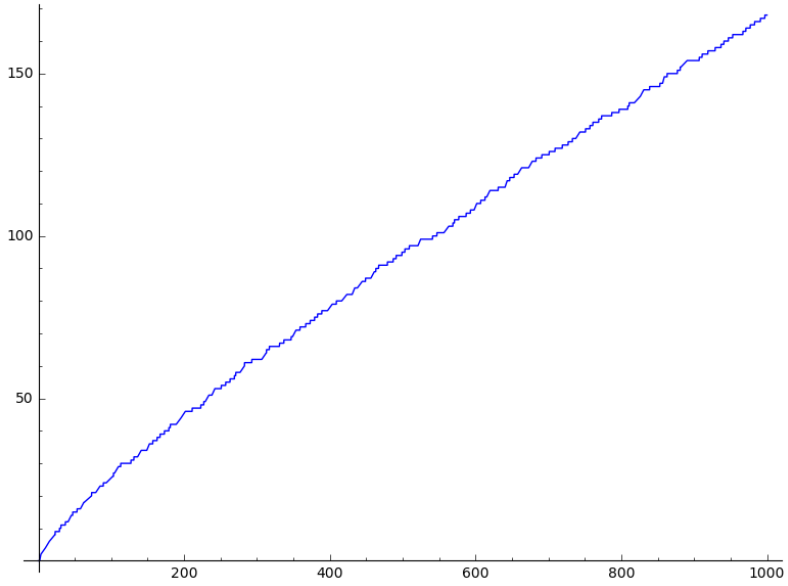
Hatice Boylan and Nils-Peter Skoruppa

Elementary Number Theory



Lecture Notes
İstanbul Üniversitesi
and Universität Siegen

Version: May 2018

Contents

Foreword

The following speech[1], given several decades ago at the occasion of the opening of one of the most famous research institutes of the world for mathematical sciences might help the reader to come closer to answer the questions *'What is number theory?'* and *'What is it good for?'*:

Why do you study number theory?

> Mathematics and German share the same disadvantage, both are universally applicable and at the same time they are the summit of artistic creation of human kind. Why do we need Goethe if we can express our wishes clearly at the market place? And for what do we need number theory if we can solve the differential equation of the heat equation numerically? Strangely enough, in this competition those domains do better which have no imaginable commercial application. One of my colleagues at Durham University was once asked by the local TV why he studies the precise dating of Crete vases, and he answers that this would be very useful for the study of the migration of

[1]This is a free translation of a German text which appeared in one of the internal publications of the Max-Planck Gesellschaft and was in turn probably translated from the original English speech. We are grateful for any hint to the English original.

the Minoan civilization. To my surprise this was accepted with respectful appreciative murmuring.

Hence our first answer to the question 'Why do you study number theory' should possibly be 'It is indispensable for the right understanding of modular forms.' After we have now put down the objections of the trifling and superficial people, we can try to answer seriously. The serious answer is, of course: 'Why not?'. Namely, beavers build dams and cuckoo borrow nests without any intent of refunding, but only humans (as far as we know) worry about the questions which prime numbers are the sum of two squares. Since we reached a partial freedom from the urgent need of surviving, the desire for knowledge and the expression of beauty were always the ultimate goal of the human race. The purpose of technology and invention is to give us time for the further study of Bach, Gauß and Goethe, and not vice versa. But it is one of the divine compensations for our existence that the compulsive quest for knowledge almost always eventually carries practical fruits.

A.O.L. Atkin,
at the occasion of a visit to the
Max-Planck-Institut for Mathematics in Bonn,
Bonn, June 1985.

Preface

These lecture notes grew out of a first course in number theory for second year students as is was given by the second author several times at the University of Siegen and by the first one in 2015/2016 at İstanbul Üniversitesi in Istanbul.

There are many books on elementary number theory, most of them in English, and with very different goals: classical, computational, theoretical, as a supplement to Algebra or from scratch. In this sense it would be unnecessary to provide a script. However, the given courses comprised each only 24 ninety minutes lecture. Hence the challenge was to reduce the contents and at the same time keep and prove rigorously key points of elementary number theory. So it might be helpful to provide a 'shortened stream-lined' version of elementary number theory as answer to the mentioned challenge. We hope that we did not do too bad when trying to reach this goal, and we hope that these lecture notes are indeed useful not only for our students, but also for our colleagues in future years.

These notes are mainly based on notes on elementary number theory which the second author collected during the past 15 years for his usage in his courses on this subject. There are some topics or treatments of such which may not be found or at least not easily found at other places. The reader will find a whole section on the basics of projective geometry since we feel that diophantine equations cannot

be treated without a certain geometric understanding. There is a whole chapter on conic sections and a natural group law on their set of rational points. This section anticipates in an elementary and easily accessible way various ideas from the theory of elliptic curves as it may be found in more advanced monographs. The theory of Pell's equation and the theory of continued fractions is here consequently explained as part of the theory of the group $\mathrm{SL}(2,\mathbb{Z})$.

Places where we want to warn the reader from wrong conclusions are marked by the first sign on the left hand side. Similarly, we occasionally leave it to the reader to find or complete an argument or computation. We indicate this by the second sign on the left side. Hints to corrections or errors are very welcome.

April 2016 Hatice Boylan and Nils-Peter Skoruppa

Chapter 1

Basics

1. The integers

We shall use $\mathbb{Z}$ for the set of integers, $\mathbb{Z}_{\geq 0}$ for the set of natural numbers[1], and $\mathbb{Q}$, $\mathbb{R}$ for the set of rational and real numbers, respectively. Recall that $\mathbb{R}$ is the smallest field containing $\mathbb{Q}$ such that every Cauchy sequence has a limit. We shall rarely deal with real numbers. Elementary number theory concerns properties of integers and rational numbers. We shall assume that the reader is acquainted with the notion of *an integer* and their basic properties, and we shall not waste time to characterize the integers axiomatically (though this would be easily possible as we shall indicate in the section "Remarks" at the end of this chapter). However, there is one property which we mention explicitly. This is the induction axiom.

Axiom (Induction Axiom)**.** *A subset of natural numbers which contains* 0*, and which contains with every number* n *also the number* $n+1$*, equals the whole set of natural numbers.*

This axiom is often applied to prove a property or identity for all natural numbers. For example one can easily prove via the induction

[1]We avoid the often used notation $\mathbb{N}$ since it is ambiguous: in the literature many authors include the number 0 in $\mathbb{N}$ whereas many others do not.

axiom that the identity

$$\sum_{k=0}^{n} k^3 = \frac{n^2(n+1)^2}{4}$$

holds true for all natural numbers.

Another important consequence, which we shall often apply is the following.

Theorem. *Every non-empty set of natural numbers has a smallest element.*

Proof. Let A be a set of natural numbers without a smallest element. We show that A is then empty. Indeed, let B be the set of natural numbers which are not in A. We have to show that B equals the set of all natural numbers, and we do this by induction. Clearly, 0 is in B since otherwise 0 would be the smallest element of A. If, for a given n, all natural $k \leq n$ are in B, then all $k \leq n+1$ are in B too since otherwise $n+1$ would be the smallest number of A. □

2. Divisibility and prime numbers

2.1. Euclid's Fundamental Theorem.

Definition. For integers a, b, we say a *divides* b, noted by $a \mid b$, if there is an $x \in \mathbb{Z}$ such that $b = ax$.

Remark. 1. The divisibility relation "'|" defines a partial ordering of $\mathbb{Z}_{\geq 0}$, i.e. "'|" is reflexive, transitive, and $a \mid b$, $b \mid a$ implies $a = b$. 2. If $d \mid a, b$, it follows $d \mid ax + by$ for all integers x and y.

Definition. A number $p \in \mathbb{Z}_{\geq 2}$ is called *prime number* (or shortly, a *prime*), if p has no other divisors than 1 and p.

Theorem (Fundamental theorem of Euclid)**.** *Every natural number possesses a unique prime factorization.*

Remark. Be prime factorization (or simply "factorization") of a number n we mean a factorization

$$n = p_1^{n_1} \cdots p_r^{n_r}$$

with prime numbers p_j and non-negative integers n_j. We can of course assume in such a writing that the exponents n_j are strictly positive and that the sequence of the p_j is strictly increasing. That such a factorization is unique means then that r, the p_j and the exponents n_j are uniquely determined by n.

