Elementary Number Theory: Some Lecture Notes

Karl-Heinz Fieseler

Uppsala 2013

Contents

1 Survey

This are supplementary lecture notes, intended to give details where we do not follow in our argumentation the textbook NZM or the LÅL-notes.

We assume the reader to have some basic knowledge of commutative rings, as for example presented in sections 3 - 7: They do not enter into the course, whose main subjects are discussed in the chapters 8, 10 - 12, 14 - 19.

The first part of the course is devoted to the solution of a polynomial equation

$$f(x) = 0,$$

in a residue class ring $\mathbb{Z}_n$, where $f \in \mathbb{Z}[X]$ is a polynomial in one variable.

By the chinese remainder theorem one can reduce that problem to prime power moduli $n = p^k$ and there we may apply in certain favourable cases a lifting procedure which creates a solution $a_k \in \mathbb{Z}_{p^k}$ for every p^k from a given solution in $\mathbb{Z}_p$. But such a chain of solutions in general does not correspond to a solution in $\mathbb{Z}$; nevertheless it turns out to be helpful to consider such chains as a new sort of numbers, called p-adic integers, see the optional section 9.

On the other hand, the polynomial

$$f = X^\ell - a$$

as the simplest polynomial of degree ℓ should be understood, i.e. one wants to understand whether or not an element $\overline{a} \in \mathbb{Z}_n$ is an ℓ-th power: We are led to the notion of a primitive root for a modulus n, see section 10.

If $\ell = 2$ there is a very convenient way to decide, whether some $a \in \mathbb{Z}_p$ is a square or not: This is the famous law of quadratic reciprocity, see section 11.

There is a short section (12) dealing with arithmetic functions. For those familiar with complex analysis, we mention the connection with Dirichlet series (not part of the course) and discuss basic features of Riemann's ζ-function in the optional section 13.

The second part of the course deals with diophantine equations

$$f(x_1, ..., x_m) = 0,\ f \in \mathbb{Z}[X_1, ..., X_m],\ m > 1$$

and one is interested in solutions $(x_1, ..., x_m) \in \mathbb{Z}^m$. We discuss in detail

1. linear ones: $a_1x_1 + ... + a_mx_m - b = 0$; we give a criterion for solvability and a recipe, how to obtain all solutions,

2. $x^2 + y^2 = n \in \mathbb{N}$; we give a criterion for solvability and count the solutions,

3. $x^2 + y^2 = z^2$, which one may regard as a special case of the previous equation when fixing $n := z^2$. All its solutions $(x, y, z) \in \mathbb{N}^3$, called pythagorean triples, are given explicitly,

4. $x^4 + y^4 = z^4$ is shown to be unsolvable following Fermat,

5. $x^2 + y^2 + z^2 + t^2 = n$ is shown to be solvable for any $n \in \mathbb{N}$,

6. Pell's equation $x^2 - dy^2 = \pm 1$ with $d \in \mathbb{N}_{>1}$ not a square. We see that all its (infinitely many) solutions can be generated from one basic solution. In order to assure the existence of solutions $(x, y) \neq (\pm 1, 0)$ we need continued fractions, section 19.

2 Distribution of primes: A short fairy tale

See also LÅL, section 2, Th.2.5 - Th.2.10. By a result of Euclid we know that there are infinitely many prime numbers:

$$|P| = \infty$$

holds for the set $P \subset \mathbb{N}$ of all prime numbers. Indeed, given $p_1, ..., p_n \in P$ we find further prime numbers, namely the prime divisors of $p_1 \cdot ... \cdot p_n + 1$.

Question: How dense are the prime numbers in $\mathbb{N}$?

They are not too sparse, indeed:

$$\sum_{p \in P} \frac{1}{p} = \infty,$$

while

$$\sum_{n \in \mathbb{N}} \frac{1}{n^2} < \infty,$$

i.e. prime numbers are more dense in $\mathbb{N}$ than squares! We shall prove the following estimate:

Proposition 2.1. *For $x \geq 3$ we have*

$$\sum_{p \leq x} \frac{1}{p} \geq \ln(\ln(x)) - 1,$$

Proof. Let $P_{\leq x} := P \cap \mathbb{R}_{\leq x} = \{p_1 = 2, p_2 = 3, ..., p_n\}$ and

$$F(n) := p_1^{\mathbb{N}} \cdot ... \cdot p_n^{\mathbb{N}}$$

be the set of all natural numbers, which are a product of the $p_i, i = 1, ..., n$. Then we find

$$\sum_{k \in F(n)} \frac{1}{k} = \sum_{\nu=0}^{\infty} \left(\frac{1}{p_1}\right)^{\nu} \cdot ... \cdot \sum_{\nu=0}^{\infty} \left(\frac{1}{p_n}\right)^{\nu}$$

$$= \frac{1}{1 - \frac{1}{p_1}} \cdot ... \cdot \frac{1}{1 - \frac{1}{p_n}}.$$

Now we have $[1, x] \cap \mathbb{N} \subset F(n)$, whence

$$\sum_{k \in F(n)} \frac{1}{k} \geq \sum_{k \leq x} \frac{1}{k} \geq \int_1^{[x]+1} \frac{dt}{t} = \ln([x] + 1) \geq \ln(x).$$

So

$$\prod_{p \leq x} \frac{1}{1 - \frac{1}{p}} \geq \ln(x)$$

and thus, applying the logarithm, we obtain

$$\sum_{p \leq x} -\ln\left(1 - \frac{1}{p}\right) > \ln(\ln(x)).$$

Now with a Taylor-MacLaurin expansion we arrive at

$$-\ln(1-t) \overset{|t|<1}{=} t + \frac{t^2}{2} + \frac{t^3}{3} + ... \overset{0 \leq t < 1}{\leq} t + \frac{t^2}{2}\left(1 + t + t^2 + ...\right) = t + \frac{t^2}{2} \cdot \frac{1}{1-t}.$$

For $0 \leq t \leq \frac{1}{2}$ this leads to

$$-\ln(1-t) \leq t + t^2.$$

Hence, with $t = \frac{1}{p}, p \leq x$, we find

$$\sum_{p \leq x} \frac{1}{p} + \sum_{p \leq x} \frac{1}{p^2} > \ln(\ln(x)).$$

Finally

$$\sum_{p \leq x} \frac{1}{p^2} \leq \sum_{k=2}^{\infty} \frac{1}{k^2} \leq \sum_{k=2}^{\infty} \frac{1}{k(k-1)} = \sum_{k=2}^{\infty} \left(\frac{1}{k-1} - \frac{1}{k} \right) = 1.$$

□

In order to get some more explicit information one introduces the prime number function

$$\pi : \mathbb{R}_{>0} \longrightarrow \mathbb{N}, \pi(x) := |P_{\leq x}|,$$

counting the number of primes below the argument $x \in \mathbb{R}_{>0}$ and investigates the prime number density function

$$\frac{\pi(x)}{x}$$

giving for $x \in \mathbb{N}$ the percentage of prime numbers in $[1, x] \cap \mathbb{N}$. One expects that prime numbers become more and more sparse.

Definition 2.2. For functions $f, g : \mathbb{R}_{>2} \longrightarrow \mathbb{R}_{>0}$ we write

$$f(x) \sim g(x) :\Longleftrightarrow \lim_{x \to \infty} \frac{f(x)}{g(x)} = 1.$$

Theorem 2.3 (Prime Number Theorem). *We have*

$$\frac{\pi(x)}{x} \sim \frac{1}{\ln(x)}.$$

With other words: Given any $q < 1$ for all sufficiently big $x \in \mathbb{R}_{>2}$ we have

$$\frac{q}{\ln(x)} \leq \frac{\pi(x)}{x} \leq \frac{1}{q \ln(x)}.$$

It has been proven first by de la Vallée-Poussin and Hadamard with complex-analytic methods (1896), see also the (optional) section about Riemann's ζ-function, and 50 years later with elementary tools by Erdös and Selberg. A more detailed investigation of the prime number function π involves Riemann's ζ-function - a factorization of it leads to a (quite sophisticated) formula for $\pi(x)$, where the factorization depends on the zeros of ζ. Indeed, the above theorem is equivalent to a statement about the zeros of ζ.

Here we can show only a partial result, a lower bound for the occurrence of prime numbers:

Proposition 2.4. *Given any $q < 1$ there are arbitrarily big $x \in \mathbb{R}_{>0}$, such that*

$$\frac{q}{\ln(x)} \leq \frac{\pi(x)}{x}.$$

First we need a formula which relates the prime density function with the sum we have just estimated:

Lemma 2.5.

$$\sum_{p \leq x} \frac{1}{p} = \frac{\pi(x)}{x} + \int_2^x \frac{\pi(t)}{t^2} dt.$$

Proof.

$$\begin{aligned}
\int_2^x \frac{\pi(t)}{t^2} dt &= \sum_{k=1}^{n-1} \int_{p_k}^{p_{k+1}} \frac{\pi(t)}{t^2} dt + \int_{p_n}^x \frac{\pi(t)}{t^2} dt \\
&= \sum_{k=1}^{n-1} \int_{p_k}^{p_{k+1}} \frac{k}{t^2} dt + \int_{p_n}^x \frac{n}{t^2} dt \\
&= \sum_{k=1}^{n-1} k \left(\frac{1}{p_k} - \frac{1}{p_{k+1}} \right) + n \left(\frac{1}{p_k} - \frac{1}{x} \right) \\
&= \sum_{k=1}^{n-1} \frac{k}{p_k} - \sum_{k=2}^{n} \frac{k-1}{p_k} + \frac{n}{p_n} - \frac{n}{x} \\
&= \sum_{k=1}^{n} \frac{1}{p_k} - \frac{\pi(x)}{x}.
\end{aligned}$$

□

Proof of Prop.2.4. Let $q < 1$. Assume the contrary, i.e. there is some $M > 0$, s.th.

$$\frac{\pi(x)}{x} \leq \frac{q}{\ln(x)}$$

holds for all $x \geq M$. We derive an estimate

$$\sum_{p \leq x} \frac{1}{p} \leq q \ln(\ln(x)) + D$$

for $x \geq M$, which contradicts the estimate

$$\ln(\ln(x)) - 1 \leq \sum_{p \leq x} \frac{1}{p},$$

since $\lim_{x \to \infty} \ln(\ln(x)) = \infty$. Indeed, the assumption implies

$$\sum_{p \leq x} \frac{1}{p} = \int_2^x \frac{\pi(t)}{t^2} dt + \frac{\pi(x)}{x} \leq \int_2^x \frac{\pi(t)}{t^2} dt + 1 \leq \int_2^M \frac{\pi(t)}{t^2} dt + \int_M^x \frac{q}{t \ln(t)} dt + 1$$

$$= C + 1 + q(\ln(\ln(x)) - \ln(\ln(M))).$$

□

Here is a further question related to the prime number theorem:

Question: How does the gaps $g_n := p_n - p_{n-1}$ between two successive prime numbers behave for $n \to \infty$? The prime number theorem tells us that the sequence of their arithmetic means satisfies

$$s_n := \frac{1}{n-1} \sum_{i=2}^{n} g_i \ \sim \ \ln(p_n);$$

indeed:

$$s_n = \frac{1}{n-1}(p_n - 2) \sim \frac{p_n}{n} = \frac{p_n}{\pi(p_n)} \sim \ln(p_n).$$

In particular the sequence $(g_n)_{n \geq 2}$ is unbounded, but that can be seen as well with an elementary argument: In the interval $[m! + 2, ..., m! + m] \cap \mathbb{N}$ there are obviously no prime numbers, so if $p_{n-1} \leq m! + 2 \leq p_n$, we have $g_n \geq m$. But from that we can not conclude that $\lim_{n \to \infty} g_n = \infty$. The following result is from 2013:

Theorem 2.6 (Zhang Yitang)**.** *There are arbitrarily big $n \in \mathbb{N}$ with*

$$p_n - p_{n-1} \leq 7 \cdot 10^7.$$

Last year the above estimate has been improved: The estimate

$$p_n - p_{n-1} \leq 246$$

holds for infinitely many $n \in \mathbb{N}$. Finally one would like to show that even

$$p_n - p_{n-1} = 2$$

holds for infinitely many $n \in \mathbb{N}$, i.e. that there are infinitely many prime twins $(p, p+2)$.

3 Rings

Number theory is concerned with the set $\mathbb{Z}$ of integers. It admits two binary operations

$$\mathbb{Z} \times \mathbb{Z} \longrightarrow \mathbb{Z},$$

addition and multiplication. But often it is quite useful also to admit other auxiliary sets, where these operations are defined, either enlarging $\mathbb{Z}$, e.g. replacing it with the rationals $\mathbb{Q} \supset \mathbb{Z}$, or "shrinking" it, e.g. replacing $\mathbb{Z}$ with the set $\mathbb{Z}/\sim$ of equivalence classes, where $\sim$ is an equivalence relation on $\mathbb{Z}$ compatible with the addition and multiplication of integers.

Hence, it turns out to be useful to introduce a general notion, that of a ring:

Definition 3.1. *A commutative ring is a triple (R, α, μ), consisting of a set R and two binary operations, i.e. maps,*

$$\alpha : R \times R \longrightarrow R, (x, y) \mapsto x + y := \alpha(x, y) \quad ,$$

the "addition", and

$$\mu : R \times R \longrightarrow R, (x, y) \mapsto xy := \mu(x, y) \quad ,$$

the "multiplication", such that

1. *For all $x, y \in R$ we have*

$$x + y = y + x, xy = yx.$$

2. *For all $x, y, z \in R$ we have*

$$(x + y) + z = x + (y + z), (xy)z = x(yz).$$

3. *For all $x, y, z \in R$ we have*

$$x(y + z) = xy + xz.$$

4. *There are elements $0, 1 \in R$, such that*

$$x + 0 = x, 1 \cdot x = x$$

holds for all $x \in R$.

5. *For all $x \in R$ there is an element $y \in R$ with*

$$x + y = 0.$$

Example 3.2. 1. The set $\mathbb{Z}$ of all integers, $\mathbb{Q}$ of rational, $\mathbb{R}$ of real and $\mathbb{C}$ of complex numbers are rings with the standard addition and multiplication of complex numbers.

2. There are further examples of the above type, i.e. subsets $R \subset \mathbb{C}$, which are additively and multiplicatively closed, contain $0, 1 \in \mathbb{C}$ and satisfy $R = -R$. We mention $R = \mathbb{Z} + \mathbb{Z}i$, the set of gaußian integers, i.e. complex numbers with integer real and imaginary part. They form a lattice in the complex plane. Another example is $R := \mathbb{Z} + \mathbb{Z}\varepsilon$ with the third root of unity $\varepsilon := \frac{1}{2}(-1 + i\sqrt{3})$. For example when dealing with Fermat's equation for $n = 3$ it can be useful to replace $\mathbb{Z}$ with R since then
$$x^3 = z^3 - y^3 = (z - y)(z - \varepsilon y)(z - \varepsilon^2 y).$$
On the other hand the ring R is sufficiently close to $\mathbb{Z}$ in order to obtain interesting information even for $\mathbb{Z}$.

Here are some features all rings share:

Remark 3.3. 1. The elements $0, 1 \in R$ are uniquely determined.

2. Assume $x + y = 0 = x + \tilde{y}$. Then

$$\tilde{y} = \tilde{y} + 0 = \tilde{y} + (x + y) = (\tilde{y} + x) + y = 0 + y = y.$$

Thus we may define

$$-x := y.$$

As an immediate consequence one obtains

$$-(-x) = x.$$

3. We define subtraction as follows:

$$x - y := x + (-y).$$

4. Next:

$$0 \cdot x = 0 = x \cdot 0$$

holds, since $x = 1 \cdot x = (1 + 0)x = x + 0 \cdot x$. Furthermore

$$(-1)x = x(-1) = -x$$

follows from $0 = (1 + (-1))x = x + (-1)x$.

5. In general there is no cancellation rule for the multiplication, since there may be nontrivial "zero divisors", i.e. elements $a \in R \setminus \{0\}$, such that

$$ab = 0$$

for some $b \neq 0$, and it can happen that

$$1 + ... + 1 = 0 \ .$$

4 Divisibility revisited

Definition 4.1. A subset $\mathfrak{a} \subset \mathbb{Z}$ is called an *ideal* if

1. $0 \in \mathfrak{a}$,

2. $x, y \in \mathfrak{a} \Longrightarrow x + y \in \mathfrak{a}$, i.e. an ideal is closed w.r.t. addition,

3. $x \in \mathfrak{a} \Longrightarrow -x \in \mathfrak{a}$, i.e. an ideal lies symmetric to the origin.

Remark 4.2. For an ideal $\mathfrak{a} \subset \mathbb{Z}$ we have

$$a \in \mathfrak{a} \Longrightarrow ka \in \mathfrak{a},\ \forall\, k \in \mathbb{Z}.$$

Example 4.3.

1. $\mathfrak{a} = \mathbb{Z}$ is an ideal, the "unit ideal".
2. $\mathfrak{a} = \{0\}$ is an ideal, the "trivial" or "zero ideal".
3. Let $n \in \mathbb{Z}$. Then
$$\mathbb{Z}n := \{kn; k \in \mathbb{Z}\}$$
is an ideal.
4. Let $a, b \in \mathbb{Z}$. Then
$$(a, b) := \mathbb{Z}a + \mathbb{Z}b = \{ka + \ell b; k, \ell \in \mathbb{Z}\}$$
is an ideal, it is called the ideal generated by a and b. Note that
$$(n, 0) = \mathbb{Z}n.$$

Theorem 4.4. *Any ideal $\mathfrak{a} \subset \mathbb{Z}$ is a principal ideal:*

$$\mathfrak{a} = \mathbb{Z}n$$

with a uniquely determined natural number $n \in \mathbb{N}$.

Proof. If $\mathfrak{a} = \{0\}$, take $n = 0$. Otherwise we have

$$\mathfrak{a}_{>0} := \{a \in \mathfrak{a}; a > 0\} \neq \emptyset$$

because of $\mathfrak{a} = -\mathfrak{a}$ and choose

$$n := \min \mathfrak{a}_{>0}.$$

Since $n \in \mathfrak{a}$ and $\mathfrak{a}$ is closed under multiplication with arbitrary integers, we get $\mathbb{Z}n \subset \mathfrak{a}$. Take now any $a \in \mathfrak{a}$ and write

$$a = qn + r,\ 0 \leq r < n.$$

Then there are two possibilities

1. $r = 0$ and thus $a \in \mathbb{Z}n$, or
2. $r = a - qn \in \mathfrak{a}_{>0}$. But then we find $r \geq \min \mathfrak{a}_{>0} = n$, a contradiction.

□

Finally we want to find explicitly the generator

$$n := \min (a, b)_{>0}.$$

Obviously there is a problem, since one does not know all $k, \ell \in \mathbb{Z}$ with $ka + \ell b > 0$. Instead we can use a reduction procedure for the generators a och b. It is based on the following remark:

Remark 4.5. If $a = qn + r$ or $a = qn - r$, then

$$(a, b) = (b, r).$$

The **Euclidean algorithm** is now an iteration of that reduction step:

1. We may asume $a \geq b \geq 0$.
2. If $a = b$ or $b = 0$, then $(a, b) = \mathbb{Z}a$.
3. If $a > b > 0$, we write $a = qb + r$ or $a = qn - r$ and the above remark gives
$$(a, b) = (b, r).$$
4. Now we can repeat that step with (b, r) instead of (a, b). If we choose $r \in \mathbb{Z}$ according to the division algorithm, we have $a > b > r \geq 0$, and arrive after finitely many steps at the situation, where $(a, b) = (n, 0) = \mathbb{Z}n$. But when admitting negative remainders, we may even require $\frac{b}{2} \geq r \geq 0$, thus making the algorithm a little bit faster.

Finally let us justify the title of this section:

Proposition 4.6. *If*

$$(a, b) = \mathbb{Z}n,$$

where $ab \neq 0$ *and* $n \in \mathbb{N}$*, we have*

$$n = \gcd(a, b).$$

Proof. We have $q := \gcd(a, b) \geq n$, since n divides both a and b. On the other hand q divides a and b, hence also $n = ra + sb$. But

$$q|n \wedge q \geq n \Longrightarrow q = n.$$

□

5 Residue class rings

We fix a natural number $n \in \mathbb{N}_{>1}$, the "modulus".

Definition 5.1. Two integers $a, b \in \mathbb{Z}$ are said to be congruent modulo n, written as:

$$a \equiv b \mod (n)$$

or

$$a \overset{n}{\equiv} b,$$

iff n divides the difference $b - a$.

Remark 5.2.

1. $.. \overset{n}{\equiv} ..$ is an equivalence relation.

2. The equivalence class

$$R_n(a) := \{b \in \mathbb{Z}; b \overset{n}{\equiv} a\}$$

of an integer $a \in \mathbb{Z}$ is

$$R_n(a) = a + \mathbb{Z}n := \{a + kn; k \in \mathbb{Z}\}.$$

3. Remember

$$R_n(a) = R_n(b) \Longleftrightarrow a \overset{n}{\equiv} b.$$

4. For $0 \leq r < n$ the equivalence class $R_n(r)$ consists of all integers giving remainder r after division with n, i.e. $a = qn + r$, this explains why one often speaks of residue classes mod n. We thus obtain a partition

$$\mathbb{Z} = \bigcup_{i=0}^{n-1} R_n(i).$$

5. The set of all residue classes mod n is denoted:

$$\mathbb{Z}_n := \{R_n(i); i = 0, ..., n-1\}.$$

Making $\mathbb{Z}_n$ a ring: We want to define addition and multiplication

$$\mathbb{Z}_n \times \mathbb{Z}_n \longrightarrow \mathbb{Z}_n$$

for residue classes mod n. We start with the set theoretic sum of subsets $A, B \subset \mathbb{Z}$, namely

$$A + B := \{a + b; a \in A, b \in B\}$$

and

$$AB := \{ab; a \in A, b \in B\}.$$

Proposition 5.3. *1.* $R_n(a) + R_n(b) = R_n(a + b)$,

2. $R_n(a) \cdot R_n(b) \subset R_n(ab)$,

3. $R_n(ab) = R_n(a) \cdot R_n(b) + \mathbb{Z}n$.

