MAS 4203 Number Theory

M. Yotov

June 15, 2017

These Notes were compiled by the author with the intent to be used by his students as a main text for the course *MAS 4203 Number Theory* taught at the Department of Mathematics and Statistics of FIU. The material included covers both the standard topics for an undergraduate course on the subject as well as some additional topics which form a glimpse into more advanced and modern facts and techniques in the area. The presentation emphasizes the algebraic nature of the results proved, and is based on four important facts: the Fundamental Theorem of the Arithmetic of $\mathbb{Z}$ (this ring is a UFD), the Chinese Remainder Theorem, the Hensel's Lifting Lemma, and the Law of Quadratic Reciprocity. The motivation for studying the topics included in the Notes comes from the need to solve equations over $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{Z}/n\mathbb{Z}$. The questions answered in discussing such equations are: when do solutions exist, and, in case there are such, how to find all solutions? In connection with these, in special sections, called Vistas,the p-adic numbers are introduced and their role in solving equations over $\mathbb{Q}$ explained. The Law of Quadratic Reciprocity is thoroughly discussed in its Legendre-Gauss, Jacobi's, Euler's, and Hilbert's versions (the latter one - in a Vista). The primitive root theory is based on the Hensel's Lifting Lemma, allowing a direct proof of existence of such for a power of an odd prime, and is treated as result on the structure of the group of units of the ring $\mathbb{Z}/n\mathbb{Z}$. A chapter of the Notes is devoted to the basic arithmetic functions, their properties, and applications to classical number theoretical problems. The last chapter of the Notes is devoted to elementary analytic methods used in studying prime numbers.

Please send comments and corrections to the author at yotovm@fiu.edu .

Contents

Chapter 1

A Motivational One

In this chapter, we are discussing some preliminaries, and giving a motivation for what we are doing in this course.

1.1 Natural Numbers

This section recalls some concepts and facts known to the students from the course of Intro to Advanced Mathematics.

1.1.1 Operations and Relations on the Natural Numbers

In this course, we will be interested in the properties of natural numbers related to the two operations on them: addition and multiplication. We denote the set of natural numbers by $\mathbb{N}$

$$\mathbb{N} = \{0, 1, 2, \dots, n, \dots\}.$$

The two operations on $\mathbb{N}$ are commutative, associative, have neutral elements, and the addition distributes over the multiplication. That is, for all $m, n, p \in \mathbb{N}$

$$m + n = n + m \qquad m + (n + p) = (m + n) + p \qquad 0 + n = n$$

$$m \cdot n = n \cdot m \qquad m \cdot (n \cdot p) = (m \cdot n) \cdot p \qquad 1 \cdot n = n.$$

Natural numbers can be compared as well: $m \leq n$ if $n = m + p$ for some $p \in \mathbb{N}$. The **relation** $\leq$ is a total order on $\mathbb{N}$, that is, for every $m, n, p \in \mathbb{N}$

$$m \leq m \qquad (m \leq n \wedge n \leq m) \Rightarrow m = n \qquad (m \leq n \wedge n \leq p) \Rightarrow m \leq p \qquad m \leq n \vee n \leq m.$$

The relation $\leq$ on $\mathbb{N}$ is a **well order**. That is, $\leq$ has the property, called **The Least Element Principle (LEP)**, that every non-empty subset of $\mathbb{N}$ has a least element. In more professional notations

$$(\forall \Sigma \subseteq \mathbb{N})(\Sigma \neq \emptyset \Rightarrow (\exists s_0 \in \Sigma)(\forall s \in \Sigma)(s_0 \leq s)).$$

As we know from Intro to Advanced Math, the LEP lays the ground for using the method of proof by (finite) induction. We will often be using this method of proof in this course.

Example 1.1.1 The first example of proof by induction one can see in any book discussing induction is the following one. Prove that

$$(\forall n \in \mathbb{N})(0 + \dots + n = n(n + 1)/2).$$

The proof by induction includes two steps: first we show that the claim is true for the smallest natural number, that is for $n = 0$ (the so called **base case**), and then we prove that if the proposition is

true for n, it is true for $n+1$ as well (the so called **inductive step**). For the base case, we have to verify that $0 = 0(0+1)/2$ which is obviously true. For the inductive step, assuming the claim for n is true, we have

$$0 + \cdots + n + (n+1) = (0 + \cdots + n) + (n+1)$$
$$= n(n+1)/2 + (n+1) = ((n+1)/2) \cdot (n+2).$$

So, indeed,

$$0 + \cdots + n + (n+1) = (n+1)((n+1)+1)/2$$

and therefore the claim is true for $n+1$ as well. □

The relation $\leq$ defines a **strict** total order on $\mathbb{N}$: for $m, n \in \mathbb{N}$ we define

$$m < n \qquad \text{if} \qquad m \leq n \ \wedge \ m \neq n.$$

This strict order is **trichotomous**: for every $m, n \in \mathbb{N}$ exactly one of the following propositions is true

$$m < n \qquad n < m \qquad m = n.$$

1.1.2 Prime and Composite Natural Numbers

The first interesting thing one observes related to the operation multiplication, $\cdot$, is that some natural numbers **divide** others.

Definition 1.1.1 *For $m, n \in \mathbb{N}$ we say that n* **divides** *m, or that m is* **divisible** *by n, if there is a natural number p such that $m = n \cdot p$. In such a case, we write $n \,|\, m$. The number n is called the* **divisor***, and the number m is called the* **dividend** *.*

Exercise 1.1 *1) Prove that if $a \,|\, b$ and $a \,|\, c$, then $a \,|\, (b+c)$.*
[We have: $b = ab_1, c = ac_1$, and so $b + c = a(b_1 + c_1)$.]
2) Suppose $a, b, c \in \mathbb{N}$ such that $a \neq 0$. Prove that $ab \,|\, ac$ if, and only if, $b \,|\, c$.
[($\Rightarrow$) $ab \,|\, ac$ means that $ac = abd$ for some natural number d which, since $a \neq 0$, is equivalent to $c = bd$, that is, equivalent to $b \,|\, c$.]
3) Suppose $a, b \in \mathbb{N}$ such that $b < a$. Prove that $a \,|\, b$ if, and only if, $b = 0$.

4) Prove that for every three natural numbers m, n, p the following holds true

$$n \,|\, n \qquad (n \,|\, m \ \wedge \ m \,|\, n) \Rightarrow m = n \qquad (m \,|\, n \ \wedge \ n \,|\, p) \Rightarrow m \,|\, p.$$

In other words, divisibility of natural numbers defines a partial order on $\mathbb{N}$. This partial order is **not** *a total order (Why?)* □

A number n is called **even** if $2 \,|\, n$, and is called **odd** otherwise. It's obvious, for instance, that of two consecutive natural numbers, exactly one is even (Why is that true?). We can do the same using 3 instead of 2, and define all the numbers divisible by 3. It is also obvious that of three consecutive natural numbers, one is always divisible by 3.

Exercise 1.2 *It is easy to do the same for any natural number n instead of 2 or 3. Prove that of any n consecutive numbers, one is divisible by n. Furthermore, prove that there is* **only one** *such number.*
[Hint: Consider n consecutive natural numbers: $m + 0, m + 1, \cdots, m + (n-1)$ where $m \in \mathbb{N}$. Prove by induction on m that for some $0 \leq i \leq n-1$ the number $m + i$ is divisible by n. For the uniqueness, notice that if n divides $m + i$ and $m + j$ for $0 \leq i < j \leq n-1$, then n divides $(m + j) - (m + i) = j - i$. But the latter divisibility is impossible, because $j - i$ is a too small positive integer. Fill in the gaps of this argument.] □

We will prove soon that a non-zero natural number is divisible by finitely many natural numbers only. With this in mind we see that the number 0 is special: it is divisible by **all** natural numbers. The number 1 is also special in this respect. Note that if $1 < n$, then n has at least **two** divisors, while the number 1 has only one! So, the neutral elements of the operations $+$ and $\cdot$ stand out. By the way, these two have other features which make them special: the number 0 divides **only one** natural number, 0, while the number 1 divides **all** natural numbers. Every integer which is bigger than 1 divides infinitely many natural numbers, but not all of them. So, from the point of view of divisibility, the natural numbers bigger than 1 are more interesting. Here is how we label these natural numbers.

Definition 1.1.2 *The natural number n is called* **prime** *if it has only two divisors. It is called* **composite** *if it is bigger than 1 and is not prime.*

The first several prime numbers are: $2, 3, 5, 7, 11, 13, 17, 19, \ldots$. Only one of the prime numbers is even, of course. Some people say that the even prime number is the oddest prime number! There is a reason for that saying, but we will not be able to appreciate it in our (quite elementary) course on Number Theory.

Exercise 1.3 *1) Is the number 999991 prime?*
2) Is it true that 3 never divides $n^2 + 1$, and that 5 never divides $n^2 + 2$? Give arguments for your answer.
[RAA: suppose $3 \mid n^2 + 1$; then $n^2 + 1 = 3m$ for some natural number m. From $n^2 = 3m - 1$ we get that $3 \nmid n^2$, and so, $3 \nmid n$. Therefore $n = 3k + 1$ or $n = 3k + 2$. We have that $n^2 = 3A + 1$ where $A = 3k^2 + 2k$ in the former case, and $n^2 = 3B + 1$ where $B = 3k^2 + 4k + 1$ in the latter case. Since neither $3A + 1 = 3m - 1$ nor $3B + 1 = 3m - 1$ is possible, we get a contradiction. The second claim of the Exercise is treated similarly.]
3) Let N be a two digit natural number, and let M be the number obtained by reordering the digits of N. Show that $9 \mid |M - N|$, and find all N such that $|N - M| = 18$.
4) Show that the product of three consecutive integers is divisible by 6, and that the product of four consecutive integers is divisible by 24.
5) Prove that, for $n \in \mathbb{N}$, we have that $6 \mid n(n-1)(2n-1)$.
[Notice that $n(n-1)(2n-1) = n(n-1)(2n-4+3) = n(n-1)(2n-4) + 3n(n-1) = 2n(n-1)(n-2) + 3n(n-1)$, and that the summands in the last expressions are divisible by 6.]
6) Prove that if neither $2 \mid n$ nor $3 \mid n$, then $24 \mid (n^2 + 23)$.
[Prove first that 3 and 8 divide $n^2 - 1$ for n non-divisible by 2 and 3.]
7) Find all natural numbers n such that $n + 1 \mid n^2 + 1$.
8) How many natural numbers, less than 100, are there such that neither 2, nor 3, nor 5 divides them?
9) Find ten consecutive composite natural numbers. For every $n > 1$, find n consecutive composite natural numbers.

1.1.3 Other Special Types of Natural Numbers

The natural numbers have been partitioned in subsets according to a variety of features that they may or may not posses. Examples of such are prime/composite numbers or even/odd numbers. We list in this sub-section other examples.

• **Squares** (numbers of the type n^2 for a natural n), **cubes** (n^3 for $n \in \mathbb{N}$), **k**th powers (n^k for $k, n \in \mathbb{N}$);

•**Square-free** numbers, that is numbers which are not divisible by the square of any non-identity natural number. For instance, 6 is square-free, while 12 is not;

• **Triangular**, and **square** numbers. These have more geometric flavour in their definitions. Arithmetically, the former are natural numbers of the form $n(n + 1)/2$, and the latter - natural numbers of the form n^2 where $n \in \mathbb{N}$;

• **Perfect numbers**: these have a property which "balances" the two operations, addition and multiplication, on them. More specifically, these are natural numbers which are the sum of their

proper divisors! Notice that divisors are defined by using multiplication, and the perfect numbers are the sum of their proper divisors! Examples are $6 = 1 + 2 + 3$ and $28 = 1 + 2 + 4 + 7 + 14$;

• **Prime numbers of type "one-mod-four" (1-mod-4)**. These are numbers which appear naturally in different considerations in Number Theory.

• **Fibonacci numbers**. The consideration of these numbers is inspired by natural sciences, such as Biology. The story goes that Fibonacci got interested in these numbers watching how the number of rabbits in consecutive generations grows. Later on Fibonacci numbers were associated with many other patterns appearing in Nature. Arithmetically, these numbers are defined inductively as follows

$$F_0 = 0, \quad F_1 = 1, \qquad\qquad F_{n+2} = F_{n+1} + F_n.$$

Often numbers are considered in pairs or in triplets. Here are some examples.

• **Twin primes**. Looking in the sequence of primes,

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots,$$

we see that there are pairs of prime numbers the difference between which is 2. Examples are:

$$(3, 5), (5, 7), (17, 19), (29, 31), \quad \text{etc.}$$

A prime number p is called a **twin prime** if either $p - 2$ or $p + 2$ is a prime number as well.

• **Prime triplets**. These are triplets, such as $(3, 5, 7)$, consisting of two pairs of twin primes.

• **Pythagorean triplets**. These are triplets of natural numbers (a, b, c) such that $a^2 + b^2 = c^2$. There are infinitely many such triplets. The most celebrated among these is, may be, $(3, 4, 5)$ for the reason that, using it, people in ancient Egypt could construct a right angle, and divide the fertile lands around Nile River into rectangular pieces, and return it to the people working on it, after the seasonal flooding of the river. The importance of the Pythagorean triplets stems also from the way they inspired Pierre de Fermat to consider triplets (a, b, c) such that $a^n + b^n = c^n$ for a natural number $n \geq 3$. The attempts to prove that there are NO such triplets (fact claimed by Fermat, and known as **The Last Fermat Theorem**), led to the development of modern Algebra, and Algebraic Geometry!

Exercise 1.4 *1) Prove that if x and y are odd natural numbers, then $x^2 + y^2$ is never a perfect square.*

[We have $x = 2u+1, y = 2v+1$. Therefore $x^2+y^2 = 4(u(u+1)+v(v+1))+2$ which is an even number not divisible by 4...]

2) Prove that no natural number, bigger than 1, whose expression in base 10 has only 1s as digits is a perfect square.

[If the number has n digits, then it is equal to $1 + 10 + \cdots + 10^{n-1} = (10^n - 1)/9$. Prove that $10^n = (3m)^2 + 1$ is an impossible relation for any natural numbers m, and $n \geq 2$.]

3) Prove that every odd prime number can be written as a difference of two squares: $p = a^2 - b^2$. Prove also that this presentation is unique. Is such a presentation possible if p is just an odd natural number? Can 2 be represented this way?

4) Prove that for no $n \in \mathbb{N}$ is the number $3n^2 - 1$ a perfect square.

5) Find all natural numbers m, n such that $1/m + 1/n$ is a natural number.

6) Show that the product of any two, three, four, or five consecutive positive integers is never a perfect square. (It is true that no product on $n \geq 2$ consecutive positive integers is a square (P. Erdös, 1935). Moreover, no such a product is even a perfect $m \geq 2$ power (Erdös-Selfridge, 1975).)

7) Prove that, for every $\pm 1 \neq n \in \mathbb{Z}$, the number $n^4 + 4$ is composite.

8) Compute the value of the expression

$$\frac{(10^4 + 324)(22^4 + 324)(34^4 + 324)(46^4 + 324)(58^4 + 324)}{(4^4 + 324)(16^4 + 324)(28^4 + 324)(40^4 + 324)(52^4 + 324)}.$$

1.1.4 Some Number Theoretical Questions

Despite the elementary level of our knowledge of numbers at the moment, we may still ask questions about them, some of which are quite non-elementary! Here are some examples:

Q_1 : Are the prime numbers infinitely many?

The answer is yes, and was known to Euclid. We will address this question in Chapter 4.

Q_2 : Are the twin primes infinitely many?

Q_3 : Are the triplets of primes infinitely many?

The answer to Q_3 is easily NO. On the other hand, Q_2 is still a very widely open question! The experts do not have a doubt that the twin primes are infinitely many!

Q_4 : How often do the prime numbers appear in the sequence of natural numbers?

Notice that, for any natural n, there is a sequence of n consecutive natural numbers, none of which is prime!

Q_5 : A natural question is which square numbers are triangular as well? In other words, are there non-zero natural numbers m, n such that $m^2 = n(n+1)/2$?

There is a way to find all such numbers, but we will not discuss it in this course.

Q_6 : Which natural numbers are sums of two squares? As an intermediate question: which **primes** are sums of two squares? For example, $2 = 1^2 + 1^2, 5 = 1^2 + 2^2, 13 = 2^2 + 3^2$. But $3, 7$ and 11 are NOT sums of two squares!

Q_7 : Which numbers can be the third component in a Pythagorean triplet (a, b, c)?

Chapter 10 of these Notes is devoted to answering Q_6 and Q_7 in full.

Q_8 : Are the perfect numbers infinitely many?

The last question is still open.

1.2 Common Factors, and Relatively Prime Natural Numbers

In this course we will be comparing different natural numbers from the point of view of their divisors. This leads to the notion of common divisors of two or more numbers.

Definition 1.2.1 *The natural number d is a* **common divisor** *of the natural numbers $n_1, \cdots, n_k$ if it divides all of them: $d \,|\, n_1, \ldots, d \,|\, n_k$. We say in this case that $n_1, \ldots, n_k$* **have/share a common factor**.

Obviously, the number 1 is a common divisor of any two natural numbers.

Definition 1.2.2 *The natural numbers $n_1, \ldots, n_k$ are called* **relatively prime** *if the only common divisor they have is* 1.

For instance, 5 and 26 are relatively prime, while 33 and 1001 are not (check out that 11 is a common factor!).

Definition 1.2.3 *The natural numbers $n_1, \dots, n_k$ are* **pairwise relatively prime** *if every two of them are relatively prime.*

For instance, the numbers $2, 3, 4$ are relatively prime, but are NOT pairwise relatively prime.

1.3 The Integers, the Rational Numbers

We may try to stick to using only the natural numbers in our considerations, but we will not be able to keep up with this for long! The reason is that some of the questions in Number Theory, even at the elementary level of our course, can only be answered by using techniques related to other number systems, such as integer numbers and rational numbers, as well. And even when we can do the things without the help of other number systems, using the latter makes the considerations more elegant and illuminating.

The first extension of the natural numbers leads us to the set of **integers**

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \dots, \pm n, \dots\}.$$

The two operations and the relation from $\mathbb{N}$ extends to $\mathbb{Z}$. We have in this set the neutral element for addition as well, i.e. 0. But we have much more than that! Every element has an **opposite**, that is every m has (a unique) n such that $m + n = 0$. In particular, every equation of the form

$$x + m = n$$

for $m, n \in \mathbb{N}$ has a (unique) solution in $\mathbb{Z}$. By the way, the set of integers is a **ring**. That is, the two operations are commutative and associative, they have neutral elements, the product distributes over the addition, and every element has also an opposite one. In our course the ring $\mathbb{Z}$ will play a fundamental role.

The notion of divisibility extends naturally to the set of integers. We say that the integer m **divides** the integer n, if there is an integer s such that $n = m \cdot s$. Notice that, once again, everybody divides 0, and that 0 divides only 0. (Prove that!) As it is in the case of natural numbers, the divisibility is a non-trivial relation (a partial order) between integers as well: there are non-zero integers which do not divide each other.

The total order $\leq$ on $\mathbb{N}$ extends to one on $\mathbb{Z}$ as well: by definition, every negative integer is less than every positive integer; 0 separates the positive from negative numbers, being less than the former, and bigger than the latter; if m, n are negative integers, then $m < n$ if for their opposites we have $(-n) < (-m)$. This order is no longer a well order!

The second extension of the natural numbers, which is actually an extension of the integers, is the set of **rational numbers**

$$\mathbb{Q} = \{\, m/n \,|\, m \in \mathbb{Z},\, n \in \mathbb{N} \setminus \{0\} \,\}$$

where by definition $m/n = p/q$ if $m \cdot q = n \cdot p$. The expressions like m/n for $m \in \mathbb{Z}$ and $n \in \mathbb{N} \setminus \{0\}$ are called **rational fractions**. Note that rational numbers are **classes** of equal rational fractions. Every rational fraction **represents** its rational number. For instance, $(-1)/2$ and $(-3)/6$ represent the same rational number. As we will prove in Chapter 4, every rational number has a **unique**, a.k.a. **canonical**, representation as a **reduced** rational fraction, that is a fraction m/n such that m and n share no common factors different from ± 1. In particular, every integer n, considered as a rational number, is represented by the reduced fraction $n/1$. The operations addition and multiplication are defined by

$$m/n + p/q = (m \cdot q + n \cdot p)/(n \cdot q) \qquad m/n \cdot p/q = (m \cdot p)/(n \cdot q).$$

The order on $\mathbb{Q}$, inherited from $\mathbb{Z}$, and, ultimately, from $\mathbb{N}$, is defined by (recall that $n, t \in \mathbb{N} \setminus \{0\}$)

$$m/n \leq s/t \qquad \text{if} \qquad m \cdot t \leq s \cdot n.$$

The set of rational numbers $\mathbb{Q}$ is a **field**, that is, $\mathbb{Q}$ is a ring every non-zero element of which has a reciprocal:

$$(\forall x \in \mathbb{Q})(x \neq 0 \Rightarrow ((\exists y \in \mathbb{Q})(x \cdot y = 1))).$$

Now, we can solve any linear equation

$$a \cdot x + b = c$$

where $a \neq 0, b$ and c are rational numbers.

Fields that one knows from pre-school times are the set of **real numbers**, $\mathbb{R}$, and the set of **complex numbers**, $\mathbb{C}$. We have the following inclusions

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

All inclusions are strict! All the sets are infinite. In this course, we will introduce, along the way, and study finite rings and finite fields as well.

As the reader may know, the real numbers are an extension of the rational numbers, and are constructed for the needs of Calculus. They play important role in studying different types of (continuous real valued) functions (of real variables).

What the reader might not know is that, for every prime number p, there is an extension of the rational numbers to a field denoted by $\mathbb{Q}_p$ and called **p-adic numbers** . This field plays a similar role to that of the real numbers (providing notions of limits and continuous functions), but is more number theoretical in nature. We will discuss briefly the p-adic numbers in a couple of Vistas later in the Notes.

Exercise 1.5 *1) Let $a, b, c \in \mathbb{Z}$. Prove that if $a \,|\, b$ and $a \,|\, c$, then for every $u, v \in \mathbb{Z}$ we have $a \,|\, (ub + vc)$.*
2) Let $m = ua + vb$ where $a, b, u, v \in \mathbb{Z}$. Prove that if d is a common divisor of a and b, then $d \,|\, m$. Conclude that if $ua + vb = \pm 1$, then a and b are relatively prime.
3) Suppose $a, b, c \in \mathbb{Z}$ such that $a \neq 0$. Prove that $ab \,|\, ac$ if, and only if, $b \,|\, c$.
4) Suppose $a, b \in \mathbb{Z}$ such that $|b| < |a|$. Prove that $a \,|\, b$ if, and only if, $b = 0$.
5) Prove that $\sqrt{2}$ is not a rational number. That is, there is no rational number m/n such that $(m/n)^2 = 2$. Do the same for $\sqrt{3}$ as well.

[RAA: assume $(m/n)^2 = 2$ for some rational number m/n. W.L.O.G. we may assume the fraction reduced, that is m and n share no positive factors different from 1. We have then that $m^2 = 2n^2$ which implies that m is even: $m = 2m_1$. Substituting in the equality and simplifying we get $2m_1^2 = n^2$ which in turn implies that n is even as well: $n = 2n_1$. So, m and n do share a bigger than 1 factor - a contradiction! The second part of the Exercise is treated in a similar way.]

Chapter 2

Pythagorean Triplets

As a first example of a non-trivial number theoretical problem we are solving the Pythagorean equation

$$x^2 + y^2 = z^2$$

in integers.

2.1 Reduction to Solving in Rational Numbers

Let the triplet of integers (a, b, c) be a solution of the Pythagorean equation, that is $a^2 + b^2 = c^2$. It is easy to find all such triplets for which $a \cdot b \cdot c = 0$. Indeed, all solutions in this case are given by

$$(0,\, \epsilon_1 \cdot t,\, \epsilon_2 \cdot t) \qquad \text{and} \qquad (\epsilon_1 \cdot t,\, 0,\, \epsilon_2 \cdot t)$$

where $\epsilon_1 = \pm 1$, $\epsilon_2 = \pm 1$, and $t \in \mathbb{N}$. So, without a loss of generality (W.L.O.G.), we may assume that $a \cdot b \cdot c \neq 0$. The next assumption is that the components of the triplet (a, b, c) are all **positive**. Indeed, any such triplet (a, b, c) determines eight distinct solutions,

$$(\epsilon_1 \cdot a, \epsilon_2 \cdot b, \epsilon_3 \cdot c) \qquad \text{for} \qquad \epsilon_i = \pm 1,$$

and any triplet with non-zero components is obtained this way (from the triplet $(|a|, |b|, |c|)$).

2.1.1 Reduction to Solving a Hyperbolic Equation

Observe that a triplet (a, b, c) with $a \neq 0$ determines a solution in rational numbers, $(b/a, c/a)$, to the equation

$$1 + X^2 = Y^2.$$

This equation is an equation of a hyperbola in the plane, only we have to find the **rational** points on it. We will call the equation **hyperbolic** in order to distinguish it from the Pythagorean one.
Now, a solution $(p/q, r/s)$ to the hyperbolic equation produces many different solutions to the Pythagorean equation. Indeed, if $u \in \mathbb{N}$ is such that $p/q \cdot u \in \mathbb{N}$ and $r/s \cdot u \in \mathbb{N}$, then the triplet

$$(u,\, p/q \cdot u,\, r/s \cdot u)$$

is a solution to the Pythagorean equation. For some pair $(p/q, r/s)$ and a number u we will get the original solution (a, b, c).
Let $u = q \cdot s$. Then we get the Pythagorean triplet (qs, ps, rq). The discussion in the previous paragraph can be restated as follows:

Every Pythagorean triplet (a, b, c) can be obtained from a triplet (qs, ps, rq), where $(p/q, r/s)$ is a solution to the hyperbolic equation, by multiplying all the components

thereof by a natural number, or by dividing those components by a non-zero common factor.

The moral here is that to find all Pythagorean triplets, it is enough to find all solutions in $\mathbb{Q}$ of the hyperbolic equation. To do this, as we will see below, is very easy.

2.1.2 Solving the Hyperbolic Equation

We have

$$1 = Y^2 - X^2 = (Y - X)(Y + X).$$

We are looking for rational solutions (u, v), so both $v + u$ and $v - u$ are rational numbers whose product is 1. They both are also positive, because we are looking for positive solutions to the Pythagorean equation. If we call $r = u + v$, for a rational number r, then $v - u = 1/r$. Obviously $r = u + v > v - u = 1/r$, and therefore $r > 1$. We easily solve then that

$$v = \frac{1}{2}\left(r + \frac{1}{r}\right) \qquad u = \frac{1}{2}\left(r - \frac{1}{r}\right).$$

Let $r = p/q$ for positive $p > q$. Then we have

$$v = \frac{1}{2}\left(\frac{p}{q} + \frac{q}{p}\right) = \frac{p^2 + q^2}{2pq} \qquad u = \frac{1}{2}\left(\frac{p}{q} - \frac{q}{p}\right) = \frac{p^2 - q^2}{2pq},$$

and the solutions in positive integers to the Pythagorean equation are given by

$$(a, b, c) = (u \cdot 2pq, u \cdot (p^2 - q^2), u \cdot (p^2 + q^2)) \qquad u \in \mathbb{N} \setminus \{0\}$$

or

$$(a, b, c) = (2pq/d, (p^2 - q^2)/d, (p^2 + q^2)/d) \qquad d \in \mathbb{N} \setminus \{0\}$$

for positive integers $p > q$, and d a common factor of the three components.

This result tells us how to find all Pythagorean triplets in principle. In the next subsection, we are discussing an interesting property of the Pythagorean triplets which will give us a much better way, it will actually provide us with a formula, for finding all Pythagorean triplets.

2.1.3 Primitive Pythagorean Triplets

Among all Pythagorean triplets (a, b, c) with positive components, there are "minimal" ones in the following sense. If the components of (a, b, c) share a common factor: $a = da_1, b = db_1, c = dc_1$, then the triplet (a_1, b_1, c_1) is also a solution with positive components. The new solution is naturally "smaller" than the original one. If the components of (a, b, c) do not share a bigger than 1 common factor, then it is impossible to produce a smaller one, and such are naturally minimal solutions. In fact, **any solution** (a, b, c) **can be obtained from a minimal one by multiplying the latter by an appropriate natural number**. Therefore, knowing the minimal Pythagorean triplets we can find all solutions to the Pythagorean equation. The amazing thing is that there is a formula exhibiting all minimal triplets. This subsection is devoted to obtaining that formula.

Definition 2.1.1 *A solution* (a, b, c) *to the Pythagorean equation with positive components which do not share a bigger than* 1 *common factor is called a* **primitive Pythagorean triplet** *or, for short PPT.*

Obviously, to know all solutions to the Pythagorean equation, we need to know the primitive Pythagorean triplets. Turns out, every primitive Pythagorean triplet can be expressed in the form that we know from the previous subsection.

Proposition 2.1.1 *Let (a, b, c) be a primitive Pythagorean triplet. Then, we have*

$$\textit{either} \qquad (a, b, c) = (2pq, p^2 - q^2, p^2 + q^2) \qquad \textit{or} \qquad (a, b, c) = (p^2 - q^2, 2pq, p^2 + q^2).$$

Moreover, the numbers $U = p + q$ and $V = p - q$ are odd and do not share common factors bigger than 1*.*

Proof The proof is divided in three steps.
Step 1: For a PPT (a, b, c), at least one of its components is an odd number (for, otherwise, all the numbers have to be even, and so sharing a bigger than 1 factor). Let's observe that the component c can not be even. Indeed, otherwise the components a and b need to have the same parity, and since they can not be both even (the triplet (a, b, c) is a PPT!), then $a = 2a_1 + 1$ and $b = 2b_1 + 1$ have to be odd. But then, we would have

$$c^2 = (2c_1)^2 = (2a_1 + 1)^2 + (2b_1 + 1)^2 = 4(a_1(a_1 + 1) + b_1(b_1 + 1)) + 2$$

and therefore

$$4(c_1^2 - a_1(a_1 + 1) - b_1(b_1 + 1)) = 2$$

which is impossible, because 4 does not divide 2.
Step 2: So, c is odd, and either a or b, but not both, is even. W.L.O.G., we may assume that $a = 2a_1$. We have now that

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

The numbers on the RHS are both odd. We are showing that they do not share factors bigger than one. Indeed, if $d \,|\, c - a$ and $d \,|\, c + a$, then d is an odd number which also divides $(c - a) + (c + a) = 2c$ and $c + a - (c - a) = 2a$. But, being odd, d has to divide a and c which is possible only when $d = 1$. We are showing next that both $c - a$ and $c + a$ are squares. Indeed, since $b \,|\, b^2$ we have that $b \,|\, (c - a)(c + a)$, and so, $b = b_1 b_2$ such that $b_1 \,|\, c - a$ and $b_2 \,|\, c + a$. But then

$$b_1 \cdot b_2 = b = \frac{c - a}{b_1} \cdot \frac{c + a}{b_2}$$

where the fractions on the RHS are integers. Observe now that no bigger than 1 divisor of b_1 divides $c + a$. This is, because if $d \,|\, b_1$, then $d \,|\, c - a$ as well, and, assuming $d \,|\, c + a$, we get that d is a common divisor of $c - a$ and $c + a$ which forces $d = 1$. Similarly, no bigger than 1 divisor of b_2 divides $c - a$. From the last equality above we get that

$$b_1 \,|\, \frac{c - a}{b_1} \qquad \text{and} \qquad b_2 \,|\, \frac{c + a}{b_2}.$$

But then both $(c - a)/b_1^2$ and $(c + a)/b_2^2$ are integers whose product is equal to 1. Therefore

$$c - a = b_1^2 \qquad \text{and} \qquad c + a = b_2^2$$

as claimed.
Step3: Denote $U = b_2$ and $V = b_1$, and observe that $U > V > 0$ are two odd numbers sharing no bigger than 1 divisors. Let now $p = (U + V)/2$, and $q = (U - V)/2$. simple computation shows that

$$(a, b, c) = \left(\frac{U^2 - V^2}{2}, UV, \frac{U^2 + V^2}{2} \right) = (2pq, p^2 - q^2, p^2 + q^2). \quad \square$$

Turns out, the restriction on U and V above ensures that the triplet in the middle of the last equalities is primitive!

Proposition 2.1.2 *If $U > V$ are relatively prime odd natural numbers, then*

$$(a, b, c) = \left(\frac{U^2 - V^2}{2}, UV, \frac{U^2 + V^2}{2} \right)$$

is a primitive Pythagorean triplet.

Proof Let $d > 0$ be a common factor of a, b and c. Then

$$a = da_1 \qquad b = db_1 \qquad c = dc_1,$$

and therefore

$$U^2 = a + c = d(a_1 + c_1) \qquad \text{and} \qquad V^2 = a - c = d(a_1 - c_1)$$

are divisible by d. So, U^2/d and V^2/d are integers. **We are showing next that d is square-free**. Indeed, if $d = d_1^2 d_2$ for some $d_1 > 1$, then

$$\left(\frac{U}{d_1}\right)^2 \cdot \frac{1}{d_2} \qquad \text{and} \qquad \left(\frac{V}{d_1}\right)^2 \cdot \frac{1}{d_2}$$

are integers. Therefore both U/d_1 and V/d_1 must be integers, which makes d_1 a common factor of U and V. By our assumption about U and V, they do not share bigger than 1 factors. So, $d_1 = 1$ which contradicts our assumption about d_1. The number d is square-free. Since $U^2/d = (U \cdot U)/d$ is an integer, part of d, say d', has to cancel out with a part of the first U and the rest of d, say $d'' = d/d'$, has to cancel out with the second U in the numerator of that fraction. So, both d' and d'' divide U. **We are showing next that the product $d' \cdot d''$ also divides U.** We have $U = d' \cdot U_1$ and $d'' \,|\, d' \cdot U_1$. But, since d is square-free, the numbers d' and d'' are relatively prime - share no bigger than 1 factors. Therefore $d'' \,|\, U_1$. This proves that $d' \cdot d'' \,|\, U$. That is $d \,|\, U$. A similar argument shows that $d \,|\, V$ as well, so d is a positive common factor of U and V, and therefore, $d = 1$. □

We summarize the results from this subsection in a theorem.

Theorem 2.1.3 *The primitive Pythagorean triplets are given by the formula*

$$(a, b, c) = ((U^2 - V^2)/2,\, UV,\, (U^2 + V^2)/2)$$

or, due to the fact that (b, a, c) is also a Pythagorean triplet, by the formula

$$(a, b, c) = (UV,\, (U^2 - V^2)/2,\, (U^2 + V^2)/2)$$

where $U > V > 0$ are relatively prime odd integers. Any solution (a', b', c') to the Pythagorean equation with $a'b'c' \neq 0$ has the form

$$(a', b', c') = (\epsilon_1 da, \epsilon_2 db, \epsilon_3 dc)$$

where (a, b, c) is a primitive Pythagorean triplet, d is a positive integer, and $\epsilon_1, \epsilon_2, \epsilon_3$ take on values ± 1 independently of each other. The rest of the solutions, those with $a'b'c' = 0$ have the form $(t, 0, \pm t)$ or $(0, t, \pm t)$ for $t \in \mathbb{Z}$.

Exercise 2.1 *Prove that, in the notations of the theorem, $(U^2 - V^2)/2$ is an even number, and that $(U^2 + V^2)/2$ is odd. Conclude that the two formulae for the primitive Pythagorean triplets are distinct.*

The right triangle with integer sides are called **Pythagorean triangles**. The triangle with side-lengths $(3, 4, 5)$ is arguably the most famous Pythagorean triangle (known to ancient Egyptians etc.). This triangle is special mathematically in at least two ways as the following exercise asks you to show.

Exercise 2.2 *(1) Consider the set PT of all Pythagorean triplets. For $(a, b, c) \in PT$ we assume $a \leq b \leq c$. Then,*

$$(\forall (a, b, c) \in PT)(a \cdot b \cdot c \geq 60 \quad \wedge \quad a \cdot b \cdot c = 60 \Leftrightarrow (a, b, c) = (3, 4, 5)).$$

[Hint: Prove that for every $(a, b, c) \in PT$ we have that $3 \mid a \cdot b$, and that $5 \,|\, a \cdot b \cdot c$. Conclude that $60 \,|\, a \cdot b \cdot c$. Observe also that $a \cdot b \cdot c = 30 \to (a, b, c) = (3, 4, 5)$.]

(2) Prove that if $(a, b, c) \in PT$ is such that a, b, c form an arithmetic progression, then $(a, b, c) = (3k, 4k, 5k)$ for some $k \in \mathbb{N} \setminus \{0\}$. Conclude that $(3, 4, 5)$ is the only PPT whose components form such a progression.

2.2 A Geometric Method

Many deep problems in number theory are proved by using geometric methods. A really simple case of use of geometry is explained below on the example of solving the Pythagorean equation in rationals.

2.2.1 Reducing to solving an Elliptic Equation

Going back to reducing to solving the Pythagorean equations in rationals, we assume that the solutions (a, b, c) we are after are with non-zero components, $abc \neq 0$, so we may divide by any of these components. Dividing by a we got to solving the hyperbolic equation above. The same would be the result if we divide by b. What will happen if we divide by c? The answer is straightforward: $(a/c, b/c)$ is a solution to the **elliptic** equation (the equation of a circle!)

$$X^2 + Y^2 = 1.$$

Using the arguments from the previous section, we see that to find all solutions to the Pythagorean equation, it is enough to find all solutions to the elliptic equation in **rational** numbers. Contrary to the hyperbolic case, we can not factor out $X^2 + Y^2$ in order to find the solutions. Here is where geometric methods help.

2.2.2 Solving the Elliptic Equation Using Secant Lines

The real solutions to the elliptic equation are all well known: they form a circle, $\kappa((0,0),1)$, of radius 1 centred at $(0,0)$. All such points have the form $(\cos t, \sin t)$ for $t \in [0, 2\pi)$. This equation has some obvious rational solutions: $(-1,0), (1,0), (0,1)$, and $(0,-1)$. So, let's find the rest of the rational solutions.

Suppose (x_1, y_1) is one such solution. The points $(-1,0)$ and (x_1, y_1) define a line which is a secant line to the circle κ. The point-slope equation of this secant line is

$$Y = m_1(X+1) \qquad \text{where} \qquad m_1 = \frac{y_1}{x_1+1} \in \mathbb{Q}.$$

Observe that if (x_2, y_2) is another rational point on κ, then the slopes of the secant lines determined by them are distinct $m_1 \neq m_2$. Conversely, we have

Proposition 2.2.1 *Suppose $m \in \mathbb{Q}$, and consider the line l_m through $(-1,0)$ with slope m. Then*

$$l_m \cap \kappa = \{(-1,0), P_m\}$$

where P_m has rational coordinates.

Proof Indeed, $P_m((1-m^2)/(1+m^2), (2m)/(1+m^2))$. Verify that as an exercise. □

To get from this result the known formulae for the solutions to the Pythagorean equation is an easy exercise.

Note that the geometric method is, in a sense, more powerful than the one we used in the hyperbolic equation case: we can apply the geometric method there to find the rational solutions too, while the method from that case is not (directly) applicable in the elliptic case.

Exercise 2.3 *(1) Noticing that the points $(0, \pm 1)$ are rational solutions to $1 + X^2 = Y^2$ apply the geometric method to find the rational solutions to this equation.*
(2) Find all solutions to $x^2 + y^2 = 2z^2$ in integers.
(3) Do the same as in (2) for $x^2 + y^2 = 3z^2$.

(4) Consider the equation $x^2 + y^2 = nz^2$. Prove that there are solutions in positive integers when n is a sum of two squares. When the latter is a case, find all integer solutions to this equation.
(5) Do the same as in (2) for $x^2 + 2y^2 = z^2$.
(6) Do the same as in (2) for $x^2 + 3y^2 = z^2$.
(7) Prove that for no positive natural number n is the expression $n^4 + 2n^3 + 2n^2 + 2n + 1$ a perfect square.
(8) Prove that there are infinitely many PPTs whose even component is a perfect square. Find such triplets.

2.3 Vista: The Equation $ax^2 + by^2 + cz^2 = 0$

Having solved the Pythagorean equation in the previous section, it is natural to ask whether the more general equation in the title of this section has solutions, and, when the answer is yes, how to find all solutions. It is no loss of generality if we ask only about existence of **non-trivial** solutions, that is, solutions (x_0, y_0, z_0) such that $x_0y_0z_0 \neq 0$. Similarly, since the case when $abc = 0$ is easy to handle, we may assume that $abc \neq 0$ as well. Obviously, non-trivial solutions exist only if the coefficients a, b, and c are not all of the same sign (which reduces to solving the equation with $a, b > 0$ and $c < 0$). Other reductions, some of whose obvious, others familiar from our considerations in the previous section) are the following

- a, b and c are square-free integers (why?);
- neither a, b and c nor x_0, y_0 and z_0 share positive common factor different from 1;
- actually no two of a, b and c share common factor bigger than 1.

Observe that the last condition is equivalent to abc being square free.

2.3.1 Legendre's Theorem

The name of Legendre will be used often in these Notes. Thus, we will learn about the **Legendre symbol**, and will prove the **Legendre-Gauss Law of Quadratic Reciprocity** - one of the centrepieces of these Notes, as well as of classical Number Theory.

In 1785, Legendre proved the precise conditions when the equation we are considering has a solution in integers. As André Weil puts it, this theorem is "one of Legendre's main claims to fame".

Before formulating the theorem, a bit of terminology. We say that the integer s is a **quadratic residue modulo** the integer t if there is an integer u such that $t \mid u^2 - s$.

Theorem 2.3.1 (*Legendre*, 1785) *Let a, b and c be tree integers, not all of the same sign, and such that abc is square-free. Then, the equation*

$$ax^2 + by^2 + cz^2 = 0$$

has a solution in integers, not all 0, if, and only if, $-ab, -bc$, and $-ca$ are quadratic residues modulo $|c|, |a|$, and $|b|$ respectively.

We will give a proof of this theorem in Chapter 6, after we learn about the Chinese Remainder Theorem. Chapter 8 is devoted to the theory of quadratic residues (modulo n). The name of Legendre will appear there again.

2.3.2 Finding All Solutions to $ax^2 + by^2 + cz^2 = 0$

One can find all solutions to the equation under investigation using the geometric method. Indeed, being interested in non-trivial solutions only, and having $abc \neq 0$, one reduces solving the equation $ax^2 + by^2 + cz^2 = 0$ in integers to solving the equation

$$\alpha X^2 + \beta Y^2 = 1$$

in rational numbers where $\alpha = -a/c$,and $\beta = -b/c$ are (non-zero) rational numbers, and $X = x/z, Y = y/z$. The latter can be done, as we know, using lines with rational slopes through a rational solution (X_0, Y_0) (the existence of which is checked using Legendre's theorem).

Chapter 3

The Division Algorithm in $\mathbb{Z}$, Applications

3.1 Division with Quotient and Remainder in $\mathbb{Z}$

The fact that the integers $\mathbb{Z}$ are not a field (that is, there are non-zero integers none of which divides the other one) makes Number Theory a very interesting and highly non-trivial subject. The most fundamental property of the integers is that they **can be divided with quotient and remainder**. In technical terms, this property makes them form an **Euclidean ring** (one can learn more about these rings in Topics in Algebraic Structures). As we will see below, this fact is based on the Least Element Property enjoyed by the natural numbers $\mathbb{N}$.

Theorem 3.1.1 *Suppose $m \neq 0, n \in \mathbb{Z}$. There exist unique $q, r \in \mathbb{Z}$ such that*

$$n = m \cdot q + r \qquad\qquad 0 \leq r < |m|.$$

Proof We prove first the **existence** of q and r. To this end, consider the set of integers

$$S = \{s \,|\, (\exists l \in \mathbb{Z})(s = n + l \cdot m)\}.$$

It's easy to see that S has both positive and negative elements. So, the set $\Sigma = S \cap \mathbb{N}$ is a non-empty subset of $\mathbb{N}$, and has therefore a least element $s_0 = n + l_0 \cdot m$.
The first observation is that $s_0 < |m|$. Indeed, if not

$$s_0 - |m| = n + (l_0 \pm 1) \cdot m \in \Sigma$$

the plus or minus sign depends on whether $m > 0$ or $m < 0$. By the minimality of s_0 we have that $s_0 \leq s_0 - |m|$ which, combined with the obvious $s_0 - |m| < s_0$, leads us to the absurd that $s_0 < s_0$!
Denote $q := -l_0$ and $r = s_0$ to get $n = q \cdot m + r$ as needed.
Uniqueness of q and r. Assuming there is another pair q' and r' with the same properties, we get

$$n = q \cdot m + r = q' \cdot m + r'$$

so that

$$(q - q') \cdot m = r' - r$$

and therefore

$$|q - q'| \cdot |m| = |r - r'|.$$

But $0 \leq r, r' < |m|$ enforces $|r - r'| < |m|$, so that

$$|q - q'| \cdot |m| = |r - r'| < |m|$$

and hence

$$0 < |m| \cdot (1 - |q - q'|).$$

The last inequality holds true only if $|q - q'| = 0$. So, $q = q'$ and therefore $r = r'$. The uniqueness is proved. □

Definition 3.1.1 *The number q, respectively r, from the theorem above is called the* **quotient**, *respectively - the* **remainder**, *of the division of n by m.*

Note that m divides n if, and only if, the remainder of the division of n by m is $r = 0$.

Exercise 3.1 *(1) Let $a > 1$ be an integer, and let $m \neq 0, n$ be natural numbers. Find the quotient and the remainder of the division of $a^n - 1$ by $a^m - 1$.*

[If $n < m$ the result is obvious. Suppose that $m \leq n$, and let $n = mq + r$ where $0 \leq r < m$. Show that in this case

$$a^n - 1 = (a^m - 1) \cdot a^r(a^{m(q-1)} + \cdots + a^m + 1) + a^r - 1. \quad]$$

(2) Do the same for $a^n + 1$ and $a^m + 1$.

3.2 The Greatest Common Divisor

The first application of $\mathbb{Z}$ being a Euclidean ring would be in proving that it is a **principal ideal ring** (one can learn more about these rings in Topics in Algebraic Structures). This is what we are doing in this section.

3.2.1 The Greatest Common Divisor of two integers

The main object of our discussion here is the greatest common divisor, gcd, of two integers. The considerations are important for our future work in the course. But they also serve as a toy model in our way of proving that $\mathbb{Z}$ is a principal ideal ring. We prove first the following

Theorem 3.2.1 *Let $m, n \in \mathbb{Z}$. Consider the set of integers $\{a \cdot m + b \cdot n \,|\, a, b \in \mathbb{Z}\}$. If this set contains a non-zero element, then there is a unique $d \in \mathbb{N} \setminus \{0\}$ such that*

$$\{a \cdot m + b \cdot n \,|\, a, b \in \mathbb{Z}\} = dZ := \{s \cdot d \mid s \in \mathbb{Z}\}.$$

In particular, $d|m$, $d|n$, and there are integers a_0, b_0 such that $d = a_0 \cdot m + b_0 \cdot n$.

Proof Existence of d. The set $\{a \cdot m + b \cdot n \mid a, b \in \mathbb{Z}\}$ contains a non-zero element if, and only if, at least one number m or n is non-zero. In such a case the set $\mathbb{N} \setminus \{0\} \cap \{a \cdot m + b \cdot n \mid a, b \in \mathbb{Z}\}$ is a non-empty subset of $\mathbb{N}$, and therefore has a least element d. Obviously, $d = a_0 \cdot m + b_0 \cdot n$ for some integers a_0 and b_0. Let's prove that $d \,|\, m$ and $d \,|\, n$. By way of contradiction, suppose d does not divide m. By the previous theorem, $m = q \cdot d + r$ where $0 < r < d$. Therefore,

$$r = m - q \cdot d = (1 - q \cdot a_0) \cdot m - (q \cdot b_0) \cdot n.$$

This implies that

$$r \in \{a \cdot m + b \cdot n \,|\, a, b \in \mathbb{Z}\} \cap \mathbb{N}$$

and so $d \leq r < d$ which is a contradiction! So, $d \,|\, m$. In a similar way we show that $d \,|\, n$ as well. We are ready to show that

$$\{a \cdot m + b \cdot n \,|\, a, b \in \mathbb{Z}\} = d\mathbb{Z}.$$

Indeed, $m = d \cdot m', n = d \cdot n'$, and hence every element $s = a \cdot m + b \cdot n$ of the set on the LHS has the form $s = (a \cdot m' + b \cdot n') \cdot d$, and is an element of the set on the RHS, $d\mathbb{Z}$. Conversely, every element $s \cdot d$ of the RHS has the form $(s \cdot a_0) \cdot m + (s \cdot b_0) \cdot n$ which makes it an element of the LHS. **Uniqueness of** d. It is enough to show that if $d\mathbb{Z} = d'\mathbb{Z}$ for some natural d, d', then $d = d'$. But this follows right away from the fact that, when the sets are equal, $d' \,|\, d$ and $d \,|\, d'$. □

Corollary 3.2.2 *The number d in the previous theorem has the property that*

$$(d\,|m) \quad \wedge \quad (d\,|\,n) \quad \wedge \quad (\forall s \in \mathbb{Z}\setminus\{0\})((s\,|\,m \wedge s\,|\,n) \to (s\,|\,d)).$$

In particular, if m or n is not zero, then d is the largest among the common divisors of m and n

Proof Indeed, we already know that d is a common divisor of m and n. If s is a divisor of the two numbers as well, then using the relation $d = a_0 \cdot m + b_0 \cdot n$ we immediately get that $s\,|\,d$. Suppose now that m or n is not zero. This means that $d \neq 0$, and is therefore a positive integer. As we just proved, for any common divisor s of m and n, we have $s\,|\,d$. So, $d = s \cdot d_1$. Taking absolute values of both sides, we get $|d| = |s| \cdot |d_1|$. Now, $d_1 \neq 0$, and so, $|d_1| \geq 1$. Therefore,

$$d = |d| = |s| \cdot |d_1| \geq |s| \cdot 1 = |s|.$$

We get from this that $s \leq d$ as claimed. □

We are giving now a name to the number d in the previous theorem.

Definition 3.2.1 *Suppose $m, n \in \mathbb{Z}$ are not both zero. Then their* **greatest common divisor**, *denoted by $gcd(m, n)$ or just by (m, n), is the largest integer which divides both numbers. That is,*

$$gcd(m,n)\,|\,m \quad \wedge \quad gcd(m,n)\,|\,n \quad \wedge \quad (\forall s \in \mathbb{Z})((s\,|\,m \wedge s\,|\,n) \to (s \leq gcd(m,n))).$$

By definition, $gcd(0, 0) = 0$.

The corollary above states in particular that the generator d of the Theorem **is** the gcd of m and n.

We had already the occasion to work with numbers not sharing bigger than 1 factors. Here is the official definition.

Definition 3.2.2 *The integers m, n are called* **relatively prime** *if $gcd(m, n) = 1$.*

The following theorem describes the most important properties of the greatest common divisor.

Theorem 3.2.3 *The following statements hold true.*
(1) For every two integers m, n the greatest common divisor exists and is unique. Also, $gcd(m, n) = gcd(|m|, |n|)$.
(2) For any pair of integers m, n, and for a natural number d we have $d = gcd(m, n)$ if, and only if,

$$d\,|\,m \quad \wedge \quad d\,|\,n \qquad \wedge \quad (\forall s \in \mathbb{Z})((s\,|\,m \wedge s\,|\,n) \to (s\,|\,d)).$$

(3) (Bézout's Identity) For some integers a, b

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

(4) For the integers m, n we have

$$gcd(m, n) = 1 \quad \Leftrightarrow \quad (\exists\, a, b \in \mathbb{Z})(a \cdot m + b \cdot n = 1).$$

(5) If $d = gcd(m, n) \neq 0$, then m/d and n/d are relatively prime integers:

$$gcd(m/d, n/d) = 1.$$

(6) We have $\gcd(d \cdot m, d \cdot n) = |d| \cdot \gcd(m, n)$.

Proof Item (1) is easy and is left as an exercise. For item (2) we have to show that the $\gcd(m,n)$ has the property there, and vice-versa, if d has that property, then $d = \gcd(m,n)$. That $\gcd(m,n)$ has that property is proved in the Corollary above. So, we have to show the "vice-versa" part of the claim. To this end, observe that since $d \,|\, m$ and $d \,|\, n$, it is a common divisor of the two integers, and therefore, by the Corollary above, $d \,|\, \gcd(m,n)$. We have also that $\gcd(m,n) \,|\, d$, because d satisfies the property in item (2), and because $\gcd(m,n)$ is a common divisor of m and n. Since both d and $\gcd(m,n)$ are natural numbers, the relations $d \,|\, \gcd(m,n)$ and $\gcd(m,n) \,|\, d$ imply that $d = \gcd(m,n)$. Item (3) is obvious. For item (4), if $gcd(m,n) = 1$, then, by the Bézout's identity, there are integers a, b such that $a \cdot m + b \cdot n = 1$. Conversely, if the latter relation is true, and s is a divisor of m and n, then s is a divisor of 1 as well, and so $s = \pm 1$. Therefore $gcd(m,n) = 1$. For item (5), notice that $a \cdot m + b \cdot n = d$ implies $a \cdot (m/d) + b \cdot (n/d) = 1$, and then apply item (4) to finish the argument. Item (6) is straightforward, and is left as an exercise. □

Remark 3.2.1 Item (2) of the theorem is significant: it characterizes the gcd of two integers only in terms of the divisibility operation. □

Exercise 3.2 *(1) Let a, b, c, d be integers such that $a = b \cdot c + d$. Prove that* $\gcd(a,b) = \gcd(b,d)$.
(2) Let $a, b, c \in \mathbb{Z}$
(i) Prove that if $a \,|\, bc$ and $\gcd(a,b) = 1$*, then $a \,|\, c$.*
(ii) Prove that if $a \,|\, c, b \,|\, c$ and $\gcd(a,b) = 1$*, then $ab \,|\, c$.*
[We have $c = a \cdot c_1 = b \cdot c_2$, and $u \cdot a + v \cdot b = 1$. Multiplying the latter equality by c_1 we get $u \cdot a \cdot c_1 + v \cdot b \cdot c_1 = c_1$ which can also be written as $b \cdot (u \cdot c_2 + v \cdot c_1) = c_1$. This implies that $b \,|\, c_1$, that is, $c_1 = b \cdot c_3$. We ultimately have $c = a \cdot c_1 = a \cdot b \cdot c_3$.]
(3) Let $a, m, n \in \mathbb{N}$. Show that if $(m,n) = 1$, then $(a, m \cdot n) = (a,m) \cdot (a,n)$. Show, by example, that this claim is not true when $(m,n) \neq 1$.
(4) For all $n \in \mathbb{N}$, prove that $(n^2 + 3n + 2, 6n^3 + 15n^2 + 3n - 7) = 1$.
(5) Let a and b be relatively prime integers. Show that
(a) $(a - b,\ a + b) = 1$ or 2; *(b) $(2a + b,\ a + 2b) = 1$ or 3;*
(c) $(a^2 + b^2,\ a + b) = 1$ or 2; *(d) $(a + b,\ a^2 - 3ab + b^2) = 1$ or 5.*
(6) For the integers a, b, c show that
(a) $(a, bc) = (a, (a,b)c)$; *(b) $(a, bc) = (a, (a,b)(a,c))$.*

3.2.2 Ideals of $\mathbb{Z}$, and Greatest Common Divisor of a Group of Integers

Ideals of $\mathbb{Z}$

Definition 3.2.3 *A non-empty subset I of $\mathbb{Z}$ is called an* **ideal of** $\mathbb{Z}$ *if the following holds true*
(1) $(\forall m,n)(m,n \in I \to m - n \in I)$ and
(2) $(\forall m,n)(m \in I \wedge n \in \mathbb{Z} \to mn \in I)$.

Example 3.2.1 (1) Every ideal I contains the zero element: $0 \in I$. Indeed, $I \neq \emptyset$ so there is a $m \in I$. But then, formally, $m, m \in I$, and therefore $0 = m - m \in I$.

(2) For every element m of an ideal I, the opposite, $-m$, is also in I. This is because $0, m \in I$ implies that $0 - m \in I$.

(3) $I = \mathbb{Z}$ and $I = \{0\}$ are ideals of $\mathbb{Z}$. An ideal I of $\mathbb{Z}$ which is not equal to $\mathbb{Z}$ is called a **proper ideal** of $\mathbb{Z}$. The ideal 0 is called the **trivial** ideal of $\mathbb{Z}$.

(4) Let a be an integer. The set $I = \{s \cdot a \,|\, s \in \mathbb{Z}\}$ is an ideal of $\mathbb{Z}$. It is denoted by $I = (a)$, and a is called a **generator** of I. Notice that $\{0\} = (0)$ and that $\mathbb{Z} = (1)$. Notice also that $(a) = (-a)$, so that we can assume, with no loss of generality, that generator of I is non-negative.

(5) Let $a_1, \ldots, a_s \in \mathbb{Z}$. Consider the set $\Sigma = \{u_1 a_1 + \cdots u_s a_s \,|\, u_1, \ldots u_s \in \mathbb{Z}\}$. Obviously, $0 \in \Sigma$, and so $\Sigma \neq \emptyset$. For any two elements $x = u_1 a_1 + \cdots + u_s a_s$ and $y = v_1 a_1 + \cdots + v_s a_s$ we have that

$$x - y = (u_1 - v_1)a_1 + \cdots (u_s - v_s)a_s$$

is an element of Σ. Finally, if $w \in \mathbb{Z}$, then

$$wx = (wu_1)a_1 + \cdots + (wu_s)a_s$$

is an element of Σ. Therefore Σ is an ideal of $\mathbb{Z}$. The elements $a_1, \ldots a_s$ are called **generators** of Σ, and we denote the fact that they are generators by writing $\Sigma = (a_1, \ldots, a_s)$. It is straightforward that

$$(a_1, \ldots, a_s) = (|a_1|, \ldots, |a_s|)$$

so that W.L.O.G. we may assume the generators are non-negative. □

Definition 3.2.4 *And ideal I of $\mathbb{Z}$ is called* **finitely generated** *if $I = (a_1, \ldots, a_s)$ for some integers $a_1, \ldots, a_s$. An ideal I of $\mathbb{Z}$ is called* **principal** *ideal if I has only one generator: $I = (a)$.*

In the previous subsection, where we discussed the gcd of two integers, we proved actually that **any ideal of $\mathbb{Z}$ generated by two elements is a principal ideal**. Using the same argument, we are proving next that all ideals of $\mathbb{Z}$ are principal. We say, because of that, that $\mathbb{Z}$ **is a principal ideal ring**.

Theorem 3.2.4 *The ring of integers $\mathbb{Z}$ is a principal ideal ring. In other words, every ideal I of $\mathbb{Z}$ has one generator.*

Proof Since the ideal (0) is principal, with a generator 0, we may assume that I is a non-zero ideal of $\mathbb{Z}$. Then, I has positive and negative elements as well. Consider the set $A = (\mathbb{N} \cap I) \setminus \{0\}$. The set A is a non-empty subset of $\mathbb{N}$, and therefore has a least element a. Since $a \in A \subset I$, all the multiples of a belong to I. In other words, $(a) \subseteq I$. To finish the proof, we will show next that actually $I \subseteq (a)$. Indeed, let $x \in I$. Since $a \neq 0$, we can divide x by a with a quotient and a remainder:

$$x = a \cdot q + r \qquad \text{where} \qquad 0 \leq r < a.$$

We have that $r = x - a \cdot q \in I$, and since r is smaller than the least element a of A, it can not be in A. This means that $r = 0$. But then $x = a \cdot q \in (a)$. This proves that $I \subseteq (a)$. □

Greatest Common Divisor of a Group of integers

Definition 3.2.5 *Suppose $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ are not all zero. Then their* **greatest common divisor***, denoted by $gcd(a_1, \ldots, a_n)$ or just by $(a_1, \ldots, a_n)$, is the largest integer which divides all the numbers. That is,*

$$(\forall i)(gcd(a_1, \ldots, a_n) \,|\, a_i) \quad \wedge \quad (\forall s \in \mathbb{Z})((\forall i)(s \,|\, a_i) \rightarrow (s \leq gcd(a_1, \ldots, a_n))).$$

By definition, $gcd(0, \ldots, 0) = 0$.

Obviously every common divisor of $a_1, \ldots, a_n$ is less than or equal to $|a_i|$ for every non-zero a_i. From this it follows that there is a greatest such divisor, that is, $\gcd(a_1, \ldots, a_n)$ does exist. Also, it is straightforward that

$$\gcd(a_1, \ldots, a_n) = \gcd(|a_1|, \ldots, |a_n|).$$

One value of the greatest common divisor is distinguished for many reasons in Number Theory: when $\gcd(a_1, \ldots, a_n) = 1$

Definition 3.2.6 *The numbers $a_1, \ldots, a_n \in \mathbb{Z}$ are called* **relatively prime** *if $\gcd(a_1, \ldots, a_n) = 1$. The numbers $a_1, \ldots, a_n$ are called* **pairwise relatively prime** *if every two of them are relatively prime: $\gcd(a_i, a_j) = 1$ for every $1 \leq i \neq j \leq n$.*

Example 3.2.2 The three numbers $n, n+1$ and $n+2$ are relatively prime: $\gcd(n, n+1, n+2) = 1$. Indeed, if $d > 0$ is a common divisor of the three, then $d \mid n$, and $d \mid n+1$, and so, $d \mid (n+1) - n = 1$. This implies that $d = 1$. If n is odd, the numbers $n, n+1$ and $n+2$ are also pairwise relatively prime. To see this notice that $\gcd(n, n+1) = \gcd(n+1, n+2) = 1$, since these are gcd's of consecutive integers, and that $\gcd(n, n+2) = 1$, because n and $n+2$ are consecutive **odd** integers (any their common divisor $d > 0$ should be odd and should divide $n + 2 - n = 2$). But if n is even, $\gcd(n, n+2) = 2$, and therefore the three numbers are not pairwise relatively prime. □

Exercise 3.3 *(1) Prove that if the integers $a_1, \ldots, a_n$ are pairwise relatively prime, then they are relatively prime. Moreover, prove that if* **two of the integers** *are relatively prime, say $\gcd(a_i, a_j) = 1$, then all integers are relatively prime: $\gcd(a_1, \ldots, a_n) = 1$.*
(2) Show by examples that there are groups of relatively prime integers which are not pairwise relatively prime. Moreover, it is possible to have $\gcd(a_1, \ldots, a_n) = 1$ with $\gcd(a_i, a_j) \neq 1$ for all $1 \leq i \neq j \leq n$.

The properties of the gcd of a group of integers follow from the properties of the gcd of two integers. We are proving the most important ones below.

Theorem 3.2.5 *Let $n \geq 3$ and $a_1, a_2, \ldots, a_n$ be integers. The following propositions hold true.*
(i) Let d be the non-negative generator of the ideal $(a_1, \ldots, a_n)$. Then $d = \gcd(a_1, \ldots, a_n)$.
(ii) (Bézout Identity) There are integers $u_1, \ldots, u_n$ such that

$$\gcd(a_1, \ldots, a_n) = u_1 \cdot a_1 + \cdots + u_n \cdot a_n.$$

(iii) The integers $a_1, \ldots, a_n$ are relatively prime if, and only if, there are integers $u_1, \ldots, u_n$ such that

$$u_1 \cdot a_1 + \cdots + u_n \cdot a_n = 1.$$

(iv) Let $d = \gcd(a_1, \ldots, a_n) \neq 0$. Then the integers $a_1/d, \ldots, a_n/d$ are relatively prime.
(v) $\gcd(a_1, a_2, \ldots, a_n) = \gcd(\gcd(a_1, \ldots, a_{n-1}), a_n)$.
(vi) $\gcd(d \cdot a_1, \ldots, d \cdot a_n) = |d| \cdot \gcd(a_1, \ldots, a_n)$.

Proof (i) Since $(d) = (a_1, \ldots, a_n)$, we have that d is a common divisor of the elements $a_1, \ldots, a_n$. That d is the greatest common divisor of these numbers follows from the fact that any common divisor of $a_1, \ldots, a_n$, since d is a linear combination of these numbers, divides d as well. This completes the proof of (i). Items (ii), (iii), and (iv) follow directly from (i) and are left as exercises. For item (v) we denote by $d_k = \gcd(a_1, \ldots, a_k)$ for $k = n-1, n$. We have to show that $d_n = \gcd(d_{n-1}, a_n)$. Since d_n is a common divisor of $a_1, \ldots, a_n$, it is a common divisor of $a_1, \ldots, a_{n-1}$, and of a_n. This implies that $d_n \mid d_{n-1}$, and $d_n \mid a_n$ which in turn implies that $d_n \mid \gcd(d_{n-1}, a_n)$. On the other hand $\gcd(d_{n-1}, a_n)$ divides d_{n-1} and a_n, so it divides all $a_1, \ldots, a_{n-1}, a_n$, and therefore, $\gcd(d_{n-1}, a_n) \mid d_n$. This immediately implies that $d_n = \gcd(d_{n-1}, a_n)$. Item (vi) is straightforward and is left as an exercise. The theorem is proved. □

Exercise 3.4 *(1) Prove that $a_1, \ldots, a_n$ are relatively prime if, and only if, $(a_1, \ldots, a_n) = (1)$*
(2) Prove items (ii), (iii), and (iv) from the theorem above.