Proof of the Fundamental theorem. For the 'Existence' we use induction: Let $n > 1$ be a natural number. Let p be the smallest divisor > 1 of n. If $n = p$ then n is a prime and we are done. Otherwise p and n/p possess a prime decomposition (by induction hypothesis), and so n does too.

For 'uniqueness', which we prove also by induction, we use the following fact (which is called Euclid's Lemma and whose proof will be given below): If a prime divides a product of integers, then it divides at least one of these integers. Assume that the uniqueness of prime factorization is verified for all $k < n$, and assume that n possesses prime decompositions

$$n = p_1^{n_1} \cdots p_r^{n_r} = q_1^{m_1} \cdots q_s^{m_r},$$

where $n_j, m_j \geq 1$ and $p_1 < p_2 < \cdots < p_r$ and $q_1 < q_2 < \cdots < q_s$. Then p_1 divides the product on the right hand side, hence it divides by Euclid's Lemma one of the factors, i.e. q_j for some j, and then it even equals this q_j (since q_j, as prime number, possesses only as positive divisors 1 and itself). Dividing by p_1 we obtain

$$n/p_1 = p_1^{n_1-1} \cdots p_r^{n_r} = q_1^{m_1} \cdots q_j^{m_j-1} \cdots q_s^{m_r}.$$

By induction hypothesis we find $r = s$ and $p_h = q_h$, $n_h = m_h$ for all h. □

As an immediate corollary we obtain that every positive rational number z possesses a unique prime factorization

$$z = p_1^{z_1} \cdots p_r^{z_r}$$

with primes $p_1 < \cdots < p_r$, where now, however, the integers z_j can be negative. Indeed, for seeing the existence of such a decomposition write $z = m/n$ with positive integers m and n, and replace m and n by their respective prime factorization. For the uniqueness let m be the product of all p^{z_j} with $z_j > 0$ and n be the product of all p^{-z_j}

with $z_j < 0$. Then $z = m/n$. For a given second decomposition with powers $q_j^{z_j'}$ $(0 \leq j \leq s)$ define m' and n' accordingly. Then $mn' = m'n$, and replacing here m, m', ... by the corresponding products of p^{z_j}, $q^{z_j'}$ and invoking the uniqueness of the prime factorization for integers we conclude $r = s$, $p_j = q_j$ and $z_j = z_j'$ for all j.

Let m and n be two integers, and $p_1, \ldots, p_r$ be the pairwise different primes occurring in the factorization of m and n. We can then write

$$m = p_1^{m_1} \cdots p_r^{m_r}, \quad n = p_1^{n_1} \cdots p_r^{n_r},$$

where m_j, n_j are non-negative integers, possibly equal to 0. If we have $m_j \leq n_j$ for all j, then m obviously divides n, since the quotient of m/n is a product of primes, hence an integer. The inverse is also true. If m divides n, then n/m has prime factorization $p_1^{n_1 - n_1} \cdots p_r^{n_r - m_r}$, and since it is an integer the exponents $n_j - m_j$ must all be non-negative.

The *Sieve of Eratosthenes* is an algorithm which allows to compute rapidly all primes below a given natural number n. For this one notes on a sheet of paper all natural number between 2 and n, and then one crosses out all numbers which are not primes. Namely, one starts by crossing out 2 and all multiples of 2. Then one searches for the first number which is not crossed out (which here is 3), and which is therefore a prime (since otherwise it would have a prime divisor which is smaller, but then it would be crossed out). We cross out all multiples of 3 which are strictly larger than 3. Next we look for the first number after 3 which is not crossed out (which would be 5) and which is therefore a prime (by the same argument as before). We cross out all multiples which are strictly larger. We continue in this way until we reach $\sqrt{n}$. The not crossed out numbers are then all primes $\leq n$ (since every composite number[2] $\leq n$ possesses at least a prime divisor $\leq \sqrt{n}$ and hence is already crossed out).

Theorem. *There are infinitely many prime numbers (i.e. for every integer N there exists a prime which is larger than N).*

Proof. Assume there are only finitely many prime numbers. Let P be the product of all these primes and set $n = P + 1$. Then n

[2] A positive integer is called *composite* if it is not a prime.

possesses a prime divisor p (for example, the smallest divisor of n). Since n leaves rest 1 upon dividing by any prime, but p divides n, we have a contradiction. □

A mysterious function is the *distribution of primes*

$$\pi(x) := \operatorname{card}\left(\{p \text{ prime } | p \leq x\}\right).$$

A plot of the graph of $\pi(x)$ for $0 \leq x \leq 1000$ can be found on the cover. Though $\pi(x)$ seems to follow no reasonable rule if one looks from close it seems to be rather regular in the large scale.

Theorem (Prime Number Theorem, without proof)**.** *The functions $\pi(x)$ and $\frac{x}{\log(x)}$ are asymptotically equal for $x \to \infty$ (i.e. the quotient $\pi(x)\log(x)/x$ tends to 1 for $x \to \infty$).*

2.2. Euclidean Division.

Definition (Greatest Common Divisor)**.** For integers a, b, not both zero, we call $\gcd(a,b) := \max\{d \in \mathbb{Z}_{\geq 1} : d \mid a, d \mid b\}$ the *greatest common divisor of a and b.*

Theorem. *For the* gcd *of positive a and b we have the formula*

$$\gcd(a,b) := p_1^{\min\{\alpha_1,\beta_1\}} \cdots p_r^{\min\{\alpha_r,\beta_r\}},$$

where $a = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ and $b = p_1^{\beta_1} \cdots p_r^{\beta_r}$ ($\alpha_i, \beta_i \geq 0$) denote the prime factorizations a and b.

Proof. Indeed, every divisor d of a and b must be of the form $d = p_1^{\gamma_1} \cdots p_r^{\gamma_r}$ with $\gamma_j \leq \alpha_j$ and $\gamma_j \leq \beta_j$, and vice versa, every integer of this form is a common divisor of a and b. The largest such integer is obtained by choosing γ_j equal to the minimum of α_j and β_j. This proves the theorem. □

Definition (Ideal)**.** A non-empty subset I of $\mathbb{Z}$ is called *ideal*, if for all $a, b \in I$ we have $a + b \in I$ and $a - b \in I$.

Remark. Note that 0 lies in any ideal. If a lies in an ideal, then any multiple of a also lies in the ideal.

Example. $\{0\}$ and $\mathbb{Z}$ are ideals. More generally, if d is an integer, then $\mathbb{Z}d = \{dx : x \in \mathbb{Z}\}$ is an ideal; it is called the *the principal ideal generated by d*. More generally, for arbitrary integers a_i, the set

$$\mathbb{Z}a_1 + \cdots + \mathbb{Z}a_r := \{a_1x_1 + \cdots + a_rx_r : x_1, \ldots, x_r \in \mathbb{Z}\}$$

forms an ideal.

Theorem (Principal ideal theorem)**.** *Every ideal is a principal ideal, i. e. one has $I = \mathbb{Z}d$ for a suitable d.*

For the proof we use:

Theorem (Euclidian Division)**.** *For given $m, q \in \mathbb{Z}$ and $q \neq 0$ there exist unique $x, r \in \mathbb{Z}$ such that $m = qx + r$ and $0 \leq r < |q|$.*

Example. $7 = -5 \cdot (-1) + 2$

Proof. Let r be the smallest number in

$$M := \big\{m - qx : x \in \mathbb{Z}\big\} \cap \mathbb{Z}_{\geq 0}.$$

Clearly, $m = qx+r$ and $0 \leq r$. If we had $r \geq |q|$, then $m-qx-|q| \in M$. But $m - qx - |q| < m - qx$ which is a contradiction to the minimality of $m - qx$.

The uniqueness of x and r is left as an exercise. □

Proof of the principal ideal theorem. If $I = \{0\}$ then $I = \mathbb{Z} \cdot 0$. Hence we can assume that I contains non-zero numbers. Then I contains a positive number (since with a number a it contains also $\pm a$). Let a be the smallest positive integer in I. We claim $I = \mathbb{Z}a$. Clearly, $I \supseteq \mathbb{Z}a$ (since $a \in I$). Let vice versa b in I. By Euclidean division we can write $b = xa + r$ for suitable x and $0 \leq r < a$. Writing $r = b - ax$ we see that I contains also $r < a$. But a is the smallest positive integer in I, hence $r = 0$. Therefore $b = ax$, that is $b \in \mathbb{Z}a$. □

Theorem (Bézout)**.** *For every pair $a, b \in \mathbb{Z}$, not both zero, there exist $x, y \in \mathbb{Z}$ with $ax + by = \gcd(a, b)$.*

Remark. For given a, b, c the equation $ax + by = c$ is solvable in integers x and y if and only if c divides the gcd of a and b.