Proof. We consider the sum and the product of arbitrary integers $a + kn \in R_n(a)$ and $b + \ell n \in R_n(b)$ and see that

$$(a + kn) + (b + \ell n) = (a + b) + (k + \ell)n \in R_n(a + b)$$

and

$$(a + kn)(b + \ell n) = ab + (kb + \ell a + k\ell n)n \in R_n(ab).$$

Since all the products lie in the residue class $R_n(ab)$, we can simply add $\mathbb{Z}n$ and obtain the entire residue class $R_n(ab)$. □

Remark 5.4. The set theoretic product of residue classes is not necessarily again a residue class, e.g.

$$R_n(0) \cdot R_n(0) = \mathbb{Z}n \cdot \mathbb{Z}n = \mathbb{Z}n^2 \subsetneq \mathbb{Z}n = R_n(0).$$

So the addition is set theoretic

$$\alpha : \mathbb{Z}_n \times \mathbb{Z}_n \longrightarrow \mathbb{Z}_n, (R_n(a), R_n(b)) \mapsto R_n(a+b) = R_n(a) + R_n(b),$$

while the set theoretic product has to be completed to a residue class

$$\mu : \mathbb{Z}_n \times \mathbb{Z}_n \longrightarrow \mathbb{Z}_n, (R_n(a), R_n(b)) \mapsto R_n(ab) = R_n(a) \cdot R_n(b) + \mathbb{Z}n.$$

A more convenient notation: If the modulus n is understood, we simply write

$$\overline{a} := R_n(a) = a + \mathbb{Z}n.$$

We then have

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, ..., \overline{n-1}\}$$

and addition and multiplication is as easy as possible:

$$\overline{a} + \overline{b} = \overline{a+b}$$

and

$$\overline{a} \cdot \overline{b} = \overline{ab}.$$

Note: From now on we use the standard notation for the product of residue classes

$$\overline{a} \cdot \overline{b} = R_n(a) \cdot R_n(b) + \mathbb{Z}n = \overline{ab},$$

so it is **not** any longer the set theoretic product!

6 Further properties of rings

Definition 6.1. 1. An element $a \in R$ in a commutative ring R is called a *non-zero divisor* iff $ab = 0 \Longrightarrow b = 0$.

2. A non-trivial ring (i.e. s.th. $1 \neq 0$), where all elements $a \neq 0$ are non-zero-divisors is called an *integral domain*.

Example 6.2. 1. The rings $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are integral domains.

2. Non-zero-divisors $a \in R$ can be cancelled in an equation:

$$ax = ay \Longrightarrow x = y.$$

3. A residue class $\overline{a} \in \mathbb{Z}_n$ is a non-zero-divisor iff $\gcd(a, n) = 1$. We have

$$\overline{a} \cdot \overline{b} = \overline{ab} = \overline{0} \Longleftrightarrow \exists k \in \mathbb{Z} : \ ab = kn.$$

If $\gcd(a, n) = 1$ that implies $n|b$, i.e. $\overline{b} = \overline{0}$. On the other hand, if $\gcd(a, n) = d > 1$, we have $n = n_0 d$ and $n|an_0$, i.e.

$$\overline{a} \cdot \overline{n}_0 = \overline{0},$$

though $\overline{n}_0 \neq \overline{0}$.

4. The residue class ring $\mathbb{Z}_n$ is an integral domain iff $n = p$ is a prime number, since a natural number is prime iff any number $k, 1 \leq k < n$ and n are relatively prime.

Definition 6.3. 1. An element $a \in R$ in a nontrivial commutative ring R is called *invertible* or a *unit* iff there is an element $b \in R$ with $ab = 1$.

2. The set

$$R^* := \{a \in R; \exists b \in R : \ ab = 1\}$$

of all invertible elements is called the *group of units* of the ring R.

Remark 6.4. 1. The number b with $ab = 1$ is uniquely determined if it exists, we write $a^{-1} := b$.

2. The group of units is closed w.r.t. multiplication and inversion. Indeed

$$(ab)^{-1} = b^{-1}a^{-1}, (a^{-1})^{-1} = a.$$

3. Units are non-zero-divisors: If $a \in R^*$ and $ab = 0$, we find $0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$.

4. In a finite (commutative) ring non-zero-divisors are already units, since for a non-zero-divisor $a \in R$ the map $\mu_a : R \longrightarrow R, x \mapsto ax$, is injective, and an injective map from a finite set to itself is even bijective. In particular there is some $b \in R$ with $\mu_a(b) = 1$.

5. $\mathbb{Z}^* = \{\pm 1\}$.

6. $\mathbb{Z}_n^* = \{\overline{a} \in \mathbb{Z}_n;\ \gcd(a, n) = 1\}$.

Example 6.5. In order to invert a unit $\overline{a} \in \mathbb{Z}_n^*$ we have to find integers $r, s \in \mathbb{Z}$, such that $ra + sn = 1$. Then we obtain

$$\overline{a}^{-1} = \overline{r}.$$

Let us compute $\overline{70}^{-1} \in \mathbb{Z}_{101}$. The Euclidean algoritm gives

$$101 = 1 \cdot 70 + 31,\ 70 = 2 \cdot 31 + 8,\ 31 = 4 \cdot 8 - 1$$

and thus

$$1 = 4 \cdot 8 - 31 = 4 \cdot (70 - 2 \cdot 31) - 31 = 4 \cdot 70 - 9 \cdot 31$$
$$= 4 \cdot 70 - 9 \cdot (101 - 70) = 13 \cdot 70 - 9 \cdot 101.$$

Hence $\overline{70}^{-1} = \overline{13} \in \mathbb{Z}_{101}$.

Definition 6.6. An integral domain R is called a *field* (*kropp* på svenska), if all nonzero elements are units, i.e.

$$R^* = R \setminus \{0\}.$$

Example 6.7. The rings $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ and $\mathbb{Z}_p$ with a prime number p are fields.

There are further interesting elements in a commutative ring:

Definition 6.8. Let R be a commutative ring.

1. An element $a \in R$ is called *nilpotent* if we have $a^n = 0$ for some $n \in \mathbb{N}_{>0}$.

2. An element $e \in R$ is called *idempotent* if we have $e^2 = e$.

Remark 6.9.

1. The elements $0, 1 \in R$ are idempotent. If R is an integral domain, there are no further idempotents: The equation $e^2 = e$ may also be rewritten as $0 = e^2 - e = e(e - 1)$.

2. An element $\overline{a} \in \mathbb{Z}_n$ is nilpotent iff a and n have the same prime divisors.

3. In $\mathbb{Z}_{p^n}$ an element is either a unit or nilpotent.

4. In $\mathbb{Z}_{p^n}$ there are no idempotents except $0, 1$. An idempotent element $e \neq 0, 1$ would be a zero divisor, hence nilpotent, and thus necessarily 0. Contradiction!

5. If n is not a prime power, then there are always idempotent elements, as we shall see later on in the chinese remainder theorem.

7 Groups of units

Definition 7.1. The function $\varphi : \mathbb{N}_{>0} \longrightarrow \mathbb{N}$ defined by

$$\varphi(n) := \begin{cases} 1 & , \quad \text{if } n = 1 \\ |\mathbb{Z}_n^*| & , \quad \text{if } n \geq 2 \end{cases}$$

is called *Eulers φ-function.*

Lemma 7.2. *For a prime power $n = p^k, k \geq 1$, we have:*

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

Proof. The complement of $\mathbb{Z}_{p^k}^*$ consists of the zero divisors

$$\overline{0}, \overline{p}, \overline{2p}, ..., \overline{(p^{k-1}-1)p}.$$

Thus

$$|\mathbb{Z}_{p^k}^*| = |\mathbb{Z}_{p^k}| - \text{number of zero divisors} = p^k - p^{k-1} = p^{k-1}(p-1).$$

□

In order to compute $\varphi(n)$ for arbitrary $n \in \mathbb{N}$ we regard the prime factorization

$$n = p_1^{k_1} \cdot .... \cdot p_r^{k_r}$$

and compare the residue class ring $\mathbb{Z}_n$ with the residue class rings $\mathbb{Z}_{p_i^{k_i}}, i = 1, ..., r$. The set $\mathbb{Z}_n$ has as many elements as the cartesian product $\mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_r^{k_r}}$, in particular there is a bijection

$$\psi : \mathbb{Z}_n \longrightarrow \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_r^{k_r}}.$$

We ask, if we can choose $\psi = (\psi_1, ..., \psi_r)$ in such a way, that we may determine $\psi(\mathbb{Z}_n^*)$. First of all the target can be made a ring:

Definition 7.3. Let $R_1, ..., R_s$ be rings. The direct product $R_1 \times .... \times R_s$ is the cartesian product of the sets $R_1, ..., R_s$ endowed with the componentwise addition and multplication..

Remark 7.4.

$$(R_1 \times .... \times R_s)^* = R_1^* \times .... \times R_s^*.$$

Now if we could choose ψ as a *ring homomorphism*, i.e. compatible with the ring operations on both sides:

$$\psi(x+y) = \psi(x) + \psi(y), \psi(xy) = \psi(x)\psi(y), \psi(1) = 1,$$

we would find that ψ induces a bijection

$$\mathbb{Z}_n^* \xrightarrow{\psi} \mathbb{Z}_{p_1^{k_1}}^* \times ... \times \mathbb{Z}_{p_r^{k_r}}^*.$$

Definition 7.5. Let R, S be rings. We write $R \cong S$ and say "*R is isomorphic to S*", if there is a bijective ring homomorphism (= a ring isomorphism)

$$\psi : R \longrightarrow S.$$

So what could we take as the component maps

$$\psi_i : \mathbb{Z}_n \longrightarrow \mathbb{Z}_{p_i^{k_i}}, i = 1, ..., r?$$

Remark 7.6. Let m be a divisor of n. Then

$$\mathbb{Z}_n \longrightarrow \mathbb{Z}_m, a + \mathbb{Z}n \mapsto (a + \mathbb{Z}n) + \mathbb{Z}m = a + \mathbb{Z}m$$

defines a ring homomorphism.

Finally we arrive at

Theorem 7.7 (Chinese Remainder Theorem). *Let*

$$n = p_1^{k_1} \cdot .... \cdot p_r^{k_r}.$$

Then

$$\psi : \mathbb{Z}_n \longrightarrow \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_r^{k_r}}, \overline{a} \mapsto (a + \mathbb{Z}p_1^{k_1}, ..., a + \mathbb{Z}p_r^{k_r})$$

is a ring isomorphism.

Proof. Since the domain of definition and the target of ψ have the same number of elements, it suffices to show that ψ is injective. Now a ring homomorphism is injective iff $\psi(z) = 0$ implies $z = 0$, since

$$\psi(x) = \psi(y) \Longleftrightarrow \psi(x - y) = 0.$$

Take $z = \overline{a} \in \mathbb{Z}_n$. Now $\psi(\overline{a}) = 0$ means $(p_i)^{k_i} | a$ for $i = 1, ..., r$ resp. $n|a$ resp. $\overline{a} = 0$. □

Corollary 7.8. *Let* $n = p_1^{k_1} \cdot .... \cdot p_r^{k_r}$. *Then*

$$\varphi(n) = \varphi(p_1^{k_1}) \cdot ... \cdot \varphi(p_r^{k_r}).$$

Let us come back to commutative rings! Given a unit $z \in R^*$ we want to study the sequence

$$1, z, z^2, z^3, ...$$

of its powers. Then either

1. $z^k \neq z^\ell$ for $k \neq \ell$, or
2. there are ℓ, k with $\ell > k$, such that $z^\ell = z^k$ resp. $z^m = 1$ for $m = \ell - k > 0$.

Obviously the second case applies for a finite group of units. Indeed $m = |R^*|$ does the job for any $z \in R^*$:

Theorem 7.9 (Lagrange)**.** *Let* R *be a commutative ring,* $|R^*| = q < \infty$. *Then we have*

$$z^q = 1$$

for all $z \in R^*$.

Proof. The map $\mu_z : R^* \longrightarrow R^*, x \mapsto zx$, is bijective, hence

$$\prod_{x \in R^*} x = \prod_{x \in R^*} \mu_z(x) = \prod_{x \in R^*} (zx) = z^q \prod_{x \in R^*} x.$$

Since units are non-zero divisors, we conclude $z^q = 1$. □

Corollary 7.10. *1. Euler: If* $\gcd(a, n) = 1$, *we have*

$$a^{\varphi(n)} \equiv 1 \mod (n).$$

2. *Little Fermat: If p is prime and does not divide a, then*

$$a^{p-1} \equiv 1 \mod (p).$$

Proof. Little Fermat is a special case of Euler, take $n = p$, while Euler follows from Lagrange with $R = \mathbb{Z}_n$. □

Public Key Cryptography: We shall explain here how one can encode messages using a public key, such that only receivers with an additional private key can decode them.

1. Take a number $n = pq \in \mathbb{N}$, which is the product of two different prime numbers p, q.

2. Messages, both plain text and encoded, are represented as elements $x \in \mathbb{Z}_n$.

3. If $k = 1 + r(p-1)(q-1)$, we have

$$x^k = x$$

for all $x \in \mathbb{Z}_n$: Since $\mathbb{Z}_n \cong \mathbb{Z}_p \times \mathbb{Z}_q$, it is sufficient to prove $x^k = x$ for $x \in \mathbb{Z}_p$ and $x \in \mathbb{Z}_q$. For $x = 0$, it is clear, and if, say $x \in \mathbb{Z}_p \setminus \{0\} = \mathbb{Z}_p^*$, we find

$$x^k = x \cdot (x^{p-1})^{r(q-1)} = x \cdot 1^{r(q-1)} = x.$$

4. Encoding is described by the map

$$\mathbb{Z}_n \longrightarrow \mathbb{Z}_n, x \mapsto y = x^e,$$

with an exponent e relatively prime to $\varphi(n) = (p-1)(q-1)$.

5. The pair (n, e) is public knowledge, but not the factors p, q. The point here is, that for great n an effective factorization - at least nowadays - is impossible.

6. In order to decode one needs an additonal private key, a number $d \in \mathbb{N}$, such that $ed \equiv 1 \mod (p-1)(q-1)$. An encoded message $y \in \mathbb{Z}_n$ is decoded applying the map

$$y \mapsto y^d.$$

7. In order to obtain d from the public key (n, e) one needs the factorization $n = pq$.

Example 7.11. We take $n = 77, e = 23$, and want to decode $y = \overline{4} \in \mathbb{Z}_{77}$. Luckily we know $77 = 7 \cdot 11$ and $\varphi(77) = 60$ and compute

$$\mathbb{Z}_{60} \ni \overline{d} = \overline{23}^{-1} = \overline{47}.$$

Thus we obtain the decoded message

$$x = \overline{4}^{47} \in \mathbb{Z}_{77} \cong \mathbb{Z}_7 \times \mathbb{Z}_{11}.$$

In order to avoid tedious computations, we use the chinese remainder theorem (of course, in a realistic situation, that is not possible, since the factorization $n = pq$ is. presumably, unknown to the receiver)

$\mathbb{Z}_{77}$	$\mathbb{Z}_7 \times \mathbb{Z}_{11}$
$\overline{22}$	$(\overline{1}, \overline{0})$
$\overline{56}$	$(\overline{0}, \overline{1})$
$\overline{4}$	$(\overline{4}, \overline{4})$
$\overline{4}^{47}$	$(\overline{4}^5, \overline{4}^7)$
$\overline{16}$	$(\overline{2}, \overline{5})$

,

so we get finally $\overline{16}$.

8 Polynomial equations over $\mathbb{Z}_n$

Any polynomial $f = \sum a_\nu X^\nu \in \mathbb{Z}[X]$ induces a function

$$R \longrightarrow R, \xi \mapsto f(\xi) := \sum a_\nu \xi^\nu$$

for any (commutative) ring R. We want to investigate what can be said about the zeros of that function in the case, where

$$R = \mathbb{Z}_n = \{\overline{0}, \overline{1}, ...-, \overline{n-1}\}$$

is the residue class ring $\mathbb{Z}_n := \mathbb{Z}/\mathbb{Z}n$. Then we have for $\xi := \overline{x} \in \mathbb{Z}_n$ the following equivalence

$$f(\xi) = 0 \Longleftrightarrow n|f(x) :\Longleftrightarrow f(x) \equiv 0 \mod (n).$$

Furthermore remember that

$$\mathbb{Z}_n^* = \{\overline{a}; \gcd(a,n) = 1\}$$

is the group of units of $\mathbb{Z}_n$.

1.) Linear polynomials: The first case we study is that of a linear polynomial

$$f = aX + b.$$

The solutions $\xi \in \mathbb{Z}_n$ are the residue classes $\xi = \overline{x}$, such that there is some $y \in \mathbb{Z}$ with

$$ax + b = yn.$$

Such a pair (x, y) exists, iff $d := \gcd(a, n)$ divides the constant term b. The condition is obviously necessary, but also sufficient: Writing $a_0 = a/d, b_0 = b/d, n_0 = n/d$ we have

$$ax + b \equiv 0 \mod (n) \Longleftrightarrow a_0 x + b_0 \equiv 0 \mod (n_0).$$

Furthermore $\overline{a}_0 \in \mathbb{Z}_{n_0}^*$ and for $\zeta \in \mathbb{Z}_{n_0}$ and $f_0 := a_0 X + b_0$ we obtain :

$$f_0(\zeta) = 0 \Longleftrightarrow \zeta = -(\overline{a}_0)^{-1} \cdot \overline{b}_0.$$

It follows that

$$f(\xi) = 0 \Longleftrightarrow \xi \mapsto \zeta = \overline{z},$$

where $\mapsto$ denotes the natural map $\mathbb{Z}_n \longrightarrow \mathbb{Z}_{n_0}, x + \mathbb{Z}n \mapsto x + \mathbb{Z}n_0$. Hence

$$\xi = \overline{z + \nu n_0} \in \mathbb{Z}_n, \nu = 0, ..., d-1$$

are the zeros of $f = aX + b$ in $\mathbb{Z}_n$. In order to find $\zeta = \overline{z}$ we may either invert $\overline{a}_0 \in \mathbb{Z}_{n_0}^*$ or solve directly the congruence

$$a_0 z \equiv -b \mod (n_0).$$

For an explicit calculation we refer to LÅL, Ex.5.1.

2.) Reduction to prime power moduli: Assume

$$n = p_1^{k_1} \cdot .... \cdot p_r^{k_r},$$

and denote

$$\psi : \mathbb{Z}_n \longrightarrow \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_r^{k_r}},$$

$$\xi = \overline{x} \mapsto (\xi_1, ..., \xi_r), \quad \xi_i := x + \mathbb{Z}p_i^{k_i},$$

the chinese remainder isomorphism. Then we have

$$f(\xi) = 0 \Longleftrightarrow f(\xi_i) = 0, i = 1, ..., r.$$

Hence we have to look for $\xi_i \in \mathbb{Z}_{p_i^{k_i}}$ with $f(\xi_i) = 0$ and may take

$$\xi = \psi^{-1}(\xi_1, ..., \xi_r).$$

Remark 8.1. Here we discuss how to evaluate

$$\psi^{-1} : \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_r^{k_r}} \longrightarrow \mathbb{Z}_n$$

or, equivalently, how to find a simultaneous representative x for $\xi_1 = \overline{x}_1, ..., \xi_r = \overline{x}_r$. Denote

$$\mathbf{e}_i := (\overline{0}, ..., \overline{0}, \overline{1}, \overline{0}, ..., \overline{0}) \in \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_r^{k_r}}$$

the "i-th unit vector" $\mathbf{e}_i := (\overline{\delta}_{i,1}, ..., \overline{\delta}_{i,r})$. Since

$$(\xi_1, ..., \xi_r) = \sum_{i=1}^{r} x_i \mathbf{e}_i,$$

we have

$$\psi^{-1}(\xi_1, ..., \xi_r) = \sum_{i=1}^{r} x_i \psi^{-1}(\mathbf{e}_i)$$

and it suffices to find $\overline{b}_i = \psi^{-1}(\mathbf{e}_i) \in \mathbb{Z}_n$ for $i = 1, ..., r$ as follows: Writing

$$n = n_i (p_i)^{k_i}$$

we have

$$b_i = c_i n_i,$$

where $t = c_i$ is a solution of the congruence

$$n_i t \equiv 1 \mod ((p_i)^{k_i}).$$

Or simply guess the correct value by checking all integer multiples $cn_i, c \in \mathbb{Z}$.

For an explicit calculation we refer to LÅL, Ex.6.1.

3.) Raising the exponent of a prime power modulus: We consider the cmmutative diagram

$$\begin{array}{ccccc} \xi & \in & \mathbb{Z}_{p^{k+1}} & \xrightarrow{f} & \mathbb{Z}_{p^{k+1}} \\ & & \downarrow & & \downarrow \\ \varrho(\xi) & \in & \mathbb{Z}_{p^k} & \xrightarrow{f} & \mathbb{Z}_{p^k} \end{array},$$

where the vertical arrows are the natural map $\varrho : \mathbb{Z}_{p^{k+1}} \longrightarrow \mathbb{Z}_{p^k}$. Obviously, if $\xi \in \mathbb{Z}_{p^{k+1}}$ is a zero of f, then $\varrho(\xi) \in \mathbb{Z}_{p^k}$ is a zero of f as well. On the other hand, if $\varrho(\xi)$ is a zero of f, then

$$\varrho^{-1}(\varrho(\xi)) = \{\xi_t = \xi + t \cdot \overline{p}^k, t = 0, ..., p-1\}$$

holds for the set $\varrho^{-1}(\varrho(\xi))$ of residue classes above $\varrho(\xi)$. We want to know, for which $t, 0 \leq t < p$, the element $\xi_t \in \varrho^{-1}(\varrho(\xi))$ is a zero of f.

We remark that for $\xi = \overline{a}$ the assumption $f(\varrho(\xi)) = 0$ means that $p^k | f(a)$. Indeed a straight forward argument shows:

Proposition 8.2. *Let $\xi = \overline{a} \in \mathbb{Z}_{p^{k+1}}$. If $f(\varrho(\xi)) = 0$, then we have*

$$f(\xi_t) = 0 \Longleftrightarrow f'(a)t \equiv -\frac{f(a)}{p^k} \mod (p).$$

In particular,

1. *if $p \not| f'(a)$, there is exactly one t with $f(\xi_t) = 0$,*
2. *if $p | f'(a)$ and $p^{k+1} | f(a)$, we have $f(\xi_t) = 0$ for all t,*
3. *and finally, if $p | f'(a)$ and $p^{k+1} \not| f(a)$, we have $f(\xi_t) \neq 0$ for all t.*

Proof. We have

$$f(a + Y) = f(a) + f'(a)Y + Y^2 g(Y)$$

with a polynomial $g \in \mathbb{Z}[Y]$. Hence, substituting tp^k for Y and passing from $\mathbb{Z}$ to $\mathbb{Z}_{p^{k+1}}$ yields

$$\mathbb{Z}_{p^{k+1}} \ni 0 = f(\xi + t\overline{p}^k) = f(\xi) + tf'(\xi)\overline{p}^k \Longleftrightarrow p^{k+1} | f(a) + tf'(a)p^k,$$

the latter being equivalent to $p | f(a) + tf'(a)$. □

Theorem 8.3 (Hensel's lemma)**.** *Let $f \in \mathbb{Z}[X]$, p a prime and $\xi \in \mathbb{Z}_p$ with $f(\xi) = 0$ and $f'(\xi) \neq 0$. Then for any $k \in \mathbb{N}_{>0}$ there is a unique zero $\xi_k \in \mathbb{Z}_{p^k}$ of f lying above $\xi \in \mathbb{Z}_p$.*

Remark 8.4. Hensel's lemma provides a sequence $\xi_k \in \mathbb{Z}_{p^k}$, s.th. $f(\xi_k) = 0$ and $\varrho(\xi_k) = \xi_{k-1}$, but that does not imply that there is some simultaneous representative $x \in \mathbb{Z}$, i.e. such that $\xi_k = \overline{x}$ holds for all $k \in \mathbb{N}$. If so, we would have $f(x) = 0$ as well. In order to understand what is happening here, we write $\xi_k = \overline{x}_k$ with $0 \leq x_k < p^k$ and look at its p-adic expansion: We have

$$x_k = \sum_{\nu=0}^{k-1} t_\nu p^\nu,\ 0 \leq t_\nu < p$$

with the digits t_ν being independent of $k \in \mathbb{N}$. In particular a global representative exists iff the increasing sequence $(x_k)_{k\geq 1}$ is bounded or equivalently $t_\nu = 0$ for $\nu \gg 0$. But that need not be the case in general: Take $f = X^2 + 1$ and $p = 5$. Then we have $f(\overline{2}) = 0$ and $f'(\overline{2}) = \overline{4} \neq 0$, so Hensel's lemma applies. But f has obviously no zero in $\mathbb{Z}$.