Exercise 3.5 *(1) Let a_1, a_2, a_3 be pairwise relatively prime integers. Prove that*

$$\gcd(a_1, a_2 a_3) = \gcd(a_2, a_1 a_3) = \gcd(a_3, a_1 a_2) = 1.$$

(2) Prove that if $a_1, \ldots, a_n$ are pairwise relatively prime integers, then for every $i = 1, \ldots, n$ we have

$$\gcd\left(a_i, \frac{a_1 \cdot a_2 \cdot \cdots \cdot a_n}{a_i}\right) = 1.$$

3.2.3 Euclid's Algorithm

We know now that computing the gcd of a group of integers, $d_n = \gcd(a_1, \ldots, a_n)$, reduces to successively computing gcds of pairs of integers:

$$d_2 = \gcd(a_1, a_2), \quad d_3 = \gcd(d_2, a_3), \quad \cdots, \quad d_n = \gcd(d_{n-1}, a_n).$$

The preceding results are good theoretical results. For practical purposes though, we need an algorithm for finding the greatest common divisor of two integers as well as the coefficients in the corresponding Bézout's identity. A very effective algorithm is provided by the way Euclid proved the existence of the greatest common divisors. Not surprisingly, Euclid's algorithm is based on the division algorithm.

Let m, n be integers such that $n \neq 0$ and $n \nmid m$. Consider the following chain of divisions with quotient and remainder

$$\begin{aligned} m &= n \cdot q_0 + r_0 \\ n &= r_0 \cdot q_1 + r_1 \\ r_0 &= r_1 \cdot q_2 + r_2 \\ &\cdots \\ r_{s-2} &= r_{s-1} \cdot q_s + r_s \\ r_{s-1} &= r_s \cdot q_{s+1} + 0 \end{aligned}$$

where r_s is the last non-zero remainder in the process of divisions. Notice that this r_s does exist, because, according to the division algorithm, $|n| > r_0 > r_1 > \cdots > r_s > \cdots \geq 0$, and since there are only finitely many positive integers less than $|n|$, there should be one equal to zero. The first such one, in our notations, would be r_{s+1}.

Theorem 3.2.6 *In the notations above, $r_s = gcd(m, n)$. In addition, the coefficients of the Bézout's identity can be found by "reading" the system of relations above backward, and solving for r_s in terms of a and b, eliminating step-by-step the rest of the remainders from the system.*

Proof The proof uses induction on $s \geq 0$. **Base case**, $s = 0$. We have in this case that

$$m = n \cdot q_0 + r_0 \qquad\qquad n = r_0 \cdot q_1 + 0.$$

The second relation tells us that $r_0 \mid n$ which, combined with the first relation gives us that $r_0 \mid m$ as well. So, r_0 is a common divisor of m and n. Suppose d is another common divisor of m and n. Since the first relation can be written as

$$r_0 = m - n \cdot q_0$$

we see that $d \mid r_0$, and that r_0 is a integer linear combination of m and n. Therefore $r_0 = \gcd(m, n)$, the Bézout identity is found by reading the system backwards, and the base case is checked.
Inductive step ($s \to s + 1$). The system of relations in this case is the following one

$$\begin{aligned} m &= n \cdot q_0 + r_0 \\ n &= r_0 \cdot q_1 + r_1 \\ r_0 &= r_1 \cdot q_2 + r_2 \\ &\cdots \\ r_{s-2} &= r_{s-1} \cdot q_s + r_s \\ r_{s-1} &= r_s \cdot q_{s+1} + r_{s+1} \\ r_s &= r_{s+1} \cdot q_{s+2} + 0 \end{aligned}$$

We can apply the induction hypothesis to the last $s+2$ relations, and get that

$$r_{s+1} = \gcd(n, r_0) \qquad \text{and} \qquad r_{s+1} = u \cdot n + v \cdot r_0$$

where $u, v \in \mathbb{Z}$ are obtained reading the system of $s+2$ relations backward. The first relation of the system, $m = n \cdot q_0 + r_0$ implies that

$$\gcd(m, n) = \gcd(n, r_0) \qquad \text{and} \qquad r_{s+1} = v \cdot m + (u - vq_0) \cdot n.$$

Therefore

$$\gcd(m, n) = r_{s+1} \qquad \text{and} \qquad r_{s+1} = a \cdot m + b \cdot n$$

where $a, b \in \mathbb{Z}$ are obtained by reading the system of relations backward. □

Exercise 3.6 *(1) Find the greatest common divisor of* 123456789 *and* 987654321.
(2) Find the greatest common divisor of all nine-digit integers that can be written by using the non-zero digits only once each.
(3) Let $a > 1, m, n \in \mathbb{N}$. Prove that $\gcd(a^m - 1, a^n - 1) = a^{\gcd(m,n)} - 1$.
(4) Let $a > 1, m > n$ be integers. Prove that
(i) $a^{2^n} + 1 \mid a^{2^m} - 1$;
(ii) if $a \in 2\mathbb{Z}$, then $\gcd(a^{2^n} + 1, a^{2^m} + 1) = 1$, *and if $a \notin 2\mathbb{Z}$, then* $\gcd(a^{2^n} + 1, a^{2^m} + 1) = 2$
(5) Let $a > 1, m, n \in \mathbb{N}$. Find $\gcd(a^m + 1, a^n + 1)$.

3.3 The Least Common Multiple

The least common multiple is a dual to the greatest common divisor notion. We discuss first the least common multiple of two, and then - of any number of integers.

3.3.1 The Least Common Multiple of Two Integers

Definition 3.3.1 *Let m, n be integers. Define the* **least common multiple**, *denoted by $lcm(m, n)$ or just by $[m, n]$, to be*
(1) the least positive integer which is divisible by both m and n in case $m \cdot n \neq 0$, and
(2) 0 *in case $m \cdot n = 0$. Observe that, in this case, m and n share only one common multiple:* 0.

We are giving next a nice interpretation of the $lcm(m, n)$ using the ideals (m) and (n). Before that, we are proving an easy and important property of ideals of $\mathbb{Z}$.

Proposition 3.3.1 *Let $I_1, dots, I_n$ be ideals of $\mathbb{Z}$. Then their intersection $I_1 \cap \cdots \cap I_n$ is also an ideal of $\mathbb{Z}$.*

Proof We have to show that $I := I_1 \cap \cdots \cap I_n \neq \emptyset$, and that

$$(\forall x, y \in I)(x - y \in I) \qquad \text{and} \qquad (\forall x \in I)(\forall a \in \mathbb{Z})(a \cdot x \in I).$$

But we have that $0 \in I$, because $0 \in I_i$ for every $i = 1, \ldots, n$. Therefore, $I \neq \emptyset$. The rest of the conditions are verified immediately, because $x, y \in I$ is equivalent to $x, y \in I_i$ for $i = 1, \ldots, n$, and because $x - y, a \cdot x \in I_i$ for all $i = 1, \ldots, n$. □

Theorem 3.3.2 *Denote by l the non-negative generator of the ideal $(m) \cap (n)$. Then, $lcm(m, n) = l$*

Proof Since, for every $a \in \mathbb{Z}$, the ideal (a) consists of all integers divisible by a, then the ideal $(m) \cap (n)$ consists of all integers divisible by both m and n. In other words, $(m) \cap (n)$ consists of all common multiples of m and n. By the very definition then it follows that $lcm(m, n)$ is equal to l - the least among the non-negative elements of $(m) \cap (n)$. But this is exactly the non-negative generator of $(m) \cap (n)$. □

The main properties of the least common multiple of two integers easily follow from this theorem, and are listed in the following proposition.

Proposition 3.3.3 *Let $m, n \in \mathbb{Z}$. The following propositions hold true.*
(i) $lcm(m, n)$ exists, and $lcm(m, n) = lcm(|m|, |n|)$;
(ii) if $\gcd(m, n) = 1$, then $lcm(m, n) = |m| \cdot |n|$;
(iii) we have $lcm(d \cdot m, d \cdot n) = |d| \cdot lcm(m, n)$;
(iv) the following relation holds true $|m| \cdot |n| = gcd(m, n) \cdot lcm(m, n)$;
(v) if $s \in \mathbb{Z}$ is such that $m \mid s$ and $n \mid s$, then $lcm(m, n) \mid s$.

Proof The claim in (i) follows from the fact that

$$(lcm(m, n)) = (m) \cap (n) = (|m|) \cap (|n|) = (lcm(|m|, |n|)).$$

We are proving the claim in item (ii) now. Let s be a common multiple of m and n. If $s = 0$, then $lcm(m, n) = 0$ and $m \cdot n = 0$, and therefore the formula we are proving is true. Assume now that $s \neq 0$. We have $s = m \cdot s'$, and $n \mid m \cdot s'$. Since $\gcd(m, n) = 1$ it follows (using an exercise from previous subsection) that $n \mid s'$, that is, $s' = n \cdot s''$. Therefore, any common multiple of m and n has the form $s = m \cdot n \cdot s''$. Obviously, the non-negative ones have the form $s = |m| \cdot |n| \cdot s''$ where $s'' \geq 0$. Now, since $s = lcm(m, n) \neq 0$, then $lcm(m, n) > 0$. Any positive common multiple of m and n has the form $s = |m| \cdot |n| \cdot s''$ where $s'' > 0$.The smallest such multiple is obtained when $s'' = 1$. Therefore, in case $\gcd(m, n) = 1$, we have $lcm(m, n) = |m| \cdot |n|$.
For the claim in (iii) observe that it is true of $d = 0$. Let's assume that $d \neq 0$. We have in this case that s is a common multiple of $d \cdot m$ and $d \cdot n$ if, and only if, s/d is a common multiple of m and n. Therefore, the least common multiple of $d \cdot m$ and $d \cdot n$ is the same as $|d|$ times the least common multiple of m and n.
To prove (iv), we use (ii) and (iii). We consider two cases: $\gcd(m, n) = 0$ and $\gcd(m, n) \neq 0$. In the former case, possible only when $m = n = 0$, the identity that we have to prove is true: $0 = 0$. In the latter case, denote by $d = \gcd(m, n)$. We have $m = d \cdot m'$ and $n = d \cdot n'$ where now $\gcd(m', n') = 1$. By (iii) and (ii) we have

$$lcm(m, n) = lcm(d \cdot m', d \cdot n') = d \cdot lcm(m', n') = d \cdot m' \cdot n' = \frac{(d \cdot m') \cdot (d \cdot n')}{d} = \frac{m \cdot n}{\gcd(m, n)}.$$

For item (v) we notice that $s \in (m) \cap (n) = (lcm(m, n))$. □

Exercise 3.7 *Let $l \in \mathbb{N}$, and $m, n \in \mathbb{Z}$. Prove that*

$$l = lcm(m, n) \quad \Leftrightarrow \quad (\forall s \in \mathbb{Z})(m \mid s \ \wedge \ n \mid s \ \rightarrow l \mid s).$$

Remark 3.3.1 After this exercise we can see the dual natures of the concept of gcd and lcm. Compare the statements we proved for d and l natural numbers, and $m, n \in \mathbb{Z}$

$$d = \gcd(m, n) \quad \Leftrightarrow \quad (\forall s \in \mathbb{Z})(s \mid m \ \wedge \ s \mid n \ \rightarrow \ s \mid d)$$

and

$$l = lcm(m, n) \quad \Leftrightarrow \quad (\forall s \in \mathbb{Z})(m \mid s \ \wedge \ n \mid s \ \rightarrow \ l \mid s).$$

Notice that both claims are expressed purely in terms of divisibility of integers. Both are also formulated using a universal quantification on the set of integers $\mathbb{Z}$. That's why the right-hand sides of the two claims are called **universal properties** of, respectively, the gcd and the lcm. □

3.3.2 The Least Common Multiple of a Group of Integers

The notion of least common multiple is readily extended to the case of three or more non-zero integers.

Definition 3.3.2 *Let $a_1, \ldots, a_n$ be integers. Define their* **least common multiple**, *denoted by $lcm(a_1, \ldots, a_n)$ or just by $[a_1, \ldots, a_n]$, to be*
(1) the least positive integer which is divisible by both all $a_i, i = 1, \ldots, n$ in case $a_1 \cdot a_2 \cdots a_n \neq 0$, and
(2) 0 in case $a_1 \cdot a_2 \cdots a_n = 0$.

As in the case of least common multiple of two integers we are proving the fundamental fact that

Theorem 3.3.4 *Let l be the non-negative generator of the ideal $(a_1) \cap \cdots \cap (a_n)$. Then,*

$$lcm(a_1, \ldots, a_n) = l.$$

Proof The proof is identical to the one for $n = 2$, and is left as an exercise. □

The theorem immediately gives us that the lcm of a group of integers exists, and moreover that

$$lcm(a_1, \ldots, a_n) = lcm(|a_1|, \ldots, |a_n|).$$

Here are some of the basic properties of the least common divisor of a group of numbers.

Exercise-Proposition 3.3.3 *For the integers $a_1, a_2, \ldots, a_n$, $n \geq 3$ the following propositions hold true.*

(i) If s is a common multiple of the integers, then $lcm(a_1, \ldots, a_n) \mid s$.
(ii) $lcm(d \cdot a_1, \ldots, d \cdot a_n) = |d| \cdot lcm(a_1, \ldots, a_n)$.
(iii) $lcm(a_1, \ldots, a_{n-1}, a_n) = lcm(lcm(a_1, \ldots, a_{n-1}), a_n)$.
(iv) If the numbers are pairwise relatively prime, then

$$lcm(a_1, a_2, \ldots, a_n) = a_1 a_2 \cdots a_n,$$

Proof All items follow easily from the fundamental theorem above. For item (i) we notice that s as a common multiple of $a_1, \ldots, a_n$ needs to be in

$$(a_1) \cap \cdots \cap (a_n) = (lcm((a_1, \ldots, a_n)).$$

Item (ii) is obvious if $d = 0$. If not, then again obviously s is a common multiple of $d \cdot a_1, \ldots, d \cdot a_n$ if, and only if, s/d is a common multiple of $a_1, \ldots, a_n$. Therefore, the lcm of $d \cdot a_1, \ldots, d \cdot a_n$ is $|d|$ times the lcm of $a_1, \ldots, a_n$.
Item (iii) follows directly from the set-theoretical fact that

$$(a_1) \cap \cdots \cap (a_{n-1}) \cap (a_n) = ((a_1) \cap \cdots \cap (a_{n-1})) \cap (a_n).$$

Due to the fundamental theorem above, we have

$$(lcm(a_1, \ldots, a_{n-1}, a_n)) = (lcm(a_1, \ldots, a_{n-1})) \cap (a_n) = (lcm(lcm(a_1, \ldots, a_{n-1}), a_n)).$$

Since any ideal of $\mathbb{Z}$ has a unique non-zero generator, we get the identity in claim (iii).
We are proving (iv) by induction on $n \geq 3$ (and using (iii)). **Base case**: $n = 3$. We have by (ii) that $lcm(a_1, a_2, a_3) = lcm(lcm(a_1, a_2), a_3)$. From item (ii) of the theorem in the previous sub-section we have that $lcm(a_1, a_2) = a_1 \cdot a_2$. Since a_1, a_2, a_3 are pairwise relatively prime, and from a previous exercise, we know that $\gcd(a_1 \cdot a_2, a_3) = 1$, and applying item (ii) of the mentioned theorem, we get the desired

$$lcm(a_1, a_2, a_3) = lcm(a_1 \cdot a_2, a_3) = a_1 \cdot a_2 \cdot a_3.$$

Inductive step: $n \to n+1$. Since $a_1, \ldots, a_n$ are pairwise relatively prime, by a previous exercise we know that $\gcd(a_1 \cdots a_{n-1}, a_n) = 1$. We have, using the inductive hypothesis for the second equality, and the property of lcm of two relatively prime integers for the third,

$$lcm(a_1, \ldots, a_{n-1}, a_n) = lcm(lcm(a_1, \ldots, a_{n-1}), a_n) = lcm(a_1 \cdots a_{n-1}, a_n) = a_1 a_2 \cdots a_n.$$

The claim in item (iv) is proved. □

Exercise 3.8 *(1) Let a, b, c be integers. Is it true that $(a, b) = (a, c)$ implies $[a, b] = [a, c]$? Give reasoning for your answer.*
(2) For a natural number n evaluate $(n, n+1, n+2)$ and $[n, n+1, n+2]$.
(3) For the natural numbers a, b, c prove that if $(a, b) = (b, c) = (c, a) = 1$, then $(a, b, c)[a, b, c] = abc$.
(4) Is the statement of (3) true without the restriction on a, b, and c? Explain.
(5) Prove for every $a, b, c \in \mathbb{Z}$ that

$$([a, b], c) \cdot [a, b, c] = [a, b] \cdot |c|.$$

(6) As in (5) prove that

$$(ab, (a, b)c) = (ac, (a, c)b) = (bc, (b, c)a).$$

If $abc \neq 0$ prove that the three expressions above are equal to $abc/[a, b, c]$.
(7) Let $a > 1, m, n$ be natural numbers. We know that $\gcd(a^n - 1, a^m - 1) = a^{\gcd(n,m)} - 1$. When is it true that

$$lcm(a^n - 1, a^m - 1) = a^{lcm(n,m)} - 1 \quad ?$$

(8) Let $s = a_1 a_2 \cdots a_n \neq 0$. Prove that
(i) $|a_1 a_2 \cdots a_n| = lcm(a_1, a_2, \ldots, a_n) \cdot \gcd(s/a_1, s/a_2, \ldots, s/a_n)$.
(ii) Suppose $m \in \mathbb{N}$ is a common multiple of $a_1, a_2, \ldots, a_n$. Then

$$m = lcm(a_1, a_2, \ldots, a_n) \quad \Leftrightarrow \quad \gcd(m/a_1, m/a_2, \ldots, m/a_n) = 1.$$

Chapter 4

The Fundamental Theorem of $\mathbb{Z}$

In this chapter, we are learning about one other important property of the prime numbers: every integer different from 0 and ± 1 is, up to a sign, an essentially unique product of finitely many prime numbers! This shows that the prime numbers are the building blocks for constructing all the integers (with the exception of $0, \pm 1$ of course).

Let's first note that it is enough to prove this result for natural numbers: the negative integers are the opposites of the natural numbers.
The proof of that theorem has two parts: existence of such a presentation, and uniqueness of that presentation. The former is based, again!, on the least element property of the natural numbers, while the latter is based on the divisibility properties of the prime numbers.

4.1 Existence

Theorem 4.1.1 *Let $n \in \mathbb{N}$ be different from 0 and 1. Then, there are finitely many prime numbers, $p_1, p_2, \ldots, p_k$, not necessarily distinct, such that $n = p_1 p_2 \cdots p_k$.*

Proof Consider the set

$$\Sigma = \{m \in \mathbb{N} \,|\, (m > 1) \wedge (m \quad \text{is not a product of finitely many primes})\}.$$

We want to show that $\Sigma = \emptyset$. Arguing by contradiction, assume $\Sigma \neq \emptyset$. Since $\Sigma \subseteq \mathbb{N}$, there is an $s \in \Sigma$ which is the least element there. Now, s is a natural number bigger than 1, so it is either a prime number or a composite number. But if s is prime, then it **is** a product of finitely many (of one) primes: $s = s$. Since $s \in \Sigma$, it is not a product of finitely many primes, and therefore s must be composite: $s = s_1 \cdot s_2$ where both s_1 and s_2 are bigger than 1. SO, both, s_1 and s_2 are less than s (!!). The latter fact implies in turn that BOTH s_1 and s_2 are NOT in Σ. And since they are both bigger than 1, they both ARE products of finitely many primes:

$$s_1 = p_1 \cdots p_l \qquad s_2 = q_1 \cdots q_m.$$

But then,$s = s_1 \cdot s_2 = p_1 \cdots p_l \cdot q_1 \cdots q_m$ is a product of finitely many primes as well - a contradiction! Therefore $\Sigma = \emptyset$, and we are done. $\square$

4.2 Uniqueness

To prove this part of the theorem we need some preparation related to divisibility of integers.

Proposition 4.2.1 *Let $a, b, c \in \mathbb{Z}$. Suppose $a \,|\, b \cdot c$. If $gcd(a, b) = 1$, then $a \,|\, c$*

Proof Since $gcd(a, b) = 1$, by the Bézout's identity, there are integers u, v such that $u \cdot a + v \cdot b = 1$. Therefore $(u \cdot a + v \cdot b) \cdot c = c$, so

$$(u \cdot c) \cdot a + v \cdot (b \cdot c) = c.$$

By $a \,|\, b \cdot c$ we get that the LHS of the last equality is divisible by a. So, the RHS of that equality, c, is divisible by a as well. □

The next Corollary is known as Euclid's Lemma and reveals a very important property of the prime numbers. Before we formulate it - an exercise:

Exercise 4.1 *Let p be a prime number and let a be an integer. Then* $\gcd(p, a) = 1$ *if* $p \nmid a$, *and* $\gcd(p, a) = p$ *otherwise.*

Corollary 4.2.2 *Let p be a prime number. Prove that*

$$(p \,|\, (b \cdot c) \ \wedge \ p \nmid b) \quad \rightarrow \quad (p \,|\, c).$$

Proof Notice that $p \nmid b$ is equivalent with $\gcd(p, b) = 1$, and apply the Proposition above. □

Exercise 4.2 *Prove that $2 \leq a \in \mathbb{N}$ is a prime number if, and only if,*

$$(\forall b, c \in \mathbb{Z})(a \,|\, b \cdot c \rightarrow (a \,|\, b \vee a \,|\, c)).$$

We will need a generalization of the last Corollary's claim.

Proposition 4.2.3 *Let p be a prime number and $a_1, a_2, \ldots, a_l \in \mathbb{Z}$. If $p \,|\, a_1 a_2 \cdots a_l$, then there is an $i \in \{1, 2, \ldots, l\}$ such that $p \,|\, a_i$.*

Proof By way of RAA, assume that this claim is not true. So, there is a prime p, and integers $a_1, \ldots, a_l$ for which p divides the product $a_1 \cdots a_l$, but no a_i is divisible by p. W.L.O.G. we may assume that l is the least positive natural number for which the claim is not true (for this fixed p). Obviously, $l \geq 2$ (why?), and, we have

$$p \,|\, (a_1 \cdots a_{l-1})a_l \quad \wedge \quad \neg(p \,|\, a_l).$$

By the Corollary above we get

$$p \,|\, a_1 \cdots a_{l-1}.$$

The last relation together with $p \nmid a_1, \ldots, p \nmid a_{l-1}$ gives that l is NOT the least positive integer for which the claim is not true (for our fixed p). This is the needed contradiction. Therefore, the assumption that the claim is not true is wrong, and the Lemma is proved. □

Theorem 4.2.4 *For every integer $n > 1$, there are unique prime numbers*

$$p_1 \leq p_2 \leq \cdots \leq p_k$$

such that $n = p_1 p_2 \cdots p_k$.

Proof We need to show that if $p_1 \leq p_2 \leq \cdots \leq p_k$ and $q_1 \leq q_2 \leq \cdots \leq q_s$ are two groups of prime numbers such that $p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_s$, then $k = s$ and $p_i = q_i$ for every $i = 1, \ldots, k$. Arguing again by contradiction, assume there are such groups of primes for which the conclusion is not true. Obviously, there are two such groups for which the total of the primes involved, $k + s$, is the smallest possible. Notice that $k, s \geq 1$ (Why?). We have

$$p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_s$$

Since p_1 divides $q_1 q_2 \cdots q_s$, by the lemma above, there is an $i \in \{1, 2, \ldots, s\}$ such that $p_1 \,|\, q_i$. But q_i , being prime, has only one divisor bigger than 1: q_i itself. So, $p_1 = q_i$. Cancelling out p_1 on the left by q_i on the right we get two new groups of primes for which the conclusion of the theorem is not true, and with smaller total number of primes involved - a contradiction! So, our assumption from the beginning of the proof of the theorem is wrong, and the theorem is proved. □

4.3 Canonical Decomposition, Applications

The Fundamental Theorem of Arithmetic of $\mathbb{Z}$ means that the integers have the form

$$0, \quad \pm 1, \quad \pm p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$$

for unique $p_1 < p_2 < \cdots < p_k$ and $\alpha_i \geq 1$. The presentation of the integers different from $0, \pm 1$ is called the **canonical decomposition of the integers as a product of primes**. This theorem reveals the importance of the prime numbers in $\mathbb{Z}$: every non-zero integer has, up to a sign, a unique representation (a.k.a. factorization) as a product of powers of distinct primes. This is why $\mathbb{Z}$ is called a **unique factorization ring**.

Remark 4.3.1 There are many unique factorization rings. One such is the ring of polynomials of one indeterminate with rational (or real, or complex) coefficients. More about all this can be learned in the course Topics in Algebraic Structures. □

Note that if we have two, or any finitely many, non-zero integers we may, without a loss of generality, assume that the prime numbers involved in their presentations are the same. This is done by allowing some, or all, of the exponents α_i to be 0. In particular, ± 1 also have such presentations: all exponents are 0. The only integer without such a presentation is 0. Of course, allowing zero exponents, we violate the uniqueness of the presentation. This violation will not be too harmful to our considerations though.

It is very easy, and illuminating, to write the greatest common divisor and the least common multiple of two numbers using their canonical presentations. You should verify as an exercise that if

$$a = \pm p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k} \quad \text{and} \quad b = \pm p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k},$$

then

$$gcd(a,b) = \prod_{i=1}^{k} p_i^{min\{\alpha_i, \beta_i\}} \quad \text{and} \quad lcm(a,b) = \prod_{i=1}^{k} p_i^{max\{\alpha_i, \beta_i\}}.$$

As pretty as these expressions look, they have mostly theoretical importance. The reason for that is that it is very hard to find the canonical decomposition of a random integer. We will see later in the course that this fact is effectively used to encrypt information (the so called "public key cryptosystem").

Exercise 4.3 *(1) Let $a_1, \ldots, a_n$ be non-zero integers. Using the canonical decompositions of these numbers as products of powers of primes, find a formula for $\gcd(a_1, \ldots, a_n)$ and for $lcm(a_1, \ldots, a_n)$. (2) Let p be a prime number, and $a, b, c \in \mathbb{Z}$. Decide if the following statements are true or false. Give reasons for your answers: if a statement is true, give a proof or otherwise - give a counterexample.*

(i) $(p \,|\, (a^2 + b^2) \ \wedge \ p \,|\, (b^2 + c^2)) \ \rightarrow \ (p \,|\, (a^2 - c^2))$

(ii) $(p \,|\, (a^2 + b^2) \ \wedge \ p \,|\, (b^2 + c^2)) \ \rightarrow \ (p \,|\, (a^2 + c^2))$

(iii) $(p \,|\, a^3) \ \rightarrow \ (p \,|\, a)$ *(iv)* $(p \,|\, a \ \wedge \ p \,|\, (a^2 + b^2)) \ \rightarrow \ (p \,|\, b)$

(v) $(a^3 \,|\, b^3) \ \rightarrow \ (a \,|\, b)$ *(vi)* $(a^3 \,|\, b^2) \ \rightarrow \ (a \,|\, b)$ *(vii)* $(a^2 \,|\, b^3) \ \rightarrow \ (a \,|\, b)$.

(3) This is a nice exercise to test our intuition about prime numbers and unique factorization. Consider the set of even integers

$$2\mathbb{Z} = \{2n \,|\, n \in \mathbb{Z}\}$$

The total order in $\mathbb{Z}$ determines one in $2\mathbb{Z}$, so we can define positive (the bigger than 0), and negative (the smaller than 0 elements in $2\mathbb{Z}$. The addition and multiplication in $\mathbb{Z}$ determine addition and multiplication in $2\mathbb{Z}$ as well: the sum of two even numbers is even, and the product of two even numbers is even. The operations in $2\mathbb{Z}$ are quite similar to the ones in $\mathbb{Z}$: they are commutative, associative, the addition distributes over the multiplication, there is an additive neutral element, 0,

every element has an opposite, and the additive cancellation property is satisfied. There are differences between $\mathbb{Z}$ *and* $2\mathbb{Z}$ *too:* **there is no multiplicative neutral element***:* $1 \notin 2\mathbb{Z}$*. Nevertheless, the multiplicative cancellation property is satisfied:*

$$2a \cdot 2b = 2a \cdot 2c \quad \wedge \quad 2a \neq 0 \quad \rightarrow \quad 2b = 2c.$$

Furthermore, in analogy with $\mathbb{Z}$*, we can discuss questions of divisibility and factorization in* $2\mathbb{Z}$*:*

$$\textit{for} \quad x, y \in 2\mathbb{Z} \quad \textit{we say that} \quad x \,|\, y \quad \textit{if} \quad (\exists z \in 2\mathbb{Z})(y = x \cdot z).$$

(Note, for example, that in $2\mathbb{Z}$*,* 2 *does not divide* $6 = 2 \cdot 3$*, but* 2 *does divide* $4 = 2 \cdot 2$*). We can also introduce the notion of* **primes in** $2\mathbb{Z}$*: these are positive numbers of* $2\mathbb{Z}$ *who* **do not** *have positive divisors in* $2\mathbb{Z}$*. (Note that since* $1 \notin 2\mathbb{Z}$*, we ignore it as a possible divisor, in* $2\mathbb{Z}$*, of the elements of* $2\mathbb{Z}$*.)*

(i) List some primes of $2\mathbb{Z}$*. Describe all the primes of* $2\mathbb{Z}$*.*

(ii) Prove that every positive element of $2\mathbb{Z}$ *can be expressed as a product of primes of* $2\mathbb{Z}$*.*

(iii) Show that this factorization into primes **need not be unique***.*

(iv) What about negative primes and the factorization of negative elements of $2\mathbb{Z}$*?*

4.4 Infinitude of the Prime Numbers

After proving the Fundamental Theorem of Arithmetic of $\mathbb{Z}$, an interesting question is: how many are the prime numbers? A simpler question is whether they are finitely or infinitely many. The answer to the latter question was known to Euclid, and is given below. The former question is much harder. It was answered in the end of 19th century (in 1896), and the answer is known as the **Prime Number Theorem**. This theorem is discussed in the last chapter of these Notes.

Theorem 4.4.1 *The odd prime numbers are infinitely many.*

Proof We follow the classical Euclid's proof of this theorem. Arguing by contradiction, assume that the prime numbers are finitely many, and that $2, p_1, p_2, \ldots, p_k$ is the list of all primes. Consider the number $E = 2 \cdot p_1 p_2 \cdots p_k + 1$. The number E is bigger than 1, so it has a canonical decomposition as a product of primes, as we just proved. So, there is a prime number (necessarily - from the list above) which divides E: say $p_i \,|\, E = 2p_1 \cdots p_i \cdots p_k + 1$. Therefore. $p_i \,|\, E - 2 \cdot p_1 \cdots p_i \cdots p_k = 1$. That is a contradiction! $\square$

Now, half of the odd numbers have the form $4n + 1$ and the other half have the form $4n + 3$. There are prime numbers of both types, of course. The question is: are the prime numbers of each of these types infinitely many? The answer to this question is "yes", but with our current knowledge in Number Theory, we are able to prove only that the primes of type $4n + 3$ are infinitely many. In a week or so, we will be able to tackle the second question successfully too.

Theorem 4.4.2 *The prime numbers of type* $4n + 3, n \in \mathbb{N}$ *are infinitely many.*

Proof The argument here is a minor modification of the one Euclid has used. Arguing again by contradiction, assume that there is a finite list of all such prime numbers: $p_1, p_2, \ldots, p_s$. Consider the number $F = 4p_1 \cdots p_k - 1$. This number is bigger than 1, and therefore has a presentation as a product of primes

$$F = q_1 q_2 \cdots q_l.$$

Since F is an odd number, all primes $q_1, \ldots, q_l$ are also odd. The crucial observation now is that there is at least one $i \in \{1, \ldots, l\}$ such that $q_i = 4n_i + 3$. Indeed, otherwise we would have

$$F = 4p_1 \cdots p_k - 1 = 4(p_1 \cdots p_k - 1) + 3 = 4A + 3$$

and

$$F = q_1 \cdots q_l = (4n_1 + 1) \cdots (4n_l + 1) = 4B + 1$$

so that

$$4A + 3 = 4B + 1$$

which is of course impossible! But this means that there is a p_j dividing F, which is another impossibility. So the assumption that the primes of type $4n+3$ are finitely many is wrong, and the theorem is proved. □

Exercise 4.4 *Prove that the prime numbers of type* $6m + 5, m \in \mathbb{N}$ *are infinitely many.*
[Hint: Mimic the argument in the proof of the previous theorem.]

It is obvious that the remainder of an odd prime number when divided by 4 can only be 1 or 3. This means that the odd prime numbers are of the form $4n + 1$ or $4n + 3$. Similarly, the remainders the odd prime numbers can have when divided by 6 can only be 1 or 5: the odd primes are of the form $6m + 1$ and $6m + 5$. We know that the odd prime numbers are infinitely many, and that the prime numbers of type $4n+3$ and $6m+5$ are infinitely many. We will prove soon, for which we will need a deeper knowledge of Number Theory, that the primes of type $4n + 1$ and the primes of type $6m + 1$ are also infinitely many. The natural question to ask here is if the prime numbers of type $an + b$ for $n \in \mathbb{N}$ are infinitely many. Of course, we need to have $\gcd(a, b) = 1$ in order to have at least one such prime. The affirmative answer to this question was given by the German mathematician P.-G. L. Dirichlet who proved in 1837 his **Theorem About the Primes in an Arithmetic Progression** stated below. The proof of this result goes far beyond the scope of our course.

Theorem 4.4.3 (*Dirichlet's Theorem about the prime numbers in an arithmetic progression*) *Let* $a, b \in \mathbb{N}$ *be relatively prime. Then the set*

$$\{\, an + b \,|\, n \in \mathbb{N}\}$$

contains infinitely many prime numbers.

Remark 4.4.1 Dirichlet's theorem claims in effect that every **linear polynomial with integer coefficients** which are relatively prime contains infinitely many prime numbers. It is natural to ask if there are polynomials of higher degree which have the same property. The answer to this question is not known. We will say more about polynomials, and other "reasonable" functions,taking on many values which are prime numbers in the last Chapter of these Notes.

4.5 An Interpretation through $\mathbb{Q}$

In this section, we give an interpretation of the fundamental theorem of the arithmetic of $\mathbb{Z}$ through the non-zero elements of the field of rational numbers.

4.5.1 The Group $\mathbb{Q}_{>0}$

The positive rational numbers can be added and multiplied without getting out of their realm: sum and product of positive numbers is positive. The addition has no neutral element though, so, that operation is not of much interest (at least to us). The multiplication does have a neutral element: $1 \in \mathbb{Q}_{>0}$, and moreover, every positive rational number has a positive reciprocal element. So we have

$$\cdot : \mathbb{Q}_{>0} \times \mathbb{Q}_{>0} \to \mathbb{Q}_{>0}$$

$$x \cdot y = y \cdot x \qquad x \cdot (y \cdot z) = (x \cdot y) \cdot z \qquad 1 \cdot x = x \quad (\forall x)(\exists y)(x \cdot y = 1).$$

In technical terms (to be developed in detail in Algebraic Structures), this means that $(\mathbb{Q}_{>0}, \cdot)$ is **an Abelian group**.

We also know that every rational number has a unique representation as a reduced rational fraction

$$(\forall x \in \mathbb{Q}_{>0})(\exists!\, m, n \in \mathbb{N})(x = m/n \ \wedge\ \gcd(m, n) = 1).$$

We know that m, n have unique expression as a product of powers of primes with non-zero exponents - that's the statement of the fundamental theorem of arithmetic of $\mathbb{Z}$. We know also that we can find a presentation of m and n in terms of the same primes allowing some the exponents to be zero. There are many such presentations, but if we consider only those who have sum of the exponents corresponding to a prime to be positive, then the presentation is unique. But then, **the rational number itself has a unique expression as a product of non-zero integer powers of primes**

$$1 \neq x = p_1^{\alpha_1} \cdots p_k^{\alpha_k} \qquad p_1 < \cdots < p_k \qquad \alpha_1, \ldots, \alpha_k \in \mathbb{Z} \setminus \{0\}.$$

Example Let $x = 75/296$. Then $m = 75 = 3 \cdot 5^2$, $n = 296 = 2^3 \cdot 37$, and we can (in a unique way) write

$$m = 2^0 \cdot 3 \cdot 5^2 \cdot 37^0 \qquad n = 2^3 \cdot 3^0 \cdot 5^0 \cdot 37.$$

Therefore

$$x = 2^{-3} \cdot 3 \cdot 5^2 \cdot 37^{-1}.$$

The positive integers in $\mathbb{Q}_{>0}$ are given by such products with **positive** exponents

$$1 \neq x \in \mathbb{N} \setminus \{0\} \subseteq \mathbb{Q}_{>0} \qquad \Leftrightarrow \qquad x = p_1^{\alpha_1} \cdots p_k^{\alpha_k}, \quad p_1 < \cdots < p_k, \quad \alpha_1, \ldots, \alpha_k \in \mathbb{N} \setminus \{0\}.$$

The moral here is that the fundamental theorem can be, equivalently, interpreted via the uniqueness in the presentation of the positive rational numbers just explained. Again in technical terms (to be explained in detail in Algebraic Structures), the uniqueness of the presentation means that $(\mathbb{Q}_{>0}, \cdot)$ is a **free Abelian group with generators - the prime numbers**. Since the generators are as many as the elements of $\mathbb{N}$, we write in this case

$$\mathbb{Q}_{>0} \cong \mathbb{Z}^{\oplus \mathbb{N}}.$$

4.5.2 The Group $\mathbb{Q}^\times$

The set $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$ of non-zero rational numbers is also an Abelian group with the operation multiplication. We can apply the fundamental theorem in full to understand the structure of this group. By the previous sub-section, and applying the theorem to any non-zero m now, we see that every different from ± 1 non-zero rational number x has a unique expression as

$$x = \pm p_1^{\alpha_1} \cdots p_k^{\alpha_k} \qquad p_1 < \cdots < p_k \qquad \alpha_1, \ldots, \alpha_k \in \mathbb{Z} \setminus \{0\}.$$

And similarly to $\mathbb{N} \setminus \{0\} \subseteq \mathbb{Q}_{>0}$ we can identify the non-zero integers on $\mathbb{Q}^\times$ to those who have non-negative exponents $\alpha_1, \ldots, \alpha_k$.
Contrary to the previous case, the group $\mathbb{Q}^\times$ is **not** a free Abelian group. This is due to the $\pm$ sign in front of the rational numbers! In professional notations we have actually that

$$\mathbb{Q}^\times \cong \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}^{\oplus \mathbb{N}}.$$

Exercise 4.5 *(1) Let k be a positive integer. Prove that*

$$(gcd(a, b) = 1 \ \wedge \ a \cdot b = c^k) \quad \rightarrow \quad ((\exists u, v \in \mathbb{Z})(a = u^k \ \wedge \ b = v^k)).$$

(2) Let p be a prime number, and $a, b \in \mathbb{Z}$ be such that $(a, b) = p$. Find all possible values of

$$(a^2, b); \quad (a^2, b^2); \quad (a^3, b); \quad (a^3, b^2).$$

(3) Let $a, b, c \in \mathbb{N}$. Show that

$$abc = (a, b, c)[ab, bc, ca] = (ab, bc, ca)[a, b, c].$$

(4) For $a, b, c \in \mathbb{N} \setminus \{0\}$ show that

$$(abc = (a, b, c)[a, b, c]) \quad \rightarrow \quad ((a, b) = (b, c) = (c, a) = 1).$$

(5) Let $a, b, c \in \mathbb{N} \setminus \{0\}$. Prove the relations

$$(a, b, c) = \frac{(a, b)(b, c)(c, a)}{(ab, bc, ca)} \qquad [a, b, c] = \frac{abc(a, b, c)}{(a, b)(b, c)(c, a)}.$$

(6) Let $a, b, c \in \mathbb{N} \setminus \{0\}$. Prove the relation

$$[(a, b), (b, c), (c, a)] = ([a, b], [b, c], [c, a]).$$

Chapter 5

Linear Diophantine Equations

Consider the equation

$$a \cdot x = b$$

where $a, b \in \mathbb{Z}$. It is immediate to see when this equation has a solution (provided $a \neq 0$). Namely, solution exists, and is unique if, and only if, $a \mid b$.
We move on and consider the equation

$$a \cdot x + b \cdot y = c$$

asking for all its solutions in integers. Such equations are called linear Diophantine equations (of two unknowns).

Use of Geometry can put our considerations in right perspective. Observe that the same equation considered over the real numbers defines a line in the Euclidean plane. The solutions in $\mathbb{R}$ correspond to the points on that line, and are uncountably many. The rational solutions to that equation correspond to the points on the line which have rational coordinates, the so called $\mathbb{Q}$-rational points. It is easy to show that there are infinitely, but countably, many such points. The integer solutions to the equation correspond to points on the line which have integers as coordinates called integral or $\mathbb{Z}$-rational points. As we will see, the existence of $\mathbb{Z}$-rational points on the line $ax + by = c$ is a delicate fact: such points may or may not exist! This chapter is devoted to solving the linear Diophantine equations.

The most general linear Diophantine equations has the form

$$a_1 \cdot x_1 + \cdots + a_k \cdot x_k = b$$

where the coefficients, not all 0, and b are integers.

This chapter is devoted to finding the solutions to the linear Diophantine equations. We begin with the case $k = 2$ - the basic and more important one.

5.1 The Equation $a \cdot x + b \cdot y = c$

Theorem 5.1.1 *The Diophantine equation $a \cdot x + b \cdot y = c$, with $|a| + |b| > 0$, has solutions in integers if, and only if $gcd(a, b) \mid c$. In that case, if $d = gcd(a, b)$ and (x_0, y_0) is a solution to the equation, then all the solutions to the equations are given by*

$$x = x_0 + (b/d) \cdot k \qquad\qquad y = y_0 - (a/d) \cdot k$$

where $k \in \mathbb{Z}$.

Proof Denote $d = gcd(a, b)$. By our assumption about a and b, $d \geq 1$. Obviously, for any integers x', y', the number $a \cdot x' + b \cdot y' \in d\mathbb{Z}$. So, if there is a solution $a \cdot x_0 + b \cdot y_0 = c$, then $c \in d\mathbb{Z}$, and so $d \,|\, c$. Conversely, if $d \,|\, c$, so that $c = d \cdot c'$, the Bézout's identity gives integers x', y' such that

$$a \cdot x' + b \cdot y' = d.$$

Then we obviously have that

$$a \cdot (x' \cdot c') + b \cdot (y' \cdot c') = (d \cdot c')$$

so that $x_0 = x' \cdot c'$ and $y_0 = y' \cdot c'$ provide a solution to the equation. This completes the proof of the first part of the claim of the Theorem.
Further, assuming $d \,|\, c$, it is obvious that the given formulae provide solutions to the linear Diophantine equation. So, it is enough to show that in that case ALL solutions are given by those formulae. We are doing this now. Note that W.L.O.G. we may assume that both a and b are non-zero. Suppose (x_0, y_0) is a fixed solution, and that (x', y') is any other solution. We have

$$a \cdot x_0 + b \cdot y_0 = c = a \cdot x' + b \cdot y'$$

which gives us

$$a \cdot (x_0 - x') = b \cdot (y' - y_0).$$

Dividing by d both sides, and denoting $a' = a/d$ and $b' = b/d$, we get

$$a' \cdot (x_0 - x') = b' \cdot (y' - y_0)$$

where, as we know, $gcd(a', b') = 1$. We have now that

$$b' \,|\, a' \cdot (x_0 - x') \quad \text{and} \quad \gcd(a', b') = 1.$$

This implies that $b' \,|\, (x' - x_0)$. So, $x' - x_0 = b' \cdot k$ for an integer number k. Substituting this back in the identity above we get

$$a' \cdot b' \cdot (-k) = b' \cdot (y' - y_0)$$

and cancelling out b' (remember - both a and b are non-zero!), we get $y' = y_0 - a' \cdot k$ as needed.This completes the proof of the second part of the claim of the Theorem. □

Example 5.1.1 Let's solve $3x + 7y = 26$ in integers.
We first check that solutions exist, that is, that $\gcd(3, 7) \,|\, 26$. Since $\gcd(3, 7) = 1$, solutions exist. We have to find next a particular solution to the equation. Often such a solution can be guessed. In the case at hand, obviously $3 \cdot (-3) + 7 \cdot 5 = 26$, so $x_0 = -3, y_0 = 5$ is a particular solution. Then we apply the formulae for the general solution

$$x = -3 + 7k \qquad y = 5 - 3k \quad k \in \mathbb{Z}.$$

There are more "theoretical" ways to find particular solutions (following the proof of the theorem above). For instance, knowing that $\gcd(3, 7) = 1$, we can find $u_o, v_0 \in \mathbb{Z}$ such that $3 \cdot u_0 + 7 \cdot v_0 = 1$ Then a particular solution is given by $x_0 = 26 \cdot u_0, y_0 = 26 \cdot v_0$. The question now is how to find (u_0, v_0)? This can again be done by guessing (a method applicable when the coefficients of the equation are not too big integers). In our case obviously $3 \cdot (-2) + 7 \cdot 1 = 1$. So a particular solution is $x_0 = -52, y_0 = 26$, and the general solution is given by

$$x = -52 + 7s \qquad y = 26 - 3s \quad s \in \mathbb{Z}.$$

Notice that, at a first glance, the two general solution formulae, corresponding to different particular solutions, look differently. But they **have to produce** the same sets of solutions to the equation (we proved a theorem!). Verify that the sets of solutions are really the same

$$\{(-3 + 7k, 5 - 3k) \,|\, k \in \mathbb{Z}\} = \{(-52 + 7s, 26 - 3s) \,|\, s \in \mathbb{Z}\}.$$

Finally, one can exclude the guessing completely from finding particular solutions! This is done by using Euclid's Algorithm. Recall, it says that integers u_0, v_0 such that $u_0 \cdot m + v_0 \cdot n = \gcd(m, n)$ can be found by "reading" the system of relations in the algorithm backward. In our case we have

$$7 = 3 \cdot 2 + 1 \qquad q_0 = 2, r_0 = 1$$

$$3 = 1 \cdot 3 + 0 \qquad q_1 = 2, r_1 = 0$$

so, as we already know, $\gcd(3, 7) = 1$, and $3 \cdot (-2) + 7 \cdot 1 = 1$. We find $u_0 = -2, v_0 = 1$. □

Remark 5.1.1 As we know from the proof of the theorem above, one easily finds a particular solution to $ax + by = c$ once they know a pair (u_0, v_0) such that $a \cdot u_0 + b \cdot v_0 = \gcd(a, b)$. Such a pair can be found, as we proved, by reading off the Euclid's algorithm backward. In many cases though, mainly when $|a|$ and $|b|$ are reasonably small, one can guess-find some u_0 and v_0. The reason for this is that, as exercises below teach us, there always is (u_0, v_0) with $|u_0| \leq |b|$ and $|v_0| \leq |a|$. □

Exercise 5.1 *(1) Solve the equations*

(i) $91x + 33y = 147$ (iv) $24x + 30y = 14$ (vii) $30x - 43y = 97$

(ii) $93x - 81y = 15$ (v) $17x + 646y = 51$ $(viii)$ $91x + 56y = 0$

(iii) $874x - 19y = 1052$ (vi) $84x - 91y = 11$ (ix) $4147x + 10,672y = 58$

(2) Find all solutions in **positive** *integers to the equations*

(i) $18x + 7y = 302$ (iii) $54x - 38y = 82$ (v) $10x + 28y = 1240$

(ii) $18x - 7y = 302$ (iv) $11x + 13y = 47$ (vi) $17x + 646y = 51$

(3) Find two rational fractions having denominators 5 *and* 7 *respectively, and whose sum is* $26/35$.
(4) Find a number that leaves the remainder 16 *when divided by* 39*, and the remainder* 27 *when divided by* 56.
(5) In this exercise we discuss the size of the coefficients in a Bézout's identity. Let $a, b \in \mathbb{N} \setminus \{0\}$ *with* $(a, b) = d$. *Consider* $u_0, v_0 \in \mathbb{Z}$ *such that* $a \cdot u_0 + b \cdot v_0 = d$. *Show that* $u_0 \cdot v_0 < 0$.

(i) In the case $d = 1$*, prove that there is only one pair* (u_0', v_0') *such that*

$$a \cdot u_0' + b \cdot v_0' = 1 \quad -b < u_0' < 0 \quad \textit{and} \quad 0 < v_0' < a.$$

Similarly, prove that there is only one pair (u_0'', v_0'') *such that*

$$a \cdot u_0'' + b \cdot v_0'' = 1 \quad 0 < u_0'' < b \quad \textit{and} \quad -a < v_0'' < 0.$$

Show also that

$$-u_0' + u_0'' = b \qquad \textit{and} \qquad v_0' - v_0'' = a.$$

(ii) Study the case of* $d \geq 1$ *in general: what can you say of the smallest possible* u_0 *and* v_0 *in this case?*

[Hint to (i): Let $a \cdot u_0 + b \cdot v_0 = 1$. *Divide* u_0 *by* b *with quotient and remainder:* $u_0 = b \cdot q' + r'$ *where* $0 \leq r' < b$. *Similarly,* $v_0 = a \cdot q'' + r''$ *where* $0 \leq r'' < a$. *We have then that*

$$a \cdot r' + b \cdot r'' - 1 = -ab(q' + q'').$$

But since $(u_0, v_0) = 1$*, we have* $r' + r'' > 0$*, and therefore*

$$0 < a \cdot r' + b \cdot r'' - 1 < 2ab,$$

which implies that

$$a \cdot r' + b \cdot r'' - 1 = ab.$$

In other words,

$$a(r' - b) + b \cdot r'' = 1 \qquad \text{and} \qquad a \cdot r' + b(r'' - a) = 1$$

and we can write

$$u_0' = r' - b, \quad v_0' = r'' \qquad \text{and} \qquad u_0'' = r', \quad v_0'' = r'' - a.$$

This proves the existence of the pairs (u_0', v_0') and (u_0'', v_0') together with the relations $u_0'' - u_0' = b$ and $v_0' - v_0'' = a$. To prove the uniqueness, assume that (u, v) has the properties of (u_0', v_0'). Then, we have

$$a \cdot u_0' + b \cdot v_0' = 1 = a \cdot u + b \cdot v$$

and so

$$a(u_0' - u) = b(v - v_0').$$

Since $\gcd(a, b) = 1$, we have that $a \mid v - v_0'$ and $b \mid u_0' - u$ which immediately implies, due to the inequalities u_0', u, v_0' and v satisfy, that $u_0' = u$ and $v_0' = v$. Ans so on...]

(6) Let $a, b \in \mathbb{N} \setminus \{0\}$, and $c \in \mathbb{Z}$ be such that $\gcd(a, b) = 1$. Prove that the equation $ax - by = c$ has infinitely many solutions in positive (respectively - negative) integers.

[Hint: Arithmetic proof is preferable, but a geometric approach would help as well.]

(7) Suppose $a, b \in \mathbb{N} \setminus \{0\}, c \in \mathbb{Z}$ and $\gcd(a, b) = 1$. Consider the equation $ax + by = c$. Prove that

(i) The equation does not have infinitely many solutions in (positive (respectively - negative) integers.

(ii) If $c > ab$ (respectively $c < -ab$), then the equation does have solutions in positive (respectively - negative) integers.

(iii) Prove that the inequality in (ii) is sharp in the following sense: the largest positive (respectively - the smallest negative) number c such that the equation has **no** *solution in positive (respectively - negative) integers is ab, (respectively $-ab$).*

[Hint. In light of item (ii), it is enough to show that $c = ab$ is not representable as a positive linear combination of a and b. RAA: assume $a \cdot u + b \cdot v = ab$ for $u, v > 0$. We have then that

$$b \cdot v = a(b - u) \qquad \text{and} \qquad a \cdot u = b(a - v).$$

So, since $\gcd(a, b) = 1$, we have that $a \mid v$ and $b \mid u$. Therefore, $u = b \cdot u_1, v = a \cdot v_1$ and

$$a \cdot u + b \cdot v = ab(u_1 + v_1) = ab.$$

We have then that $1 = u_1 + v_1 \geq 1 + 1 = 2$ - a contradiction.]

(iv) For every $n \in \mathbb{N}$ there are a, b, c such that the equation has exactly n solutions in positive (respectively - negative) integers.

(8) Find the smallest and the biggest integers c such that the equation $5x + 7y = c$ has exactly nine solutions in positive (respectively - negative) integers.

(9) State and prove a necessary and sufficient condition that the equation $ax + by = c$ have infinitely many solutions in positive (respectively - negative) integers.

(10) Let $a, b \in \mathbb{Z}$ be such that $(a, b) = d$. Verify that if $x_0, y_0 \in \mathbb{Z}$ are such that $a \cdot x_0 + b \cdot y_0 = d$, then $(x_0, y_0) = 1$.

5.2 The General Linear Diophantine Equation

We are addressing here the question of solving the general linear Diophantine equation

$$a_1 x_1 + a_2 x_2 + \cdots + a_k x_k = b$$

where $b \in \mathbb{Z}$ and the coefficient, not all zero, are integers as well: $a_1, a_2, \ldots, a_k \in \mathbb{Z}$. Denote by d the greatest common divisor of the coefficients: $d = gcd(a_1, \ldots, a_k)$. Here is the theorem explaining the precise conditions when the general linear equation has solutions, and how they look like.

Theorem 5.2.1 *The general linear Diophantine equation above has solutions if, and only if,* $d \mid b$.

Proof The proof of this statement easily follows from the properties of the gcd of a group of integers, including the Bézout's identity, and is left as an exercise. □

The question of how to find the solutions to the general linear equation takes a bit longer to answer. The solution has to depend on $k-1$ independent parameters (for $k=2$ we had one parameter). It turns out that solving the case of two variables is enough to solve the general case. Here is how it goes.

Suppose, by induction, that we can solve the general linear equation for $k \geq 2$. We are showing how to find the solutions to the equation when the unknowns are $k+1$. So let the equation be

$$a_1x_1 + \cdots a_kx_k + a_{k+1}x_{k+1} = b.$$

If any of the coefficients is zero, say $a_{k+1} = 0$, we see that any integer can be in the place of x_{k+1} in a solution to the equation: $(x_1^0, \ldots, x_k^0, x_{k+1}^0)$. Also, the k-tuple $(x_1^0, \ldots, x_k^0,)$ is a solution to the equation

$$a_1x_1 + \cdots a_kx_k = b$$

which we know how to solve. So, in this case, the solutions of the original equation are formed by the solutions of the last equation (depending on $k-1$ parameters), and an independent parameter for the value of x_{k+1}^0 (so, the solution depends on the total of k parameters).
Let's assume now that all coefficients are non-zero. Denote by d_k the greatest common divisor of the first k coefficients: $d_k = gcd(a_1, \ldots, a_k)$. Consider the equation

$$d_ky + a_{k+1}x_{k+1} = b$$

By exercise (6) of section 3.2 we have that $d = gcd(d_k, a_{k+1})$. Assuming that $d \mid b$, we get that the last linear equation does have solutions. Moreover, we know how to find them all, by the previous section. Namely, if (y^0, x_{k+1}^0) is a particular solution, then the general one is given by

$$y' = y^0 + (a_{k+1}/d) \cdot s_k \qquad x'_{k+1} = x_{k+1}^0 - (d_k/d) \cdot s_k$$

for s_k - an integer parameter. To find the solution to the original equation it is enough to solve for $x_1, \ldots, x_k$ the equation

$$(a_1/d_k) \cdot x_1 + \cdots (a_k/d_k) \cdot x_k = y'.$$

Since the coefficients of this equation are relatively prime - share no factors bigger than 1 (by exercise (7) of section 3.2), we can solve it, and thus get the solutions for the original linear equation. Note that the solution will depend on k parameters: s_k and $k-1$ more coming from the solution to the last equation.

Example 5.2.1 Let's solve $6x + 15y + 30z = 21$
We check out first if $\gcd(6, 15, 30) \mid 21$. We have $\gcd(6, 15, 30) = \gcd(\gcd(6, 15), 30) = \gcd(3, 30) = 3$. Since $3 \mid 21$ there are solutions to the equation. Denote by $t = 2x + 5y$. For t and z we have the equation

$$3t + 30z = 21$$

the solutions to which we know how to find. We have

$$t = -3 + 10k \qquad z = 1 - k \qquad k \in \mathbb{Z}.$$

To find x and y we have to solve, for every integer k, the equation

$$2x + 5y = -3 + 10k.$$

This we do the same well known way: we find a particular solution, and then use the formulae for the general solution. Obviously, $x_0 = 2 + 5k, y_0 = -1$ is a particular solution. Then we get

$$x = 2 + 5k + 5s \quad y = -1 - 2s \qquad s \in \mathbb{Z}.$$

So the general solution to the original equation is given by

$$x = 2 + 5k + 5s \qquad y = -1 - 2s \qquad z = 1 - k \qquad k, s \in \mathbb{Z}.$$

As expected, the solution depends on two parameters. □

Exercise 5.2 *(1) Solve in integers the equation* $6x + 15y + 20z = 2$.
(2) The same for any equation of three variables you can think of.

Chapter 6

Modular Arithmetic, the Ring $\mathbb{Z}/n\mathbb{Z}$

In studying natural numbers (and the integers), it is often useful to make arguments by using divisibility of integers by a fixed natural number. Solving Diophantine equations is where making such arguments is especially valuable. Consider for instance solving in $\mathbb{Z}$ the equation known to us from Chapter 2

$$x^2 + y^2 = 3z^2.$$

Are there any solutions apart from the obvious one, $(0, 0, 0)$? Geometric arguments tell us that if there is one non-zero solution, then we will find all solutions. The idea behind using divisibility to check if solutions exist is simple and natural. For any solution (a, b, c) we have that the two natural numbers $a^2 + b^2$ and $3c^2$ are equal. But then, for every positive rational number n we would have that the remainder of $a^2 + b^2$ and of $3c^2$ when divided by n are equal. Now, divisibility by 3 immediately shows that non-zero solutions do not exist.

This method of constructing arguments has evolved into what we call Modular Arithmetic. It is based on the concept of **congruences modulo a positive integer** introduced explicitly by Euler in 1850.

In this Chapter, we develop Modular Arithmetic proving the main theorems which will be used later on in the course. Professionally (that is - algebraically) speaking, we are constructing a ring, of n elements, for every natural number $n \geq 1$. When n is a prime number, the corresponding ring turns out to be a field. The first two sections of the chapter are devoted to the definition of these rings, and to establishing some of their most basic, and very important, properties. In the third section of the Chapter, we prove the three standard theorems: Fermat's Little theorem, its generalization due to Euler, and Wilson's theorem. Finally, we prove a property of the rings $\mathbb{Z}/n\mathbb{Z}$ known as the Chinese Remainder Theorem. This theorem will be instrumental in the rest of the course.

6.1 Congruences Modulo $n \in \mathbb{N}$

Let $n \in \mathbb{N}$.

Definition 6.1.1 *We say that the integers a, b are* **congruent modulo** *n, and write $a \equiv b \mod n$, if $n \mid a - b$.*

Exercise 6.1 *When are two integers congruent modulo 0? When are two integers congruent modulo 1?*

Having in disposition the division with quotient and remainder, we easily get the following statement

Exercise 6.2 *Let $n > 0$ be a natural number. For the integers a, b let*

$$a = n \cdot q_1 + r_1 \qquad b = n \cdot q_2 + r_2$$

where $0 \leq r_1, r_2 < n$. Prove that

$$a \equiv b \mod n \quad \Leftrightarrow \quad r_1 = r_2.$$

Proposition 6.1.1 *If $a \equiv b \mod n \quad \wedge \quad a' \equiv b' \mod n$, then*

$$a + a' \equiv b + b' \mod n, \quad \text{and} \quad a \cdot a' \equiv b \cdot b' \mod n.$$

Proof The proof is straightforward. We have $n \,|\, a-b$ and $n \,|\, a'-b'$, so $a-b = n \cdot A$ and $a'-b' = n \cdot B$. But then

$$(a + a') - (b + b') = n \cdot (A - B) \quad \text{and} \quad a \cdot a' - b \cdot b' = n \cdot (b \cdot B + b' \cdot A + n \cdot A \cdot B)$$

which finishes the proof. □

Observe that the cancellation property modulo n

$$ab \equiv ac \mod n \quad \rightarrow \quad b \equiv c \mod n$$

is not true in general, even if $\neg(a \equiv 0 \mod n)$. Here is the right statement.

Proposition 6.1.2 *Let $a, b, c, n \geq 1$ be integers. Then we have*

$$ab \equiv ac \mod n \quad \Leftrightarrow \quad b \equiv c \mod n/\gcd(a, n).$$

In particular

$$(\gcd(a, n) = 1) \quad \rightarrow \quad (ab \equiv ac \mod n \quad \Leftrightarrow \quad b \equiv c \mod n).$$

Proof The second claim follows from the first one in a straightforward way. We are proving the first claim now. Let $d = \gcd(a, n)$. Observe that, since $n \neq 0$, we have also that $d \neq 0$. Denote $n = d \cdot n_1$ and $a = d \cdot a_1$, and recall that $\gcd(a_1, n_1) = 1$. We obviously have the following

$$\begin{aligned} ab \equiv ac \mod n \quad &\Leftrightarrow \quad n \,|\, a(b-c) \quad \Leftrightarrow \quad dn_1 \,|\, da_1(b-c) \\ &\Leftrightarrow \quad n_1 \,|\, a_1(b-c) \quad \Leftrightarrow \quad n_1 \,|\, (b-c) \quad \Leftrightarrow \quad b \equiv c \mod n_1. \end{aligned}$$

Exercise 6.3 *(1) If $a \in \mathbb{Z}$ is even, then $a^2 \equiv 0 \mod 4$. If $a \in \mathbb{Z}$ is odd, then $a^2 \equiv 1 \mod 8$.*
(2) Let $n \in \mathbb{N}$ be an $(n+1)$-digit number (in base 10) with digits $a_0, \ldots a_n$

$$N = a_n 10^n + \cdots + a_1 10 + a_0.$$

Prove that

$$(i)\ N \equiv a_0 \mod 2 \text{ or } 5 \qquad (ii)\ N \equiv a_0 - a_1 + \cdots + (-1)^n a_n \mod 11$$

$$(iii)\ N \equiv a_0 + a_1 + \cdots + a_n \mod 3 \text{ or } 9.$$

(3) Let $n = n_1 n_2 \cdots n_k$ where $n_1, \ldots, n_k$ are pairwise relatively prime. Prove for the integers a, b that

$$a \equiv b \mod n \qquad \Leftrightarrow \qquad (\forall i)(a \equiv b \mod n_i).$$

(4) Let $f(X) = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_1 X + a_0$ be a polynomial with integer coefficients, $a_i \in \mathbb{Z}$, let $a, b \in \mathbb{Z}$, and let $n \in \mathbb{N}$. Prove that if $a \equiv b \mod n$, then $f(a) \equiv f(b) \mod n$.
(5) Let p be an odd prime number, $0 \neq k \in \mathbb{N}$, and $a \in \mathbb{Z}$. Prove that

$$(a^2 \equiv 1 \mod p^k) \qquad \Leftrightarrow \qquad (a \equiv 1 \mod p^k) \vee (a \equiv -1 \mod p^k).$$

(6) Let a and k be as in (5). Prove that
(i) $a^2 \equiv 1 \mod 2$ if, and only if, $a \equiv 1 \mod 2$;
(ii) $a^2 \equiv 1 \mod 2^2$ if , and only if, $a \equiv \pm 1 \mod 2^2$;
(iii) For $k \geq 3$ we have $a^2 \equiv 1 \mod 2^k$ if, and only if, $a \equiv \pm 1 \mod 2^k$ or $a \equiv 2^{k-1} \pm 1 \mod 2^k$.

6.2 The Ring $\mathbb{Z}/n\mathbb{Z}$

The notion of congruence modulo n can be organized in a better construct. Here is how this goes.

Theorem 6.2.1 *The relation on $\mathbb{Z}$ defined by $a \sim_n b$ if $a \equiv b \mod n$ is an equivalence relation. That is*

$$\forall a \in \mathbb{Z}(a \sim_n a); \qquad a \sim_n b \quad \to \quad b \sim_n a; \qquad a \sim_n b \ \wedge \ b \sim_n c \quad \to \quad a \sim_n c.$$

Proof Do it as an easy exercise. □

By the general theory of equivalence relations, we may consider the set of equivalence classes, the quotient set of $\mathbb{Z}$ modulo the relation $\sim_n$,

$$\mathbb{Z}/\sim_n \quad :=: \quad \mathbb{Z}/n\mathbb{Z} := \{[a]_n \,|\, a \in \mathbb{Z}\}$$

where $[a]_n$ stands for the equivalence class of $a \in \mathbb{Z}$ w.r.t. $\sim_n$, that is

$$[a]_n = \{\, b \in \mathbb{Z} \,|\, a \sim_n b \,\}.$$

Whenever there is no danger of confusion, we will write $[a]$ instead of $[a]_n$.