Bézout's theorem is an immediate consequence of

Theorem. $\mathbb{Z}a + \mathbb{Z}b = \mathbb{Z}\gcd(a, b)$.

Proof. By the principal ideal theorem we know that the ideal $\mathbb{Z}a+\mathbb{Z}b$ is principal, that is $I := \mathbb{Z}a + \mathbb{Z}b = \mathbb{Z}g$ for a suitable positive integer g. Since a and b are in I we conclude that g divides both numbers, hence the $\gcd(a, b)$. Vice versa, the $\gcd(a, b)$ divides a and b, and hence g (since g equals $ax + by$ for suitable integers x and y). It follows $g = \gcd(a, b)$. □

Theorem (Euklid's Lemma)**.** *Let p be a prime and $a, b \in \mathbb{Z}$. Then $p \mid ab$ implies $p \mid a$ or $p \mid b$.*

Proof. Assume p does not divide a. Then $\gcd(p, a) = 1$ and hence, by Bézout's Theorem, $1 = px + ay$ for suitable x and y. Multiplying by b we obtain $b = pbx + aby$. Since p divides ab we conclude $p|b$. □

Remark. Inductively we obtain from the theorem the slightly more general statement: Is p prime, $p \mid a_1 \cdots a_r$, then $p \mid a_j$ for at least one j.

Consequence. *Note that this completes the proof of the uniqueness of the prime factorization of natural numbers.*

2.3. Euclid's algorithm. The most effective algorithm for computing the gcd of given integers is provided by *Euclid's Algorithm.* The simplest variant is based on the following lemma.

Lemma. *For all integers a, b and x, one has*

$$\gcd(a, b) = \gcd(a, b + ax).$$

Proof. Indeed, if g divides the left hand side it divides a and b and hence also a and $b+ax$, and hence the right hand side. If g divides the right hand side it divides a and $a+bx$, hence a and $b = (b+ax)-ax$, hence the left hand side. So both sides have the same divisors and are positive, so they are equal. □

Example. Successive application of the lemma (and the obvious rules $\gcd(a, b) = \gcd(b, a)$ and $\gcd(a, 0) = a$) yields an effective algorithm

for calculating the gcd of two numbers.

$$\begin{aligned}\gcd(102,27) &= \gcd(21 = 102 - 27\cdot 3, 27)\\ &= \gcd(21, 6 = 27 - 21\cdot 1)\\ &= \gcd(3 = 21 - 6\cdot 3, 6)\\ &= \gcd(3, 0 = 6 - 3\cdot 2) = 3.\end{aligned}$$

This is easy to put into a program[3]

Algorithm: Computation of the gcd of two positive integers

```
def my_first_gcd( a, b):
    while b > 0:
        c = b; b = a%b; a = c
    return a
```

We can do this also using recursion:

```
def my_second_gcd( a,b):
    return a if 0 == b else my_second_gcd(
        b, a%b)
```

If we keep track of all division steps of the preceding algorithm we can obtain at the same time also solutions x, y as in Bézout's Theorem, i.e. solutions of the equation $ax + by = \gcd(a, b)$. Namely, we start at the bottom of the last calculation, which tells us that the gcd of 102 and 27 is 3, and go up replacing at each level the remainder by the linear combination of the two preceding remainders. In our example this goes as follows:

[3]For describing algorithms we use the programming language *Python*. If you want to test or experiment with the code of this script you can easily install Python, which is freely available for almost any platform. You can, for example, install it in your Android cellphone (search in the Playstore for the app *QPython*). More advanced and also useful for other courses, you might want to use *Sage*, which is Python with mathematical libraries covering almost all parts of mathematics. If you would like to run the examples with Sage in your Web-Browser you might want to open your own Sage notebook in the *SageMathCloud* at https://cloud.sagemath.com/.

Example.

$$\begin{aligned}
3 &= 21 - \underline{6} \cdot 3 \\
&= \underline{2}1 - (27 - \underline{2}1 \cdot 1) \cdot 3 \\
&= (102 - 27 \cdot 3) - (27 - (102 - 27 \cdot 3) \cdot 1) \cdot 3 \\
&= 102 \cdot [1 + 1 \cdot 3] + 27 \cdot [-3 - 3 - 3 \cdot 1 \cdot 3] \\
&= 102 \cdot 4 + 27 \cdot (-15).
\end{aligned}$$

Replacing the remainders successively by a linear combination of the two preceding remainders is a recursive procedure. Therefore it is again extremely easy to put this into an algorithm. This can be done as follows.

Algorithm: Solving $ax + by = \gcd(a, b)$

```
def my_Bezout( a, b):
    if 0 == b: return 1,0
    x,y = my_Bezout( b, a%b)
    return y, x-(a//b)*y
```

It is sometimes useful to describe this extended Euclidean algorithm using matrices. For this we record the successive Euclidean divisions as follows:

$$\begin{aligned}
a &= a_0 b + r_1 & \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] &= \left[\begin{smallmatrix} a_0 & 1 \\ 1 & 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} b \\ r_1 \end{smallmatrix}\right] \\
b &= a_1 r_1 + r_2 & \left[\begin{smallmatrix} b \\ r_1 \end{smallmatrix}\right] &= \left[\begin{smallmatrix} a_1 & 1 \\ 1 & 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} r_1 \\ r_2 \end{smallmatrix}\right] \\
r_1 &= a_2 r_2 + r_3 & \left[\begin{smallmatrix} r_1 \\ r_2 \end{smallmatrix}\right] &= \left[\begin{smallmatrix} a_2 & 1 \\ 1 & 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} r_2 \\ r_3 \end{smallmatrix}\right] \\
&\vdots & &\vdots \\
r_{n-1} &= a_n r_n + 0 & \left[\begin{smallmatrix} r_{n-1} \\ r_n \end{smallmatrix}\right] &= \left[\begin{smallmatrix} a_n & 1 \\ 1 & 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} r_n \\ 0 \end{smallmatrix}\right],
\end{aligned}$$

where $n > r_1 > r_2 > \cdots > r_n > 0$, and where $r_n = \gcd(a, b)$. We then have

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} r_n \\ 0 \end{bmatrix}.$$

The matrix on the right is of the form $\begin{bmatrix} a/r_n & u \\ b/r_n & v \end{bmatrix}$. Its determinant is $(-1)^{n+1}$ (since the determinant of a matrix of the form $\left[\begin{smallmatrix} a & 1 \\ 1 & 0 \end{smallmatrix}\right]$ is -1). In other words, we have

$$a(-1)^{n+1}u + b(-1)^n v = \gcd(a, b),$$

which provides us with solutions x, y as in Bézout's theorem.

Definition (Least Common Multiple)**.** For $a, b \in \mathbb{Z}$, not both zero, we call the number

$$\operatorname{lcm}(a, b) := \min\{d \in \mathbb{Z}_{>0} : a|d,\ b|d\}$$

the *least common multiple of a and b.*

Theorem. *For the lcm of numbers a and b one has the formula*

$$\operatorname{lcm}(a, b) = p_1^{\max\{\alpha_1,\beta_1\}} \cdots p_r^{\max\{\alpha_r,\beta_r\}},$$

where $a = p_1^{\alpha_1} \cdot \cdots \cdot p_r^{\alpha_r}$ and $b = p_1^{\beta_1} \cdot \cdots \cdot p_r^{\beta_r}$ denote the prime factorizations of a and b.

As consequence of the formulas for the gcd and the lcm in terms of prime decompositions and the formula

$$\min\{\alpha_j, \beta_j\} + \max\{\alpha_j, \beta_j\} = \alpha_j + \beta_j$$

one obtains

Theorem. $\gcd(a, b) \cdot \operatorname{lcm}(a, b) = ab$.