Indeed, there are uncountably many sequences ξ_k reap. x_k, so the overwhelming majority has no global representative.

In the optional section 10 we describe, how one can, analogous to the creation of real numbers from rational ones, produce p-adic integers, which admit infinite expansions

$$x = \sum_{\nu=0}^{\infty} t_\nu p^\nu,\ 0 \leq t_\nu < p.$$

Proof of Th.8.3. We use induction on $k \in \mathbb{N}_{>0}$. For $k = 1$ the claim is true by assumption. Assume $\overline{a}_k \in \mathbb{Z}_{p^k}$ has been found. Since $a_k \equiv a \mod (p)$, we have as well $f'(a_k) \equiv f'(a) \mod (p)$ resp. $p \not| f'(a_k)$. So we may apply Prop. 8.2. □

Example 8.5. We take $p = 3$ and look for zeros of $f = 7X^6 + 4X + 12$ in $\mathbb{Z}_{27}$.

1. For $\xi \in \mathbb{Z}_3$ we have $f(\xi) = \xi^6 + \xi = \xi^2 + \xi = \xi(\xi + 1)$ and find the zeros $\xi = \overline{0}, \xi = -\overline{1}$. We have $f' = 42X^5 + 4$, hence $f'(\xi) = 1$ for all $\xi \in \mathbb{Z}_3$. Hence we may apply Hensel's lemma and find that there are two zeros of f in $\mathbb{Z}_{27}$.

2. Now let us look for the zeros of f in $\mathbb{Z}_9$. One is of the form $t\overline{3}$ with

$$t \overset{(3)}{\equiv} -\frac{f(0)}{3} = -\frac{12}{3} = -4,$$

hence $t = -1$ is possible and $-\overline{3} \in \mathbb{Z}_9$ is the zero of f above $\overline{0} \in \mathbb{Z}_3$.

The other one is of the form $-\overline{1} + t\overline{3}$, where

$$t \overset{(3)}{\equiv} -\frac{f(-1)}{3} = -\frac{15}{3} = -5,$$

hence $t = 1$ is possible and $-\overline{1} + \overline{3} = \overline{2} \in \mathbb{Z}_9$ is the zero of f above $-\overline{1} \in \mathbb{Z}_3$.

3. Finally we have to lift the zeros $-\overline{3}, \overline{2} \in \mathbb{Z}_9$. The first zero is $-\overline{3} + t\overline{9} \in \mathbb{Z}_{27}$ with

$$t \overset{(3)}{\equiv} -\frac{f(-3)}{9} = -\frac{7 \cdot 3^6}{9} = -7 \cdot 3^4,$$

hence $t = 0$ is possible and $-\overline{3} \in \mathbb{Z}_{27}$ is the zero of f above $-\overline{3} \in \mathbb{Z}_9$.

The other one is of the form $\overline{2} + t\overline{9}$, where

$$t \overset{(3)}{\equiv} -\frac{f(2)}{9} = -\frac{90}{9} = -10,$$

hence $t = -1$ is possible and $\overline{2} - \overline{9} = -\overline{7} \in \mathbb{Z}_{27}$ is the zero of f above $\overline{2} \in \mathbb{Z}_9$.

4.) Prime moduli: There is no general recipe how to solve $f(\xi) = 0$ over a field $\mathbb{Z}_p$. In any case we may assume that $\deg f < p$: The polynomial

$$h := X^p - X$$

satisfies $h(\xi) = 0$ for all $\xi \in \mathbb{Z}_p$ according to Th.7.9, so if we write

$$f = gh + r$$

with $\deg r < p$, the remainder polynomial r and f have the same zeros in $\mathbb{Z}_p$. The polynomial $\overline{h} \in \mathbb{Z}_p[X]$ induced by $h \in \mathbb{Z}[X]$ satisfies

$$\overline{h} = \prod_{a \in \mathbb{Z}_p} (X - a).$$

After division with X we obtain

$$X^{p-1} - 1 = \prod_{a \in \mathbb{Z}_p^*} (X - a),$$

in particular

$$-1 = (-1)^{p-1} \prod_{a \in \mathbb{Z}_p^*} a.$$

Thus we have found

Theorem 8.6 (Wilson's theorem)**.** *For a prime number p we have*

$$(p-1)! \equiv -1 \mod (p).$$

Proof. Indeed

$$\prod_{a \in \mathbb{Z}_p^*} a = \overline{(p-1)!}.$$

□

9 The ring of p-adic integers (optional)

Hensel's lemma provides solutions of polynomial congruences with arbitrarily high prime powers as moduli. In order to express this fact in a more concise and satisfactory way, one introduces new numbers, the p-adic integers, which play relative to usual integers a rôle analogous to that one of real numbers in relation to rational numbers. But there are also a lot of strange features.

Definition 9.1. Let p be a prime number. A p-adic integer is a sequence (or family)

$$\mathbf{a} = (a_\nu)_{\nu \in \mathbb{N}} \in \prod_{\nu=0}^{\infty} \mathbb{Z}_{p^{\nu+1}},$$

such that the immediate successor $a_{\nu+1} \in \mathbb{Z}_{p^{\nu+2}}$ of $a_\nu \in \mathbb{Z}_{p^{\nu+1}}$ lies above (or is a lift of) a_ν, i.e.

$$\mathbb{Z}_{p^{\nu+2}} \in a_{\nu+1} \mapsto a_\nu \in \mathbb{Z}_{p^{\nu+1}}$$

holds for all $\nu \in \mathbb{N}_{>0}$. We denote

$$\mathbb{Z}_{(p)} \subset \prod_{\nu=0}^{\infty} \mathbb{Z}_{p^{\nu+1}}$$

the set of all p-adic integers. We add and multiply p-adic integers componentwise.

Remark 9.2. 1. There is a natural injection

$$\mathbb{Z} \hookrightarrow \mathbb{Z}_{(p)}, n \mapsto \mathbf{n} := (n + \mathbb{Z}p^{\nu+1})_{\nu\in\mathbb{N}},$$

i.e. usual integers are also p-adic integers. From now on we treat the above injection as an inclusion and write n for $\mathbf{n}$.

2. The set $\mathbb{Z}_{(p)}$ forms a commutative ring.

3. We want to define an absolute value

$$|\ \ | : \mathbb{Z}_{(p)} \longrightarrow \mathbb{R}_{\geq 0}.$$

To do so we start with a valuation

$$v : \mathbb{Z}_{(p)} \longrightarrow \mathbb{N} \cup \{\infty\}.$$

Namely, for $\mathbf{a} := (a_\nu)_{\nu\in\mathbb{N}}$ let

$$v(\mathbf{a}) := \min\{\nu; a_\nu \neq 0\}.$$

Then the following holds

(a) $v(\mathbf{a} + \mathbf{b}) \geq \min(v(\mathbf{a}), v(\mathbf{b}))$
(b) $v(\mathbf{ab}) = v(\mathbf{a}) + v(\mathbf{b})$
(c) $v(\mathbf{a}) = \infty \Longleftrightarrow \mathbf{a} = 0.$

Now the absolute value of a p-adic integer is

$$|\mathbf{a}| := p^{-v(\mathbf{a})}$$

(with $p^{-\infty} := 0$). It satisfies

(a) $|\mathbf{a} + \mathbf{b}| \leq \max(|\mathbf{a}|, |\mathbf{b}|)$

(b) $|\mathbf{ab}| = |\mathbf{a}| \cdot |\mathbf{b}|$

(c) $|\mathbf{a}| = 0 \Longleftrightarrow \mathbf{a} = 0$.

In particular the rule (b) shows that $\mathbb{Z}_{(p)}$ is an integral domain.

4. We note that always
$$|\mathbf{a}| \leq 1$$
and that for $m = ap^n \in \mathbb{Z} \subset \mathbb{Z}_{(p)}$ with a not divisible with p we have
$$|m| = p^{-n}.$$

5. Using the absolute we may define the convergence of a sequence of p-adic integers:
$$\mathbf{a}_n \to \mathbf{a} :\Longleftrightarrow |\mathbf{a}_n - \mathbf{a}| \to 0.$$
We see that
$$p^n \to 0$$
and that any series
$$\sum_{n=0}^{\infty} \mathbf{a}_n p^n$$
converges. Indeed any p-adic integer has such an expansion with coefficients
$$\mathbf{a}_n = t_n \in [0, p-1] \cap \mathbb{N}.$$
Namely if $\mathbf{a} = (\bar{\ell}_\nu)_{\nu \in \mathbb{N}}$ with $0 \leq \ell_\nu < p^{\nu+1}$, then
$$\mathbf{a} = \lim_{\nu \to \infty} \ell_\nu.$$
Now we expand ℓ_ν w.r.t. the basis p and obtain
$$\ell_\nu = \sum_{n=0}^{\nu} t_n p^n,$$
where the digits t_n satisfy $0 \leq t_n < p$ and do not depend on ν due to $\ell_{\nu+1} \equiv \ell_\nu \mod (p^{\nu+1})$. Finally we arrive at an infinite p-adic expansion
$$\mathbf{a} = \sum_{n=0}^{\infty} t_n p^n$$
with unique digits $t_n, 0 \leq t_n < p$.

6. In order to expand a real number we can also work with the prime p as basis (instead of 10), but get series in the opposite direction

$$\mathbb{R} \ni x = \sum_{n=\ell}^{\infty} t_n p^{-n}.$$

7. Let us return to p-adic integers: We have $1 - p \in \mathbb{Z}_{(p)}^*$ and

$$(1-p)^{-1} = \sum_{n=0}^{\infty} p^n,$$

so a negative rational number is the infinite sum of positive numbers: This shows that p-adic integers do not admit an order relation compatible with ring operations.

8. p-adic numbers vs. real and complex numbers: If $p \equiv 1 \mod (4)$, the polynomial $X^2 + 1$ has a zero in $\mathbb{Z}_p$, e.g. for $p = 5$ we could take $\overline{2} \in \mathbb{Z}_5$. Now Hensel's lemma tells us that there is a p-adic integer $\mathbf{a} \in \mathbb{Z}_{(p)}$ with $\mathbf{a}^2 = -1$. Hence it is not possible to find an embedding $\mathbb{Z}_{(p)} \hookrightarrow \mathbb{R}$ such that the addition and multiplication of real numbers coincides with the p-adic operations. One can show that there exists such an embedding $\mathbb{Z}_{(p)} \hookrightarrow \mathbb{C}$, but can't describe it explicitly, so that result is not particularly helpful.

Theorem 9.3 (Hensel's lemma)**.** *Let $f \in \mathbb{Z}[X]$. Then any simple zero $a \in \mathbb{Z}_p$ of f (i.e. $f'(a) \neq 0 \in \mathbb{Z}_p$) can be lifted to a unique zero $\mathbf{a} := (a_\nu)_{\nu \in \mathbb{N}} \in \mathbb{Z}_{(p)}$, i.e. $a_0 = a$.*

We conclude this section with some "geometric" considerations:

Proposition 9.4. *1. The group of units of the ring of p-adic integers is the p-adic unit sphere:*

$$\mathbb{Z}_{(p)}^* = \{\mathbf{a} = (a_\nu)_{\nu \geq 0}; a_0 \neq 0\} = \{\mathbf{a} \in \mathbb{Z}_{(p)}; |\mathbf{a}| = 1\}.$$

2. *Any* $\mathbf{a} = (a_\nu) \in \mathbb{Z}_{(p)} \setminus \{0\}$ *can uniquely be written as a product*

$$\mathbf{a} = p^{v(\mathbf{a})}\mathbf{e},\ \ \mathbf{e} \in \mathbb{Z}_p^*.$$

In particular $\mathbb{Z}_{(p)}$ *resembles an onion: The set*

$$\mathbb{Z}_{(p)} = \{0\} \cup \bigcup_{n=0}^{\infty} p^n \mathbb{Z}_{(p)}^*,$$

is the countable union of the spheres (or shells?) $p^n\mathbb{Z}_{(p)}^*$ *of radius* p^{-n} *and their common center* $0 \in \mathbb{Z}_{(p)}$.

Remark 9.5. Let $p > 2$. Though $\mathbb{Z}_{p^m}^*$ is cyclic for every $m \in \mathbb{N}$, the group of units $\mathbb{Z}_{(p)}^*$ is not. Indeed it is uncountable!

Proof. The first part is obvious. Let $n := v(\mathbf{a})$, so $a_\nu = 0$ for $\nu < n$ and $a_n \neq 0$, then, writing $a_\nu = k_\nu + \mathbb{Z}p^{\nu+1}$, we have $k_{n+\nu} = p^n\ell_\nu$ for $\nu \geq 0$, and may take $\mathbf{e} = (\ell_{n+\nu} + \mathbb{Z}p^{\nu+1})_{\nu\in\mathbb{N}}$. □

p-adic number fields: Finally one wants to enlarge $\mathbb{Z}_{(p)}$ to a field by adding new shells to the onion $\mathbb{Z}_{(p)}$: One takes

$$\mathbb{Q}_p := \left\{ \frac{\mathbf{a}}{p^n}; \mathbf{a} \in \mathbb{Z}_{(p)}, n \in \mathbb{N} \right\},$$

where fractions with a p-adic integer as numerator and a p-power as denominator are added and multiplied in the usual way. We thus get a field, whose elements are called p-adic numbers; it is an onion

$$\mathbb{Q}_p = \{0\} \cup \bigcup_{n=-\infty}^{\infty} p^n \mathbb{Z}_{(p)}^*$$

of infinite radius. (The absolute value for p-adic integers extends in a unique way to $\mathbb{Q}_p$.)

10 Primitive roots

In this section we discuss for a prime power modulus $n = p^r$ the zeros of the two term polynomial

$$f = X^\ell - a,$$

i.e. we have to decide, whether $\overline{a} \in \mathbb{Z}_{p^r}$ is an ℓ-th power or not. First of all we may reduce everything to the case where $\overline{a} \in \mathbb{Z}_n^*$ is a unit. In order to see that we observe that $\mathbb{Z}_{p^r}$ is a sort of onion: It is the disjoint union

$$\mathbb{Z}_{p^r} = \bigcup_{k=0}^{r-1} \mathbb{Z}_{p^r}^* \cdot \overline{p}^k \cup \{0\},$$

of the shells $\mathbb{Z}_{p^r}^* \cdot \overline{p}^k$ and the center $\{0\}$ (the group of units being the outer shell). The k-th shell of the onion admits a bijection:

$$\mathbb{Z}_{p^r}^* \cdot \overline{p}^k \longrightarrow \mathbb{Z}_{p^{r-k}}^*, \quad e\overline{p}^k \mapsto \varrho(e),$$

where $\varrho : \mathbb{Z}_{p^r} \longrightarrow \mathbb{Z}_{p^{r-k}}$ denotes as usual the natural map. In particular the shells are shrinking with increasing k. We find:

1. If $\overline{a} = 0$, the elements $\xi = e\overline{p}^s, \ell s \geq r$, are the solutions of the equation $f(\xi) = 0$.

2. On the other hand, if $\overline{a} = c\overline{p}^k, k < r$, there is a zero in $\mathbb{Z}_{p^r}$, iff $k = \ell s, s \in \mathbb{N}$, and $\varrho(c) = e^\ell$ with some $e \in \mathbb{Z}_{p^{r-k}}^*$. The solutions are then of the form
$$\xi = \tilde{e}\overline{p}^s, \quad \mathbb{Z}_{p^r}^* \ni \tilde{e} \mapsto e \in \mathbb{Z}_{p^{r-k}}^*.$$

Thus we are left with the problem to determine how many ℓ-th roots a given unit admits. A possible strategy could be to apply Hensels lemma. Indeed:

Proposition 10.1. *Let $f = X^\ell - a$ with $\gcd(p, a) = 1$..*

1. *If p does not divide ℓ the zeros of f in $\mathbb{Z}_{p^r}$ correspond one-to-one to the zeros of f in the field $\mathbb{Z}_p$.*

2. *If $\gcd(\ell, p-1) = 1$ the map $\mathbb{Z}_p^* \longrightarrow \mathbb{Z}_p^*, \xi \mapsto \xi^\ell$ is bijective. In particular there is a unique zero of f in $\mathbb{Z}_p$.*

Proof. We have $f' = \ell X^{\ell-1}$, hence $f'(\xi) = 0$ for all zeros of f in $\mathbb{Z}_p$. For the second part, choose $r, s \in \mathbb{Z}$ with $r\ell + s(p-1) = 1$. Then $\xi \mapsto \xi^r$ is the map $\mathbb{Z}_p^* \longrightarrow \mathbb{Z}_p^*$ inverse to $\xi \mapsto \xi^\ell$ because of $\xi^{p-1} = 1$ for $\xi \in \mathbb{Z}_p^*$. □

For a more detailed study we need a sort of logarithm for units: A primitive element is an element $a \in \mathbb{Z}_n^*$, such that any unit b is a power $b = a^\nu$ with

some $\nu \in \mathbb{Z}$, the "logarithm of b with basis a". Indeed the exponent $\nu \in \mathbb{Z}$ is unique only up to a multiple of $\varphi(n) = |\mathbb{Z}_n^*|$. Hence, a power equation

$$x^\ell = b$$

is, writing $x = a^t, b = a^\nu$, transformed into the linear congruence:

$$\ell t \equiv \nu \mod (\varphi(n)).$$

In the following we consider a commutative ring R with finite group of units R^*, denote

$$q := |R^*|$$

its order. We are hunting for an element $a \in R^*$ with

$$R^* = a^{\mathbb{Z}} := \{a^\nu; \nu \in \mathbb{Z}\};$$

and we are obviously done, if we find some $a \in R^*$ with

$$|a^{\mathbb{Z}}| = q.$$

Let $a \in R^*$ be any element. Since there are only finitely many units, we have $a^\mu = a^\nu$ for suitable exponents $\mu > \nu \geq 0$, in particular $k = \mu - \nu > 0$ satisfies $a^k = 1$. The least such exponent deserves a name:

Definition 10.2. The order $\mathrm{ord}(a) \in \mathbb{N}_{>0}$ of an element $a \in R^*$ is defined as

$$\mathrm{ord}(a) := \min\{k \in \mathbb{N}_{>0}; a^k = 1\}.$$

Example 10.3.

1. The elements $\overline{3}, \overline{5}, \overline{7} \in \mathbb{Z}_8^*$ have order 2.
2. The element $\overline{2} \in \mathbb{Z}_{11}^*$ has order 10, its sucessive powers being

$$\overline{2}, \overline{4}, \overline{8}, \overline{5}, \overline{10} = -\overline{1}, -\overline{2} = \overline{9}, -\overline{4} = \overline{7}, -\overline{8} = \overline{3}, -\overline{5} = \overline{6}, \overline{1}.$$

Furthermore $\mathrm{ord}(\overline{4}) = 5$ and $\mathrm{ord}(\overline{10}) = 2$.

Lemma 10.4. *For an element $a \in R^*$ of order $d := \mathrm{ord}(a)$ we have:*

1. *$a^{\mathbb{Z}} = \{1, a, ..., a^{d-1}\}$, where the powers $1, a, ..., a^{d-1}$ are pairwise different, and*

2. $a^k = 1 \Longleftrightarrow d|k$. *In particular* $d|q$.

Proof. We have $a^k = a^r$, if $k = sd + r$ with $0 \leq r < d$. This gives the nontrivial inclusion in the first part of the statement as well as the second part: If $a^k = 1$, we have as well $a^r = 1$ and that is possible only for $r = 0$, since $d = \mathrm{ord}(a)$. Furthermore $k = q$ is possible, since $a^q = 1$ holds according to Th. 7.9. Finally $a^\mu = a^\nu$ with $0 \leq \mu < \nu < d$ would give $a^{\nu-\mu} = 1$ with $0 < \nu - \mu < d$, a contradiction. □

We are looking for exponents ℓ, that "kill" all elements in R^* simultaneously:

$$a^\ell = 1, \ \forall a \in R^*.$$

We know that $\ell = q$ is such an exponent. Indeed, the least possible choice of such an exponent is $\ell = n$ with

$$n := \mathrm{lcm}\left\{\mathrm{ord}(a); a \in R^*\right\}.$$

Proposition 10.5. *For an integral domain* R *with* $q = |R^*| < \infty$ *we have*

$$q = \mathrm{lcm}\left\{\mathrm{ord}(a); a \in R^*\right\}.$$

Proof. The polynomial $f = X^n - 1 \in R[X]$ with $n := \mathrm{lcm}\left\{\mathrm{ord}(a); a \in R^*\right\}$ vanishes on R^* and has at most n zeros, hence $q \leq n$. Since by Th. 7.9 we have $n|q$ it follows $q = n$. □

Example 10.6. For a non-integral domain the statement does not hold:

1. For $R = \mathbb{Z}_8$ we have
$$R^* = \left\{\overline{1}, \overline{3}, \overline{5}, \overline{7}\right\},$$
hence $q = 4$ and $n = 2$.

2. Let p, r be two different odd primes. The group of units
$$\mathbb{Z}_{pr}^* \cong \mathbb{Z}_p^* \times \mathbb{Z}_r^*$$
has $q = (p-1)(r-1)$ elements, but
$$a^\ell = 1$$
holds already for $\ell = (p-1)(r-1)/2$, since ℓ is divisible both with $p-1$ and $r-1$. Thus $n|\frac{q}{2}$.