The quotient set $\mathbb{Z}/n\mathbb{Z}$ is finite:

Theorem 6.2.2 *The set $\mathbb{Z}/n\mathbb{Z}$ has n elements*

$$\mathbb{Z}/n\mathbb{Z} = \{\, [0], [1], \ldots, [n-1] \,\}.$$

Proof Indeed, if $a = n \cdot q + r$ where $0 \leq r < n$, then $a \sim_n r$, so that $[a] = [r]$. On the other hand, if $0 \leq r_1, r_2 < n$, then $r_1 \sim_n r_2$ if, and only if, $r_1 = r_2$. Therefore, $\mathbb{Z}/\sim_n$ has as many elements as are the remainders modulo n. □

The important feature of the set $\mathbb{Z}/n\mathbb{Z}$ is that we can define two operations, addition and multiplication, on its elements, in such a way that the result be a ring!

Theorem 6.2.3 *For $[a], [b] \in \mathbb{Z}/n\mathbb{Z}$ define*

$$[a] + [b] := [a + b] \quad \textit{and} \quad [a] \cdot [b] := [a \cdot b].$$

We have then that
(1) the two operations are well defined, that is: do not depend on the choices of representatives of the classes;
(2) the two operations are associative and commutative;
(3) the class $[0]$ is the neutral element for the addition $+$, while the class $[1]$ is neutral element for the multiplication $\cdot$;
(4) every element has an opposite one: $\forall [a] \exists [b] ([a] + [b] = [0])$;
(5) the multiplication distributes over the addition: $[a] \cdot ([b] + [c]) = [a] \cdot [b] + [a] \cdot [c]$.
Therefore, $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ is a (commutative) ring (with identity).

Proof (1). This item states the same as Proposition 6.1.1 above. Indeed, we have to show that if

$$a_1, a_2 \in [a], \quad \text{and} \quad b_1, b_2 \in [b],$$

then

$$[a_1] + [b_1] = [a_2] + [b_2] \quad \text{and} \quad [a_1] \cdot [b_1] = [a_2] \cdot [b_2].$$

But this is equivalent to saying that

$$a_1 \equiv a_2 \mod n \quad \text{and} \quad b_1 \equiv b_2 \mod n$$

implies that

$$a_1 + b_1 \equiv a_2 + b_2 \mod n \quad \text{and} \quad a_1 \cdot b_1 \equiv a_2 \cdot b_2 \mod n$$

which is exactly the content of Proposition 6.1.1.
Items (2) and (3) are obvious. As to item (4), here is a proof

$$[a] \cdot ([b] + [c]) = [a] \cdot [b + c] = [a \cdot (b + c)]$$

$$= [a \cdot b + a \cdot c] = [a \cdot b] + [a \cdot c] = [a] \cdot [b] + [a] \cdot [c]. \quad \square$$

Notice that it is possible to have $[a] \neq [0] \neq [b]$ and still $[a] \cdot [b] = [0]$. Indeed, if $n = n_1 \cdot n_2$ is composite ($n_1, n_2 > 1$), then $[n_1] \neq [0] \neq [n_2]$, but $[n_1] \cdot [n_2] = [n_1 \cdot n_2] = [n] = [0]$. In the congruence notations, this looks like

$$\neg(a \equiv 0 \mod n) \wedge \neg(b \equiv 0 \mod n), \quad \text{but} \quad a \cdot b \equiv 0 \mod n.$$

In professional, that is - algebraic, terms, this means that $\mathbb{Z}/n\mathbb{Z}$ has **non-zero zero divisors** when n is composite. It turns out that the case when $n = p$ is a prime number is much gentler.

Theorem 6.2.4 *Let $n = p$ be a prime number, and let $[0] \neq [a] \in \mathbb{Z}/p\mathbb{Z}$. Then, there is (a unique) $[b] \in \mathbb{Z}/p\mathbb{Z}$ such that $[a] \cdot [b] = [1]$, that is, $[b]$ is the reciprocal of $[a]$ in $\mathbb{Z}/p\mathbb{Z}$. In other, professional, words, $\mathbb{Z}/p\mathbb{Z}$ is a field.*

Proof We have $[a] \neq [0] \leftrightarrow p \nmid a$, and so $gcd(p, a) = 1$. The Bézout's identity gives us $b, c \in \mathbb{Z}$ such that $a \cdot b + p \cdot c = 1$. Therefore, $p \mid a \cdot b - 1$, and according to the definition of $\sim_p$, $[a \cdot b] = [1]$. So, $[a] \cdot [b] = [1]$. On the other hand, if $[b]$ and $[b']$ are reciprocals of $[a]$, then

$$[b] = [1] \cdot [b] = ([b'] \cdot [a]) \cdot [b] = [b'] \cdot ([a] \cdot [b]) = [b'] \cdot [1] = [b']. \quad \square$$

As a consequence of this theorem, we get the cancellation property of $\mathbb{Z}/p\mathbb{Z}$:

$$([a] \cdot [b] = [a] \cdot [c] \quad \wedge \quad [a] \neq [0]) \quad \rightarrow \quad [b] = [c].$$

We know already that this property is not true for $\mathbb{Z}/n\mathbb{Z}$ for a composite n. The cancellation property in this case reads as follows.

Theorem 6.2.5 *In $\mathbb{Z}/n\mathbb{Z}$, every $[a]$ with $gcd(n, a) = 1$ has a (unique) reciprocal, that is $[b]$ such that $[a] \cdot [b] = [1]$. Moreover, we have the cancellation property*

$$([a] \cdot [b] = [a] \cdot [c] \quad \wedge \quad gcd(n, a) = 1) \quad \rightarrow \quad [b] = [c].$$

Proof Do the proof as an exercise, using the Bézout's identity as in the proof of Theorem 6.2.4 above. $\square$

Remark 6.2.1 The definition of the classes modulo n can be extended from **integers**, $[a], a \in \mathbb{Z}$, to **rational fractions**, $[a/b]$ where $a \in \mathbb{Z}, b \in \mathbb{N}$, and $\gcd(b, n) = 1$. By definition, the class $[a/b]$ is such that $[b][a/b] = [a]$ in $\mathbb{Z}/n\mathbb{Z}$. The cancellation property modulo n ensures that this definition is correct: the class exists and is unique. With this definition, we can write

$$a/b \equiv c \mod n$$

when $c \in [a/b]$. We will use this extension, and the notations, in the exercises to follow. $\square$

Exercise 6.4 *(1) For which natural numbers n is the number $3^n + 1$ divisible by 10?*
(2) Find the remainder of the division of $1! + 2! + \cdots + 50!$ by 7.
(3) Is it true that 36 divides $n^4 + n^2 + 4$ for infinitely many natural numbers n? Explain!
(4) What are the possible values of the last digit of $4^m, m \in \mathbb{N}$?

(5) Show that a three digit natural number written as abc is divisible by 7 if, and only if, $2a+3b+c$ is divisible by 7.
(6) Let p be an odd prime number, and let $k \in \{1, \ldots, p-1\}$. Prove that

$$\binom{p}{k} \equiv 0 \mod p \qquad \binom{p-1}{k} \equiv (-1)^k \mod p.$$

(7) Let p be an odd prime number. Prove that

$$1/1 + \cdots + 1/k + \cdots + 1/(p-1) \equiv 0 \mod p.$$

(8) Let p be an odd prime number. Prove that $p \,|\, 2^p - 2$, and that

$$1/1 - 1/2 + \cdots + (-1)^{k-1}1/k + \cdots - 1/(p-1) \equiv (2^p - 2)/p \mod p.$$

6.3 Fermat, Euler, and Wilson

We are proving here three properties of $\mathbb{Z}/n\mathbb{Z}$ known as Fermat's Little Theorem, Euler's Theorem, and Wilson's Theorem.

6.3.1 Fermat's Little Theorem

Lemma 6.3.1 *Let $[0] \neq [a] \in \mathbb{Z}/p\mathbb{Z}$. We have the following equality of sets*

$$\{[1], [2], \ldots, [p-1]\} = \{[a]\cdot[1], [a]\cdot[2], \ldots, [a]\cdot[p-1]\}.$$

Proof Indeed, by the cancellation property in $\mathbb{Z}/p\mathbb{Z}$, we have that

$$[a]\cdot[i] = [a]\cdot[j] \quad \rightarrow \quad [i] = [j].$$

Therefore, the RHS-set has $p-1$ pairwise distinct non-zero elements. This means that it contains all the non-zero classes modulo p. □

Theorem 6.3.2 (*Fermat's Little Theorem (FLT), (1640)*) *Let p be a prime number. For every $[a] \neq [0]$ in $\mathbb{Z}/p\mathbb{Z}$, we have*

$$[a]^{p-1} = [1].$$

Proof By the Proposition above,

$$([a]\cdot[1])\cdot([a]\cdot[2])\cdots([a]\cdot[p-1]) = [1]\cdot[2]\cdots[p-1].$$

Therefore,

$$[a]^{p-1}\cdot[(p-1)!] = [(p-1)!]$$

which, after cancelling out $[(p-1)!] \neq [0]$ gives the result. □

Exercise 6.5 *(1) Prove that, for any two integers a and b, and a prime number p,*

$$(a+b)^p \equiv a^p + b^p \mod p.$$

Prove also that if $a \equiv b \mod p$, then $a^p \equiv b^p \mod p^2$.
(2) Let p be a prime number. Show that $\binom{2p}{p} \equiv 2 \mod p$.
(3) Let p be a prime number and let $a \in \mathbb{Z}$. Show that $p \,|\, (a^p + a(p-1)!)$.
(4) Let p be an odd prime number. Show that

$$1^p + 2^p + \cdots (p-1)^p \equiv 0 \mod p,$$

and that

$$1^{p-1} + 2^{p-1} + \cdots + (p-1)^{p-1} \equiv -1 \mod p.$$

6.3.2 Euler's Theorem, Reduced Residue System, Euler's Totient Function

The idea of proving Fermat's Little Theorem can easily be generalized to the case of powers of elements of $\mathbb{Z}/n\mathbb{Z}$ for any $n \in \mathbb{N}$. Indeed, notice first of all that, as in the case of $\mathbb{Z}/p\mathbb{Z}$, we have equality of sub-sets of $\mathbb{Z}/n\mathbb{Z}$:

$$\{[a][1], [a][2], \ldots, [a][n-1]\} = \{[1], [2], \ldots, [n-1]\}$$

for any $[a]$ such that $gcd(a, n) = 1$. (Verify this! It is based on the corresponding cancellation property.) So, as in the case of $\mathbb{Z}/p\mathbb{Z}$, we have that

$$[a]^{n-1}[(n-1)!] = [(n-1)!] \quad \text{in} \quad \mathbb{Z}/n\mathbb{Z}.$$

One may be tempted, applying the cancellation property, to infer from here that $[a]^{n-1} = [1]$ in $\mathbb{Z}/n\mathbb{Z}$. There is a problem though! The cancellation property is not applicable for composite numbers n: in fact, $gcd((n-1)!, n) > 1$ in this case.

The right thing to do in order to use the idea from the proof of FLT is to consider the set

$$(\mathbb{Z}/n\mathbb{Z})^\times := \{[a] \in \mathbb{Z}/n\mathbb{Z} \mid \gcd(a, n) = 1\}.$$

Definition 6.3.1 *The cardinality of set $(\mathbb{Z}/n\mathbb{Z})^\times$ is denoted by $\varphi(n)$ and is called* **Euler's phi-function** *or* **Euler's totient function**.

Obviously, $\varphi(n)$ is the number of all natural numbers, which are less than n and are relatively prime with n. Also, for a prime number p, we have $\varphi(p) = p-1$. Very soon, in the context of the **Chinese Remainder Theorem**, we will establish a formula for $\varphi(n)$ for all $n \in \mathbb{N}$.

Definition 6.3.2 *The integers $b_1, \ldots, b_n$ represent a* **complete residue system** *modulo n if they are pairwise incongruent modulo n.*
The integers $a_1, \ldots, a_{\varphi(n)}$ represent a **reduced residue system** *modulo n if they are relatively prime to n, and are pairwise incongruent modulo n.*

Let $(\mathbb{Z}/n\mathbb{Z})^\times = \{[a_1], [a_2], \ldots, [a_{\varphi(n)}]\}$.

Proposition 6.3.3 *(1) For every two integers a and b we have*

$$\gcd(ab, n) = 1 \quad \Leftrightarrow \quad \gcd(a, n) = 1 \wedge \gcd(b, n) = 1.$$

(2) For every $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ we have the equality of sets

$$\{[a_1], [a_2], \ldots, [a_{\varphi(n)}]\} = \{[a] \cdot [a_1], [a] \cdot [a_2], \ldots, [a] \cdot [a_{\varphi(n)}]\}.$$

(3) The set $(\mathbb{Z}/n\mathbb{Z})^\times$ consists of all elements of $\mathbb{Z}/n\mathbb{Z}$ which have reciprocal elements

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[a] \in \mathbb{Z}/n\mathbb{Z} \mid (\exists b \in \mathbb{Z})([a][b] = [1])\}.$$

Proof (1) ($\Rightarrow$) $\gcd(ab, n) = 1$ so, there are integers u, v such that $u(ab) + vn = 1$. This identity can be read in two ways:

$$(ub)a + vn = 1 \quad \text{and} \quad (ua)b + vn = 1$$

which in turn imply that $\gcd(a, n) = 1$ and $\gcd(b, n) = 1$. ($\Leftarrow$) The relations $\gcd(a, n) = 1$ and $\gcd(b, n) = 1$ imply that, for some $u_1, u_2, v_1, v_2 \in \mathbb{Z}$,

$$u_1 a + v_1 n = 1 \quad \text{and} \quad u_2 b + v_2 n = 1.$$

We have then that

$$1 = (u_1 a + v_1 n)(u_2 a + v_2 n) = (u_1 u_2) \cdot ab + (u_1 a v_2 + u_2 b v_1 + v_1 v_2 n) \cdot n$$

which implies that $gcd(ab, n) = 1$ as needed.
(2) By item (1) we have that

$$\{[a] \cdot [a_1], [a] \cdot [a_2], \ldots, [a] \cdot [a_{\varphi(n)}]\} \subseteq \{[a_1], [a_2], \ldots, [a_{\varphi(n)}]\}.$$

By the cancellation property in $\mathbb{Z}/n\mathbb{Z}$ we have that the elements of the LHS are pairwise disjoint: $[a][a_i] = [a][a_j] \to [a_i] = [a_j]$. Since the LHS and the RHS sets are finite and of same cardinality, they are equal.
(3) We have to show that

$$\{[a] \in \mathbb{Z}/n\mathbb{Z} \,|\, \gcd(a, n) = 1\} = \{[a] \in \mathbb{Z}/n\mathbb{Z} \,|\, (\exists b \in \mathbb{Z})([a][b] = [1])\}.$$

Showing that $LHS \subseteq RHS$. Let $[a] \in LHS$, that is $a \in \mathbb{Z}$ such that $\gcd(a, n) = 1$. By Bézout, there are $u, v \in \mathbb{Z}$ such that $ua + vn = 1$. This means that $[ua] = [1]$, and therefore $[a][u] = [1]$, and the element $[a]$ has a reciprocal in $\mathbb{Z}/n\mathbb{Z}$. This implies that $[a] \in RHS$. Showing that $RHS \subseteq LHS$. Left as an exercise! So, $LHS = RHS$ as needed. □

Corollary 6.3.4 *If $[a]$ and $[b]$ are elements of $(\mathbb{Z}/n\mathbb{Z})^\times$, then so are $[a][b]$ and the reciprocals of $[a]$ and $[b]$.*

Proof An easy exercise. □

Remark 6.3.1 In professional, that is Algebraic, terminology the claim of the Proposition means that $(\mathbb{Z}/n\mathbb{Z})^\times$ is a group with respect to the multiplication operation. It is called the **group of units** of $\mathbb{Z}/n\mathbb{Z}$. □

Theorem 6.3.5 (*Euler (1736)*) *For any $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$*

$$[a]^{\varphi(n)} = [1] \quad \textit{in} \quad \mathbb{Z}/n\mathbb{Z}.$$

Remark 6.3.2 In congruence notations, Euler's theorem sounds like this. For every $a \in \mathbb{Z}$ such that $gcd(a, n) = 1$ we have that

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

Proof of Theorem 6.3.5 Using item (2) of the previous proposition, we get that

$$[a]^{\varphi(n)} \cdot ([a_1][a_2] \cdots [a_{\varphi(n)}]) = [a_1][a_2] \cdots [a_{\varphi(n)}] \quad \text{in} \quad \mathbb{Z}/n\mathbb{Z}$$

and we now **may** cancel out $[a_1][a_2] \cdots [a_{\varphi(n)}]$, because it, being an element of $(\mathbb{Z}/n\mathbb{Z})^\times$, is subject to the cancellation property of $\mathbb{Z}/n\mathbb{Z}$. □

Exercise 6.6 *(1) If p and q are distinct prime numbers, is it true that we always have*

$$p^{q-1} + q^{p-1} \equiv 1 \mod pq \quad ?$$

More generally, if $m, n \in \mathbb{N}$ are relatively prime, is it true that

$$n^{\varphi(m)} + m^{\varphi(n)} \equiv 1 \mod nm \quad ?$$

(2) Show that $(\forall n \in \mathbb{N})\,(3^{2n+2} \equiv 8n + 9 \mod 64)$.
(3) Let m, and $a > 1$ be natural numbers such that $(a, m) = (a - 1, m) = 1$. Show that

$$1 + a + a^2 + \cdots + a^{\varphi(m)-1} \equiv 0 \mod m.$$

6.3.3 Wilson's Theorem

Wilson's theorem reveals one more property of the elements of the group $(\mathbb{Z}/p\mathbb{Z})^\times$.

Theorem 6.3.6 (*Wilson (1770)*) *For any prime number p we have*

$$[1][2]\cdots[p-1] = [-1] \quad in \quad \mathbb{Z}/p\mathbb{Z}.$$

Or, in congruence notations,

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

Proof We know that in $(\mathbb{Z}/p\mathbb{Z})^\times$ every element has a reciprocal:

$$\forall [a]\ \exists [b]([a][b] = [1]).$$

So, for all $[a]$ for which $[b] \neq [a]$ their product in $[1][2]\cdots[p-1]$ will produce a $[1]$. Therefore, the product $[1][2]\cdots[p-1]$ reduces to the product of only the elements $[a]$ with reciprocals equal to themselves, that is to all of which $[a][a] = [1]$. But $[a]^2 = [1]$ is equivalent to $p \,|\, (a^2-1)$, that is to $p \,|\, (a-1)(a+1)$. Since p is prime this latter is in turn equivalent to $p \,|\, a-1$ or $p \,|\, a+1$. Therefore, there are only two elements of $(\mathbb{Z}/p\mathbb{Z})^\times$ which coincide with their reciprocals: $[a] = [1]$ and $[a] = p-1$. We have now that

$$[1][2]\cdots[p-1] = [1][p-1] = [p-1] = [-1]. \quad \square$$

Keeping up with the idea of generalizing results modulo p to the general case of modulo n, the next step here would be to generalize Wilson's theorem as well, and show that $[a_1]\cdot[a_2]\cdots[a_{\varphi(n)}] = -[1]$ in $\mathbb{Z}/n\mathbb{Z}$. This however is not true in general: the answer depends on n.

Example 6.3.1 The natural numbers less than 15 and relatively prime with 15 are $1, 2, 4, 7, 8, 11, 13$, and 14. So, in particular, $\varphi(15) = 8$. It is easy to check that in $\mathbb{Z}/15\mathbb{Z}$

$$[1][2][4][7][8][11][13][14] = [1] \quad \square$$

Exercise 6.7 * *Prove, nevertheless, that actually*

$$\prod_{[a]\in(\mathbb{Z}/n\mathbb{Z})^\times} [a] = \pm[1] \quad in \quad \mathbb{Z}/n\mathbb{Z}.$$

Prove, more precisely, that the value of the product is $[-1]^{s/2}$ where s is the number of elements $[a] \in \mathbb{Z}/n\mathbb{Z}$ such that $[a]^2 = [1]$. (In particular, s is even!) [Hint: Study the proof of Wilson's theorem, and adopt the idea from there.]

To find the exact value of the product of elements of a reduced system of remainders modulo n, equivalently - to find s from the last Exercise, is harder. Later on in this course we will find two different ways to do that.

Exercise 6.8 *(1) Let p be an odd prime number. Show that*

$$\sum_{k=1}^{p-1} (k-1)!(p-k)!k^{p-1} \equiv 0 \mod p.$$

(2) Let p be a prime number. Show that, for every integer k such that $0 < k < p$,

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

6.4 The Chinese Remainder Theorem

This theorem, which was really known to the ancient Chinese mathematicians, has numerous generalizations and vast applications throughout modern Algebra. Our treatment of this theorem, while kept elementary, will point into its algebraic nature, and will prepare the ground to some of its generalizations in the course Topics in Algebraic Structures.

6.4.1 Naive Set-up

Suppose, we want to find an integer x which satisfies two congruences simultaneously, say

$$x \equiv a \mod m \qquad \text{and} \qquad x \equiv b \mod n.$$

We find x by making it satisfy them one at a time, of course:

$$x = a + k \cdot m, \quad k \in \mathbb{Z} \qquad \text{from the first congruence}$$

and then, substituting in the second, we find all k for which x satisfies the second congruence as well

$$k \cdot m - a \equiv b \mod n \qquad \text{or better} \qquad m \cdot k \equiv b - a \mod n.$$

The last congruence means that, for some integer s we have $m \cdot k + n \cdot s = b - a$, or equivalently, k is the X-component of the solution to the linear Diophantine equation

$$mX + nY = b - a.$$

We know very well now that the equation has a solution precisely when $\gcd(m, n) \mid b - a$. So, when the latter happens, and only then, we find the needed x.

The right question to ask here would be the following: **for what m and n does such an x always (that is, for all integers a and b) exist**? The answer obviously is: whenever $gcd(m, n) = 1$.

Exercise 6.9 *Assuming that $gcd(m, n) = 1$, find the solutions to problem above, and show it is unique modulo mn.*

We can formulate our finding as follows.

Theorem 6.4.1 (**Chinese Remainder Theorem**) *Let $m, n \in \mathbb{N}$ be relatively prime. Then, for every $a, b \in \mathbb{Z}$, there is a unique modulo mn integer x satisfying the congruences*

$$x \equiv a \mod m \qquad x \equiv b \mod n.$$

6.4.2 Set Theoretical Approach

A step closer to the "professional" take of the Chinese Remainder Theorem is to look ate it from set-theoretical point of view. Let $m, n \in \mathbb{N}$ be positive. Consider the sets $\mathbb{Z}/(mn)\mathbb{Z}, \mathbb{Z}/m\mathbb{Z}$, and $\mathbb{Z}/n\mathbb{Z}$. The truth of the matter is that there are maps

$$f : \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \quad \text{and} \quad g : \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$$

such that

$$f([a]_{mn}) = [a]_m \quad \text{and} \quad g([a]_{mn}) = [a]_n.$$

(**Verify that these are well defined as an exercise**!) But then, there is a map

$$F : \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

defined by

$$F([a]_{mn}) = (f([a]_{mn}), g([a]_{mn})) = ([a]_m, [a]_n).$$

This map is neither one-to-one, nor onto in general. What is true, and is equivalent to the Chinese Remainder Theorem, is the following

Theorem 6.4.2 (*Chinese Remainder Theorem*) *The map $F : \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is one-to-one if, and only if, it is onto, if, and only if, $gcd(m, n) = 1$.*

Proof The equivalence of the properties of F being one-to-one and onto follows from the fact that the domain and the co-domain of F have the same finite cardinality: mn. The equivalence of F being one-to-one and $gcd(m, n) = 1$ can be seen as follows. If $gcd(m, n) = 1$, let $F([a]_{mn}) = F([b]_{mn})$. This means that $[a]_m = [b]_m$ and $[a]_n = [b]_n$, or, equivalently, that $m \,|\, a - b$ and $n \,|\, a - b$. But since $gcd(m, n) = 1$, we get that $mn \,|\, a - b$ which means that $[a]_{mn} = [b]_{mn}$. So, F is one-to-one. In the opposite direction, assuming that F is one-to-one, observe that $F([lcm(m, n)]_{mn}) = ([0]_m, [0]_n)$, and therefore, $[lcm(m, n)]_{mn} = [0]_{mn}$, that is $mn \,|\, lcm(m, n)$. We know that $lcm(m, n) = (mn)/gcd(m, n)$, and so $gcd(m, n) = 1$. □

Remark 6.4.1 Notice that **the statement that F is a bijection is equivalent to the statement in the Chinese Remainder Theorem (CRT)**. Indeed, the CRT says that for every $a, b \in \mathbb{Z}$, there is a unique $c \in \mathbb{Z}$ modulo mn such that $c \equiv a \mod m$ and $c \equiv b \mod n$. This means precisely that there is a unique class $[c]_{mn}$ such that $[c]_m = [a]_m$ and $[c]_n = [b]_n$, that is, F is a bijection. □

Exercise 6.10 *These exercises show two ways how to find $[c]_{mn}$ such that $[c]_m = [a]_m$ and $[c]_n = [b]_n$.*
(1) Let $m, n \in \mathbb{N}$ be relatively prime, and let

$$F : \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

be the map from the CRT above. Prove that for any $a, b \in \mathbb{Z}$

$$F(a \cdot n^{\varphi(m)} + b \cdot m^{\varphi(n)}) = ([a]_m, [b]_n).$$

(2) Keep the notation of (1). Prove that there are $A, B \in \mathbb{Z}$ be such that

$$An \equiv 1 \mod m \qquad \text{and} \qquad Bm \equiv 1 \mod n.$$

Prove further that

$$F([aAn + bBm]_{mn}) = ([a]_m, [b]_n).$$

6.4.3 More General Case

An easy induction by the number of the factors, $k \geq 2$, of $n = n_1 n_2 \cdots n_k$ gives the following generalization of the Chinese Remainder Theorem.

Theorem 6.4.3 (*Chinese Remainder Theorem*) *If $n_1, n_2, \cdots, n_k$ are pairwise relatively prime, and if $n = n_1 n_2 \ldots n_k$, then the map*

$$F : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$$

given by $F([a]_n) = ([a]_{n_1}, [a]_{n_2}, \ldots, [a]_{n_k})$ is a bijection.

Proof The proof goes by induction on $k \geq 1$. The case $k = 1$, and $n = n_1$, is trivial: $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/n_1\mathbb{Z}$, and F is the identity map. The case $k = 2$ is the statement of the Chinese Remainder Theorem discussed and proved in the previous subsection. Assume, by way of induction, that the statement is true for $k = m \geq 2$. We are showing that it is true for $k = m + 1$ as well. Let $n_1, \ldots, n_m, n_{m+1}$ be pairwise relatively prime. Then, the first m of them are also pairwise relatively prime, and by the induction hypothesis the map

$$F' : \mathbb{Z}/n'\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_m\mathbb{Z}$$

given by $[a]_{n'} \mapsto ([a]_{n_1}, \ldots, [a]_{n_m})$, where $n' = n_1 \cdots n_m$, is a bijection. On the other hand, $gcd(n', n_{m+1}) = 1$ as well, and the CRT for $k = 2$ gives the bijection

$$F'' : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n'\mathbb{Z} \times \mathbb{Z}/n_{m+1}\mathbb{Z}$$

where $G([a]_n) = ([a]_{n'}, [a]_{n_{m+1}})$ where $n = n'n_{m+1} = n_1 \cdots n_m \cdot n_{m+1}$. We can combine the maps F' and F'' to finish the proof. Consider the composition of maps

$$\mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n'\mathbb{Z} \times \mathbb{Z}/n_{m+1}\mathbb{Z} \longrightarrow \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_m\mathbb{Z} \times \mathbb{Z}/n_{m+1}\mathbb{Z}$$

where the first arrow is the map F'' while the second is the map $(F', Id_{\mathbb{Z}/n_{m+1}\mathbb{Z}})$. Notice that both these maps are bijections, and as such their composition $F = (F', Id_{\mathbb{Z}/n_{m+1}\mathbb{Z}}) \circ F''$ is a bijection as well. It is straightforward to check that F maps $[a]_n$ to $([a]_{n_1}, \ldots, [a]_{n_m}, [a]_{n_{m+1}})$ and is the map we wanted to show is a bijection for $k = m + 1$. The theorem is proved. □

In the notations of congruences, the theorem says that, under the restrictions on $n_1, \ldots, n_k$, for any integers $a_1, \ldots, a_k$ there is a unique modulo $n_1 \cdots n_k$ integer c such that

$$c \equiv a_1 \mod n_1 \qquad c \equiv a_2 \mod n_2 \qquad \cdots \qquad c \equiv a_k \mod n_k.$$

Exercise 6.11 *Let $n_1, \ldots, n_k$ be pairwise relatively prime natural numbers. For $n = n_1 \cdots n_k$, consider the map*

$$F : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$$

as in CRT. We know that for any $a_1, \ldots, a_k \in \mathbb{Z}$, there is a unique modulo n integer c such that $F([c]_n) = ([a_1]_{n_1}, \ldots, [a_k]_{n_k})$. Design two formulae for c. [Hint: Mimic the formulae for $k = 2$ established in the previous subsection.]

6.4.4 A Professional (that is, Algebraic) Approach to the Above

Turns out, we have proved so far in this section more than claimed. In this final subsection, we are giving the things we saw above the right names. We are giving them the right treatment as well.

To start off, recall the $\mathbb{Z}/n\mathbb{Z}$ is a ring with the operations addition and multiplication. We also know the set $(\mathbb{Z}/n\mathbb{Z})^\times$ is a group with the operation multiplication. Indeed, the product of every two elements of $(\mathbb{Z}/n\mathbb{Z})^\times$ belongs there as well, the inverse (reciprocal) of every element of $(\mathbb{Z}/n\mathbb{Z})^\times$ belongs there too, and the class $[1]_n$ is the neutral element with respect to the operation. This group is also commutative: $[a][b] = [b][a]$ for every two elements of $(\mathbb{Z}/n\mathbb{Z})^\times$.

Now, the Cartesian product $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ of the two rings $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$ is also a ring with the following two (component-wise) operations

$$([a]_m, [b]_n) + ([a_1]_m, [b_1]_n) := ([a]_m + [a_1]_m, [b]_n + [b_1]_n)$$

and

$$([a]_m, [b]_n) \cdot ([a_1]_m, [b_1]_n) := ([a]_m \cdot [a_1]_m, [b]_n \cdot [b_1]_n).$$

Verifying this, that is checking that the two operations are associative, commutative, have neutral elements, every element has an opposite, and that multiplication distributes over addition, is a bit lengthy, but absolutely routine. Generalizing this construction to the make a ring out of the Cartesian product of any (finite) number of rings

$$\mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z},$$

with component-wise operations addition and multiplication, is straightforward. It is left as an exercise.

The map that we designed in the previous subsections, for $n = n_1 n_2 \cdots n_k$,

$$F : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$$

respects the operations of both the domain and the co-domain sending sum to sum and product to product:

$$F([a]_n + [b]_n) = F([a]_n) + F([b]_n) \quad \text{and} \quad F([a]_n \cdot [b]_n) = F([a]_n) \cdot F([b]_n),$$

and also sends the identity of the domain to the identity of the co-domain

$$F([1]_n) = ([1]_{n_1}, [1]_{n_2}, \ldots, [1]_{n_k}).$$

In professional terms, this means that F is a **ring homomorphism**. Moreover, as we know, when $n_1, n_2, \ldots, n_k$ are pairwise relatively prime, the map F is also a **bijection**. This, in addition to F being a ring homomorphism, makes it a **ring isomorphism**. This means that in this case the domain and the co-domain are identifiable, via F, **as rings**. Every property of rings one can prove for any one of these two rings, will automatically be true for the other one! In particular, when $n_1 = p_1^{\alpha_1}, n_2 = p_2^{\alpha_2}, \ldots, n_k = p_k^{\alpha_k}$ where $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ is the canonical factorization of n, we get

$$F : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \mathbb{Z}/p_2^{\alpha_2}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_k^{\alpha_k}\mathbb{Z}$$

is a ring isomorphism, and to understand the properties of the domain, one might try to use the properties of each of the factors of the co-domain! For instance, an invertible element of $\mathbb{Z}/n\mathbb{Z}$ corresponds to an invertible element in product of rings, which in turn corresponds to an ordered k-tuple of invertible elements, one for each factor $\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}$. We will use this observation in the next section to establish an important property of the totient function φ.

Exercise 6.12 *(1) Find all integers x, y, z such that $2 \le x \le y \le z$ and*

$$xy \equiv 1 \mod z, \quad yz \equiv 1 \mod x, \quad zx \equiv 1 \mod y.$$

(2) Let $p \ge 5$ be a prime number. Compute $\gcd(p!, (p-2)! - 1)$. [Hint: Wilson.]
(3) Show that there do not exist natural numbers m, n such that $1 + n + n^2 = m^2$.
(4) Let $n = n_1 \cdots n_k$ be a product of pairwise relatively prime positive integers. Let s_i be the number of $[b] \in \mathbb{Z}/n_i\mathbb{Z}$ such that $[b]^2 = [1]$ in $\mathbb{Z}/n_i\mathbb{Z}$. Prove that $s_1 \cdots s_k$ is the number of $[a] \in \mathbb{Z}/n\mathbb{Z}$ such that $[a]^2 = [1]$ in $\mathbb{Z}/n\mathbb{Z}$.*
*(5)** Applying the result in (4) to the canonical representation of n conclude that*

$$\prod_{[a] \in (\mathbb{Z}/n\mathbb{Z})^\times} [a] = -[1] \quad \text{in} \quad \mathbb{Z}/n\mathbb{Z}$$

if, and only if, $n = 2, 4, p^\alpha$ or $2p^\alpha$ where p is an odd prime, and $\alpha \ge 1$. This is a generalization of Wilson's theorem due to Gauss (1801).

[Hint: Exercises (5) and (6) of Sect 6.1, and Exercise (3) from 6.3.3 could be helpful.]

6.5 Vista: Proof of Legendre's Theorem

Recall the Legendre's theorem 2.3.1:

Theorem 6.5.1 (*Legendre*, 1785) *Let a, b and c be tree integers, not all of the same sign, and such that abc is square-free. Then, the equation*

$$ax^2 + by^2 + cz^2 = 0$$

has a solution in integers, not all 0, if, and only if, $-ab, -bc$, and $-ca$ are quadratic residues modulo $|c|, |a|$, and $|b|$ respectively.

The necessity of the conditions on a, b, and c is easy to see. Indeed, if there is a non-trivial solution, then, since a, b, and c are pairwise relatively prime, there is a solution with pairwise relatively prime components: (x_0, y_0, z_0). Consider the relation $ax_0^2 + by_0^2 + cz_0^2 = 0$ modulo, say, c. We have

$$ax_0^2 + by_0^2 \equiv 0 \mod c.$$

Since $\gcd(x_0, c) = 1$, we have also that

$$ab + b^2y_0^2t_0^2 \equiv 0 \mod c$$

where t_0 is the reciprocal of x_0 modulo c. Then, obviously, $-ab$ is a quadratic residue modulo c. In a similar way, we prove the other two conditions are true as well.

To prove the sufficiency of these conditions, we need to work more.

6.5.1 Factoring $ax^2 + by^2 + cz^2$ in $\mathbb{Z}/(abc)\mathbb{Z}$

We are proving now that the conditions on the coefficients: abc - square free, and $-ab, -bc$, and $-ca$ are quadratic residues modulo $|c|, |a|$, and $|b|$ respectively, implies that the polynomial $ax^2+by^2+cz^2$ is factorizable modulo abc.

Theorem 6.5.2 *With the conditions on a, b and c listed above, there are integers a_1, a_2, b_1, b_2, c_1 and c_2 such that*

$$(a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z) \equiv ax_2 + by^2 + cz^2 \mod abc.$$

Proof We do the proof in two steps.

- **Factoring $ax^2 + by^2 + cz^2$ in $\mathbb{Z}/a\mathbb{Z}$, $\mathbb{Z}/b\mathbb{Z}$, and $\mathbb{Z}/c\mathbb{Z}$**

Consider first $\mathbb{Z}/a\mathbb{Z}$. By our assumption, $-bc \equiv u^2 \mod a$. From this it follows that $b \equiv -c(u/c)^2 \mod a$, and therefore that

$$ax^2 + by^2 + cz^2 \equiv -c(\frac{u}{c})^2y^2 + cz^2 \equiv (0x - uy + cz)(0x + \frac{u}{c}y + z) \mod a.$$

In a similar way we get that

$$ax^2 + by^2 + cz^2 \equiv (-vx + 0y + cz)(\frac{v}{c}x + 0y + z) \mod b$$

and

$$ax^2 + by^2 + cz^2 \equiv (-wx + by + 0z)(\frac{w}{b}x + y + 0z) \mod c.$$

- **Factoring $ax^2 + by^2 + cz^2$ in $\mathbb{Z}/(abc)\mathbb{Z}$**

With the help of the CRT, we find integers a_1, b_1 and c_1 such that

$$a_1 \equiv 0 \mod a, \quad a_1 \equiv -v \mod b, \quad a_1 \equiv -w \mod c$$

$$b_1 \equiv -u \mod a, \quad b_1 \equiv 0 \mod b, \quad b_1 \equiv 1 \mod c$$

$$c_1 \equiv c \mod a, \quad c_1 \equiv c \mod b, \quad c_1 \equiv 0 \mod c.$$

We immediately conclude that

$$a_1x + b_1y + c_1z \equiv 0x - uy + cz \mod a$$

$$a_1x + b_1y + c_1z \equiv -vx + 0y + cz \mod b$$

$$a_1x + b_1y + c_1z \equiv -wx + by + 0z \mod c.$$

In a similar way we find integers a_2, b_2 and c_2 such that

$$a_2x + b_2y + c_2z \equiv 0x + \frac{u}{c}y + z \mod a$$

$$a_2x + b_2y + c_2z \equiv \frac{v}{c}x + 0y + z \mod b$$
$$a_2x + b_2y + c_2z \equiv \frac{w}{b}x + y + 0z \mod c.$$

With these choices we obviously have the needed

$$(a_1x + b_1y + c_1z)(a_2x + b_2y + c_2z) \equiv ax^2 + by^2 + cz^2 \mod abc.$$

Denote $f(x, y, z) := a_1x + b_1y + c_1z$, and $g(x, y, z) := a_2x + b_2y + c_2z$.

6.5.2 Sufficiency of the Conditions on a, b and c

Consider the function $f(x, y, z) = a_1x + b_1y + c_1z$ with integer and x, y, and z in the parallelepiped

$$\{x^2 < |bc|\} \times \{y^2 < |ac|\} \times \{z^2 < |ab|\}.$$

It is straightforward that the cardinality of this domain is $(2s + 1)(2t + 1)(2r + 1)$ where

$$s = \lfloor\sqrt{|bc|}\rfloor \quad t = \lfloor\sqrt{|ac|}\rfloor \quad r = \lfloor\sqrt{|ab|}\rfloor.$$

Exercise 6.13 *Let $A \in \mathbb{N}$, and $d = \lfloor\sqrt{A}\rfloor$. Prove that $d^2 + 2d \geq A$. Conclude that $d \geq \sqrt{A+1} - 1$.*

From this exercise we derive that

$$\begin{aligned}(2s + 1)(2t + 1)(2r + 1) &\geq (2\sqrt{|bc| + 1} - 1)(2\sqrt{|ac| + 1} - 1)(2\sqrt{|ab| + 1} - 1)\\ &= \frac{(4|bc| + 3)(4|ac| + 3)(4|ab| + 3)}{(2\sqrt{|bc| + 1} + 1)(2\sqrt{|ac| + 1} + 1)(2\sqrt{|ab| + 1} + 1)}.\end{aligned}$$

We have furthermore that

$$2\sqrt{|bc| + 1} + 1 < 4\sqrt{|bc|}, \quad 2\sqrt{|ac| + 1} + 1 < 4\sqrt{|ac|}, \quad 2\sqrt{|ab| + 1} + 1 < 4\sqrt{|ab|},$$

and therefore that

$$(2s + 1)(2t + 1)(2r + 1) \geq \frac{(4|bc| + 3)(4|ac| + 3)(4|ab| + 3)}{4\sqrt{|bc|}4\sqrt{|ac|}4\sqrt{|ab|}} > |abc|.$$

We have proven the following

Proposition 6.5.3 *In the notations above, there are two distinct triplets*

$$(x_1, y_1, z_1), (x_2, y_2, z_2) \in \{x^2 < |bc|\} \times \{y^2 < |ac|\} \times \{z^2 < |ab|\}$$

such that $f(x_0, y_0, z_0) \equiv f(x_1, y_1, z_1) \mod abc$.

As a direct consequence we get that

Corollary 6.5.4 *There is a non-zero triplet $(x_0, y_0, z_0) \in \{x^2 < |bc|\} \times \{y^2 < |ac|\} \times \{z^2 < |ab|\}$ such that $f(x_0, y_0, z_0) \equiv 0 \mod abc$.*

We are ready now to finish the proof of the sufficiency of the conditions in Legendre's theorem. Indeed, we have that

$$ax_0^2 + by_0^2 + cz_0^2 \equiv f(x_0, y_0, z_0)g(x_0, y_0, z_0) \equiv 0 \mod abc.$$

We have obviously that $|ax_0^2| < |abc|, |by_0^2| < |abc|$, and $|cz_0^2| < |abc|$. Since a, b and c are not all of the same sign, we get (assuming that two of the coefficients are positive) that

$$-|abc| < ax_0^2 + by_0^2 + cz_0^2 < 2|abc|$$

which, in combination with $ax_0^2 + by_0^2 + cz_0^2 \equiv 0 \mod abc$, implies that there are two options

$$ax_0^2 + by_0^2 + cz_0^2 = 0 \text{ or } ax_0^2 + by_0^2 + cz_0^2 = |abc| = -abc.$$

In the first case, we are done. In the second, we check that

$$a(x_0z_0 + by_0)^2 + b(y_0z_0 - ax_0)^2 + c(z_0^2 + ab)^2 = 0$$

which finishes the proof.

6.6 Multiplicativity of φ

In this section, we are using the CRT to establish a very important property of the totient function: this property relates nicely $\varphi(mn)$ to $\varphi(m)$ and $\varphi(n)$. =

6.6.1 The Induced Map $G : (\mathbb{Z}/(mn)\mathbb{Z})^\times \to (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$

Let m, n be positive natural numbers. Consider the associated with the CRT map

$$F : \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

As we know, F is a ring homomorphism which sends the identity element $[1]_{mn}$ of its domain to the identity element $([1]_m, [1]_n)$ of its co-domain. This means that

$$[a]_{mn}[b]_{mn} = [1]_{mn} \quad \Rightarrow \quad F([a]_{mn}[b]_{mn}) = F([1]_{mn}) \quad \Rightarrow \quad F([a]_{mn})F([b]_{mn}) = ([1]_m, [1]_n).$$

In other words, **a unit of $\mathbb{Z}/mn\mathbb{Z}$ is sent by F to a unit of $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$**. Decorating the notation of the latter ring in a usual way to denote the group of units thereof, we get that the ring homomorphism F induces a map between groups (with operation multiplication)

$$G : (\mathbb{Z}/(mn)\mathbb{Z})^\times \to (\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z})^\times$$

defined by $G([a]_{mn}) = F([a]_{mn})$. It follows from the definition of G that it is a **group homomorphism**, that is for all $[a]_{mn}, [b]_{mn} \in (\mathbb{Z}/mn\mathbb{Z})^\times$ we have

$$G([a]_{mn}[b]_{mn}) = G([a]_{mn})G([b]_{mn}).$$

Recalling that

$$gcd(a, mn) = 1 \qquad \Leftrightarrow \qquad gcd(a, m) = 1 \wedge gcd(a, n) = 1,$$

we identify the units of $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ to the set $(\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$. This latter set is the Cartesian product of two groups, and is a group with the component-wise multiplication. Observe that this is the same operation which the Cartesian product inherits from the ring under the inclusion

$$(\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times \subseteq \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$

As a result, we get that F induces a group homomorphism

$$G : (\mathbb{Z}/(mn)\mathbb{Z})^\times \to (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times .$$

6.6.2 Multiplicativity of φ and a Formula for $\varphi(n)$

Recall that in general the map $F : \mathbb{Z}/(mn)\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ is neither injective, nor surjective. The same applies to G as well. When m and n are relatively prime, however, the situation is much better.

Theorem 6.6.1 *Let m and n be relatively prime positive integers, and consider the homomorphism*

$$G : (\mathbb{Z}/(mn)\mathbb{Z})^\times \to (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$$

as above. Then G is an isomorphism, that is, it is a homomorphism which is a bijection as well.

Proof Since $\gcd(m, n) = 1$, the map F is a bijection, that is - an injection and a surjection. Since G is induced by F, then G is an injection as well. We have to show also that G is a surjection. To this end, let $([a]_m, [b]_n) \in (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$ So, there are $[a']_m$ and $[b']_n$ such that

$$[a]_m[a']_m = [1]_m \qquad \text{and} \qquad [b]_n[b']_n = [1]_n.$$

that is $([a']_m, [b']_n) \in (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$ as well, and

$$([a]_m, [b]_n) \cdot ([a']_m, [b']_n) = ([1]_m, [1]_n).$$

Since F is a surjection, there are $x, y \in (\mathbb{Z}/(mn)\mathbb{Z})^\times$ such that

$$F(x) = ([a]_m, [b]_n) \qquad \text{and} \qquad F(y) = ([a']_m, [b']_n).$$

But then we have also that

$$F(xy) = F(x)F(y) = ([a]_m, [b]_n) \cdot ([a']_m, [b']_n) = ([1]_m, [1]_n) \qquad \text{and} \qquad F([1]_{mn}) = ([1]_m, [1]_n),$$

and so, $F(xy) = F([1]_{mn})$. Since F is an injection, it follows that $xy = [1]_{mn}$, and that $x, y \in (\mathbb{Z}/mn\mathbb{Z})^\times$ Finally, by the definition of G, we have $([a]_m, [b]_n) = F(x) = G(x)$. This proves that G is a surjection. □

This theorem has an interpretation through the totient function.

Corollary 6.6.2 *The totient function $\varphi : \mathbb{N} \to \mathbb{N}$ is multiplicative, that is, for any two relatively prime natural numbers m and n we have*

$$\varphi(mn) = \varphi(m) \cdot \varphi(n).$$

Proof For $\gcd(m, n) = 1$, the map $G : (\mathbb{Z}/(mn)\mathbb{Z})^\times \to (\mathbb{Z}/m\mathbb{Z})^\times \times (\mathbb{Z}/n\mathbb{Z})^\times$ is a bijection. Comparing the cardinalities of the two sets, we get the result. □
This corollary helps us compute the values of φ whenever we know the canonical factorization of the argument. Here is the result

Proposition 6.6.3 *Let $n > 1$ be an integer, and let $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ be its canonical decomposition as a product of powers of primes. We have the formula*

$$\varphi(n) = n \cdot \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_k}\right).$$

Proof The result follows from the above corollary, by a straightforward induction, and by the easily verifiable formula, for every prime number p, and for any positive integer α,

$$\varphi(p^\alpha) = p^\alpha \cdot \left(1 - \frac{1}{p}\right).$$

The last formula can be proved as follows. By definition, $\varphi(p^\alpha)$ is the number of elements in a reduced residue system modulo p^α. To find this number, one can find the number of residues which are **not** relatively prime with p^α instead, and subtract it from the total number, p^α, of residues modulo p^α. Now, a residue is not relatively prime with p^α if, and only if, it is divisible by p. Obviously, among the natural numbers less than p^α the ones which are divisible by p have the form pm, and are such that $0 \leq pm < p^\alpha$. In other words, $0 \leq m < p^{\alpha-1}$. The total of such numbers is $p^{\alpha-1}$. therefore, $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$. The formula is proven. □

As a matter of fact, this proposition can be seen from more conceptual point of view. Indeed, consider, as we did in the end of the last section, $n = n_1 n_2 \cdots n_k$ where the $n_1, n_2, \ldots, n_k$ are pairwise relatively prime. As we already know, the corresponding map

$$F : \mathbb{Z}/n\mathbb{Z} \ \to \ \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$$

is a bijection. It is again straightforward to see that F restricted to the subset $(\mathbb{Z}/n\mathbb{Z})^\times \subseteq \mathbb{Z}/n\mathbb{Z}$ has values in $(\mathbb{Z}/n_1\mathbb{Z})^\times \times (\mathbb{Z}/n_2\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/n_k\mathbb{Z})^\times$. If we denote, as above, by

$$G : (\mathbb{Z}/n\mathbb{Z})^\times \ \to \ (\mathbb{Z}/n_1\mathbb{Z})^\times \times (\mathbb{Z}/n_2\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/n_k\mathbb{Z})^\times$$

the map obtained by restricting the domain and the co-domain of F to the ones in the previous sentence, we immediately get that it is a bijection. Comparing the cardinalities of the two sets gives, modulo the formula for $\varphi(p^\alpha)$, the result about φ in the Proposition.

Example 6.6.1 We are showing that there is no natural number n such that $\varphi(n) = 14$.
Let $n = 2^\alpha p_1^{\beta_1} \cdots p_k^{\beta_k}$ be the canonical decomposition of n. Then we have

$$\varphi(n) = 2^{\alpha-1} p_1^{\beta_1 - 1}(p_1 - 1) \cdots p_k^{\beta_k - 1}(p_k - 1) = 2 \cdot 7.$$

We see that $\alpha \leq 2$, and $\alpha = 2$ is impossible: there has to be an odd prime divisor of n which will increase the exponent of 2 dividing $\varphi(n)$. So, n is either odd $n = 2m + 1$, or twice an odd number, $n = 2(2m + 1)$. In both cases $\varphi(n) = \varphi(2m + 1)$. Observe that $\varphi(n) = 14$ implies also that $k \leq 1$ (why?). So, $2m + 1 = p^\beta$. Finally, $\varphi(n) = 14$ implies that $\beta \leq 2$ (how?). So, $2m + 1 = 1, p$, or p^2. It is easy to check that none of these cases works. □

Exercise 6.14 *(1) Prove that,* $n \geq 3 \quad \to \quad \varphi(n) \in 2\mathbb{N}$.
(2) Prove that $\varphi(n) = \varphi(2n)$ *if, and only if,* n *is an odd number.*
(3) Find all n *such that* $4 \nmid \varphi(n)$.
(4) Find all $n \in \mathbb{N}$ *for which*

$$(i)\ \varphi(n) = n/2 \qquad (ii)\ \varphi(n) = n/3 \qquad (iii)\ \varphi(n) = n/6.$$

(5) What can you say about n *if* $\varphi(n)$ *is a prime number? The same question for when* $\varphi(n)$ *is a square of a prime number.*
(6) Find the smallest positive integers a *which is not in the range of the totient function. That is, find the smallest* $a \in \mathbb{Z}_+$ *such* $\varphi(n) \neq a$ *for every* $n \in \mathbb{Z}_+$.
(7) There are ten natural numbers n *for which* $\varphi(n) = 24$. *Find them.*
(8) Find all pairs of natural numbers m, n *for which* $\varphi(mn) = \varphi(m) + \varphi(n)$.
(9) For p *a prime number, and* k *a positive integer, prove that*

$$\varphi(1) + \varphi(p) + \cdots \varphi(p^k) = p^k.$$

Prove further that for any positive integer n

$$\sum_{d>0,\, d|n} \varphi(d) = n.$$

(10) Let $S_n = \{m \mid 1 \leq m \leq n,\ \gcd(m, n) = 1\}$. *Prove that*

$$\sum_{m \in S_n} m = (n\varphi(n))/2.$$

(11) Prove that the function $\psi : \mathbb{N} \to \mathbb{N}$ *defined by* $\psi(n) = n - \varphi(n)$ *has the property*

$$d \mid n \ \wedge\ d < n \qquad \to \qquad \psi(d) < \psi(n).$$

(12) Prove that for any two positive integers m, n *with* $d = \gcd(m, n)$ *we have*

$$\varphi(mn) = \frac{d \cdot \varphi(m)\varphi(n)}{\varphi(d)}.$$

(13) Prove that, for positive integers m *and* n,

$$\varphi(mn) = \varphi(n) \qquad \Leftrightarrow \qquad (m = 1) \vee (m = 2 \wedge 2 \nmid n).$$

(14) Prove that, for the positive integers m, n,

$$\varphi(mn) = \varphi(m)\varphi(n) \qquad \Leftrightarrow \qquad \gcd(m, n) = 1.$$

(15) Let $m = m_1 m_2 \cdots m_r$ *where* $m_1, m_2, \ldots, m_r$ *are pairwise relatively prime natural numbers. Show that*

$$m_1^{\varphi(m)/\varphi(m_1)} + m_2^{\varphi(m)/\varphi(m_2)} + \cdots + m_r^{\varphi(m)/\varphi(m_r)} \equiv r - 1 \mod m.$$

Chapter 7

Polynomial Equations Modulo n

7.1 Linear Equations in $\mathbb{Z}/n\mathbb{Z}$

Following the general philosophy from Algebra, once we have a ring, we would like to be able to solve linear equations with coefficients in that ring. The ring at hand is $\mathbb{Z}/n\mathbb{Z}$, and we want to study the solutions to an equation of the form

$$[a] \cdot X = [b]$$

where $[a] \neq [0]$. In the notations of congruences, we want to solve the **congruence equation**

$$a \cdot X \equiv b \mod n.$$

Theorem 7.1.1 *Consider the equation $[a] \cdot X = [b]$ in $\mathbb{Z}/n\mathbb{Z}$. The following is true*
(1) If $[a] = [0]$, then the equations has solutions if, and only if, $[b] = [0]$ in which case it has n distinct solutions (every element of $\mathbb{Z}/n\mathbb{Z}$ is a solution to $[0] \cdot X = [0]$);
(2) If $[a] \neq [0]$, let $d = gcd(a, n)$. The equation above has a solution if, and only if, $d \,|\, b$ in which case it has exactly d distinct solutions.

Before giving a proof to the theorem - an exercise! The exercise shows that item (2) of the theorem does make sense!

Exercise 7.1 *Let $[a] \neq [0]$ in $\mathbb{Z}/n\mathbb{Z}$. Prove that*
(1) if $[a] = [b]$, then $gcd(a, n) = gcd(b, n)$;
(2) if $d \,|\, n$ and $[a] = [b]$, then $d \,|\, a \Leftrightarrow d \,|\, b$.

Again, in the congruence notations item (2) of the theorem says that if $\neg(a \equiv 0 \mod n)$ and $d = gcd(a, n)$, then the congruence equation

$$a \cdot X \equiv b \mod n$$

has solutions if, and only if, $d \,|\, b$ in which case there are exactly d **incongruent modulo** n solutions. Note that the solutions in integers are infinitely many in that case! All they are organized in d distinct classes modulo n.

Proof of Theorem 7.1.1. The item (1) is obvious. So, let's prove item (2). The crucial observation here is that $[a] \cdot X = [b]$ has a solution $[x_0]$ if, and only if, $n \,|\, a \cdot x_0 - b$, if, and only if, $a \cdot x_0 - b = n \cdot y_0$, if, and only if, the equation

$$a \cdot X + n \cdot Y = b$$

has a solution in $\mathbb{Z}$. But we already know that the linear Diophantine equation has a solution if, and only if, $gcd(a,n) \mid b$ in which case all solutions are given by the formulae

$$x' = x_0 + (n/d)\cdot k \qquad y' = y_0 - (a/d)\cdot k$$

where $k \in \mathbb{Z}$, $d = gcd(a,n)$, and (x_0, y_0) is a particular solution of that equation. This proves half of the claim of item (2): the existence of solutions. For the number of distinct solutions in $\mathbb{Z}/n\mathbb{Z}$, we need to figure out when two solutions

$$x' = x_0 + (n/d)\cdot k' \qquad \text{and} \qquad x'' = x_0 + (n/d)\cdot k''$$

determine the same element in $\mathbb{Z}/n\mathbb{Z}$, that is, when $[x'] = [x'']$. But since

$$x' - x'' = (n/d)(k' - k'') = n\cdot(k' - k'')/d,$$

we see that $n \mid x' - x''$, that is $[x'] = [x'']$ in $\mathbb{Z}/n\mathbb{Z}$, if, and only if, $(k' - k'')/d$ is an integer, if, and only if, $d \mid k' - k''$, if, and only if, $[k'] = [k'']$ in $\mathbb{Z}/d\mathbb{Z}$. Therefore the distinct solutions to $[a]\cdot X = [b]$ in $\mathbb{Z}/n\mathbb{Z}$ are as many as are the elements of $\mathbb{Z}/d\mathbb{Z}$. This finishes the proof of item (2). □

As usual, the case when $n = p$ is a prime number is very simple: if $[a] \neq [0]$ in $\mathbb{Z}/p\mathbb{Z}$, then $gcd(a,p) = 1$, so the solution to $[a]\cdot X = [b]$ is unique. Of course, if $[a']$ is the reciprocal of $[a]$, which exists since $[a] \neq [0]$, then, that (unique) solution is given by

$$[x_0] = [a']\cdot[a].$$

Exercise 7.2 *(1) Solve the congruences*

(i) $4X \equiv 2 \mod 6$ (iii) $4X \equiv 4 \mod 6$ (v) $256X \equiv 179 \mod 337$

(ii) $4X \equiv 1 \mod 6$ (iv) $3X \equiv 800 \mod 11$ (vi) $1215X \equiv 560 \mod 2755$.

(2) Let p be a prime number, and a, b be integers such that $1 \leq a \leq p-1$. Prove that the congruence $aX \equiv b \mod p$ has a solution

$$x \equiv b\cdot(-1)^{a-1}\cdot\frac{1}{a}\cdot\binom{p-1}{a-1} \mod p.$$

7.2 Equations of Higher Degree in $\mathbb{Z}/n\mathbb{Z}$

7.2.1 General Remarks and Notations

Let $f(X) = a_dX^d + \cdots + a_1X + a_0$ be a polynomial with integer coefficients, and let $n \in \mathbb{N}$ be a positive number. We are interested in finding the solutions to the equation

$$f(X) \equiv 0 \mod n.$$

By definition this means that we want to find all integers $s \in \mathbb{Z}$ such that

$$f(s) \equiv 0 \mod n.$$

Let's observe the following.

(i) If $s, s' \in \mathbb{Z}$ such that $s \equiv s' \mod n$, then $f(s) \equiv f(s') \mod n$.

(ii) Suppose $g(X) = a'_dX^d + \cdots a'_1X + a'_0$ is a polynomial with integer coefficients for which

$$a_0 \equiv a'_0 \mod n, \quad a_1 \equiv a'_1 \mod n, \quad \ldots \quad a_d \equiv a'_d \mod n.$$

(**We write in such a case** $f(X) \equiv g(X) \mod n$.) If $s \in \mathbb{Z}$, then $f(s) \equiv g(s) \mod n$. □

Our observations mean that solving $f(X) \equiv 0 \mod n$ is the same as solving the equation

$$[f(X)]_n := [a_d]_nX^d + \cdots + [a_1]_nX + [a_0]_n = [0]_n \quad \text{in} \quad \mathbb{Z}/n\mathbb{Z}.$$

The polynomial $[f(X)]_n$ with coefficients in $\mathbb{Z}/n\mathbb{Z}$ is called **the reduction modulo** n of the polynomial $f(X)$.

7.2.2 Reduction of the Modulus Using the Chinese Remainder Theorem

Suppose now that $n = n_1 n_2$ where $\gcd(n_1, n_2) = 1$. If $s \in \mathbb{Z}$ is a solution to $f(X) \equiv 0 \mod n$, that is if $f(s) \equiv 0 \mod n$, then obviously s is a solution to both equations

$$f(X) \equiv 0 \mod n_1 \qquad \text{and} \qquad f(X) \equiv 0 \mod n_2.$$

This means that the solution $[s]_n$ to the equation $[f(X)]_n = [0]_n$ produces solutions $[s]_{n_1}$, respectively $[s]_{n_2}$, to $[f(X)]_{n_1} = [0]_{n_1}$, respectively $[f(X)]_{n_2} = [0]_{n_2}$.

By using the CRT (recall that $\gcd(n_1, n_2) = 1$), we see that if $[s_1]_{n_1}$ is a solution to $[f(X)]_{n_1} = [0]_{n_1}$ and if $[s_2]_{n_2}$ is a solution to $[f(X)]_{n_2} = [0]_{n_2}$, then there is a unique $[s]_n$ such that $[s]_{n_1} = [s_1]_{n_1}$ and $[s]_{n_2} = [s_2]_{n_2}$, and moreover, since $f(s) \equiv f(s_1) \mod n_1$ and $f(s) \equiv f(s_2) \mod n_2$, we have also that $f(s) \equiv 0 \mod n$.

We can interpret the result we just established in a useful way using the CRT map $F : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$. Denote by $M \subseteq \mathbb{Z}/n\mathbb{Z}$ the set of solutions to $[f(X)]_n = [0]_n$, and by $M_i \subseteq \mathbb{Z}/n_i\mathbb{Z}$ the set of solutions to $[f(X)]_{n_i} = [0]_{n_i}$ for $i = 1, 2$. Then, by the above considerations we have that

$$F_{|M} : M \to \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$$

has a range exactly $M_1 \times M_2$. As a consequence we get that F induces a bijection

$$\tilde{F} : M \to M_1 \times M_2.$$

Our discussion so far helps us prove the following theorem.

Theorem 7.2.1 *Suppose $f(X)$ is a polynomial with integer coefficients, and $n \in \mathbb{N}$ is a positive number such that $n = n_1 n_2 \cdots n_k$ where $n_1, n_2, \ldots, n_k$ are pairwise relatively prime. Denote by $M_i \subseteq \mathbb{Z}/n_i\mathbb{Z}$ the set of solutions to $[f(X)]_{n_i} = [0]_{n_i}$ for $i = 1, \ldots, k$, and by $M \subseteq \mathbb{Z}/n\mathbb{Z}$ the set of solutions to $[f(X)]_n = [0]_n$. Then, the CRT map $F : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$ induces a bijection between M and $M_1 \times \cdots \times M_k$.*

Proof We argue by induction on $k \geq 2$. The discussion preceding the theorem proves the base case $k = 2$. Let's verify the induction step: $(n \to n + 1)$. Suppose $n = n_1 \cdots n_k n_{k+1}$ is a product of pairwise relatively prime positive integers. We want to show that the CRT map

$$F : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z} \times \mathbb{Z}/n_{k+1}\mathbb{Z}$$

induces a bijection

$$\tilde{F} : M \to M_1 \times \cdots \times M_k \times M_{k+1}.$$

Let $n' = n/n_{k+1}$. We have that $n' = n_1 \cdots n_k$ is a product of k pairwise relatively prime positive integers, and therefore we can apply the induction hypothesis to the solutions of $[f(X)]_{n'} = [0]_{n'}$. Denoting by $M' \subseteq \mathbb{Z}/n'\mathbb{Z}$ the set solutions and by $F' : \mathbb{Z}/n'\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$ the corresponding CRT map, we get that F' induces a bijection $\tilde{F'} : M' \to M_1 \times \cdots \times M_k$.
Applying the base step to $n = n'n_{k+1}$ we get the bijection $\tilde{F''} : M \to M' \times M_{k+1}$ where $F'' : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n'\mathbb{Z} \times \mathbb{Z}/n_{k+1}\mathbb{Z}$ is the CRT map relevant here.
Observe now that F is the composition of the following maps

$$F'' : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n'\mathbb{Z} \times \mathbb{Z}/n_{k+1}\mathbb{Z}$$

and

$$(F', Id_{\mathbb{Z}/n_{k+1}\mathbb{Z}}) : \ \mathbb{Z}/n'\mathbb{Z} \times \mathbb{Z}/n_{k+1}\mathbb{Z} \ \to \ \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z} \times \mathbb{Z}/n_{k+1}\mathbb{Z},$$

that is, $F = (F', Id_{\mathbb{Z}/n_{k+1}\mathbb{Z}}) \circ F''$. Restricting both maps, on the left and on the right of the equality, to M we get the composition $\tilde{F} = (\tilde{F'}, Id_{M_{k+1}}) \circ \tilde{F''} : M \to M_1 \times \cdots \times M_k \times M_{k+1}$. Since both, $(\tilde{F'}, Id_{M_{k+1}})$ and $\tilde{F''}$ are bijections, then so is their composition. Therefore, $\tilde{F}$ is a bijection as

claimed. This completes the proof of the induction step. □

This theorem gives in particular an estimate of the number of solutions to $[f(X)]_n = [0]_n$: in the notation of the theorem, it is $|M| = |M_1| \cdots |M_k|$.