Definition (GCD and LCM of more than two numbers)**.** For integers $a_1, \ldots, a_r$, not both zero, one defines

$$\begin{aligned} \gcd(a_1, \ldots, a_r) &:= \max\{d \in \mathbb{Z}_{\geq 0} : d \mid a_1, \ldots, d \mid a_r\} \\ \operatorname{lcm}(a_1, \ldots, a_r) &:= \min\{d \in \mathbb{Z}_{>0} : a_1 \mid d, \ldots, a_r \mid d\} \end{aligned}$$

Theorem. *One has the formulas*

$$\begin{aligned} a_1\mathbb{Z} + \cdots + a_r\mathbb{Z} &= \gcd(a_1, \ldots, a_r)\mathbb{Z} \\ a_1\mathbb{Z} \cap \cdots \cap a_r\mathbb{Z} &= \operatorname{lcm}(a_1, \ldots, a_r)\mathbb{Z} \end{aligned}$$

Proof. The first formula we saw already above for $r = 2$. The general case follows then on induction. The second formula we leave as an exercise. □

Remark. For $r \geq 3$ we have in general

$$\gcd(a_1 \cdots a_r) \cdot \operatorname{lcm}(a_1 \cdots a_r) \neq a_1 \cdots a_r.$$

The Euclidean algorithm as explained above can also easily be extended to more than two integers. Given a list of (say, positive) integers. one searches for the smallest one, replaces all integers by the remainder upon division by the smallest one, and then repeats this step until all but one integer are zero. An implementation could look like this.

Algorithm: Simultaneous computation of the gcd of a list of positive integers

```
def my_third_gcd( v):
    v = [ a for a in v if a != 0]
    v.sort()
    return v[0] if 1 == len(v) else
        my_third_gcd( [v[0]] + [a%v[0] for
        a in v[1:]])
```

Sometimes one can improve this algorithm slightly by modifying the Euclidean Division: instead of $b = aq+r$ with $0 \leq r < |a|$ one uses the modified Euclidean Division $b = aq' + s$ with $-|a|/2 < s \leq |a|/2$.

The careful reader might ask why we prefer the Euclidean algorithm for computing the gcd over the usual one that everybody performs in his head, namely using directly our definition of the gcd via the prime factorization of the integers in question. Here, for example $127 = 2 \cdot 3 \cdot 17$ and $27 = 3^3$, whence $\gcd(102, 27) = 3$ as we see immediately. This seems to be much shorter than the example calculation using the Euclidean algorithm which we gave above.

The answer is that this factorizing and then applying the defining formula for the gcd is, of course, the preferred method — for small

numbers though! If numbers become bigger there is a deep problem arising. Namely, we cannot factor too big numbers. There are many ingenious algorithms for factoring, but none of them is capable to surely factor an integer with a few thousand decimal digits. The most naive algorithm would try to factor a positive integer n by trying to find a divisor starting with 2, 3, etc. We can stop our search once we reach $\lfloor\sqrt{n}\rfloor$ since a composite number must obviously have at least one divisor smaller or equal to $\sqrt{n}$. If the number would be, for instance, a square of a prime we would have to perform indeed $\sqrt{n}$ many divisions before the factorization succeeds. Hence if n has 1000 decimal digits it could turn out that we have to try approximately $\sqrt{10^{1000}} = 10^{500}$ divisions before success. This is an incredibly large number as one learns if one asks a physicist. He would explain that there are approximately 10^{11} stars in our milky way. Hence it is much much faster to count the stars in our milky way than to factor a number with 1000 digits using the described naive algorithm. In fact, there are faster algorithms, but they are still all exponential in the number of digits of the given candidate for factoring. The conclusion is that we cannot be certain to compute the gcd of two numbers with several thousand digits using factorization.

But Euclid's algorithm can very well compute the gcd of two such numbers. Indeed, if we apply the Euclidean algorithm to numbers $b > a > 0$ then we expect that the remainder after the first division is in the average $a/2$. We then have to apply Euclidean division to numbers of the magnitude $a, a/2$, and we expect the remainder to be in the average $a/4$. So then we have in the third division to apply Euclidean division to numbers of the magnitude $a/2, a/4$, and we expect the remainder to be around $a/8$ etc. So we expect that the Euclidean algorithm terminates after $\log_2 a$ steps. In other words the number of divisions needed to compute the gcd of a, b is in the order of the number of binary digits of a. If a has 1000 decimal digits we would need approximately $1000 \log_2 10 \approx 3,300$ Euclidean divisions, which even the slowest smart phone would be able to perform the in a few seconds. The heuristic arguments given here can indeed be turned into a rigorous upper bound for the number of necessary divisions, and it turns out that the bound is in fact in the order of the number of binary digits of a.

We shall see later that computing gcds of big integers is indeed of practical interest, for example in cryptography based communication.

3. Congruences

3.1. Computing with congruences.

Definition. Let $a, b, m \in \mathbb{Z}$. We call a *congruent to* b *modulo* m (as formula: $a \equiv b \bmod m$ or $a \equiv_m b$), if $m \mid (a - b)$.

Remark. One has $a \equiv b \bmod 0$ if and only if $a = b$. Therefore, in the following, *the module* m will often be a non-zero integer, which allows to obtain identities modulo m which would not hold as identities for integers. Moreover, obviously $a \equiv b \bmod m$ if and only if $a \equiv b \bmod -m$. Hence, we do not loose any substantial part of theory when assuming occasionally that the module is positive.

Theorem. *For given* m *the relation* $\equiv_m$ *is an equivalence relation.*

We leave it to the reader to check for the given relation the axioms of an equivalence relation, namely that the relation is reflexive, symmetric and transitive.

Definition. The set of equivalence classes of the relation "congruent modulo m" is denoted by $\mathbb{Z}/m\mathbb{Z}$.

Theorem. *Assume* $m \neq 0$. *Then* $a \equiv b \bmod m$ *if and only if* a *and* b *leave the same remainder upon Euclidean Division by* m.

Proof. Write $a = mq + r$ and $b = mq' + r'$ with integers q, q' and $0 \leq r, r' < |m|$. Then $a - b = m(q - q') + r - r'$. From this it is clear that if m divides $a - b$ then it divides $r - r'$, and vice versa. But m divides $r - r'$ if and only if $r = r'$ as follows from $|r - r'| < |m|$. □

Hence, for $m \neq 0$, every equivalence class of the equivalence relation "congruent mod m" contains exactly one integer $0 \leq r < |m|$. Therefore, we have as many equivalence classes as residues, namely $|m|$ many. Moreover, an integer is congruent mod m to a given integer a if it is of the form $a + mx$, i.e. if it is contained in the set

$$a + m\mathbb{Z} := \{a + mx : x \in \mathbb{Z}\}.$$

Vice versa every number in this set is equivalent to a mod m. In particular, we have

$$a + m\mathbb{Z} = r + m\mathbb{Z},$$

where r is the remainder (or residue) of a after division by m. The equivalence classes of the relation "congruent mod m" are usually called *residue classes modulo* m, and the class containing a given integer a is called *residue class mod* m *of* a. We summarize:

Theorem. *Assume* $m \neq 0$.

(i) *The equivalence classes in* $\mathbb{Z}/m\mathbb{Z}$ *are of the form*

$$a + m\mathbb{Z} = \{a + mx : x \in \mathbb{Z}\}.$$

Vice versa, every such set is an equivalence class modulo m.

(ii) *One has* $\mathbb{Z}/m\mathbb{Z} = \{r + m\mathbb{Z} : 0 \leq r < m\}$. *The set* $\mathbb{Z}/m\mathbb{Z}$ *is in particular finite, one has* $\operatorname{card}(\mathbb{Z}/m\mathbb{Z}) = |m|$.

"Computing with congruences" is based on the following rules.

Theorem. *Let* m *be an integer' and assume that' for given integers* a, a', b *and* b'*' one has* $a \equiv a' \bmod m$ *and* $b \equiv b' \bmod m$. *Then*

(i) $a + b \equiv a' + b' \bmod m$, *and*

(ii) $ab \equiv a'b' \bmod m$.

Proof. By assumption we know that m divides $a - a'$ and $b - b'$. For proving (i) we write

$$(a + b) - (a' + b') = (a - a') + (b - b'),$$

from which it is obvious that the left hand side is divisible by m.

For (ii) we have to be a bit more tricky, namely we write

$$ab - a'b' = (a - a')b + a'(b - b'),$$

which again makes (ii) obvious. □

Example. We give an example for how to use the preceding rules. The reader has probably seen already the following fact:

> A given positive integer n is divisible by 9 if and only if its digit sum[4]. is divisible by 9.