Proposition 10.7. *1. Let* $\mathrm{ord}(a) = m$. *Then*

$$\mathrm{ord}(a^k) = \frac{m}{\gcd(k,m)}.$$

2. If $m = \mathrm{ord}(a)$ *and* $n = \mathrm{ord}(b)$ *are relatively prime, we have*

$$\mathrm{ord}(ab) = mn.$$

Proof. We have

$$1 = (a^k)^\ell = a^{k\ell} \Longleftrightarrow m | k\ell \Longleftrightarrow \frac{m}{\gcd(k,m)} | \ell.$$

Let

$$1 = (ab)^\ell \Longrightarrow a^\ell = b^{-\ell} \Longrightarrow \mathrm{ord}(a^\ell) = \mathrm{ord}(b^{-\ell})$$
$$\Longrightarrow \frac{m}{\gcd(\ell,m)} = \frac{n}{\gcd(\ell,n)},$$

i.e. $1 = 1$ and

$$\gcd(\ell,m) = m, \gcd(\ell,n) = n \Longrightarrow \ell \in \mathbb{Z} \cdot mn.$$

□

Theorem 10.8. *Assume that the group of units* R^* *of the integral domain* R *is finite:* $q := |R^*|$. *Then there is a primitive root, i.e. an element* $a \in R^*$ *with*

$$\mathrm{ord}(a) = q.$$

resp.

$$R^* = a^{\mathbb{Z}} = \{1, a, ..., a^{q-1}\}.$$

Proof. We have to find an element $a \in R^*$ with $\mathrm{ord}(a) = q$. Let

$$q = p_1^{k_1} \cdot ... \cdot p_s^{k_s}$$

be the prime factorization of q. For each $i = 1, ..., s$ we construct an element $a_i \in R^*$ of order $\mathrm{ord}(a_i) = p_i^{k_i}$ and set

$$a = a_1 \cdot ... \cdot a_s.$$

Now for any $c \in R^*$ write

$$\mathrm{ord}(c) = p_1^{k_1(c)} \cdot ... \cdot p_s^{k_s(c)} \quad \text{with } 0 \leq k_i(c) \leq k_i.$$

As a consequence of Prop. 10.5 we see, that

$$k_i = \max\left\{k_i(c); c \in R^*\right\},\ i = 1, ..., s.$$

In particular there is an element c_i, whose order is divisible with $p_i^{k_i}$, say $\mathrm{ord}(c_i) = \ell_i p_i^{k_i}$. We now take $a_i := c_i^{\ell_i}$. □

Corollary 10.9. *Let p be a prime number. Then the group of units $\mathbb{Z}_p^*$ of $\mathbb{Z}_p$ admits a primitive root.*

Let us now consider prime power moduli $p^r, r \geq 2$. We have to investigate certain *subgroups* $G \subset \mathbb{Z}_{p^r}^*$ of the group of units $\mathbb{Z}_{p^n}^*$.

Definition 10.10. A non-empty subset $G \subset R^*$ is called a subgroup, if it is closed w.r.t. multiplication, i.e.

$$x, y \in G \Longrightarrow xy \in G.$$

It is called *cyclic* if it admits a primitive element $a \in G$, i.e. such that

$$G = a^{\mathbb{Z}}.$$

Remark 10.11. Since R^* consists only of elements of finite order, a subgroup $G \subset R^*$ is even closed w.r.t. inversion:

$$a^{-1} = a^{\mathrm{ord}(a)-1}$$

and contains the neutral element $1 = aa^{-1}$.

Example 10.12. For $R = \mathbb{Z}_{p^r}$ and $n \geq 1$ we find the subgroups

$$G_n := \overline{1} + p^n \mathbb{Z}_{p^r} \subset \mathbb{Z}_{p^r}^*.$$

We have

$$G_1 \supset G_2 \supset ... \supset G_r = \{\overline{1}\}$$

and

$$|G_n| = p^{r-n} \text{ for } n \leq r.$$

For $p = 2$ we find $G_1 = \mathbb{Z}_{2^r}^*$.

In the next proposition we study the action of the Frobenius map

$$\mathbb{Z}_{p^r} \longrightarrow \mathbb{Z}_{p^r}, x \mapsto x^p,$$

on those subgroups:

Proposition 10.13. *We have*

$$a \in G_n \Longrightarrow a^p \in G_{n+1}.$$

More precisely:

$$(G_n \setminus G_{n+1})^p \subset G_{n+1} \setminus G_{n+2}$$

holds for

1. $n = 1, ..., r-2$ *and* $p > 2$,
2. $n = 2, ..., r-2$ *and* $p = 2$.

Hence, for $p > 2$ *and* $a \in G_1$ *we have*

$$\mathrm{ord}(a) = p^{r-1} = |G_1| \Longleftrightarrow a \notin G_2,$$

while for $p = 2$ *and* $a \in G_2$ *we get*

$$\mathrm{ord}(a) = p^{r-2} = |G_2| \Longleftrightarrow a \notin G_3.$$

Proof. We have

$$(1 + p^n t)^p = 1 + p^{n+1}t + \sum_{\nu=2}^{p} \binom{p}{\nu} (pt)^{\nu n} \equiv 1 \mod (p^{n+1}).$$

If $p > 2$ or $p = 2$ and $n \geq 2$, the sum is even divisible with p^{n+2}, hence for $p \not| t$ the RHS is not congruent $1 \mod (p^{n+2})$. □

Corollary 10.14. 1. *For* $p > 2$ *the subgroup* G_1 *is cyclic, indeed the set* $G_1 \setminus G_2$ *consists of the possible primitive roots.*

2. *For* $p = 2$ *the subgroup* G_2 *is cyclic, and the set* $G_2 \setminus G_3$ *consists of the possible primitive roots for the subgroup* G_2. *In particular* $\overline{5}$ *satisfies*

$$G_2 = \overline{5}^{\mathbb{Z}}.$$

Finally we obtain:

Theorem 10.15. *1. For a prime $p > 2$ and $r \geq 1$ the group of units $\mathbb{Z}_{p^r}^*$ of $\mathbb{Z}_{p^n}$ is cyclic. Indeed, if $r \geq 2$ and $\sigma : \mathbb{Z}_{p^r} \longrightarrow \mathbb{Z}_{p^2}$ denotes the natural map, we have*

$$\mathbb{Z}_{p^r}^* = a^{\mathbb{Z}} \Longleftrightarrow \mathbb{Z}_{p^2}^* = \sigma(a)^{\mathbb{Z}}.$$

2. For $r \geq 1$ we have

$$\mathbb{Z}_{2^r}^* = \pm\overline{5}^{\mathbb{Z}}.$$

Proof. Let $p > 2$. We consider the diagram

$$\begin{array}{ccccc} & & \mathbb{Z}_{p^n} & \xrightarrow{\varrho} & \mathbb{Z}_p \\ & & \cup & & \cup \\ \varrho^{-1}(\mathbb{Z}_p^*) & = & \mathbb{Z}_{p^n}^* & \xrightarrow{\varrho} & \mathbb{Z}_p^* \\ \cup & & \cup & & \cup \\ \varrho^{-1}(\overline{1}) & = & G_1 & \longrightarrow & \{\overline{1}\} \end{array},$$

where the map ϱ is the natural map. Now, given a primitive root $b \in \mathbb{Z}_p^*$, there is a unique element $c \in \mathbb{Z}_{p^n}^*$ lying above b with $\mathrm{ord}(c) = p-1 = \mathrm{ord}(b)$. Uniqueness follows from the fact that any other element above b is of the form $c(\overline{1} + p\overline{t})$ and thus has order $p-1$ iff the second factor (whose order is a prime power p^s) equals $\overline{1}$.

In order to see that there is such an element, take any $d \in \mathbb{Z}_{p^n}$ above $b \in \mathbb{Z}_p$. Its order is a multiple of $p-1$ (since $\mathrm{ord}(b) = p-1$) and divides the order $p^{n-1}(p-1)$ of the group of units $\mathbb{Z}_{p^n}$, hence of the form $p^s(p-1)$. Now we choose $c := d^{p^s}$. Indeed c lies above $b^{p^s} = b$.

Now it follows from Prop.10.13 that $a = c(\overline{1}+p\overline{t}) \in \mathbb{Z}_{p^n}^*$ is a primitive root iff p does not divide t - and the same criterion applies to $\sigma(a) = \sigma(c)(\overline{1}+p\overline{t})$.

Finally the second part follows also immediately from Prop.10.13 since

$$\mathbb{Z}_{2^n}^* = \overline{1} + 2\mathbb{Z}_{2^n} = (\overline{1} + 4\mathbb{Z}_{2^n}) \cup (-\overline{1} + 4\mathbb{Z}_{2^n})$$

and

$$-\overline{1} + 4\mathbb{Z}_{2^n} = -(\overline{1} + 4\mathbb{Z}_{2^n}).$$

□

Example 10.16. We determine a primitive root for $\mathbb{Z}_{343}^* = \mathbb{Z}_{7^3}^*$.

1. First we look for a primitive root in $\mathbb{Z}_7^*$. The possible orders of a unit $\neq \overline{1}$ are $2, 3, 6$. The residue class $\overline{2}$ is not because of $\overline{2}^3 = \overline{1}$, but $\overline{3}$ is, since both $\overline{3}^2 = \overline{9} \neq \overline{1}$ and $\overline{3}^3 = -\overline{1} \neq \overline{1}$.

2. Let us now take $\overline{3}$ as a candidate for a primitive root in $\mathbb{Z}_{49}$. We have to check that $\overline{3}^6 \neq \overline{1}$ - indeed

$$\mathbb{Z}_{49} \ni \overline{3}^6 = \overline{81} \cdot \overline{9} = -\overline{17} \cdot \overline{9} = -\overline{153} = -\overline{6} \neq \overline{1}.$$

   Now $\overline{3} \in \mathbb{Z}_{343}^*$ is a primitive root as well.

Finally we come back to our original problem, the computation of ℓ-th roots:

Theorem 10.17. *Let R be a ring with finite cyclic group of units, $|R^*| = q$, and $a \in R^*, \ell \in \mathbb{N}_{>0}$. The following conditions are equivalent:*

1. *There is an element $b \in R^*$ with $a = b^\ell$.*
2. $$a^k = 1 \text{ holds for } k := \frac{q}{\gcd(\ell, q)}.$$

Proof. "$\Longrightarrow$": We have

$$\ell k = \frac{\ell q}{\gcd(\ell, q)} = q \cdot \frac{\ell}{\gcd(\ell, q)},$$

hence $a = b^\ell$ gives

$$a^k = b^{\ell k} = (b^q)^{\cdots} = 1,$$

since $b^q = 1$ according to Th.7.9.
"$\Longleftarrow$": Assume $R^* = c^{\mathbb{Z}}$ and write $a = c^n$. The condition $a^k = 1$ is equivalent to

$$q | \frac{nq}{\gcd(\ell, q)} \Longleftrightarrow \gcd(\ell, q) | n,$$

the condition equivalent to the solvability of the congruence

$$\ell t \equiv n \mod (q).$$

Take now $a = c^t$ with a solution of that congruence. □

Corollary 10.18. *The residue class $-\overline{1} \in \mathbb{Z}_p^*$ is a square iff $p \equiv 1 \mod (4)$.*

Proof. Take $q = p - 1, \ell = 2$. Then $-\overline{1} \in \mathbb{Z}_p^*$ is a square iff $2 | \frac{p-1}{2}$ iff $p \equiv 1 \mod (4)$. □

11 Quadratic reciprocity

For a quadratic polynomial $f \in \mathbb{Z}[X]$ there is a straight forward procedure to check whether it has zeros in $\mathbb{Z}_p$ or not. The case $p = 2$ is easy: If we denote $\overline{f} \in \mathbb{Z}_2[X]$ the induced polynomial, we have no solution for $\overline{f} = X^2 + X + 1$, while in the remaining cases they are obvious. If p is odd, we may write

$$\overline{f} = X^2 + 2\overline{b}X + \overline{c} = (X + \overline{b})^2 - (\overline{b}^2 - \overline{c})$$

and thus are left with the question, whether or not $\overline{b}^2 - \overline{c}$ is a square.

Remark 11.1. A residue class $\overline{a} \in \mathbb{Z}_p$ is a square iff $\overline{a}^{(p-1)/2} = -\overline{1}$. This is an immediate consequence of Th.10.17.

But to compute the $(p-1)/2$-th power of a residue class might require a lot of time. Indeed, there is a considerably simpler algorithm. In order to formulate it we need the following notation:

Definition 11.2. Denote $P_{>2}$ the set of all odd primes. The Legendre symbol is the map

$$\left(\frac{-}{-}\right) : \mathbb{Z} \times P_{>2} \longrightarrow \{0, \pm 1\}, (a, p) \mapsto \left(\frac{a}{p}\right),$$

where

$$\left(\frac{a}{p}\right) := \begin{cases} 1 & , \text{ if } \overline{a} \in \mathbb{Z}_p^* \text{ is a square} \\ -1 & , \text{ if } \overline{a} \in \mathbb{Z}_p^* \text{ is not a square} \\ 0 & , \text{ if } p|a \text{ in } \mathbb{Z}_p \end{cases} .$$

Remark 11.3. 1. We define

$$\left(\frac{\overline{k}}{p}\right) := \left(\frac{k}{p}\right)$$

for $\overline{k} \in \mathbb{Z}_p$.

2. If $c \in \mathbb{Z}_p^*$ is a generator of the (cyclic) multiplicative group $\mathbb{Z}_p^*$, i.e., $\mathbb{Z}_p^* = c^{\mathbb{Z}}$, we have

$$\left(\frac{c^\nu}{p}\right) = (-1)^\nu.$$

3. If K is a field s.th. $1+1 \neq 0$, we may regard the Legendre symbol as a function taking values in K, since then the elements $0, 1, -1 \in K$ are pairwise different.

4. With that convention we may write

$$\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \in \mathbb{Z}_p$$

for all $a \in \mathbb{Z}_p$.

5. The Legendre symbol is multiplicative in the "numerator":

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

for $a, b \in \mathbb{Z}$ as well as $a, b \in \mathbb{Z}_p$. In particular it is sufficient to compute the Legendre symbol for a being a prime as well.

Theorem 11.4. *Let $p \in P_{>2}$ be an odd prime.*

1. *We have*

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & , \quad \text{if } p \equiv \pm 1 \mod (8) \\ -1 & , \quad \text{if } p \equiv \pm 3 \mod (8) \end{cases},$$

or, more briefly

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}.$$

2. *Furthermore*

$$\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}} = \begin{cases} 1 & , \quad \text{if } p \equiv 1 \mod (4) \\ -1 & , \quad \text{if } p \equiv 3 \mod (4) \end{cases}.$$

3. *The law of quadratic reciprocity: For a prime $q \in P_{>2}$ different from p we have*

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{p}{q}\right).$$

With other words

$$\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right)$$

if $p \equiv q \equiv 3 \mod (4)$, *and*

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$$

otherwise.

Example 11.5. We compute

$$\left(\frac{42}{61}\right) = \left(\frac{2}{61}\right)\left(\frac{3}{61}\right)\left(\frac{7}{61}\right) = (-1)\cdot\left(\frac{61}{3}\right)\left(\frac{61}{7}\right)$$

$$= (-1)\cdot\left(\frac{1}{3}\right)\left(\frac{-2}{7}\right) = -\left(\frac{-1}{7}\right)\left(\frac{2}{7}\right) = -(-1)\cdot 1 = 1.$$

In the proof of Th.11.4 we take for granted the existence of a field $K \supset \mathbb{Z}_p$, such that there is an element $\zeta \in K^*$ of order 8 resp. an element $\eta \in K^*$ of order q. In problem 2.4 we have constructed a field $K \supset \mathbb{Z}_p$ of order p^2. Hence $\zeta := a^{(p^2-1)/8}$, where $a \in K^*$ is a primitive root, is a possible choice. For the second case see the explanations at the end of this section. For the proof itself we need only to know that the *Frobenius map*

$$K \longrightarrow K, x \mapsto x^p,$$

is compatible with addition and multiplication:

$$(x+y)^p = x^p + y^p, (xy)^p = x^p y^p,$$

and that $\mathbb{Z}_p \subset K$ is its fixed point set:

$$\mathbb{Z}_p = \{x \in K; x^p = x\}.$$

The inclusion "$\subset$" follows from the little Fermat theorem, while the right hand side is the zero set in K of the polynomial $X^p - X \in K[X]$, hence has at most p elements.

Proof of Th.11.4. 1.) For $\beta := \zeta + \zeta^{-1}$ we have

$$\beta^2 = \zeta^2 + 2 + \zeta^{-2} = \zeta^2 + 2 - \zeta^2 = 2,$$

since $\zeta^4 = -1$. As a consequence 2 is a square in $\mathbb{Z}_p$ iff $\beta \in \mathbb{Z}_p$ iff $\beta = \beta^p$. Now

$$\beta^p = \zeta^p + \zeta^{-p} = \zeta^r + \zeta^{-r},$$

where $p = 8d + r$. Thus for $r = \pm 1$ we find $\beta^p = \beta$, while $r = \pm 3$ yields $\beta^p = -\beta$.
2.): Follows from Rem.11.3.4.
3.): In order to simplify notation we write

$$\chi(a) := \left(\frac{a}{q}\right).$$

We need the following auxiliary lemma:

Lemma 11.6. *Let K be a field containing an element $\eta \in K \setminus \{1\}$ with $\eta^q = 1$. Then the square of the* Gauß' sum

$$\gamma := \sum_{i=0}^{q-1} \chi(i)\eta^i \in K$$

satisfies

$$\gamma^2 = \chi(-1)q \in K.$$

Remark 11.7. We add without (the demanding) proof: If $K = \mathbb{C}, \eta = e^{2\pi i/q}$, we have

$$\gamma = \begin{cases} \sqrt{q} & , \text{ if } q \equiv 1 \mod (4) \\ i\sqrt{q} & , \text{ if } q \equiv 3 \mod (4) \end{cases}.$$

Let us first finish the proof of the Th.11.4: We take a field $K \supset \mathbb{Z}_p$ containing an element $\eta \neq 1$ with $\eta^q = 1$ and apply the Frobenius map $K \longrightarrow K, x \mapsto x^p$, to our Gauß' sum:

$$\gamma^p = \sum_{i=0}^{q-1} \chi(i)^p \eta^{ip} = \sum_{i=0}^{q-1} \chi(i)\eta^{ip}$$

$$= \sum_{i=0}^{q-1} \chi(p^2 i)\eta^{ip} = \chi(p) \sum_{i=0}^{q-1} \chi(ip)\eta^{ip} = \chi(p)\gamma,$$

using $\chi(i)^p = \chi(i) = \chi(p^2 i)$ and the fact that $\mathbb{Z}_q \longrightarrow \mathbb{Z}_q, i \mapsto ip$, is a bijection. Since $\gamma \neq 0$, we may conclude

$$\left(\frac{p}{q}\right) = \gamma^{p-1} = (\gamma^2)^{\frac{p-1}{2}} = (\chi(-1)q)^{\frac{p-1}{2}}$$

$$= ((-1)^{\frac{q-1}{2}})^{\frac{p-1}{2}} q^{\frac{p-1}{2}} = (-1)^{\frac{q-1}{2}\frac{p-1}{2}} \left(\frac{q}{p}\right),$$

where we have used Remark 11.3.1 with respect to $a = -1$ and the prime q as well as $a = q$ and the prime p. □

Proof of the lemma. First of all we may understand the exponents of η as well as the argument of χ as elements in $\mathbb{Z}_q$. Using that convention we remark the following identities:

$$\sum_{i \in \mathbb{Z}_q} \chi(i) = 0, \ \sum_{i \in \mathbb{Z}_q} \eta^i = 0.$$

The first one follows from the fact that $\chi(0) = 0$ and there are $\frac{q-1}{2}$ quadratic residues as well as $\frac{q-1}{2}$ quadratic non-residues in $\mathbb{Z}_q$, for the second one notes that the sum not changing when multiplied by η has to be $= 0$.

Now

$$\begin{aligned} \gamma^2 &= \sum_{(i,j) \in (\mathbb{Z}_q)^2} \chi(i)\chi(j)\eta^{i+j} \\ &= \sum_{\ell \in \mathbb{Z}_q} \left(\sum_{i+j=\ell} \chi(i)\chi(j) \right) \eta^\ell. \end{aligned}$$

Let us now compute the inner sums. The case $\ell = 0$ yields

$$\sum_{i+j=0} \chi(i)\chi(j) = \sum_{i \in \mathbb{Z}_q} \chi(i)\chi(-i) = \sum_{i \in \mathbb{Z}_q} \chi(-i^2) = \chi(-1)(q-1),$$

since $\chi(-i^2) = \chi(i^2)\chi(-1) = \chi(-1)$ for $i \neq 0$, while $\chi(-0^2) = 0$.

Finally we treat the case $\ell \neq 0$ and get

$$\begin{aligned}\sum_{i+j=\ell} \chi(i)\chi(j) &= \sum_{i\in\mathbb{Z}_q^*} \chi(i)\chi(\ell-i) \\ &= \sum_{i\in\mathbb{Z}_q^*} \chi(i^{-1})\chi(\ell-i) \\ &= \sum_{i\in\mathbb{Z}_q^*} \chi(\ell i^{-1}-1) \\ &= \sum_{i\in\mathbb{Z}_q\setminus\{-1\}} \chi(i) = -\chi(-1),\end{aligned}$$

since $\chi(0)=0, \chi(i^{-1})=\chi(i)$ and $\{\ell i^{-1}-1; i\in\mathbb{Z}_q^*\} = \mathbb{Z}_q \setminus \{-1\}$. Hence

$$\gamma^2 = \chi(-1)(q-1) - \chi(-1)\sum_{\ell\in\mathbb{Z}_q^*} \eta^\ell$$

$$= \chi(-1)(q-1) - \chi(-1)\cdot(-1) = \chi(-1)q.$$

□

Construction of extension fields $K \supset \mathbb{Z}_p$ admitting a primitive q-th root of unity $\eta \in K$: If $q|(p-1)$ we may take $K=\mathbb{Z}_p$ and $\eta = c^{(p-1)/q}$, where $c\in\mathbb{Z}_p$ is a primitive root. Otherwise there is no q-th root of unity in $\mathbb{Z}_p$ except 1, the situation we have to deal with.