Corollary 7.2.2 *Let $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ be the canonical decomposition of $n \geq 2$, and let $f(X)$ be a polynomial with integer coefficients. Denote by M the number of solutions to $[f(X)]_n = [0]_n$, and by M_i - the number of solutions to $[f(X)]_{p_i^{\alpha_i}} = [0]_{p_i^{\alpha_i}}$ for $i = 1, \ldots, k$. Then, $M = M_1 \cdots M_k$.*

Example 7.2.1 Let $n = 2^\alpha p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, where $2 < p_1 < \cdots < p_k$, be the canonical decomposition of n as a product of powers of primes. Here we allow $\alpha \geq 0$, but $\alpha_i \geq 1$ for $i = 1, \ldots, k$. For the number m of solutions to $[X^2 - 1]_n = [0]_n$ we have
(i) $m = 2^k$ if $\alpha \leq 1$;
(ii) $m = 2^{k+1}$ if $\alpha = 2$;
(iii) $m = 2^{k+2}$ if $\alpha \geq 3$.
This follows from Exercise 6.3 items (5) and (6). □

Example 7.2.2 We are solving $f(X) = X^3 + 4X + 7 \equiv 0 \mod 84$. For this, we need to solve

$$f(X) \equiv 0 \mod 4, \qquad f(X) \equiv 0 \mod 3, \qquad f(X) \equiv 0 \mod 7$$

first. We check directly that the first equation has solutions $X \equiv 1 \mod 4$, the second $X \equiv 1 \mod 3$, and the third has $X \equiv 0 \mod 7$. So, the solution to $f(X) \equiv 0 \mod 84$ is unique, and is given by the system of three linear equations above. Since $\gcd(3, 4) = 1$ the system $X - 1 \equiv 0 \mod 4$ and $X - 1 \equiv 0 \mod 3$, has a solution $X \equiv 1 \mod 12$. Together with the third linear equation we have

$$X = 7X_1, \qquad 7X_1 \equiv 1 \mod 12$$

which is the same as

$$7X_1 \equiv 49 \mod 12.$$

Since $\gcd(7, 12) = 1$ we cancel out a 7, and get $X_1 \equiv 7 \mod 12$. So, $X_1 = 12k + 7$ and $X = 7X_1 = 7(7 + 12k) = 49 + 84k$. The solution to $f(X) \equiv 0 \mod 84$ is given by $X \equiv 49 \mod 84$. □

7.3 Equations of Higher Degree in $\mathbb{Z}/p^k\mathbb{Z}$

The last theorem of the previous section reduces solving the equation $[f(X)]_n = [0]_n$ to solving the system of equations $[f(X)]_{p_i^{\alpha_i}} = [0]_{p_i^{\alpha_i}}$ where $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ is the canonical decomposition of n. So, we may assume, W.L.O.G., that $n = p^k$. Turns out the situation in this case is even better: solving $f(X) \equiv 0 \mod p^k$ almost entirely depends on solving $f(X) \equiv 0 \mod p$ - a fact discovered in its utmost generality by Kurt Hensel (in the early 1900's). His result is what we are proving in this section.

7.3.1 The Derivative of a Polynomial

The concept of derivative of a function is known from Calculus. This definition still works when we consider a polynomial over the fields $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$. What we need here is a definition of derivative of a polynomial over a ring, and even a finite ring. No matter that our approach is not related to finding limits, as it is in Calculus, the resulting formula for the derivative will be the same.

Consider the polynomial $f(X) = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_1 X + a_0$ where the coefficients belong to some ring (such as $\mathbb{Z}, \mathbb{Q}, \mathbb{R}\,\mathbb{C}, \mathbb{Z}/n\mathbb{Z}$ or any other ring). We can express this polynomial in terms of $X - a$ for any a in the ring.

$$f(X) = a_d(X - a + a)^d + a_{d-1}(X - a + a)^{d-1} + \cdots + a_1(X - a + a) + a_0$$

-

$$= b_d(X-a)^d + b_{d-1}(X-a)^{d-1} + \cdots + b_1(X-a) + b_0 =: g(X-a).$$

Obviously, $b_0 = f(a)$. Let's compute b_1.

Exercise 7.3 *Prove that* $b_1 = d \cdot a_d \cdot a^{d-1} + (d-1) \cdot a_{d-1} \cdot a^{d-2} + \cdots + 2 \cdot a_2 \cdot a + a_1$.

By the expression above, we see that b_1 is the coefficient at the linear term of $g(X-a)$, and $b_1(X-a) + b_0 = b_1(X-a) + f(a)$ is therefore similar to the "linear approximation" of $f(X)$ at a

$$f(X) \sim f(a) + b_1(X-a).$$

This is the motivation for the following definition.

Definition 7.3.1 *In the notations above, the polynomial*

$$d \cdot a_d X^{d-1} + (d-1) \cdot a_{d-1} X^{d-2} + \cdots + 2 \cdot a_2 X + a_1$$

is called the derivative of $f(X)$, and is denoted by $f'(X)$. We define the kth derivative, $f^{(k)}(X)$, of $f(X)$ for every natural number k: $f^{(0)}(X) = f(X), f^{(1)}(X) = f'(X)$, and for every $k \in \mathbb{N}$ we set $f^{(k+1)}(X) = (f^{(k)}(X))'$.

In these notations we have that $b_1 = f'(a)$, and we obtain the identity

$$f(X) = f(a) + f'(a)(X-a) + b_2(X-a)^2 + \cdots + b_d(X-a)^d.$$

Obviously, if $k > d$, then $f^{(k)}(X) = 0$.

7.3.2 Solving Equations Modulo p^k: Hensel's Lemma

Suppose that $s \in \mathbb{Z}$ is a solution to $f(X) \equiv 0 \mod p^k$ for $k \geq 2$. That is

$$f(s) \equiv 0 \mod p^k.$$

Then, for every $1 \leq l < k$ we have that $f(s) \equiv 0 \mod p^l$, that is, s is a solution to $f(X) \equiv 0 \mod p^l$ for every such l. Turns out that often we can reverse the process: starting with a solution to $f(X) \equiv 0 \mod p^l$, we can determine a solution to $f(X) \equiv 0 \mod p^{l+1}$. This process is called **lifting the solution from mod p^l to mod p^{l+1}**. Repeating this process, one ends up with a solution mod p^k. Here is the official definition of lifting solutions.

Definition 7.3.2 *Suppose $m, n \in \mathbb{N}$ such that $m \mid n$, and that x_0 is a solution to $f(X) \equiv 0 \mod m$. The integer x_1 is called a* **lift** $\mod n$ **of** x_0 *if*
(i) $x_1 \equiv x_0 \mod m$, and
(ii) $f(x_1) \equiv 0 \mod n$

In particular, a lift $\mod p^{l+1}$ of a solution x_0 to $f(X) \equiv 0 \mod p^l$ is an integer x_1 such that

$$x_1 \equiv x_0 \mod p^l \qquad \text{and} \qquad f(x_1) \equiv 0 \mod p^{l+1}.$$

Notice that (see the example below) there are cases in which several incongruent solutions to mod p^k restrict to the same solution to $\mod p^l, (l < k)$.

Example 7.3.1 Let $p = 2$ and $k = 3$. The equation $X^2 - 1 \equiv 0 \mod 8$ has four solutions: $[1]_8, [3]_8, [5]_8$ and $[7]_8$. When restricting to $\mod 4$, we see that $[1]_8$ and $[5]_8$ restrict to $[1]_4$, and $[3]_8$ and $[7]_8$ restrict to $[3]_4$. □

Since every solution to $f(X) \equiv 0 \mod p^{l+1}$ restricts to a solution $\mod p^l$, the lifting of solutions from $\mod p^l$ to $\mod p^{l+1}$ will produce **all** solutions $\mod p^{l+1}$.

The derivative of the polynomial $f(X)$ plays an important role in the lifting process. Here is why. Suppose S is a solution to $f(X) \equiv 0 \mod p^l$ which we want to lift to solutions $\mod p^{l+1}$. Any lifted solution will have the form $s' = s + t \cdot p^l$ for some $t \in \mathbb{Z}$, and will satisfy the congruence $f(s') \equiv 0 \mod p^{l+1}$. Since we are working mod p^{l+1}, the possible lifts of s to mod p^{l+1} are given by $0 \leq t \leq p-1$. More explicitly

$$f(s') = a_d(s+tp^l)^d + a_{d-1}(s+tp^l)^{d-1} \cdots + a_1(s+tp^l) + a_0 \equiv 0 \mod p^{l+1}.$$

Substituting $s' = s + t \cdot p^l$ for X, and s for a in the last formula of the previous subsection we can express the previous equation as

$$f(s') = f(s) + f'(s) \cdot (tp^l) + b_2 \cdot (tp^l)^2 + \cdots + b_d \cdot (tp^l)^d \equiv 0 \mod p^{l+1}.$$

Disregarding the terms with p^r for $r \geq l+1$, since they are $0 \mod p^{l+1}$, we simplify the last congruence to (this calculation is straightforward and routine, and should be done by every student on their own)

$$f(s) + tf'(s)p^l \equiv 0 \mod p^{l+1}.$$

Since $p^l \mid f(s)$, we see that the values of t we are looking for satisfy the congruence

$$\frac{f(s)}{p^l} + f'(s)t \equiv 0 \mod p.$$

So, t, being an integer bounded by 0 and $p-1$, is a solution to the linear equation

$$f'(s) \cdot X \equiv -\frac{f(s)}{p^l} \mod p.$$

By the first section of this chapter we know how to solve this equation. Namely, a solution exists if, and only if, $\gcd(f'(s), p) \mid f(s)/p^l$ in which case the number of the incongruent solutions $\mod p$ is $\gcd(f'(s), p)$.
There are two cases to consider here:

$$(i) \quad \gcd(f'(s), p) = 1 \quad \text{when} \quad p \nmid f'(s) \qquad \text{and} \qquad (ii) \quad \gcd(f'(s), p) = p \quad \text{when} \quad p \mid f'(s).$$

In the case when $\gcd(f'(s), p) = 1$ s is liftable to mod p^{l+1} in a unique way: $s' = s + t_0 p^l$ where t_0 is the unique solution to the equation mod p above.
In the case when $\gcd(f'(s), p) = p$ the equation reduces to

$$0 \cdot X \equiv -\frac{f(s)}{p^l} \mod p.$$

Accordingly, we have two sub-cases here: (ii') $p \nmid f(s)/p^l$, and (ii") $p \mid f(s)/p^l$. In the first sub-case no solutions exist, and the lifting of s from $\mod p^l$ to $\mod p^{l+1}$ is not possible. In the second sub-case, all classes $\mod p$ can be t, that is, any $t = 0, \ldots, p-1$, and s has p lifts to mod p^{l+1}.

We have proven now item (1) of the following theorem.

Theorem 7.3.1 (*Hensel's Lifting Lemma*) *Let $f(X)$ be a polynomial with integers coefficients, let p be a prime number, and l a positive integer.*
(1) Suppose that the integer s is a solution to $f(X) \equiv 0 \mod p^l$. Then we have
(i) if $p \nmid f'(s)$ then there is a unique $\mod p^{l+1}$ number $s' \in \mathbb{Z}$ such that $s' \equiv s \mod p^l$ and $f(s') \equiv 0 \mod p^{l+1}$;
(ii) if $p \mid f'(s)$ and $p^{l+1} \nmid f(s)$, then there is no s' with properties as in (i);
(iii) if $p \mid f'(s)$ and $p^{l+1} \mid f(s)$, then there are p incongruent $\mod p^{l+1}$ integers $s'_1, \ldots, s'_p$ such that $s'_i \equiv s \mod p^l$ and $f(s'_i) \equiv 0 \mod p^{l+1}$.
(2) If $f(s) \equiv 0 \mod p$ and $p \nmid f'(s)$, then for every $k \geq 1$ there is a unique $\mod p^k$ integer s_k such that $s_k \equiv s \mod p$, and $f(s_k) \equiv 0 \mod p^k$.

Proof The discussion before the theorem provides the proof of item (1). To prove item (2) we argue by induction on $k \geq 2$. The base step $k = 2$ is the case $l = 1$ in item (1), and is therefore verified. We are proving the inductive step: $(k \to k+1)$. Let the extension s_k of s to $\mod p^k$ be unique. We have to show that there is an extension of s to $\mod p^{k+1}$ and that this extension is unique.
(Existence.) Notice that, since $s_k \equiv s \mod p$, we have $f'(s_k) \equiv f'(s) \mod p$, and so $p \nmid f'(s_k)$. By item (1), we get that there is a unique $\mod p^{k+1}$ extension s'_k of s_k. This means in particular that $f(s'_k) \equiv 0 \mod p^{k+1}$ and that $s'_k \equiv s_k \mod p^k$. But then, $s'_k \equiv s_k \equiv s \mod p$ as well, and we can choose $s_{k+1} = s'_k$: a $\mod p^{k+1}$ extension of s does exist.
(Uniqueness.) For the uniqueness of the $\mod p^{k+1}$ extension of s, notice that any such extension is also a $\mod p^{k+1}$ extension of s_k as well, and so, as we proved, is unique. This completes the proof of item (2). □

Remark 7.3.1 To solve $f(X) \equiv 0 \mod p^k$ we must start with solving $f(x) \equiv 0 \mod p$, and then extend, whenever possible, the solutions to $\mod p^k$. Hensel's lemma tells us how to do that step by step. It, moreover, teaches us that every solution $\mod p$ which does not annihilate $f'(X) \mod p$ has a unique extension to $\mod p^k$. In particular, if the latter happens for all solutions $\mod p$, then $f(X) \equiv 0 \mod p^k$ has as many solutions as $f(X) \equiv 0 \mod p$. □

Example 7.3.2 Consider the equation $f(X) = X^2 - 1 \equiv 0 \mod 16$. The solution to $f(X) \equiv 0 \mod 2$ is only one: $s = 1$. Since $f'(X) = 2X$ we have that $f'(1) \equiv 0 \mod 2$. Since $4 \mid f(1)$, the theory says that there will be two lifts of $s = 1$ to solutions $\mod 4$. Indeed $f(X) \equiv 0 \mod 4$ has two solutions: $s_1^{(1)} = 1$ and $s_1^{(2)} = 3$. As before, $f'(1) \equiv 0 \mod 2$ and $f'(3) \equiv 0 \mod 2$. Since in this case $f(s_1^{(j)}) \equiv 0 \mod 8$, for $j = 1, 2$, the solutions $s_1^{(j)}$ lift to two solutions $\mod 8$ each. Indeed, these are

$$s_2^{(1)} = 1, s_2^{(2)} = 5, s_2^{(3)} = 3, s_2^{(4)} = 7$$

the first two extending $s_1^{(1)}$ and the last two extending $s_1^{(2)}$. To find the solutions $\mod 16$, we have to extend, whenever possible, the solutions $\mod 4$. In this case

$$f(1) = 0, f(3) = 8, f(5) = 24, f(7) = 48$$

so that only the first and the last are divisible by 16. Therefore, only $s_2^{(1)}$ and $s_2^{(4)}$ can be lifted to solutions $\mod 16$, and each of them has two lifts. In particular, the equation $x^2 - 1 \equiv 0 \mod 16$ has four solutions. □

Example 7.3.3 How many solutions does $f(X) = X^3 + 2X + 1 \equiv 0 \mod 1800$ have? Solve the equation.
There is no problem in principle to find all solutions by brute force: trying all possible 1800 values for X, and select the ones which solve the equations. The more so if we have a computer handy. We will work here using theory, not computers.
Since $1800 = 2^3 3^2 5^2$ we have to solve three equations

$$f(X) \equiv 0 \mod 2^3, \qquad f(X) \equiv 0 \mod 3^2, \qquad f(X) \equiv 0 \mod 5^2.$$

(i) Solving $f(X) \equiv 0 \mod 2^3$. Again, we can find the solutions directly, checking out the eight possible values for X. Since this is obvious how to do, we will show how to find the solutions using the theory developed so far in this chapter. To this end, we solve $f(X) \equiv 0 \mod 2$ first. There is one such solution: $s_1 =$. We use Hensel's lemma to lift this solution to $\mod 2^2$. We have that $f'(X) = 3X^2 + 2$. Since $f'(s_1) = f'(1) = 5 \equiv 1 \mod 2 \neq 0 \mod 2$, this solution lifts to $\mod 4$. It lifts to $\mod 8$ as well, as we know, and $f(X) \equiv 0 \mod 8$ has a unique solution. Let's find it. We have $s'_1 = s_1 + 2t = 1 + 2t$ so that $f(s_1) + 2tf'(s_1) \equiv 0 \mod 4$. This means $4 + 10t \equiv 0 \mod 4$, so $t = 0$ (recall that $t \in \{0, 1\}$), and $s'_1 = 1$. To extend s'_1 to $\mod 8$ we do the same procedure: $s''_1 = s'_1 + 4t = 1 + 4t$ so that $f(s'_1) + 4tf'(s'_1) \equiv 0 \mod 8$. We get $4 + 20t \equiv 0 \mod 8$, so $t = 1$ and the only solution to $f(X) \equiv 0 \mod 8$ is given by $s''_1 = 5$.

(ii) Solving $f(X) \equiv 0 \mod 9$. The equation has no solutions $\mod 3$. This means that the original equation has no solutions either. □

Exercise 7.4 *Solve the equations*
(1) $7X^2 + 19X + 25 \equiv 0 \mod 27$.
(2) $9X^2 + 29X + 62 \equiv 0 \mod 64$.
(3) $X^3 + 2X + 2 \equiv 0 \mod 125$.
(4) $X^4 + 4X^3 + 2X^2 + 2X + 12 \equiv 0 \mod 625$.
(5) $6X^3 + 27X^2 + 17X + 20 \equiv 0 \mod 30$.
(6) $31X^4 + 57X^3 + 96X + 191 \equiv 0 \mod 225$.

Exercise 7.5 *(1) Suppose* $d \in \mathbb{N}$ *is such that* $\gcd(d, p) = 1$, *and let* $k > 0$ *and* a *be integers with* $\gcd(a.p) = 1$. *Prove that the equation* $X^d - a \equiv 0 \mod p^k$ *has as many solutions as the equation* $X^d - a \equiv 0 \mod p$.
(2) Let a *and* $k > 0$ *be integers such that* $\gcd(a, p) = 1$. *Prove that* $X^{p-1} - a \equiv 0 \mod p^k$ *has* $p - 1$ *solutions when* $a \equiv 1 \mod p$, *and* 0 *solutions otherwise.*

7.4 Vista: p-adic Numbers

Hensel's lemma discussed above is related to number systems $\mathbb{Q}_p$, extensions of the rational numbers, one for each prime number p, and called p-adic numbers. These number systems were invented by Kurt Hensel ($\sim$ 1897) with the aim of making methods of power series expansion, so powerful in analytic function theory, available to Number Theory as well.

This section is devoted to the definition and the discussion of the most basic properties of the p-adic numbers. In the second Vista about the p-adic numbers, in the next Chapter, we will discuss some number theoretical questions which show the importance of these extensions of the rational numbers.

7.4.1 The p-adic Integers $\mathbb{Z}_p$

We begin with the definition of p-adic integers. By definition, a p-adic integer is a sequence

$$< [a_n]_{p^n} >_{n\geq 1} := \{[a_1]_p, [a_2]_{p^2}, \dots, [a_n]_{p^n}, \dots\}$$

such that

$$a_{n+1} \equiv a_n \mod p^n$$

for every $n \geq 1$. The class $[a_n]_{p^n}$ is called the nth component of the p-adic integer $x =< [a_n]_{p^n} >$. Naturally, the p-adic integers form a subset $\mathbb{Z}_p$ of the Cartesian product

$$\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \cdots \times \mathbb{Z}/p^n\mathbb{Z} \times \cdots =: \prod_{n\geq 1} \mathbb{Z}/p^n\mathbb{Z}$$

The defining relation between neighbouring components of x suggest the notation $\varprojlim \mathbb{Z}/p^n\mathbb{Z}$ for the set of sequences defining p-adic integers. This notation is not reserved for this particular family of sequences, and is used in more general situations as well. In those $\varprojlim$ denotes the so called **projective** or **inverse** limit. So that we have

$$\mathbb{Z}_p = \varprojlim \mathbb{Z}/p^n\mathbb{Z}.$$

The p-adic integers form a ring with component-wise addition and multiplication: for $x =< [a_n]_{p^n} >$ and $y =< [b_n]_{p^n} >$ we define

$$x + y :=< [a_n + b_n]_{p^n} >, \qquad x \cdot y :=< [a_n \cdot b_n]_{p^n} > .$$

Exercise 7.6 *(1) Prove that the p-adic integers* $\mathbb{Z}_p$ *with the two operations just defined is a commutative ring the* **zero element** $([0]_p, [0]_{p^2}, \dots, [0]_{p^n}, \dots)$ *being the neutral element w.r.t. addition,*

and the **identity element** $([1]_p, [1]_{p^2}, \dots, [1]_{p^n}, \dots)$ *being the neutral element w.r.t. multiplication. These are denoted by* 0 *and* 1 *as usual.*
(2) Let $x = ([a_1]_p, \dots [a_n]_{p^n}, \dots)$ *be a p-adic integer. Prove that the components* $[a_1]_p, \dots, [a_k]_{p^k}$ *are uniquely determined by* $[a_{k+1}]_{p^{k+1}}$*. In particular, if* $[a_k]_{p^k} = [0]_{p^k}$*, then*

$$[a_1]_p = [0]_p, \quad [a_2]_{p^2} = [0]_{p^2}, \quad \dots, \quad [a_{k-1}]_{p^{k-1}} = [0]_{p^{k-1}}.$$

Conclude that a non-zero element $x \in \mathbb{Z}_p$ *either has no zero components,* $(\forall n \geq 1)([a_n]_{p^n} \neq [0]_{p^n})$*, or has the form*

$$([0]_p, \dots, [0]_{p^k}, [a_{k+1}]_{p^{k+1}}, \dots)$$

where $(\forall s \geq 1)([a_{k+s}]_{p^{k+s}} \neq [0]_{p^{k+s}})$*.*

There is a natural map $\varphi_p : \mathbb{Z} \to \mathbb{Z}_p$ defined on $k \in \mathbb{Z}$ by

$$\varphi_p(k) = \{[k]_p, [k]_{p^2}, \dots, [k]_{p^n}, \dots\}.$$

This map is obviously one-to-one, and is a homomorphism (of rings). So, we can naturally consider $\mathbb{Z}$ a subset of $\mathbb{Z}_p$.

What is more interesting here is that some non-integer rational numbers are also naturally a p-adic integers. Indeed, let $x = a/b$ where $\gcd(b, p) = 1$. As we know, since b is invertible modulo any positive power of p, the number a/b determines a class in $\mathbb{Z}/p^n\mathbb{Z}$ for every $n \geq 1$: the only class $[a_n] \in \mathbb{Z}/p^n\mathbb{Z}$ for which $a \equiv b \cdot a_n \mod p^n$. We will denote $[a_n]$ by $[a/b]_{p^n}$. It is straightforward to check (do that!) that the sequence $< [a/b]_{p^n} >$ is a p-adic integer.

A bit more "globally" now, denote by $\mathbb{Z}_{(p)}$ the set of all rational numbers whose reduced representation has denominator not divisible by p:

$$\mathbb{Z}_{(p)} = \{x \in \mathbb{Q} \,|\, x = a/b,\ a > 0, b \in \mathbb{Z},\ \gcd(a, b) = 1,\ p \nmid b\}.$$

The set $\mathbb{Z}_{(p)}$ with the operations addition and multiplication inherited from $\mathbb{Q}$ forms a ring (verify this!). The ring of integers $\mathbb{Z}$ is naturally a subset (identifying an integer m with the rational fraction $m/1$), and actually - a subring of $\mathbb{Z}_{(p)}$. Finally, the map $\varphi_p : \mathbb{Z} \to \mathbb{Z}_p$ extends to a one-to-one homomorphism, denoted the same way,

$$\varphi_p : \mathbb{Z}_{(p)} \to \mathbb{Z}_p, \quad \varphi_p(a/b) = ([a/b]_p, [a/b]_{p^2}, \dots, [a/b]_{p^n}, \dots).$$

So, we may also assume that $\mathbb{Z}_{(p)}$ is a subset (a sub-ring) of $\mathbb{Z}_p$.

Exercise 7.7 *(1) Suppose the p-adic integer x has the form*

$$x = ([0]_p, \dots, [0]_{p^k}, [a_{k+1}]_{p^{k+1}}, \dots)$$

with $[a_{k+1}]_{p^{k+1}} \neq [0]_{p^{k+1}}$*, and* $k \geq 1$*. Prove that there is a unique p-adic integer y such that* $x = \varphi_p(p^k) \cdot y$*. More specifically, let* $b_{k+l} = a_{k+l}/p^k$ *for* $l \geq 1$*. (*$b_{k+l} \in \mathbb{Z}$ *- why?) Prove that the unique y the existence of which is claimed above has the form*

$$y = ([b_{k+1}]_p, \dots, [b_{k+1}]_{p^{k+1}}, \dots, [b_{k+l}]_{p^{k+l}}, \dots).$$

(2) Prove that $x = ([a_1]_p, \dots, [a_n]_{p^n}, \dots) \in \mathbb{Z}_p$ *has a reciprocal, i.e.,* $(\exists y \in \mathbb{Z}_p)(x \cdot y = 1)$*, if, and only if,* $[a_1]_p \neq [0]_p$*.*

The last exercise teaches us that the group of units $(\mathbb{Z}_p)^\times$ of the ring of p-adic integers consists of all $x = ([a_1]_p, \dots, [a_n]_{p^n}, \dots)$ with $p \nmid a_1$.

Exercise 7.8 *Prove that* $\mathbb{Z}_p$ *has no non-zero zero divisors: if* $x \neq 0$ *and* $y \neq 0$*, then* $x \cdot y \neq 0$*. (Rings with this property are called* **integral domains***.)*

7.4.2 The p-adic Numbers $\mathbb{Q}_p$

The definition of $\mathbb{Q}_p$ is pretty formal from algebraic point of view, and resembles the the construction of the rational numbers $\mathbb{Q}$ starting from the integers $\mathbb{Z}$: rational numbers are fractions of integers. In the case at hand, we want to consider fractions of p-adic numbers instead. This is a construction which can be carried over for any ring R without non-zero zero divisors. The result is a field, called the field of fractions of the ring R and denoted by $Frac(R)$. In other words, $\mathbb{Q} = Frac(\mathbb{Z})$, and, by definition, $\mathbb{Q}_p = Frac(\mathbb{Z}_p)$.

The field of fractions of a ring is studied thoroughly in the course Topics in Algebraic Structures. We only sketch one way to construct it here.

Suppose R is a ring with no non-zero zero divisors. Denote

$$Frac(R) = \{(r_1, r_2) \,|\, r_1, r_2 \in R, r_2 \neq 0\}/\sim$$

where $\sim$ is a relation on $R \times (R \setminus \{0\})$ defined by

$$(r_1, r_2) \sim (r_1', r_2') \quad \text{if} \quad r_1 \cdot r_2' = r_1' \cdot r_2.$$

The important fact here is that $Frac(R)$ has a structure of a field with the operations

$$(r_1, r_2) + (r_1', r_2') = (r_1 \cdot r_2' + r_1' \cdot r_2, r_2 \cdot r_2') \quad \text{and} \quad (r_1, r_2) \cdot (r_1', r_2') = (r_1 \cdot r_1', r_2 \cdot r_2').$$

It is customary to write r_1/r_2 instead of (r_1, r_2). Notice that if $R = \mathbb{Z}$, the construction of $Frac(\mathbb{Z})$ is exactly the one that is usually described in the course Intro to Advanced Math.

Exercise 7.9 *Prove that every p-adic number $\alpha \in \mathbb{Q}_p$ can uniquely be represented in the form $\alpha = \varphi_p(p^n) \cdot x$ where $n \in \mathbb{Z}$ and $x \in (\mathbb{Z}_p)^\times$. In addition, $\mathbb{Z}_p$ consists of all such α with $n \in \mathbb{N}$, and $(\mathbb{Z}_p)^\times$ consists of all such α with $n = 0$.*

Based on the exercise above, it is straightforward to verify that the map $\varphi_p : \mathbb{Z}_{(p)} \to \mathbb{Z}_p$ extends to a map $\varphi_p : \mathbb{Q} \to \mathbb{Q}_p$ which is still a homomorphism of rings (so, a homomorphism of fields). Using this map, we identify $\mathbb{Q}$ to its image in $\mathbb{Q}_p$.

7.4.3 Hensel's Lifting Lemma and p-adic Integers

Let's go back to solving polynomial equations modulo powers of prime numbers. Suppose $f(X) \in \mathbb{Z}[X]$, and we want to solve it modulo the powers of a prime number p. Hensel's lifting lemma tells us what happens here: every solution modulo p^{k+l} is an extension of a solution modulo p^k, and we know the mechanism of lifting a solution modulo p^k to modulo p^{k+l}.

Now, we can evaluate a polynomial with integer coefficients at a p-adic integer, and get a p-adic integer as value. To explain this, suppose $x = ([a_1]_p, \dots, [a_n]_{p^n}, \dots) \in \mathbb{Z}_p$. Then, by definition,

$$f(x) = ([f(a_1)]_p, \dots, [f(a_n)]_{p^n}, \dots).$$

Exercise 7.10 *Check that, if $x \in \mathbb{Z}_p$ is as above, then, for every $n \geq 1$, $f(a_{n+1}) \equiv f(a_n) \mod p^n$. Conclude that $f(x) \in \mathbb{Z}_p$.*

The exercise shows that any such polynomial can be considered as a map, a polynomial function, $f : \mathbb{Z}_p \to \mathbb{Z}_p$. We may, therefore, ask about the solutions to $f(X) = 0$ in $\mathbb{Z}_p$. Any such solution $x \in \mathbb{Z}_p$ consists of a sequence $< a_n >_{n \geq 1}$ such that, for $n \geq 1$, $a_{n+1} \equiv a_n \mod p^n$, and $f(a_n) \equiv 0 \mod p^n$. But this is exactly what it means to have a_1 a solution to $f(X) \equiv 0 \mod p$ which is liftable to any $\mod p^n$.

Exercise 7.11 *Prove that $f(X) \equiv 0 \mod p^n$ has a solutions for every $n \geq 1$ if, and only if, there is a solution to $f(X) \equiv 0 \mod p$ which is liftable to $\mod p^n$ for every $n \geq 1$.*

In other words, $f(X) \equiv 0 \mod p^n$ is solvable for every $n \geq 1$ if, and only if, it is solvable in $\mathbb{Z}_p$.

One of the items in Hensel's lifting lemma ensures that if $f(a) \equiv 0 \mod p$, and if $p \nmid f'(a)$, then $f(X) = 0$ has a solution in $\mathbb{Z}_p$. Notice that if $f(a) = 0$ in $\mathbb{Z}$, then $\varphi_p(a) \in \mathbb{Z}_p$ is a solution of $f(X) = 0$ in $\mathbb{Z}_p$ for **every** p. But if $f(X) = 0$ has no integer solutions, then any solution in $\mathbb{Z}_p$ will be in $\mathbb{Z}_p \setminus \varphi_p(\mathbb{Z})$.

An interesting question arising here is the following: **Is it true that $f(X) = 0$ would have solutions in $\mathbb{Z}$ if it has solutions in $\mathbb{Z}_p$ for every prime number p?** The answer is negative in general. We will address this question in more detail in the end of next Chapter.

Naturally, one can consider polynomial functions $f : \mathbb{Q}_p \to \mathbb{Q}_p$ as well where

$$f(X) = b_d X^d + b_{d-1} X^{d-1} + \cdots b_1 X + b_0 \in \mathbb{Z}[X].$$

Indeed, let $\alpha \in \mathbb{Q}_p$ have its unique expression $\alpha = \varphi_p(p^k) \cdot x$ where $x \in \mathbb{Z}_p$, and $k \leq 0$ is an integer. By definition

$$f(\alpha) = \varphi_p(b_d)\varphi_p(p^k)^d x^d + \varphi_p(b_{d-1})\varphi_p(p^k)^{d-1} x^{d-1} + \cdots + \varphi_p(b_1)\varphi_p(p^k)x + \varphi_p(b_0) \in \mathbb{Q}_p.$$

To be saved from dealing with such long and ugly, but correct!, expressions we identify $\mathbb{Q}$ with its image $\varphi_p(\mathbb{Q}) \subset \mathbb{Q}_p$, and write s instead od $\varphi_p(s)$ for any $s \in \mathbb{Q}$. With this simplification, we have $\alpha = p^k x$, and

$$f(\alpha) = b_d p^{kd} x^d + b_{d-1} p^{k(d-1)} x^{d-1} + \cdots + b_1 p^k x + b_0.$$

7.4.4 Hensel's Definition of p-adic Numbers

There is a natural and intuitive way to encode any $x = ([a_1]_p, \ldots, [a_n]_{p^n}, \ldots) \in \mathbb{Z}_p$ using powers of p and a fixed system of residues modulo p. Classically, the residue system is chosen to be $\{0, 1, \ldots, p-1\}$, but we can work with any other fixed system. We choose to be classical in this subsection.

As we know, any class $[a]_{p^n}$ has a unique representative $b \in [a]_{p^n}$ such that $0 \leq b < p^n$. Moreover, this representative has a unique expression to the base p:

$$b = b_0 + b_1 p + \cdots b_{n-1} p^{n-1}, \quad 0 \leq b_0, b_1, \ldots, b_{n-1} < p.$$

Considering $x \in \mathbb{Z}_p$ above, we may assume that for all $n \geq 1$ we have chosen $0 \leq a_n < p^n$. Then, because of the relations $a_{n+1} \equiv a_n \mod p^n$, we obviously have a sequence $< b_n >_{n \geq 1}$, $0 \leq b_n < p$ of integers such that

$$a_1 = b_0, \quad a_2 = b_0 + b_1 p, \quad a_3 = b_0 + b_1 p + b_2 p^2, \quad \cdots, \quad a_n = b_0 + b_1 p + \cdots + b_{n-1} p^{n-1} \quad n \geq 1.$$

So, a p-adic integer is the same as a **formal** power series

$$x = b_0 + b_1 p + \cdots + b_n p^n + \cdots, \quad 0 \leq b_n < p, \quad n \in \mathbb{N},$$

because the components a_n uniquely determine the series, and vice-versa - the series uniquely determines the components a_n. Recalling that any p-adic number α has the form $\alpha = p^{-k} \cdot x$ for unique $x \in \mathbb{Z}_p$ and $k \geq 0$, we see that α can be written as a formal Laurent series

$$\alpha = \frac{b_{-k}}{p^k} + \cdots + \frac{b_{-1}}{p} + b_0 + b_1 p + \cdots b_n p^n + \cdots, \quad 0 \leq b_i < p, \quad i \geq -k.$$

It is an interesting and amusing question about how to recognize $\mathbb{Z}$ and $\mathbb{Q}$ in these notations.

Exercise 7.12 *(1) Prove that, for any $a \in \mathbb{N}$, its expression as a p-adic integer coincides with its expression in base p: $a = b_0 + b_1 p + \cdots + b_s p^s$.*
(2) Prove that the formal power series representing -1 has all its coefficients equal to $p-1$:

$$-1 = (p-1) + (p-1)p + \cdots + (p-1)p^n + \cdots .$$

Prove moreover that every negative integer is expressed by an infinite formal power series.
(3) Prove that in terms of formal power series of $1/(1-p)$ has all coefficients equal to 1:

$$\frac{1}{1-p} = 1 + p + p^2 + \cdots + p^n + \cdots .$$

(4) (A^{++}-students) Prove that $\alpha \in \mathbb{Q}_p$ is a rational number, that is, belongs to the subset $\mathbb{Q}$ of $\mathbb{Q}_p$, if, and only if, the sequence $b_{-k}, \ldots, b_{-1}, b_0, b_1, \ldots, b_n, \ldots$ of the coefficients of its formal Laurent series is periodic from some point onwards.

As stated in the title of this sub-section, this is how Hensel defined the p-adic numbers. Having these numbers as power series, although formal, he wanted to employ the methods of Calculus in studying number theoretical problems. There is a way to consider these formal series as actually convergent series! We will explain this briefly in the second Vista devoted to the p-adic numbers (in the next Chapter). As a final remark here, note that the operations addition and multiplication should be performed in these notations following the rule of "carrying over" familiar from the operations in base 10 system.

7.5 Equations of Higher Degree in $\mathbb{Z}/p\mathbb{Z}$

We learned in the previous section that solving equations $\mod n$ reduces to solving equations $\mod p$. The case of a prime modulus is particularly attractive, for $\mathbb{Z}/p\mathbb{Z}$ is a field, like $\mathbb{R}$ and $\mathbb{C}$, and the polynomials with coefficients in it exhibit properties, similar to the ones the polynomials with real and complex coefficients have. The fact of the matter is that even the case of a degree two equation in $\mathbb{Z}/p\mathbb{Z}$ is non-trivial! Some of the results on such equations bear the names of great mathematicians such as Lagrange, Gauss, and Jacobi... We will devote several lectures to this case in our course. In the current section, we are proving the following fact (which we know well about polynomials with real and complex coefficients).

Theorem 7.5.1 (*Lagrange*(~ 1780)) *Consider the degree d polynomial*

$$[f(X)]_p = [a_d]_p \cdot X^d + [a_{d-1}]_p \cdot X^{d-1} + \cdots + [a_1]_p \cdot X + [a_0]_p$$

with coefficients in $\mathbb{Z}/p\mathbb{Z}$. (This means that $[a_d]_p \neq [0]$.) This polynomial has no more than d roots in $\mathbb{Z}/p\mathbb{Z}$, counted with multiplicities. Equivalently (contra-positively), if a polynomial as above has more than d roots in $\mathbb{Z}/p\mathbb{Z}$, then it is the zero polynomial, that is

$$[a_d]_p = [a_{d-1}]_p = \cdots = [a_0]_p = [0]_p.$$

In the **congruence** notations this theorem says that, if p does not divide a_d, then the congruence equation

$$f(X) = a_d \cdot X^d + a_{d-1} \cdot X^{d-1} + \cdots + a_1 \cdot X + a_0 \equiv 0 \mod p$$

has no more than d incongruent solutions, counted with multiplicity.
Recall that if a polynomial is non-zero, its degree is the largest index of its non-zero coefficients. Hence the degree of a non-zero polynomial is $d \geq 0$. The zero polynomial, by definition, has degree $-\infty$.

Before demonstrate that the theorem is true, we have to define what multiplicity of a root is. We begin with a result which is true for any $n \geq 1$.

Lemma 7.5.2 (*Bézout*) *In the notation above, with $[a_d]_n \neq [0]_n$, assume that $d \geq 1$, and that*

$$f([x_0]_n) := [f(x_0)]_n = [0]_n.$$

Then,

$$[f(X)]_n = (X - [x_0]_n)([a_d]_n X^{d-1} + [b_{d-2}]_n X^{d-2} + \cdots + [b_1]_n X + [b_0]_n)$$

for some integers $b_{d-2}, \ldots, b_1, b_0$. In other words, in such a case

$$[f(X)]_n = (X - [x_0]_n)[g(X)]_n$$

where the degree of $g(X)$ is one less than the degree of $f(x)$, and has the same leading coefficient as $f(x)$.

Proof We have that $[f(x_0)]_n = [a_d]_n [x_0]_n^d + \cdots [a_1]_n [x_0]_n + [a_0]_n = [0]_n$. Therefore

$$[f(X)]_n = [f(X)]_n - [f(X_0)]_n$$

$$= (\,[a_d]_n \cdot X^d + \cdots + [a_1]_n \cdot X + [a_0]_n\,) \;-\; (\,[a_d]_n \cdot [x_0]_n^d + \cdots + [a_1]_n \cdot [x_0]_n + [a_0]_n\,)$$

$$= [a_d]_n \cdot (X^d - [x_0]_n^d) + \cdots + [a_1]_n \cdot (X - [x_0]_n) + ([a_0]_n - [a_0]_n)$$

and by the exercise after this proof we get

$$= (X - [x_0]_n) \cdot ([a_d]_n \cdot X^{d-1} + \cdots) = (X - [x_0]_n) \cdot [g(X)]_n$$

as promised. □

Exercise 7.13 *Prove that* $X^k - [x']_n^k = (X - [x']_n)(X^{k-1} + X^{k-2}[x']_n + \cdots + X[x']_n^{k-2} + [x']_n^{k-1})$.

Using Bézout's lemma, we easily get a corollary (again for ebery $n \geq 1$.

Corollary 7.5.3 *In the notations above, if $[f(X)]_n$ has a root (in $\mathbb{Z}/n\mathbb{Z}$), then*

$$[f(X)]_n = (X - [x_1]_n) \cdots (X - [x_k]_n) \cdot [h(X)]_n$$

where $h(X)$ has degree $d - k$, has a leading coefficient a_d, and doesn't have roots (in $\mathbb{Z}/n\mathbb{Z}$). In particular, since $[h(X)]_n \neq 0$ as a polynomial ($[a_d]_n \neq [0]_n$), its degree is a natural number, that is, $d - k \geq 0$.

Proof By the Bézout's lemma $[f(X)]_n = (X - [x_1]_n) \cdot [g_1(X)]_n$ where $[x_1]_n$ is a root of $[f(X)]_n$, and $g_1(X)$ has degree one less than the degree of $f(X)$, and leading coefficient a_d. If $[g_a(X)]_n$ has no roots in $\mathbb{Z}/n\mathbb{Z}$, we are done. If it does have, then again by the Bézout's lemma we have $[g_1(X)]_n = (X - [x_2]_n) \cdot [g_2(X)]_n$ where $g_2(X)$ has the corresponding degree and leading coefficient. Continuing this way in no more than d steps we will get to a $[g_k(X)]_n$ which has no roots in $\mathbb{Z}/n\mathbb{Z}$, has degree $d-k$ and leading coefficient a_d. We denote $h(X) := g_k(X)$. The corollary is proved. □

Remark 7.5.1 Unfortunately, in general when $n \geq 1$, the classes in the presentation

$$[f(X)]_n = (X - [x_1]_n) \cdots (X - [x_k]_n) \cdot [h(X)]_n$$

are not uniquely defined. Consider for example $f(X) = X^3 - X$, and $n = 6$. We obviously have

$$[f(X)]_6 = (X - [0]_n)(X - [1]_6)(X - [5]_6) = (X - [1]_6)(X - [2]_6)(X - [3]_6)$$

where, in both presentations, $h(X) = 1$. The picture is much better looking when $n = p$ is a prime number! □

Proposition 7.5.4 *Let p be a prime number, and suppose*

$$[f(X)]_p = (X - [x_1]_p) \cdots (X - [x_k]_p) \cdot [h(X)]_p$$

where $[h(X)]_p$ has no roots in $\mathbb{Z}/p\mathbb{Z}$. Then the classes $[x_1]_p, \ldots, [x_k]_p$ constitute the roots (may be with repetition) of the polynomial $[f(X)]_p$. They are, therefore, uniquely determined.

Proof We have to show that $[f(x')]_p = [0]_p$ if, and only if, $[x']_p = [x_i]_p$ for some $i = 1, \ldots, k$. But

$$[0]_p = [f(x')]_p = ([x']_p - [x_1]_p) \cdots ([x']_p - [x_k]_p) \cdot [h(x')]_p$$

if, and only if,

$$p \,|\, (x' - x_1) \cdots (x' - x_k) \cdot h(x')$$

which, since p is a prime number, is equivalent to p dividing one of the $k+1$ factors involved. But since $[h(X)]_p$ has no roots in $\mathbb{Z}/p\mathbb{Z}$, the factor $h(x')$ is **not** divisible by p, it remains to have $p \,|\, x' - x_i$ for some i, which gives the result. □

In the presentation $[f(X)]_p = (X - [x_1]_p) \cdots (X - [x_k]_p) \cdot [h(X)]_p$ we combine the like linear terms together and, after re-indexing the classes involved, rewrite the presentation as

$$[f(X)]_p = (X - [x_1]_p)^{m_1} \cdots (X - [x_s]_p)^{m_s} \cdot [h(X)]_p$$

where $m_1 + \cdots m_s = k$, and $[x_i]_p \neq [x_j]_p$ for $1 \leq i < j \leq s$. In this presentation, the classes $[x_i]_p$ are well defined: all the distinct roots of $[f(X)]_p$ in $\mathbb{Z}/p\mathbb{Z}$. We are proving next that the exponents m_i are also well defined: they are uniquely determined by $f(X)$ itself. In other words, the part

$$(X - [x_1]_p)^{m_1} \cdots (X - [x_s]_p)^{m_s}$$

of the presentation of $[f(X)]_p$ is uniquely determined.

Proposition 7.5.5 *Let $[f(X)]_p$, with roots $[x_1]_p, \ldots, [x_s]_p$, have the following two presentations*

$$(X - [x_1]_p)^{m_1} \cdots (X - [x_s]_p)^{m_s} \cdot [h(X)]_p \;=\; (X - [x_1]_p)^{m'_1} \cdots (X - [x_s]_p)^{m'_s} \cdot [h_1(X)]_p$$

where $[h(X)]_p$ and $[h_1(X)]_p$ have no roots in $\mathbb{Z}/p\mathbb{Z}$. Then,

$$m_1 = m'_1, \ldots, m_s = m'_s \quad \textit{and} \quad [h(X)]_p = [h_1(X)]_p.$$

Proof We are proving first that $m_1 = m'_1, \ldots, m_s = m'_s$. Arguing by RAA, and obviously W.L.O.G., we may assume that $m_1 < m'_1$. Then

$$[0]_p = [f(X)]_p - [f(X)]_p = (X - [x_1]_p)^{m_1} \cdot ([g_1(X)]_p - [g_2(X)]_p)$$

is the zero polynomial (al coefficients are $[0]_p$) where

$$[g_1(X)]_p = (X - [x_2]_p)^{m_2} \cdots (X - [x_s]_p)^{m_s} \cdot [h(X)]_p$$

and

$$[g_2(X)]_p = (X - [x_1]_p)^{(m'_1 - m_1)} \cdots (X - [x_s]_p)^{m'_s} \cdot [h_1(X)]_p.$$

But the product of any two non-zero polynomials with coefficients in $\mathbb{Z}/p\mathbb{Z}$ is non-zero, because the leading coefficient of the product is the product of the leading coefficients of the two polynomials, and is therefore non-zero. It follows from this that $[g_1(X)]_p - [g_2(X)]_p = [0]_p$ is the zero polynomial, or that $[g_1(X)]_p$ and $[g_2(X)]_p$ are equal polynomials. So, we have that

$$(X - [x_2]_p)^{m_2} \cdots (X - [x_s]_p)^{m_s} \cdot [h(X)]_p \;=\; (X - [x_1]_p)^{(m'_1 - m_1)} \cdots (X - [x_s]_p)^{m'_s} \cdot [h_1(X)]_p,$$

and therefore that $[x_1]_p$ is a root of the LHS. Since this is not possible, we get a contradiction which completes the proof that $m_1 = m'_1, \ldots, m_s = m'_s$.

Using the same argument, we easily get that $[h(X)]_p = [h_1(X)]_p$ as well: just consider the zero polynomial

$$[0]_p = [f(X)]_p - [f(X)]_p = (X - [x_1]_p)^{m_1} \cdots (X - [x_s]_p)^{m_s} \cdot ([h(X)]_p - [h_1(X)]_p).$$

The proposition is proved. □

Now we are ready to give the definition of multiplicity of a root. Notice that we can do this only when $n = p$ is a prime number!

Definition 7.5.1 *Let $[x_1]_p, \ldots, [x_s]_p$ be all distinct roots of $[f(X)]_p$ in $\mathbb{Z}/p\mathbb{Z}$, and let*

$$[f(X)]_p = (X - [x_1]_p)^{m_1} \cdots (X - [x_s]_p)^{m_s} \cdot [h(X)]_p$$

with $[h(X)]_p$ having no roots in $\mathbb{Z}/p\mathbb{Z}$. Then we say that $[x_i]_P$ is a root of $f(X)]_p$ **of multiplicity** *m_i. We say that $[x_i]_p$ is a* **simple root** *if $m_i = 1$.*

Exercise 7.14 *Prove that $[x']_p$ is a root of $[f(X)]_p$ of multiplicity m' if, and only if,*

$$[f(X)]_p = (X - [x']_p)^{m'} \cdot [h(X)]_p$$

where $[h(x')]_p \neq [0]_p$.

Proof of Theorem 7.5.1 The proof is now obvious. If $[f(X)]_p$ is non-zero, then its degree is a natural number, and if it has no roots, then we are done ($0 \leq 0$). If $[f(X)]_p$ does have roots, by the presentation

$$[f(X)]_p = (X - [x_1]_p)^{m_1} \cdots (X - [x_s]_p)^{m_s} \cdot [h(X)]_p$$

where $[h(X)]_p$ has no roots we get that $m_1 + \cdots m_s$ is the number of roots of $[f(X)]_p$ counted with multiplicities. On the other hand, $[h(X)]_p$ is non-zero as well, and has degree $d - (m_1 + \cdots + m_s)$ which should be a natural number as well. So, $m_1 + \cdots m_s \leq d$. The theorem is proved. □

Exercise 7.15 *(1) For any positive integer k, find n and a linear equation with coefficients in $\mathbb{Z}/n\mathbb{Z}$ which has k solutions in $\mathbb{Z}/n\mathbb{Z}$.*
(2) Suppose $f(X) = a_0 + a_1X + \cdots + a_dX^d$ has integer coefficients. Prove that $f(X) \equiv 0 \mod p$ has p solutions if, and only if, $[f(X)]_p = (X^p - X) \cdot [g(X)]_p$ where $g(X)$ is a polynomial with integer coefficients.
(3) Let $d > 0$ be an integer. Prove that $X^d - 1 \equiv 0 \mod p$ has $p-1$ solutions if, and only if, $p-1 \mid d$.

7.6 Two Important Examples

We are considering in this section four polynomials - two are related to Fermat's Little Theorem, and the other two are related to Euler's generalization of that theorem.

7.6.1 The polynomials $e_p(X) = X^{p-1} - 1$ and $f_p(X) = (X-1)\cdots(X-(p-1))$

The main character of this chapter is the polynomial

$$X^{p-1} - [1] \quad \text{in} \quad \mathbb{Z}/p\mathbb{Z}.$$

The following theorem describes completely this polynomial.

Theorem 7.6.1 *We have*

$$X^{p-1} - [1] = (X - [1]) \cdot (X - [2]) \cdot \cdots \cdot (X - [p-1]) \quad \textit{in} \quad \mathbb{Z}/p\mathbb{Z}.$$

Proof It's enough to show that the difference

$$X^{p-1} - [1] - (X - [1])(X - [2]) \cdot \cdots \cdot (X - [p-1])$$

is the zero polynomial. But this difference is a polynomial of degree strictly less than $p-1$, and has, by the FLT, $p-1$ roots: $[1], [2], \ldots, [p-1]$. Therefore, by Lagrange's theorem, this polynomial has degree $-\infty$, that is, it is the zero polynomial. □

Recall that two polynomials, with coefficients in a ring (such as $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}/n\mathbb{Z}$ etc.) are equal, by definition, if their coefficients are the same. The theorem above means then that, after eliminating the parentheses of $(X-[1])\cdot(X-[2])\cdots(X-[p-1])$, we will get coefficients the same as in $X^{p-1}-[1]$. This gives us the following important information about those coefficients.

Proposition 7.6.2 *Let p be an odd prime number. Write*

$$(X - [1]) \cdot (X - [2]) \cdots (X - [p-1]) = X^{p-1} + [A_{p-2}] \cdot X^{p-2} + \cdots + [A_1] \cdot X + [A_0]$$

where

$$A_0 = (p-1)!, \quad A_1 = -\sum_{i=1}^{p-1} \frac{(p-1)!}{i}, \quad \ldots, \quad A_{p-3} = \sum_{1 \le i < j \le p-1} ij, \quad A_{p-2} = -\sum_{i=1}^{p-1} i.$$

Then, $[A_0] = -[1]$, $[A_1] = \cdots = [A_{p-2}] = [0]$. In particular,

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

an identity we know as Wilson's Theorem.

Proof That's obvious. □

An important corollary, used in the exercises below, is the following one.

Corollary 7.6.3 *In the notations of the proposition, $A_1 \equiv 0 \mod p^2$ for $p \ge 5$.*

Proof The polynomial with integer coefficients $f_p(X) = (X-1)(X-2)\cdots(X-(p-1))$ has coefficients $A_0, A_1, \ldots, A_{p-2}$ and 1. So, we have

$$f_p(X) = (X-1)(X-2)\cdots(X-(p-1)) = X^{p-1} + A_{p-2}X^{p-2} + \cdots + A_1X + A_0 \quad \text{in} \quad \mathbb{Z}.$$

Evaluating the polynomials at p we get

$$(p-1)! = f_p(p) = p^{p-1} + A_{p-2}p^{p-2} + \cdots + A_2p^2 + A_1p + A_0.$$

Since $A_0 = (p-1)!$, and after simplifications, we get

$$A_1 + A_2p + \cdots + A_{p-2}p^{p-3} + p^{p-2} = 0 \quad \text{in} \quad \mathbb{Z}.$$

The result follows. □

Exercise 7.16 * *(Wolstenholme's Theorem) Prove that if $p \ge 5$ is a prime number, then*

$$1 + 1/2 + \cdots + 1/(p-1) \equiv 0 \mod p^2.$$

7.6.2 The Polynomials $e_n(X) = X^{\varphi(n)} - 1$ and $f_n(X) = \prod_{a \in R(n)}(X - a)$

Denote by $R(n)$ the natural numbers less than n and relatively prime with n.

$$R(n) = \{m \,|\, 1 \leq m < n, \gcd(m, n) = 1\}.$$

Obviously, the set $R(n)$ is a reduced residue system for n.

Consider the polynomial with integer coefficients $e_n(X) := X^{\varphi(n)} - 1$.
By Euler's generalization of FLT, we have that $R(n)$ is a set of solutions to $e_n(X) \equiv 0 \mod n$.
Consider now $f_n(X) = \prod_{a \in R(n)}(X - a)$ over the integers. Obviously, $R(n)$ consists of solutions of $f_n(X) \equiv 0 \mod n$ as well.
More is actually true.

Proposition 7.6.4 *All solutions of $e_n(X) \equiv 0 \mod n$ and of $f_n(X) \equiv 0 \mod n$ belong to $R(n)$.*

Proof Exercise! □

So, $e_n(X)$and $f_n(X)$ are two polynomials having the same $\varphi(n)$ pairwise distinct roots. When $n = p$ is a prime number, as we know, the two polynomials coincide. But this is not at all true in general.

Exercise 7.17 *Verify that the difference $e_n(X) - f_n(X)$ is a polynomial of degree strictly less than $\varphi(n)$, and that it has $\varphi(n)$ distinct roots. Can you conclude from this that*

$$X^{\varphi(n)} - [1] = \prod_{[a] \in (\mathbb{Z}/n\mathbb{Z})^\times} (X - [a])\,?$$

Give arguments to your answer: that is, either prove they are equal, or give a counterexample if they aren't.

Remark 7.6.1 It is a natural question for which composite n do the polynomials $[e_n(X)]_n$ and $[f_n(X)]_n$ coincide. The answer is quite exciting: $n = 2F_k$ where F_k is a Fermat prime number, that is, a prime number of the form $F_k = 2^{2^k} + 1$. (M. Hernandez - M. Yotov, 2014.) □

Chapter 8

Quadratic Equations Modulo n

We are presenting in this chapter the theory of solving quadratic equations modulo a positive natural number. This theory turns out to be very rich mathematically and historically led to many discoveries and generalizations in modern Number Theory. Its development was initiated by Fermat, Euler, Lagrange, and Legendre, and culminated in the work of Gauss who included it in his book *Disquisitiones Arithmeticae* (1801), Arithmetical Investigations, which he wrote at the age of 21, and published three years later. The first text book on Number Theory was written by A.-M. Legendre and titled *Essai sur la théorie des nombres* (1798). But the work that made Number Theory a systematic science is Gauss's *Disquisitiones*. After this chapter we will have covered the first four of the eight chapters of Gauss's book.

The central theorem of this chapter is a beautiful result, conjectured by Euler and Legendre, and proved by Gauss: the **Law of Quadratic Reciprocity**. The Law expresses a relation between pairs of congruences

$$X^2 \equiv q \mod p \qquad \text{and} \qquad X^2 \equiv p \mod q$$

where p and q are distinct odd prime numbers. The proof of the Law we chose to present here is the simplest known (and the third of the eight given by Gauss!). It has two ingredients: **Gauss's Lemma** (computing the Legendre symbol as $\left(\frac{a}{p}\right) = (-1)^{\mu(a,p)}$), and the **ingenious observation of Eisenstein**, a student of Gauss's, (showing that $\mu(p,q) + \mu(q,p) \equiv (p-1)(q-1)/2 \mod 2$).

8.1 General Remarks

8.1.1 Reduction to $X^2 - a \equiv 0 \mod n$ with $\gcd(a,n) = 1$

We want to solve a degree two equation

$$aX^2 + bX + c \equiv 0 \mod n$$

where $n \geq 2, a, b, c \in \mathbb{Z}, [a]_n \neq [0]_n$. This equation is equivalent to the following

$$4a^2X^2 + 4abX + 4ac \equiv 0 \mod 4an$$

which can be given a simpler form

$$(2aX + b)^2 - (b^2 - 4ac) \equiv 0 \mod 4an.$$

Denoting by $Y = 2aX + b$ and $D = b^2 - 4ac$, we see that the original equation reduces to solving for $X \mod n$

$$Y^2 - D \equiv 0 \mod 4an \qquad \text{and} \qquad 2aX + b \equiv Y \mod 4an.$$

The degree two equation we got has a very special form. It is called a **binomial quadratic equation**. One can easily recognize in D the discriminant of a quadratic polynomial from elementary

mathematics. Solving the binomial equation $\mod 4an$ is equivalent to figuring out if "$\sqrt{D}$" exists in $\mathbb{Z}/4an\mathbb{Z}$. To solve for X the system of two equations, we have to first solve the quadratic one. After having solved for Y, finding X is a matter of solving a linear equation which we know how to do. This chapter is therefore devoted to solving binomial quadratic equations

$$f(X) = X^2 - a \equiv 0 \mod n.$$

A very straightforward case to solve is when $a \equiv 0 \mod n$. Indeed, let M be the largest integer whose square divides n so that $n = M^2 \cdot n'$. Note that n' is a square-free natural number: if $n' > 1$, then $n' = p_1 \cdots p_l$ with $p_1 < \cdots < p_l$.

Exercise 8.1 *Prove that $X^2 \equiv 0 \mod n$ if, and only if $X \equiv 0 \mod Mn'$. So, the solutions to the binomial equation are given by the classes $\mod n$ of the numbers $\{Mn'q \,|\, 0 \leq q \leq M-1\}$.*

Having dealt with the case $n \,|\, a$, **we assume in what follows that** $n \nmid a$.

We are showing now that the binomial equation can be reduced in this case (with $n \nmid a$) to a very specific one in which $\gcd(a, n) = 1$. To this end, let $d = \gcd(a, n)$, and let $a = da_1$ and $n = dn_1$. If $x_0 \in \mathbb{Z}$ is a solution to $X^2 \equiv a \mod n$, then, since $d \,|\, x_0^2 - a$ and $d \,|\, a$, we see that $d \,|\, x_0^2$. In other words, x_0 provides a solution to $X^2 \equiv 0 \mod d$ as well. Denoting by m the largest integer whose square divides d, and expressing $d = m^2 d'$, we get that $md' \,|\, x_0$, and therefore that $x_0 = md'x_0'$. Now, for x_0' we compute that

$$(md'x_0')^2 \equiv m^2 d' a_1 \mod m^2 d' n_1 \qquad \Leftrightarrow \qquad d'(x_0')^2 \equiv a_1 \mod n_1,$$

and therefore x_0' provides a solution to

$$d'X^2 \equiv a_1 \mod n_1$$

for relatively prime a_1 and n_1.

Exercise 8.2 *Prove that the existence of a solution to the last equation forces $\gcd(d', n_1) = 1$ to hold true.*

Since $\gcd(d', n_1) = 1$, we can cancel out d', and express the equation as

$$X^2 \equiv a_1 (d')^{-1} \mod n_1.$$

The RHS of this equation is relatively prime with the modulus. Every solution x_0 to $X^2 \equiv a \mod n$ has the form $x_0 = md'x_0'$ where x_0' is a solution to $X^2 \equiv a_1 (d')^{-1} \mod n_1$.

The moral from the last discussion is that W.L.O.G. we may consider solving binomial equations

$$X^2 \equiv a \mod n$$

for $\gcd(a, n) = 1$, and $n \geq 2$.

Exercise 8.3 *Consider, as above, the equation $aX^2 + bX + c \equiv 0 \mod n$, where $n \nmid a$. Let s be the maximal integer such that $s \,|\, 4a$ and $\gcd(s, n) = 1$. Denote $d = 4a/s$. Prove that*

$$aX^2 + bX + c \equiv 0 \mod n \qquad \Leftrightarrow \qquad 4a^2X^2 + 4abX + 4ac \equiv 0 \mod dn.$$

In particular, if $\gcd(2a, n) = 1$, then the original equation is equivalent to $4a^2X^2 + 4abX + 4ac \equiv 0 \mod n$.

As we know from the previous chapter: solving equations $\mod n$ reduces to solving them modulo the prime divisors of n, and then applying the Hensel's Lifting lemma. In this lemma the derivative, $f'(X) = 2X$, of the polynomial of the equation plays an important role: if x_0 is a solution to $f(X) \equiv 0 \mod p$, and if $p \nmid f'(x_0)$, the solution has a unique lift to $f(X) \equiv 0 \mod p^k$ for every positive integer k. Since, in our case, this derivative is zero $\mod 2$ we will have to consider this case separately.

8.1.2 The Equation $X^2 - a \equiv 0 \mod 2^\alpha$ with $\gcd(a, 2) = 1$

The case $p = 2$ needs no deep theory to handle it. The diligent student should have done this already for the particular case of $a = 1$ in an exercise related to the generalized Wilson's theorem.
Notice that if $X^2 \equiv a \mod 2^\alpha$ for $\gcd(a, 2) = 1$ is solvable, then the number of solutions are as many as they are for $a = 1$. Indeed, if $x_0^2 \equiv a \mod 2^\alpha$, then x_0 is odd, and the equations

$$X^2 \equiv a \mod 2^\alpha \qquad \text{and} \qquad X^2 \equiv 1 \mod 2^\alpha$$

are equivalent via associating with any solution $[x]$ to the first equation the solution $[xx_0^{-1}]$ to the second. So, the main question here is whether there are solutions at all. Here is the corresponding general result.

Theorem 8.1.1 *Suppose a is an odd integer, and $\alpha \geq 1$. Consider the equation $X^2 - a \equiv 0 \mod 2^\alpha$. We have the following cases.*
(1) If $\alpha = 1$, there is one solution: $x \equiv 1 \mod 2$.
(2) If $\alpha = 2$, solutions exist if, and only if, $a \equiv 1 \mod 4$. If that is the case, there are two solutions: $x \equiv \pm 1 \mod 4$.
(3) If $\alpha \geq 3$, solutions exist if, and only if, $a \equiv 1 \mod 8$. If that is the case, there are four solutions.