[4]By *digit sum of* n one means the sum of the digits of the decimal expansion of n

For example, 123456789 is divisible by 9 since $1+2+\cdots+9 = 45$ is so. For proving the given rule we note that $10 \equiv 1 \bmod 9$. By (successive) application of (ii) this gives $10 \cdot 10 \equiv 1 \bmod 9$, $10 \cdot 10 \cdot 10 \equiv 1 \bmod 9$ and so forth. If z_l, z_{l-1}, $\ldots$, z_0 are the decimal digits of n, then

$$n = z_l \cdot 10^l + z_{l-1} \cdot 10^{l-1} + \cdots + z_1 \cdot 10 + z_0.$$

By (successive) application of (i) and (ii) we can replace here modulo 9 all powers of 10 by 1, i.e.

$$n \equiv z_l + z_{l-1} + \cdots + z_1 + z_0 \bmod 9.$$

The given rule is now obvious. Note that we have actually proved more, namely that a number leaves the same remainder upon Euclidean division by 9 as its digit sum.

We encourage the reader to work out similar division rules for division by 3, 11, 2, 4, 8, 5, 25.

We saw that we can work with congruences like with identities: We can replace in congruences modulo m left hand sides of another congruence mod m by its right hand side, we can multiply or add respective side of congruences modulo m so to obtain another congruence modulo m. However, the cancellation law does not hold true in general, i.e. from a given congruence $ka \equiv kb \bmod m$ we can in general not deduce $a \equiv b \bmod m$. For example, we have $2 \equiv -2 \bmod 4$, whereas $1 \not\equiv -1 \bmod 4$. The following theorem shows that under a certain assumption we can still apply cancellation.

Theorem. *Assume* $\gcd(k, m) = 1$. *Then* $ka \equiv kb \bmod m$ *implies* $a \equiv b \bmod m$.

Proof. By assumption and Bézout's Theorem there exist integers x and y such that $kx + my = 1$. We have to show that m divides $a - b$, and we know that m divides $k(a - b)$. For this we write

$$a - b = 1 \cdot (a - b) = (kx + my)(a - b) = kx(a - b) + my(a - b).$$

Since the right hand side is divisible by m, the claim follows. □

A residue class modulo m is called *primitive* if its members are relatively prime to m. Note that we only have to check for one member

that it is relatively prime to m to be certain that this holds true for all. We note an important special case of the preceding theorem:

Corollary. *Let p be a prime number. Then $ka = kb \bmod p$ implies $a = b \bmod p$, provided $k \not\equiv 0 \bmod p$.*

This rule resembles very much that fact that we may divide both sides of an identity for, say, rational or real numbers by a nonzero number. The deeper reason for this is revealed in algebra, where one learns that the set $\mathbb{Z}/p\mathbb{Z}$ of residue classes modulo a prime p form a *field*.

The preceding theorem can also be obtained as a consequence of the following stronger statement.

Theorem. *For given integers m and k there exists an integer k' such that $kk' \equiv 1 \bmod m$ if and only if $\gcd(k, m) = 1$.*

Proof. Assume $kk' \equiv 1 \bmod m$ for some integer k'. This means that $kk' = 1 + my$ for suitable k' and y, which implies that k and m are relatively prime. Vice versa, if $\gcd(k, m) = 1$ then by Bézout's Theorem $1 = kx+my$ for suitable x and y, and so, $kx \equiv 1 \bmod m$. □

Again we have as special case:

Corollary. *Let p be a prime number. Then, for every integer k which is not divisible by p, there exists an integer k' such that $kk' \equiv 1 \bmod p$.*

It is often necessary to calculate an inverse of an integer modulo a given m. For small m the easiest way is to try. For example, if we want to invert 2 modulo 5, we use that there are only 4 primitive residue classes modulo 5. Thea are represented by the possible non-zero residues modulo 5, i.e. by 1, 2, 3 and 4. Hence we multiply 2 by each of these until the result equals 1 modulo 5. We leave it to the reader to find the inverse of 2 mod 5 in this way. For larger modules m trying will not be possible. However, as we saw in the last proof computing the inverse modulo m of an integer amounts essentially to solve Bézout's equation $kx + my = 1$, which in turn is done by the extended Euclidean Algorithm. amounts essentially to

Algorithm: Computation of the inverse modulo m

```
def inv( k, m):
    x,y = my_Bezout( k,m) # see
        preceding section
    return x
```

There are obvious other rules for computing with congruences whose discovery and proof we leave to the reader (he will stumble over them once he starts to compute with congruences by himself). However, we mention the following. If $a \equiv b \bmod m$, then, for every integer k, we have $ak \equiv bk \bmod mk$. And vice versa, if $ka \equiv kb \bmod m$, k divides m and $k \neq 0$, then $a \equiv b \bmod m/k$.

3.2. The Chinese remainder theorem.

Theorem (Chinese Remainder Theorem)**.** *Let $m_1, \ldots, m_r$ be pairwise relatively prime positive integers. Let $a_1, \ldots, a_r \in \mathbb{Z}$. Then there is a solution x of the simultaneous congruences*

$$x = a_j \bmod m_j \qquad (1 \leq j \leq r).$$

Such a solution is modulo $m := m_1 \cdots m_r$ unique (i.e. if x' is another solution; then $x \equiv x' \bmod m$.

This theorem was indeed as far as one knows first written up in ancient China several thousand years ago. This is not surprising since the theorem, and in particular its proof, is of quite practical interest. Think of periodically recurring events (like star or planet constellations) which occur every m_1, m_2, ...years, respectively. If the first event occurred in year a_1, the second in year a_2, ..., is then there a year where all occur at the same time? The Chinese Remainder Theorem gives an affirmative answer if the periods are pairwise relatively prime. And what is the closest year in the future when all events do occur at the same time. Again, the proof of the Chinese Remainder Theorem will show how to compute this year.

Proof of the Chinese Remainder Theorem. For finding an integer x as in the theorem we set up a table as follows:

$$\begin{array}{ccccc} a_1 & m_1 & m/m_1 & m_1' & a_1(m/m_1)m_1' \\ a_2 & m_2 & m/m_2 & m_2' & a_2(m/m_2)m_2' \\ \vdots & \vdots & & & \vdots \\ a_r & m_r & m/m_r & m_r' & a_r(m/m_r)m_r' \end{array}$$

We let x be the sum of the entries of the last column, i.e. we set

$$x := a_1(m/m_1) \cdot m_1' + a_2(m/m_2) \cdot m_2' + \cdots + a_r(m/m_r) \cdot m_r'.$$

Here m_j' is an inverse modulo m_j of m/m_j, respectively, i.e. as solution of $m_j' \cdot m/m_j \equiv 1 \bmod m_j$. Note that such m_j' exist since m_j and m/m_j are relatively prime by assumption. We also know how to compute m_j' effectively as we learned in the last section. We leave it to the reader to verify that the so constructed x satisfies $x \equiv a_1 \bmod m_1$, $x \equiv a_2 \bmod m_2$ etc..

The uniqueness is easy to see: if x' is another solution, then $x \equiv x' \bmod m_j$ for all j. Therefore x-x' is divisible by all m_j, and since the m_j are pairwise relatively prime, we conclude that $x - x'$ must be divisible by the product of all m_j. □

Following the procedure described in the proof it is easy to let a computer find an x as in the theorem.

Algorithm: Solving simultaneous congruences

```
def prod(lst):
    """
    Return the product of the objects in
       the list lst.
    """
    pr = 1
    for x in lst:
        pr *= x
    return pr
```

```
def my_Chinese( d):
    """
    Return the smallest positive
        simultaneous solution x
    to the congruences
         x = d[n] mod n
    where n runs through the keys of the
        dictionary d.
    The keys must be pairwise relatively
        prime.
    """
    m = prod( d.keys())
    table = [( m/n, inv(m/n,n), d[n]) for n
         in d]
    x = sum( [a*b*c for a,b,c in table])
    return table,m,x,x%m

d = {3:2, 5:4, 7:6}
my_Chinese(d)
```

We note a theoretical consequence which is extremely important for solving congruences and for counting the solutions modulo m of a given congruence modulo m.

Let $f(x_1, \ldots, x_s)$ be a polynomial in s variables with integral coefficients, and let $m > 0$ be an integer. Assume that for each prime power $p^\alpha \parallel m$[5] we have a solution $\vec{x}_p \in \mathbb{Z}^s$ of the congruence

$$f(\vec{x}_p) \equiv 0 \bmod p^\alpha.$$

By the Chinese Remainder Theorem there exists an $\vec{x} \in \mathbb{Z}^r$, such that for every prime power $p^\alpha \parallel m$ we have $\vec{x} \equiv \vec{x}_p \bmod p^\alpha$. (These congruences are to be read and solved component by component.)

[5] We write $t \parallel m$ and call t an *exact divisor of* m, if t is a divisor of m such that $\gcd(t, m/t) = 1$.