We consider a finite dimensional $\mathbb{Z}_p$-vector space V and realize

$$K \subset \mathrm{End}(V)$$

as a subring of the endomorphism ring of V. Namely we take some linear operator (or endomorphism) $A: V \longrightarrow V$ and consider the set

$$K := \mathbb{Z}_p[A] := \left\{ f(A) \in \mathrm{End}(V); f\in\mathbb{Z}_p[X] \right\}$$

of all polynomials

$$f(A) := \lambda_m A^m + ... + \lambda_1 A + \lambda_0 E, f = \sum_{\mu=0}^{m} a_\mu X^\mu \in \mathbb{Z}_p[X]$$

in A. Here E denotes the identity operator

$$E := \mathrm{id}_V.$$

Since $A^r A^s = A^{r+s} = A^{s+r} = A^s A^r$, the ring $\mathbb{Z}_p[A]$ is even commutative. Furthermore we want that it is a field and that A plays the rôle of η, i.e. that

$$A^q = E, A \neq E.$$

Since our first choice does not match our expectations we write W instead of V and B instead of A. We choose

$$W := (\mathbb{Z}_p)^q$$

with the cyclic shift operator $B : W \longrightarrow W$ satisfying

$$B(e_i) = e_{i+1}, i = 1, ..., q-1,\ B(e_q) = e_1.$$

Here $e_1, ..., e_q \in (\mathbb{Z}_p)^q$ denotes the standard basis. Unfortunately $\mathbb{Z}_p[B]$ is not yet a field: We have to replace W with a minimal B-invariant subspace $V \subset W$, s.th.

$$A := B|_V \neq \mathrm{id}_V$$

and may apply the below proposition 11.8. In particular, we obtain a surjective, but in general not injective ring homomorphism

$$\mathbb{Z}_p[B] \longrightarrow \mathbb{Z}_p[A], C \mapsto C|_V.$$

It remains to show that the condition $B|_V \neq \mathrm{id}_V$ can be realized: First of all we note that

$$\mathbb{Z}_p(1, ..., 1)$$

is the eigenspace of B for the eigenvalue 1. It admits the complementary B-invariant subspace

$$W_0 := \{(x_1, ..., x_n); x_1 + ... + x_n = 0\},$$

where we need $q \neq p$ in order to exclude the possibility $(1, ..., 1) \in W_0$. Now we choose $V \subset W_0$ as a minimal nontrivial B-invariant subspace. Since $Bu \neq u$ for all $u \in W_0 \setminus \{0\}$, we are done. We remark that we have $\dim V > 1$, since otherwise it would consist of eigenvectors of B belonging to an eigenvalue $\lambda \in \mathbb{Z}_p \setminus \{\overline{1}\}$ satisfying $\lambda^q \neq \overline{1}$.

Proposition 11.8. *Let F be a field and V a finite dimensional F-vector space. If the linear operator $A : V \longrightarrow V$ does not admit nontrivial proper invariant subspaces, the ring*

$$K := F[A] \subset \mathrm{End}(V)$$

is a field.

Example 11.9. If $V = \mathbb{R}^2$ and $A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is the counterclockwise rotation with an angle of 90 degrees, then we obtain $\mathbb{R}[A] \cong \mathbb{C}$.

Proof of Prop. 11.8. We have to show that any operator $f(A) \neq 0$ is invertible and $f(A)^{-1} \in F[A]$. Indeed, its kernel $\ker(f(A)) \subsetneqq V$ is a proper A-invariant subspace, since

$$f(A)u = 0 \Longrightarrow f(A)Au = Af(A)u = A(0) = 0,$$

hence trivial. So $f(A)$ is invertible. In particular

$$F[A] \longrightarrow F[A], C \mapsto f(A)C,$$

is an injective endomorphism of the finite dimensional vector space $F[A]$, hence an isomorphism and $E \in f(A)F[A]$. □

12 Arithmetic functions

Definition 12.1. An arithmetic function is a complex-valued function

$$f : \mathbb{N}_{>0} \longrightarrow \mathbb{C}.$$

It is called multiplicative iff $f(1) = 1$ and

$$f(ab) = f(a)f(b),$$

whenever $a, b \in \mathbb{N}_{>0}$ are relatively prime.

Remark 12.2. A multiplicative arithmetic function f is uniquely determined by its values $f(p^n)$ at prime power arguments, and those values can be prescribed arbitrarily.

Example 12.3. We list here some multiplicative arithmetic functions:

1. the power functions $p_k(n) = n^k$,

2. the Dirac function

$$\delta(n) = \begin{cases} 1 & , \text{ if } n = 1 \\ 0 & , \text{ if } n > 1 \end{cases},$$

3. Eulers φ-function,

4. the sum

$$\sigma_k(n) := \sum_{d|n} d^k$$

of the k-th powers of all positive divisors of $n \in \mathbb{N}$, in particular

   (a) $\tau(n) := \sigma_0(n)$ is the number of positive divisors of n and

   (b) $\sigma(n) := \sigma_1(n)$ their sum,

5. the Möbius function

$$\mu(n) := \begin{cases} (-1)^{\omega(n)} & , \text{ if } n \text{ is square free} \\ 0 & , \text{ otherwise} \end{cases},$$

where $\omega(n)$ denotes the number of different prime divisors of n (i.e. without multiplicities!).

The following remark is optional, it is intended for readers familiar with complex analysis:

Remark 12.4. Given an arithmetic function $f : \mathbb{N}_{>0} \longrightarrow \mathbb{C}$. one considers the *Dirichlet series*

$$\Phi_f(z) := \sum_{n=1}^{\infty} \frac{f(n)}{n^z}$$

with the entire functions

$$n^z := e^{\ln(n)z}.$$

Since

$$|n^{-z}| = e^{-\ln(n)\Re(z)}$$

is a decreasing function of $\Re(z)$, we easily deduce: If

$$\sum_{n=1}^{\infty} \frac{|f(n)|}{n^a} < \infty$$

holds for some $a \in \mathbb{R}$, then $\Phi_f(z)$ converges uniformly in the right half plane $R_{\geq a} := \{z \in \mathbb{C}; \Re(z) \geq a\}$, hence defines there a continuous function, holomorphic in $R_{>a}$. But we can do even better: If for some $z_0 \in R_{\geq 0}$ the series $\Phi_f(z_0)$ converges (not necessarily absolutely), then Φ_f converges uniformly in any angular segment

$$z_0 + \mathbb{R}_{\geq 0} \exp\left(i\left[\frac{\pi}{2} - \varepsilon, -\frac{\pi}{2} + \varepsilon\right]\right),\ \varepsilon > 0,$$

with vertex z_0. In particular it defines a holomorphic function in $R_{>\Re(z_0)}$.

So there is a unique minimal number $a \in \mathbb{R} \cup \{\pm\infty\}$, such that Φ_f defines a holomorphic function in $R_{>a}$, called the abscissa of convergence of Φ_f. Let us note that this does not imply absolute convergence in $R_{>a}$. Instead one can assure absolute convergence a priori only in $R_{>a+1}$, e.g. for $f(n) = (-1)^n$ one has $a = 0$ (take $z_0 = q \in \mathbb{R}_{>0}$ and note that n^{-q} is a strictly decreasing sequence tending to 0), while absolute convergence requires $R_{>1}$. The most famous example is the series

$$\Phi_1 : R_{>1} \longrightarrow \mathbb{C}$$

belonging to $f \equiv 1$. It provides a partial definition of *Riemann's ζ-function*, see the next (optional) section. It has abscissa of convergence $a = 1$ and it converges even absolutely in $R_{>1}$. The latter is also true for Φ_μ, indeed its abscissa a of convergence satisfies

$$\frac{1}{2} \leq a \leq 1,$$

but its exact value is not known.

For a multiplicative arithmetic function f we may factorize $\Phi_f(z)$ as an infinite product

$$\Phi_f(z) = \prod_{p \in P} F_p(z)$$

with

$$F_p(z) := \sum_{\nu=0}^{\infty} \frac{f(p^\nu)}{p^{\nu z}},$$

valid in the right half plane, where Φ_f converges absolutely. E.g. for $f = \mu$ we have

$$F_p(z) = 1 - p^{-z}$$

and for $f \equiv 1$ we obtain

$$F_p(z) = \sum_{\nu=0}^{\infty} \frac{1}{p^{\nu z}} = \frac{1}{1-p^{-z}}.$$

We obtain thus factorizations

$$\Phi_\mu(z) = \prod_{p \in P} \left(1 - p^{-z}\right), \ z \in R_{>1},$$

and

$$\Phi_1(z) = \prod_{p \in P} \left(\frac{1}{1-p^{-z}}\right), \ z \in R_{>1}. \tag{13.1}$$

Obviously we have

$$\Phi_f + \Phi_g = \Phi_{f+g}.$$

The *convolution* $f * g$ of two arithmetic functions f, g is defined in order to make the equality

$$\Phi_f \cdot \Phi_g = \Phi_{f*g}$$

hold in the open half plane, where both Φ_f and Φ_g converge absolutely.

Definition 12.5. The *convolution* $f * g$ of two arithmetic functions f, g is defined by

$$(f * g)(n) := \sum_{d|n} f(d) g\left(\frac{n}{d}\right) = \sum_{k\ell=n} f(k) g(\ell).$$

Remark 12.6.

1. $f * g = g * f$
2. $f * (g * h) = (f * g) * h$
3. $\delta * f = f$
4. If $f, g : \mathbb{N}_{>0} \longrightarrow \mathbb{C}$ are multiplicative, so is $f * g$.

Proof. For $\gcd(a, b) = 1$ we find

$$(f * g)(ab) = \sum_{d|ab} f(d) g\left(\frac{ab}{d}\right) = \sum_{k|a, \ell|b} f(k\ell) g\left(\frac{a}{k} \cdot \frac{b}{\ell}\right)$$

$$= \sum_{k|a,\ell|b} f(k)f(\ell)g\left(\frac{a}{k}\right)g\left(\frac{b}{\ell}\right)$$

$$= \sum_{k|a} f(k)g\left(\frac{a}{k}\right) \cdot \sum_{\ell|b} f(\ell)g\left(\frac{b}{\ell}\right) = (f*g)(a)\cdot(f*g)(b).$$

□

5. The following identities for multiplicative arithmetic functions hold, since they hold for prime power arguments.
   (a) $1 * \mu = \delta$.
   (b) $p_k * 1 = \sigma_k$.
   (c) $\varphi * 1 = p_1$. (Remember that $p_1(n) = n$.)

Definition 12.7.

$$\hat{f} := f * 1, \tilde{f} := f * \mu.$$

Proposition 12.8. *1. If f is multiplicative, so are $\hat{f}$ and $\tilde{f}$.*

2. The transformations $f \mapsto \hat{f}$ and $f \mapsto \tilde{f}$ are inverse one to the other.

Proof.

$$(f * 1) * \mu = f * (1 * \mu) = f * \delta = f$$

and

$$(f * \mu) * 1 = f * (\mu * 1) = f * \delta = f.$$

□

Example 12.9. 1. $\hat{p_k} = \sigma_k$.

2. $\hat{\varphi} = p_1 = \mathrm{id}_{\mathbb{N}_{>0}}$.

3. $\hat{\mu} = \delta$.

Remark 12.10. Obviously the set of all arithmetic functions together with argumentwise addition and the convolution as multiplication becomes a commutative ring. We have interpreted it using formal Dirichlet series, but we can instead also look at formal power series in countably many indeterminates X_p, indexed for convenience by prime numbers $p \in P$. The elements in the ring

$$\mathbb{C}[[X_p, p \in P]]$$

are formal sums with complex coefficients of finite monomials

$$X^\alpha := \prod_{p\in P}(X_p)^{\alpha(p)},$$

where the exponent is a function $\alpha : P \longrightarrow \mathbb{N}$ with finite support $\alpha^{-1}(\mathbb{N}_{>0})$. An arithmetic function $f : \mathbb{N}_{>0} \longrightarrow \mathbb{C}$ corresponds to the series

$$\sum_\alpha f(\mathbf{p}^\alpha)X^\alpha,$$

where

$$\mathbf{p}^\alpha := \prod_{p\in P} p^{\alpha(p)} \in \mathbb{N}_{>0}.$$

13 Riemann's zeta function (optional)

Riemann's ζ-function

$$\zeta : \mathbb{C} \setminus \{1\} \longrightarrow \mathbb{C}$$

is a holomorphic extension of the Dirichlet series $\Phi_1 : R_{>1} \longrightarrow \mathbb{C}$, i.e.

$$\zeta|_{R_{>1}} = \Phi_1.$$

Before we discuss how to find an expression for ζ outside the right half plane $R_{>1}$, we note that ζ is unique if it exists and that $\zeta(z) \neq 0$ for all $z \in R_{>1}$. This is obvious for $z \in \mathbb{R}_{>1}$ and follows in general from the factorization (13.1) and the fact that the locally uniform limit $f : G \longrightarrow \mathbb{C}$ of a sequence of nowhere vanishing holomorphic functions $f_n : G \longrightarrow \mathbb{C}$ on a domain G is either $f \equiv 0$ or has itself no zeros.

The expressions for ζ involve parameter dependent integrals, either an improper "real" integral or a path integral in the complex plane:
1.) For every $n \in \mathbb{N}$ there is an expression of the first type on the right half plane $R_{>-n}$. We discuss only the case $n = 0$. Indeed, on $R_{>1}$ we may write

14.1
$$\Phi_1(z) = \frac{z}{z-1} - z\int_1^\infty \frac{t-[t]}{t^{z+1}}dt$$

with the Gauß bracket $[t] = \max \mathbb{Z}_{\leq t}$ and

$$t^z := e^{z\ln t}.$$

Now the integral defines a holomorphic function even in $R_{>0}$ and thus the RHS of 14.1 a meromorphic extension of Φ_1 to $R_{>0}$. The proof is an easy exercise: Write

$$\int_1^N \frac{[t]dt}{t^{z+1}} = \sum_{n=1}^{N} \int_n^N \frac{dt}{t^{z+1}}.$$

2.) The second approach requires complex integration. We start with the entire function

$$I(z) := \frac{1}{2\pi i} \int_C \frac{(-\eta)^z}{e^\eta - 1} \frac{d\eta}{\eta} \quad , z \in \mathbb{C},$$

where for fixed $z \in \mathbb{C}$ the integrand is regarded as a holomorphic function in η on $\mathbb{C} \setminus (\mathbb{R}_{\geq 0} \cup 2\pi i\mathbb{Z})$ in the following way: Take

$$(-\eta)^z := e^{\log(-\eta)z}$$

with the principal branch $\log : \mathbb{C} \setminus \mathbb{R}_{\leq 0} \longrightarrow \mathbb{C}$ of the logarithm, such that there are upper and lower limits along $\mathbb{R}_{\geq 0}$. The path of integration is

$$C = (\infty, \varepsilon] \cup \partial D_\varepsilon(0) \cup [\varepsilon, \infty)$$

with $\varepsilon < 2\pi$. Along the first part $(\infty, \varepsilon]$ we take the upper limit

$$(-\eta)^z := e^{(\ln(|\eta|) - \pi i)z},$$

while along the third part $[\varepsilon, \infty)$ we take the lower limit

$$(-\eta)^z := e^{(\ln(|\eta|) + \pi i)z}.$$

Since the integrand is holomorphic in $D_{2\pi}(0) \setminus [0, 2\pi)$, we see that $I(z)$ does not depend on the choice of ε.

Note that I is an entire function and that for $z \in R_{>1}$ the circle integral tends to 0 for $\varepsilon \to 0$, hence

$$\begin{aligned} I(z) &= \frac{1}{2\pi i} \cdot \lim_{\varepsilon \to 0} \int_\varepsilon^\infty \frac{\exp(z(\ln(\eta) + \pi i)) - \exp(z(\ln(\eta) - \pi i))}{(e^\eta - 1)\eta} d\eta \\ &= \frac{\sin(\pi z)}{\pi} \lim_{\varepsilon \to 0} \int_\varepsilon^\infty \frac{\eta^{z-1}}{e^\eta - 1} d\eta \\ &= \frac{\sin(\pi z)}{\pi} \lim_{\varepsilon \to 0} \int_\varepsilon^\infty \left(\sum_{n=1}^{\infty} \eta^{z-1} e^{-n\eta} \right) d\eta \end{aligned}$$

$$= \frac{\sin(\pi z)}{\pi} \lim_{\varepsilon \to 0} \sum_{n=1}^{\infty} \int_{\varepsilon}^{\infty} \eta^{z-1} e^{-n\eta} d\eta$$

$$= \frac{\sin(\pi z)}{\pi} \lim_{\varepsilon \to 0} \sum_{n=1}^{\infty} \frac{1}{n^z} \cdot \int_{\varepsilon/n}^{\infty} t^{z-1} e^{-t} dt$$

14.2
$$= \frac{\sin(\pi z)}{\pi} \cdot \Gamma(z) \cdot \Phi_1(z)$$

with

$$\Gamma(z) := \int_0^{\infty} t^{z-1} e^{-t} dt, z \in R_{>0},$$

an expression defining a holomorphic function $\Gamma : R_{>0} \longrightarrow \mathbb{C}$. One easily checks $\Gamma(1) = 1$, and partial integration leads to the functional equation

$$\Gamma(z+1) = z\Gamma(z).$$

It gives a successive meromorphic extension

$$\Gamma(z) = \frac{\Gamma(z+n)}{z(z+1) \cdot ... \cdot (z+(n-1))}$$

to any right half plane $R_{>-n}$. So eventually Γ is a meromorphic function on the entire plane with simple poles in $-\mathbb{N}$. Now it would be good to have a formula for Γ valid simultaneously for all $z \in \mathbb{C}$. Unfortunately the products in the denominator do not form a convergent sequence. Instead we normalize the constant terms and add to each factor an exponential in order to guarantee convergence of the infinite product

$$\Delta(z) := z \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-z/n}.$$

The remaining numerator is an exponential as well: We finally get

$$\Gamma(z) = \frac{e^{-\gamma z}}{\Delta(z)}$$

with the Euler-Mascheroni constant

$$\gamma := \lim_{m \to \infty} \left(\sum_{n=1}^{m} \frac{1}{n} - \ln(m+1) \right) = 0,57721....$$

needed in order to compensate the exponential factors in the infinite product.

The auxiliary function $\Delta(z)$ has the points in $-\mathbb{N}$ as simple zeros, while $\Delta(1-z)$ has its zeros in $\mathbb{N}_{>0}$; so $\Delta(z)\Delta(1-z)$ has the integers as its set of (simple) zeros. The same is true for the function

$$\sin(\pi z) = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right),$$

and indeed

$$\Delta(z)\Delta(1-z) = \pi^{-1} e^{-\gamma} \sin(\pi z),$$

or equivalently

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}. \tag{14.3}$$

Solving 14.3 for $\Gamma(z)$ and substituting the RHS into the expression 14.2 for $I(z), z \in R_{>1}$, leads us to the following definition of Riemann's ζ-function

$$\zeta(z) := \Gamma(1-z) I(z). \tag{14.4}$$

Note that at integer arguments $z = \ell \in \mathbb{Z}$ the integrals in the expression for $I(z)$ over the straight parts of C cancel and thus

$$I(\ell) = (-1)^{\ell} \cdot \mathrm{Res}_0 \left(\frac{\eta^{\ell-1}}{e^{\eta} - 1}\right). \tag{14.5}$$

In particular $I(\ell) = 0$ for $\ell \geq 2$ (as it should be, since ζ is holomorphic in $R_{>1}$ and $\Gamma(1-z)$ has poles at the points in $\mathbb{N}_{>1}$) and for $\ell = -2k, k \in \mathbb{N}_{>0}$, as well. The latter fact follows, since

$$\frac{\eta^{-2k-1}}{e^{\eta} - 1} = \eta^{-2k-2} \left(-\frac{\eta}{2} + g(\eta)\right)$$

with an even function $g(\eta)$. Furthermore $\zeta(0) = -\frac{1}{2}$ because of $\Gamma(1) = 1$.

Finally we try to evaluate $I(z)$ for $\Re(z) < 0$. In that case we have

$$\lim_{n\to\infty} \frac{1}{2\pi i} \int_{\partial D_{(2n+1)\pi}(0)} \frac{(-\eta)^z}{e^{\eta} - 1} \frac{d\eta}{\eta} = 0,$$

and we may regard the integral over C as the limit of the integrals over the loops

$$C_n := [(2n+1)\pi, \varepsilon] \cup \partial D_{\varepsilon}(0) \cup [\varepsilon, (2n+1)\pi] \cup \partial D_{(2n+1)\pi}(0)^{-1},$$

to be understood as the negatively oriented boundary ∂G^{-1} of the domain $G := \{\eta; \varepsilon < |\eta| < (2n+1)\pi\} \setminus (\varepsilon, (2n+1)\pi)$. With other words, for $\Re(z) < 0$ we have

$$I(z) = \lim_{n\to\infty} \frac{1}{2\pi i} \int_{C_n} \frac{(-\eta)^z}{e^\eta - 1} \frac{d\eta}{\eta},$$

and, according to the residue theorem

$$\frac{1}{2\pi i} \int_{C_n} \frac{(-\eta)^z}{e^\eta - 1} \frac{d\eta}{\eta} = - \sum_{0<|\nu|<n+\frac{1}{2}} \mathrm{Res}_{2\pi i\nu} \left(\frac{(-\eta)^z}{\eta(e^\eta - 1)} \right)$$

$$= - \sum_{0<|\nu|<n+\frac{1}{2}} \frac{(-2\pi i\nu)^z}{2\pi i\nu} = 2^z \pi^{z-1} \sin\left(\frac{\pi z}{2}\right) \sum_{\nu=1}^{n} \frac{1}{\nu^{1-z}}.$$

So we obtain

$$I(z) = 2^z \pi^{z-1} \sin\left(\frac{\pi z}{2}\right) \Phi_1(1-z), z \in L_{<0}.$$

If we multiply with $\Gamma(1-z)$, we get the following functional equation for the ζ-function:

$$\zeta(z) = 2\Gamma(1-z)(2\pi)^{z-1} \sin\left(\frac{\pi z}{2}\right) \zeta(1-z).$$

It follows that there are no further zeros of ζ in $L_{<0} \cup R_{>1}$ than the points in $-2\mathbb{N}_{>0}$, also called the *trivial zeros* of ζ.

On the other hand we obtain new interesting information about special values of Φ_1. Taking $z = 1 - 2k, k \in \mathbb{N}$, and using the above residue formula 14.3, we obtain explicit formulae for the infinite series

$$\Phi_1(2k) = \sum_{n=1}^{\infty} \frac{1}{n^{2k}}.$$

In the previous section we have established a factorization

$$\zeta(z) = \prod_{p\in P} \left(\frac{1}{1 - p^{-z}} \right), z \in R_{>1}.$$

One could now try to look for a global factorization of ζ, hoping that it might contain some information about prime numbers as well. For this one has to know all zeros of ζ, i.e. also those ones in the gap

$$[0,1] + i\mathbb{R} = \mathbb{C} \setminus (L_{<0} \cup R_{>1})$$

between the two half planes $L_{<0}$ and $R_{>1}$, usually called the *critical strip*

Though one has up to now not really succeeded with that, there are estimates for the number of zeros in the rectangles $[0,1]+i[-T,T]$, which allow to assure the convergence of a product with factors corresponding to pairs of nontrivial zeros of ζ and to prove that it is the missing part needed for a factorization.

In order to describe how this works we denote

$$Z \subset [0,1]+i\mathbb{R}$$

the set of zeros of ζ in the critical strip $[0,1]+i\mathbb{R}$. We use the auxiliary entire function

$$\xi(z) := \Gamma\left(\frac{z}{2}+1\right)\cdot(z-1)\pi^{-z/2}\zeta(z)$$

with zero set Z and functional equation

$$\xi(z) = \xi(1-z).$$

Since furthermore $\xi(\overline{z}) = \overline{\xi(z)}$, it follows that its zero set Z is symmetric w.r.t. both the real line and the line $\Re(z)=\frac{1}{2}$.