Proof Item (1) is obvious. Item (2) follows from the fact that any solution should be odd, since a is odd, and that the square of an odd number is congruent to $1 \mod 8$. We are proving, by induction on $\alpha \geq 3$, a refined version of item (3): (i) if a solution exists, then $a \equiv 1 \mod 8$, and (ii) if $a \equiv 1 \mod 8$, then for every $\alpha \geq 3$, there are four solutions, $x_\alpha^{(i)}, i = 1, 2, 3, 4$ and only two of which, we call them $x_\alpha^{(1)}, x_\alpha^{(2)}$ have lifts $\mod 2^{\alpha+1}$.
Proving (3, i). Since a is odd, any solution to $X^2 - a \equiv 0 \mod 2^\alpha$ would be an odd number, and therefore its square is congruent to 1 modulo 8. So, since $\alpha \geq 3$, we have that $a \equiv 1 \mod 8$ is a necessary condition for solving the equation.
Proving (3, ii). We argue by induction on $\alpha \geq 3$. Before doing that recall that according to Hensel's lifting Lemma, any lift of a solution x to $X^2 - a \equiv 0 \mod 2^\alpha$ to a solution $\mod 2^{\alpha+1}$ has the form $x' = x + 2^\alpha t$ where

$$f(x) + 2^\alpha f'(x)t = f(x) + 2^{\alpha+1}xt \equiv 0 \mod 2^{\alpha+1},$$

that the lift exists if and only if $2^{\alpha+1} \mid f(x)$, and when this is the case x has two lifts, for $t = 0$ and for $t = 1$.
Base case ($\alpha = 3$). The equation $f(X) = X^2 - a \equiv 0 \mod 8$ has obviously four solutions: $x \equiv 1, 3, 5$, and $7 \mod 8$. We are proving next that only two of these has a lift $\mod 16$. Indeed, any such lift would have the form $x' = x + 8t$ where t is a solution to the equation

$$f(x) + 8tf'(x) \equiv 0 \mod 16, \quad \text{that is} \quad x^2 - a + 16xt \equiv 0 \mod 16.$$

So, solutions for t exist (and will be two such: for $t \equiv 0 \mod 2$ and for $t \equiv 1 \mod 2$) if, and only if $16 \mid x^2 - a$. Observe that since $a \equiv 1 \mod 8$, then $a \equiv 1 \mod 16$ or $a \equiv 9 \mod 16$. In the former case only $x \equiv 1, 7$ have lifts, while in the latter case - only $x \equiv 3, 5$ do. Denoting the liftable solutions by $x_3^{(1)}$ and $x_3^{(2)}$ we find the solutions $\mod 16$ to be

$$x' \equiv x_3^{(1)}, \quad x_3^{(1)} + 8, \quad x_3^{(2)}, \quad x_3^{(2)} + 8 \mod 16.$$

Inductive step ($\alpha \Rightarrow \alpha + 1$). Suppose there are only for solutions $x_\alpha^{(i)}, i = 1, 2, 3, 4$, to $f(X) \equiv 0 \mod 2^\alpha$, that only the first two of them have lifts $\mod 2^{\alpha+1}$. So the solutions $\mod 2^{\alpha+1}$, being lifts of solutions to $f(X) \equiv 0 \mod 2^\alpha$ are four and have the form

$$x_\alpha^{(1)}, \quad x_\alpha^{(1)} + 2^\alpha, \quad x_\alpha^{(2)}, \quad x_\alpha^{(2)} + 2^\alpha \qquad \mod 2^{\alpha+1}.$$

It remains to show that only two of these solutions have lifts $\mod 2^{\alpha+2}$. For the four solutions $\mod 2^{\alpha+1}$ listed above we have

$$f(x_\alpha^{(i)}) = (x_\alpha^{(i)})^2 - a = 2^{\alpha+1}A_{\alpha+1}^{(i)} \qquad \text{and}$$

$$f(x_\alpha^{(i)} + 2^\alpha) = (x_\alpha^{(i)} + 2^\alpha)^2 - a = (x_\alpha^{(i)})^2 - a + 2^{\alpha+1}x_\alpha^{(i)} + 2^{2\alpha}$$
$$= 2^{\alpha+1}(A_{\alpha+1}^{(i)} + x_\alpha^{(i)}) + 2^{2\alpha} \equiv 2^{\alpha+1}(A_{\alpha+1}^{(i)} + x_\alpha^{(i)}) \mod 2^{\alpha+2}.$$

Since $x_\alpha^{(i)}$ are odd, the integers, for a fixed $i = 1, 2$,

$$x_\alpha^{(i)} \qquad \text{and} \qquad A_{\alpha+1}^{(i)} + x_\alpha^{(i)}$$

have different parity. This means that among the four values computed, $f(x_\alpha^{(i)}), f(x_\alpha^{(i)}+2^\alpha), i = 1, 2$, **only two** are divisible by $2^{\alpha+2}$. Therefore, only two of the solutions $\mod 2^{\alpha+1}$ lift to solutions $\mod 2^{\alpha+2}$. This finishes the proof of the inductive step. The theorem is proved. □

Example 8.1.1 Find the solutions to $X^2 - 33 \equiv 0 \mod 64$.
Solution $64 = 2^6$, so solution exists if, and only if, $33 \equiv 1 \mod 8$ which is actually the case. Also, the solutions have to be four. We are finding them now. We start with the solutions $\mod 8$: $1, 3, 5$, and 7. Among the numbers $1^2 - 33, 3^2 - 33, 5^2 - 33$ and $7^2 - 33$ only the first and the last are divisible by 16, so 1 and 7 lift to solutions $\mod 16$, and all solutions $\mod 16$ are $1, 7, 9$, and 15. In a similar fashion we see that the solutions who have lifts to $\mod 32$ are 1 and 15, and the solutions $\mod 32$ are $1, 15, 17$, and 31. Once more we find the two liftable solutions: 17 (b/c $(16+1)^2 - 33 = 64 + 32 + 1 - 33 = 64$), and 31 (b/c, for instance, 1 and 15 are not liftable!). So the solutions $\mod 64$ are: $17, 31, 49$, and 63. All these $\mod 64$ of course. □

8.1.3 The General Theorem on Solving $X^2 - a \equiv 0 \mod n$

In this subsection we are summarizing the results on solving $X^2 - a \equiv 0 \mod n$ for $\gcd(a, n) = 1$ obtained so far.

Theorem 8.1.2 *Let $1 < n \in \mathbb{N}$ have the canonical decomposition $n = 2^\alpha p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ where $\alpha \geq 0$ and $\alpha_1, \ldots, \alpha_k \geq 1$. Let $a \in \mathbb{Z}$ be such that $\gcd(a, n) = 1$. Denote by $M_i, i = 1, \ldots, k$ the number of solutions to $X^2 - a \equiv 0 \mod p_i$. Then, for every $i = 1, \ldots, k$, either $M_i = 0$ or $M_i = 2$. Denote by M the number of solutions to $X^2 - a \equiv 0 \mod n$. Then we have*
(i) if $\alpha = 0$ or 1, then $M = M_1 \cdots M_k$ (so, $M = 0$ or $M = 2^k$);
(ii) if $\alpha = 2$, then $M = 2M_1 \cdots M_k$ if $a \equiv 1 \mod 4$, and $M = 0$ otherwise (so, $M = 0$ or $M = 2^{k+1}$);
(iii) if $\alpha \geq 3$, then $M = 4M_1 \cdots M_k$ if $a \equiv 1 \mod 8$, and $M = 0$ otherwise (so, $M = 0$ or $M = 2^{k+2}$).

Proof We know that the number of solutions to $X^2 - a \equiv 0 \mod n$ is a product of the numbers of solutions to the same equation modulo the maximal powers of the primes dividing n. But since the derivative of $X^2 - a$ is $2X$, and since $\gcd(a, n) = 1$ we get that the number of solutions $\mod p_i^{\alpha_i}$ is the same as the number of solutions $\mod p_i$ (every solution $\mod p_i$ has a unique lift to a solution $\mod p_i^s$ for any positive integer s). Considering $X^2 - a \equiv 0 \mod p_i$, we know that there are no more than two solutions (the degree of the equation!). Also, if the equation at hand has one solution, it will have a second one as well (why?). So $M_i = 0$ if $X^2 - a \equiv 0 \mod p_i$ has no solutions, and $M_i = 2$ if it does. The formualae for M follow now from what we proved about the solutions to $X^2 - a \equiv 0 \mod 2^\alpha$ in the previous subsection. □

8.2 Quadratic Residues Modulo p

We are studying in this section the quadratic polynomials of the form $X^2 - [a]$ in $\mathbb{Z}/p\mathbb{Z}$ where $[a] \neq [0]$. We want to understand when such an equation does have a solution. Since the solutions can only be non-zero in $\mathbb{Z}/p\mathbb{Z}$, we will effectively be studying $(\mathbb{Z}/p\mathbb{Z})^\times$ by solving quadratic equations in $\mathbb{Z}/p\mathbb{Z}$. For a reason discussed in the previous section, we are considering only odd primes p in what follows.

8.2.1 Quadratic Residues and Non-Residues Modulo n

We begin with a definition concerning $\mod n$ for any positive integer.

Definition 8.2.1 *The element $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ is called a* **quadratic residue modulo** n (QR_n) *if the equation $X^2 - [a]_n = [0]_n$ has a solution in $\mathbb{Z}/n\mathbb{Z}$. The element $[a]$ is called a* **non-residue modulo** n (NR_n) *if it is not a QR_n. The set of all quadratic residues modulo n is denoted by $(\mathbb{Z}/n\mathbb{Z})^{\times 2}$.*

As usually happens, whenever no confusion is feared, we will write QR and NR instead of QR_n and NR_n respectively. By the definition it follows that

$$(\mathbb{Z}/n\mathbb{Z})^{\times 2} = \{[a_1]^2, [a_2]^2, \ldots, [a_{\varphi(n)}]^2\}$$

where the classes we square form a reduced residue system $\mod n$:

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[a_1], [a_2], \ldots, [a_{\varphi(n)}]\}.$$

We know that $((\mathbb{Z}/n\mathbb{Z})^\times, \cdot)$ is a group. This fact makes the QRs and NRs behave well under multiplication. On the contrary, the other operation in $\mathbb{Z}/n\mathbb{Z}$ messes up the QRs and NRs, as you will see doing some of the exercises later on.

Theorem 8.2.1 *Let p be an odd prime, and let $k > 0$ be an integer. Then there is equal number of QRs and NRs modulo p^k. That is, the set $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2}$ constitutes half of the set $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$.*

Proof The set of quadratic residues consists of the squares of elements of $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$. We need to know when two such squares are identical. To this end observe that if p is odd, $p \,|\, a-b$ and $p \,|\, a+b$, then $p \,|\, 2a$, and $p \,|\, 2b$, and therefore $p \,|\, a$ and $p \,|\, b$. From this it follows that, for $[a], [b] \in \left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$,

$$[a]^2 = [b]^2 \quad \leftrightarrow \quad p^k \,|\, (a-b)(a+b) \quad \leftrightarrow \quad (p^k \,|\, a-b \ \vee \ p^k \,|\, a+b) \quad \leftrightarrow \quad ([a] = [b] \ \vee \ [a] = [-b]).$$

This means that every quadratic residue comes as the square of EXACTLY TWO distinct elements of $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times$. The theorem is proved. $\square$

Exercise 8.4 *(1) Show by examples that the claim of the theorem above is not true of $n \neq p^k$ for an odd prime p, and positive integer k.*
(2) Prove that $\left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2} = \{[a]^2 \,|\, 1 \leq a \leq (p^k-1)/2 \quad \wedge \quad \gcd(a,p) = 1\}$. Conclude that

$$|\left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2}| = (p^k - p^{k-1})/2 = \varphi(p^k)/2.$$

The next theorem says, in professional terms, that the set $(\mathbb{Z}/n\mathbb{Z})^{\times 2}$ of QRs forms a sub-group of $(\mathbb{Z}/n\mathbb{Z})^\times$, and because of the previous theorem - when $n = p^k$ it is a subgroup of index 2. We denote the latter fact by

$$[\left(\mathbb{Z}/p^k\mathbb{Z}\right)^\times : \left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2}] = 2.$$

In the end of this chapter, we will be able to compute the index of the sub-group of quadratic residues $\mod n$ for any n. More about subgroups and their indices can be learned in Algebraic Structures!

Theorem 8.2.2 *Let n be a positive integer. The following hold true*
(1) If $[a]$ is a quadratic residue modulo n, then so is its reciprocal $[a]^{-1}$;
(2) a product of two quadratic residues modulo n is a quadratic residue modulo n: $(QR \times QR = QR)$;
(3) a product of a quadratic residue with a non-residue modulo n is a non-residue modulo n: $(QR \times NR = NR)$;
(4) product of two quadratic non-residues modulo p^k is a quadratic residue modulo p^k: $(NR_{p^k} \times NR_{p^k} = QR_{p^k})$.

Proof Item (1) is obvious: $[a] = [x]^2 \Leftrightarrow [a]^{-1} = ([x]^{-1})^2$.
The items (3) and (4) follow from (1), (2), and the previous theorem.
For (2), observe that if $[a] = [a_1]^2$ and $[b] = [b_1]^2$, then $[a][b] = [a_1 b_1]^2$, so, $[a][b]$ is a quadratic residue modulo n as well.
For (3), assume $[a] = [a_1]^2$ and $[c] = [c_1]^2$, and that $[a][b] = [c]$. Then, since $[b] = [a]^{-1}[c]$, and by items (1) and (2), the element $[b]$ is a quadratic residue modulo n. Equivalently, if $[a]$ is a quadratic residue, and $[b]$ is a non-residue modulo n, then the product $[a][b]$ is a non-residue modulo n.
We are proving (4) now. Obviously $(\mathbb{Z}/n\mathbb{Z})^\times \setminus (\mathbb{Z}/n\mathbb{Z})^{\times 2} \neq \emptyset$, there are quadratic non-residues modulo n. Let $[b]$ be a NR_n. By (3), the set

$$[b]\,(\mathbb{Z}/n\mathbb{Z})^{\times 2} = \{[b][a] \,|\, [a] \in (\mathbb{Z}/n\mathbb{Z})^{\times 2}\}$$

is disjoint from $(\mathbb{Z}/n\mathbb{Z})^{\times 2}$, and has the same cardinality. By the preceding theorem, when $n = p^k$,

$$[b]\left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2} \cup \left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2} = \left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times}.$$

Therefore

$$[b]\left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2} = \left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times} \setminus \left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2}$$

and the non-residues modulo p^k constitute the set $[b]\left(\mathbb{Z}/p^k\mathbb{Z}\right)^{\times 2}$. Let now $[b_1]$ and $[b_2]$ be NR_{p^k}. According to the last thing we proved $[b_1] = [b][a_1]$ and $[b_2] = [b][a_2]$ for some quadratic residues $[a_1]$ and $[a_2]$. Therefore $[b_1][b_2] = [b]^2[a_1][a_2]$ which is a quadratic residue as a product of three such.
The theorem is proved. □

Exercise 8.5 *(1) Does there exist a perfect square of the form* $1! + 2! + \cdots + n!$ *where* $n > 3$*?*
(2) Prove that for no $n > 1$ *is the sum* $(1!)^2 + (2!)^2 + \cdots + (n!)^2$ *a perfect square.*

8.2.2 The Legendre Symbol, and the Euler's Criterion

The following theorems are true for $n = p$ is an odd prime number. As we know, this case is enough in order to study the theory of solving binomial equations modulo any n.

Definition 8.2.2 *Let* p *be an odd prime, and* $a \in \mathbb{Z}$ *be relatively prime with* p *(that is,* $[a] \in (\mathbb{Z}/p\mathbb{Z})^\times$*). Define* $\left(\frac{a}{p}\right)$*, the* **Legendre symbol of** a **modulo** p*, as follows*

$$\left(\frac{a}{p}\right) = 1 \quad \textit{if} \quad [a] \in (\mathbb{Z}/p\mathbb{Z})^{\times 2}$$

and

$$\left(\frac{a}{p}\right) = -1 \quad \textit{if} \quad [a] \notin (\mathbb{Z}/p\mathbb{Z})^{\times 2}.$$

Notice that, the Legendre symbol of a modulo p depends actually on the class of a modulo p:

$$a \equiv b \mod p \quad \rightarrow \quad \left(\frac{a}{p}\right) = \left(\frac{b}{p}\right).$$

In these notations, the theorem above gets the form

Theorem 8.2.3 *Let* p *be an odd prime, and let* $a, b \in \mathbb{Z}$ *be non-zero modulo* p*. Then*

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

Proof Obvious. □

In professional terms, the Legendre symbol is a **homomorphism** from the group $(\mathbb{Z}/p\mathbb{Z})^\times$ to the group $(\{-1, 1\}, \cdot)$ whose **kernel** is $(\mathbb{Z}/p\mathbb{Z})^{\times 2}$.

Exercise 8.6 *(1) Let p and q be two odd prime numbers such that $p = q + 4a$. Is it true that*

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

(2) Let p be an odd prime number. Prove that

$$\sum_{j=1}^{p-1} \left(\frac{j}{p}\right) = 0.$$

(3) Let $p > 3$ be a prime number. Show that

$$\sum_{1 \le j \le p-1, \left(\frac{j}{p}\right)=1} j \quad \equiv \quad 0 \mod p.$$

It is obvious that, to answer if the polynomial $X^2 - [a], [a] \neq [0]$, has a root in $\mathbb{Z}/p\mathbb{Z}$ we have to compute the Legendre symbol $\left(\frac{a}{p}\right)$. The following result gives us a means to do so.

Theorem 8.2.4 (**Euler's Criterion**) *Let p be an odd prime number, and let $a \in \mathbb{Z}$ be a non-zero modulo p integer. Then*

$$a^{(p-1)/2} \equiv \left(\frac{a}{p}\right) \mod p.$$

Proof Since p is an odd prime, we have the factorization of polynomials in $\mathbb{Z}/p\mathbb{Z}$

$$X^{p-1} - [1] = (X^{(p-1)/2} - [1])(X^{(p-1)/2} + [1]).$$

We know, by Fermat's Little Theorem, that the LHS polynomial has $p-1$ roots in $\mathbb{Z}/p\mathbb{Z}$, and that they constitute the set $(\mathbb{Z}/p\mathbb{Z})^\times$. Therefore the RHS has the same roots. Each of the factors there has at most $(p-1)/2$ roots in $\mathbb{Z}/p\mathbb{Z}$, and since their total is $p-1$, each of the factors has $(p-1)/2$ roots in $\mathbb{Z}/p\mathbb{Z}$. But, obviously, every quadratic residue modulo p is a root of the first factor (b/c if $[a] = [b]^2$, then $[a]^{(p-1)/2} = [b]^{p-1} = [1]$ by FLT!), so that $(\mathbb{Z}/p\mathbb{Z})^{\times 2}$ consists of **all** the roots of $X^{(p-1)/2} - [1]$. Therefore, the non-residues modulo p must **all** be roots of the second factor: $X^{(p-1)/2} + [1]$. Put another way, this means that

$$a^{(p-1)/2} \equiv 1 \mod p \quad \leftrightarrow \quad [a] \in (\mathbb{Z}/p\mathbb{Z})^{\times 2} \quad \leftrightarrow \quad \left(\frac{a}{p}\right) = 1$$

and this is exactly what the Euler's Criterion states. □

8.2.3 Computing $\left(\frac{-1}{p}\right)$ and $\left(\frac{2}{p}\right)$

The statements about the values of the two cases of Legendre symbols written in the title of this subsection are known as **supplementary laws of quadratic reciprocity**.

Theorem 8.2.5 *Let p be an odd prime number. Then $[-1] \in (\mathbb{Z}/p\mathbb{Z})^{\times 2}$ if, and only if, $p \equiv 1 \mod 4$. Informally, $\sqrt{-1}$ exists in $\mathbb{Z}/p\mathbb{Z}$ if, and only if, p is $1 \mod 4$.*

Proof This immediately follows from the Euler's Criterion:

$$(-1)^{(p-1)/2} \equiv 1 \mod p \quad \leftrightarrow \quad (p-1)/2 \text{ is even.} \quad \square$$

This was an easy, and not quite interesting application of Euler's Criterion to compute a Legendre symbol! The next one, done for the first time by Gauss, is much more interesting.

Theorem 8.2.6 *Let p be an odd prime. Then*

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

Proof The goal is to compute $2^{(p-1)/2}$ modulo p. Here is the ingenious way Gauss did that! Consider

$$2^{(p-1)/2}(\frac{p-1}{2})! = 2 \cdot 4 \cdots (p-3) \cdot (p-1)$$

instead. Some of the factors of the RHS of the last identity are less than or equal to $(p-1)/2$, some are bigger. We will change the latter ones substituting any such $2a$ with $2a - p$ (which is an odd number!). Notice that

$$-(p-1)/2 \leq 2a - p < 0,$$

so, these are negative numbers. Let's denote the amount of these numbers by $\mu(2, p)$. On the other hand, $2a \equiv p - 2b \mod p$ if, and only if, $2a + 2b \equiv 0 \mod p$ which is impossible to happen if $0 < a, b \leq (p-1)/2$. So, after changing every $2a$ in the range $(p-1)/2 < x \leq p-1$ with $(-1)(p-2a)$ we get

$$2 \cdot 4 \cdots (p-3) \cdot (p-1) \equiv (-1)^{\mu(2,p)} \cdot 1 \cdot 2 \cdots \frac{p-3}{2} \cdot \frac{p-1}{2},$$

and, therefore,

$$2^{(p-1)/2} \cdot (\frac{p-1}{2})! \equiv (-1)^{\mu(2,p)} \cdot (\frac{p-1}{2})! \mod p.$$

Cancelling the factor $(\frac{p-1}{2})!$ from both sides of the last identity, we get that

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

Claim 8.2.7 *We have*

$$p = 8k+1 \quad \to \quad \mu(2,p) = 2k \qquad\qquad p = 8k+3 \quad \to \quad \mu(2,p) = 2k+1;$$

$$p = 8k+5 \quad \to \quad \mu(2,p) = 2k+1 \qquad\qquad p = 8k+7 \quad \to \quad \mu(2,p) = 2k+2.$$

Proof of Claim The proof of this claim is a simple counting! (do it as an Exercise!). □

To finish the proof of the Theorem, we need to check that $(p^2-1)/8$ is even exactly when p is 1 or 7 mod 8. This is also straightforward (and comprises another Exercise!). The theorem is proved. □

Suppose $a = \pm p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ is the canonical presentation of $a \in \mathbb{Z} \setminus \{0, \pm 1\}$, and that $gcd(a, p) = 1$. Then, by the properties of the Legendre symbol, we have that

$$\left(\frac{a}{p}\right) = \left(\frac{\pm p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}}{p}\right) = \left(\frac{\pm 1}{p}\right) \left(\frac{p_1}{p}\right)^{\alpha_1} \left(\frac{p_2}{p}\right)^{\alpha_2} \cdots \left(\frac{p_k}{p}\right)^{\alpha_k}.$$

So, learning how to compute the Legendre symbol in general, one needs to know how to compute the Legendre symbol $\left(\frac{p}{q}\right)$ where q is a prime number different from p. We know already how to compute $\left(\frac{2}{p}\right)$, so what is left is to deal with odd prime numbers q. This we will do in the next section. We are finishing this one with an application of what we know so far and with an exercise.

Theorem 8.2.8 *The prime numbers* $1 \mod 4$ *are infinitely many.*

Proof Assuming, by contradiction, these primes are finitely many: $p_1 < p_2 < \cdots < p_k$, consider the number $(2p_1p_2\cdots p_k)^2 + 1$. This is a bigger than one odd integer, so it has a divisor an odd prime, say p. We have for this prime that $-1 \in (\mathbb{Z}/p\mathbb{Z})^{\times 2}$, because $x_0 = 2p_1p_2\cdots p_k$ is a solution to

$$X^2 + 1 \equiv 0 \mod p.$$

So, $\left(\frac{-1}{p}\right) = 1$ and, therefore, $p \equiv 1 \mod 4$. This means that, according to our assumption, $p \in \{p_1, p_2, \ldots, p_k\}$. This immediately leads to a contradiction! □

Exercise 8.7 *(1) Let p be an odd prime. Prove that*

$$\left(\frac{(p-1)/2}{p}\right) = (-1)^{(p-1)(p+5)/8}.$$

(2) Does there exist a perfect square of the form $(-1) \mod 1997$?
(3) Let $n \in \mathbb{N}$. Prove that the odd prime divisors of $n^2 + 1$ have the form $1 \mod 12$ or $5 \mod 12$.
(4) Does the congruence $X^2 \equiv 2 \mod 231$ have any solutions? If yes, what are they? If not, explain why not?
(5) Is it true that there are infinitely many natural numbers n such that $23 \mid n^2 + 14n + 47$? Explain.
(6) Let $q = 2p + 1$ and $p = 4m + 3 \geq 7$ be prime numbers. Prove that $q \mid M_p = 2^p - 1$.

8.3 Gauss's Lemma

Let $a \in \mathbb{Z}$ be relatively prime with the prime number p. We want to compute the Legendre symbol $\left(\frac{a}{p}\right)$. By Euler's Criterion, $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \mod p$. To simplify our notations, we denote $P := (p-1)/2$.

As in the case of $a = 2$, consider the product

$$(1 \cdot a)(2 \cdot a)\cdots(P \cdot a) = a^P \cdot P!.$$

If $1 \leq k \leq P$, then both a and k are relatively prime with p, and so p does not divide ka. This means that, by the division with quotient and remainder,

$$ka = p \cdot q_k + r_k, \qquad 0 < r_k < p.$$

There are two options for r_k: either $0 < r_k \leq P$, or $P < r_k < p$. In the latter case $P - p < r_k - p < 0$, and so, in that case,

$$-P = P - (p-1) \leq r_k - p < 0.$$

This means that, in both cases above, we can replace ka with an integer a_k such that

$$ka \equiv a_k \mod p, \quad \text{and} \quad -P \leq a_k \leq P.$$

Definition 8.3.1 *In the notations above, denote by $\mu(a, p)$ the amount of negative a_ks. That is*

$$\mu(a, p) := |\{\, a_k \mid a_k < 0 \,\}|.$$

Obviously,

$$(1 \cdot a)(2 \cdot a)\ldots(P \cdot a) \equiv a_1 \cdot a_2 \cdots a_P = (-1)^{\mu(a,p)}|a_1| \cdot |a_2| \cdots |a_P| \mod p.$$

Lemma 8.3.1 *In the notations above,* $|a_1| \cdot |a_2| \cdots |a_P| = P!$.

Proof Since $-P \leq a_k \leq P$, and since $a_k \neq 0$ for every $k = 1, 2, \ldots P$, we have that

$$0 < |a_k| \leq P \quad \text{for all} \quad k = 1, 2, \ldots, P.$$

It is enough, therefore, to prove that $|a_i| \neq |a_j|$ for all $1 \leq i < j \leq P$. But we have that $a_k \equiv ka \mod p$, and hence

$$|a_i| = |a_j| \quad \rightarrow \quad a_i = \pm a_j \quad \rightarrow \quad ia \equiv \pm ja \mod p.$$

The latter congruence means that

$$(i \pm j) \cdot a \equiv 0 \mod p$$

which, because of $gcd(a, p) = 1$, implies that $i \pm j \equiv 0 \mod p$. Recall that $1 \leq i < j \leq P$. This means that $0 < |i \pm j| < P$, and so, the number $i \pm j$ is **never** divisible by p. This contradiction rules out the possibility of having $|a_i| = |a_j|$ for any $1 \leq i < j \leq P$, and completes the proof of lemma. □

Theorem 8.3.2 (**Gauss's Lemma**) *For an odd prime p and a relatively prime with p integer a, we have*

$$\left(\frac{a}{p}\right) = (-1)^{\mu(a,p)}.$$

Proof. We have that

$$a^P \cdot P! \equiv (-1)^{\mu(a,p)} |a_1| \cdot |a_2| \cdots |a_P| = (-1)^{\mu(a,p)} \cdot P! \mod p,$$

and so,

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

By the Euler's criterion, therefore, we immediately get that

$$\left(\frac{a}{p}\right) \equiv (-1)^{\mu(a,p)} \mod p$$

which means that $p \mid \left(\frac{a}{p}\right) - (-1)^{\mu(a,p)}$. But the latter difference can be equal to ± 2 or 0 only. Since p is an odd prime, the divisibility is possible only if the difference is 0, that is, only if

$$\left(\frac{a}{p}\right) = (-1)^{\mu(a,p)}. \quad \square$$

Exercise 8.8 *1) Prove that, for any a with $(a, p) = 1$ we have*

$$\mu(a, p) = \left| \left[\left(\frac{p}{2}, p\right) \cup \left(\frac{3p}{2}, 2p\right) \cup \cdots \cup \left(\frac{2b-1}{2} p, bp\right) \right] \cap \{a, 2a, \ldots, Pa\} \right|$$

where $b = \lfloor a/2 \rfloor$ and $P = (p-1)/2$.
2)Prove that, in the notations of 1),

$$\mu(a, p) = \left| \left[\left(\frac{p}{2a}, \frac{p}{a}\right) \cup \left(\frac{3p}{2a}, \frac{2p}{a}\right) \cup \cdots \cup \left(\frac{2b-1}{2a} p, \frac{bp}{a}\right) \right] \cap \{1, 2, \ldots, P\} \right|,$$

and conclude that

$$\mu(a, p) = \left| \left[\left(\frac{p}{2a}, \frac{p}{a}\right) \cup \left(\frac{3p}{2a}, \frac{2p}{a}\right) \cup \cdots \cup \left(\frac{2b-1}{2a} p, \frac{bp}{a}\right) \right] \cap \mathbb{Z} \right|.$$

8.4 Eisenstein's Theorem

Theorem 8.4.1 (**Eisenstein's Theorem**) *Let p and q be distinct odd prime numbers. Then, modulo 2, $\mu(p,q)+\mu(q,p)$ is the number of points with integer coordinates on the rectangle*

$$\{\,(x,y)\,|\,1/2 < x \leq (p-1)/2),\quad 1/2 < y \leq (q-1)/2\,\},$$

and, therefore,

$$\mu(p,q)+\mu(q,p) \equiv \frac{p-1}{2}\cdot\frac{q-1}{2} \mod 2.$$

Proof To prove the theorem, we need an additional observation to make.

Lemma 8.4.2 *Let $[a] \neq [0]$ in $\mathbb{Z}/p\mathbb{Z}$, and let $a = p\cdot m + r$ where $0 < r < p$. We have (with $P := (p-1)/2$*
(1) $0 < r \leq P \quad\rightarrow\quad \lfloor\frac{2a}{p}\rfloor = 2m$;
(2) $P < r \leq p-1 \quad\rightarrow\quad \lfloor\frac{2a}{p}\rfloor = 2m+1$.
Therefore,

$$\mu(a,p) \equiv \sum_{j=1}^{P}\lfloor\frac{2ja}{p}\rfloor \mod 2.$$

Proof of Lemma An easy exercise! □ The following is a very important technical result on the way of proving the theorem.

Claim 8.4.3 *For odd $a \in \mathbb{Z}$, and for odd prime numbers p we have that*

$$\mu(a,p) \equiv \sum_{j=1}^{P}\lfloor\frac{ja}{p}\rfloor \mod 2.$$

Proof of Claim It follows from this lemma that, for any integer a not divisible by p,

$$\left(\frac{a}{p}\right) = (-1)^{\mu(a,p)} = (-1)^{\sum_{j=1}^{P}\lfloor\frac{2ja}{p}\rfloor}.$$

When a is also odd, so that $(a+p)/2$ is an integer, we can say more:

$$\left(\frac{2}{p}\right)\cdot\left(\frac{a}{p}\right) = \left(\frac{2a}{p}\right) = \left(\frac{\frac{4\cdot(a+p)}{2}}{p}\right) = \left(\frac{4}{p}\right)\cdot\left(\frac{(a+p)/2}{p}\right) = \left(\frac{(a+p)/2}{p}\right) = (-1)^{\mu(\frac{a+p}{2},p)},$$

and, by lemma above,

$$\left(\frac{(a+p)/2}{p}\right) = (-1)^{\sum_{j=1}^{P}\lfloor\frac{2j(a+p)/2}{p}\rfloor}.$$

A simple computation is in order now:

$$\sum_{j=1}^{P}\lfloor\frac{2j(a+p)/2}{p}\rfloor = \sum_{j=1}^{P}\lfloor\frac{j\cdot(a+p)}{p}\rfloor = \sum_{j=1}^{P}\lfloor\frac{ja+jp}{p}\rfloor = \sum_{j=1}^{P}\left(\lfloor\frac{ja}{p}\rfloor + j\right)$$

$$= \sum_{j=1}^{P}\lfloor\frac{ja}{p}\rfloor + \sum_{j=1}^{P} j = \sum_{j=1}^{P}\lfloor\frac{ja}{p}\rfloor + \frac{P(P+1)}{2} = \sum_{j=1}^{P}\lfloor\frac{ja}{p}\rfloor + \frac{p^2-1}{8}.$$

Recalling that

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

we get that

$$\left(\frac{2}{p}\right)\cdot\left(\frac{a}{p}\right) \quad = \quad \left(\frac{2}{p}\right)\cdot(-1)^{\sum_{j=1}^{P}\lfloor\frac{ja}{p}\rfloor},$$

and, therefore, that

$$(-1)^{\mu(a,p)} = \left(\frac{a}{p}\right) = (-1)^{\sum_{j=1}^{P}\lfloor\frac{ja}{p}\rfloor}.$$

The last equalities are equivalent to

$$\mu(a,p) \equiv \sum_{j=1}^{P}\lfloor\frac{ja}{p}\rfloor \mod 2.$$

The Claim is proved. □

Returning to the proof of the theorem, specifying the result in the Claim, we get

$$\mu(p,q)+\mu(q,p) \equiv \sum_{i=1}^{P}\lfloor\frac{iq}{p}\rfloor + \sum_{j=1}^{Q}\lfloor\frac{jp}{q}\rfloor \mod 2.$$

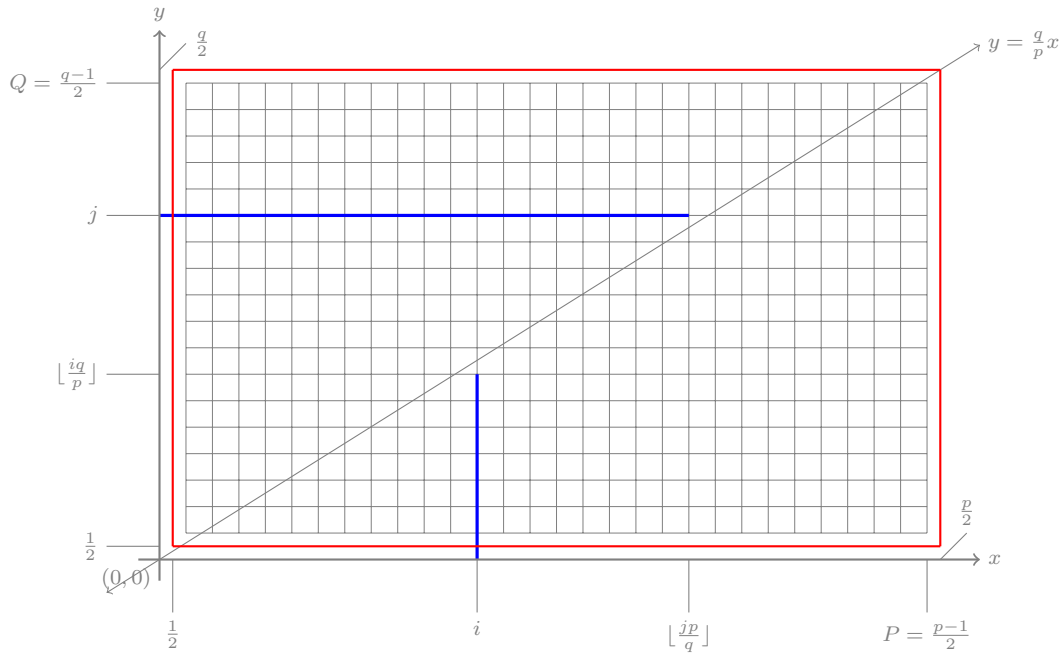
To finish the proof we are establishing, through an ingenious geometric argument, the important relation (here $Q := (q-1)/2$)

$$\sum_{i=1}^{P}\lfloor\frac{iq}{p}\rfloor + \sum_{j=1}^{Q}\lfloor\frac{jp}{q}\rfloor = P\cdot Q.$$

Let's observe now that the RHS of this equality has a nice geometric meaning. In the following diagram we have plotted the graph l of $y = (q/p)x$ in a closed proximity of the rectangle with vertices

$$(0,0),(p/2,0),(p/2,q/2),(0,q/2).$$

The grid shown in the diagram has as vertices the integral points in that rectangle. Notice that, since $\gcd(p,q)=1$, no points of this grid lie on the line l.



Let $1 \le i \le P$ be fixed. The point $(i, iq/p)$ belongs to l, and $\lfloor iq/p \rfloor$ is the length of the segment connecting $(i,0)$ to $(i,\lfloor iq/p\rfloor)$, which is equal to the number of the nodes of the grid above $(i,0)$,

and below l. Therefore the sum

$$\sum_{i=1}^{P} \lfloor \frac{iq}{p} \rfloor$$

is equal to the number of nodes of the grid lying below the line l. In a similar way, we have that for every fixed $1 \leq j \leq Q$, the point $(jp/q, j)$ belongs to l, and the number $\lfloor jp/q \rfloor$ is equal to the length of the segment connecting $(0, j)$ to $(\lfloor jp/q \rfloor, j)$. This length is equal to the number of nodes to the right of $(0, j)$, and to the left of l. Therefore, the sum

$$\sum_{j=1}^{Q} \lfloor \frac{jp}{q} \rfloor$$

is equal to the number of the nodes of the grid above the line l. Noticing that the total number of nodes in the rectangle with vertices

$$(1/2, 1/2), (p/2, 1/2), (p/2, q/2), (1/2, q/2)$$

is $P \cdot Q$, we get that

$$\sum_{i=1}^{P} \lfloor \frac{iq}{p} \rfloor + \sum_{j=1}^{Q} \lfloor \frac{jp}{q} \rfloor = P \cdot Q.$$

This completes the proof of Eisenstein's theorem. □

Exercise 8.9 *1) Let $(x, y) \subset \mathbb{R}$ be an open interval. Prove that, for every integer $n \in \mathbb{Z}$, we have*

$$|\,(x + n, y + n) \cap \mathbb{Z}\,| = |\,(x, y) \cap \mathbb{Z}\,|.$$

2) Let $n \in \mathbb{N}$, and let (x, y) be as in 1). Prove that

$$|\,(x, y + n) \cap \mathbb{Z}\,| = |\,(x - n, y) \cap \mathbb{Z}\,| = |\,(x, y) \cap \mathbb{Z}\,| + n.$$

3) For (x, y) as in 1) show that

$$|(x, y) \cap \mathbb{Z}| = |(-y, -x) \cap \mathbb{Z}|\,.$$

4) If (x, y) is as in 1), and $x, y \notin \mathbb{Z}$, then for every $n \in \mathbb{N}$ such that $y < x + n$ we have

$$|(x, y) \cap \mathbb{Z}| + |(y, x + n) \cap \mathbb{Z}| = n.$$

Exercise 8.10 *1) Let $(a, p) = 1$, and let $p = 4ak + r$ with $0 < r < 4a$. Using the results in Exercises 8.8 and 8.9 prove that*

$$\mu(a, p) \equiv \left| \left[\left(\frac{r}{2a}, \frac{r}{a} \right) \cup \left(\frac{3r}{2a}, \frac{2r}{a} \right) \cup \cdots \cup \left(\frac{(2b-1)r}{2a}, \frac{br}{a} \right) \right] \cap \mathbb{Z} \right| \mod 2$$

where $b = \lfloor a/2 \rfloor$ as before.
2) Let $(a, p) = 1$, and let $-p = 4ak + r$ with $0 < r < 4a$. Using Exercises 8.8 and 8.9 prove that

$$\begin{aligned}
\mu(a, p) &\equiv \left| \left[\left(-2 - \frac{r}{2a}, -\frac{r}{a} \right) \cup \left(-2 - \frac{3r}{2a}, -\frac{2r}{a} \right) \cup \cdots \cup \left(-2 - \frac{(2b-1)r}{2a}, -\frac{br}{a} \right) \right] \cap \mathbb{Z} \right| \mod 2 \\
&\equiv \left| \left[\left(\frac{r}{a}, \frac{r}{2a} + 2 \right) \cup \left(\frac{2r}{a}, \frac{3r}{2a} + 2 \right) \cup \cdots \cup \left(\frac{br}{a}, \frac{(2b-1)r}{2a} + 2 \right) \right] \cap \mathbb{Z} \right| \mod 2 \\
&\equiv \left| \left[\left(\frac{r}{2a}, \frac{r}{a} \right) \cup \left(\frac{3r}{2a}, \frac{2r}{a} \right) \cup \cdots \cup \left(\frac{(2b-1)r}{2a}, \frac{br}{a} \right) \right] \cap \mathbb{Z} \right| \mod 2
\end{aligned}$$

where, as before, $b = \lfloor a/2 \rfloor$.

8.5 The Law of Quadratic Reciprocity

The celebrated Law states the following

Theorem 8.5.1 (*Legnedre-Gauss*) *Suppose p and q are two distinct odd prime numbers. We have the following about the Legendre symbols.*
(1) $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$ *and* $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$
(2) $\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$.

Item (1) consists of the supplementary laws of quadratic reciprocity, while item (2), the Law of Quadratic Reciprocity proper, expresses a relation between the solutions of the pair of equations

$$X^2 - p \equiv 0 \mod q \qquad \text{and} \qquad X^2 - q \equiv 0 \mod p.$$

Namely, the Law says that (this is the form of the Law formulated in Gauss' *"Discuisitiones Arithmeticae")*

> **if either p or q is $1 \mod 4$, then the equations have solutions at the same time,**
> *and that*
> **if p and q are $3 \mod 4$, then one equation has a solution exactly when the other one doesn't.**

But besides this theoretical aspect of the Law, it, in combination with the supplementary Laws, provides also a very effective method for computing the Legendre symbol $\left(\frac{p}{q}\right)$ as we will see in the examples below.

Remark 8.5.1 Legendre, together with Euler, Lagrange, and Gauss, was one of the discoverers of the Law of Quadratic Reciprocity. He tried to prove the Law too, and almost succeeded. As a matter of fact, Legendre based his arguments on his theorem about solvability of $aX^2 + bY^2 + cZ^2 = 0$ we discussed earlier in these Notes, and to complete them he needed to use that there are infinitely many prime numbers in an arithmetic progression $\{an + b \mid \gcd(a, b) = 1, n \in \mathbb{N}\}$. Legendre was not able to prove this claim, and formulated it as a conjecture. The conjecture was eventually proved by P.G.L. Dirichlet.

Proof of Theorem 8.5.1 *Only item (2) needs a proof, item (1) being a subject of the previous section. But, according to the Gauss's Criterion,*

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

which, because of Eisenstein's theorem, gives immediately the result. □

Example 8.5.1 Here are some computations illustrating the way the Law of Quadratic Reciprocity is applied.

$$\left(\frac{7}{29}\right) = (-1)^{(7-1)/2\cdot(29-1)/2}\left(\frac{29}{7}\right) = \left(\frac{1}{7}\right) = 1.$$

$$\left(\frac{17}{103}\right) = \left(\frac{103}{17}\right) = \left(\frac{1}{17}\right) = 1.$$

$$\left(\frac{17}{97}\right) = \left(\frac{97}{17}\right) = \left(\frac{12}{17}\right) = (\left(\frac{2}{17}\right))^2 \cdot \left(\frac{3}{17}\right)$$

$$= ((-1)^{(17^2-1)/8})^2\left(\frac{17}{3}\right) = \left(\frac{-1}{3}\right) = (-1)^{(3-1)/2} = -1. \quad \square$$

Exercise 8.11 (*Euler's form of the LQR*) *Let p and q be distinct odd primes, and let $a \in \mathbb{N}$ be such that $p \nmid a$ and $q \nmid a$. Prove, using Exercise 8.10, that if $q \equiv \pm p \mod 4a$, then $\left(\frac{a}{p}\right) = \left(\frac{a}{q}\right)$.*

8.5.1 The General Theorem on Solving $X^2 - a \equiv 0 \mod n$ Revisited

We are restating here an augmented version of the General Theorem using the notations we introduced after we proved that theorem in the end of the first section of this chapter. Before that, denote by A_n the number of quadratic residues modulo n:

$$A_n = |\,(\mathbb{Z}/n\mathbb{Z})^{\times 2}\,|.$$

Recall that $(\mathbb{Z}/n\mathbb{Z})^{\times 2}$ is a subgroup of the group $(\mathbb{Z}/n\mathbb{Z})^{\times}$. As mentioned already, every subgroup H of a group G has an index as such, denoted by $[G : H]$. By definition, when G is a finite set, $[G : H] = |G|/|H|$ is the quotient of the cardinalities of G and H. By a theorem of Lagrange (proved in Algebraic Structures), the index of a subgroup is always an integer. In our case we have

$$[(\mathbb{Z}/n\mathbb{Z})^{\times} : (\mathbb{Z}/n\mathbb{Z})^{\times 2}] = |\,(\mathbb{Z}/n\mathbb{Z})^{\times}\,|/|\,(\mathbb{Z}/n\mathbb{Z})^{\times 2}\,| = \varphi(n)/A_n.$$

We know already that $A_{p^k} = (p^k - p^{k-1})/2 = \varphi(p^k)/2$. In the augmented General Theorem, we are giving a formula for A_n in general.

Theorem 8.5.2 *Let $1 < n \in \mathbb{N}$ have the canonical decomposition $n = 2^{\alpha} p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ where $\alpha \geq 0$ and $\alpha_1, \ldots, \alpha_k \geq 1$. Let $a \in \mathbb{Z}$ be such that $\gcd(a, n) = 1$. Then the equation*

$$X^2 - a \equiv 0 \mod n$$

has solutions if , and only if, the following restrictions are met

$$\left(\frac{a}{p_1}\right) = \cdots = \left(\frac{a}{p_k}\right) = 1 \quad \wedge \quad (\alpha = 2 \to a \equiv 1 \mod 4) \quad \wedge \quad (\alpha \geq 3 \to a \equiv 1 \mod 8).$$

When this is the case, for the number M of solutions to this equation we have
(i) $M = 2^k$ if $\alpha \leq 1$;
(ii) $M = 2^{k+1}$ if $\alpha = 2$;
(iii) $M = 2^{k+2}$ if $\alpha \geq 3$.
In addition to that, we have that, for every n, $A_n \cdot M = \varphi(n)$.

Proof *We have to prove only the addendum to the General Theorem: $A_n \cdot M = \varphi(n)$. To this end, consider that map*

$$\psi : (\mathbb{Z}/n\mathbb{Z})^{\times} \to (\mathbb{Z}/n\mathbb{Z})^{\times 2} \qquad [a] \mapsto \psi([a]) = [a]^2.$$

This map is a surjection with fibres, $(\mathbb{Z}/n\mathbb{Z})^{\times}_{[a]} = \psi^{-1}([a])$, $[a] \in (\mathbb{Z}/n\mathbb{Z})^{\times 2}$, of cardinality M. Indeed we have that $\psi([b_1]) = \psi([b_2]) = [a]$ if, and only if $[b] = [b_1][b_2]^{-1}$ is a solution to $X^2 - 1 \equiv 0 \mod n$. So, the cardinality of $(\mathbb{Z}/n\mathbb{Z})^{\times}_{[a]}$ is the same as the cardinality of $(\mathbb{Z}/n\mathbb{Z})^{\times}_{[1]}$. But $[1]$ is a quadratic residue modulo n, so that the cardinality of the latter fibre is M. Since $(\mathbb{Z}/n\mathbb{Z})^{\times}$ is a disjoint union of the fibres of ψ, then the cardinality of the latter set is a sum of the cardinalities of those fibres, and we have

$$\varphi(n) = \sum_{[a] \in (\mathbb{Z}/n\mathbb{Z})^{\times 2}} |\,(\mathbb{Z}/n\mathbb{Z})^{\times}_{[a]}\,| = |\,(\mathbb{Z}/n\mathbb{Z})^{\times 2}\,| \cdot M = A_n \cdot M$$

as needed. The theorem is proved. □

Remark 8.5.2 The proof of the theorem above is a particular instance of application of the First Isomorphism Theorem in Group Theory (proved in Algebraic Structures). It implies in general that if $\psi : G_1 \to G_2$ is surjective map which is a group homomorphism, these are called **(group) epimorphisms**, and if G_1 is finite, then

$$|G_1| = |G_2| \cdot |ker(\psi)|$$

where $ker(\psi) := \{g \in G_1 \,|\, \psi(g) = e_{G_2}\}$, the **kernel** of ψ, is a subgroup of G_1, and consists of all elements mapped to the identity element of G_2. It is straightforward to check that $\psi : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/n\mathbb{Z})^{\times 2}$ is a homomorphism, so it is an epimorphism, and that $ker(\psi)$ consists of the square roots of identity of $(\mathbb{Z}/n\mathbb{Z})^\times$. So, the formula we obtained in the proof is a consequence of the general algebraic result. □

Exercise 8.12 *(1) Compute* $\left(\frac{503}{773}\right)$ *and* $\left(\frac{501}{773}\right)$. *(All integers involved are prime).*
(2) Prove that $\left(\frac{-7}{p}\right) = 1$ *if, and only if,* $p \equiv 1, 2$ *or* $4 \mod 7$.
(3) Show that the statement in the Law of Quadratic Reciprocity can be written (in the form Gauss did) as

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

(4) Show that 5 *is a quadratic non-residue modulo any prime number of the type* $p = 6^n + 1$.
(5) (Euler's version of the Law of Quadratic Reciprocity) Using Gauss-Legendre LQR prove that if q *and* $p = \pm q + 4a$ *are odd prime numbers with* $a \in \mathbb{N}$, *then*

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

(6) Let a *be an integer such that* $gcd(a, p) = 1$. *Determine all prime numbers* p *such that*

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

(7) Let p *be a prime number of type* $1 \mod 4$. *Prove that*

$$\sum_{j=1}^{(p-1)/2} \left(\frac{j}{p}\right) = 0.$$

(8) Find all primes p *for which the equation has a solution*

$$(i)\ X^2 - 11 \equiv 0 \mod p, \qquad (iii)\ X^2 - 6 \equiv 0 \mod p,$$

$$(ii)\ X^2 - 10 \equiv 0 \mod p \qquad (iv)\ X^2 - 14 \equiv 0 \mod p.$$

(9) Find all prime numbers p *for which both equations*

$$X^2 - 2 \equiv 0 \mod p \qquad \textit{and} \qquad X^2 - 3 \equiv 0 \mod p$$

have solutions.
(10) Prove that the congruence

$$X^2 + 3 \equiv 0 \mod n$$

has no solutions when $n = 7^2 \cdot 19^2 \cdot 23$, *and has eight solutions when* $n = 7^2 \cdot 19^2 \cdot 31$.
(11) Suppose $\gcd(a, b) = 1$ *where* $b \in \mathbb{N}$ *is an odd number, and* $a \in \mathbb{Z}$. *Prove that the equation* $X^2 - a \equiv 0 \mod b$ *has a solution if, and only if, for every prime divisor* p *of* b *we have* $\left(\frac{a}{p}\right) = 1$. *If this is the case, how many solutions does the equation have?*
(12) Let $a, b \in \mathbb{Z}$, *and let* p *be an odd prime number with* $\gcd(a, p) = 1$.

(i) Prove that if a *and* b *are quadratic residues* $\mod p$, *then the equation* $aX^2 - b \mod p$ *does have a solution.*

(ii) What if both a *and* b *are non-residues* $\mod p$*?*

(iii) What if one of a *or* b *is a quadratic residue and the other is a non-residue?*

Give a concrete example of each type, and exhibit a solution when it exists.

(13) Let p be an odd prime number, and $a, b, c \in \mathbb{Z}$ be such that $p \nmid a$. Denote by $D = b^2 - 4ac$ the discriminant of the quadratic polynomial $f(X) = aX^2 + bX + c$. Prove for the equation

$$f(X) \equiv 0 \mod p$$

that

(i) there are no solutions if $p \nmid D$ and $\left(\frac{D}{p}\right) = -1$;

(ii) there is one solution when $p \mid D$;

(iii) there are two solutions when $p \nmid D$ and $\left(\frac{D}{p}\right) = 1$.

8.6 An Example

This short section is devoted to an interesting example related to the application of Modular Arithmetic to solving Diophantine equations. We know that using divisibility modulo appropriate natural numbers one can show that certain Diophantine equations have no interesting solutions (in $\mathbb{Z}$). Recall the example of $X^2 + Y^2 = 3Z^2$ having no non-zero integer solutions.

On the other hand, if a Diophantine equation does have solutions, then, considering the equation modulo **any** *positive integer n, each of those would be solutions modulo n as well. So, the interesting question here is the following.*

> **Suppose a Diophantine equation does have (interesting) solutions modulo** *every* **positive integer n. Is it true that this equation should have (interesting) solutions in integers as well?**

The example below gives negative answer to this question!

Example 8.6.1 The polynomial $f(X) = (X^2 + 1)(X^4 - 4)(X^2 + X + 2)$ does have solutions for every positive modulus, but does not have integer solutions.
Proof Do this as an exercise. Prove that $f(X) \equiv 0 \mod p$ does have solutions for every prime number p, and show that Hensel's lifting lemma allows one to find a solution for every power of a prime number, and therefore, by the CRT, for every positive modulus n. □

Remark 8.6.1 For the fans of equations of more than one variable: it can be shown (T.Nagell) that the Diophantine equation $2X^2 - 219Y^2 + 1 = 0$ has no solutions in $\mathbb{Z}$, but does have solutions in $\mathbb{Z}/n\mathbb{Z}$ for every positive integer n. □

8.7 Vista: Local-to-Global Principle

We reveal in this section that the analytic nature of the p-adic numbers (these are completions of $\mathbb{Q}$ with respect to p-adic norms), and discus a very important fact related to proving the existence of non-trivial integer solutions of quadratic equations of any number of variables through solving them in $\mathbb{R}$ and in $\mathbb{Q}_p$ for every prime number p (the Local-to-Global Principle).

8.7.1 $\mathbb{Q}_p$ as a Completion of $\mathbb{Q}$

The algebraic construction of p-adic numbers we gave in the previous chapter does not reveal an important "analytic" property of these numbers which makes them close "relatives" with the real numbers. As the reader might know (from the course of Introduction to Advanced Mathematics for instance), the real numbers $\mathbb{R}$ are constructed as a space containing the rational numbers $\mathbb{Q}$, and with a metric which extends the metric on the rational numbers in such a way that the rational numbers are a dense subset of the real numbers, and every Cauchy sequence of rational numbers has a limit in that extension. It is a consequence of this that any Cauchy sequence of real numbers has a limit in

$\mathbb{R}$ as well. This last property means that $\mathbb{R}$ is a complete metric space. Since $\mathbb{R}$ is a complete metric space containing $\mathbb{Q}$ as a dense subset, it is called a **completion** *of $\mathbb{Q}$. The metric on $\mathbb{Q}$ which is extended to a metric on $\mathbb{R}$ is the absolute value of the difference of two numbers*

$$d(x,y) = |x-y|.$$

This means that it is actually the absolute value, or the norm, function on $|\circ| : \mathbb{Q} \to \mathbb{Q}_{\geq 0}$ which determines the metric, and then the completion of $\mathbb{Q}$. Norm functions on $\mathbb{Q}$ are by definition functions $\|\circ\|\colon \mathbb{Q} \to \mathbb{Q}$ having the properties

(N_1) $(\forall x \in \mathbb{Q})(\| x \| \geq 0)$, *and* $(\| x \| = 0 \to x = 0)$;
(N_2) $(\forall x, y \in \mathbb{Q})(\| xy \| = \| x \| \| y \|$;
(N_3) $(\forall x, y \in \mathbb{Q})(\| x + y \| \leq \| x \| + \| y \|$.

Given a norm $\| \circ \|$ on $\mathbb{Q}$, one can define $d : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ setting $d(x,y) = \| x - y \|$. It is straightforward to check that the newly defined function has the properties

(M_1) $(\forall x, y \in \mathbb{Q})(d(x,y) \geq 0)$, *and* $d(x,y) = 0$ *only if* $x = y$;
(M_2) $(\forall x, y \in \mathbb{Q})(d(x,y) = d(y,x))$;
(M_3) $(\forall x, y, z \in \mathbb{Q})(d(x,z) \leq d(x,y) + d(x,z))$.

A function d satisfying the last three properties is called a **metric** *on $\mathbb{Q}$. When the metric is defined using a norm, then we say that it is* **the associated (with the norm) metric**. *For instance, the absolute value function $|\circ|$ is a norm on $\mathbb{Q}$, denoted by $\|\circ\|_\infty$, and called* **the infinity norm on $\mathbb{Q}$** *. It's associated metric is the classical one on $\mathbb{Q}$ we use in Calculus.*

Given a metric d on $\mathbb{Q}$, not necessarily associated with a norm, one defines **Cauchy sequences** *of rational numbers w.r.t. the metric as follows: the sequence $< x_n > \subseteq \mathbb{Q}$ is Cauchy if*

$$(\forall N \in \mathbb{N} \setminus \{0\})(\exists n_0)(\forall m, n)(m, n \geq n_0 \to d(x_n, x_m) < 1/N).$$

As mentioned above, the real numbers $\mathbb{R}$ form a field with a metric on it $\tilde{d} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. This metric has the properties $(M_1) - (M_3)$ above with $\mathbb{Q}$ replaced with $\mathbb{R}$, it is an extension of the metric d on $\mathbb{Q}$ in the sense that, for $x, y \in \mathbb{Q} \subseteq \mathbb{R}$ we have

$$\tilde{d}(x,y) = d(x,y),$$

it has $\mathbb{Q}$ as a dense subset in the sense that for every $\epsilon > 0$, and for every $x \in \mathbb{R}$, there is a $y \in \mathbb{Q}$ such that $\tilde{d}(x,y) < \epsilon$, and is such that every Cauchy sequence in $\mathbb{R}$ has a limit (in $\mathbb{R}$). This last property is abbreviate to saying that $(\mathbb{R}, \tilde{d})$ is a complete metric space. All the properties of $(\mathbb{R}, \tilde{d})$ with respect to $\mathbb{Q}$ above are abbreviated to saying that $(\mathbb{R}, \tilde{d})$ is a **completion** *of $(\mathbb{Q}, d)$.*

The real numbers are instrumental in defining what a metric on **any** *set, not only on $\mathbb{Q}$ and $\mathbb{R}$, is. Thus we have that $(X, d : X \times X \to \mathbb{R})$ is a metric space if*

(M_1) $(\forall x, y \in X)(d(x,y) \geq 0)$, *and* $d(x,y) = 0$ *only if* $x = y$;
(M_2) $(\forall x, y \in X)(d(x,y) = d(y,x))$;
(M_3) $(\forall x, y, z \in X)(d(x,z) \leq d(x,y) + d(x,z))$.

We know that $\mathbb{R}$ is a completion of $\mathbb{Q}$. But $\mathbb{Q}$ can have many completions, each depending on a metric on $\mathbb{Q}$. More precisely, the metric space $(K, \tilde{d'})$ is a completion of $(\mathbb{Q}, d')$ if K is a field which is a complete metric space containing $\mathbb{Q}$ as a dense subset, and the metric $\tilde{d'}$ is an extension of the metric d'.

Turns out that for every prime number p, there is a norm $\| \circ \|_p$ on $\mathbb{Q}$, called the p-adic norm on $\mathbb{Q}$. Every such norm, through its associated (p-adic) metric, defines a completion of $\mathbb{Q}$ which is exactly the field of p-adic numbers. This subsection is devoted to explaining how this is done.

p-adic Valuations on $\mathbb{Q}$

*We have an interpretation of the Fundamental Theorem of the Arithmetic of $\mathbb{Z}$ through the group of non-zero rational numbers $\mathbb{Q}^\times$. Namely, every $x \in \mathbb{Q}^\times$ has a unique expression in the form (***canonical expression of** *x)*

$$x = \pm \prod_{k\in\mathbb{N}} p_k^{\alpha_k}$$

where $p_0 = 2, p_1 = 3, \ldots, p_k, \ldots$ is the sequence of prime numbers, and only finitely many α_k are non-zero.

For a fixed prime number $p = p_k$ define the function

$$v_p : \mathbb{Q} \to \mathbb{Z} \cup \{\infty\}$$

by $v_p(x) = \alpha_k$ when $x = \pm\prod_{k\in\mathbb{N}} p_k^{\alpha_k}$, and $v(0) = \infty$. This function has the properties that, for every $x, y \in \mathbb{Q}$,

(V_1) $v_p(xy) = v_p(x) + v_p(y)$;
(V_2) $v_p(x + y) \geq min\{v_p(x), v_p(y)\}$;
(V_3) if $v_p(x) \neq v_p(y)$, then $v_p(x + y) = min\{v_p(x), v_p(y)\}$

where, by definition, for any $n \in \mathbb{Z}$ we have

$$\infty + n = n + \infty = \infty, \quad \infty + \infty = \infty, \quad \infty \geq n, \quad \infty \geq \infty.$$

Exercise 8.13 *Verify that the properties $(V_1), (V_2)$ and (V_3) are satisfied by the function v_p.*

Exercise 8.14 *Prove that if $x \neq 0$, then $|x| = \prod_p p^{v_p(x)}$. Conclude that the rational number x is an integer if, and only if, for every p we have $v_p(x) \geq 0$, and that a rational number is a non-zero integer if, and only if, for every prime number p, we have $v_p(x) \in \mathbb{N}$.*

We abbreviate the fact that $(V_1) - (V_3)$ are satisfied by v_p by saying that v_p is a **discrete valuation** *on $\mathbb{Q}$. The valuation v_p is referred to also as p-***adic valuation on** *$\mathbb{Q}$.*

p-adic Norms on $\mathbb{Q}$

Let p be a prime number. The p-adic norm on $\mathbb{Q}$

$$\| \circ \|_p \colon \mathbb{Q} \to \mathbb{Q}$$

is defined by $\| x \|_p = p^{-v_p(x)}$. In other words, if $x = 0$, then $\| 0 \|_p = 0$, and if $x = \pm\prod_p p^{v_p(x)}$, then $\| x \|_p = p^{-v_p(x)}$.

Exercise 8.15 *Prove that $\| \circ \|_p$ satisfies the properties $(N_1) - (N_3)$, that is - it is a norm indeed. As matter of fact, this norm satisfies a stronger property than (N_3): $\| x+y \|_p \leq max\{\| x \|_p, \| y \|_p\}$ with equality if $\| x \|_p \neq \| y \|_p$. We will denote this property by (N_3').*

*The norm $\| \circ \|_p$ is referred to as the p-***adic norm on** *$\mathbb{Q}$.*

Exercise 8.16 *Prove that*

$$\mathbb{Z} = \{x \in \mathbb{Q} \,|\, (\forall p)(\| x \|_p \leq 1)\}.$$

Prove also that $\{-1, 1\} = \{x \in \mathbb{Q} \,|\, (\forall p)(\| x \|_p = 1\}$.

The fact that integers have p-adic norms bounded by 1 is very different from what we have about the norm $\| \circ \|_\infty$. More precisely, the rational numbers with the latter norm have the property that for any $x \in \mathbb{Q}$, there is a $n \in \mathbb{Z}$ such that $\| x \|_\infty < \| n \|_\infty$ while with the p-norms, this property is not satisfied (check this out!). Since Archimedes was one of the first mathematicians who used this property in their work, a norm having it is called an **Archimedean norm**. *The rest are called* **non-Archimedean norms**. *As a matter of fact, the norm boundedness of the integers by 1 is equivalent to the stronger property (N_3').*

Proposition 8.7.1 *Suppose $\| \circ \|$ is a norm on $\mathbb{Q}$. Then the norms of the integers are bounded by 1 if, and only if, the norm satisfies (N_3').*

Proof *The "if" direction of the statement is straightforward. For the "only if" one, consider the binomial formula, for any natural number n and any two rational numbers x and y,*

$$(x+y)^n = \binom{n}{0}x^n + \cdots + \binom{n}{k}x^{n-k}y^k + \cdots + \binom{n}{n}y^n.$$

Taking the norm from both sides, and using the boundedness of the norms of the integers by 1, we get

$$\| x+y \|^n \leq \| x \|^n + \cdots + \| x \|^{n-k} \| y \|^k + \cdots + \| y \|^n \leq (n+1)\max\{\| x \|^n, \| y \|^n\}.$$

Therefore, taking n-th root from the left most and the right most sides of these inequalities, we get

$$\| x+y \| \leq \sqrt[n]{n+1}\max\{\| x \|, \| y \|\}.$$

Letting n go to infinity, we see that the norm satisfies the property (N_3'). □

Exercise 8.17 *Make sense of the following expression: the rational number x is "small" in the p-adic norm if an, and only if, it is divisible by a large power of p.*

There is a very important relation between the different norms on $\mathbb{Q}$. It is revealed in the following exercise.

Exercise 8.18 *Prove that, for every non-zero $x \in \mathbb{Q}$, there are only finitely many prime numbers p such that $\| x \|_p \neq 1$. Prove also that we have*

$$\left(\prod_p \| x \|_p\right) \cdot \| x \|_\infty = 1.$$

p-adic Metric on $\mathbb{Q}$

The p-adic metric d_p on $\mathbb{Q}$ is the one associated with the p-adic norm $\| \circ \|_p$.

$\mathbb{Q}_p$ as a Completion of $\mathbb{Q}$

Having defined the p-adic metric d_p on $\mathbb{Q}$, we can find the completion of $\mathbb{Q}$ with respect to this metric. It is a general fact that the completions always exist and are **essentially unique**. *To explain the last, recall that we wanted originally to have $\mathbb{Q}$ as a dense subset of its completion K such that the metric d_p' on K extends the metric on $\mathbb{Q}$. What this means is actually that there is a homomorphism of fields $\varphi_K : \mathbb{Q} \to K$ which is an isometry (for every $x, y \in \mathbb{Q}$ we have $d_p'(\varphi_K(x), \varphi_K(y)) = d_p(x,y)$), and such that $Ran(\varphi_K) \subseteq K$ is a dense subset. Well, the completions are not unique as fields, but they are identifiable in a nice way. More precisely, if (L, d_p'') is another completion of $(\mathbb{Q}, d_p)$, there is a unique map $\psi : K \to L$ such that $\varphi_L = \varphi_K \circ \psi$. Moreover, this map is an isomorphism of fields, which is an isometry thereof. This is what essential uniqueness is.*

Now, we know from the above that the completion of $(\mathbb{Q}, d_p)$ exists and is essentially unique. This means that knowing a completion, we know all of them. We are ending this subsection by proving that $\mathbb{Q}_p$ with an appropriate metric on it, is a completion of $(\mathbb{Q}, d_p)$.