For such an x we than have also

$$f(\vec{x}) \equiv 0 \bmod m.$$

Moreover, every solution $\vec{x}$ of the preceding congruence is obtained in such a way.

If we set

$$a(m) := \operatorname{card}\big\{(x_1, \ldots, x_r) \in \mathbb{Z} : \\ 0 \le x_1, \ldots, x_r < m, f(x_1, \ldots, x_r) \equiv 0 \bmod m\big\},$$

then the considerations of the last paragraph show

$$a(m) = \prod_{p^\alpha \| m} a(p^\alpha).$$

This formula has to be understood in the sense that the p^α run over all prime powers exactly dividing m.

It is sometimes useful to rewrite the Chinese Remainder Theorem in terms of maps. For this we define the *reduction map from* $\mathbb{Z}/m\mathbb{Z}$ *to* $\mathbb{Z}/n\mathbb{Z}$ for divisors $n|m$ as the map

$$\operatorname{red}_{m,n} : \mathbb{Z}/m\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}, \quad a + m\mathbb{Z} \mapsto a + n\mathbb{Z}.$$

The reader should verify that this map is *well-defined*. This means the following: If we take another element b in $C := a + m\mathbb{Z}$, then $a+m\mathbb{Z} = b+m\mathbb{Z}$. Therefore, we have suddenly two definitions for $red_{m,n}(C)$, namely $a + n\mathbb{Z}$ and $b + n\mathbb{Z}$, and our definition make sense only if these two expressions define the same residue class modulo n.

Using the reduction map the Chinese Remainder Theorem can be restated as follows:

Theorem (Chinese Remainder Theorem, map theoretical formulation)**.** *Let $m_1, \ldots, m_r$ be pairwise relatively prime positive integers, and set $m = m_1 \cdots m_r$. The map*

$$\operatorname{red}_{m,m_1} \times \cdots \times \operatorname{red}_{m,m_r} : \mathbb{Z}/m\mathbb{Z} \longrightarrow \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_r\mathbb{Z} \\ a + m\mathbb{Z} \mapsto (a + m_1\mathbb{Z}, \ldots, a + m_r\mathbb{Z}).$$

is bijective.

Indeed, the first part of the classical formulation of the Chinese Remainder Theorem says that our map is surjective, whereas the second part says it is injective.

3.3. Algebraic congruences mod p^n. As we saw in the discussion succeeding the Chinese remainder theorem in Section 3.2, given a polynomial $f(x_1, \dots, x_s)$ in s variables with integral coefficients, any congruence $f(x_1, \dots, x_s) \equiv 0 \bmod m$ can be reduced to the corresponding congruences modulo the prime powers p^n exactly dividing m. However, a congruence modulo a prime power p^n can often be reduced to the congruence modulo p. The advantage lies at hand. For finding a solution

$$f(x_1, \dots, x_s) \equiv 0 \bmod p$$

there is in general no better method than trying systematically all possible elements of $(\mathbb{F}_p)^s$ for being a solution. If p and s are sufficiently small such a search can be done, whereas a corresponding search modulo p^2 would already square the amount of trials. The mentioned method of reduction is an adaption of Newton's method for finding real roots of a polynomial in one variable. It is explained in the proof of the following theorem.

Theorem (Newton's method)**.** *Let f be a polynomial in s variables with integral coefficients, and p^n $(n \geq 1)$ a prime power. Assume*

$$f(x_1, \dots, x_s) \equiv 0 \bmod p^n, \quad \nabla f(x_1, \dots, x_s) \not\equiv 0 \bmod p.$$

Then there exists a solution

$$f(y_1, \dots, y_s) \equiv 0 \bmod p^{n+1} \quad \textit{with} \quad y_1, \dots, y_s \equiv x_1, \dots, x_s \bmod p^n.$$

Here ∇f is the vector of length s whose jth entry is the partial derivative of f with respect to the jth variable.

Proof of the theorem. Write $\vec{y}$ for $(y_1, \dots, y_s)$ and similar for the vector of the x_j. For the desired solution $\vec{y} \equiv \vec{x} \bmod p^n$ we make the ansatz

$$\vec{y} = \vec{x} + p^n \vec{t}$$

with a vector $\vec{t}$ so that we have to solve

$$f(\vec{x} + p^n \vec{t}) \equiv 0 \bmod p^{n+1}.$$

Here $\vec{y} = (y_1, \ldots, y_s)$ and $\vec{x} = (x_1, \ldots, x_s)$.

We expand f around $\vec{x}$ and observe that all higher terms apart from the constant and linear ones vanish modulo p^{n+1}:

$$f(\vec{x} + p^n\vec{t}) \equiv f(\vec{x}) + p^n \nabla f(\vec{x}) \cdot \vec{t} \bmod p^{n+1},$$

where the dot on the right is the usual scalar product of row vectors. The congruence of our ansatz becomes therefore

$$-\frac{1}{p^n} f(\vec{x}) \equiv \nabla f(\vec{x}) \cdot \vec{t} \bmod p.$$

But this congruence is solvable in $\vec{t}$ since we assumed that $\nabla f(\vec{x})$ is not zero modulo p. Note that the solutions $\vec{t}$ form an affine subspace of $\mathbb{F}_p^s$ of co-dimension 1. In particular, $\vec{t}$ is unique if $s = 1$. □

As we saw in the proof the case $s = 1$ is especially interesting.

Corollary. *Let f be a polynomial in one variable with integral coefficients, and p a prime number. Assume $f(y_1) \equiv 0 \bmod p$ and $f'(y_1) \not\equiv 0 \bmod p$ Then, for any n, there exists exactly one solution y_n modulo p^n of $f(y_n) \equiv 0 \bmod p^n$ with $y_n \equiv y_1 \bmod p$.*

From the uniqueness we deduce $y_{n+1} \equiv y_n \bmod p^n$. If we set $y_{n+1} = y_n + t_n p^n$, and $t_0 = y_1$, then $y_{n+1} = \sum_{\nu=0}^{n} t_\nu p^\nu$. Note that we can assume that the y_n have been chosen so that $0 \le t_\nu < p$. The sums look like the partial sums of a p-adic expansion of some object, and it is natural to ask what object this might be. The interested reader can find the answer in Section 4.2.

3.4. Primitive residue classes.

Definition. A residue class modulo m is called *primitive* if all its elements are relatively prime to m. We denote the set of primitive residue classes modulo m by $(\mathbb{Z}/m\mathbb{Z})^*$.

Remark. The reader should verify that a residue class modulo m is primitive if at least one of its elements is relatively prime to m.

Definition. *Euler's ϕ-function*[6] φ is defined on the set of positive integers and its values are

$$\varphi(m) := \operatorname{card}\left((\mathbb{Z}/m\mathbb{Z})^*\right) \qquad (m \ge 1).$$

[6] Euler's φ-function is sometimes also called *Euler's totient function*

In other words, since every residue class is represented by an integer $0 \leq r < m$, we have

$$\varphi(m) = \operatorname{card}\left(\left\{0 \leq r < m : \gcd(r, m) = 1\right\}\right),$$

or, equivalently, that $\varphi(m)$ equals the number of fractions $0 \leq x < 1$ whose denominator in shortest form is m.

Example. The first values of Euler's *phi*-function are

m	1	2	3	4	5	6	7	8	9	10	11	12	24
$\varphi(m)$	1	1	2	2	4	2	6	4	6	4	10	4	8

The table suggests the following theorem:

Theorem. *Let $m_1, \ldots, m_r$ be pairwise relatively prime positive integers, set $m = m_1 \cdots m_r$. Then $\varphi(m) = \varphi(m_1) \cdots \varphi(m_r)$.*

Proof. For this one checks that the map

$$\operatorname{red}_{m,m_1} \times \cdots \times \operatorname{red}_{m,m_r}$$

from the preceding section defines after restriction a bijection

$$(\mathbb{Z}/m\mathbb{Z})^* \longrightarrow (\mathbb{Z}/m_1\mathbb{Z})^* \times \cdots \times (\mathbb{Z}/m_r\mathbb{Z})^*.$$

From this the claimed formula is obvious. □

Lemma. *For prime powers p^α one has $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}$.*

Proof. The primitive residue classes modulo p^α are represented by those numbers from the list $0, 1, \ldots, p^\alpha - 1$ which are not divisible by p. But there are exactly $p^{\alpha-1}$ numbers in the list which are divisible by p, namely the numbers $0, p, 2p, 3p, \cdots, (p^{\alpha-1} - 1) \cdot p$. If we suppress these from the list, exactly $p^\alpha - p^{\alpha-1}$ number remain. This proves the lemma. □

The last theorem and the last lemma imply the following formula for $\varphi(m)$:

Theorem. *For any positive integer m, one has the formula*

$$\varphi(m) = m \prod_{p \mid m} \left(1 - \frac{1}{p}\right).$$

Here p runs through the (pairwise different) prime divisors of m.