Now taking $Z_+ = Z\cap([0,1]+i\mathbb{R}_{>0})$ - there are no zeros of ζ on $[0,1)$ - and pairing the linear factors belonging to $\varrho\in Z_+$ and $1-\varrho$ together, we find

$$\xi(z) = \frac{1}{2}\prod_{\varrho\in Z_+}\left(1-\frac{z^2-z}{\varrho^2-\varrho}\right) = \frac{1}{2}\prod_{\varrho\in Z_+}\left[\left(1-\frac{z}{\varrho}\right)\left(1-\frac{z}{1-\varrho}\right)\right],$$

with an absolutely convergent infinite product. Here points in Z_+ have to be counted with multiplicities - just to be on the safe side: It is conjectured that all zeros are simple ones.

Finally the formula

$$\zeta(z) = \frac{\pi^{z/2}}{(z-1)}\cdot e^{C(z/2+1)}\cdot\Delta\left(\frac{z}{2}+1\right)\cdot\frac{1}{2}\prod_{\varrho\in Z_+}\left(1-\frac{z^2-z}{\varrho^2-\varrho}\right)$$

gives a factorization of ζ.

Here are some comments on the set $Z = Z_+\cup(1-Z_+)$:

1. There are infinitely many zeros $\varrho_n = \frac{1}{2}\pm i\tau_n$ on the line $\Re(z)=\frac{1}{2}$. The first 20 approximative values are given in the following table:

n	τ_n
1	14,1347251417346937...
2	21,022039638771554993...
3	25,010857580145688763...
4	30,424876125859513210...
5	32,935061587739189690...
6	37,586178158825671257...
7	40,918719012147495187...
8	43,327073280914999519...
9	48,005150881167159727...
10	49,773832477672302181
11	52,970321477714460644...
12	56,446247697063394804...
13	59,347044002602353079...
14	60,831778524609809844...
15	65,112544048081606660...
16	67,079810529494173714...
17	69,546401711173979252...
18	72,067157674481907582...
19	75,704690699083933168...
20	77,144840068874805372...

2. There are no zeros on the boundary lines $\Re(z) = 0, 1$. This statement is equivalent to the prime number theorem Th.2.3.

3. For a number $a \in \left[\frac{1}{2}, 1\right)$ the following statements are equivalent:

   (a) $Z \subset [1-a, a]$,

   (b) The Dirichlet series $\Phi_\mu(z)$ belonging to the Möbius μ-function converges in $R_{>a}$,

   (c) For all $\varepsilon > 0$ we have $\pi(x) - \mathrm{li}(x) = O(x^{1/2+\varepsilon})$

The implication "$(b) \Longrightarrow (a)$" is easy: We have

$$\zeta(z)\Phi_\mu(z) = \Phi_1(z)\Phi_\mu(z) = \Phi_{1*\mu}(z) = \Phi_\delta(z) = 1$$

for $z \in R_{>1}$, hence by analytic continuation

$$\zeta(z)\Phi_\mu(z) = 1$$

for $z \in R_{>a} \setminus \{1\}$, in particular $\zeta(z) \neq 0$. But, unfortunately, it is not clear at all whether there is some $a \in \left[\frac{1}{2}, 1\right)$ satisfying the above three equivalent conditions.

The famous **Riemann hypothesis** now states that $a = \frac{1}{2}$ is possible or equivalently

$$Z \subset \frac{1}{2} + i\mathbb{R}.$$

Finally, for those interested in more details we state the formula for the (at the jumps modified) prime number function

$$F(x) := \frac{1}{2} \cdot \left(|P_{\leq x}| + |P_{<x}|\right),$$

which Riemann derives from the factorization of ζ. Actually we give a formula for the step function

$$J(x) = \sum_{n=1}^{\infty} \frac{1}{n} \cdot F(\sqrt[n]{x}),$$

which has a jump of height $1/n$ at every prime power p^n, and can then apply the inversion formula

$$F(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} J(\sqrt[n]{x}).$$

Here it is:

$$J(x) = \int_0^x \frac{dt}{\ln(t)} - \sum_{\varrho \in Z_+} \mathrm{Ei}((\varrho \ln(x)) + \mathrm{Ei}((1-\varrho)\ln(x))$$

$$- \ln(2) + \int_x^{\infty} \frac{dt}{t(t^2-1)\ln(t)}, \quad x > 1.$$

The sum is not absolutely convergent: It has to be evaluated with increasing $\Im(\varrho)$. Furthermore

$$\mathrm{Ei} : \mathbb{C} \setminus \mathbb{R}_{\geq 0} \longrightarrow \mathbb{C}$$

is the *exponential integral function*, characterized by

$$\mathrm{Ei}'(z) = \frac{e^z}{z}, \quad \lim_{x \to -\infty} \mathrm{Ei}(x) = 0.$$

Remark 13.1. The following "universality result" by Voronin shows that it definitely is not easy to control the behaviour of the ζ-function in the critical strip. Namely, given

1. a compact set $K \subset (\frac{1}{2}, 1) + i\mathbb{R}$ with connected complement $\mathbb{C} \setminus K$,
2. a continuous function $g : K \longrightarrow \mathbb{C}$ holomorphic in the interior of K,
3. and some $\varepsilon > 0$,

there is some $t \in \mathbb{R}$, such that

$$|g(z) - \zeta(z + it)| < \varepsilon$$

holds for all $z \in K$.

14 Linear Diophantine Equations

Theorem 14.1. *The diophantine equation*

$$a_1 x_1 + ... + a_n x_n = b$$

is solvable in $\mathbb{Z}^n$ *iff* $\gcd(a_1, ..., a_n) | b$.

Proof. Indeed, solvability means nothing but

$$b \in \mathbb{Z}a_1 + ... + \mathbb{Z}a_n = \mathbb{Z}\gcd(a_1, ..., a_n).$$

□

Remark 14.2. Here is a recipe how to find all solutions: Consider the $(n+1) \times n$-matrix

$$A := \begin{pmatrix} a_1, ..., a_n \\ I_n \end{pmatrix} \in \mathbb{Z}^{n+1,n},$$

transform it with column operations over $\mathbb{Z}$ (addition of a multiple of a column to some other column, multiplication of a column with ± 1 and exchange of two columns) to a matrix

$$C = \begin{pmatrix} d, 0, ..., 0 \\ C_0 \end{pmatrix} \in \mathbb{Z}^{n+1,n},$$

where $d := \gcd(a_1, ..., a_n)$. Then the solutions of our equation are of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = C_0 \begin{pmatrix} b/d \\ s_2 \\ \vdots \\ s_n \end{pmatrix}$$

where $s_2, ..., s_n \in \mathbb{Z}$.

Proof. We note first that

$$C\mathbb{Z}^n = A\mathbb{Z}^n,$$

since column operations correspond to multiplication from the right hand side with elementary matrices and on $\mathbb{Z}^n$ such matrices act as isomorphisms. Let

$$\mathcal{S} := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{Z}^n; a_1x_1 + ... + a_nx_n = b \right\}.$$

Then we have

$$\{b\} \times \mathcal{S} = A\mathbb{Z}^n \cap (\{b\} \times \mathbb{Z}^n) = C\mathbb{Z}^n \cap (\{b\} \times \mathbb{Z}^n) = \{b\} \times C_0 \cdot (\{b/d\} \times \mathbb{Z}^{n-1}).$$

□

If $n = 2$, we may use the euclidean algorithm in order to find integers $s_1, s_2 \in \mathbb{Z}$ with $s_1a_1 + s_2a_2 = d$, write $a_i = dc_i$ and obtain with

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{b}{d} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} + k \begin{pmatrix} c_2 \\ -c_1 \end{pmatrix}, k \in \mathbb{Z}$$

all solutions.

15 Sums of two squares and pythagorean triples

In this section we discuss our first non-linear diophantine equation: We consider the equation

$$x^2 + y^2 = n \in \mathbb{N}$$

and ask for which $n \in \mathbb{N}$ there is a solution $(x, y) \in \mathbb{Z}^2$. Our strategy is the following: We rewrite our equation in the form

$$(x + iy)(x - iy) = n,$$

write the RHS $n \in \mathbb{Z} \subset \mathbb{Z}[i]$ as a product of prime factors in the ring of gaussian integers

$$\mathbb{Z}[i] := \mathbb{Z} + \mathbb{Z}i \subset \mathbb{C}$$

and then check how the factors can be arranged in order to form a pair $z, \overline{z}$ of conjugate complex numbers. So let us first investigate "gaussian prime factorization".

First of all, the group of units consists of the shortest gaussian integers $\neq 0$, indeed

$$\mathbb{Z}[i]^* = \{z \in \mathbb{Z}[i]; |z| = 1\} = \{\pm 1, \pm i\}.$$

Definition 15.1. A non-zero gaussian integer $z \in \mathbb{Z}[i]$ is called (a) "g-prime", if it satisfies

1. $|z| > 1$, i.e. z is not a unit,
2. and $z = uv \Longrightarrow |u| = 1 \vee |v| = 1$, i.e. one of the factors is a unit.

Remark 15.2.

1. If $|z|^2 = p$ is prime, z is a g-prime.
2. A prime $p \in \mathbb{N}$ need not be a g-prime, e.g.

$$2 = (1 + i)(1 - i), 5 = (2 + i)(2 - i).$$

3. $1 + i, 1 - i, 2 + i, 2 - i$ are gaussian primes.
4. A prime $p = 4k+3$ is even a gaussian prime. Assume $p = zw, |z|, |w| > 1$. That implies $p = |z|^2 = |w|^2$. Let $z = x+iy$ and consider the residue classes

$$a := \overline{x}, b := \overline{y} \in \mathbb{Z}_p^*.$$

They satisfy $a^2 + b^2 = 0$ resp. $c^2 = -1$ holds for $c := ab^{-1}$, i.e. $c \in \mathbb{Z}_p^*$ has order 4 and thus 4 divides $p - 1$ or, equivalently, $p \equiv 1 \mod (4)$.

Definition 15.3. Two gaussian integers $z, w \neq 0$ are called associated iff they differ only by a unit, or equivalently if $|z| = |w| > 0$ and angle(z, w) equals 0, $\pi/2$ or π.

Example 15.4. $(1-i) = (-i)(1+i)$, while $2+i, 2-i$ are not associated.

Question: Does the FTA (**F**undamental **T**heorem of **A**rithmetics) hold for gaussian integers?
Answer: Obviously any gaussian integer is a finite product of gaussian primes, since in any nontrivial factorization the lengths decrease. For uniqueness up to order and association we need that for a gaussian prime z the implication

$$z|uv \Longrightarrow z|u \vee z|v$$

holds.
Idea: Create a gcd in $R = \mathbb{Z}[i]$, more technically, given $u, v \in R$, find $d \in R$ with

$$Ru + Rv = Rd.$$

Then we define the greatest common divisor of u and v as

$$\gcd(u, v) := d.$$

Here d, if it exists, is determined up to association. Sets of the form $Ru + Rv$ are ideals:

Definition 15.5. A non-empty subset $\mathfrak{a} \subset R$ of a ring R is called an ideal if

1. $\mathfrak{a} + \mathfrak{a} \subset \mathfrak{a}$, i.e. $b, a \in \mathfrak{a} \Longrightarrow b + a \in \mathfrak{a}$.

2. $R \cdot \mathfrak{a} \subset \mathfrak{a}$, i.e. $b \in R, a \in \mathfrak{a} \longrightarrow ba \in \mathfrak{a}$.

Example 15.6. Let $u, v \in R$.

$\mathfrak{a} = Ru$, a principal ideal, in particular $\mathfrak{a} = \{0\}$, the trivial or zero ideal, and $\mathfrak{a} = R$, the unit ideal.

$\mathfrak{a} = Ru + Rv$.

Proposition 15.7. *Any ideal $\mathfrak{a} \subset \mathbb{Z}[i]$ is a principal ideal:*

$$\mathfrak{a} = \mathbb{Z}[i]d$$

with a gaussian integer d unique up to multiplication with a unit.

Proof. If $\mathfrak{a} = \{0\}$, take $d = 0$. Otherwise take $d \in \mathfrak{a} \setminus \{0\}$ of minimal length. Now

$$d \in \mathfrak{a} \Longrightarrow \mathbb{Z}[i]d \subset \mathfrak{a}.$$

On the other hand given some element $u \in \mathfrak{a}$, approximate the complex number $ud^{-1} \in \mathbb{C}$ by a gaussian integer $b \in \mathbb{Z}[i]$, such that

$$|ud^{-1} - b| < 1.$$

That is possible, since any complex number lies in a unit square spanned by lattice points. Then the distance of any point in that square to the (or rather a) nearest lattice point is $\leq 1/\sqrt{2} < 1$. In particular

$$|u - bd| < |d|.$$

Since $u - bd \in \mathfrak{a}$, that means $u - bd = 0$ resp. $u = bd$. □

Theorem 15.8. *For a gaussian prime z we have*

$$z|uv \Longrightarrow z|u \vee z|v.$$

Proof. If $z \not| u$, we remember Prop.15.7 and write

$$\mathfrak{a} := \mathbb{Z}[i]z + \mathbb{Z}[i]u = \mathbb{Z}[i]d.$$

Since z is prime and $d|z$, we have either $z = ed$ with a unit $e \in \mathbb{Z}[i]^*$ or d itself is a unit, then w.l.o.g. $d = 1$. The first case implies $z|u$, contrary to our assumption. So the second case applies and we may write

$$1 = az + bu$$

resp.

$$v = vaz + b(uv),$$

where both summands are divisible with z. Hence so is their sum v. □

Theorem 15.9. *Let $p \equiv 1 \mod (4)$ with a prime p.*

1. *There is a unique g-prime $z_p = a + ib \in \mathbb{Z}[i]$ with $|z_p|^2 = p, 0 < b < a$. Up to association z_p and $\overline{z}_p$ are the only g-primes z with $|z|^2 = p$.*
2. *The equation $x^2+y^2 = p, (x, y) \in \mathbb{N}^2$, has exactly two solutions $(a, b), (b, a)$.*

Proof. First we look for a solution $(x, y) \in \mathbb{N}^2$. We know that there is a residue class $j \in \mathbb{Z}_p$ with $j^2 = -\overline{1}$ and consider the map

$$\psi : \mathbb{Z}[i] \longrightarrow \mathbb{Z}_p, x + yi \mapsto \overline{x} + \overline{y}j.$$

It is a ring homomorphism, i.e. respects both addition and multiplication. We consider its kernel

$$\ker(\psi) := \{u \in \mathbb{Z}[i]; \psi(u) = 0\} \supsetneqq \mathbb{Z}[i]p,$$

the inclusion being proper, since otherwise the p^2 elements

$$\mathbb{Z}_p \ni \psi(x + yi), 0 \leq x, y < p$$

would be pairwise different: Two elements with the same ψ-image differ by an element in the kernel. Now $\ker(\psi) \subset \mathbb{Z}[i]$ is an ideal, hence according to Prop. 15.7

$$\ker(\psi) = \mathbb{Z}[i]z.$$

In particular

$$p = zw.$$

and $p^2 = |z|^2 \cdot |w|^2$. Since $\mathbb{Z}[i]p \subsetneqq \mathbb{Z}[i]z \subsetneqq \mathbb{Z}[i]$, we have $z, w \notin \mathbb{Z}[i]^*$, hence

$$|z|^2 = p = |w|^2.$$

In particular $x^2 + y^2 = p$ for $z = x + yi$ and $w = \overline{z}$, furthermore $x \neq y$, since p is odd. Uniqueness follows with Th.15.8. □

Theorem 15.10. *The gaussian primes are, up to multiplication with $\pm 1, \pm i$, as follows:*

1. $1 + i$,
2. *for each prime number $p = 4k + 1$ there are two gaussian primes $z = a + bi$ with $0 < b < a, a^2 + b^2 = p$, and $\overline{z}$,*
3. *ordinary prime numbers $p = 4k + 3$.*

The given gaussian primes are pairwise non-associated, i.e. two different do not only differ by a factor $\pm 1, \pm i$; the first ones are

$$1+i, 3, 2+i, 2-i, 7, 11, 3+2i, 3-2i, 4+i, 4-i, 19, 23, 5+4i, 5-4i, 31, ...$$

Denoting the above sequence $(z_\nu)_{\nu\in\mathbb{N}}$ *any gaussian integer can uniquely be factorized*

$$z = ez_1^{k_1} \cdot ... \cdot z_r^{k_r}$$

with $e \in \{\pm 1, \pm i\}$ *and exponents* $k_1, ..., k_r \geq 0, k_r > 0$.

Proof. The given gaussian integers z_ν are prime, since either $|z_\nu| = p$ is prime or $|z_\nu| = p^2$ and there is no gaussian integer z with $|z|^2 = p$. That any gaussian integer $z \in \mathbb{Z}$ actually is a product of the given ones, follows by first factorizing $z\overline{z} \in \mathbb{N}$ and then identifying z with a partial product, using Th.15.8. □

Remark 15.11. Here are some hints how to find the prime factorization of gaussian integers: Let $z = x + iy \in \mathbb{Z}[i]$. We start writing

$$z = \gcd(x, y)z_0,$$

factorize $\gcd(x, y) \in \mathbb{N}$ in the usual way and then the prime divisors $p \equiv 1 \mod (4)$ as $p = \pi\overline{\pi}$ och $2 = -i(1 + i)^2$. —

From now on we assume $\gcd(x, y) = 1$, in particular no prime number $p \equiv 3 \mod (4)$ divides z. Then we know

$$z\overline{z} = |z|^2 = 2^\ell \prod_{\nu=1}^{r} p_\nu^{k_\nu}$$

with prime numbers $p_\nu \equiv 1 \mod (4)$. The factorization of z in gaussian primes now takes the form

$$z = e(1 + i)^\ell \prod_{\nu=1}^{r} \rho_\nu^{k_\nu},$$

where $e \in \{\pm 1, \pm i\}$ and, depending on ν, we have $\rho_\nu = \pi_\nu$ or $\rho_\nu = \overline{\pi}_\nu$ - since $\gcd(x, y) = 1$, the g-primes π_ν and $\overline{\pi}_\nu$ can not both divide z.

Example 15.12. We factorize $z = 201 + 43i$. We have $\gcd(201, 43) = 1$ and

$$z\overline{z} = |a|^2 = 42250 = 2 \cdot 5^3 \cdot 13^2.$$

Hence

$$a = e \cdot (1 + i)(\rho_1)^3(\rho_2)^2,$$

where $\rho_1 \in \{2 \pm i\}$ and $\rho_2 \in \{3 \pm 2i\}$. We start the factorization with

$$a = (1+i)(122 - 79i).$$

We check that $122 - 79i$ is divisible with $3 + 2i$ and factorize

$$a = -(1+i)(2+11i)(3+2i)^2.$$

We check that $2 + 11i$ is divisible with $(2 + i)$; and obtain the factorization

$$a = -(1+i)(2+i)^3(3+2i)^2.$$

Finally:

Theorem 15.13. *Let $n \in \mathbb{N}_{>1}$ be a natural number.*

1. *The equation*
$$|z|^2 = n$$
can be solved with a gaussian integer $z \in \mathbb{Z}[i]$ iff all prime divisors $p \equiv 3 \mod (4)$ of n have even multiplicity.

2. *If $n = a^2 m$, where all prime divisors of a are of the form $p = 4k+3$ and all prime divisors of m are of the form $p = 2$ or $p = 4k+1$, then any solution of the equation $|z|^2 = n$ is of the form $z = aw$ with $|w|^2 = m$.*

3. *If all prime divisors of n are of the form $p = 2$ or $p = 4k+1$, any solution of the equation*
$$|z|^2 = n = p_1 \cdot ... \cdot p_r$$
can be written as a product
$$z = z_1 \cdot ... \cdot z_r$$
of solutions $z_i \in \mathbb{Z}[i]$ of the equation $|z_i|^2 = p_i, i = 1, ..., r$. Here the primes $p_1, ..., p_r$ are not assumed to be pairwise different.

Proof of Th.15.13. Take $z \in \mathbb{Z}[i]$ with $|z|^2 = n$ and compare the prime factorizations of z and n. □

Eventually we want to determine *pythagorean triples*:

Definition 15.14. A triple $(a, b, c) \in (\mathbb{N}_{>0})^3$ is called a pythagorean triple if $a^2 + b^2 = c^2$. It is called primitive if $\gcd(a, b, c) = 1$.

So pythagorean triples correspond to right angled triangles with integer side lengths.

Remark 15.15. 1. Obviously any pythagorean triple can be written

$$(a, b, c) = \lambda(a_0, b_0, c_0)$$

with a primitive pythagorean triple (a_0, b_0, c_0).

2. For a primitive pythagorean triple the first two components a and b have different parity, and in particular c is odd. Clearly, they can not be both even, while odd a and b are impossible as well: This is easily seen by passing from $\mathbb{Z}$ to $\mathbb{Z}_4$: The only squares in $\mathbb{Z}_4$ are the residue classes $\overline{0}, \overline{1}$. Hence

$$\overline{c}^2 = \overline{a}^2 + \overline{b}^2 = \overline{1} + \overline{1} = \overline{2},$$

a contradiction!

Here is a recipe how to create all primitive pythagorean triples:

Theorem 15.16. *Let $p, q \in \mathbb{N}_{>0}$ be natural numbers such that*

1. $\gcd(p, q) = 1$,
2. $p > q$,
3. *p and q have different parity.*

Then the triple $(a, b, c) \in \mathbb{N}^3$ with

$$a = p^2 - q^2, b = 2pq, c = p^2 + q^2,$$

or, equivalently

$$a + bi = (p + iq)^2, c = |p + iq|^2,$$

is a primitive pythagorean triple with even b, and every such triple can be written in that way with uniquely determined numbers p, q.

Proof. Obviously, given p, q as above, the triple (a, b, c) is primitive and pythagorean.
Uniqueness: The equation $a+bi = w^2$ has $w = \pm(p+iq)$ as its only solutions.
Existence: We show that $z = a + bi$ admits a square root in $\mathbb{Z}[i]$, i.e.

$$a + bi = (p + iq)^2,$$

where we may assume $p > 0$. The remaining properties for p, q now follow easily from $a, b > 0$ and $\gcd(a, b) = 1$.

Since $\gcd(a, b) = 1$, we see that $\gcd(z, \overline{z}) = 1$ holds in $\mathbb{Z}[i]$. Indeed

1. If $(1 + i)|z$, then $2|z\overline{z} = c^2$, a contradiction.
2. If u is a gaussian prime, s.th. $\overline{u}$ is not associated to u and it divides both z and $\overline{z}$, then $\overline{u}|z$ and thus $u\overline{u}|z$, i.e. $u\overline{u} \in \mathbb{N}$ divides both a and b, a contradiction.