Recall from subsections 7.1.1 and 7.1.2 that the non-zero p-adic numbers have the form

$$\alpha = \varphi_p(p)^n \cdot x$$

where $n \in \mathbb{Z}$, and x is an invertible p-adic integer, that is,

$$x = (a_1, a_2, \dots, a_k, \dots), \quad a_i \in \mathbb{Z},\ a_i \equiv a_{i+1} \mod p^i \quad (i = 1, 2, \dots), \qquad [a_1]_p \neq [0]_p.$$

Here $\varphi_p : \mathbb{Z} \to \mathbb{Z}_p$ defined by $\varphi_p(a) = (a, a, \dots, a, \dots)$.

We first notice that the p-adic norm on $\mathbb{Q}$ is extended to a norm on $\mathbb{Q}_p$ as well.

Exercise 8.19 *For $0 \neq \alpha \in \mathbb{Q}_p$, define $\| \alpha \|_p = 1/p^n$ where $\alpha = \varphi_p(p)^n \cdot x$ with $x \in \mathbb{Z}_p^\times$, and set $\| 0 \|_p = 0$. Prove that $\| \circ \|_p : \mathbb{Q}_p \to \mathbb{R}$ is a non-Archimedean norm which extends the one on $\mathbb{Q}$. The latter means that $\| \varphi_p(y) \|_p = \| x \|_p$ for every $y \in \mathbb{Q}$.*

We are proving next that $\varphi_p(\mathbb{Q}) \subseteq \mathbb{Q}_p$ is a dense subset with respect to this norm. We have to show that for every $\alpha \in \mathbb{Q}$, and for every $\epsilon > 0$, there is a $y \in \mathbb{Q}$ such that $\| \alpha - \varphi_p(y) \|_p < \epsilon$. Since $\varphi_p(0) = 0$, we may assume that $\alpha \neq 0$, and therefore that $\alpha = \varphi_p(p)^n \cdot (a_1, a_2, \dots, a_k, \dots)$ as above. Notice that

$$\varphi_p(p^n \cdot a_k) = \varphi_p(p)^n \cdot (a_1, a_2, \dots, a_{k-1}, a_k, \dots, a_k, \dots)$$

and that

$$\alpha - \varphi_p(p^n \cdot a_k) = \varphi_p(p)^{n+k}(a_{k+1} - a_k, \dots, a_{k+s} - a_k, \dots).$$

Therefore,

$$\| \alpha - \varphi_p(p^n \cdot a_k) \|_p \leq \frac{1}{p^{n+k}} < \epsilon$$

for $k \in \mathbb{N}$ big enough. The density property is established.

Finally, we have to show that $\mathbb{Q}_p$ with this norm is a complete metric space: every Cauchy sequence in $\mathbb{Q}_p$ is convergent in $\mathbb{Q}_p$.

Let $< \alpha_k >$ be a Cauchy sequence of p-adic numbers.

Exercise 8.20 *Prove that if $< \alpha_k >$ has infinitely many zero elements, then $< \alpha_k > \to 0$.*

After this exercises, we may assume that the sequence has finitely many zero elements. By a standard argument know from Calculus, the convergence of a sequence doesn't depend on any finite number of elements of that sequence. So, W.L.O.G., we may assume that $\alpha_k \neq 0$ for every $k \in \mathbb{N}$. As we know, in such a case, for every $k \in \mathbb{N}$, we have

$$\alpha_k = \varphi_p(p)^{n_k} \cdot x_k, \qquad n_k \in \mathbb{Z}, \quad x_k \in \mathbb{Z}_p^\times.$$

Exercise 8.21 *Assuming that $< \alpha_k >$ is Cauchy sequence of non-zero p-adic numbers, prove that the sequence of exponents $< n_k >$ is bounded below. That is, there is an integer m such that $m \leq n_k$ for every $k \in \mathbb{N}$.*

This integer m will help us reduce our task to working with a sequence of p-adic integers.

Exercise 8.22 *In the setting of the previous exercise denote $\beta_k = \varphi_p(p)^{-m}\alpha_k$. Prove that the sequence $< \beta_k >$ is a Cauchy sequence of p-adic integers.*

Again a standard argument from Calculus shows that $< \beta_l >$ is convergent if, and only if $< \alpha_k >$ is, and that in case of convergence we have

$$\lim_k \alpha_k = \varphi_p(p)^m \cdot \lim_k \beta_k.$$

So, it remains to prove that a Cauchy sequence of p-adic integers is convergent (to a p-adic integer). We will prove this by using a version of the diagonal method invented by Cantor, and widely used in Mathematics, but before that - the following straightforward fact.

Exercise 8.23 *Suppose $x = (a_1, \dots, a_k, \dots)$ and $x' = (a'_1, \dots, a'_k, \dots)$ are p-adic integers such that $\| x - x' \|_p < 1/p^s$ for $s \in \mathbb{N}$. Prove that $a_s \equiv a'_s \mod p^s$.*

Let now $< \alpha_k >$ be a Cauchy sequence of p-adic integers. We know that, for every $k \in \mathbb{N}$ there is an index m_k such that $\| \alpha_s - \alpha_t \|_p < 1/p^k$ as long as $s, t \geq m_k$. Without a loss of generality, we may assume that $< m_k >$ is a strictly increasing sequence. For every $k \in \mathbb{N}$ we have

$$\alpha_k = (a_{1k}, a_{2k}, \dots, a_{sk}, \dots).$$

Consider the map $f : \mathbb{N} \to \mathbb{Z}$ given by $f(t) = m_k$ if $m_{k-1} \leq t < m_k$, where $m_{-1} := 0$. The following exercises exhibit a limit of our sequence.

Exercise 8.24 *(1) Prove that the function f constructed above is a p-adic number. That is, show that $(f(0), f(1), \dots, f(k), \dots) \in \mathbb{Z}_p$. Call this number α.*
(2) Prove that α is a limit of $< \alpha_k >$.

8.7.2 Ostrowski's Theorem

We know that given a metric on $\mathbb{Q}$, the completion of $\mathbb{Q}$ with respect to this metric is (essentially) unique. But different metrics may define the same completions on $\mathbb{Q}$.

Exercise 8.25 *Prove that the completions of $\mathbb{Q}$*

$$\{\mathbb{R}, \mathbb{Q}_p \,|\, p \text{ prime number}\}$$

are all pairwise non-isomorphic. [Hint: For any two completions find a sequence of rational numbers which is convergent in one completion, and divergent in the other.]

A concept that allows us to compare norms is the concept of **equivalent norms**. *Given norms $\| \circ \|$ and $\| \circ \|'$ on $\mathbb{Q}$, we say that they are equivalent if there exist two positive (rational) numbers α and β such that, for every $x \in \mathbb{Q}$ we have*

$$\alpha \| x \| \leq \| x \|' \leq \beta \| x \| .$$

Exercise 8.26 *Prove that equivalent norms on $\mathbb{Q}$ define isomorphic completions of $\mathbb{Q}$.*

A remarkable theorem reveals that, up to equivalence of norms, we already know all norms on $\mathbb{Q}$.

Theorem 8.7.2 (*A. Ostrowski, 1916*) *Let $\| \circ \|$ be a norm on $\mathbb{Q}$. If this norm is Archimedean, then it is equivalent to $\| \circ \|_\infty$. If the norm is non-Archimedean, then there is a (unique) prime number p such that the norm is equivalent to $\| \circ \|_p$.*

In other words, all completions of $\mathbb{Q}$ with respect to a metric coming from a norm are the fields of real numbers $\mathbb{R}$, and the fields of p-adic numbers $\mathbb{Q}_p$.

8.7.3 Hasse-Minkowski's Principle

After explaining the analytic nature of the algebraically constructed p-adic numbers, a Number Theory application is in order.

Suppose that $F(x_1, \dots, x_k) = 0$ is a Diophantine equation that we want to solve. The first question to ask here is whether this equation has solutions at all. We know that non-existence can be proved by using modular methods (considering the equation modulo n for every $n \in \mathbb{N}$) The Chinese Remainder Theorem tells us that it is actually enough to work modulo the powers of prime numbers: $n = p^m$. Obviously, if there is no solution to the equation for some power of a prime, then it has no solutions in $\mathbb{Z}$ either. Similarly, if $F(x_1, \dots, x_k) = 0$ has no solutions in $\mathbb{R}$, then it has no integer solutions either.

One can restate the results of this discussion as follows:

If $F(x_1, \ldots, x_k) = 0$ has solutions in $\mathbb{Z}$, then it has solutions in all completions $\mathbb{R}$ and $\mathbb{Q}_p$ for p a prime number.

The natural question to ask here is whether the converse is true as well. The example we discussed in the previous section shows that this is not true in general. Indeed,

$$f(X) = (X^2 + 1)(X^4 - 4)(X^2 + X + 2) = 0$$

has solutions in every $\mathbb{Q}_p$, and in $\mathbb{R}$, but has no integer solutions. The following deep theorem gives a positive answer in an important case.

Theorem 8.7.3 (*Hasse-Minkowski's Principle*) *Let $F(x_1, \ldots, x_n) = \sum_{1 \le i \le j \le n} a_{ij} x_i x_j$ be a non-trivial homogeneous polynomial of degree 2 with integer coefficients. Then $F = 0$ has non-zero solution in $\mathbb{Q}$ if, and only if, it has non-zero solutions in $\mathbb{R}$ and in $\mathbb{Q}_p$ for every prime number p.*

Observe that, since F is homogeneous, existence of solutions in $\mathbb{Z}$ is equivalent to having solutions in $\mathbb{Q}$. Another important (Linear Algebra) fact is that with a linear change of the variables one can express any quadratic form in "diagonal form"

$$F(x_1, \ldots, x_n) = \sum_{1 \le i < j \le n} a_{ij} x_i x_j$$

$$= a_1 y_1^2 + \cdots + a_k y_k^2 =: G(y_1, \ldots, y_n)$$

where $k \le n$, $\quad a_1, \ldots, a_k \in \mathbb{Q}$, and $a_1 a_2 \cdots a_k \neq 0$.
Solving $F(x_1, \ldots, x_n) = 0$ is equivalent to solving $G(y_1, \ldots, y_n) = 0$, and non-trivial solutions of the former correspond to non-trivial solutions of the latter.
Obviously if $k < n$, there is a non-trivial solution of $G(a_1, \ldots, y_n) = 0$, and the Hasse-Minkowski's Principle is trivially true. So, the interesting case is when $k = n$, that is when the quadratic form is called **non-degenerate**. *We will prove now the theorem when the form is non-degenerate, and $n = 3$. This case is familiar to us from previous Vistas: the Legendre theorem 2.3.1 gives necessary and sufficient conditions for solvability of the form of a different type.*

Proof of Hasse-Minkowski's Principle for $F(x, y, z) = ax^2 + by^2 + cy^2$ with $a, b, c \in \mathbb{Q}^\times$.

We are proving that solubility of $F = 0$ in all completions of $\mathbb{Q}$ implies its solubility in $\mathbb{Q}$ the other direction of the claim being obvious. Since $F = 0$ is solvable in $\mathbb{R}$, not all coefficients of F are of the same sign. W.L.O.G. we may assume that $F(x, y, z) = ax^2 + by^2 - z^2$ where either $a > 0$ or $b > 0$. Standard arguments reduce the equation to one with **integer** *coefficients a and b with ab a square-free number. We will prove the theorem by induction on $n = |a| + |b| \ge 2$.*
If $n = 2$, then $F(x, y, z) = \pm x^2 \pm y^2 - z^2$ with not all coefficients negative numbers, and has obvious solution. Assume the theorem is true for all $n \le k$, and consider the case $n = k + 1$. W.L.O.G. we may assume $|a| \le |b|$. Since $n = k + 1 > 2$, we have that $|b| = p_1 \cdots p_s$ is a product of s distinct primes. Since $ax^2 + by^2 - z^2$ is soluble in every $\mathbb{Q}_{p_i}$ for $i = 1, \ldots, s$, we have that, for every $i = 1, \ldots, s$ the congruence $ax^2 \equiv z^2 \mod p_i$ has a solution. Since we may assume that $\|x\|_{p_i} = \|y\|_{p_i} = \|z\|_{p_i} = 1$ for the solution in $\mathbb{Q}_{p_i}$, we conclude that a is a quadratic residue $\mod p_i$ for every $i = 1, \ldots, s$. By the CRT then we get that a is a quadratic residue $\mod |b|$ as well, and we have

$$r^2 - a = bc, \qquad 0 \le r \le \frac{|b|}{2}, \quad c \in \mathbb{Z}.$$

Using the well known from the proof of Legendre's theorem 2.3.1 trick, we have the identity

$$a(rx + z)^2 + b(cy)^2 - (ax + rz)^2 = bc(ax^2 + cy^2 - z^2).$$

This together with the inequalities

$$|c| = |\frac{r^2-1}{b}| \leq |\frac{r^2}{b}| + |\frac{a}{b}| \leq \frac{|b|}{4} + 1 < |b|,$$

so that $|a|+|c| \leq k$, allows us to use induction, and get prove that the case of $n = k+1$ is also true. □

This theorem is an example of a "Local-to-Global Principle". Solving an equation in $\mathbb{Q}$ is considered a global event, while solving it in $\mathbb{R}$ and $\mathbb{Q}_p$ is considered a local event. The theorem says that if we can solve the equation locally, then we can solve it globally.
Of course, if a Diophantine equation, homogeneous or not, is solvable in $\mathbb{Z}$, it is also solvable in any of the completions of $\mathbb{Q}$. This is an obvious fact which can be interpreted as a Global-to-Local Principle. It is always valid. Contrary to that, the Local-to-Global Principle is not valid if the degree of F is bigger than 2 even if the equation remains homogeneous!

8.7.4 The Groups $\mathbb{Q}_p^{\times 2}$

Let $\alpha \in \mathbb{Q}_p^{\times}$. We want to characterize those α which are in $\mathbb{Q}_p^{\times 2}$. In other words, we want to find the necessary and sufficient conditions on α under which the equation

$$x^2 = \alpha$$

is solvable in $\mathbb{Q}_p^{\times}$. We know that $\alpha = \varphi_p(p^s)a$ where $a \in \mathbb{Z}_p^{\times}$. Obviously, if $\alpha = \beta^2$ for some $\beta \in \mathbb{Q}_p$, then s is an even number: $s = 2s_1$. Now solving

$$x^2 = \varphi_p(p^{2s_1})a$$

in $\mathbb{Q}_p$ is equivalent to solving

$$x^2 = a$$

in $\mathbb{Z}_p$.

Case of an Odd Prime p

Let $a = (a_1, a_2, \dots, a_n, \dots) \in \mathbb{Z}_p^{\times}$. The equation $x^2 = a$ is solvable only if, and only if, the equations $y^2 \equiv a_n \mod p^n$ are solvable for every $n \geq 1$. By the Hensel's lifting lemma, these equations either are solvable, and have two solutions each, precisely when $y^2 \equiv a_1 \mod p$ is solvable. Moreover the solutions of the equations form two elements of $\mathbb{Z}_p$ x_1 and $x_2 = -x_1$ which are the (two) solutions to the original equation. We conclude fro this discussion that

Proposition 8.7.4 *Let p be an odd prime. Then $\mathbb{Q}_p^{\times 2} = \{\varphi_p(p^n)a \,|\, n \in 2\mathbb{Z}$ and $a \in \mathbb{Z}_p^{\times 2}\}$ where*

$$\mathbb{Z}_p^{\times 2} = \{a = (a_1, \dots) \in \mathbb{Z}_p \,|\, \left(\frac{a_1}{p}\right) = 1\}.$$

Case of $p = 2$

Let now $a = (a_1, a_2, \dots, a_n, \dots) \in \mathbb{Z}_2^{\times}$. This means in particular that all components of a are odd integers. As before, solving $x^2 = a$ in $\mathbb{Z}_2^{\times}$ is equivalent to solving the equations $y^2 =\equiv a_n \mod 2^n$ for all $n \geq 1$. As we know, the latter is possible if, and only if, $a_n \equiv 1 \mod 8$ for $n \geq 3$ which reduces to just $a_1 \equiv 1 \mod 8$. In this case, the original equation $x^2 = a$ has two solutions. We have proven the following

Proposition 8.7.5 *We have $\mathbb{Q}_2^{\times 2} = \{\varphi_2(2^n)a \,|\, n \in 2\mathbb{Z},$ and $a \in \mathbb{Z}_2^{\times 2}\}$ where*

$$\mathbb{Z}_2^{\times 2} = \{a = (a_1, \dots) \in \mathbb{Z}_2 \,|\, a_1 \equiv 1 \mod 8\}.$$

Exercise 8.27 *Examine the solutions of $y^2 \equiv a_n \mod 2^n$ for $a = (a_1, a_2, \dots, a_n, \dots) \in \mathbb{Z}_2^{\times}$, and figure out how the solutions $\mod 2^n$ for every n organize in* **two** *solutions in $\mathbb{Q}_2$.*

8.8 The Generalized Law of Quadratic Reciprocity

In order to use the Law of Quadratic Reciprocity for calculating the Legendre symbol $\left(\frac{a}{p}\right)$, we have to factor a out as a product of primes. This, as already mentioned, is a hard to do in general. Turns out, a fact noticed by Jacobi, there is no need for such a factorization!

Definition 8.8.1 *Let $Q = p_1 p_2 \cdots p_k$ be an odd number with its representation as product of primes $(2 < p_1 \leq p_2 \leq \cdots \leq p_k)$. If a is an integer, relatively prime with Q, define the* **Jacobi symbol of a modulo Q** *to be*

$$\left(\frac{a}{Q}\right) := \left(\frac{a}{p_1}\right) \cdot \left(\frac{a}{p_2}\right) \cdots \left(\frac{a}{p_k}\right).$$

Exercise 8.28 *Show that, for every odd $Q \in \mathbb{N}$ and for every $a, b \in \mathbb{Z}$, we have*

$$\left(\frac{ab}{Q}\right) = \left(\frac{a}{Q}\right) \cdot \left(\frac{b}{Q}\right) \qquad \text{and} \qquad (a \equiv b \mod Q) \ \rightarrow \ \left(\left(\frac{a}{Q}\right) = \left(\frac{b}{Q}\right)\right).$$

The amazing fact is that the Jacobi symbol satisfies similar Law of Quadratic Reciprocity as Legendre symbol does!

Theorem 8.8.1 (**Generalized Law of Quadratic Reciprocity**) *Suppose P and Q are relatively prime odd natural numbers. The following formulae hold true for the Jacobi symbol.*
(1) $\left(\frac{-1}{P}\right) = (-1)^{\frac{P-1}{2}}$ *and* $\left(\frac{2}{P}\right) = (-1)^{\frac{P^2-1}{8}}$
(2) $\left(\frac{Q}{P}\right) \cdot \left(\frac{P}{Q}\right) = (-1)^{\frac{P-1}{2} \cdot \frac{Q-1}{2}}$.

Proof *Let $P = p_1 p_2 \cdots p_k$ and $Q = q_1 q_2 \cdots q_l$ where $p_1 \leq p_2 \leq \cdots \leq p_k$ and $q_1 \leq q_2 \leq \cdots \leq q_l$ are prime numbers. By assumption $\{p_1, p_2, \ldots, p_k\} \cap \{q_1, q_2, \ldots, q_l\} = \emptyset$. Obviously, proving the formulae in (1) is equivalent to showing that*

$$\frac{P-1}{2} \equiv \sum_{i=1}^{k} \frac{p_i - 1}{2} \mod 2 \qquad \text{and} \qquad \frac{P^2-1}{8} \equiv \sum_{i=1}^{k} \frac{p_i^2 - 1}{8} \mod 2,$$

while proving the formula in (2), we have to show that

$$\frac{P-1}{2} \cdot \frac{Q-1}{2} \equiv \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{p_i - 1}{2} \cdot \frac{q_j - 1}{2} \mod 2.$$

We are proving the first equality above. The number P is odd, so $p_i = 2n_i + 1$ for every $i = 1, 2, \ldots, k$. Then

$$LHS_1 = \frac{1}{2} \cdot \left(\prod_{i=1}^{k} (2n_i + 1) - 1\right) = \frac{4 \cdot A + 2\sum_{i=1}^{k} n_i}{2} \equiv \sum_{i=1}^{k} n_i \mod 2,$$

and

$$RHS_1 = \sum_{i=1}^{k} \frac{2n_i + 1 - 1}{2} = \sum_{i=1}^{k} n_i.$$

Obviously $LHS_1 \equiv RHS_1 \mod 2$. The first equality is proved.
For the second equality, we have

$$LHS_2 = \frac{P^2-1}{8} = \frac{\prod_{i=1}^{k}(2n_i+1)^2 - 1}{8} = \frac{\prod_{i=1}^{k}(4(n_i^2+n_i)+1) - 1}{8}$$

$$= \frac{16 \cdot B + 4\sum_{i=1}^{k}(n_i^2 + n_i)}{8} \equiv \frac{\sum_{i=1}^{k}(n_i^2+n_i)}{2} \mod 2,$$

and

$$RHS_2 = \sum_{i=1}^{k} \frac{(2n_i+1)^2-1}{8} = \sum_{i=1}^{k} \frac{n_i^2+n_i}{2}.$$

Obviously $LHS_2 \equiv RHS_2 \mod 2$, *and the second equality is proven too.*
For the third equality, since Q *is odd, and therefore* $q_j = 2m_j + 1$ *for every* $j = 1, 2, \ldots l$, *we have*

$$LHS_3 = \frac{\prod_{i=1}^{k}(2n_i+1)-1}{2} \cdot \frac{\prod_{j=1}^{l}(2m_j+1)-1}{2}$$

$$= \frac{4 \cdot C + 2\sum_{i=1}^{k} n_i}{2} \cdot \frac{4 \cdot D + 2\sum_{j=1}^{l} m_j}{2} \equiv \left(\sum_{i=1}^{k} n_i\right) \cdot \left(\sum_{j=1}^{l} m_j\right) \mod 2,$$

and

$$RHS_3 = \sum_{i=1}^{k}\sum_{j=1}^{l}(n_i)\cdot(m_j) = \left(\sum_{i=1}^{k} n_i\right) \cdot \left(\sum_{j=1}^{l} m_j\right).$$

We have once more time $LHS_3 \equiv RHS_3 \mod 2$, *and the third equality is proved.*
The theorem is proved as well. □

The practical value of the Jacobi symbol for this course is in allowing one to compute Legendre symbols without needing to have the arguments of the symbol prime numbers! On the other hand, knowing that Legendre symbol is 1 if, and only if, the corresponding congruence equation has a solution, one may be tempted to use the Jacobi symbol as a criterion for having solutions to the congruence

$$X^2 \equiv Q \mod P$$

where P *is an odd natural number, and* $gcd(P, Q) = 1$. *One has to be more careful here! The exercise below explains why.*

Of course, it is not the computation of the Legendre symbol which the Jacobi symbol is used **only** *for! As a matter of fact, there are many symbols (cubic, biquadratic, quintic, and so on) who bear the names of their inventors (Kronecker, Dirichlet, Eisenstein, Furtwängler, Hilbert, Artin, Hasse, ... just to mention some!) which were designed to capture in bigger generality the relationship (reciprocity in a general sense) between numbers detected in its easiest form by the Law of Quadratic Reciprocity. A bit more about all this can be found in the following Vista. In Chapter 11, we will see interesting (and at a more elementary level) applications of the Jacobi symbol in answering non-trivial number theoretical questions.*

Exercise 8.29 *Let* P *and* Q *be integers such that* P *is odd, and* $\gcd(P, Q) = 1$. *Consider the congruence equation* $X^2 \equiv Q \mod P$. *Prove that*
(1) If the equation has a solution, then $\left(\frac{Q}{P}\right) = 1$;
(2) Show, by an example, that the converse of (1) is not true in general.

The following exercises offer equivalent formulations of the Law of Quadratic Reciprocity

Exercise 8.30 *(1) (Eisenstein's form of the Generalized Law of Quadratic Reciprocity) Prove that, for positive odd integers* P, Q, P', Q' *such that* $\gcd(P, Q) = 1$ *and* $\gcd(P', Q') = 1$, *if* $P \equiv P' \mod 4$ *and* $Q \equiv Q' \mod 4$, *then*

$$\left(\frac{P}{Q}\right)\left(\frac{Q}{P}\right) = \left(\frac{P'}{Q'}\right)\left(\frac{Q'}{P'}\right).$$

(2) Conversely, prove that Eisenstein's form of the Generalized Law of Quadratic Reciprocity from (1) is true, then the Gauss-Legendre form of that Law is true as well.

Exercise 8.31 *(1) (Euler's version of the Law of Quadratic Reciprocity) Prove that, for a fixed integer a, the Legendre symbol $\left(\frac{a}{p}\right)$ depends only on $[\pm p]_{4a}$.*

[We have to show that if $q \equiv \pm p \mod 4a$, then $\left(\frac{a}{p}\right) = \left(\frac{a}{q}\right)$. If $q = p + 4ak$, then $\left(\frac{a}{p}\right) = \left(\frac{4ak}{p}\right)\left(\frac{k}{p}\right) = \left(\frac{q}{p}\right)\left(\frac{k}{p}\right)$, and similarly that $\left(\frac{a}{q}\right) = \left(\frac{-p}{q}\right)\left(\frac{k}{q}\right)$. Using the properties of the Jacobi symbol, and the GLQR, we reduce to the case k odd or $k = 2k_1$ with k_s odd, and compute that $\left(\frac{q}{p}\right)\left(\frac{k}{p}\right) = \left(\frac{-p}{q}\right)\left(\frac{k}{q}\right)$.]

(2) Conversely, prove that if the Legendre symbol has the property of (1), then it satisfies the Law of Quadratic Reciprocity.

[Let $p < q$ be two odd primes. Then either $p + q \equiv 0 \mod 4$ or $p - q \equiv 0 \mod 4$. In the former case, let $d = (p+q)/4$. We have $\left(\frac{p}{q}\right) = \left(\frac{p+q}{q}\right) = \left(\frac{4d}{q}\right) = \left(\frac{d}{q}\right)$, and similarly $\left(\frac{q}{p}\right) = \left(\frac{d}{p}\right)$. Since $p + q = 4d$ we have $\left(\frac{d}{p}\right) = \left(\frac{d}{q}\right)$, and so, $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$. On the other hand, $(p-1)(q-1)/4 = (p-1)(4d-p-1)/4 = (4d(p-1) + 1 - p^2)/4 \equiv 0 \mod 2$, and therefore $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = 1 = (-1)^{(p-1)(q-1)/4}$. In the latter case, let $d = (q-p)/4$. We have in this case that $\left(\frac{q}{p}\right) = \left(\frac{d}{p}\right)$ and $\left(\frac{p}{q}\right) = \left(\frac{-d}{q}\right) = \left(\frac{-1}{q}\right)\left(\frac{d}{q}\right)$. Therefore, $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = \left(\frac{-1}{q}\right) = (-1)^{(q-1)/2} = (-1)^{(q-1)(p-1)/4}$.]

Remark 8.8.1 (Second Proof of the LQR) Having established that Gauss-Legendre and Euler's versions of the LQR are equivalent statements, and using Exercise 8.11, we get a second proof of the LQR. □

8.9 Vista: Laws of Reciprocity

As Hecke put it, "The modern Number Theory begins with the discovery of the Law of Quadratic Reciprocity". In a very strong sense, the modern Algebra and Algebraic Number Theory was developed in order to find the most general form of the law of reciprocity. In this section we briefly explain the history and the place of the Reciprocity Laws.

8.9.1 Laws of Reciprocity in Math

The Law of Quadratic Reciprocity (LQR) in the form of Legendre and Gauss, reveals a relationship between pairs of (odd) distinct primes in regard with solving quadratic congruences. A fact which goes beyond the scope of this course is that solving such congruences leads to studying **quadratic extensions** *of the field $\mathbb{Q}$. In its generalized form, using the Jacobi symbol, this Law relates pairs of numbers not necessarily prime (but still prime to each other).*

The LQR was rigorously proven for the first time by Gauss in "Disquisitiones Arithmeticae". In his lifetime, Gauss gave eight proofs of this Law (two of which in the "Disquisitiones") every one of which revealed different ways the LQR is connected to math in general. It was also Gauss who first realised that the LQR can be explained by considering extensions of $\mathbb{Q}$ different than quadratic, and in his studies of such extensions, he discovered the Biquadratic Reciprocity Law.

The generalizations of the LQR take the form of relationship between pairs of elements of extensions of $\mathbb{Q}$ of higher degree: three, four, ..., n. This led to the development of the Local and Global Class Field Theories in the first half of the last century. The list of people who made generalizations of the LQR consists of the names of many great mathematicians of that time: Gauss, Jacobi, Eisenstein, Kummer, Kronecker, Hilbert, Furtwängler, Takagi, Artin, Hasse, etc. Thus, for instance, it was possible to discover and prove the Law of Quintic Reciprocity only after Kummer introduced and developed the theory of ideal numbers. It was Hilbert who included finding the most general form of the laws of reciprocity, as number 9, in the list of (23) most important problems for Mathematics

in the Twentieth Century (he did this at the First Congress of the Mathematicians held in Paris in 1900). This most general form was found by E. Artin in 1927 (Artin's Reciprocity Law), and pertains to what is called Abelian Class Field Theory.

The generalization of the laws of reciprocity didn't stop there. In recent years the Abelian setup of the Class Field Theory was vastly generalized to non-Abelian Class Field Theory following the ideas and the vision of mathematicians like Shimura and Langlands. It was the development of this non-Abelian case which led to the powerful results of A. Wiles.

The discussion of any of the mentioned generalizations of the LQR goes far beyond the scope of these Lecture Notes. To complete our discussion of the LQR, in the next subsection, we give a treatment of the LQR and its close relation to solving quadratic Diophantine equations, and to the theory of p-adic numbers using the Hilbert's norm symbol.

8.9.2 Hilbert's Norm-Symbol and the Law of Quadratic Reciprocity

The Hilbert's Norm Symbol $(a,b)_p$

Suppose $a, b \in \mathbb{Q}^\times$. *Consider the equation* $ax^2 + by^2 = z^2$ *in* $\mathbb{Q}_p$.

Definition 8.9.1 *We define the Hilbert's norm symbol,* $(a,b)_p$, *by setting*
$(a,b)_p = 1$ *if the equation has a non-trivial solution, and*
$(a,b)_p = -1$ *if the equation has no such a solution.*

So, obviously, $(a,b)_p = (b,a)_p$.
Although we defined the Hilbert norm symbol for non-zero rational numbers, its computation reduces to the case of integer square-free integers. Indeed, the non-zero rational numbers a *and* b *can be written as*

$$a = a_1 \cdot x_1^2 \qquad b = b_1 \cdot y_1^2$$

where a_1 *and* b_1 *are square-free integers, and* x_1 *and* y_1 *are rational numbers. It is obvious that* $(a,b)_p = (a_1,b_1)_p$. *So, from this moment on we will work with integers* a *and* b.

To see the reason why the function $(\bullet,\bullet)_p$ *is called norm symbol, lets observe that*

Proposition 8.9.1 *We have* $(a,b)_p = 1$ *if, and only if,* $z^2 - by^2 = a$ *has (non-trivial) solutions in* $\mathbb{Q}_p$.

Proof *Indeed, the solution* (y_0, z_0) *of* $z^2 - by^2 = a$ *defines a solution* $(1, y_0, z_0)$ *of* $ax^2 + by^2 = z^2$, *and therefore* $(a,b)_p = 1$. *Suppose now* $(a,b)_p = 1$, *and let* (x_0, y_0, z_0) *be a non-trivial solution of* $ax^2 + by^2 = z^2$. *If* $x_0 \neq 0$, *we are done:* $(y_0/x_0, z_0/x_0)$ *is a (non-trivial) solution of* $z^2 - by^2 = 1$. *If* $x_0 = 0$, *then* $y_0 z_0 \neq 0$, *and* $b = (z_0/y_0)^2$. *Therefore*

$$z^2 - by^2 = (z - (z_0/y_0)y)(z + (z_0/y_0)y).$$

Since the system

$$z - (z_0/y_0)y = a \qquad z + (z_0/y_0)y = 1$$

is solvable in $\mathbb{Q}_p$, *we find a solution to* $z^2 - by^2 = a$ *as well.* □

In technical terms (from Galois Theory) the expression $z_0^2 - by_0^2 = a$ *means that* a *is the norm of the element* $z_0 + \sqrt{b}y_0 \in \mathbb{Q}_p(\sqrt{b})$. *In other words,* $(a,b)_p = 1$ *if, and only if,* a *is a norm of an element of* $\mathbb{Q}_p(\sqrt{b})$.

*Here are some basic (***local***) properties of the Hilbert's norm symbol.*

(1) The Hilbert's norm symbol is multiplicative with respect to its arguments

$$(a_1a_2,b)_p=(a_1,b)_p(a_2,b)_p \qquad (a,b_1b_2)_p=(a,b_1)_p(a,b_2)_p\,.$$

To see this we consider several cases. (i) Let $(a_1,b)_p=(a_2,b)_p=1$. *Then* $a_1=z_1^2-by_1^2$ *and* $a_2=z_2^2-by_2^2$. *It is obvious that*

$$a_1a_2=(z_1^2-by_1^2)(z_2^2-by_2^2)=(z_1z_2+by_1y_2)^2-b(z_1y_2+y_1z_2)^2,$$

so that $(a_1a_2,b)_p=1$. *(ii) Let* $(a_1,b)_p(a_2,b)_p=-1$. *Assume, by contradiction, that* $(a_1a_2,b)_p=1$. *W.L.O.G. we have* $(a_1,b)_p=1$. *Then by (i)*

$$(a_1a_2,b)_p(a_1,b)_p=(a_1a_2a_1,b)_p=(a_2,b)_p=-1$$

which is an absurdity, because $(a_1a_2,b)_p(a_1,b)_p=1\cdot 1=1$. *So,* $(a_1a_2,b)_p=-1$. *(iii) Let* $(a_1,b)_p=(a_2,b)_p=-1$. *Observe that, if* p *is an odd prime, then* $\left(\frac{a_1}{p}\right)=\left(\frac{a_2}{p}\right)=-1$. *This follows from the fact that, by the Hensel's lifting lemma, for any integer* c *we have* $\left(\frac{c}{p}\right)=1$ *implies that* $c=z_0^2$ *for* $z_0\in\mathbb{Q}_p$. *But then* $\left(\frac{a_1a_2}{p}\right)=1$ *and therefore* $a_1a_2=z_0^2-b\cdot 0^2$. *We conclude that* $(a_1a_2,b)_p=1$ *as needed. When* $p=2$ *we have that none of the square-free integers* a_1,a_2, *and* b *are squares in* $\mathbb{Q}_2$. *This is equivalent to having them congruent to* $2,3,5,6$ *or* 7 *modulo* 8. *Moreover, W.L.O.G. we may assume, modulo* $\mathbb{Q}_2^{\times 2}$ *that* a_1,a_2, *and* b *are equal to* $2,3,-3,-2$, *or* -1. *Let's fix* $b=-1$ *say, and lets find all possible values of* a_1 *so that* $(a_1,-1)_2=-1$. *Of the equations*

$$-x^2-y^2=z^2,\quad 2x^2-y^2=z^2,\quad -2x^2-y^2=z^2,\quad ,3x^2-y^2=z^2,\quad -3x^2-y^2=z^2$$

only the first the third and the fourth do not have non-trivial solutions in $\mathbb{Q}_2$. *So* a_1, *and* a_2 *for that matter, can be equal to* $-1,-2$ *or* 3. *But then* $a=a_1a_2$ *is either a square, or equal to* $2,-3$, *or* -6. *In all these cases* $ax^2-y^2=z^2$ *does have non-trivial solutions in* $\mathbb{Q}_2$. *We just proved that*

$$(a_1,-1)_2(a_2,-1)_2=(a_1a_2,-1)_p\,.$$

The rest of the cases: $b=2,3,-3$, *and* -2 *are treated in a similar way proving the multiplicity of the Hilbert norm symbol for* $p=2$ *as well.* □

The multiplicativity property of Hilbert's norm symbol reduces the computation of the symbol to cases such as $(\pm p,\pm q)_r$ *where* p,q,r *are, not necessarily distinct, prime numbers..*

(2) For every odd prime p**, and for every** $a\in\mathbb{Z}$ **with** $\gcd(a,p)=1$ **we have** $(a,p)_p=\left(\frac{a}{p}\right)$

Indeed, if the equation $a=z^2-py^2$ *is solvable in* $\mathbb{Q}_p$, *then* $a\equiv z^2 \mod p$ *is solvable as well. Therefore,* $(a,p)_p=1$ *implies* $\left(\frac{a}{p}\right)=1$. *Conversely, if* $\left(\frac{a}{p}\right)=1$, *then, as we know,* $a\in\mathbb{Z}_p^{\times 2}$, *and therefore* $a=z^2-py^2$ *has a non-trivial solution. The latter means that* $(a,p)_p=1$. *The claim is proved.* □

(3) For every odd prime p **we have** $(p,p)_p=\left(\frac{-1}{p}\right)$

This is obvious by looking at when solutions of $p=z^2-py^2$ *exist (do this as an exercise).* □

(4) For every integer m**, and for any prime number** p **we have** $(-m,m)_p=1$

Another obvious fact (exercise!). □

(5) if p **is an odd prime, and if** $\|a\|_p=\|b\|_p=1$ **(that is, if** $p\nmid ab$**) , then** $(a,b)_p=1$

This is a simple counting: Consider $ax^2 + by^2 \equiv z^2 \mod p$. *We may assume that* a, b *are not quadratic residues* $\mod p$. *Consider* $a + by^2 \mod p$ *for* $y = 0, 1, \ldots, (p-1)/2$. *At least one of these elements of* $\mathbb{Z}/p\mathbb{Z}$ *is a quadratic residue* $\mod p$, *because the nonresidues are* $(p-1)/2$. *So, we have* $a + by_0^2 \equiv z_0^2 \mod p$. *Using Hensel's lifting lemma, we find* $a + by_0^2 = z_1^2$ *in* $\mathbb{Q}_p$.

We are missing the values of the Hilbert's norm symbol for $p = 2$. *Here are the corresponding results.*

(6) If $\|a\|_2 = \|b\|_2 = 1$**, then** $(a, b)_2 = (-1)^{(a-1)(b-1)/4}$

Verification of this fact is quickly reduced to the cases when neither a, *nor* b *is a square in* $\mathbb{Q}_2$. *Therefore, one may assume that* a *and* b *are equal to* $3, 5$, *or* 7. *Direct check in which of the cases a solution to* $ax^2 + by^2 = z^2$ *exists, verifies the formula. (Exercise!)*
For the missing value 2 *of* a *or* b, *we have*

(7) For every odd integer a **we have** $(a, 2)_2 = (-1)^{(a^2-1)/8}$

As before, we may assume that a *is not a square in* $\mathbb{Q}_2$. *So,* $a = 3, -3$ *or* -1. *But obviously* $(-1, 2)_2 = 1$, $(3, 2)_2 = -1$ *(using arguments* $\mod 8$*), and* $(-3, 2)_2 = -1$ *(using arguments* $\mod 8$*). The formula is verified.* $\square$

The Hilbert's norm symbol is defined in connection with solving a homogeneous quadratic equation in all non-Archimedean completions of $\mathbb{Q}$. *Of course, it should be defined for the Archimedean completion as well. Here is the definition.*

Definition 8.9.2 *Define* $(a, b)_\infty = 1$ *if* $ax^2 + by^2 = z^2$ *has a non-trivial solution, and* $(a, b)_\infty = -1$ *if otherwise.*

Obviously, $(a, b)_\infty = -1$ *if, and only if, both* a *and* b *are negative rational numbers. The symbol* $(a, b)_\infty$ *satisfies the properties (1) and (4) above.*

We can prove now a **global** *property of the Hilbert's norm symbol. This property is called the* **Hilbert's Reciprocity Law**

Theorem 8.9.2 (*Hilbert's Reciprocity Law*) *Let* v *denote a prime number or the symbol* ∞. *Then, for every* $a, b \in \mathbb{Q}^\times$ *we have*

$$\prod_v (a, b)_v = 1.$$

As a matter of fact, the above formula is equivalent to the Legendre-Gauss's LQR.

Proof *Using the multiplicativity property of the symbol, we reduce* $(a, b)_v$ *to the case when* a, b *are,* -1, *or primes, while* v *is a prime equal to* a *or* b, *or is* ∞. *Let* $a = p$, $b = q$ *be distinct odd primes. Then*

$$\prod_v (p, q)_v = (p, q)_p (p, q)_q (p, q)_2 = \left(\frac{p}{q}\right)\left(\frac{q}{p}\right)(-1)^{(p-1)(q-1)/4}$$

which is equal to 1 *according to the LQR. Let* $a = -1$ *and* $b = q$ *an odd prime.*

$$\prod_v (-1, q)_v = (-1, q)_q (-1, q)_2 = \left(\frac{-1}{q}\right)(-1)^{(-1-1)(q-1)/4} = 1$$

by one of the supplementary parts of the LQR. We have obviously that

$$\prod_v (-1, -1)_v = (-1, -1)_2 (-1, -1)_\infty = (-1)(-1) = 1.$$

Let now $a = p$ be an odd prime. We have

$$\prod_v (p,2)_v = (p,2)_p (p,2)_2 = \left(\frac{2}{p}\right)(-1)^{(p^2-1)/8} = 1$$

by the second supplementary part of the LDR. Finally, for $a = b = 2$ we obviously have $(2,2)_v = 1$. The formula is proved. □

Exercise 8.32 *Prove that the the Hilbert's Reciprocity Law implies the Legendre-Gauss LQR.*

Chapter 9

Binomial Equations mod n, the Structure of $(\mathbb{Z}/n\mathbb{Z})^\times$

We are discussing in this chapter solving $\mod n$ *a binomial equations of any degree*

$$X^d - a \equiv 0 \mod n \qquad \gcd(a, n) = 1.$$

When there is a solution, the number $a \in \mathbb{Z}$ *is called* d**-the power residue** $\mod n$ *. In the previous chapter, we found all* **quadratic** *residues* $\mod n$*, and also all solutions to a fixed binomial quadratic equation. Our goal here is to do the same for the binomial equation of degree* d*: we will find all* d*-th power residues, and we will find all numbers whose* d*-th power is that power residue.*

Before going any deeper in the theory, let's show (as we did in the previous chapter) that considering $\gcd(n, a) = 1$ *is not a loss of generality: solving binomial equations without this restriction reduces to solving equations with it.*
Indeed, let $s = \gcd(a, n)$*, so that* $n = s \cdot n_1$ *and* $a = s \cdot a_1$ *with* $\gcd(n_1, a_1) = 1$*. Let further* x_0 *be a solution to* $X^d \equiv a \mod n$*. From* $x_0^d \equiv a \equiv 0 \mod s$ *we get that* $s \mid x_0^d$*. Let* s_1 *be the smallest positive integer such that*

$$s \mid s_1^d \qquad \text{and} \qquad s_1 \mid x_0.$$

Such an integer exists, because it has to be no bigger than x_0*. Also, the prime divisors of* s_1 *are the same as the prime divisors of* s*: otherwise, disregarding the ones not dividing* s*, we can find a smaller than* s_1 *number having the needed properties. Although* s_1 *is defined using a solution* x_0*, by the exercise below, it actually depends only on* s*.*
Denoting $x_0 = s_1 \cdot x_1$*, and* $s_1^d = s \cdot s_2$*, we can rewrite the relation* $x_0^d \equiv a \mod n$ *as follows*

$$(s_1 \cdot x_1)^d \equiv s \cdot a_1 \mod s \cdot n_1,$$

and after simplifications as

$$s_2 \cdot x_1^d \equiv a_1 \mod n_1.$$

Observe now that $\gcd(s_2, n_1) \mid a_1$ *which together with* $\gcd(s_2, n_1) \mid n_1$ *gives us that*

$$\gcd(s_2, n_1) \mid \gcd(a_1, n_1) = 1.$$

Therefore, x_1 *is a solution to*

$$X^d \equiv a_1 \cdot (s_2)^{-1} \mod n_1$$

with $\gcd(a_1 \cdot (s_2)^{-1}, n_1) = 1$*.*

Exercise 9.1 *Let in the notations above* $s = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$*. Prove that* $s_1 = p_1^{\beta_1} \cdots p_k^{\beta_k}$ *where for every* i *the number* β_i *satisfies* $\alpha_i \leq d \cdot \beta_i$*, and are the least numbers with this property.*

So, the solutions to the original equation $X^d \equiv a \mod n$ *are obtained solving* $X^d \equiv a_1 \cdot (s_2)^{-1} \mod n_1$*, and then multiplying them with* s_1*.*

From this point on, we are considering $\gcd(a,n) = 1$*. To keep up with the professional (algebraic) interpretation of what we are doing, denote the set of all d-the power residues* $\mod n$ *by* $(\mathbb{Z}/n\mathbb{Z})^{\times d}$*. That is,*

$$(\mathbb{Z}/n\mathbb{Z})^{\times d} = \{[a] \in (\mathbb{Z}/n\mathbb{Z})^\times \mid (\exists [b] \in (\mathbb{Z}/n\mathbb{Z})^\times)\ ([a] = [b]^d)\}.$$

*It is easy to verify (***do that as an exercise***) that a product of two d-th power residues is a d-th power residue, and that the inverse of every such residue is a d-th power residue as well. This means that, as it was in the case* $d = 1$*, the set* $(\mathbb{Z}/n\mathbb{Z})^{\times d}$ *is a subgroup of* $(\mathbb{Z}/n\mathbb{Z})^\times$*. Following the analogy with* $d = 2$*, consider the map*

$$\psi_d : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/n\mathbb{Z})^{\times d} \qquad [a] \mapsto \psi_d([a]) = [a]^d.$$

*This map is obviously a surjection, and is easy to check (***an exercise!***) that it is a group homomorphism. So,* ψ_d *is an epimorphism. Our goal can be formulated as "understanding" the map* ψ_d*: describing the structure of its co-domain* $(\mathbb{Z}/n\mathbb{Z})^{\times d}$*, and of its kernel* $ker(\psi_d) = \{[b] \mid \psi_d([b]) = [1]\}$*.*

To motivate the way we go about reaching our goal, let's make some observations.

(1) Observe that, since $\gcd(a,n) = 1$*, any solution,* x_0*, to* $X^d - a \equiv 0 \mod n$ *is relatively prime with n, and is therefore invertible* $\mod n$*. So, the class* $\mod n$ *of any such solution is naturally an element of* $(\mathbb{Z}/n\mathbb{Z})^\times$*.*
(2) Therefore, if x_1 *and* x_2 *are two solutions, then they both are invertible* $\mod n$*, and* $y = x_1(x_2)^{-1} \mod n$ *is a solution to*

$$Y^d - 1 \equiv 0 \mod n$$

that is, y **is a** *d***th root of unity** $\mod n$*.*
(3) Vice-versa, given a solution x_1*, any other such has the form* $x_2 = x_1 y \mod n$ *for a d-th root of unity* $\mod n$*.*

From (1), (2), and (3) it follows that
(i) to find if a is a d-th power residue $\mod n$*, we have to see if there is an element of* $(\mathbb{Z}/n\mathbb{Z})^\times$ *whose d-th power is a, and*
(ii) to find all elements like in (i) is equivalent to finding all d-th roots of identity $\mod n$*,*

These two conclusions lead to considering **orders of elements** $\mod n$*, and, ultimately, to finding the structure of the group* $(\mathbb{Z}/n\mathbb{Z})^\times$*. This is what we are doing in this chapter.*

9.1 Orders Modulo n

Consider the group $(\mathbb{Z}/n\mathbb{Z})^\times$*. We know, by Euler's theorem, that*

$$(\forall [a] \in (\mathbb{Z}/n\mathbb{Z})^\times)\left([a]^{\varphi(n)} = [1]\right).$$

So, we can speak of the least positive integer d such that $[a]^s = [1]$*.*

Definition 9.1.1 *For any* $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ *define the* **order of** $[a]$ *to be the least natural number s such that* $[a]^s = [1]$*. The order of* $[a]$ *is denoted by* $|[a]| = ord([a]) = ord_n(a)$ *. The number is also called* **the order of** a **modulo** n*.*

It is straightforward that the order modulo n exists. Here is the important theorem summarizing the properties of orders modulo n.

Theorem 9.1.1 *Let $[a], [b] \in (\mathbb{Z}/n\mathbb{Z})^\times$, and let $m \in \mathbb{Z}$. We have the following.*
(1) $[a]^m = [1]$ if, and only if, $|[a]| \,|\, m$;
(2)$|[a]^{-1}| = |[a]|$;
(3) $|[a]^m| = |[a]|/\gcd(|[a]|, m)$;
(4) $\gcd(|[a]|, |[b]|) = 1 \quad \Rightarrow \quad |[a][b]| = |[a]| \cdot |[b]|$.

Proof. *Denote by s the order of $[a]$, and by t - the order of $[b]$.*
(1): The "if" part is obvious: $m = s \cdot m'$ implies that $[a]^m = [a]^{s\cdot m'} = ([a]^s)^{m'} = ([1])^{m'} = [1]$. For the "only if" part, assume $[a]^m = [1]$. We want to show that $s \,|\, m$. Dividing m by s, we get $m = s \cdot q + r$ where $0 \le r < s$. We have that $[1] = [a]^m = [a]^{s\cdot q} \cdot [a]^r = [a]^r$, and so, if $r \neq 0$ we would have $s \le r < s$ which is an absurd. Therefore, $r = 0$, and m $s \cdot q$ as needed.
(2): It is enough to show that $([a]^m = [1]) \quad \Leftrightarrow \quad ([a^{-1}]^m = [1]$. But, by the cancellation property mod n, *$[a]^m = [1]$ is equivalent to $[b] \cdot [a]^m = [b]$ for any $[b] \in (\mathbb{Z}/n\mathbb{Z})^\times$. Choosing $[b] = [a]^{-1}$ we get what we wanted to prove.*
(3): Let $s' = |[a]^m|$, and $d = \gcd(s, m)$. We need to show that $s' = s/d$. By definition, s' is the least among the positive integers l such that $([a]^m)^l = [1]$. The integers l have the property $[a]^{m\cdot l} = [1]$, or equivalently (by (1)), $s \,|\, m \cdot l$. But we have that $m = d \cdot m_1, s = d \cdot s_1$ and $\gcd(s_1, m_1) = 1$. So, l has the property that $d \cdot s_1 \,|\, d \cdot m_1 \cdot l$ which is equivalent to $s_1 \,|\, m_1 \cdot l$. By $\gcd(s_1, m_1) = 1$ we get that l is distinguished by $s_1 \,|\, l$. The least such l is obviously $l = s_1$. Therefore, $s' = s_1 = s/d$ which is what we wanted to prove.
(4): Denote by $s' = |[a]|$, and by $s'' = |[b]|$. We have that $\gcd(s', s'') = 1$. By definition, the number $|[a][b]|$ is the least among the positive numbers l such that $([a][b])^l = [1]$. Equivalently, l satisfies $[a]^l = ([b]^{-1})^l$. But this means that $|[a]^l| = |([b]^{-1})^l|$ which, by (3), translates into $d = s'/(s', l) = s''/(s'', l)$. Notice that $d \,|\, s'$ and that $d \,|\, s''$. So, $d \,|\, \gcd(s', s'') = 1$. Therefore, $d = 1$. The latter can happen only if $(s', l) = s'$ and $(s'', l) = s''$, that is, only when $s' \,|\, l$ and $s'' \,|\, l$. The relative primeness of s' and s'' gives that $s' \cdot s'' \,|\, l$. The least positive such l is, of course $s' \cdot s''$. This proves that $|[a][b]| = |[a]| \cdot |[b]|$.
The theorem is proved. □

As an immediate consequence, using Euler's theorem, we get that

Corollary 9.1.2 *For every element $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ we have $|[a]| \,|\, \varphi(n)$.*

The following is a key fact about the group $(\mathbb{Z}/n\mathbb{Z})^\times$ It is true also for every finite Abelian group!.

Proposition 9.1.3 *Let $[a_0] \in (\mathbb{Z}/n\mathbb{Z})^\times$ be an element of* **maximum order:** *$\forall [a](|[a]| \le |[a_0]|$. Then,*

$$\forall [a]\ (|[a]| \,|\, |[a_0]|.$$

Proof. *Let $[a_0]$ be the element of the group $(\mathbb{Z}/n\mathbb{Z})^\times$ of largest order s_0, and let $[b]$ be any other element of that group. Denoting by t the order of $[b]$, we have to show that $t \,|\, s_0$. Assume by RAA, that $t \nmid s_0$. This, in particular, means that $1 < t, s_0$. Our assumption implies that there is a prime number p whose exponent in the presentation of t is bigger than the one in s_0. If we denote these exponents by β and α respectively, we get that $p^\alpha \,|\, s_0, \quad p^{\alpha+1} \nmid s_0$, that $p^\beta \,|\, t, \quad p^{\beta+1} \nmid t$, and that $0 \le \alpha < \beta$. By item (3) of the previous theorem we have that $|[a_0]^{p^\alpha}| = s_0/p^\alpha$, and that $|[b]^{t/p^\beta}| = p^\beta$. Denoting $[a'] = [a_0]^{p^\alpha}$ and $[b'] = [b]^{t/p^\beta}$, we get, by item (4) of the previous theorem, $|[a'][b']| = (s_0/p^\alpha) \cdot p^\beta = s_0 \cdot p\beta - \alpha \ge s_0 \cdot p > s_0$ which is impossible due to the extremal property of s_0. The Proposition is proved.* □

Definition 9.1.2 *The element $[g] \in (\mathbb{Z}/n\mathbb{Z})^\times$ is called a* **generator of** *$(\mathbb{Z}/n\mathbb{Z})^\times$ if $|[g]| = \varphi(n)$. We also say in such a case that the number g is a* **primitive root modulo** *n.*

9.2 Primitive Roots Modulo n

Obviously,when g is a primitive root modulo n, the group $(\mathbb{Z}/n\mathbb{Z})^\times$ consists of all powers of $[g]$

$$(\mathbb{Z}/n\mathbb{Z})^\times = \{[g]^0, [g]^1, \ldots, [g]^{\varphi(n)-1}\}.$$

It is very easy in such a case to find all d-th power residues $\mod n$, that is, all $[a]$ such that the equation

$$X^d - [a] = [0]$$

has solutions. Indeed, when $[g]$ is a generator of $(\mathbb{Z}/n\mathbb{Z})^\times$, the element $[a]$ is a d-th power in $(\mathbb{Z}/n\mathbb{Z})^\times$ exactly when it has the form $[a] = [g]^{ds}$ for some integer s. We will return to this again soon. So, it is interesting to know if $(\mathbb{Z}/n\mathbb{Z})^\times$ has a generator. The bad news is that the answer to that question is negative in general. The reason for this is very simple, and is easily seen by using the group isomorphism induced by the CRT. Namely, recall that we have

$$G_n : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z})^\times \times (\mathbb{Z}/p_2^{\alpha_2}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_k^{\alpha_k}\mathbb{Z})^\times$$

where $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$. As we have discussed this already, the two sides of this isomorphism are groups which can be identified. In particular, the LHS has an element of order $\varphi(n)$ if, and only if, the RHS has one. Suppose n is divisible by at least two **odd** *primes. Then, in such a case the following is true*

Exercise 9.2 *For every $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ we have that $|[a]| \,|\, \varphi(n)/2$.*
[Hint: Show that any element of the RHS raised to the power $\varphi(n)/2$ is equal to $([1]_{p_1^{\alpha_1}}, \ldots, [1]_{p_k^{\alpha_k}})$. You may have to use also the multiplicativity of the totient function, and may have to apply Euler's theorem as many times as needed.]

So, $(\mathbb{Z}/n\mathbb{Z})^\times$ could have a generator only when $n = 2^\alpha p^\beta$ where $\alpha, \beta \geq 0$. Applying an argument similar to the one in the hint above, one can (and should) prove that

Exercise 9.3 *If $n = 2^\alpha p^\beta$ where $\alpha \geq 2$ and $\beta \geq 1$, then the orders of the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$ divide $\varphi(n)/2$.*

So, the only cases of n having two prime divisors are $n = 2p^\beta$. Of course, there are also the cases when n is divisible by one prime only: $n = 2^\alpha, \alpha \geq 1$ or $n = p^\beta, \beta \geq 1$. The rest of this section is devoted to these three cases.

Remark 9.2.1 Recall from Subsection 7.6.2 the polynomials $e_n(X) = X^{\varphi(n)} - 1$ and $f_n(X) = \prod_{a \in R(n)}(X - a)$. One may think that existence of primitive roots $\mod n$ is related, at least in some weak sense, to $[e_n(X)]_n$ and $[f_n(X)]_n$ being equal as elements of $\mathbb{Z}/n\mathbb{Z}[X]$. As we pointed out in Remark 7.6.1, the coincidence of these polynomials is extremely rare: apart from the case of $n = p$ being a prime number, this happens only when $n = 2F_k$ where F_k is a Fermat prime. Our main approach to this question in the next subsection, is based on Proposition 9.1.3 and on Hensel's Lifting Lemma. Other approaches, which can be found in the literature, are discussed in the exercises to the same subsection.

9.2.1 Primitive Roots modulo $n = p^\beta$ for Odd p

Theorem 9.2.1 *Let p be an odd prime number, and $n = p^\beta$ for $\beta \geq 1$. Then the group $(\mathbb{Z}/p^\beta\mathbb{Z})^\times$ has a generator $[g]$. The number of distinct generators of the group is $\varphi(\varphi(p^\beta))$.*

Proof *We have to show first that $\varphi(p^\beta)$ is the order of an element in the group $\left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$. According to Proposition 9.1.3, there is an element g of $\left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$ of order d such that the order of every element of that group divides d. This implies that the elements of $\left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$ satisfy the equation*

$$f(X) = X^d - 1 \equiv 0 \mod p^\beta.$$

We are proving below that since this equation has $\varphi(p^\beta)$ distinct solutions, $\varphi(p^\beta) \mid d$. Since $|g| = d \mid \varphi(p^\beta)$ as well, we get that $d = \varphi(p^\beta)$, and therefore, g is a generator of the group $\left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$. To complete the proof of the theorem, showing also that the generators are $\varphi(\varphi(n))$ many, we notice that an element $x = g^s$ is a generator if, and only if, $\gcd(s, \varphi(p^\beta)) = 1$.

(1) Proving that $p-1 \mid d$. Since $d \mid \varphi(p^\beta)$, it has the form $d = d' \cdot p^\alpha$ with $d' \mid p-1$, and $\alpha \leq \beta - 1$. According to the general theory, the solutions to $f(X) = X^d - 1 \equiv 0 \mod p^\beta$ are all lifts of solutions to the same equation $\mod p$: solutions $\mod p$ which is liftable $\beta - 1$ times comprise the solutions $\mod p^\beta$. Hensel's lifting lemma tells us that a liftable from $\mod p^k$ to $\mod p^{k+1}$ solution produces either one or p solutions. So, to have eventually $\varphi(P^\beta)$ solutions, we need to have $p-1$ solutions $\mod p$, each of which is liftable to a solution $\mod p^\beta$.
Observe now that

$$[f(X)]_p = X^d - [1]_p = (X^{d'})^{p^\alpha} - [1]_p = (X^{d'} - [1]_p)^{p^\beta} \quad in \quad \mathbb{Z}/p\mathbb{Z}[X]$$

and so, by Lagrange's theorem 7.5.1, has no more than d' distinct solutions in $\mathbb{Z}/p\mathbb{Z}$. This implies that $d' = p - 1$. (The students who do all exercises diligently should have noticed that the last argument could have been replaced by Exercise 7.15 (3).)

(2) Proving that $p^{\beta-1} \mid d$. It is enough to consider $\beta \geq 2$ here. Since every solution to $f(X) \equiv 0 \mod p$ has to be liftable to $p^{\beta-1}$ solutions to $f(X) \equiv 0 \mod p^\beta$. This implies two important facts about $f(X)$ and its solutions.

(i) The derivative $[f'(X)]_p = [d]_p X^{d-1}$ has to be identically 0 in $\mathbb{Z}/p\mathbb{Z}[X]$ (equivalently, $d = d' \cdot p^\alpha \equiv 0 \mod p$ forcing $\alpha \geq 1$).

(ii) For every $k = 1, \ldots, \beta - 1$, and for every integer a such that $f(a) \equiv 0 \mod p^k$ we need to have also that $f(a) \equiv 0 \mod p^{k+1}$ (if the latter is not true, according to the Hensel's Lifting Lemma, a will not be liftable from solution $\mod p^k$ to solution $\mod p^{k+1}$).

Notice that item (ii) is equivalent to saying that for every a such that $\gcd(a, p) = 1$ we have $f(a) \equiv 0 \mod p^\beta$, and that this is exactly the assumption we are working under ($f(X) \equiv 0 \mod p^\beta$ has $\varphi(p^\beta)$ distinct solutions).

Now, since $\beta \geq 2$ the integer $1 + p$ should be a solution to $f(X) \equiv 0 \mod p^\beta$. An easy calculation by induction on k shows that (exercise)

$$(1+p)^{p^k} = 1 + p^{k+1} \cdot A_k \qquad \gcd(p, A_k) = 1.$$

Therefore $f(1+p) = p^{\alpha+1} \cdot A_\alpha$ with $p \nmid A_\alpha$, and since by assumption $f(1+p) \equiv 0 \mod p^\beta$, we have that $\beta \leq \alpha + 1$. This combined with $\alpha \leq \beta - 1$ implies that $\alpha = \beta - 1$ and that $p^{\beta-1} \mid d$ as claimed. □

The proof of this theorem can be made much shorter using the following steps.

Exercise 9.4 *According to Proposition 9.1.3 the maximal order d of an element of $\left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$ is divisible by the orders of all elements of that group. So, to show that $\varphi(p^\beta) \mid d$ it is enough to show that $p - 1 \mid d$ and that $p^{\beta-1} \mid d$. The former can be proved using Exercise 7.15 (iii) (or using the argument in item 1) of the proof of the previous theorem). The latter is equivalent to having an element of $\left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$ of order $p^{\beta-1}$. To finish the proof of the theorem is enough therefore to show that the order of $[1+p]_{p^\beta}$ is $p^{\beta-1}$. Do that!*

The next step after proving that primitive roots $\mod p^\beta$ exist would be to find a way to construct such roots. This is not an easy task in general. There is a good news and a bad news here! The good news is that finding a primitive root $\mod p^\beta$ with $\beta \geq 1$ reduces to the case of $\beta = 1$. The following exercises shows one way to show that.

Exercise 9.5 *Let g_0 be a primitive root $\mod p$. Then $g = (1+p)g_0^{p^\beta}$ is a primitive root $\mod p^\beta$.*

[Hint: Show that the order of $[g]_{p^\beta}$ is $\varphi(p^\beta)$ using the fact that $[1+p]_{p^\beta}$ has order $p^{\beta-1}$, and showing that the order of $g_0^{p^\beta}$ is $p-1$.]

The bad news is that finding primitive roots $\mod p$ *is not an easy task (when the prime number is big, of course!). In some of the exercises below, you will see cases in which a primitive root can be easily determined. In general, there are conjectures about what the primitive roots are, and how they behave. For instance, Gauss conjectured that the number* 10 *is a primitive root modulo infinitely many prime numbers. A very famous conjecture, formulated by E. Artin (1898-1962) in 1927, is that every integer which is not a square, and is not* -1 *is a primitive root for infinitely many prime moduli. A significant progress in direction of proving Artin's conjecture is the result that* **all but two prime numbers are primitive roots of infinitely many prime moduli**. *It has to be mentioned here that nobody knows the two exceptions!*

Exercise 9.6 *The following exercises suggest a way of proving the existence of primitive roots* $\mod p^\beta$ *which can be found in most of the books on Number Theory. Here, $p>2$ is a prime number.*
(1) Prove that there is a primitive root $\mod p$ *using either*
(i) Proposition 9.1.3, and Exercise 7.15 (3)
or
(ii) (Gauss) denoting, for every divisor d of $p-1$, by $\psi(d)$ the number of elements of $(\mathbb{Z}/p\mathbb{Z})^\times$ of order d, and showing that for every such d we have $\psi(d)=\varphi(d)$. In particular, $\psi(p-1)=\varphi(p-1)\neq 0$.
(2) Let g be a primitive root $\mod p$*. Prove that there is $s_0\in\{1,2,\dots,p-1\}$ such that the order of $g+s_0p$ modulo p^2 is bigger than $p-1$. (Equivalently (why?), $g+s_0p$ is not a solution to $X^{p-1}-1\equiv 0 \mod p^2$.) Conclude that $(g+s_0p)^{p-1}=1+p\cdot A$ with $p\nmid A$.*
(3) Prove that,for every $k\geq 0$,

$$(g+s_0p)^{(p-1)p^k}=1+p^{k+1}A_k \qquad p\nmid A_k.$$

Conclude that, for every $\beta\geq 2$, the integer $g+s_0p$ is a primitive root modulo p^β.