Note that, for any m, we have $\varphi(m) \leq m-1$, with equality if and only if m is a prime.

After having determined the number of primitive residue classes modulo a given number m we study now a bit deeper the structure provided by these classes. The first thing to remark is that the product of two numbers which define a primitive residue class modulo m does so too. We shall tacitly apply this in the following.

Theorem. *Let m be a positive integer and a be an integer which is relatively prime to m. There exist an $n > 0$ such $a^n \equiv 1 \bmod m$.*

Proof. The powers a^k, where k runs through the positive integers cannot all be pairwise incongruent modulo m since there are at most m residue classes. Therefore there exist integers positive $k < l$ such that $a^k \equiv a^l \bmod m$. Choose an inverse a' of a modulo m, and multiply the last identity by a'^k. It follows $1 \equiv a^{l-k} \bmod m$. □

Definition. The smallest positive integer n such that $a^n \equiv 1 \bmod n$ is called the *order of a modulo m*.

Theorem. *Let m be a positive integer, a relatively prime to m and n be the order of a modulo m. Then*

$$\{a^k + m\mathbb{Z} : k \in \mathbb{Z}_{\geq 0}\} = \{a^k + m\mathbb{Z} : 0 \leq k < n\}.$$

This theorem is an immediate consequence of

Theorem. *Let a and $m > 0$ be relatively prime integers, and let n be the order of a modulo m. Then $a^k \equiv a^l \bmod m$ if and only if $k \equiv l \bmod n$.*

Proof. If $k = l + nx$ then clearly $a^k \equiv a^l \bmod m$. Assume the latter congruence. Without loss of generality we may assume $k < l$. Multiplying the congruence by a'^k for an inverse a' modulo m of a, we obtain $a^{l-k} \equiv 1 \bmod m$. Let r be the remainder of $l-k$ after division by n. Again it follows $a^r \equiv 1 \bmod m$. Since $r < n$ and n is the smallest positive integer such that $a^n \equiv 1 \bmod m$ we deduce $r = 0$, i.e. that n divides $l-k$. □

Definition. Let m be a positive integer. An integer w is called *primitive root modulo* m , if

$$\{w^n + m\mathbb{Z} : n \in \mathbb{Z}\} = (\mathbb{Z}/m\mathbb{Z})^*.$$

From the preceding theorem we deduce that a is a primitive root modulo m if and only if the order of a modulo m equals $\varphi(m)$. However, it is not at all clear whether a primitive root modulo m exists at all.

Example. The number $a := 10$ is a primitive root modulo 7: $10 \equiv_7 3$, $100 \equiv_7 2$, $1000 \equiv_7 6$, $10000 \equiv_7 4$, $100000 \equiv_7 5$, $1000000 \equiv_7 1$,

Example. The reader should check that, for every odd number a, one has $a^2 \equiv 1 \bmod 8$. Therefore the order of an odd number modulo 8 is 1 or 2. But $(\mathbb{Z}/8\mathbb{Z})^*$ has four elements. Therefore there exists no primitive root modulo 8.

We shall come back to the question which m possess primitive roots. However, for this and only because its interesting for its own sake we study, first of all, the notion of "order modulo m".

Theorem (Fermat's Little Theorem)**.** *Let p be a prime. For every integer x one has $x^p \equiv x \bmod p$.*

Proof. The claimed congruence is obviously correct if x is divisible by p. So assume that p does not divide x. We have

$$x^{p-1} \prod_{j=1}^{p-1} j = \prod_{j=1}^{p-1} (xj).$$

If we reduce both sides modulo p we observe that the product on the right hand side is congruent modulo p to the product $P := \prod_{j=1}^{p-1} j$ since the set of all numbers xj ($1 \le j \le p-1$ represents also all residue classes modulo p (indeed, if $xj \equiv xj' \bmod p$ then $j \equiv j' \bmod p$ since x is not divisible by p). It follows $x^{p-1}P \equiv P \bmod p$, and since P is not divisible by p, then $x^{p-1} \equiv 1 \bmod p$, or equivalently $x^p \equiv x \bmod p$. □

As an immediate consequence one obtains the binomial theorem modulo p.

Corollary. *For any two integers x, y, one has*

$$(x+y)^p \equiv x^p + y^p \bmod p.$$

Note that, by the usual binomial theorem, the corollary is equivalent to the statement that $\binom{p}{k}$ is divisible by p for every $1 \leq k \leq p-1$. It is not hard, however, to prove this directly by using the formula

$$\binom{p}{k} = \frac{p(p-1)\cdots(p-k+1)}{k!},$$

and noting that $k!$ is not divisible by p.

Fermat's Little Theorem should not be confused with Fermat's last Theorem, whose proof was a long outstanding problem in number theory for several hundred years and which was finally proved 20 years ago.

Theorem (Fermat's Last Theorem)**.** *The equation $a^n + b^n = c^n$ for $n > 2$ does not possess any integral solutions with $abc \neq 0$.*

Fermat's Little Theorem generalizes to arbitrary modules. Note that it implies that $x^{p-1} \equiv 1 \bmod p$ if x is not divisible by p (in fact, we used this in the proof). In this form it can be quickly generalized to arbitrary modules m.

Theorem (Euler)**.** *Let x and $m > 0$ be relatively prime integers. Then $x^{\varphi(m)} \equiv 1 \bmod m$.*

Recall that $\phi(m) = m - 1$ if m is a prime, in which case Euler's theorem becomes Fermat's Little Theorem. The proof of Euler's theorem is almost identical to the proof of Fermat's Little Theorem, and we leave the details as an exercise.

As immediate consequence we obtain:

Theorem. *Let a and $m > 0$ be relatively prime integers. Then the order of a modulo m divides $\varphi(m)$.*

Fermat's Little Theorem implies a (probabilistic) primality test: Given a positive integer m, check randomly chosen x which are relatively prime to m whether they satisfy $x^{m-1} \equiv 1 \bmod m$. If some x does not possess this test, then m cannot be a prime number. This

primality test fails however for *Carmichael numbers*. These are composite numbers m which satisfy $x^{m-1} \equiv 1 \bmod m$ for all x relatively to m. Such numbers do indeed exist. We leave it to the reader to find the first Carmichael Number.

We come back to the question which modules m possess primitive roots. We start with prime modules. Here the answer is easy.

Theorem. *Every prime p possesses a primitive root modulo p.*

For the proof we need two lemmas, which are interesting for its own sake.

Lemma. *Let p be a prime, and let $a_n, \ldots, a_0$ be integers not all divisible by p. The congruence*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 \equiv 0 \bmod p$$

possesses at most n solutions modulo p.

Proof. The left hand side of the congruence in question can be viewed as a polynomial, which we denote by $f(x)$. We proceed by induction over n. If $n = 0$, then $f(x) = a_0$, and since by assumption $a_0 \not\equiv 0 \bmod p$ there is no solution. Suppose now that f is a polynomial of degree $\leq n+1$ and $f(x_0) \equiv 0 \bmod p$. We shall show that f has at most n-many solutions modulo p which are different from x_0. For this, we write

$$\begin{aligned} f(x) &\equiv f(x) - f(x_0) \bmod p \\ &\equiv \sum_{k=0}^{n+1} a_k x^k - \sum_{k=0}^{n+1} a_k x_0^k \bmod p \\ &\equiv \sum_{k=0}^{n+1} a_k (x^k - x_0^k) \bmod p. \end{aligned}$$

using the formula

$$(x^k - x_0^k) = (x - x_0)(x^{k-1} + x^{k-2} x_0 + \ldots + x_0^{k-1}),$$

we obtain

$$f(x) \equiv (x - x_0)\big[a_0 + \sum_{k=1}^{n+1} a_k (x^{k-1} + x^{k-2} x_0 + \ldots + x_0^{k-1})\big] \bmod p.$$

Denote the sum on the right by $g(x)$. Then $g(x)$ defines a polynomial of degree $\leq n$. If $x_1 \not\equiv x_0 \bmod p$ and $f(x_1) \equiv 0 \bmod p$, then clearly $g(x_1) \equiv 0 \bmod p$. However, by induction hypothesis the congruence $g(x) \equiv 0 \bmod p$ possesses at most n solutions modulo p. □

Lemma. *One has*

$$\sum_{d|m} \varphi(d) = m.$$

Here the sum is over all positive divisors d of m.

