Graduate Texts in Mathematics, vol. 131, Springer-Verlag, New York, 2001. MR1838439

[9] ———, *Lectures on modules and rings*, Graduate Texts in Mathematics, vol. 189, Springer-Verlag, New York, 1999. MR1653294

[10] Donald S. Passman, *A course in ring theory*, The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1991. MR1096302

[11] Joseph Rotman, *Galois theory*, 2nd ed., Universitext, Springer-Verlag, New York, 1998. MR1645586

[12] Louis H. Rowen, *Ring theory*, Student edition, Academic Press, Inc., Boston, MA, 1991. MR1095047

[13] H. Weber, *Leopold Kronecker*, Math. Ann. **43** (1893), no. 1, 1–25, DOI 10.1007/BF01446613 (German). MR1510799