9.2.2 Primitive Roots Modulo $n=2p^\beta$

Theorem 9.2.2 *Let g be a primitive root modulo p^β. If g is odd, then it is a primitive root modulo $2p^\beta$. If g is even, then $g+p^\beta$ is a primitive root modulo $2p^\beta$.*

Proof. *A very easy Exercise.* □

9.2.3 Primitive Roots Modulo $n=2^\alpha$

Theorem 9.2.3 *(1) The group $(\mathbb{Z}/2\mathbb{Z})^\times$ has only one element, so it has a generator;*
(2) The group $(\mathbb{Z}/4\mathbb{Z})^\times$ has a generator: the element $[3]_4$;
(3) For $\alpha\geq 3$, the maximum order of the elements of $(\mathbb{Z}/2^\alpha\mathbb{Z})^\times$ is $2^{\alpha-2}$, so the group has no generators.

Proof. *Items(1) and (2) being obvious are left as an (easy) exercise. Let's prove (3).*
We know that there is an element h of $[a]\in(\mathbb{Z}/2^\alpha\mathbb{Z})^\times$ which has a maximal order d. Since $d\,|\,2^{\alpha-1}=\varphi(2^\alpha)$, we have that $d=2^\gamma$ for $\gamma<\alpha$. To find this maximal order d, we notice that for every odd integer $h=1+2a$ we have that $h^2=1+8b_1$. Using induction, it is straightforward to show that for any positive natural number k

$$h^{2^k}=1+2^{k+2}b_k,$$

and that if b_1 is odd, then b_k is odd for every $k\geq 1$. This implies that $h^{2^{\alpha-2}}\equiv 1 \mod 2^\alpha$, and that $\gamma\leq\alpha-2$. Moreover, since

$$3^2=1+8\cdot 1,$$

the class of 3 *in* $(\mathbb{Z}/2^\alpha\mathbb{Z})^\times$ *is the maximal possible*

$$|[3]_{2^\alpha}| = 2^{\alpha-2}.$$

(Notice that the order of 3 *modulo* 4 *is also the maximal one,* $|[3]_4| = 2$*.)* □

Proposition 9.2.4 *Let* $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2^{\alpha-2}\mathbb{Z}$ *be the group with operation addition coordinate-wise. Consider the map*

$$\psi_\alpha : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{\alpha-2}\mathbb{Z} \to (\mathbb{Z}/2^\alpha\mathbb{Z})^\times$$

defined by $([a]_2, [b]_{2^{\alpha-2}}) \mapsto [(-1)^a 3^b]_{2^\alpha}$*. Then,* ψ_α *is a group isomorphism, that is, it is a bijection such that*

$$(\forall x, y)(\psi_\alpha(x + y) = \psi_\alpha(x)\psi_\alpha(y)).$$

Proof. *We need to show first that the map* ψ_α *is well defined: no choices of* a *and* b *affect where the elements are sent via* ψ_α*. But that's obvious in view of the fact that the order of* 5 *modulo* 2^α *is* $2^{\alpha-2}$*. Next, we show that*

$$(\mathbb{Z}/2^\alpha\mathbb{Z})^\times = \{[\pm 3^k]_{2^\alpha} \,|\, 0 \le k \le 2^{\alpha-2}\}$$

which is also straightforward (here we use that $\alpha \ge 3$*). Since* ψ_α *is onto, it is a bijection. Finally* ψ_α *respects the operations of the domain and co-domain again in an obvious way.* □

9.2.4 Main Theorem

We summarize the results in this section in the following theorem.

Theorem 9.2.5 *Let* n *be a positive integer. Then the group* $(\mathbb{Z}/n\mathbb{Z})^\times$ *has a generator, in other words - there is a primitive root modulo* n*, if, and only if,* $n = 2, 4, p^\beta$ *and* $2p^\beta$ *where* $\beta \ge 1$*.*

Exercise 9.7 *(1) Let* p *be an odd prime number, and let* k *be a natural number. Compute* $\mod p$ *the sum*

$$1^k + 2^k + \cdots (p-1)^k$$

(2) Compute the order of 2 *modulo* n *for* $n = 11, 13, \ldots, 19$*. Do the same for the order of* 3 *modulo* n *where* $n = 10, 11, 13, 14, 16, 17, 19$*.*
(3) Let p *be an odd prime number. Prove that a primitive root* $\mod p^\alpha$ *is also a primitive root* $\mod p$*.*
(4) Suppose p *is an odd prime, and that* g *is a primitive root* $\mod p$*. Prove that if* $g^{p-1} - 1 \equiv 0 \mod p^2$*, then* g *is not a primitive root* $\mod p^\alpha$ *for* $\alpha \ge 2$*.*
(5) Suppose p *is an odd prime, and that* g *is an integer. Prove that if* p *is* $1 \mod 4$*, then* g *is a primitive root* $\mod p$ *if, and only if,* $-g$ *is, and that if* p *is* $3 \mod 4$*, then* g *is a primitive root* $\mod p$ *if, and only if, the order* $\mod p$ *of* $-g$ *is* $(p-1)/2$*.*
(6) Suppose that p *and* $q = (p-1)/2$ *are both prime numbers. Prove that if* q *is* $1 \mod 4$*, then* 2 *is a primitive root* $\mod p$*, and if* q *is* $3 \mod 4$*, then* -2 *is a primitive root* $\mod p$*. Prove also that if* $q > 3$*, then* -3 *is a primitive root* $\mod p$*.*
(7) (Chebychev) (i) Show that 3 *is a primitive root* $\mod p$ *for any prime* $p = 2^n + 1 > 3$*. (ii) Let* $p = 4q + 1$ *and* $q > 2$ *be primes. Prove that* 3 *is a primitive root modulo* p*.*
(8) Let $p > 3$ *be a prime number, and let* $PR_p = \{[a]_p \,|\, ord_p([a]) = p - 1\}$ *be the set of primitive roots* $\mod p$*. Prove that*

$$\prod_{[a]\in PR_p} [a] = [1].$$

(9) Let p *be an odd prime number, and let* g *be a primitive root* $\mod p$*. Is it true that*

$$[g] \in (\mathbb{Z}/p\mathbb{Z})^{\times 2}\,?$$

9.3 The Structure of $(\mathbb{Z}/n\mathbb{Z})^\times$

In this section, we are designing an isomorphism between the group $(\mathbb{Z}/n\mathbb{Z})^\times$ and a product of finitely many cyclic groups, depending only on n, revealing this way the structure of that group $(\mathbb{Z}/n\mathbb{Z})^\times$. This isomorphism will allow us to reduce solving binomial equations like

$$X^d - a \equiv 0 \mod n \quad \textit{for} \quad \gcd(a,n) = 1$$

to solving a system of **linear** *equations of type*

$$dY - b \equiv 0 \mod n'.$$

In other words, this isomorphism replaces **exponentiation** *with* **addition***: the same way the logarithmic function in Calculus does. The isomorphism is not canonical: it depends on choices of integers. These integers form a* **system of indices modulo** n *, and play the role of bases for the logarithmic functions in Calculus. So, informally and intuitively, the isomorphism provides, for the fixed n, a system of* **discrete logarithmic functions** *which help us solve the higher degree equation.*

Recall the isomorphism

$$G_n : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/2^\alpha\mathbb{Z})^\times \times (\mathbb{Z}/p_1^{\alpha_1}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_k^{\alpha_k}\mathbb{Z})^\times .$$

We have already thoroughly studied the group factors on the RHS of this isomorphism. In particular, we know that for all p_i the group $(\mathbb{Z}/p_i^{\alpha_i}\mathbb{Z})^\times$ has a generator (in the other terminology - there is a primitive root modulo $p_i^{\alpha_i}$). We also know how the group $(\mathbb{Z}/2^\alpha\mathbb{Z})^\times$ looks like depending on α. Using this knowledge of ours, we are expressing below the group $(\mathbb{Z}/n\mathbb{Z})^\times$ as a product of groups of type $\mathbb{Z}/s\mathbb{Z}$ for appropriate natural numbers s.

9.3.1 Indices Modulo $p^\beta,\ p > 2$

Let g be a primitive root modulo p^β. As we know, since the order of g modulo p^β is $\varphi(p^\beta)$,

$$g^A \equiv g^B \mod p^\beta \qquad \textit{if, and only if,} \qquad A \equiv B \mod \varphi(p^\beta).$$

Since $[g] \in \left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$ generates the group,

$$(\mathbb{Z}/p^\beta\mathbb{Z})^\times = \{[g]^0, [g]^1, \ldots, [g]^{\varphi(p^\beta)-1}\},$$

every element $[a] \in \left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times$ is a power of $[g]$: $[a] = [g]^s$. By the above, the number s is unique modulo $\varphi(p^\beta)$.

Using this we define the **index map modulo p^β to the base** *g as follows. We use in the definition below two different moduli, p^β and $\varphi(p^\beta)$, so we are careful with the notations! (Nevertheless, as usual, we will be simplifying the notations whenever this will not easily lead to misunderstanding!)*

Definition 9.3.1 *Let g be a primitive root modulo p^β. The map*

$$I_g^{(p^\beta)} : \left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times \to \mathbb{Z}/\varphi(p^\beta)\mathbb{Z} \qquad [a]_{p^\beta} \mapsto I_g^{(p^\beta)}([a]_{p^\beta}) := [i]_{\varphi(p^\beta)}$$

where $g^i \equiv a \mod p^\beta$, is called the **index map modulo p^β to the base** *g. Any integer $s \in I_g([a])$ is called* **the index of** *a* **modulo p^β to the base** *g.*

By the above discussion, we see that the map is well defined: the class $[i]_{\varphi(p^\beta)}$ is uniquely determined. Obviously, s is an index of a modulo p^β to the base g, if, and only if, $g^s \equiv a \mod p^\beta$.

In particular, if $\beta = 1$, we have the map

$$I_g^{(p)} : (\mathbb{Z}/p\mathbb{Z})^\times \to \mathbb{Z}/(p-1)\mathbb{Z}.$$

The map $I_g^{(p^\beta)}$ is a bijection, and, by the very definition of the map, its inverse map is the exponentiation, called **the exponential function modulo p^β to the base g**

$$\exp_g^{(p^\beta)} : \mathbb{Z}/\varphi(p^\beta)\mathbb{Z} \to \left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times \qquad [i]_{\varphi(p^\beta)} \mapsto \exp_g^{(p^\beta)}([i]_{\varphi(p^\beta)}) = [g]_{p^\beta}^i.$$

Exercise 9.8 *(1) Verify that the definition of $\exp_g$ is correct, that is $\exp_g([i])$ does not depend on the choice of $s \in [i]_{\varphi(p^\beta)}$.*
(2) Prove moreover that the exponentiation can be extended to a map

$$\exp^{(p^\beta)} : \left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times \times \mathbb{Z}/\varphi(p^\beta)\mathbb{Z} \to \left(\mathbb{Z}/p^\beta\mathbb{Z}\right)^\times \quad ([a]_{p^\beta}, [b]_{\varphi(p^\beta)}) \mapsto \exp_{[a]_{p^\beta}}^{(p^\beta)}([b]_{\varphi(p^\beta)}) = ([a]_{p^\beta})^b.$$

Verify that this map has the familiar properties of exponentiation:

$$([a]_{p^\beta})^{[b]_{\varphi(p^\beta)}+[c]_{\varphi(p^\beta)}} = ([a]_{p^\beta})^{[b]_{\varphi(p^\beta)}} \cdot ([a]_{p^\beta})^{[c]_{\varphi(p^\beta)}} \qquad \textit{and}$$

$$(([a]_{p^\beta})^{[b]_{\varphi(p^\beta)}})^{[c]_{\varphi(p^\beta)}} = ([a]_{p^\beta})^{[b]_{\varphi(p^\beta)}\cdot[c]_{\varphi(p^\beta)}}.$$

Very importantly, the map I_g respects the operations in the groups it relates: the product in the domain is transformed to the sum in the co-domain.

Proposition 9.3.1 *For every two elements $[a], [b] \in (\mathbb{Z}/p^\beta\mathbb{Z})^\times$ we have*

$$I_g([a][b]) = I_g([a]) + I_g([b]) \quad \textit{in} \quad \mathbb{Z}/\varphi(p^\beta)\mathbb{Z}.$$

In particular, for every k

$$I_g([a]^k) = kI_g([a]) \quad \textit{in} \quad \mathbb{Z}/\varphi(p^\beta)\mathbb{Z}.$$

Proof. *An easy* **Exercise**. *We have used simplified notations here: no moduli are mentioned!* □

Using a professional language, the index map is an **isomorphism** *between the groups it relates. This means, we can (and will) identify the two groups via this map. Keep in mind that* **this identification is possible only after choosing a primitive root***! So there are as many as $\varphi(\varphi(p^\beta))$ such identifications. It is good to know how identifications corresponding to incongruent primitive roots are related.*

Proposition 9.3.2 *Let g and g_1 be primitive roots modulo p^β. Then the corresponding index maps I_g and I_{g_1} are related by*

$$I_{g_1}([a]) = I_g([a]) \cdot I_{g_1}([g]) \quad \textit{in} \quad \mathbb{Z}/\varphi(p^\beta)\mathbb{Z}.$$

Proof *The claim follows from the previous proposition, and from the obvious relations*

$$[g_1]^{I_{g_1}([a])} = [a] = [g]^{I_g([a])} = ([g_1]^{I_{g_1}([g])})^{I_g([a])} = [g_1]^{I_{g_1}([g])\cdot I_g([a])}. \quad \square$$

Obviously, the index map modulo p^β associated with a primitive root g has the properties of the familiar logarithmic function from Calculus.

Example 9.3.1

9.3.2 Indices Modulo 2^α

When $n = 2^\alpha$ for $\alpha \geq 1$, we have to consider two cases: (i) $\alpha \leq 2$, and (ii) $\alpha \geq 3$.

In case (i), we have that $(\mathbb{Z}/2^\alpha\mathbb{Z})^\times$ is either a trivial group, that is with only one element in it, or has two elements one of which is a generator (3 is the only primitive root modulo 4). We have in this case unique isomorphisms (no choices used!), the index map

$$I^{(2^\alpha)} : (\mathbb{Z}/2^\alpha\mathbb{Z})^\times \to \mathbb{Z}/\varphi(2^\alpha)\mathbb{Z}.$$

In case (ii), we established an isomorphism

$$\psi_\alpha : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{\alpha-2}\mathbb{Z} \to (\mathbb{Z}/2^\alpha\mathbb{Z})^\times$$

the inverse of which we call the index map

$$I^{(2^\alpha)} : (\mathbb{Z}/2^\alpha\mathbb{Z})^\times \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{\alpha-2}\mathbb{Z}.$$

Due to the fact that the maps I^{2^α} are group homomorphisms, they send products to sums, and therefore have the properties of a discrete logarithmic function.

Definition 9.3.2 *The map $I^{(2^\alpha)}$ described above is called the* **index map modulo** 2^α*. A number s, when $\alpha \leq 2$, or a pair if numbers (s,t), when $\alpha \geq 3$, in the class $I^{(2^\alpha)}([a])$ is called the* **index of** a **modulo** 2^α *when $\alpha \leq 2$, and the* **system of indices of** a **modulo** 2^α *when $\alpha \geq 3$.*

Example 9.3.2

9.3.3 System of Indices Modulo n

Let

$$n = 2^\alpha p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k} \qquad \text{for} \quad 2 < p_1 < \cdots < p_k, \quad \alpha \geq 0, \quad k \geq 0.$$

Using the index maps $I^{(2^\alpha)}$ and $I_{g_i}^{(p_i^{\beta_i})}$ we construct a map

$$H_n : (\mathbb{Z}/2^\alpha\mathbb{Z})^\times \times \left(\mathbb{Z}/p_1^{\beta_1}\mathbb{Z}\right)^\times \times \cdots \times \left(\mathbb{Z}/p_k^{\beta_k}\mathbb{Z}\right)^\times \to A \times \mathbb{Z}/\varphi(p_1^{\beta_1})\mathbb{Z} \times \cdots \times \mathbb{Z}/\varphi(p_k^{\beta_k})\mathbb{Z},$$

where $A = \mathbb{Z}/1\mathbb{Z} \cong (0)$ when $\alpha \leq 1$, $A = \mathbb{Z}/2\mathbb{Z}$ when $\alpha = 2$, and $A = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{\alpha-2}\mathbb{Z}$ when $\alpha \geq 3$, which sends

$$([a_0], [a_1], [a_2], \ldots, [a_k]) \quad \text{to} \quad (I^{(2^\alpha)}([a_0]), I_{g_1}([a_1]), I_{g_2}([a_2]), \ldots, I_{g_k}([a_k])).$$

This map is a bijection, because it is so component-wise. Note that the domain and the co-domain of H_n are groups. The domain - with component-wise multiplication, and the co-domain with component-wise addition.

Exercise 9.9 *Prove that H_n is a group homomorphism, that is, for any $x, y \in Dom(H_n)$, we have*

$$H_n(x \cdot y) = H_n(x) + H_n(y).$$

The composition of

$$H_n : (\mathbb{Z}/2^\alpha\mathbb{Z})^\times \times \left(\mathbb{Z}/p_1^{\beta_1}\mathbb{Z}\right)^\times \times \cdots \times \left(\mathbb{Z}/p_k^{\beta_k}\mathbb{Z}\right)^\times \to A \times \mathbb{Z}/\varphi(p_1^{\beta_1})\mathbb{Z} \times \cdots \times \mathbb{Z}/\varphi(p_k^{\beta_k})\mathbb{Z},$$

with

$$G_n : (\mathbb{Z}/n\mathbb{Z})^\times \to (\mathbb{Z}/2^\alpha\mathbb{Z})^\times \times \left(\mathbb{Z}/p_1^{\beta_1}\mathbb{Z}\right)^\times \times \cdots \times \left(\mathbb{Z}/p_k^{\beta_k}\mathbb{Z}\right)^\times$$

defines a map

$$I_n = H_n \circ G_n : (\mathbb{Z}/n\mathbb{Z})^\times \to A \times \mathbb{Z}/\varphi(p_1^{\beta_1})\mathbb{Z} \times \cdots \times \mathbb{Z}/\varphi(p_k^{\beta_k})\mathbb{Z}$$

sending $[a] \in (\mathbb{Z}/n\mathbb{Z})^\times$ to

$$I_n([a]) = (I^{(2^\alpha)}([a]), I_{g_1}([a]), I_{g_2}([a]), \ldots, I_{g_k}([a])).$$

Since I_n is a composition of two bijections, then it is a bijection as well. On the other hand, I_n is a group homomorphism, (the Exercise below), so I_n is an isomorphism. Keep in mind that the map I_n depends on choices if $k \geq 1$: the primitive roots modulo $p_i^{\beta_i}$. We have suppressed this dependence in our notations!

Exercise 9.10 *Verify that indeed I_n is a homomorphism. That is, prove that for every $[a], [b] \in (\mathbb{Z}/n\mathbb{Z})^\times$ we have*

$$I_n([a][b]) = I_n([a]) + I_n([b]).$$

Definition 9.3.3 *The map I_n constructed above is called an* **index map modulo** n. *The components of the k- or $k+1$-tuple $I_n([a])$ are called a* **system of indices of** $[a]$ **modulo** n.

9.4 Solving $X^d \equiv a \mod n$

We are answering in this section the questions on when do solutions to the equation in the title, and how many are they when exist? This we do using the theory of indices developed in this chapter.

9.4.1 General Results

The index map associated with $n = 2^\alpha p_1^{\beta_1} \cdots p_k^{\beta_k}$, and a choice of primitive roots modulo $p_i^{\beta_i}$ for $i = 1, \ldots, k$,

$$I_n : (\mathbb{Z}/n\mathbb{Z})^\times \to A \times \mathbb{Z}/\varphi(p_1^{\beta_1})\mathbb{Z} \times \cdots \times \mathbb{Z}/\varphi(p_k^{\beta_k})\mathbb{Z}$$

where $A = \mathbb{Z}/2^0\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z}$ or $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{\alpha-2}\mathbb{Z}$ depending on whether α is ≤ 1, $= 2$ or ≥ 3 respectively, is a group isomorphism. That is, I_n sends a product of elements into the sum of their images, $I_n(xy) = I_n(x) + I_n(y)$. From this it immediately follows that for every natural number d, and for every integer a such that $\gcd(a, n) = 1$,

$$I_n(a^d) = dI_n(a).$$

Suppose now that x is an integer. Then

$$\begin{aligned} x \quad &\textit{is a solution to} \quad X^d - a \equiv 0 \mod n \\ \Leftrightarrow \quad & x^d \equiv a \mod n \\ \Leftrightarrow \quad & dI_n(x) = I_n([a]_n) \end{aligned}$$

the first equivalence being true by definition of solution to an equation, and the second - due to the property of I_n of sending products to sums. The last is actually an equality of two vectors

$$d(I^{(2^\alpha)}(x), I_{g_1}(x), \ldots, I_{g_k}(x)) = (I^{(2^\alpha)}([a]_{2^\alpha}), I_{g_1}([a]_{p_a^{\beta_1}}), \ldots, I_{g_k}([a]_{p_k^{\beta_k}})).$$

In other words, if we introduce the unknown vector $\overrightarrow{X} = (X_0, X_1, \ldots, X_k)$, the original degree d equation is equivalent to the linear one

$$d\overrightarrow{X} = (I^{(2^\alpha)}([a]_{2^\alpha}), I_{g_1}([a]_{p_a^{\beta_1}}), \ldots, I_{g_k}([a]_{p_k^{\beta_k}})).$$

Let $\overrightarrow{x} = (x_0, x_1, \ldots, x_k)$ be a solution to the linear equation (keep in mind that $x_0 \in A$ and $x_i \in \mathbb{Z}/\varphi(p_i^{\beta_i})\mathbb{Z}$ for $1 \leq i \leq k$). Then the corresponding solution to the original equation is given by

$$\begin{aligned} I_n^{-1}(\overrightarrow{x}) &= (H_n \circ G_n)^{-1}(\overrightarrow{x}) = G_n^{-1}(H_n^{-1}(\overrightarrow{x})) \\ &= G_n^{-1}((\psi_\alpha(x_0), [g_1]^{x_1}, \ldots, [g_k]^{x_k})) \in (\mathbb{Z}/n\mathbb{Z})^\times. \end{aligned}$$

Notice that

$$(\psi_\alpha(x_0), [g_1]^{x_1}, \ldots, [g_k]^{x_k}) \in (\mathbb{Z}/2^\alpha\mathbb{Z})^\times \times \left(\mathbb{Z}/p_1^{\beta_1}\mathbb{Z}\right)^\times \times \cdots \times \left(\mathbb{Z}/p_k^{\beta_k}\mathbb{Z}\right)^\times.$$

Now, since $d\overrightarrow{X} = (dx_0, dx_1, \ldots, dx_k)$ we see that solving $d\overrightarrow{X} = I_n([a_n])$ is equivalent to solving $k, k+1$ or $k+2$ equations depending on whether $\alpha \leq 1, = 2$ or ≥ 3:

$$dX_0 = I^{(2^\alpha)}([a]), \quad dX_1 = I_{g_1}([a]), \quad \cdots, \quad dX_k = I_{g_k}([a]),$$

that is, to a system of linear congruences which we know how to deal with.

Put all this another way, we want to solve

$$X^d - a \equiv 0 \mod n, \quad \gcd(a,n) = 1, \quad n = 2^\alpha p_1^{\beta_1} \cdots p_k^{\beta_k}.$$

As we know, the (degree d) equation is equivalent to the system of (degree d) equations

$$X^d - a \equiv 0 \mod 2^\alpha, \quad X^d - a \equiv 0 \mod p_1^{\beta_1}, \quad \cdots, \quad X^d - a \equiv 0 \mod p_k^{\beta_k}.$$

Using the theory of indices, we solve an equivalent system of **linear** *equations instead*

$$dY_0 = I^{(2^\alpha)}([a]), \quad dY_1 = I_{g_1}([a]), \quad \cdots, \quad dY_k = I_{g_k}([a]).$$

Then using the inverse of I_n map we get the solution we wanted.

As an immediate corollary of this discussion, and of the developed theory we get the following.

Theorem 9.4.1 *Let $n = 2^\alpha p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$ be an odd number,and let $g_1, g_2, \ldots, g_k$ be primitive roots modulo $p_1^{\beta_1}, p_2^{\beta_2}, \ldots, p_k^{\beta_k}$ respectively. Let d be a natural number, and a - an integer relatively prime with n. Denote by $d_i = \gcd(d, \varphi(p_i^{\beta_i}))$. Consider the equation*

$$x^d - a \equiv 0 \mod n.$$

The following holds true.
(i) If $\alpha \leq 1$ that is, if $n = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$ or if $n = 2p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$, then solutions exist if, and only if,

$$\forall i \in \{1, 2, \ldots, k\} \ (d_i \mid I_{g_i}([a]_{p_i^{\beta_i}})$$

in which case the number of incongruent modulo n solutions is $d_1 d_2 \cdots d_k$.
(ii) If $\alpha = 2$, that is, if $n = 4p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$, letting $d_0 = \gcd(d, 2)$, solutions exist if, and only if,

$$d_0 \mid I^{(4)}([a]_4) \quad \wedge \quad \forall i \in \{1, 2, \ldots, k\} \ (d_i \mid I_{g_i}([a]_{p_i^{\beta_i}})$$

in which case the number of incongruent solutions is equal to $d_0 d_1 \cdots d_k$.
(iii) If $\alpha \geq 3$, let $d_0' = \gcd(d, 2)$, $d_0'' = \gcd(d, 2^{\alpha-2})$, and $I^{(2^\alpha)}([a]_{2^\alpha}) = ([s]_2, [t]_{2^{\alpha-2}}) \in A$. Then, solutions exist if, and only if

$$d_0' \mid s \quad \wedge \quad d_0'' \mid t \quad \wedge \quad (\forall i \in \{1, 2, \ldots, k\}) \ (d_i \mid I_{g_i}([a]_{p_i^{\beta_i}}))$$

in which case the number of incongruent solutions is equal to $d_0' d_0'' d_1 \cdots d_k$. □

This theorem is a powerful theoretical tool. Let's use it to draw some interesting conclusions regarding the group $(\mathbb{Z}/n\mathbb{Z})^\times$.

Observe first that the equation

$$X^d - 1 \equiv 0 \mod n$$

always has solutions. So, we get the following result.

Corollary 9.4.2 *Denote by $u(n,d)$ the number if d-th roots of unity in $(\mathbb{Z}/n\mathbb{Z})^\times$, that is,*

$$u(n,d) = |\{[a]_n \mid a^d - 1 \equiv 0 \mod n\}|.$$

In the notations of the theorem above we have
(i) If $\alpha \leq 1$, then $u(n,d) = d_1 d_2 \cdots d_k$.
(ii) If $\alpha = 2$, then $u(n,d) = d_0 d_1 \cdots d_k$.
(iii) If $\alpha \geq 3$, then $u(n,d) = d_0' d_0'' d_1 \cdots d_k$. □

Following the reasoning in Theorem 8.5.2, we can conclude now information about the size of the group $(\mathbb{Z}/n\mathbb{Z})^{\times d}$, of d-th power residues $\mod n$. (Or, equivalently, about the index of this group in $(\mathbb{Z}/n\mathbb{Z})^\times$.) We have the following

Corollary 9.4.3 *In the notations above we have $|\ (\mathbb{Z}/n\mathbb{Z})^{\times d}\ | = \varphi(n)/u(n,d)$. Or, in professional notations,*

$$[(\mathbb{Z}/n\mathbb{Z})^\times : (\mathbb{Z}/n\mathbb{Z})^{\times d}] = u(n,d). \quad \square$$

Exercise 9.11 (*) *Derive the result from Theorem 8.5.2 using the last two corollaries.*
[Hint: You need to prove two things here. First - that if solutions exist in the case when $d = 2$ and $n \geq 3$, then all $d_0', d_0'', d_0, d_1, \ldots, d_k$ need to be equal to 2. And second - that, for $i \geq 1$, we have $d_i \neq 0$ if, and only if, $\left(\frac{a}{p_i}\right) = 1$.]

The situation is particularly simple when $d_0' = d_0'' = d_0 = d_i = 1$ for every $i = 1, 2, \ldots, k$. We have then that, for any a with $\gcd(a,n) = 1$, the solution to $X^d - a \equiv 0 \mod n$ exists and is unique.

Exercise 9.12 *Prove that $d_0' = d_0'' = d_0 = d_i = 1$ if, and only if, $\gcd(d, \varphi(n)) = 1$.*

In the professional terminology, we have

Corollary 9.4.4 *Suppose the positive integer d is relatively prime with $\varphi(n)$. We have*

$$(\mathbb{Z}/n\mathbb{Z})^\times = (\mathbb{Z}/n\mathbb{Z})^{\times d}.$$

We are formulating here, just for the record, the results above when $n = p$ is an odd prime number.

Theorem 9.4.5 *Consider the equation $X^d \equiv a \mod p$. If $a \equiv 0 \mod p$, then the only solution (of multiplicity d) is $X \equiv 0 \mod p$. If $[a]_p \neq [0]_p$, then a solution exists if, and only if, the linear equation*

$$d \cdot Y \equiv I_g([a]_p) \mod (p-1)$$

has a solution (that is if, and only if, $gcd(d, p-1) \mid I_g([a]_p)$), in which case the solutions to $X^d \equiv a \mod p$ are given by $[g]_p^{[y_0]_{p-1}}$ where $[y_0]_{p-1}$ is a solution to the linear equation. In particular, when solvable, the equation $X^d \equiv a \mod p$ has $gcd(n, p-1)$ incongruent solutions modulo p. $\square$

Example 9.4.1

9.4.2 Solving $X^d - a \equiv 0 \mod n$ When $\gcd(d, \varphi(n)) = 1$

The situation in the title of this subsection is particularly easy to handle (and needs no theory of indices!). As we know (from the theory of indices(!)),there is a unique solution in this case. Here is how it looks like.

Proposition 9.4.6 *Let $\gcd(d, \varphi(n)) = 1$, and $\gcd(a,n) = 1$. Then the (unique) solution to the equation $X^d - a \equiv 0 \mod n$ is given by*

$$x_0 \equiv a^u \mod n \qquad \text{where} \qquad d \cdot u \equiv 1 \mod \varphi(n).$$

Proof *An easy exercise. (Do it!).* $\square$

We will see soon, in chapter 12, how to use the fact just proved for purposes of encoding information. But before that we have to address the important issue of computing powers modulo n. We need this for reasons not only of finding the solution to our equation in the simplest case of this subsection, but also for computing orders modulo n, finding primitive roots whenever they exist, and computing indices.

Exercise 9.13 *Suppose d, n and m are positive integers such that d is the largest relatively prime with $\varphi(n)$ divisor of m. Let $d \cdot u \equiv 1 \mod \varphi(n)$. Prove that, for every $a \in \mathbb{Z}$, the equation $X^m - a \equiv 0 \mod n$ is equivalent to (i.e., has the same solutions as) the equation $X^{m/d} - a^u \equiv 0 \mod n$.*

9.4.3 Powers Modulo n, Successive Squaring

Computing dth roots modulo n is equivalent to finding the solutions to a congruence of the type

$$x^d - a \equiv 0 \mod n.$$

The theory of indices modulo n teaches us in principle how to solve such in the case $\gcd(a, n) = 1$. *The index map,* I_n*,which helps us linearise the congruence above depends on the primitive roots chosen, and on knowing the system of indices for every* $[a]_n \in (\mathbb{Z}/n\mathbb{Z})^\times$. *This in turn is related to raising the primitive roots to all powers from* 0 *to* $\varphi(p^\beta)$ *modulo* p^β*, for odd p, and all the powers* 5^s *for* $0 \leq s \leq 2^{\alpha-2}$ *modulo* 2^α*, for* $\alpha \geq 3$. *We need, obviously, an effective way to compute any power of any number modulo n. One such way is using* **successful squaring**.

The idea of the method is simple. If k is the exponent we want to raise a modulo n, then we express

$$k = i_0 + i_1 \cdot 2 + i_2 \cdot 2^2 + \cdots + i_m \cdot 2^m$$

in base-two system, that is, all the coefficients in that expression are 0 *or* 1. *Then,*

$$a^k = a^{i_0 + i_1 \cdot 2 + i_2 \cdot 2^2 + \cdots + i_m \cdot 2^m} = a^{i_0} \cdot (a^2)^{i_1} \cdots (a^{2^m})^{i_m}.$$

Notice that all the terms in the last product which correspond to coefficients $i_j = 0$ *are equal to* 1, *and hence the expression often may not be as big as it seems to be at a first glance. So that*

$$a^k = \prod_{i_j = 1} a^{2^j},$$

and if we know $a^{2^s} \equiv b_s \mod n$, *then*

$$a^k \equiv \prod_{i_j = 1} b_j \mod n.$$

It is for finding b_j, $j = 1, \ldots m$ *where we use the successful squaring: notice that, for* $j \geq 1$,

$$b_{j+1} \equiv b_j a \mod n \qquad \text{where} \qquad b_1 \equiv a \mod n.$$

This recurrence process, together with $|a|, |b_j| \leq n/2$ *which we can assume W.L.O.G., allows a quick computation of the numbers* b_j.

Example 9.4.2

Chapter 10

Sums of Two Squares

This chapter is devoted to answering the question about which natural numbers are representable as sums of two squares. A motivation to study this question comes in particular from what we know about Pythagorean triplets. Recall that these are triplets (a, b, c) of natural numbers such that $a^2+b^2 = c^2$. In the beginning of this course we proved that all such can be produced by using primitive Pythagorean triplets. For these the integers a, b and c are pairwise relatively prime natural numbers. For such a triplet (a, b, c), as we know, there are two odd, relatively prime natural numbers $u > v$ such that

$$c = \frac{u^2 + v^2}{2}.$$

One interesting question that arises here is: which natural numbers can play the role of c in a Pythagorean triplet? Obviously, this question reduces to the case of primitive Pythagorean triplets, which in turn is equivalent to asking which (even) numbers are sums of the squares of two relatively prime integers?
Another interesting question is how many Pythagorean triplets have third component equal to a fixed c?
These questions can be restated as

(1) For what n does the Diophantine equation

$$X^2 + Y^2 = n$$

have solutions?

(2) When solutions exist, how many are they?

These are the questions we answer in this Chapter.

10.1 Primes Representable as a Sum of Two Squares

The prime number 2 is obviously a sum of the squares of two natural numbers. The interesting case is when the prime number is odd. We are proving here the fundamental theorem due to Pierre de Fermat that every prime number type $1 \mod 4$ is a sum of squares of two natural numbers. By necessity, these two numbers are relatively prime, and are unique up to reordering.
The theorem is proved by using the method of descent invented by Fermat to study Diophantine equations. The realization of this method in our proof below can be summarized as follows. We show first that a positive multiple mp of a prime p of type $1 \mod 4$ is a sum of the squares of two natural numbers: $mp = a^2 + b^2$. Then we prove that if $m > 1$, there is a positive natural number $m' < m$ such that $m'p$ is also a sum of the squares of two natural numbers. If $m' > 1$, we can repeat the process finding a descending sequence of positive natural numbers $m > m' > \cdots$. This chain cannot descend ad infinitum, and the only way for it to be finite is that some of the natural numbers m' we get in the process is 1. Therefore p itself is a sum of the squares of two natural numbers.

Theorem 10.1.1 (*Fermat's theorem on primes presentable as sums of two squares*) *An odd prime number p is a sum of the squares of two natural numbers if, and only if, $p \equiv 1 \mod 4$. The presentation is unique, up to reordering of the summands.*

Proof *If $p = a^2 + b^2$, then $a^2 + b^2 \equiv 0 \mod p$, and therefore $a^2 \equiv -b^2 \mod p$. Both a and b are relatively prime with p, so we can write the relation of Legendre symbols*

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

which immediately implies that

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

We know that last relation holds true only if $p \equiv 1 \mod 4$. So

$$p = a^2 + b^2 \quad \Rightarrow \quad p \equiv 1 \mod 4.$$

Assume now that $p \equiv 1 \mod 4$. We want to show that $p = a^2 + b^2$ for some natural numbers a and b. We will prove this is true by using induction (in the form of the least element property for non-empty subsets of $\mathbb{N}$). To this end, consider the set

$$\Sigma := \{m \in \mathbb{N} \,|\, (\exists A, B \in \mathbb{N})(mp = A^2 + B^2)\} \subseteq \mathbb{N}.$$

Claim 1 *The set Σ is non- empty.*

Indeed, the congruence $x^2 \equiv -1 \mod p$ has a solution x_0. So that $x_0^2 + 1 = m_0 p$, and hence $m_0 \in \Sigma$. **Notice** *that x_0 can be taken to be small: $|x_0| < p/2$. That's true, because we need a solution* **modulo** *p.*

The claim ensures that Σ has a least element, $s_0 \geq 1$. **Notice** *that $s_0 \leq x_0$, so $\gcd(s_0, p) = 1$. We want to show that*

Claim 2 $s_0 = 1$.

The proof of this claim is the heart of the argument due to Fermat to prove the theorem. It is known as **the method of infinite descent**. *Here is how it goes. Assuming that $s_0 > 1$, we, following Fermat, will construct an $s' \in \Sigma$ such that $s' < s_0$. Then the double inequality $s_0 \leq s' < s_0$ provides the contradiction needed to finish the proof.*
We have

$$s_0 p = A_0^2 + B_0^2.$$

Let A_1 and B_1 be such that

$$A_1 \equiv A_0 \mod s_0, \qquad B_1 \equiv B_0 \mod s_0, \qquad |A_1|, |B_1| \leq \frac{s_0}{2}.$$

We have that $A_1^2 + B_1^2 \leq (s_0)^2/4$, and that

$$A_0 A_1 + B_0 B_1 \equiv 0 \mod s_0, \qquad A_0 B_1 - A_1 B_0 \equiv 0 \mod s_0.$$

Therefore

$$s_0 p(A_1^2 + B_1^2) = (A_0^2 + B_0^2)(A_1^2 + B_1^2) = (A_0 A_1 + B_0 B_1)^2 + (A_0 B_1 - A_1 B_0)^2.$$

The rightmost sum is divisible by s_0^2, because each summand is. So, we have that

$$\frac{s_0 p(A_1^2 + B_1^2)}{s_0^2} = \left(\frac{A_0 A_1 + B_0 B_1}{s_0}\right)^2 + \left(\frac{A_0 B_1 - A_1 B_0}{s_0}\right)^2.$$

So, $s_0 \mid p(A_1^2 + B_1^2)$*, and, since* $\gcd(s_0, p) = 1$*, we have that* $s' = (A_1^2 + B_1^2)/s_0$ *is an integer. But we also have that* $(A_1^2 + B_1^2)/s_0 \le s_0/4 < s_0$*, and that*

$$s'p = \left(\frac{A_0A_1 + B_0B_1}{s_0}\right)^2 + \left(\frac{A_0B_1 - A_1B_0}{s_0}\right)^2.$$

The natural number s' *is the promised one that leads to a contradiction. Our assumption,* $s_0 > 1$*, was wrong, so there are natural numbers* a, b *such that* $p = a^2 + b^2$*.* □

As a matter of fact, Fermat applied his method of descent to prove his theorem differently. As he wrote to his friend Carcavi in 1650, he uses the method to show that if a prime number of type $1 \mod 4$ *is not a sum of two squares, then there is a smaller prime number of the same type which is not a sum of two squares either. Obviously, this process can be repeated infinitely many times, which, on the other hand, is impossible, because there is no infinite strictly descending sequence of natural numbers. This contradiction proves that there shouldn't be a prime* $1 \mod 4$ *which is not a sum of two squares. The following exercises lead to proving Theorem 10.1.1 in the spirit of Fermat.*

Exercise 10.1 *(1) Let* $a, b, u, v \in \mathbb{Z}$*, and* q *be a [prime number. Suppose* $\gcd(a, b) = \gcd(u, v) = 1$*, and* $q \mid a^2 + b^2$ *and* $q \mid u^2 + v^2$*. Prove that either* $q \mid ua + bv$ *and* $q \mid va - ub$ *or* $q \mid ua = vb$ *and* $q \mid va + ub$*.*

[Hint: Notice that $a^2 + b^2 \equiv 0 \mod q$ *and* $u^2 + v^2 \equiv 0 \mod q$ *imply that*

$$(a/b)^2 \equiv -1 \equiv (u/v)^2 \mod q.$$

Therefore, either $a/b \equiv u/vmodq$ *or* $a/b \equiv -u/v \mod q$*. That is, either* $va - ub \equiv 0 \mod q$ *or* $va + ub \equiv 0 \mod q$*. On the other hand, we have* $(a^2 + b^2)(u^2 + v^2) \equiv 0 \mod q$ *as well, which, combined with* $(a^2 + b^2)(u^2 + v^2) = (ua + vb)^2 + (va - ub)^2 = (ua - vb)^2 + (va + ub)^2$*, gives the result.]*

(2) Suppose $m = a^2 + b^2$ *where* $\gcd(a, b) = 1$*, the prime* q *divides* m*, and* $q = u^2 + v^2$*. Show that* m/q *is a sum of two squares as well.*

[Hint: Use the identities

$$\frac{m}{q} = \frac{a^2 + b^2}{q} = \frac{q(a^2 + b^2)}{q^2} = \frac{(u^2 + v^2)(a^2 + b^2)}{q^2}$$

$$= \frac{(ua + bv)^2 + (va - ub)^2}{q^2} = \frac{(ua - vb)^2 + (va + ub)^2}{q^2}$$

and so

$$\frac{m}{q} = \left(\frac{ua + vb}{q}\right)^2 + \left(\frac{va - ub}{q}\right)^2 = \left(\frac{ua - vb}{q}\right)^2 + \left(\frac{va + ub}{q}\right)^2.$$

Then, use the previous exercise to get the result.]

(3) Suppose p *is a prime number of type* $1 \mod 4$ *which is not a sum of squares of two numbers. Prove that there is an integer* m *such that* $4 \nmid m$ *and* $mp = a^2 + b^2$ *with* $\gcd(a, b) = 1$*.*

[Hint: Use ideas from the proof of Theorem 10.1.1.]

(4) Prove that if $2n = a^2 + b^2$ *where* $\gcd(a, b) = 1$*, then* $n = c^2 + d^2$ *for* $\gcd(c, d) = 1$*.*

[Hint: Since a *and* b *have the same parity (why?), there are both odd and distinct. Observe then that* $4n = (1 + 1)(a^2 + b^2) = (a + b)^2 + (a - b)^2$ *and both* $a + b$ *and* $a - b$ *are even. Finish the proof showing that* $\gcd((a + b)/2, (a - b)/2) = 1$*.]*

(5) Suppose $p \equiv 1 \mod 4$ *is a prime number which is not a sum of squares of two integers. Prove that there is a prime number* $q \equiv 1 \mod 4$ *which is not a sum of squares of two integers, and such that* $q < p$*. Conclude that every prime* $p \equiv 1 \mod 4$ *is a sum of squares of two integers.*

[Hint: Use (3) and (4) to show that there is an odd integer $m > 1$ *such that* $mp = a^2 + b^2$ *with* $\gcd(a, b) = 1$*. W.L.O.G. we may assume that* m *is the least such integer. Show that* $m < p$*. Let* q *be a prime divisor of* m*. Show that* $q < p$*, and that, by (1),* q *is not a sum of squares of two integers.]*

10.2 Natural Numbers Which are Sums of Two Squares

Let $m = A^2 + B^2$ for non-zero integers A and B. If $d = \gcd(A, B)$, then $A = dA_1, B = dB_1$, $\gcd(A_1, B_1) = 1$ and

$$m = d^2(A_1^2 + B_1^2).$$

This means that $m = d^2 m_1$ such that $m_1 = A_1^2 + B_1^2$ is a sum of the squares of two **relatively prime** *natural numbers.*

Claim *If m_1 is even, then m_1 is not divisible by 4. Every odd divisor p of m_1 satisfies $p \equiv 1 \mod 4$.*

Proof *Do that as an* **exercise**. □

We get from the claim that if $m = A^2 + B^2$ for natural numbers A, B, then the exponents of the primes of type "3-mod-4" in the canonical decomposition of m as a product of powers of primes, are all even. In particular

Corollary 10.2.1 *If $m = A^2 + B^2$ where $A, B \in \mathbb{N}$, and $\gcd(A, B) = 1$, then either $m = 2$, or $m = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$, or $m = 2p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$, where $p_1 < p_2 < \cdots < p_k$ and $p_i \equiv 1 \mod 4$ for every $i = 1, 2, \ldots, k$.*

The converse to the statement in the corollary is also true.

Theorem 10.2.2 *If $m = 2$, or $m = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$, or $m = 2p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$ where $p_1 < p_2 < \cdots < p_k$ and $p_i \equiv 1 \mod 4$ for every $i = 1, 2, \ldots, k$, then $m = A^2 + B^2$ for natural numbers A, B such that $\gcd(A, B) = 1$.*

Proof *The proof is done in two steps: (1) $p \equiv 1 \mod 4$ and $k \in \mathbb{N}$, then $p^k = a_k^2 + b_k^2$ with $\gcd(a_k, b_k) = 1$; (2) $M = a^2 + b^2$, $N = c^2 + d^2$ with $\gcd(a, b) = \gcd(c, d) = \gcd(M, N) = 1$, then $MN = u^2 + v^2$ with $\gcd(u, v) = 1$. From these two steps the theorem follows in an obvious way.*

Claim 1 *Let $p \equiv 1 \mod 4$ and $k \in \mathbb{N}$. Then $p^k = a_k^2 + b_k^2$ for some $a_k, b_k \in \mathbb{N}$ with $\gcd(a_k, b_k) = 1$.*

Proof *Induction on $k \geq 1$. Base case: $k = 1$. We have $p = a_1^2 + b_1^2$, because of the fundamental theorem of Fermat. If $d = \gcd(a_1, b_1)$, then $d^2 \mid p$ so that $d = 1$. The base case is verified. Assume now that for some $k \in \mathbb{N}$*

$$p^k = a_k^2 + b_k^2, \qquad \gcd(a_k, b_k) = 1.$$

We have the following obvious identities

$$\begin{aligned} p^{k+1} = p^k p &= (a_k^2 + b_k^2)(a_1^2 + b_1^2) \\ &= (a_k a_1 + b_k b_1)^2 + (a_k b_1 - a_1 b_k)^2 \\ &= (a_k a_1 - b_k b_1)^2 + (a_k b_1 + a_1 b_k)^2. \end{aligned}$$

We are proving next that at least the pair of one of the last two lines consists of relatively prime numbers. That is

$$d_1 = \gcd(a_k a_1 + b_k b_1, a_k b_1 - a_1 b_k) = 1 \quad \vee \quad d_2 = \gcd(a_k a_1 - b_k b_1, a_k b_1 + a_1 b_k) = 1.$$

Assume, by way of contradiction, that $d_1 > 1$ and $d_2 > 1$. Since $d_i^2 \mid p^{k+1}$, each of the summands is divisible by p. In particular

$$a_k a_1 + b_k b_1 \equiv 0 \mod p \qquad \text{and} \qquad a_k a_1 - b_k b_1 \equiv 0 \mod p.$$

But then, $2a_k a_1 \equiv 0 \mod p$ *which, for* p *is odd implies* $p \,|\, a_k a_1$*. Since* $a_1 < p$*, we must have* $p \,|\, a_k$*. By the identity* $p^k = a_k^2 + b_k^2$*, we get that* $p \,|\, b_k$ *as well, so that* $\gcd(a_k, b_k) > 1$ *- a contradiction!* □

Claim 2 *Let* $M = a^2 + b^2$*,* $N = c^2 + d^2$ *with* $\gcd(a, b) = \gcd(c, d) = \gcd(M, N) = 1$*. Then,* $MN = u^2 + v^2$ *with* $\gcd(u, v) = 1$*.*

Proof *We again use the identities*

$$MN = (a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2 = (ac - bd)^2 + (ad + bc)^2.$$

Denote $d_1 = \gcd(ac + bd, ad - bc)$ *and* $d_2 = \gcd(ac - bd, ad + bc)$*. We are showing next that* $d_1 = d_2 = 1$*. Arguing by contradiction and W.L.O.G., assume that there is a prime number* $p \,|\, d_1$*. Since* $d_1^2 \,|\, MN$*, then* $p \,|\, MN$ *and therefore* p *divides either* M *or* N*. Suppose* $p \,|\, M$*. Recall that* $\gcd(M, N) = 1$*. Therefore* p **doesn't divide** N*.*
We have

$$ac + bd \equiv 0 \mod p \qquad ad - bc \equiv 0 \mod p$$

so that

$$c(ac + bd) + d(ad - bc) \equiv 0 \equiv d(ac + bd) - c(ad - bc) \mod p$$

and therefore

$$a(c^2 + d^2) \equiv 0 \equiv b(d^2 + c^2) \mod p.$$

Canceling out the factor $c^2 + d^2$*, we get*

$$a \equiv 0 \equiv b \mod p$$

which implies that $\gcd(a, b) > 1$ *- a contradiction!*
The case when $p \,|\, N$ *instead is dealt with the same way. Claim 2 is proved.* □

Corollary 10.2.3 *The odd natural number* c *can be the third component of a primitive Pythagorean triplet if, and only if, all prime divisors of* c *are of type "*1*-mod-*4*".*

We are in a position now to describe all natural numbers which are sums of the squares of two natural numbers.

Theorem 10.2.4 *We have* $(\exists A, B \in \mathbb{N})(n = A^2 + B^2)$ *if, and only if,* $n = d^2 \cdot m$ *where* m *is either* 2*, or has prime divisors of type* 1*-mod-*4 *only, or is twice such a number. The numbers* A, B *are relatively prime if, and only if, the number* d *can be taken to be* 1*.*

10.3 Number of Presentations as a Sum of Two Squares

We are addressing in this section the question on how many solutions the equation

$$X^2 + Y^2 = n$$

has.

10.3.1 Presentations as Sums of Squares of Relatively Prime Integers

We are finding the number of presentations of a natural number n *as a sum of squares of two relatively prime integers*

$$n = x^2 + y^2, \qquad \gcd(x, y) = 1.$$

Since the answer is obvious for $n = 0$*, and* $n = 1$*, we are assuming in what follows that* $n \geq 2$*. Notice that in this case* $xy \neq 0$*, and if* $n > 2$*, we have also that* $x \neq y$*.*
Consider the set

$$\Sigma_n = \{(x, y) \in \mathbb{N} \times \mathbb{N} \,|\, \gcd(x, y) = 1 \ \wedge \ x^2 + y^2 = n\}.$$

Denote $N(n) = |\Sigma_n|$.
Notice that, for every $(x, y) \in \Sigma_n$, *since* $\gcd(x, y) = 1$, *we have* $x \neq y$. *Therefore, if* $(x, y) \in \Sigma_n$, *then the set of four distinct elements*

$$\{(x, y), (x, -y), (-x, y), (-x, -y)\}$$

consists of solutions to the equation above. So the number of all solutions is $4N(n)$.

Suppose now $(x, y) \in \Sigma_n$. *Since* $x^2 + y^2 = n$, *we have that*

$$x^2 + y^2 \equiv 0 \mod n,$$

and , since $\gcd(y, n) = 1$, *we have* $x/y \mod n$ *is a solution to the equation*

$$Z^2 + 1 \equiv 0 \mod n.$$

Recall that $x/y \mod n$ *stands for the class* $[x]_n[y]_n^{-1} \in \mathbb{Z}/n\mathbb{Z}$.

Theorem 10.3.1 *We have that* $N(n)$ *is the number of solution to* $Z^2 + 1 \equiv 0 \mod n$.

Proof *Denote by* Λ_n *the solutions to* $Z^2 + 1 \equiv 0 \mod n$. *We have a map*

$$h : \Sigma_n \to \Lambda_n \qquad (x, y) \mapsto h((x, y)) = x/y \mod n.$$

We will show that h *is a bijection.*

h **is an injection**. *Let* (x, y) *and* (x', y') *be elements of* Σ_n *such that* $h((x, y)) = h((x', y'))$. *This means that* $x/y \equiv x'/y' \mod n$ *which in turn means that*

$$xy' - x'y \equiv 0 \mod n.$$

So, $xy' - x'y = nk$ *for some* $k \in \mathbb{N}$ *(assuming, obviously W.L.O.G., that* $xy' \geq x'y$). *We have*

$$n^2 = (x^2 + y^2)((x')^2 + (y')^2) = (xy' - x'y)^2 + (xx' + yy')^2 = k^2n^2 + (xx' + yy')^2,$$

and therefore $n^2 \mid (xx' + yy')^2$, *that is,* $xx' + yy' = nl$ *for some* $l \in \mathbb{N}$. *Returning to the computation in the previous line, we get*

$$n^2 = k^2n^2 + l^2n^2$$

which immediately implies that either (i) $k = 1, l = 0$, *or (ii)* $k = 0, l = 1$. *In case (i) we get*

$$xy' - x'y = 1 \qquad \text{and} \qquad xx' + yy' = 0.$$

so, we obviously have $x \mid yy'$ *and* $y \mid x'x$. *This together with* $\gcd(x, y) = 1$ *implies that* $x \mid y'$ *and* $y \mid x'$. *That is* $y' = d_1x$ *and* $x' = d_2y$ *for some* $d_1, d_2 \in \mathbb{N}$. *In a similar way we get* $y' \mid x$ *and* $x' \mid y$, *that is* $x = d'y'$ *and* $y = d''x'$ *for some* $d', d'' \in \mathbb{N}$. *Since* $n \geq 2$, *none of the integers* x, x', y *and* y' *is zero. Therefore,* $d_1d' = d_2d'' = 1$ *which implies that* $d_1 = d' = d_2 = d'' = 1$. *So,* $(x, y) = (y', x')$. *But then*

$$0 = xx' + yy' = x^2 + y^2 = n$$

a contradiction! So, case (i) is not possible. For case (ii) we have in a similar way that

$$x \mid x' \quad y \mid y' \quad x' \mid x \quad y' \mid y$$

and so $(x, y) = (x', y')$.

h **is a surjection**. *Let* $[a]_n \in \Lambda_n$. *We want to prove that there is* $(x, y) \in \Sigma_n$ *such that*

$$x/y \equiv a \mod n.$$

Let $n = 2^\alpha p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ *be the canonical presentation of* n *with* $\alpha \in \mathbb{N}$ *and* $\alpha_i \in \mathbb{N}_{>0}$. *The relation* $a^2 + 1 \equiv 0 \mod n$ *implies that* $\alpha \leq 1$ *and that* $p_i \equiv 1 \mod 4$ *for all* $i = 1, \ldots, k$. *We know from*

the previous section that the numbers 2 *and* p^l*, for* $p > 2$*, are sums of squares of relatively prime numbers:* $2 = 1^2 + 1^2$*, and* $p^l = s^2 + t^2$*. From Chapters* 7 *and* 8 *we know that the equation*

$$Z^2 + 1 \equiv 0 \mod p^l,$$

when $p \equiv 1 \mod 4$*, has only two solutions, and that they are reciprocal to each other modulo* p^l *(verify that as an exercise!). So, if* $a^2 + 1 \equiv 0 \mod p^l$ *and* $s^2 + t^2 = p^l$*, then either* $a \equiv s/t \mod p^l$ *or* $a \equiv t/s \mod p^l$*. As a result we get that for every* $i = 1, \ldots, k$ *there are* (s_i, t_i) *such that*

$$s_i^2 + t_i^2 = p_i^{\alpha_i}, \qquad (s_i, t_i) = 1, \qquad s_i/t_i \equiv a \mod p_i^{\alpha_i}.$$

The following exercise implies, using simple induction, that we have also

$$n = p_1^{\alpha_1} \cdots p_k^{\alpha_k} = (s_1^2 + t_1^2) \cdots (s_k^2 + t_k^2) = s^2 + t^2$$

and

$$n = 2p_1^{\alpha_1} \cdots p_k^{\alpha_k} = (1^2 + 1^2)(s_1^2 + t_1^2) \cdots (s_k^2 + t_k^2) = (s')^2 + (t')^2$$

with

$$\gcd(s, t) = 1 \qquad s/t \equiv a \mod n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$$

and

$$\gcd(s', t') = 1 \qquad s'/t' \equiv a \mod n = 2p_1^{\alpha_1} \cdots p_k^{\alpha_k}.$$

So, $h((s, t)) = [a]_n$ *when* n *is odd, and* $h((s', t')) = [a]_n$ *when* n *is even. The theorem is proved.* □

Exercise 10.2 *Let* n_1 *and* n_2 *be relatively prime positive integers, and suppose the pairs of integers* (s_1, t_1) *and* (s_2, t_2) *are such that*

$$\gcd(s_i, t_i) = 1, \quad s_i^2 + t_i^2 = n_i, \quad s_i/t_i \equiv a \mod n_1, \quad i = 1, 2.$$

Then

$$n_1 n_2 = (s_1^2 + t_1^2)(s_2^2 + t_2^2) = s^2 + t^2, \quad \gcd(s, t) = 1, \quad s/t \equiv a \mod n_1 n_2.$$

Corollary 10.3.2 *The natural number* $n > 1$ *is representable as a sum of the squares of two relatively prime numbers if, and only if, either* n *or* $n/2$ *is an odd number and all odd prime divisors of* n *are* $1 \mod 4$*. In such a case, if* k *is the number of the distinct odd prime divisors of* n*, then* n *has* $4 \cdot N(n) = 2^{k+2}$ *distinct presentations as a sum of the squares of two relatively prime integers.*

Proof *This follows directly from the fact about the number of solutions to* $Z^2 + 1 \equiv 0 \mod n$*.* □

Exercise 10.3 *(1) Let* c *can be the third component of a primitive Pythagorean triplet. How many are the primitive Pythagorean triplets the third component of which is* c*?*
(2) We know, from Chapter 2, that the equations $x^2 + y^2 = z^2$ *and* $x^2 + y^2 = 2z^2$ *do have solutions in positive integers. We also know that* $x^2 + y^2 = 3z^2$ *has no such solutions. Make a conjecture describing the positive integers* n *for which the equation* $x^2 + y^2 = nz^2$ *has solutions in positive integers. Prove your conjecture. Find all solutions to such an equation.*

10.3.2 Presentations as Sums of Squares of Two Integers

Suppose that $n > 1$ *is representable as a sum of two non-zero squares*

$$n = x^2 + y^2$$

and let $d = \gcd(x, y) > 1$*. Then* $x = d \cdot x_1, y = d \cdot y_1$ *with* $\gcd(x_1, y_1) = 1$*, and*

$$n = d^2(x_1^2 + y_1^2).$$

If $n = (x')^2 + (y')^2$ is another presentation with $d' = \gcd(x', y')$, and $x' = d' \cdot x'_1,\ y' = d' \cdot y'_1$, then

$$\text{either} \quad d \neq d' \quad \text{or} \quad (x_1, y_1) \neq (x'_1, y'_1).$$

This implies that the number of presentations of n as a sum of the squares of two integers is equal to the sum

$$\tilde{N}(n) = \sum_{d^2 \mid n,\ d^2 < n} N(n/d^2).$$

Let

$$n = 2^\alpha p_1^{\alpha_1} \cdots p_k^{\alpha_k} q_1^{\beta_1} \cdots q_l^{\beta_l}$$

be the canonical presentation of $n > 1$ with $\alpha, \alpha_i, \beta_j \geq 0$, the primes $p_i \equiv 1 \mod 4$, and the primes $q_j \equiv 3 \mod 4$. Then

$$\tilde{N}(n) \neq 0 \quad \Leftrightarrow \quad \beta_1, \ldots, \beta_l \in 2\mathbb{N}$$

and

$$\tilde{N}(n) = \tilde{N}(2^\alpha p_1^{\alpha_1} \cdots p_k^{\alpha_k}) = \tilde{N}(p_1^{\alpha_1} \cdots p_k^{\alpha_k}),$$

where, to include the case when $\alpha = \alpha_1 = \cdots = \alpha_k = 0$, we set $\tilde{N}(1) = 1$. (Verify the last line as an exercise!)

Exercise 10.4 *(1) Prove that $\tilde{N}(p_1^{\alpha_1}) = (\alpha_1 + 1)$.*
(2) Prove that $\tilde{N}(p_1^{\alpha_1} p_2^{\alpha_2}) = (\alpha_1 + 1)(\alpha_2 + 1)$.
(3) Make a conjecture about a formula computing $\tilde{N}(p_1^{\alpha_1} \cdots p_k^{\alpha_k})$, and prove it.
(4) Prove that if we consider presentations $x^2 + y^2 = n$ and $y^2 + x^2 = n$ as the same, then the number of presentations of n as a sum of two squares is equal to $\lfloor (\tilde{N}(n) + 1)/2 \rfloor$.
(5) With the identification of the previous exercise, find the smallest integer which has 2 different presentations as a sum of squares of two positive numbers. Do the same for 3 and 4 instead of 2 as well.

10.4 An Application of the Method of Descent

We are closing this chapter by illustrating the power of the Method of Descent on treating some Diophantine equations of higher degree. We used the method to prove **existence** *results. But the method can be quite effective in proving* **non-existence** *results as well.*

10.4.1 The Equation $X^4 + Y^4 = Z^4$

The first recorded proof of the following result belongs to Frenicle, one of the correspondents of Fermat. The proof uses the Method of Descent, and it is believed that the idea for it originates with Fermat.

Theorem 10.4.1 *The equation $X^4 + Y^4 = Z^2$ has no solutions in non-zero integers.*

Proof *We proceed by contradiction. Since the signs of the components of a solution do not matter, let's assume that (a, b, c) is a solution with a, b, and c being positive integers. Of all such solutions, there is one with smallest c. We will show, using the descent method, that there is another such solution with even smaller third component which will be the needed contradiction.*

(1) Let's observe first that the numbers a, b, and c are pairwise relatively prime. Indeed, if a prime p divides a and b, then p^2 divides c, and $(a/p, b/p, c/p^2)$ is a solution to the equation. This is impossible though, because, by the choice of (a, b, c), the inequality $a \leq a/p$ should be satisfied, which it is not. If we assume that p divides a and c, say, than p should divide b as well, and as we know this leads to an absurd. So, (a, b, c) is a primitive triplet of positive numbers.

(2) Observe now that (a^2, b^2, c) *is a primitive Pythagorean triplet. Assuming* a^2 *is even, then there are odd positive integers* $u > v$ *with* $\gcd(u, v) = 1$ *such that*

$$a^2 = \frac{u^2 - v^2}{2}, \quad b^2 = uv, \quad c = \frac{u^2 + v^2}{2}.$$

Since $\gcd(u, v) = 1$*, we get that* $u = s^2$ *and* $v = t^2$*, and so* $u \equiv v \equiv 1 \mod 8$.

(3) The second observation we do is that $\gcd(u - v, u + v) = 2$*, and that* $(u + v)/2$ *is odd. The last fact implies that* $\gcd(u - v, (u + v)/2) = 1$*. But then, by the first equality above, we get that both* $u - v$ *and* $(u + v)/2$ *are squares. Since* $u + v$ *is even, we conclude that*

$$u - v = (2a_1)^2, \qquad (u + v)/2 = a_2^2, \qquad (2a_1)^2 \cdot a_2^2 = a^2.$$

In particular, $u = 2a_1^2 + a_2^2$*, and* $v = -2a_1^2 + a_2^2$*. Therefore*

$$b^2 = uv = -(2a_1^2)^2 + (a_2^2)^2$$

and $(2a_1^2, b, a_2^2)$ *is another primitive Pythagorean triplet. So, there are odd positive integers* $g > h$ *with* $\gcd(g, h) = 1$ *such that*

$$2a_1^2 = (g^2 - h^2)/2, \qquad b = gh, \qquad a_2^2 = (g^2 + h^2)/2.$$

(4) Since $\gcd(g, h) = 1$ *and both* g *and* h *are odd, we have that* $\gcd(g - h, g + h) = 2$*, and so* $\gcd((g - h)/2, (g + h)/2) = 1$*. Since* $(g - h)/2 \cdot (g + h)/2 = a_1^2$ *we get that*

$$g - h = 2P^2, \qquad g + h = 2Q^2$$

for positive integers P, Q *with* $\gcd(P, Q) = 1$*. Computing from here that*

$$g = P^2 + Q^2, \qquad h = P^2 - Q^2$$

and therefore

$$a_2^2 = (g^2 + h^2)/2 = P^4 + Q^2.$$

So, finally, we got a new solution (P, Q, a_2) *to the original equation. By our assumption about* (a, b, c) *we need to have* $c \leq a_2$*. On the other hand,* $a_2 < a < \sqrt{c} < c$*. Therefore, we have* $a_2 < c \leq a_2$ *which is impossible.* □

Corollary 10.4.2 *The equation* $X^4 + Y^4 = Z^4$ *has no solutions in integers with* $XYZ \neq 0$.