So

$$z\overline{z} = c^2$$

implies that all prime divisors of z have even multiplicity. As a consequence, z or iz is a square in $\mathbb{Z}[i]$ - but the latter is not possible, since $\Im(w^2)$ is even for $w \in \mathbb{Z}[i]$, while $\Im(iz) = a$ is odd. □

16 Fermat's Equation for $n > 2$

Theorem 15.16 describes all the solutions of the pythagorean equation

$$x^2 + y^2 = z^2.$$

Now given an arbitrary exponent $n \in \mathbb{N}$, we wonder what can be said about the solutions $(x, y, z) \in (\mathbb{N}_{>0})^3$ of *Fermat's equation*

$$x^n + y^n = z^n.$$

Unfortunately for $n > 2$ we do not find any solutions by inspection. The following theorem, conjectured by Fermat in 1637, has finally been proved in 1993:

Theorem 16.1 (Wiles)**.** *For $n \in \mathbb{N}_{>2}$ there are no triples $(x, y, z) \in (\mathbb{N}_{>0})^3$ satisfying*

$$x^n + y^n = z^n.$$

Of course it suffices to consider the exponent $n = 4$ or $n = p$, an odd prime number. The case $n = 3$ has already been settled by Euler, and for "regular prime numbers" the above theorem has been shown by Kummer in 1846 (Here are some of them: $p = 3$, 5, 7, 11, 13, 17, 19, 23, 29, 31, 41, 43, 47, 53, 61, 71, 73, 79, 83, 89, 97, 107, 109, 113, 127, 137, 139, 151, 163, 167, 173, 179, 181, 191, 193, 197, 199, ... — it is not known, whether there are infinitely many regular prime numbers or not.)

Following Fermat, we treat the case $n = 4$. It follows from

Theorem 16.2 (Fermat)**.** *There are no triples* $(x, y, z) \in (\mathbb{N}_{>0})^3$ *satisfying*

$$x^4 + y^4 = z^2.$$

Proof. The reasoning is as follows:

1. Assuming that there is a solution, we can even find a solution $(x, y, z) \in (\mathbb{N}_{>0})^3$ with minimal third component z.

2. Starting with a solution $(x, y, z) \in (\mathbb{N}_{>0})^3$ we construct (or "descend to") an other solution $(\tilde{x}, \tilde{y}, \tilde{z}) \in (\mathbb{N}_{>0})^3$ with $\tilde{z} < z$.

The two points are contradictory, hence the assumption in the first point is wrong.

Now let us explain the descent procedure: We apply twice the parametrization of primitive pythagorean triples.

1. If $d = \gcd(x, y, z) > 1$ we have $d^2 | z$ and may take $\tilde{x} = x/d, \tilde{y} = y/d, \tilde{z} = z/d^2$.

2. If $\gcd(x, y, z) = 1$, we obtain with (x^2, y^2, z) a primitive pythagorean triple. Assuming that x is even, we may write it

$$x^2 = 2pq, y^2 = p^2 - q^2, z = p^2 + q^2.$$

3. Then (q, y, p) is an other primitive pythagorean triple, in particular p is an odd number and q even, y being odd. Thus we may write

$$q = 2\ell k, y = \ell^2 - k^2, p = \ell^2 + k^2.$$

4. Since $2pq = x^2$ and p is odd and $\gcd(p, q) = 1$, we find that $p = \xi^2, q = 2\eta^2$ with suitable $\xi, \eta \in \mathbb{N}_{>0}$.

5. We obtain a further primitive pythagorean triple (ℓ, k, ξ).

6. The equality $2\eta^2 = q = 2\ell k$ together with $\gcd(\ell, k) = 1$, yields that $\ell = L^2, k = K^2$ with $L, K \in \mathbb{N}_{>0}$.

7. Finally $\tilde{x} := L, \tilde{y} := K, \tilde{z} := \xi$ is the solution we are looking for. Indeed

$$z = p^2 + q^2 = \xi^4 + q^2 > \xi^4 \geq \xi = \tilde{z}.$$

□

17 The Four Squares Theorem

Theorem 17.1. *A natural number $n \in \mathbb{N}$ is the sum of three squares*

$$x^2 + y^2 + z^2 = n$$

if and only if it is not of the form $n = 4^k(8m + 7)$.

We show that the given condition is necessary:

1. The case $k = 0$ follows from the fact, that the only squares in $\mathbb{Z}_8$ are $\overline{0}, \overline{1}$, and that $\overline{7}$ obviously is not the sum of three residue classes which either equal $\overline{0}$ or $\overline{1}$.

2. If $n = 4\ell$ is a sum of three squares, so is ℓ itself. It suffices to show that $x = 2a, y = 2b, z = 2c$ are even, so $a^2 + b^2 + c^2 = \ell$. But that is obvious: In $\mathbb{Z}_4$ a sum of three squares, not all $= \overline{0}$, never equals $\overline{0}$.

3. So $n = 4^k(8m+7)$ is never a sum of three square, since otherwise $8m+7$ would be as well.

If we even allow four squares there is no restriction anymore:

Theorem 17.2. *Any natural number is the sum of 4 squares of integers.*

Proof. As a consequence of the below four squares lemma we see that it suffices to write every prime number as a sum of four squares. Since

$$2 = 1^2 + 1^2 + 0^2 + 0^2,$$

we may concentrate on primes $p > 2$.

Lemma 17.3 (Four squares lemma)**.** *Let* $(x_1, ..., x_4), (y_1, ..., y_4) \in \mathbb{R}^4$*. Then we have*

$$(x_1^2 + ... + x_4^2)(y_1^2 + ... + y_4^2) = (Q_1^2 + ... + Q_4^2)$$

with

$$\begin{aligned} Q_1 &= \textstyle\sum_{\nu=1}^4 x_\nu y_\nu, \\ Q_2 &= -x_1y_2 + x_2y_1 - x_3y_4 + x_4y_3, \\ Q_3 &= -x_1y_3 + x_3y_1 - x_4y_2 + x_2y_4, \\ Q_4 &= -x_1y_4 + x_4y_1 - x_2y_3 + x_3y_2. \end{aligned}$$

Proof. Check yourself or use the interpretation of four vectors as quaternions, see the below remark. Then we have

$$(x_1 + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k})(y_1 - y_2\mathbf{i} - y_3\mathbf{j} - y_4\mathbf{k}) = Q_1 + Q_2\mathbf{i} + Q_3\mathbf{j} + Q_4\mathbf{k}$$

and use that the square of the euclidean norm

$$||x_1 + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k}||^2 = \sum_{\nu=1}^4 x_\nu^2$$

is multiplicative. □

Remark 17.4. For those not familiar with quaternions we give here a short introduction. We realize them as complex matrices: The set

$$\mathbb{H} := \left\{ \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}; z, w \in \mathbb{C} \right\} \subset \mathbb{C}^{2,2}$$

forms a real vector subspace of the complex vector space $\mathbb{C}^{2,2}$, it is multiplicatively closed and all nonzero matrices in $\mathbb{H}$ are invertible:

$$\mathbb{H} \setminus \{0\} \subset GL_2(\mathbb{C}).$$

Furthermore the euclidean norm

$$|| \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} || := \sqrt{z\overline{z} + w\overline{w}}$$

satisfies

$$|| \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} || = \sqrt{\det \begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}},$$

thus is obviously multiplicative:

$$||AB|| = ||A|| \cdot ||B||.$$

This is what we have used in the proof of Lemma 17.3. Let us explain that: A basis of the real vector space $\mathbb{H}$ is given by the matrices

$$E,\ I,\ J,\ K,$$

with the unit matrix $E \in \mathbb{C}^{2,2}$ and

$$I := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, J := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, K := \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

The products of the matrices K, I, J are as follows

$$I^2 = J^2 = K^2 = -E, IJ = K = -JI, JK = I = -KJ, KI = J = -IK.$$

In classical notation one writes

$$x_1 + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k} = x_1E + x_2I + x_3J + x_4K,$$

and calls the expression on the LHS a quaternion, since it is determined by four real parameters. The products are computed using the distributive law and the analogues of the above relations:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \mathbf{i} \cdot \mathbf{j} = \mathbf{k} = -\mathbf{j} \cdot \mathbf{i}, \mathbf{j} \cdot \mathbf{k} = \mathbf{i} = -\mathbf{k} \cdot \mathbf{j}, \mathbf{k} \cdot \mathbf{i} = \mathbf{j} = -\mathbf{i} \cdot \mathbf{k}.$$

This is the way how the "creator" of the quaternions, the irish mathematician William Rowan Hamilton, presented them. He was looking for some way how to define a product with nice properties on $\mathbb{R}^3$, say, making it a field, but did not succeed. Finally it worked for $\mathbb{R}^4$ - one gets almost a field, abandoning only the commutative law for the multiplication. Indeed, this is only possible on $\mathbb{R}^2$ - then one obtains $\mathbb{C}$ - and $\mathbb{R}^4$.

Though we do not really need it, we give some comments on the multiplicative structure of $\mathbb{H}$. There is a conjugation of quaternions:

$$A \mapsto A^* := \overline{A}^T,$$

satisfying

$$AA^* = \det(A)E.$$

Since conjugation is involutive, $\mathbb{H}$ is the direct sum

$$\mathbb{H} = \mathbb{H}_1 \oplus \mathbb{H}_{-1}$$

of the eigenspaces to the eigenvalues ± 1. Indeed

$$\mathbb{H}_1 = \mathbb{R}E$$

and

$$\mathbb{H}_{-1} = \mathbb{R}I + \mathbb{R}J + \mathbb{R}K.$$

For $A \in \mathbb{H}_{-1}$ of length $||A|| = 1$ we have $A^2 = -E$, hence there are infinitely many different possibilities to realize $\mathbb{C}$ in $\mathbb{H}$, namely as $\mathbb{R}E \oplus \mathbb{R}A$. Finally there is also a relation to three dimensional geometry on $\mathbb{H}_{-1} \cong \mathbb{R}^3$: If we denote

$$P : \mathbb{H} = \mathbb{H}_1 \oplus \mathbb{H}_{-1} \longrightarrow \mathbb{H}_{-1}$$

the projection onto $\mathbb{H}_{-1}$, the map

$$\mathbb{H}_{-1} \times \mathbb{H}_{-1} \longrightarrow \mathbb{H}_{-1}, (A, B) \longrightarrow P(AB)$$

describes the vector product.

Theorem 17.2 now follows from two lemmata, the first one being the start point for a descent procedure leading after finitely many steps to the desired decomposition of the prime p as a sum of 4 squares.

Lemma 17.5. *Let $p > 2$ be a prime. Then there is a natural number $h, 1 \leq h < p$, such that hp is a sum of 4 squares.*

Lemma 17.6. *Let $p > 2$ be a prime. Given $h \in \mathbb{N}, 1 < h < p$, such that hp is a sum of four squares, there is sone $q \in \mathbb{N}, 1 \leq q < h$, such that qp is a sum of four squares as well.*

Proof of Lemma 17.5. We consider the set

$$X := \left\{ \bar{i}^2 \in \mathbb{Z}_p; 0 \leq i \leq \frac{p-1}{2} \right\} \subset \mathbb{Z}_p.$$

It contains $\frac{p-1}{2} + 1$ elements, since if $\bar{j}^2 = \bar{i}^2$ with $0 \leq i < j \leq \frac{p-1}{2}$, we obtain

$$p|(j^2 - i^2) = (j+i)(j-i)$$

with

$$1 \leq j - i < j + i \leq p - 1,$$

a contradiction. Now $|X| + |-1 - X| = 2|X| = p + 1 > |\mathbb{Z}_p|$ implies that X and $-1 - X$ have at least one common element, so there are $1 \leq i, j \leq \frac{p-1}{2}$ with

$$\bar{i}^2 = -1 - \bar{j}^2$$

resp.

$$p|1 + i^2 + j^2.$$

But

$$0 < 1 + i^2 + j^2 \leq 1 + 2 \cdot \frac{(p-1)^2}{4} < 1 + \frac{p^2}{2} < p^2,$$

whence

$$0^2 + 1^2 + i^2 + j^2 = hp, 1 \leq h < p.$$

□

Proof of Lemma 17.6. Let

$$hp = \sum_{i=1}^{4} n_i^2.$$

We distinguish two cases:

1. If the number h is even, we may replace h with $q = \frac{h}{2}$ as follows: Then either none, two or all of the numbers $n_1, ..., n_4$ are odd: In any case we may assume that n_1, n_2 as well as n_3, n_4 have the same parity. Hence the numbers

   $$m_1 := \frac{n_1 + n_2}{2}, \; m_2 := \frac{n_1 - n_2}{2}$$

   and

   $$m_3 := \frac{n_3 + n_4}{2}, m_4 := \frac{n_3 - n_4}{2}$$

   are integers and satisfy for $q := \frac{h}{2}$ the desired equality

   $$qp = \sum_{i=1}^{4} m_i^2.$$

2. The number h is odd. In that case there are unique integers $\ell_1, ..., \ell_4$, satisfying

$$\ell_i \equiv n_i \mod (h) \ , \ -\frac{h-1}{2} \leq \ell_i \leq \frac{h-1}{2}.$$

Since not all n_i are divisible with h - otherwise hp would be divisible with h^2 - , we have $\ell_i \neq 0$ for some i. Hence

$$0 < \sum_{i=1}^{4} \ell_i^2 \leq 4\frac{(h-1)^2}{4} < h^2,$$

and thus

$$\sum_{i=1}^{4} \ell_i^2 \equiv \sum_{i=1}^{4} n_i^2 \equiv 0 \mod (h)$$

gives

$$\sum_{i=1}^{4} \ell_i^2 = qh, \ 1 \leq q < h.$$

On the other hand, by Lemma 17.3 we have

$$\sum_{i=1}^{4} Q_i^2 = \left(\sum_{i=1}^{4} n_i^2\right) \cdot \left(\sum_{i=1}^{4} \ell_i^2\right) = hp \cdot qh,$$

where all the Q_i are divisible with h: For $Q_2, .., Q_4$ that follows with the formulae of 17.2 and the fact that $\ell_i \equiv n_i \mod (h)$, then use that the entire sum is divisible with h^2, hence Q_1 is as well. Finally for m_i with $Q_i = hm_i$ we find

$$\sum_{i=1}^{4} m_i^2 = qp.$$

□

□

18 Pell's Equation

The second quadratic diophantine equation in two variables $(x, y) \in \mathbb{Z}^2$ we shall study is *Pell's equation*:

$$x^2 - d\,y^2 = n \in \mathbb{N},$$

where $d \in \mathbb{N}_{>0}$. If $d = c^2$ we can factorize its left hand side and get for any factorization $n = ab$ two linear equations

$$x + cy = a, x - cy = b$$

easy to be solved or noticed to be unsolvable. From now on we shall assume that $d > 0$ is not a square in $\mathbb{N}$. The main emphasis is on the case $n = \pm 1$, i.e.

$$x^2 - d\,y^2 = \pm 1.$$

We try the same strategy as in the previous section: We look at the ring

$$\mathbb{Z}\left[\sqrt{d}\right] := \mathbb{Z} + \mathbb{Z}\sqrt{d} \subset \mathbb{R}$$

of "*d-quadratic integers*" (no standard terminology!) - indeed it lies dense on the real line - and replace complex conjugation with the ring automorphism

$$\sigma : \mathbb{Z}\left[\sqrt{d}\right] \longrightarrow \mathbb{Z}\left[\sqrt{d}\right], x + y\sqrt{d} \mapsto x - y\sqrt{d}.$$

Though looking quite harmless, it is not from the topological point of view: It is nowhere continuous! E.g. if $u_n := x_n + y_n\sqrt{d} \to 0$ and $u_n \neq 0$ for all $n \in \mathbb{N}$, we have $|x_n|, |y_n| \to \infty$, and x_n, y_n have different sign for $n \gg 0$. It follows $|\sigma(u_n)| \to \infty$.

The function

$$N : \mathbb{Z}\left[\sqrt{d}\right] \longrightarrow \mathbb{Z}, u \mapsto N(u) := u\sigma(u),$$

associates to $u \in \mathbb{Z}\left[\sqrt{d}\right]$ its "norm" $N(u) \in \mathbb{Z}$. It is multiplicative, i.e. satisfies

$$N(uv) = N(u)N(v),$$

but in contrast to the ring of gaussian integers it may take both positive and negative values. In general there is no unique prime factorization available; indeed, we have to content ourselves with the computation of the group of units of $\mathbb{Z}[\sqrt{d}]$. Here is a first observation:

Proposition 18.1. *The group of units* $\mathbb{Z}\left[\sqrt{d}\right]^*$ *of the ring of d-quadratic integers is*

$$\mathbb{Z}\left[\sqrt{d}\right]^* = \{u \in \mathbb{Z}\left[\sqrt{d}\right]; N(u) = \pm 1\}.$$

Proof. For a d-quadratic integer $u \in \mathbb{Z}\left[\sqrt{d}\right]^*$ its norm $N(u) \in \mathbb{Z}$ is a unit as well because of $N(1) = 1$, hence $N(u) = \pm 1$. On the other hand, if $N(u) = \pm 1$, then $u^{-1} = N(a)\sigma(u) \in \mathbb{Z}\left[\sqrt{d}\right]$. □

So solving Pell's equation with $n = \pm 1$ turns out to be equivalent to determining the group of units

$$\mathbb{Z}\left[\sqrt{d}\right]^* \subset \mathbb{Z}\left[\sqrt{d}\right].$$

We want to appeal to geometric intuition and identify the lattice

$$\Lambda_d := \{(x, y\sqrt{d}) \in \mathbb{R}^2; x, y \in \mathbb{Z}\}$$

with the ring of d-quadratic integers: Consider the map

$$\pi : \mathbb{R}^2 \longrightarrow \mathbb{R}, (\xi, \eta) \mapsto \xi + \eta,$$

the projection onto the x-axis along the lines parallel to $\mathbb{R}(1, -1)$. It induces a bijection

$$\pi|_{\Lambda_d} : \Lambda_d \longrightarrow \mathbb{Z}\left[\sqrt{d}\right],$$

since on any line parallel to $\mathbb{R}(1, -1)$ there is at most one lattice point.

Remark 18.2. We have

$$\mathbb{Z}\left[\sqrt{d}\right]^* = \pi(\Lambda_d \cap H)$$

with the set

$$H := \{(\xi, \eta); \xi^2 - \eta^2 = \pm 1\},$$

the union of two hyperbolas composed of a left and right branch H_ℓ, H_r resp. an upper and a lower branch H_+, H_-. The map $\pi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ induces homeomorphisms

$$H_+ \xrightarrow{\cong} \mathbb{R}_{>0}, \; H_r \xrightarrow{\cong} \mathbb{R}_{>0}$$

and

$$H_- \xrightarrow{\cong} \mathbb{R}_{<0},\ H_\ell \xrightarrow{\cong} \mathbb{R}_{<0},$$

and the kernel $\pi^{-1}(0) = \mathbb{R}(1,-1)$ is one of the two asymptotic lines of the hyperbolas. Thus, $\Lambda_d \subset \mathbb{R}^2$ being discrete, it follows that

$$\mathbb{Z}\left[\sqrt{d}\right]^* \subset \mathbb{R}^*$$

is a discrete subset of the punctured line - the only possible point of accumulation is the origin. In any case we have

$$(\pm 1, 0) \in \Lambda_d \cap H,$$

corresponding to $\pm 1 \in \mathbb{Z}\left[\sqrt{d}\right]^*$, but do not know yet whether there are more lattice points on H than these two.

From the above geometric considerations we derive that the group of units $\mathbb{Z}\left[\sqrt{d}\right]^*$ admits up to sign a primitive root, also called the *basic unit*, as it is the case with $\mathbb{Z}_{2^n}^* = \pm\overline{5}^{\mathbb{Z}}$, only here it has infinite order:

Lemma 18.3. *If* $\mathbb{Z}\left[\sqrt{d}\right]^* \supsetneqq \{\pm 1\}$, *we have*

$$\mathbb{Z}\left[\sqrt{d}\right]^* = \pm a^{\mathbb{Z}}$$

with the "basic unit"

$$a := \min\left(\mathbb{Z}\left[\sqrt{d}\right]^* \cap \mathbb{R}_{>1}\right).$$

Indeed $a = \alpha + \beta\sqrt{d}$ *with positive integers* $\alpha, \beta \in \mathbb{N}_{>0}$.

Corollary 18.4. *If* $a = \alpha + \beta\sqrt{d}$ *is the basic unit of* $\mathbb{Z}\left[\sqrt{d}\right]$, *any solution* $(x, y) \in \mathbb{N}^2$ *of Pell's equation*

$$x^2 - d\,y^2 = \pm 1$$

is of the form

$$x + y\sqrt{d} = (\alpha + \beta\sqrt{d})^n$$

with some $n \in \mathbb{N}$.

Proof. Given $b \in \mathbb{Z}\left[\sqrt{d}\right]^*, b \neq \pm 1$, one of the numbers $\pm b^{\pm 1}$ lies in $\mathbb{R}_{>1}$. (Note that in terms of the lattice that means taking reflections w.r.t. the coordinate axes!) So the assumption implies that $(\mathbb{Z}\left[\sqrt{d}\right]^*)_{>1}$ is non-empty and thus the basic unit is well defined. It remains to show that every unit $b \in \mathbb{Z}\left[\sqrt{d}\right]^*$ is of the form $b = \pm a^n$ with some integer $n \in \mathbb{Z}$. Again we may assume $b > 1$: Then choose $n \in \mathbb{N}$ with $a^n \leq b < a^{n+1}$. We have $a^{-n}b \in \mathbb{Z}\left[\sqrt{d}\right]^*, 1 \leq a^{-n}b < a$, hence $a^{-n}b = 1$ resp. $b = a^n$. □

Remark 18.5. For the basic unit $a = \alpha + \beta\sqrt{d}$ we have $\alpha, \beta \geq 1$. As a consequence the sequences α_n, β_n with

$$a^n = \alpha_n + \beta_n\sqrt{d}$$

are strictly increasing:

$$\alpha_{n+1} = \alpha\alpha_n + d\beta\beta_n > \alpha\alpha_n \geq \alpha_n, \beta_{n+1} = \alpha\beta_n + \beta\alpha_n > \alpha\beta_n \geq \beta_n.$$

So in order to show that some d-quadratic unit $a = \alpha + \beta\sqrt{d}$ really is the basic unit, it is sufficient to show that there is no solution (x, y) of Pell's equation with $1 \leq x < \alpha, 1 \leq y < \beta$.

Example 18.6.
1. "$d = 2$": We find $a = 1 + \sqrt{2}$, $N(a) = -1$.
2. "$d = 3$": We find $a = 2 + \sqrt{3}$, $N(a) = 1$.
3. "$d = 5$": We find $a = 2 + \sqrt{5}$, $N(a) = -1$.

Remark 18.7. Note that $N(a) = -1$ iff $H_+ \cap \Lambda_d \neq \emptyset$.