The last equation is interesting, because it is a particular case of the Fermat's Last Theorem(FLT) claiming that for $n \geq 3$ *the equation*

$$X^n + Y^n = Z^n$$

has no solutions in non-zero integers. The result proved above, for $n = 4$*, settles the FLT for all* n *divisible by* 4*. For the rest of cases for* n*, that is* $n > 2$ *and* $4 \nmid n$*, there is an odd prime* p *which divides* n*. Obviously, to prove FLT in general, one can restrict themselves to considering* $n = p$ *an odd prime number. An (incomplete) proof of this claim for* $n = 3$ *was given, for the first time, by Euler in 1770. It was based on the Method of Descent. The case* $n = 5$ *was dealt with by Dirichlet and Legendre in 1825, and the case* $n = 7$ *in 1839 by Lamé. All proofs were done using the Method of Descent. An attempt to prove FLT in general lead Ernst Kummer to develop the theory of ideal numbers (later on evolving into the theory of ideals in Dedekind's work). Using his theory, and the Method of Descent, Kummer managed to prove FLT for (almost) all primes less than* 100*. This was the greatest achievement in this direction in the Ninetieth Century. The history of attempts to prove FLT is long and rich (encompassing 350 years!). Many prominent mathematicians tried to do find a proof, and arguably it was the proof of FLT which was often the driving force behind the development of modern Algebra. For the statement of FLT is purely algebraic, people expected that algebraic methods solely would be enough to handle it. Ironically, the only known proof today is based on highly non-trivial geometric ideas. This proof was achieved in 1993-1994 by A. Wiles with the help of R. Taylor. The proof is more than 200 pages long, and uses methods of Algebraic Geometry.*

Using the Method of Descent we are proving next the following.

Proposition 10.4.3 *The equation $X^4 + Y^2 = Z^4$ has no solution in non-zero integers.*

Proof *Using contradiction, assume that (a, b, c) is a solution with positive components and with smallest possible c. It is straightforward that the triplet should be primitive, and that c should be odd. Consider two cases: b even, and b odd.*

(1) Assume b is even. Then a and c are odd, and relatively prime. It is immediate to see that $c^2 + a^2, c + a$ and $c - a$ pairwise share only 2 as a divisor. Since $c^4 - a^4 = (c^2 + a^2)(c + a)(c - a) = b^2$, we have

$$c^2 + a^2 = 2u^2, \quad c + a = 2v^2, \quad c - a = 2w^2$$

for pairwise relatively prime u, v, and w. We easily compute from the last two equalities that $c = v^2 + w^2$ and $a = v^2 - w^2$, and , after substituting in the first equality we get $v^4 + w^4 = u^2$ which, by the theorem above is impossible!

(2) Assume b is odd. Obviously (a^2, b, c^2) is a primitive Pythagorean triplet, with odd b. There are therefore odd positive integers $u > v$ with $\gcd(u, v) = 1$ and such that

$$a^2 = (u^2 - v^2)/2, \quad b = uv, \quad c^2 = (u^2 + v^2)/2.$$

Since $\gcd((u - v)/2, (u + v)/2) = 1$ we have that either

$$(u - v)/2 = 2a_1^2, \qquad (u + v)/2 = a_2^2$$

or

$$(u - v)/2 = a_1^2, \qquad (u + v)/2 = 2a_2^2.$$

In the first case $c^2 = 4a_1^4 + a_2^4$, while in the second $c^2 = a_1^4 + 4a_2^4$. Both cases are treated the same way, so we'll consider the first one only. There are positive odd integers $s > t$ with $\gcd(s, t) = 1$ such that

$$2a_1^2 = (s^2 - t^2)/2, \qquad a_2^2 = st, \qquad c = (s^2 + t^2)/2.$$

From this we get that $s = P^2$ and $t = Q^2$ with $\gcd(P, Q) = 1$ and $PQ = a_2$, and that

$$Q^4 + (2a_1)^2 = P^4.$$

So, the triplet $(Q, 2a_1, P)$ is a solution to the original equation, and $P^2 \leq a_2^2 < a^2 < c^2$. This is impossible, because by our assumption on (a, b, c), we should have $c \leq P$. This completes the proof. □

10.4.2 Some Amusing Related Results

There is a series of exercises which are based on the claims in the previous subsection which are quite amusing. Here are some of them.

Exercise 10.5 *(1) Call a right triangle Pythagorean if it has integer side lengths. Prove that at most one side of a Pythagorean triangle can be a perfect square.*
(2) Prove that no side of a Pythagorean triangle can be the hypotenuse of a Pythagorean triangle.
(3) Prove that, over the integers, it is impossible to "square" a right triangle: there is no square of integer side-lengths whose area is equal to the area of a Pythagorean triangle. Prove also that the area of a Pythagorean triangle can not be twice the area of such a square either.

10.4.3 Some More Diophantine Equations

Using the Method of Descent, or the proven facts above, one can study some related Diophantine equations. Here are some popular examples.

Exercise 10.6 *(1) Prove that the equation $X^4 - 4Y^4 = Z^2$ has no non-zero integer solutions. [Hint: Show that if there is a non-zero solution, (a, b, c), then there is a primitive one, with pairwise relatively prime components, and with a and c odd. Use then that $a^2 - c = 2b_1^4$ and that $a^2 + c = 2b_2^4$ with $b = b_1 \cdot b_2$. A second proof is the following. Squaring both sides of $a^4 - 4b^4 = c^2$, and rearranging terms, one gets that $(2ab)^4 + t^4 = (a^4 + 4b^4)^2$...]*
(2) Prove the same for $X^4 - 2Y^2 = 1$ and for $X^2 - 2Y^4 = 1$ [Hint: $X^4 - 2Y^2 = 1$ is equivalent to $X^4 + Y^4 = (Y + 1)^2$. For the second equation, use that 2 is not a difference of two fourth powers of integers.]
(3) Prove that the only non-zero integer solutions to $X^4 - 2Y^2 = -1$ are given by $|X| = |Y| = 1$. What can you say about the solutions to $X^2 - 2Y^4 = -1$?

Chapter 11

Arithmetic Functions; Applications

11.1 Arithmetic Functions

The notion of arithmetic function is very general: **any function with domain $\mathbb{N}$ and co-domain $\mathbb{C}$ is called arithmetic**. *Among all such functions there are some which are very useful in Number Theory. Important examples are given below.*

Definition 11.1.1 *An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is called* **multiplicative**, *if $f \neq 0$, and for every two natural numbers $m, n \in \mathbb{N}$*

$$\gcd(m, n) = 1 \qquad \Rightarrow \qquad f(mn) = f(m)f(n).$$

Exercise 11.1 *Prove that for any multiplicative arithmetic function f we have $f(1) = 1$.*

Most of the functions we will be working with in this course are multiplicative. For instance, the Euler's phi-function is a multiplicative arithmetic function, but some of them are **additive**.

Definition 11.1.2 *An arithmetic function $f : \mathbb{N} \to \mathbb{C}$ is called* **additive** *if for every two natural numbers $m, n \in \mathbb{N}$*

$$\gcd(m, n) = 1 \quad \Rightarrow \quad f(mn) = f(m) + f(n).$$

Since the co-domain of arithmetic functions is a ring, actually - a field, the operations addition and multiplication of such functions is well defined. But for the theory of numbers an operation introduced by P.-G. L. Dirichlet among the arithmetic functions is very important and useful.

Definition 11.1.3 *The* **Dirichlet product**, *$f * g$, of two arithmetic functions, f and g is defined by*

$$(f * g)(n) := \sum_{d \,|\, n} f(d)g(n/d).$$

Exercise 11.2 *(1) Prove that, for every two multiplicative functions f and g, their Dirichlet product $f * g$ is also a multiplicative function.*
(2) Prove that the Dirichlet product is commutative and associative, that is

$$f * g = g * f, \qquad (f * g) * h = f * (g * h).$$

(3) Let f be a multiplicative function. Define the arithmetic function F by

$$F(n) = \sum_{d \,|\, n} f(d).$$

Prove that F is a multiplicative arithmetic function as well.
(4) Suppose f is a multiplicative arithmetic function. Define F by $F(n) = f(n)/n$ for every $n \in \mathbb{N}$. Prove that F is a multiplicative function as well.
(5) Suppose f is an additive real valued arithmetic function, and let $a > 0$ be a real number. Define F by $F(n) = a^{f(n)}$ for all $n \in \mathbb{N}$. Prove that F is a multiplicative arithmetic function. Why do we need a to be a positive real number?
(5') Suppose that f is an additive function with range a subset of $\mathbb{N}$, and let $a \neq 0$ be any **complex** *number. Define f by $f(n) = a^{f(n)}$ for all $n \in \mathbb{N}$. Prove that F is a multiplicative arithmetic function.*

11.2 Important Arithmetic Functions

Here is a list of arithmetic functions that are widely used in Math.

(1) The **Euler's phi-function** $\varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|$

(2) The **Euler's sigma-function** $\sigma(n) = \sum_{d \mid n} d$, *that is the sum of all positive divisors of n*

(2') A generalization of the sigma-function: $\sigma_r(n) := \sum_{d \mid n} d^r$ *where $r \in \mathbb{N}$.*

(3) $\tau(n)$: **the number of positive divisors of** n. *Formally, $\tau(n) = \sigma_0(n)$.*

(4) The **total number of the distinct prime divisors of** $n \geq 2$: $\omega(n) = \sum_{p \mid n} 1$, *and* $\omega(1) = 0$.

(5) The **total number of the positive divisors of** $n \geq 2$ *which are powers of a prime:* $\Omega(n) = \sum_{p^k \mid n} 1$, *and* $\Omega(1) = 0$.

(6) The **kernel of** n, $n \geq 2$, $\gamma(n) = \prod_{p \mid n} p$, *and* $\gamma(1) = 1$.

(7) The **Liouville function** $\lambda(n) = (-1)^{\Omega(n)}$.

(8) The **Möbius function** $\mu(n) = (-1)^{\omega(n)}$ *if* $n = \gamma(n)$, *and* $\mu(n) = 0$ *if* $n \neq \gamma(n)$.

Exercise 11.3 *(1) The functions ω and Ω are additive arithmetic functions. All the rest of the listed functions are multiplicative arithmetic functions. Verify that.*
(2) Show that $n = \gamma(n)$ if, and only if, $\mu^2(n) = 1$.
(3) If $n = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$, then

$$\tau(n) = \prod_{i=1}^{k} (\beta_i + 1) \qquad \sigma(n) = \prod_{i+1}^{k} \frac{p_i^{\beta_i+1} - 1}{p_i - 1}.$$

(4) Prove that, for every natural n,

$$\sum_{d \mid n} \varphi(d) = n.$$

(5) Prove that, for every natural n,

$$\sum_{d \mid n} \frac{\mu(d)}{d} = \prod_{p \mid n} \left(1 - \frac{1}{p}\right),$$

and conclude that

$$\varphi(n) = n \sum_{d \mid n} \frac{\mu(d)}{d}.$$

The last two exercises are notable in the following sense. Denote by I the **identity arithmetic function**, *and by* 1 *-the* **constant function taking on values** 1

$$I : \mathbb{N} \to \mathbb{N} \quad I(n) = n, \qquad 1 : \mathbb{N} \to \mathbb{N} \quad 1(n) = 1.$$

According to Exercise (4)(verify that!),

$$I = 1 * \varphi,$$

and according to Exercise (5) (verify that as well!),

$$\varphi = \mu * I.$$

In other words, we can "solve" for φ, expressing it via I, by using the Möbius function. This example is a particular case of a more general fact.

Theorem 11.2.1 (*Möbius Inversion Formula*) *Let F and f be two arithmetic functions. Then*

$$F(n) = \sum_{d \,|\, n} f(d) \quad \Leftrightarrow \quad f(n) = \sum_{d \,|\, n} \mu(d) F(n/d).$$

In other words,

$$F = 1 * f \quad \Leftrightarrow \quad f = \mu * F.$$

Proof *The theorem follows from the associativity of the Dirichlet product, the fact that*

$$e = 1 * \mu$$

is a function for which $e(1) = 1$ and $e(n) = 0$ for $n > 1$, and that for every arithmetic function f

$$f * e = f \quad \square$$

Exercise 11.4 *(1) Show that*

$$1 * 1 = \tau, \quad I * I = I \cdot \tau, \quad 1 * I = \sigma.$$

(2) Prove that

$$\mu * \tau = 1, \quad \mu * \sigma = I.$$

11.2.1 Groups of Arithmetic Functions

We are demonstrating here the importance of the Dirichlet product for the set of arithmetic functions. Denote by AF_1 the set of all arithmetic functions f such that $f(1) \neq 0$. Notice that all multiplicative arithmetic functions form a subset of AF_1!

Theorem 11.2.2 *The pair $(AF_1, *)$ is a commutative group with identity element the function e defined above. Moreover, the set MAF of multiplicative arithmetic functions, with the Dirichlet product, forms a subgroup of $(AF_1, *)$.*

Proof *The only axioms that remains to be checked in order to prove the first part of the theorem is that*

$$(\forall f \in AF_1)(\exists g \in AF_1)(f * g = e).$$

This latter is done by induction. Namely since $f(1) \neq 0$, we can define $g(1) = 1/f(1)$. Assume that we have defined already $g(1), g(2), \ldots, g(n)$. Lets define $g(n+1)$. We want to have $(f * g)(n+1) = e(n+1) = 0$, so we have

$$\sum_{d \,|\, n+1} f(d) g((n+1)/d) = 0$$

which we solve for $g(n+1)$.
The proof of the second part of the Theorem is straightforward noticing that MAF is closed under the Dirichlet product $*$, the function e is multiplicative, and that the reciprocal of a multiplicative arithmetic function is multiplicative as well. $\square$

11.2.2 The Algebra of Arithmetic Functions

What we actually proved in the previous sub-sections is that the triplet $(AF, +, *)$ is a commutative $\mathbb{C}$-algebra with **group of unit elements** *$AF^{\times} = AF_1$. Moreover, this algebra is local: it has only one maximal ideal: the set of non-invertible (w.r.t. $*$) functions. More about these things - in Topics in Algebraic Structures.*

11.3 Applications

11.3.1 Some Special Numbers

Since ancient times, the people have been attracted to natural numbers with certain specific properties. May be the most attractive among these, with many questions answered and unanswered about them, are the **prime** *numbers. A* **twin prime** *is a prime number p such that either $p-2$ or $p+2$ is prime as well. Other numbers which we have already mentioned are, for instance, the* **triangular** *numbers, that is, numbers*

$$n = 1 + 2 + \cdots + m$$

for some natural m. As we know, the triangular numbers are geometric in nature. Other examples of special numbers are the **perfect** *numbers. These are numbers n which are the sum of their proper, that is, strictly smaller than n, divisors. Recalling the definition of the arithmetic function σ, we see that n* **is perfect if, and only if,** *$n = \sigma(n) - n$. A generalization of perfect numbers is the* **aliquot** *numbers for which the sum of their (proper) divisors are integer multiples of the numbers, that is $\sigma(n) - n = kn$ for some $k \in \mathbb{N}$. The number n is called* **k-perfect** *if $\sigma(n) = kn$. A related to the concept of perfect numbers is the concept of amicable numbers. Two natural numbers $m, n \in \mathbb{N}$ are called* **amicable** *if*

$$\sigma(m) - m = n \quad \wedge \quad \sigma(n) - n = m.$$

Further, a number n of the form $n = F_k = 2^{2^k} + 1$ is called a **Fermat number**, *and a number n of the form $n = M_p = 2^p - 1$ where p is a prime number, is called a* **Mersenne number**. *When F_k and M_p are prime, they are called a* **Fermat prime** *and a* **Mersenne prime** *respectively. Considerations of Fermat and Mersenne numbers are motivated by arithmetic questions.*

Exercise 11.5 *(1) Consider numbers of the type $n = 2^m + 1$. Prove that such an n is prime only if $n = F_k$ for some $k \in \mathbb{N}$.*
(2) Consider numbers of the type $n = a^m - 1$ where $a > 1$ is a natural number. Prove that such an n is prime only if $n = M_p$ for some p.

As a matter of fact though, Fermat primes have a beautiful geometry hidden behind them! This feature of the Fermat primes was revealed by the teenage Gauss, showing that a regular 17-gon can be constructed by using a straight edge and compass.

Many are the (unanswered yet) questions about the mentioned above special numbers. For instance, are there infinitely many twin primes, are there infinitely many Fermat and Mersenne primes, are there odd perfect numbers etc.

In the next subsection, using the properties of the function σ, and following Euclid and Euler, we are showing that the **even** *perfect numbers are as many as are the Mersenne primes.*

11.3.2 Even Perfect Numbers

Theorem 11.3.1 (*Euclid -Euler's theorem*) *An even natural number n is perfect if, and only if, $n = 2^{p-1} \cdot M_p$ where M_p is a Mersenne prime number.*

Proof *The divisors of* $2^{p-1} \cdot M_p$ *are*

$$1, 2, \ldots, 2^{p-1}, M_p, 2 \cdot M_p, \ldots, 2^{p-1} \cdot M_p.$$

So that

$$\sigma(2^{p-1} \cdot M_p) = (M_p + 1) \cdot (1 + 2 + \cdots + 2^{p-1}) = 2^p \cdot (2^p - 1) = 2 \cdot (2^{p-1} \cdot M_p).$$

Let now n *be a perfect even natural number. So,* $n = 2^s \cdot m$ *where* $s \geq 1$ *and* m *is odd, and* $\sigma(n) = 2n$*. Obviously,* $m \neq 1$*. We know that* σ *is a multiplicative arithmetic function. So, we have further that*

$$2^{s+1} \cdot m = \sigma(2^s) \cdot \sigma(m) = (2^{s+1} - 1) \cdot \sigma(m).$$

Since $\gcd(2^{s+1}, 2^{s+1} - 1) = 1$*, it follows that* $\sigma(m) = 2^{s+1} \cdot u$*, and therefore*

$$m = (2^{s+1} - 1) \cdot u.$$

If we assume that $m \neq 2^{s+1} - 1$*, then*

$$2^{s+1} \cdot u = \sigma(m) \geq 1 + u + m = 1 + u + (2^{s+1} - 1) \cdot u = 1 + 2^{s+1} \cdot u$$

so that $2^{s+1} \cdot u \geq 1 + 2^{s+1} \cdot u$ *which is absurd! So,* $m = 2^{s+1} - 1$*,* $u = 1$*, and* $\sigma(m) = 2^{s+1}$*. But that means that* m *has only two divisors:* $2^{s+1} = 1 + (2^{s+1} - 1)$*, and therefore, it is a (Mersenne) prime number,* $m = M_{s+1}$*. But* $s + 1$ *has to be then a prime number:* $s + 1 = p$*, and we get the result:* $n = 2^{p-1} \cdot M_p$*, where* M_p *is prime.* □

The fact that the numbers of type $2^{p-1} \cdot (2^p - 1)$ *with* $2^p - 1$ *prime are perfect was known to Euclid! The fact that those are* **all even** *perfect numbers was proved by Euler. It is not known if the even perfect numbers are infinitely many. No odd perfect number is known so far!*

11.3.3 On Presenting Natural Numbers as Sums of Two Squares

In this subsection, we are showing how using the Legendre and Jacobi symbols one can construct arithmetic functions which have applications in the theory of presenting numbers as sums of squares of two numbers.

A Warm Up: the Function L_p

Recall that the Legendre symbol $\left(\frac{a}{p}\right)$ *is defined for an odd prime number* p *and an integer* a *relatively prime with* p*. We can extend the definition of the symbol to all integers by setting that*

$$\left(\frac{a}{p}\right) = 0 \quad \textit{if} \quad p \,|\, a.$$

This extended function $\left(\frac{\bullet}{p}\right) : \mathbb{N} \to \mathbb{C}$*, with* $Ran(\left(\frac{\bullet}{p}\right)) = \{0, 1, -1\}$*, is* **completely multiplicative***, that is*

$$\forall a, b \in \mathbb{Z} \quad \left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

By what we know from Section 1 of this chapter the function

$$L_p(n) = \sum_{d \,|\, n} \left(\frac{d}{p}\right)$$

is a multiplicative function as well. So, $L_p(1) = 1$*, and if* $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ *is the canonical decomposition of* n*, then*

$$L_p(n) = L_p(p_1^{\alpha_1}) \cdots L_p(p_k^{\alpha_k}).$$

Exercise 11.6 *Prove that $L_p(p^\alpha) = 1$. Suppose q is a prime number different from p. Then*

$$L_p(q^\alpha) = (1 - (-1)^{\alpha+1})/2 \quad \textit{if} \quad \left(\frac{q}{p}\right) = -1,$$

and

$$L_p(q^\alpha) = \alpha + 1 \quad \textit{if} \quad \left(\frac{q}{p}\right) = 1.$$

So, $L_p(n) \neq 0$ if, and only if, the non-residues $\mod p$ among the prime divisors of n enter in odd degree in the canonical decomposition of n.

The Interesting Case: the Function J_P

We want to do a similar thing as in the previous sub-subsection, but with a varying "denominator" of the Legendre symbol. So, we have to use the Jacobi symbol instead.

Recall the the Jacobi symbol $\left(\frac{P}{Q}\right)$, is defined for $(P,Q) = 1$ and Q - an odd number. We naturally extend the definition to relax the first condition:

$$\left(\frac{P}{Q}\right) = 0 \quad \textit{if} \quad \gcd(P,Q) \neq 1.$$

So, for every fixed non-zero integer P we define

$$\left(\frac{P}{\bullet}\right) : \mathbb{N} \setminus \{0\} \to \mathbb{C} \qquad n \mapsto \left(\frac{P}{\bullet}\right)(n) = \left(\frac{P}{n_1}\right)$$

where $n = 2^\beta n_1$ for n_1 an odd natural number. We call n_1 the **odd content of** *n. By definition*

$$\left(\frac{P}{1}\right) = 1.$$

Obviously the function is completely multiplicative, and has $Ran(\left(\frac{P}{\bullet}\right)) = \{0, -1, 1\}$ (verify this as an exercise!). Define the function

$$J_P : \mathbb{N} \setminus \{0\} \to \mathbb{C} \qquad n \mapsto J_P(n) = \sum_{d \,|\, n_1} \left(\frac{P}{\bullet}\right)(d) = \sum_{d \,|\, n,\ d\ odd} \left(\frac{P}{d}\right).$$

It is almost straightforward that it is multiplicative (verify that!). If $n = 2^\alpha p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ is the canonical decomposition of n, then

$$J_P(n) = J_P(p_1^{\alpha_1}) \cdots J_P(p_k^{\alpha_k}).$$

Now, the expected exercise!

Exercise 11.7 *Let q be an odd prime number, and let $\beta \geq 0$. If $q \,|\, P$, then $J_P(q^\beta) = 1$. If $q \nmid P$, then*

$$J_P(q^\beta) = (1 - (-1)^{\beta+1})/2 \quad \textit{if} \quad \left(\frac{P}{q}\right) = -1,$$

and

$$J_P(q^\beta) = \beta + 1 \quad \textit{if} \quad \left(\frac{P}{q}\right) = 1.$$

Example 11.3.1 (1) $J_1(n) = (\alpha_1 + 1) \cdots (\alpha_k + 1)$ equals the number of **odd** divisors of n.
(2) We know that $\left(\frac{-1}{q}\right) = (-1)^{(q-1)/2}$. So, if $n = p_1^{\alpha_1} \cdots p_k^{\alpha_k} q_1^{\beta_1} \cdots q_l^{\beta_l}$ is the canonical decomposition of n with $p_i \equiv 1 \mod 4$, and $q_j \equiv 3 \mod 4$, then $J_{-1}(n) = 0$ if, and only if, there is an odd β_j. Equivalently, $J_{-1}(n) \neq 0$ if, and only if, $\beta_1, \ldots, \beta_l$ are all even. If this is the case, we have

$$J_{-1}(n) = (\alpha_1 + 1) \cdots (\alpha_k + 1). \quad \square$$

Exercise 11.8 *Recalling that, for any odd Q, $\left(\frac{-1}{Q}\right) = (-1)^{(Q-1)/2}$, prove that*

$$\forall n \in \mathbb{N} \setminus \{0\} \quad J_{-1}(n) = D_1 - D_3$$

where D_i, $i = 1, 3$, is the number of $i \mod 4$ divisors of n.

Exercise-Proposition 11.3.4 *Prove that the function $\tilde{N}$ assigning to any $n > 0$ the number of distinct presentations of n as a sum of the squares of two non-negative numbers (see Chapter 10), and the function J_{-1} are identical. Conclude that*

the number of presentations of n as a sum of the squares of two non-negative numbers is equal to the difference of the number of $1 \mod 4$ divisors of n and the number of the $3 \mod 4$ divisors of n.

[Hint: Do the exercises in the end of sect. 10.3.]

Exercise 11.9 *Describe the functions J_2 and J_{-2}.*

Chapter 12

Applications to Designing Cryptosystems

In this Chapter, we are describing some applications of the methods developed in the Lecture Notes. To illustrate these applications on realistic examples, doing all the calculations involved, we would need the help of especially designed software, such as SAGE. Since our course is not computations oriented, we will give here descriptions of the applications in principle referring the interested readers to more specialized literature.

12.1 General Remarks

Transmitting information securely has been an important issue since very ancient times. It has always been absolutely paramount to be able to send messages to friends and/or allies in a way enemy doesn't know the content even when they detect the transmission of message. To achieve this, people often use cryptosystems. These are methods designed to transform a plain-text (such as usual human text) into a cipher-text (encryiption/enciphering process), and then, after transmitting the latter as it is, transform it back to a plain-text (decryption/deciphering process). To make these transformations, one needs a cipher. Ideally, the cipher has to be designed in such a way that encryption/decryption processes are easy to perform, but "breaking" the cipher by adversaries, and therefore recovering the plain-text by unauthorized parties, is hard.
Very often properties of integers have been used to design ciphers. Therefore, the theory of numbers naturally has played an important role in designing cryptosystems.
In securely transmitting the encrypted information, it is important to prevent it from being intercepted to start with. This has always been causing complications in the process of communication. In the late 1970s, techniques from Number Theory, some of which - very simple, made possible transmission of secret messages under the assumption that (almost) all the communication is intercepted and read by the unfriendly party: the **public key cryptosystems** *were developed. We are discussing in what follow some of these systems.*

12.2 Exponential Ciphers

Exponential Ciphers were developed in 1976 by M. Hellman. The idea is very simple. The enciphering goes in three steps :

(1) replace the symbols of the alphabet we are using (the English alphabet with or without additional symbols, for instance) with strings of digits of the same length;
(2) after replacing the symbols of the alphabet used in the plain-text to be transmitted with the strings of digits from (1), and in the order they appear in the plain-text, consider the plain-text as a string of digits itself, and cut it in a sequence of numbers $\{a_1, a_2, \ldots, a_n\}$*, consisting of the same amount*

of digits;
(3) choose a prime number p bigger than all the numbers $a_1, a_2, \ldots, a_n$, and an **enciphering key**, *i.e., a natural number $e < p$ with $\gcd(e, p-1) = 1$, and produce a sequence obtained by replacing every number of the one in (2) with*

$$b_1 \equiv a_1^e \mod p,\ b_2 \equiv a_2^e \mod p,\ \cdots,\ b_n \equiv a_n^e \mod p$$

where for simplicity b_i may be chosen to be the smallest possible.

The transmitted message is the sequence $\{b_1, b_2, \ldots, b_n\}$.

The deciphering process, that is, recovering the string of numbers $\{a_1, a_2, \ldots, a_n\}$ step (1), and eventually the original plain-text, goes as follows.

(i) Find an integer f, the **deciphering key**, *such that $e \cdot f \equiv 1 \mod p-1$*
(ii) Recover a_i as the residue of $b_i^f \mod p$ which belongs to $[0, p)$, for $i = 1, 2, \ldots, n$.

All the exponentiation can be done the way we discussed in Subsection 9.4.3. To make the process secure enough, we have to keep our choices of p and e secret.

A modification of this method which makes it more flexible is the following. Here is how it works in a realistic scenario.
Suppose there are two people, H_1 and H_2, trying to exchange messages. H_1 and H_2 can not meet privately, so everything they communicate to each other should be done by transmitting information through secure channels.
H_1 and H_2 first chose a prime number p, and a positive integer $s < p$ with $\gcd(s, p-1) = 1$ which they communicate to each other.
Choosing an enciphering key (the same for both of them) can be done as follows. Each of H_1 and H_2 chooses a positive integer less than p, say h_1 and h_2, and keep it secret (do not transmit it). Then H_1 sends to H_2 the residue $\mod p$ of s^{h_1}, and H_2 sends to H_1 the residue $\mod p$ of s^{h_2}. The enciphering key is $e \equiv (s^{h_1})^{h_2} \equiv (s^{h_2})^{h_1} \mod p$, which both compute. Notice that $\gcd(e, p-1) = 1$.
To apply the method in steps (1), (2), and (3), the participants have to choose the lengths of numbers on step (2) to be such that the numbers a_i are all less than p.
Notice also that the enciphering key can change from instance to instance of transmitting messages between the participants, which makes the method more secure.

This method was developed by Diffie and Hellman in 1976. The cipher just explained is much more secure than the one before it, because it is quite safe even if s, p, s^{h_1}, and s^{h_2}, the numbers originally transmitted, become known to the unauthorized parties who want to decipher the message! Indeed, the deciphering process depends on the key f. To obtain it, one should know $e \mod p$, so ultimately, one needs to know h_1 and/or h_2. If these are kept really secret, then, assuming s, p, s^{h_1}, and s^{h_2} have been intercepted, one has to try to find them using the values of $s^{h_1} \mod p$ and/or $s_{h_2} \mod p$. When p is small, this can be done using regular computers. But if p is large, which is standardly the case, the computations are so long and heavy, that it would be impossible to crack the enciphered message and react to it in time, the mission will be accomplished before the enemy understands what it is about!

12.3 The RSA Encryption System

The RSA criptosystem is a more flexible alternative of Diffie-Hellman key exchange criptosystem, and is named after its developers Rivest, Shamir, and Adleman in 1978. This system has not one, but many enciphering kays, one for each of the participants in the exchange of (encripted) information. Every participant **publishes** *their key in the public register of such keys. This is how it goes.*

Let H be a person who wants to be receiving messages using the RSA cryptosystem. The person H chooses two distinct prime numbers p and q, and a positive integer s such that $\gcd(s, \varphi(pq)) = 1$. The enciphering key of H is the ordered pair (r, s) where $r = p \cdot q$. H publishes (r, s) in the public registry of such keys.
If a person H', not necessarily having an enciphering key, wants to send securely a message to H, they do the following. First they produce the sequence of integers $\{a_1, a_2, \ldots, a_n\}$ going through steps (1) and (2) from the Diffie-Hellman cryptosystem replacing the symbols of the alphabet with numbers by using an encoding known to H. The sequence to be transmitted as an open text from H' to H is $\{b_1, b_2, \ldots, b_n\}$ where $b_i \equiv a_i^r \mod r$.
Deciphering the information is done by H by first choosing a deciphering key, that is, a positive number f such that $s \cdot f \equiv 1 \mod \varphi(r)$ (which can be done solely by H), and following steps (i) and (ii) from the Diffile-Hellman cryptosystem.
If H wants to send a message to H' safely, then he has to use the enciphering key of H' for that.

To crack this cipher, one needs to figure out a deciphering key. To get such, knowing s, one needs to know $\varphi(r)$. This last is easily computed once one knows the primes p and q. But if they are kept secret, it takes a long time to figure them out: factoring integers having large prime factors is virtually impossible in a reasonable time.

The RSA encryption system has proven to be so useful and secure that if one makes a money transfer on line, chances are that the encription system used to make the transaction secure is very close to the RSA or the Diffie-Hellman systems described here.

Remark 12.3.1 As we saw above, the RSA criyptosystem is very easy to implement. This is due partially to the fact that the theory of numbers used for that is quite elementary. What makes the system really good to use, and also widespread, is that cracking it is hard. Naturally, in order to crack the RSA cryptosystem, one needs effective factorization methods. Many such (for instance, Pollard's ρ-method, Pollard's $p-1$-method, the quadratic sieve factorization method, Lenstra's elliptic curve factorization method, and the number field sieve), based on beautiful and much deeper theory than the one presented in these Notes, have been developed. Discussing these is, unfortunately, far beyond the scope of these Notes, and are not discussed here.

Chapter 13

A Bit More on Primes

We are discussing in this chapter properties of the prime numbers which are proved by using methods from Math Analysis rather than from Algebra. We will be as gentle as possible in using such methods: our arguments will be as "elementary", that is, not using hard analysis, as possible. At the same time, we will prove important and non-trivial results. Some of these results can be proved very quickly, using really ingenious tricks (chapter one of the book Problems from the BOOK by M. Aigner and G. Ziegler contains wonderful examples of such proofs). We will not use such proofs. Instead, we will try to explain the nature of methods and the results in this area of Number Theory.

13.1 Infinitude of Primes Revisited

We already know that the prime numbers are infinitely many: we have proved earlier in the course even that prime numbers of certain types, such as $1 \mod 4, 3 \mod 4, 1 \mod 6$*, and* $5 \mod 6$*, are infinitely many. The natural question now would be how many are the primes with respect to other types of integers? We will address this question by studying the function counting the prime numbers less than a given (real) number.*

Definition 13.1.1 *Denote by* $\mathbf{P}$ *the set of all prime numbers. Define, for every* $x \in \mathbb{R}$

$$\pi(x) = |\{p \,|\, p \in \mathbf{P} \wedge p \leq x\}|.$$

The function $\pi(s)$ *is determined by its values at the natural numbers, and could therefore be considered as an arithmetic function. We prefer to have it defined for all real numbers, because we study its properties by using methods of Math Analysis. Obviously,* $\pi(x) = 0$ *for all numbers* $x < 2$*. So, in our discussion of* $\pi(x)$*, we will focus on arguments in the interval* $x \geq 2$*.*

Notice that the infinitude of prime numbers is equivalent to saying that $\lim_{x\to\infty} \pi(x) = \infty$*. Obviously,* $\pi(x) < x$*. One natural question is whether* $\pi(x) < x^2$ *for example. This question actually asks if the prime numbers in the interval* $(0, x]$ *are less than the squares in that interval (do you see why?). Similarly, we can try to compare* $\pi(x)$ *to any real-valued function* $f : \mathbb{R}_{\geq 2} \to \mathbb{R}$*. Naturally, the interesting functions we do the comparison with need to have* $f(x) \to_{x\to\infty} \infty$*. Examples of such functions are* $f(x) = x, f(x) = x^2$*, and* $f(x) = \ln x$*.*
Suppose, $f(x)$ *is a function we want to compare* $\pi(x)$ *to. It is often natural to study the behaviour of the quotient* $\pi(x)/f(x)$*. When* $f(x) = x$*, the quotient* $\pi(x)/x$ *is the* **density function** *of the prime numbers in the positive real numbers. If we restrict* x *to be an integer, that is* $x = n$*, then* $\pi(n)/n$ *can also be interpreted as the probability of randomly chosen natural number in the interval* $[0, n]$ *to be prime.*
More generally, the quantity $(\pi(x + a) - \pi(x))/a$ *measures the density of the prime numbers in an interval of length* $a > 0$ *and left end-point* x*. When* $a = n$ *and* x *are integers (not less than* 2*), the above quantity measures the probability of having a prime number in the interval* $(x, x + n]$*.*

13.1.1 Some Notations from the Theory of Functions

A piece of terminology and notations is in order, so that we can state the results in this section in a more or less professional way.

Definition 13.1.2 *Let $f, g : (x_0, \infty) \to \mathbb{R}$ be two functions. We say that* **the function f is big oh of the function g as $x \to \infty$,** *and write $f = O(g)$, if there are constants $x' \geq x_0$ and c such that $|f(x)| < c|g(x)|$ for all $x > x'$.*
We say that f **and g have the same order of magnitude as $x \to \infty$,** *and write $f \asymp g$, if $f = O(g)$ and $g = O(f)$*

It follows from the definition that $f = O(1)$ means $|f(x)| < c$ for some c as $x \to \infty$. Notice that the same order of magnitude of two functions f and g doesn't mean they are close to each other (as $x \to \infty$): the difference $|f(x) - g(x)|$ may tend to infinity with x. Consider $f(x) = x^2 + x$ and $g(x) = x^2$ for example. What is true though is that.

Exercise 13.1 *Show that $f \asymp g$ is equivalent to $\ln|f(x)| - \ln|g(x)| = O(1)$. In other words, two functions have the same order of magnitude, as $x \to \infty$, iff the difference $|\ln|f(x)| - \ln|g(x)||$ is bounded, as $x \to \infty$.*

Remark 13.1.1 In Number Theory, the big oh notation is often replaced by the notation introduced by the Russian number theorist I.M. Vinogradov. We write $f \ll g$ if $f = O(g)$. In these notations, $f \asymp g$ whenever $f \ll g$ and $g \ll f$. □

There is a refinement of the notion of same order of magnitude.

Definition 13.1.3 *Let $f, g : (x_0, \infty) \to \mathbb{R}$ be two functions such that g is nowhere vanishing in that interval. We say that f* **and g are asymptotically equivalent**, *and write $f \sim g$, if $\lim_{x\to\infty} f(x)/g(x) = 1$.*

Exercise 13.2 *Show that $f \sim g$ implies that for every $0 < c_1 < 1 < c_2$ there is an $x' \geq x_0$ such that*

$$c_1|g(x)| \leq |f(x)| \leq c_2|g(x)|$$

for $x \geq x'$. In particular, we have $f \asymp g$.

13.1.2 The Prime Number Theorem

The function $\pi(x)/x$ was studied (empirically) by Legendre and Gauss. Gauss went on to study the density of the primes in intervals of fixed length. He considered in particular the function

$$\Delta(x) = \frac{\pi(x) - \pi(x - 1000)}{1000},$$

and noticed that it behaves as if

$$\Delta(x) \asymp \frac{1}{\ln x}.$$

Observe that the function $\Delta(x)$ is a slope of a secant line of the graph of $\pi(x)$. So, $\pi(x)$ should be close to the integral of $1/\ln x$. Naturally, Gauss introduced the **Logarithmic integral** *function*

$$Li(x) = \int_2^x \frac{1}{\ln t} dt,$$

and, after some empirical evidence, conjectured that

Theorem 13.1.1 (**Prime Number Theorem (PNT)**)

$$\pi(x) \sim Li(x) \quad as \quad x \to \infty.$$

Gauss made his conjecture in 1792-1793, being barely 16 year old. The first public record of the conjecture was made by Legendre in 1798 in his Essai sur la Théorie des Nombres.

The first significant progress toward proving the PNT is due to the Russian mathematician Pafnuty Chebyshev, in 1850 (the year of Gauss' passing away), who proved that

- $\pi(x) \asymp x/\ln x$ *for* $x \to \infty$. *More precisely, for* $x > 2$,

$$\frac{7}{8} \cdot \frac{x}{\ln x} \leq \pi(x) \leq \frac{9}{8} \cdot \frac{x}{\ln x}.$$

- *If* $\lim_{x\to\infty} \pi(x)/(x/\ln x)$ *exists, then* $\pi(x) \sim x/\ln x$.

Since $Li(x) \sim x/\ln x$ *as* $x \to \infty$ *(check this out using L'Hôpital's rule as an easy exercise!), and since the asymptotic equivalence is an equivalence relation, we see that the results of Chebyshev are closely related to the PNT. Very often, the Prime Number Theorem is stated, in an equivalent way, as*

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln x} = 1.$$

The most important ingredient toward proving the PNT was the suggestion by the German mathematician, and a student of Gauss', Bernhard Riemann , in 1859, who showed that the zeta function $\zeta(s)$, *to be introduced below, can be used to attack the PNT. It was in 1896 when the French mathematician Jacque Hadamard and the Belgian mathematician Charles de la Vallée Poussin independently realized the idea of Riemann, and proved the PNT (showing that* $\zeta(s)$ *has no zeros on the line* $Re(s) = 1$*).*

The Prime Number Theorem is considered central in the theory of numbers. There are several proofs of it known. Some of them are short, but use deeper knowledge of complex analysis, other are "elementary" in the sense that they do not use heavy analytical machinery, but are long. Either way, the proof of PNT goes far beyond the scope of these lectures, and we are not proving it here. Instead, in the following subsections, we are proving an easier version of Chebyshev's results, and are discussing, on quite elementary level, the relationship between $\zeta(s)$, *and* $\pi(x)$. *This, we hope, will give a good idea of the methods used in this part of Number Theory (called Analytic Number Theory).*

The Function $\zeta(s)$

It was Euler who first recognized the importance of the series

$$\frac{1}{1^s} + \frac{1}{2^s} + \cdots + \frac{1}{n^s} + \cdots$$

where $s \in \mathbb{N} \setminus \{0\}$, *for the theory of prime numbers. One of his observations was that, for* $s \geq 2$, *we have*

$$\sum_{n\geq 1} \frac{1}{n^s} = \prod_{p\in\mathbf{P}} \left(\sum_{k\geq 0} \frac{1}{(p^s)^k} \right) = \prod_{p\in\mathbf{P}} \frac{1}{1-\frac{1}{p^s}}.$$

The equality of the two extreme terms above is known as **Euler's product**, *and the factors on the right-hand side are called* **Euler's factors**.
Euler also knew that the harmonic series, corresponding to $s = 1$, *is divergent. That allowed him to give a new proof of the infinitude of the prime numbers. Here is how it goes. Assume, by RAA, that the prime numbers are finitely many:* $\mathbf{P} = \{p_1, \ldots . p_m\}$. *Then obviously*

$$\prod_{p\in\mathbf{P}} \left(\sum_{k\geq 0} \frac{1}{p^k} \right) = \frac{1}{1-\frac{1}{p_1}} \times \cdots \times \frac{1}{1-\frac{1}{p_m}} \quad \in \mathbb{Q}$$

which, as can be shown, forces the harmonic series to converge

$$\sum_{n\geq 1}\frac{1}{n}=\frac{1}{1-\frac{1}{p_1}}\times\cdots\times\frac{1}{1-\frac{1}{p_m}}.$$

This proof was apparently the first in which Calculus methods were used in a significant way to establish a number theoretical result.
One of the greatest triumphs of Euler as a mathematician is him proving (among the enormous amount of other important and deep results) that

$$\frac{1}{1^2}+\frac{1}{2^2}+\cdots+\frac{1}{n^2}+\cdots=\frac{\pi^2}{6}.$$

This formula gives us one more way of proving that the primes are infinitely many. This proof wouldn't be in the reach of Euler, though! As it was proven by F. Lindemann in 1882 (99 years after Euler passed away), the number π is transcendental, that is, it is not a root of any polynomial with integer coefficients and of positive degree. From this it follows that $\pi^2/6$ is not a rational number (do you see why?). Now, if we assume that the prime numbers are finitely many, in the notations introduced above, we would have

$$\frac{\pi^2}{6}=\sum_{n\geq 1}\frac{1}{n^2}=\prod_{p\in\mathbf{P}}\left(\sum_{k\geq 0}\frac{1}{(p^2)^k}\right)=\frac{1}{1-\frac{1}{(p_1)^2}}\times\cdots\times\frac{1}{1-\frac{1}{(p_m)^2}}\quad\in\mathbb{Q}$$

which is not possible.

The two examples of usage of the sum of the sth powers of the reciprocals of the positive natural numbers motivates the introduction of the function called **Riemann's zeta function**

$$\zeta(s)=\sum_{n\geq 1}\frac{1}{n^s}$$

where s is a complex number with real part $Re(s)>1$. It is not hard to prove, but goes beyond the scope of these Lecture Notes, that the function is well defined (the right-hand side is a convergent series for such s). The function $\zeta(s)$ was introduced by Euler. Georg Friedrich Bernhard Riemann (1826-1866) studied $\zeta(s)$ thoroughly and proved (in 1859) that it can be extended to the whole complex plane $\mathbb{C}$ as a meromorphic function having a simple pole, with residue 1, at $s=1$. In other words, $\zeta(s)$ is defined for all $s\neq 1$, and the difference

$$\zeta(s)-\frac{1}{s-1}$$

is an analytic function around 0. In studying meromorphic functions, it is very important to know the zeros and the poles (with their orders) of these functions. Riemann proved that $z(s)$ has simple zeros at $s=-2,-4,\ldots$ (often called trivial zeros of $\zeta(s)$), and that it has infinitely many zeros in the region $0\leq Re(s)\leq 1$, called the critical strip). Riemann conjectures that the zeros in the critical strip, called non-trivial zeros of $\zeta(s)$, have $Re(s)=1/2$. This conjecture, which is widely believed to be true, is known as the **Riemann hypothesis**, *and is still open. Proving the Riemann hypothesis is an important goal, because a long list of important theorems can be proved knowing that the hypothesis is true. In regard to the function $\pi(x)$ for example, Riemann sketched a proof of an explicit formula for it, proven rigorously by Mangoldt in 1895, which depends on the non-trivial zeros of $\zeta(s)$, and which, if the hypothesis is true, simplifies to*

$$\pi(x)=Li(x)+O(x^{1/2+\epsilon})$$

where $\epsilon>0$ is an arbitrary constant. This immediately would imply the PNT.

A Sketch of a Proof of the Prime Number Theorem

In this subsection, we are giving a very short account, avoiding the hardest and most technical details, of a proof of the PNT borrowed from Helmut Koch's wonderful book Number Theory Algebraic Numbers and Functions, and which belongs to Don Zagier (in Newman's Short Proof of the Prime Number Theorem). The arguments, and the computations, are accessible to students with background in courses such as Advanced Calculus and Complex Variables.

We have about the logarithmic derivative $\zeta'(s)/\zeta(s)$, of $\zeta(s)$, and for $Re(s) > 1/2$,

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{d}{ds}\log(\zeta(s)) = \frac{d}{ds}\log\left(\prod_{p\in\mathbf{P}}\left(1-\frac{1}{p^s}\right)^{-1}\right) = -\sum_{p\in\mathbf{P}}\frac{d}{ds}\log\left(1-p^{-s}\right)$$

$$= -\sum_{p\in\mathbf{P}}\frac{p^{-s}\log p}{1-p^{-s}} = -\sum_{p\in\mathbf{P}}\frac{\log p}{p^s-1} = -\left(\phi(s)+\sum_{p\in\mathbf{P}}\frac{\log p}{p^s(p^s-1)}\right)$$

where

$$\phi(s) = \sum_{p\in\mathbf{P}}\frac{\log p}{p^s}.$$

Since $\zeta(s)$ is a meromorphic function, its logarithmic derivative has simple poles at the points of the zeros, and the pole of $\zeta(s)$. The residues at these poles are equal to the orders of the zeros or with minus the order of the pole. The function

$$\sum_{p\in\mathbf{P}}\frac{\log p}{p^s(p^s-1)}$$

is holomorphic, so $\phi(s)$ is a meromorphis function whose poles are simple, and are at the zeros and at the pole of $\zeta(s)$. It is the function $\phi(s)$ which allows us to establish the crucial for the proof of the PNT **fact about $\zeta(s)$: it doesn't have zeros on the line** *$Re(s) = 1$. This fact was proven by de la Valée Poussin in 1896. Assume, by RAA, that for some $t \in \mathbb{R}$ the number $s_1 = 1 + it$ is a zero of $\zeta(s)$ of order $k \geq 1$. Since $\zeta(s)$ has a pole at $s = 1$, we get that $t \neq 0$. Obviously the number $s_2 = 1 + i2t$ is a zero of $\zeta(s)$ of order $l \geq 0$. The very important functional equation (proved by Riemann in 1859) states that*

$$\zeta(1-s) = (2\pi)^{-s}2\cos\left(\frac{\pi}{2}\right)\Gamma(s)\zeta(s)$$

where $\Gamma(s)$ is the Gauss's gamma function. From this it follows that $s_3 = 1 - it$, respectively $s_4 = 1 - i2t$, is also a zero of $\zeta(s)$, and of order equal to the order of s_1, respectively s_2. Therefore, due to the minus sign in the formula relating the logarithmic derivative of $\zeta(s)$ and $\phi(s)$, the latter function has poles at s_1 and s_3 of residue $-k$, and at s_2 and s_4 of residue $-l$. A fact known from Complex Variable course about the residues at simple poles is the following

$$\lim_{\epsilon\to 0}\epsilon\phi(1+\epsilon\pm it) = -k, \quad \lim_{\epsilon\to 0}\epsilon\phi(1+\epsilon\pm i2t) = -l, \quad \lim_{\epsilon\to 0}\epsilon\phi(1+\epsilon) = 1.$$

Now - a trick which explains the need for the zeros s_2 and s_4 (of order ≥ 0). We have

$$\sum_{m=-2}^{2}\binom{4}{2+m}\epsilon\cdot\phi(1+\epsilon+imt) = \epsilon\sum_{p\in\mathbf{P}}\frac{1}{p^{1+\epsilon}}\left(p^{it/2}+p^{-it/2}\right)^4 \geq 0$$

if we choose $\epsilon > 0$. Letting $\epsilon \to 0$ with positive values we get

$$-2l - 4k + 6 \geq 0$$

which forces $k = 0$ - a contradiction.
As a consequence of the above, we get that $\phi(1+z)$ is having a simple pole at $z = 0$, and of residue 1, and is holomorphic elsewhere for $Re(z) \geq 0$. Therefore, the function

$$g(z) = \frac{\phi(1+z)}{1+z} - \frac{1}{z}$$

is holomorphic for $Re(z) \geq 0$. The following theorem is the technical hearth of the proof, and is omitted here (the reader is referred to look for it in the book by Koch cited above).

Theorem 13.1.2 *Denote $\theta(x) = \sum_{p \leq x} \log p$. We have*

$$\lim_{T \to \infty} \int_1^T \frac{\theta(x) - x}{x^2}\, dx = g(0).$$

The function $\theta(x)$ was introduced by Chebyshev. It is closely related to the function $\pi(x)$, because of the following (quite elementary) estimates. First notice that

$$\theta(x) = \sum_{p \leq x} \log p \leq \sum_{p \leq x} \log x = \pi(x) \cdot \log(x).$$

Second, for every $0 < \epsilon < 1$ we have

$$\theta(x) \geq \sum_{x^{1-\epsilon} \leq p \leq x} \log p \geq \sum_{x^{1-\epsilon} \leq p \leq x} \log(x^{1-\epsilon}) = (1-\epsilon) \sum_{x^{1-\epsilon} \leq p \leq x} \log x$$

$$= (1-\epsilon)(\pi(x) - \pi(x^{1-\epsilon})) \log x \geq (1-\epsilon)(\pi(x) - x^{1-\epsilon}) \log x,$$

and so, we get

$$\frac{\theta(x)}{x} \leq \pi(x) \cdot \frac{\log x}{x} \leq \left(\frac{1}{1-\epsilon}\right) \frac{\theta(x)}{x} + \frac{\log x}{x^\epsilon}.$$

Assume, for a moment, that $\theta(x) \sim x$. The estimates above immediately give us that, for $0 < \epsilon < 1$, for $\delta > 0$, and for $x >> 0$

$$1 - \delta \leq \pi(x) \cdot \frac{\log x}{x} \leq \left(\frac{1}{1-\epsilon}\right) \cdot (1+\delta)$$

which immediately proves that

$$\pi(x) \sim \frac{x}{\log x}.$$

We are showing now how the theorem above helps us prove that $\theta(x) \sim x$. Since the limit $\lim_{T\to\infty} \int_1^T (\theta(x) - x)/x^2\, dx$ exists, for every $\epsilon > 0$ there is $x_0 \in \mathbb{R}$ such that for any $x_0 \leq y_1 < y_2$ we have

$$-\epsilon \leq \int_{y_1}^{y_2} \frac{\theta(x) - x}{x^2}\, dx \leq \epsilon.$$

The function $\theta(x)$ is an increasing step function. So, the function $\theta(x)/x$ is continuous between every two consecutive prime numbers. As the exercise below states, we have $0 \leq \theta(x)/x \leq 2$, and therefore $\liminf_{x\to\infty} \theta(x)/x$ and $\limsup_{x\to\infty} \theta(x)/x$ exist. We want to show that both these limits are equal to 1. For this, it is enough to show that

$$\liminf_{x \to \infty} \frac{\theta(x)}{x} \geq 1 \qquad \textit{and} \qquad \limsup_{x \to \infty} \frac{\theta(x)}{x} \leq 1.$$

Let $\liminf_{x\to\infty} \theta(x)/x = m < 1$. Let λ be such that $m < \lambda < 1$. So, we have that $\theta(x) \leq \lambda x$ for $x >> 0$. In the notations above, for $x_0 \leq \lambda y$, we have

$$-\epsilon \leq \int_{\lambda y}^{y} \frac{\theta(x) - x}{x^2} dx \leq \int_{\lambda y}^{y} \frac{\lambda x - x}{x^2} dx = (1-\lambda) \log \lambda.$$

Since these inequality $-\epsilon \leq (1-\lambda)\log\lambda$ *holds true for every positive* ϵ*, we get the absurd that* $0 \leq (1-\lambda)\log\lambda < 0$*. Therefore,* $\liminf_{x\to\infty} \theta(x)/x \geq 1$ *as claimed. In a similar way, one proves that* $\limsup_{x\to\infty} \theta(x)/x \leq 1$*. This completes the argument.*

Exercise 13.3 *(Chebyshev) Prove that, for every* $x \in \mathbb{R}_{\geq 0}$*, we have* $\theta(x) \leq 2x$*.*

13.2 Bertrand, Goldbach, and Twin Primes

In this section, we address some additional classical facts and questions about the prime numbers.

13.2.1 Bertrand's Postulate

In mid 19th century, the theory of groups (considered as subgroups of the symmetric group S_n*) was in the center of interests and studies of most of the leading mathematicians of that time. The work of Abel and Galois on the subject was still freshly published. Joseph Bertrand, a prominent mathematician, and a member of the Paris Academy, of that time is known for his work in many areas of mathematics, including Algebra. When studying subgroups of small index of* S_n *in 1845, he came up with the need for having a prime number between* n *and* $2n-2$ *for any natural* $n > 3$*. He was not able to prove that this is true, but checked it out for all* $n \leq 6,000,000$*. Since he didn't find a counterexample, he formulated the claim as a postulate. It was in 1850 when the Russian mathematician Pafnuty Lvovich Chebyshev proved the claim, and made it thereby a theorem. The result is now referred to as* **Bertrand's Postulate** *or* **Chebyshev's Theorem**.
Chebyshev proved Bertrand's Postulate using the function $\theta(x)$ *he introduced in his work on the PNT. In terms of this function, the Postulate states that, for every* $n > 6$ *we have*

$$\theta(n) - \theta(\frac{n}{2}) > 0.$$

Knowing that theta$(x) \sim x$*, it is immediate to prove the positivity of the difference above* **for big enough** x*. So, this would "prove" the Postulate asymptotically. To prove the Postulate as stated by Bertrand, Chebyshev used uniform over the positive real numbers bounds of* $\theta(x)$

$$c_1 x \leq \theta(x) \leq c_2 x$$

for appropriate positive c_1 *and* c_2*, and for* $x \geq 2$*.*
An "asymptotic" proof of the Bertrand's Postulate can be done using the PNT as well. Indeed, the Postulate in terms of $\pi(x)$ *claims that* $\pi(x) - \pi(x/2) > 0$ *for* $x > 6$*. But asymptotically we have*

$$\pi(x) - \pi(x/2) \sim \frac{x(\log x - 2\log 2)}{2\log x \log(x/2)}$$

which is obviously positive for $x >> 0$*.*
There are now several profs of the Bertrand's Postulate. The most famous one belongs to P. Erdös and was published in 1932, and is widely accessible in the literature (the book Proofs from the Book is a wonderful source to read the proof from), and on the Internet. We are not reproducing it here. At about 1854 Chebyshev proved

$$(\forall \epsilon > 1/5)(\exists \xi_\epsilon)(\forall n)(n > \xi_\epsilon \to (\exists p \in \mathbf{P})(n < p \leq (1+\epsilon)n)).$$

For instance, when $\epsilon = 1$*, Bertran's Postulate claims* $\xi_1 = 3$*. Sylvester and others proved after that that the theorem is true for all* $\epsilon > 0$*. This in particular gives us that*

$$\lim_{n\to\infty} \frac{p_{n+1}}{p_n} = 1$$

where p_k *denotes the kth prime number.*
Speaking of John Sylvester (who was the first prominent European mathematician to receive an

offer to come and work in the USA, being paid in pure gold), he formulated a generalization of the Bertrand's Postulate: **for any positive** m**, and** n **such that** $m > n$**, there is a prime number** $p > n$ **dividing the product** $m(m+1)\cdots(m+n-1)$. *Bertrand's Postulate corresponds to* $m = n+1$. *This is still an open conjecture, as is the question whether there is a prime number between any two consecutive complete squares*

$$n^2 < p < (n+1)^2.$$

Closing this subsection, let mention an amusing fact which easily follows from the Bertran Postulate: For no integer $n > 1$ *is the number* $n!$ *a power of an integer. (Try to find an argument proving this claim!)*

13.2.2 Goldbach's Conjectures

In a letter to Leonard Euler in June 1742, Christian Goldbach (1690-1764) mentioned that every even number bigger than 3 *is a sum of two primes. Edward Waring (1734-1798), known for the Waring's problem for representing natural numbers as sums of* k*th powers, also declared that every number is either a prime or a sum of three primes. These claims are nowadays known as the* **binary and the ternary Goldbach problems**

(BGP): Every natural number $n \geq 4$ *is a sum of two prime numbers;*
(TGP): Every odd natural number $n \geq 7$ *is a sum of three prime numbers.*

The BGP is still open, and remains one of the hardest problems in (classical) Number Theory. The TGP was solved by the Russian mathematician Ivan Matveevich Vinogradov who proved in 1937 that there is an integer N_0 *such that every odd* $n \geq N_0$ *is a sum of three prime numbers.*

13.2.3 Twin Primes

A very old and famous question in Number Theory is if the twin primes are infinitely many. The empirical evidence follows toward positive answer to that question, and the claim that the answer is positive is known as the **Twin Prime Conjecture**. *But proof is not known as of today. What follows is a short account of what we know so far in this direction.*
First off, following the line of study of the number of primes less than a given real number, we introduce the function

$$\pi_2(x) = Card(\{p \in \mathbf{P} \,|\, p < x \wedge p+2 \in \mathbf{P}\}).$$

The so called quantitative form of the twin prime conjecture is the following claim

Twin Prime Conjecture(Quantitative Form) *Asymptotically, as* $x \to \infty$, *we have*

$$\pi_2(x) \sim 2C_2 \int_2^x \frac{dt}{(\log t)^2}$$

where

$$C_2 = \prod_{p \in \mathbf{P}, p>2} \left(1 - \frac{1}{(p-1)^2}\right).$$

Proof of this conjecture would immediately confirm that the Twin Prime Conjecture is a theorem.

A significant progress was achieved in the recent years in tackling these conjectures. It was proven by Y. Zhang in 2013 that for $N = 7 \cdot 10^7$ *we have*

$$\liminf_{n\to\infty}(p_{n+1} - p_n) < N.$$

The number N *was improved to* 246 *in the year that followed. This means that there are infinitely many consecutive primes with* $p_{n+1} < p_n + 246$.

13.3 Functions with Values Prime Numbers

In this last section, we are considering examples of "reasonable" functions taking on values which are prime numbers.

13.3.1 Polynomial Functions

Recall Dirichlet's theorem about prime numbers in an arithmetic progression: if a_0 and a_1 are relatively prime, non-zero integers, then the polynomial $f(X) = a_0 + a_1X$ considered as a function on $\mathbb{Z}$ takes on infinitely many values which are prime numbers. People asked the question if there are polynomials of greater than 1 degree which have the same property. There are known polynomials of second degree which take on many prime values when considered as functions on $\mathbb{Z}$. For instance, for $X = 0, 1, \ldots, 28$, the values of the polynomial $6X^2+6X+31$ are all prime numbers. The polynomial $2X^2 + 29$ takes on prime numbers for $X = 0, \pm1, \ldots, \pm28$. But the real champion is the polynomial suggested by L. Euler

$$f(X) = X^2 + X + 41.$$

It takes on prime values for 80 integer values of X: $0, \pm1, \pm2, \ldots, \pm39, -40$. It was proven that there is no natural number $A > 41$ such that the polynomial $X^2 + X + A$ takes on prime number values for $X = 0, 1, \ldots, A - 2$.
As a matter of fact, no polynomial with integer coefficients, and of degree greater than 1 is known which, considered as a function on $\mathbb{Z}$, has infinitely many prime numbers in its range. There are famous hypotheses in this respect. For instance, Emil Artin conjectured that $f(X) = X^2 + 1$ is one such polynomial.
Here is an elementary fact which tells us that wanting infinitely many primes in the ranges of polynomials is the most that one can get.

Proposition 13.3.1 *There is no non-constant polynomial with integer coefficients which considered as a function on $\mathbb{Z}$ has range containing only prime numbers.*

Proof *Let $f(X) = a_0 + a_1X + \cdots + a_nX^n \in \mathbb{Z}[X]$ with $n \geq 1$ and $a_n \neq 0$. First off, notice that there is an $m \in \mathbb{Z}$ such that $f(m) \neq \pm1$. Indeed otherwise one of $f(X) \pm 1$ would have infinitely many roots which is impossible. Let $d > 1$ be a divisor of $f(m)$, and consider $g(X) = f(dX + m)$. It is straightforward that $g(X)$ is a non-constant polynomial with integer coefficients which are all divisible by d. So, for any $s \in \mathbb{Z}$ the value $g(s) = f(ds + m)$ is not a prime number.* □

13.3.2 Functions on $\mathbb{Z}$ with Only Prime Number Values

When looking for "reasonable" functions which would take prime number values, it is natural to consider those with domains (a subset of) the integers $\mathbb{Z}$. Naturally, a function defined as $P(n) = p_n$ where p_n is the nth prime number can not be considered as a reasonable function, because it "looks" too artificial, and is not in a "closed form"! A way to hide the direct and rude use of prime numbers in the definition of $P(X)$, and finding actually a "closed form" for that function, is the following (borrowed from An Introduction to the Theory of Numbers by G. Hardy and E. Wright). Let

$$\alpha = \sum_{m=1}^{\infty} p_m 10^{-2^m}.$$

Obviously α exists (the series is convergent due to Bertrand's Postulate). It is immediate to see that the function

$$h(X) = \lfloor 10^{2^X}\alpha \rfloor - 10^{2^{X-1}} \lfloor 10^{2^{X-1}}\alpha \rfloor$$

is such that for every positive integer n we have $h(n) = p_n$. The problem with this function is that to get p_n we need to know α up to its 2^nth decimal places...
One less "obvious" example of such function is the following

$$f(X) = \lfloor A^{3^X} \rfloor.$$

In 1947 W. H. Mills proved that there is a constant A such that for every $s \in \mathbb{Z}$ the value $f(s)$ is a prime number.
Both last examples have mainly theoretical importance: neither α, nor A are practically known. If we relax the restriction of having **all** *values to be prime numbers, one can find more tractable examples. For instance, consider the functions*

$$X^{2^n} + k, \quad k = 1, 3, 7, 9$$

with domain $\mathbb{Z}$. It is claimed in W. Sierpinski's Elementary Theory of Numbers that for any positive integer n, the range of these functions contain infinitely many prime numbers.

Bibliography

[1] *A. Adler and J. Coury The Theory of Numbers (A Text and a Source Book of Problems), Jones and Bartlett Publishers, Inc., 1995*

[2] *P. Aluffi Algebra. Chapter 0 GSM 104, AMS, 2009*

[3] *E. Bolker Elementary Number Theory. An Algebraic Approach W.A. Benjamin, Inc., 1970*

[4] *G. Hardy and E. Wright An Introduction to the Theory of Numbers, Oxford University Press, 2008*

[5] *H. Koch Number Theory GSM 24, AMS, 2000*

[6] *F. Lemmermeyer Reciprocity Laws Springer-Verlag, 2000*

[7] *T. Nagell Introduction to Number Theory John Wiley and Sons, Inc., New York, and Almkwist and Wiksell, Stockholm, 1951*

[8] *P. Pollack Not Always Buried Deep: a Second Course in Elementary Number Theory, Monograph Book, AMS, 2009*

[9] *W. Scharlau and H. Opolka From Fermat to Minkowski, Springer-Verlag, 1984*

[10] *J.-P. Serre A Course of Arithmetic, Springer-Verlag. 1993*

[11] *W. Sierpinski Elementary Theory of Numbers, North-Holland Mathematical Library, PWN - Polish Scientific Publishers, 1988*

[12] *J. Silverman A Friendly Introduction to Number Theory, Pearson, 2013*

[13] *W. Stein Elementary Number Theory. Primes, Congruences and Secrets, UTM, Springer, 2009*

[14] *B. A. Venkov Elementary Number Theory Wolters-Noordhoff Publishing, Groningen, 1970*

[15] *I. M. Vinogradov Elements of Number Theory, Dover Publications, Inc., 1954*

[16] *A. Weil Number Theory (An Approach through History from Hammurapi to Legendre) Birkhäuser Boston, 2007*

[17] *E. Weiss First Course in Algebra and Number Theory, Academic Press, Inc. 1971*

Index