19 Continued fractions

Here we describe a natural way to approximate an irrational number

$$x_0 \in \mathbb{R} \setminus \mathbb{Q}$$

by a sequence

$$(c_n)_{n\in\mathbb{N}} \subset \mathbb{Q}$$

of rational numbers as follows: We construct a bijection

$$\mathbb{R} \setminus \mathbb{Q} \longrightarrow \mathbb{Z} \times (\mathbb{N}_{>0})^{\mathbb{N}_{>0}}, x_0 \mapsto (a_n)_{n \in \mathbb{N}},$$

between the set of all irrational numbers and the set of all sequences $(a_n)_{n \in \mathbb{N}}$ of integers, s.th. $a_n > 0$ for $n > 0$, and obtain

$$c_n = K_n(a_0, ..., a_n)$$

with a function K_n defined on the set of initial segments of length $n+1$ of our sequences (a_ν). Furthermore we investigate for which $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ the corresponding sequence is "preperiodic", i.e. is periodic for $n \gg 0$.

Finally we find an answer to the question of the previous section: If we take $x_0 = \sqrt{d}$ and write

$$c_n = \frac{h_n}{k_n}$$

as a reduced fraction, the pair $(x, y) = (h_n, k_n)$ is a solution of Pell's equation $x^2 - dy^2 = \pm 1$ for infinitely many indices $n \in \mathbb{N}$.

So let us now start with our construction: Assume the irrational number $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ is given. We take

$$c_0 := a_0 := [x_0]$$

with the integer part function

$$[x] := \max \mathbb{Z}_{\leq x}.$$

In order to improve the approximation we consider the error

$$x_0 - c_0 \in (0, 1),$$

take the part integer part

$$a_1 := \left[\frac{1}{x_1}\right]$$

of its inverse

$$x_1 := \frac{1}{x_0 - c_0} \in \mathbb{R}_{>1}$$

and replacing x_1 in the formula

$$x_0 = a_0 + \frac{1}{x_1}$$

with a_1 we obtain the second approximation

$$c_1 = a_0 + \frac{1}{a_1}.$$

Now we improve the approximation $a_1 \in \mathbb{Z}$ of $x_1 \in \mathbb{R} \setminus \mathbb{Q}$ by a better rational one

$$a_1 + \frac{1}{a_2}$$

as we did with x_0 and substitute it:

$$c_2 = a_0 + \frac{1}{a_1 + \frac{1}{a_2}}.$$

Now we continue with that procedure and hopefully get better and better approximations of x_0. But before we discuss that question we present a more digestible description of the above algorithm:

Approximation of irrational numbers by finite continued fractions: Let $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ be an irrational number.

1. We take x_0 as the first member in a sequence

   $$(x_n)_{n\in\mathbb{N}} \ \subset \mathbb{R} \setminus \mathbb{Q}$$

   of irrational numbers $x_n \in \mathbb{R}$, defined by the recursion formula

   $$x_{n+1} := \frac{1}{x_n - [x_n]}.$$

   Obviously we have $x_n > 1$ for all $n \in \mathbb{N}$.

2. We associate to it two further sequences,

   (a) the sequence of integers $(a_n)_{n\in\mathbb{N}}$ defined by $a_n := [x_n]$,

   (b) and the sequence $(c_n)_{n\in\mathbb{N}}$ of rational numbers satisfying

   $$c_n := K_n(a_0, ..., a_n),$$

   where the functions

   $$K_n : \mathbb{R} \times (\mathbb{R}_{\geq 1})^n \longrightarrow \mathbb{R}$$

   are defined by $K_0(t) := t$ and

   $$K_n(t, t_1, ..., t_n) = K_{n-1}\left(t, t_1, .., t_{n-1} + \frac{1}{t_n}\right).$$

Remark 19.1. 1. In order to simplify notation we usually write

$$K(t, t_1, ..., t_n) = K_n(t, t_1, ..., t_n).$$

2. We have

$$K(t, t_1) = t + \frac{1}{t_1},\ K(t, t_1, t_2) = t + \frac{1}{t_1 + \frac{1}{t_2}}, ...$$

3. The function $K(t, t_1, .., t_n)$ is strictly increasing w.r.t. t and t_{2i} and strictly decreasing w.r.t. t_{2i+1}.

4. By induction one proves

$$x_0 = K(a_0, a_1, ..., a_{n-1}, x_n),$$

in particular we have

$$c_{2n} = K(a_0, a_1, ..., a_{2n-1}, a_{2n}) < x_0 < c_{2n+1} = K(a_0, a_1, ..., a_{2n}, a_{2n+1}).$$

We shall prove

Theorem 19.2. *Denote $\mathbb{Z} \times (\mathbb{N}_{>0})^{\mathbb{N}_{>0}}$ the set of all sequences $(a_\nu)_{\nu\in\mathbb{N}}$ of integers with $a_\nu \geq 1$ for $\nu \geq 1$. Then the limit*

$$K(a_0, a_1, ...) := \lim_{n\to\infty} K(a_0, ..., a_n)$$

exists for any sequence $(a_\nu)_{\nu\in\mathbb{N}} \in \mathbb{Z} \times (\mathbb{N}_{>0})^{\mathbb{N}_{>0}}$ and is an irrational number. Indeed the map

$$K : \mathbb{Z} \times (\mathbb{N}_{>0})^{\mathbb{N}_{>0}} \longrightarrow \mathbb{R} \setminus \mathbb{Q}, (a_n)_{n\in\mathbb{N}} \mapsto K(a_0, a_1, ...)$$

is a bijection.

Theorem 19.3. *Denote $\mathbb{Z} \times (\mathbb{N}_{>0})^{\infty}$ with*

$$(\mathbb{N}_{>0})^{\infty} = \bigcup_{n=1}^{\infty} (\mathbb{N}_{>0})^n$$

the set of all finite sequences $a_0, ..., a_n \in \mathbb{Z}$ of integers with $a_\nu \geq 1$ for $\nu \geq 1$. The map

$$K : \mathbb{Z} \times (\mathbb{N}_{>0})^{\infty} \longrightarrow \mathbb{Q}, (a_0, ..., a_n) \mapsto K(a_0, a_1, ..., a_n)$$

is two to one, i.e. it is onto and all its fibres have order 2. Indeed:

$$K(a_0, ..., a_n) = \begin{cases} K(a_0, ..., a_{n-1} + 1) & , \quad \text{if } a_n = 1, n > 0 \\ K(a_0, ..., a_n - 1, 1) & , \quad \text{if } a_n > 1 \end{cases}.$$

Proof of Th.19.3. Assume

$$K(a_0, ..., a_n) = K(b_0, ..., b_m), m \geq n.$$

Choose $\ell \leq n$ maximal with $a_\ell = b_\ell$. If $\ell = n$, we have $m = n$, since $m > n$ would imply

$$K(a_0, ..., a_n, b_{n+1}, ..., b_m) = K(a_0, ..., a_n + \frac{1}{K(b_{n+1}, .., b_m)}) \neq K(a_0, ..., a_n).$$

If $\ell < n$, we have

$$[a_{\ell+1}, a_{\ell+1} + 1] \ni K(a_{\ell+1}, .., a_n) = K(b_{\ell+1}, ..., b_m) \in [b_{\ell+1}, b_{\ell+1} + 1],$$

with the intervals having exactly one boundary point in common. Since by assumption $n \leq m$, we find with the below remark $n = \ell + 1, m = n + 1$ and $b_m = a_n + 1$, as desired.
Surjectivity: Take $x_0 = \frac{u_0}{u_1} \in \mathbb{Q}$ and define the sequence x_ν. Then we have

$$x_0 = K(a_0, ..., a_n)$$

with $a_\nu = [x_\nu]$. □

Here are some preparatory results for the proof of Th.19.2:

Remark 19.4. For $(x_0, x_1, ..., x_n) \in \mathbb{R} \times (\mathbb{R}_{\geq 1})^n$ we have

1. $K(x_0, ..., x_n) = x_0 + \frac{1}{K(x_1, ..., x_n)}$,
2. $K(x_0, ..., x_n) \geq 1$ for $x_0 \geq 1$,
3. $x_0 \leq K(x_0, x_1, ..., x_n) \leq x_0 + 1$,
4. $x_0 = K(x_0, x_1, ..., x_n) \iff n = 0$,
5. $K(x_0, x_1, ..., x_n) = x_0 + 1 \iff n = 1, x_1 = 1$.

For $a \in \mathbb{R}$ let

$$M(a) := \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}.$$

Given $(a_\nu)_{\nu\in\mathbb{N}}$ we define recursively a sequence of matrices $(Q_n)_{n\in\mathbb{N}}$ by

$$Q_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$Q_{n+1} = Q_n M(a_n).$$

We derive from it two sequences $(h_\nu)_{\nu\geq -2}$ and $(k_\nu)_{\nu\geq -2}$ as follows:

$$Q_n = \begin{pmatrix} h_{n-1} & h_{n-2} \\ k_{n-1} & k_{n-2} \end{pmatrix}.$$

More explicitly

$$h_n = a_n h_{n-1} + h_{n-2}, k_n = a_n k_{n-1} + k_{n-2}$$

with

$$h_0 = a_0, k_0 = 1.$$

Proposition 19.5. *For $n \geq 1$ and $x \in \mathbb{R}_{\geq 1}$ we have*

$$K(a_0, ..., a_{n-1}, x) = \frac{x h_{n-1} + h_{n-2}}{x k_{n-1} + k_{n-2}},$$

in particular

$$c_n = K(a_0, ..., a_{n-1}, a_n) = \frac{h_n}{k_n}.$$

Proof. A straight forward verification. □

Remark 19.6. Here is a rough estimate for the denominators: We have

$$k_n \geq n, n \geq 0.$$

Indeed $k_0 = 1, k_1 = a_1 k_0 + k_{-1} = a_1 \geq 1$, while $a_n \geq 1$ for $n \geq 1$ yields

$$k_n = a_n k_{n-1} + k_{n-2} \geq (n-1) + 1 = n$$

for $n \geq 2$.

Lemma 19.7. *1.* $h_n k_{n-1} - h_{n-1} k_n = \det Q_{n+1} = (-1)^{n+1}$,

2. $h_n k_{n-2} - h_{n-2} k_n = (-1)^n a_n$,

3. hence

$$c_n - c_{n-1} = (-1)^{n-1} \frac{1}{k_n k_{n-1}}$$

4. and

$$c_n - c_{n-2} = (-1)^n \frac{a_n}{k_n k_{n-2}}.$$

Corollary 19.8. *The sequence $(c_n)_{n\in\mathbb{N}}$ converges, the subsequence $(c_{2n})_{n\in\mathbb{N}}$ is strictly increasing, while $(c_{2n+1})_{n\in\mathbb{N}}$ is strictly decreasing.*

Proof. By Rem. 19.6 and Lemma 19.7.3 we have

$$|c_n - c_{n-1}| \leq \frac{1}{n(n-1)},$$

hence the series $\sum_{n=1} c_n - c_{n-1}$ converges even absolutely. This implies the first part of the statement, while the second is nothing but 19.7.4. □

Proof of Lemma 19.7. 1. Follows from $\det M(a_\nu) = -1$ for $\nu = 0, ..., n$.

2. We have

$$\begin{pmatrix} h_n & h_{n-2} \\ k_n & k_{n-2} \end{pmatrix} = Q_n \cdot \begin{pmatrix} a_n & 0 \\ 1 & 1 \end{pmatrix},$$

now take determinants.

3. We compute

$$c_n - c_{n-1} = \frac{h_n}{k_n} - \frac{h_{n-1}}{k_{n-1}} = \frac{h_n k_{n-1} - h_{n-1} k_n}{k_n k_{n-1}} = \frac{(-1)^{n-1}}{k_n k_{n-1}}$$

4. and

$$c_n - c_{n-2} = \frac{h_n}{k_n} - \frac{h_{n-2}}{k_{n-2}} = \frac{h_n k_{n-2} - h_{n-2} k_n}{k_n k_{n-2}} = \frac{(-1)^n a_n}{k_n k_{n-2}}.$$

□

Proof of Th.19.2. The map K is well defined: Cor.19.8 assures convergence, and the limit is an irrational number: Assume that

$$x_0 = \lim_{n\to\infty} c_n = \frac{h}{k} \in \mathbb{Q}.$$

Then we have

$$0 < \left|\frac{h}{k} - c_n\right| < |c_{n+1} - c_n| \leq \frac{1}{k_{n+1}k_n}$$

because x_0 lies between c_n and c_{n+1}. Now we multiply

$$0 < \left|\frac{h}{k} - \frac{h_n}{k_n}\right| < \frac{1}{k_{n+1}k_n}$$

with kk_n and obtain

$$0 < |hk_n - h_nk| < \frac{k}{k_{n+1}} < 1$$

for $n \geq k$ and $hk_n - h_nk \in \mathbb{Z}$, a contradiction!
Surjectivity: Follows from Rem.19.1.4.
Injectivity: We assume $K(a_0, ...) = K(b_0, ...)$ and show $a_0 = b_0$. Since

$$K(a_0, ...) = a_0 + \frac{1}{K(a_1, ...)}, \quad K(b_0, ...) = b_0 + \frac{1}{K(b_1, ...)},$$

that implies $K(a_1, ...) = K(b_1, ...)$ and thus we can obtain by induction $a_n = b_n$ for all $n \in \mathbb{N}$. An interval $[\ell, \ell + 1]$ with $\ell \in \mathbb{Z}$ is determined by any of its interior points, thus

$$a_0 < K(a_0, ..., a_{2n}) \leq K(a_0, ...) \leq K(a_0, ..., a_{2n+1}) < a_0 + 1$$

as well as

$$b_0 < K(b_0, ...) < b_0 + 1$$

implies $a_0 = b_0$. □

Unfortunately we have no easy results relating the continued fraction expansion of a sum resp. a product to those of the summands resp. factors, but we can characterize the irrationals having an expansion, which becomes after some initial segment periodic:

Theorem 19.9. *An irrational number $x_0 \in \mathbb{R}\backslash\mathbb{Q}$ has a preperiodic expansion*

$$x_0 = K(a_0, ..., a_{r-1}, \overline{a_r, ..., a_{r-1+p}}),$$

if and only if $x_0 \in \mathbb{Q}\left[\sqrt{D}\right] := \mathbb{Q} + \mathbb{Q}\sqrt{D}$ for some non-square $D \in \mathbb{N}_{\geq 1}$.

Proof. "$\Longrightarrow$": For $r = 0$, i.e.

$$x := K(\overline{a_0, ..., a_{p-1}})$$

we have

$$x = K(a_0, ..., a_{p-1}, x) = \frac{xh_{p-1} + h_{p-2}}{xk_{p-1} + k_{p-2}},$$

i.e. x satisfies a quadratic equation with rational coefficients, hence belongs to some field $\mathbb{Q}\left[\sqrt{D}\right]$. In the general case we obtain that

$$y := K(\overline{a_r, ..., a_{r-1+p}}) \in \mathbb{Q}\left[\sqrt{D}\right]$$

and then

$$x_0 = K(a_0, ..., a_{r-1}, y) = \frac{yh_{r-1} + h_{r-2}}{yk_{r-1} + k_{r-2}} \in \mathbb{Q}\left[\sqrt{D}\right]$$

as well.
"$\Longleftarrow$": We show that the set

$$\{x_n; n \in \mathbb{N}\}$$

is finite: Then we have $x_{k+p} = x_k$ for some $k, p \in \mathbb{N}$ and thus $x_{n+p} = x_n$ for all $n \geq k$. Now, if $x_0 \in \mathbb{Q}\left[\sqrt{D}\right]$, we have as well $x_n \in \mathbb{Q}\left[\sqrt{D}\right]$ for all $n \in \mathbb{N}$. We use that fact in order to write $x_n = f(m_n, q_n)$ with integers m_n, q_n. Furthermore, once again we need the conjugation

$$\sigma : \mathbb{Q}\left[\sqrt{D}\right] \longrightarrow \mathbb{Q}\left[\sqrt{D}\right], \alpha + \beta\sqrt{D} \mapsto \alpha - \beta\sqrt{D}.$$

Indeed, we verify the following three statements:

1. For a suitable $d = \ell^2 D$, every $x_n \in \mathbb{Q}\left[\sqrt{d}\right] = \mathbb{Q}\left[\sqrt{D}\right]$ can be written in the form
$$x_n = \frac{m_n + \sqrt{d}}{q_n}$$
with integers $m_n, q_n \in \mathbb{Z}$ satisfying $d - m_n^2 \in \mathbb{Z}q_n$. Indeed they can be computed using the sequence $(a_n)_{n\in\mathbb{N}}$ as follows:
   (a) $m_{n+1} = a_n q_n - m_n$,

(b) $q_{n+1}q_n = d - m_{n+1}^2$.

2. For some $k \in \mathbb{N}_{\geq 1}$ we have

$$\sigma(x_k) < 0$$

and for $n > k$ even

$$-1 < \sigma(x_n) < 0.$$

3. If 2.) holds for the index k, we have

$$(m_n, q_n) \in \left[0, \sqrt{d}\right] \times \left[0, 2\sqrt{d}\right]$$

for $n > k+1$. Hence there are only finitely many possible lattice points (m_n, q_n) resp. values $x_n, n \in \mathbb{N}$.

The statement 1.): The case $n = 0$: We extend the fraction

$$x_0 = \frac{a + b\sqrt{D}}{c}; \quad a, c \in \mathbb{Z}, b \in \mathbb{N}_{>0},$$

with $|c|$ and obtain

$$x_0 = \frac{a \cdot |c| + \sqrt{b^2c^2D}}{c \cdot |c|},$$

i.e. we may take

$$d = b^2c^2D, \ m_0 = a \cdot |c|, \ q_0 = c \cdot |c|.$$

Indeed

$$d - m_0^2 = b^2c^2D - a^2c^2 = \frac{c}{|c|}\left(b^2D - a^2\right) q_0 \in \mathbb{Z}q_0.$$

Now let us consider the step from n to $n+1$, applying the recursion formula:

$$x_{n+1} = \frac{1}{x_n - [x_n]} = \frac{q_n}{m_n + \sqrt{d} - q_n a_n}$$

$$= \frac{q_n}{\sqrt{d} - m_{n+1}} = q_n \frac{m_{n+1} + \sqrt{d}}{d - m_{n+1}^2} = \frac{m_{n+1} + \sqrt{d}}{q_{n+1}},$$

with $q_{n+1} \in \mathbb{Z}$, since the assumption $q_n | (d - m_n^2)$ implies $q_n | (d - m_{n+1}^2)$.

The statement 2.): With Prop.19.5 we may write

$$x_0 = K(a_0, ..., a_{n-1}, x_n) = \frac{x_n h_{n-1} + h_{n-2}}{x_n k_{n-1} + k_{n-2}}$$

and

$$\sigma(x_0) = \frac{\sigma(x_n) h_{n-1} + h_{n-2}}{\sigma(x_n) k_{n-1} + k_{n-2}}.$$

Solving for $\sigma(x_n)$ we arrive at

$$\sigma(x_n) = -\frac{k_{n-2}}{k_{n-1}} \cdot \frac{\sigma(x_0) - c_{n-2}}{\sigma(x_0) - c_{n-1}},$$

and since $k_{n-1}, k_{n-2} > 0$ and the second fraction converges to 1 (note that $\sigma(x_0) - c_n \to \sigma(x_0) - x_0 \neq 0$), it follows that $\sigma(x_n) < 0$ for $n \gg 0$.

Finally, for $n \geq 1$ we have $\sigma(x_n) < 0 \Longrightarrow \sigma(x_{n+1}) \in (-1, 0)$, since $x_n > 1$ for $n \geq 1$ implies

$$\sigma(x_n) - [x_n] > -1$$

and

$$0 > \sigma(x_{n+1}) = \frac{1}{\sigma(x_n) - [x_n]} > -1.$$

The statement 3.): We substitute $x_n = ...$ in 2.) and obtain

$$-1 < \frac{m_n - \sqrt{d}}{q_n} < 0,$$

while in any case

$$1 < x_n = \frac{m_n + \sqrt{d}}{q_n}.$$

If we add, we get

$$0 < \frac{2m_n}{q_n},$$

i.e. m_n, q_n have the same sign. On the other hand $-\sigma(x_n) > 0$ yields

$$1 < x_n - \sigma(x_n) = 2\frac{\sqrt{d}}{q_n},$$

so m_n and q_n are positive and

$$0 < q_n < 2\sqrt{d}.$$

Now by 1.(b) with n instead of $n+1$ we have

$$d - m_n^2 = q_n q_{n-1} > 0,$$

hence

$$m_n < \sqrt{d}.$$

□

We note without proof:

Theorem 19.10. *The quadratic irrational $x_0 \in \mathbb{Q}\left[\sqrt{d}\right] \setminus \mathbb{Q}$ has a periodic continued fraction expansion*

$$x_0 = K(\overline{a_0, ..., a_{p-1}})$$

if and only if $x_0 > 1$ and $-1 < \sigma(x_0) < 0$.

Example 19.11. We have

$$\left[\sqrt{d}\right] + \sqrt{d} = K\left(\overline{2\left[\sqrt{d}\right], a_1, ..., a_{p-1}}\right),$$

where we usually assume the period p to be minimal. In particular

$$\sqrt{d} = K\left(\left[\sqrt{d}\right], \overline{a_1, ..., a_p}\right).$$

We come now back to Pell's equation:

Proposition 19.12. *For $x_0 = \sqrt{d}$ we have*

$$(h_n)^2 - d(k_n)^2 = (-1)^{n-1} q_{n+1}.$$

Proof of Prop.19.12. We have

$$\sqrt{d} = x_0 = \frac{x_{n+1}h_n + h_{n-1}}{x_{n+1}k_n + k_{n-1}} = \frac{(m_{n+1} + \sqrt{d})h_n + q_{n+1}h_{n-1}}{(m_{n+1} + \sqrt{d})k_n + q_{n+1}k_{n-1}},$$

multiply with the denominator and compare coefficients:

$$\alpha + \beta\sqrt{d} = \tilde{\alpha} + \tilde{\beta}\sqrt{d} \Longleftrightarrow \alpha = \tilde{\alpha}, \beta = \tilde{\beta}.$$

□

Corollary 19.13. *We have*

$$(h_{\ell p-1})^2 - d(k_{\ell p-1})^2 = (-1)^{\ell p}.$$

Proof. The equality $x_0 = \sqrt{d}$ gives $q_0 = 1$, while for $\ell > 0$ we have

$$x_{\ell p} = K(a_{\ell p}, ...) = K(a_p, ...) = \left[\sqrt{d}\right] + \sqrt{d},$$

whence $q_{\ell p} = 1$. □

Remark 19.14. Indeed, any solution of Pell's equation $x^2 - dy^2 = \pm 1$ is obtained as in Cor.19.13. In particular $x^2 - dy^2 = -1$ is solvable if and only if the minimal period of the continuous fraction expansion of $\sqrt{d}$ is odd.