Proof. Consider the m fractions

$$\frac{0}{m}, \frac{1}{m}, \frac{2}{m}, \ldots, \frac{m-1}{m}.$$

The denominator of any of these fractions in shortest form is a divisor d of m. The fractions in shortest form with denominator d are the fractions $\frac{k}{d}$, where $0 \leq k < d$ and $\gcd(k, d) = 1$. There are exactly $\varphi(d)$ many such fractions. The claimed formula is now obvious. □

Proof of the theorem on primitve roots mod primes. We need to show that there is a number w, not divisible by p, whose order modulo p equals $p - 1$. We know that the order of any a modulo p is a divisor of $p - 1$. For a divisor d of $p - 1$ let $A(d)$ denote the set of numbers $1 \leq a \leq p - 1$ whose order equals d. We shall show in a moment

$$\#A(d) \leq \varphi(d).$$

But it is clear that $\sum_{d|p-1} \#A(d) = p - 1$ (since every $1 \leq a \leq p - 1$ must occur in exactly one $A(d)$). On the other hand $\sum_{d|p-1} \varphi(d) = p-1$. Both identities are only possible if $\#A(d) = \varphi(d)$. In particular, $\#A(p - 1) = \varphi(p - 1) > 0$, which proves our theorem.

It remains to prove the claimed inequality. Assume that $A(d)$ is not empty and let a be an element of $A(d)$. We then have

$$\big\{a^k + p\mathbb{Z} : 1 \leq k \leq d\big\} = \big\{x + p\mathbb{Z} : x^d \equiv 1 \bmod p\big\}.$$

Indeed the set on the left has d elements (see theorem above), it is obviously a subset of the set on the right, and by the first lemma above the set on the right cannot have more than d elements (consider the polynomials $x^d - 1$). We conclude that $A(d)$ is contained in the left hand side, and it remains to count the powers a^k whose order modulo

p equals d. But if $a^{kt} \equiv 1 \bmod p$ if and only if d divides kt, and therefore the order of $a^k \bmod p$ equals d if and only if k is relatively prime to d. The claimed inequality follows. □

In fact, more is true:

Theorem. *For every odd prime p there exists an integer w which is a primitive root modulo any power p^α.*

Remark. The proof will actually show more. Namely, if w is a primitive root modulo p, then the order of w modulo p^2 equals $p-1$ or $p(p-1)$. In the second case w is a primitive root modulo any p^α, whereas in the first case $w+p$ is is a primitive root modulo any p^α.

Proof of the theorem. Let w be a primitive root modulo p. The order n of w modulo p^2 divides $\varphi(p^2) = p(p-1)$, and since in particular $w^n \equiv 1 \bmod p$ we see that $p-1$ divides n. Therefore $n = p(p-1)$ or $p-1$. In the latter case $(w+p)^{p-1} \equiv w^p + (p-1)w^{p-1}p \equiv 1 + (p-1)w^{p-1}p \bmod p^2$. Thus replacing w by $w+p$ if necessary we can assume that w is a primitive root modulo p^2. We claim that w is a primitive root modulo all powers of p.

For this we show by induction over α that

$$w^{\varphi(p^{\alpha-1})} = 1 + p^{\alpha-1}n_\alpha$$

with some number n_α which is not divisible by p. By choice of w this is true for $\alpha = 1$ and $\alpha = 2$. Assume it is true for some $\alpha \geq 2$. Then

$$w^{\varphi(p^\alpha)} = (1 + p^{\alpha-1}n_\alpha)^p = 1 + p^\alpha n_\alpha + \binom{p}{2} p^{2\alpha-2} n_\alpha^2 + \cdots .$$

But the third, fourth term etc. on the right is divisible by $p^{\alpha+1}$ since, for $t \geq 3$, we have $\alpha + 1 \leq t(\alpha - 1)$ (as follows from $\frac{\alpha+1}{\alpha-1} = 1 + \frac{2}{\alpha-1} \leq 3$ for $\alpha \geq 2$), and since $\alpha \leq 2\alpha - 2$ (for $\alpha \geq 2$) and $p \mid \binom{p}{2}$. (Note that the latter is not true for $p = 2$). If we write the right hand side as $1 + p^\alpha n_{\alpha+1}$ we see that $n_{\alpha+1} \equiv n_\alpha \bmod p$, i.e. that $n_{\alpha+1}$ is not divisible by p.

The claim now follows easily. Namely, let n be the order of w modulo p^α. As before we have, first of all, that n divides $\varphi(p^\alpha) =$

$p^{\alpha-1}(p-1)$, and that $p-1$ divides n. Hence $n = p^u(p-1)$ with $u \leq \alpha - 1$. Since

$$w^{p^{\alpha-2}(p-1)} = w^{\varphi(p^{\alpha-1})} = 1 + p^{\alpha-1}n_\alpha \not\equiv 1 \bmod p^\alpha$$

it follows $\alpha - 2 < u$, and therefore $u = \alpha - 1$ as was to be shown. □

The preceding theorem is not true for powers of 2 as we saw in an example above where we considered the module 8 which provides a counter example. For powers of 2 one has instead the following whose proof we leave to the ambitious reader.

Theorem. *For every odd integer a and every power 2^α there exists an $n \geq 0$ such that $a \equiv \pm 5^n \bmod 2^\alpha$.*

We finally can answer the question which positive integers m possess primitive roots.

Theorem. *A positive integer possesses a primitive root if and only if it is of the form 2, 4, p^s, or $2p^s$, where p is an odd prime and s a positive integer.*

Proof. It is obvious that 2 and 4 possess primitive roots, we proved that p^s possesses a primitive root, and any odd primitive root of p^s is one for $2p^s$.

Vice versa assume that m possesses a primitive root. Then the order of w modulo m is $\varphi(m)$. On the other hand side, by Euler's theorem $w^{\varphi(p^n)} \equiv 1 \bmod p^n$ for every prime power p^n relative prime to w. From the Chinese remainder theorem we deduce that therefore $w^N \equiv 1 \bmod m$, where N denotes the least common multiple of all $\varphi(p^n)$ ($p^n \parallel m$). Since $\varphi(m)$ is the product of all theses $\varphi(p^n)$ we conclude $N \leq \varphi(m)$, and then (since $\varphi(m)$ is the order of m) that $N = \varphi(m)$. But the latter implies that m contains at most one odd prime power (since $\varphi(p^n)$ is even for any odd p), and if m contains an odd prime power, then m cannot be divisble by 4 (since $\varphi(4) = 2$). If m is a power of 2, then it equals $m = 2$ or $m = 4$ since 8 (and accordingly any higher 2-power) possesses no primitive root. This proves the theorem. □

3.5. Sums of two squares.

Theorem (Wilson)**.** *For any prime p, one has $(p-1)! \equiv -1 \bmod p$.*

Proof. If $k \in \{1, \ldots, p-1\}$, then k is relatively prime to p, and so possesses an inverse modulo p, which after reducing modulo p is also contained in this set. We shall show in a moment that only the elements 1 and $p-1$ are their own inverses modulo p. Thus, the elements $2, \ldots, p-2$ must split up into pairs $\{x, x^{-1}\}$. It follows that their product is 1. Hence,

$$(p-1)! = 1 \cdot (p-1) \equiv -1 \bmod p.$$

It remains to prove that for $0 < k < p$, we have that $k^2 \equiv 1 \bmod p$ if and only if $k = 1$ or $k = p-1$. If $k = 1$ or $k = p-1$, then $k^2 \equiv 1 \bmod p$. Conversely, suppose that $k^2 \equiv 1 \bmod p$. Then

$$p | k^2 - 1 = (k-1)(k+1),$$

and since p is prime, $p|k-1$ or $p|k+1$. The only number in the set $\{1, \ldots, p-1\}$ which satisfies $p|k-1$ is $k = 1$, and the only number in $\{1, \ldots, p-1\}$ which satisfies $p|k+1$ is $p-1$. □

Theorem. *Let p be an odd prime. Then $x^2 \equiv -1 \bmod p$ is solvable if and only if $p \equiv 1 \bmod 4$.*

Proof. Let w denote a primitive root mod p. Recall that $-1 \equiv w^{\frac{p-1}{2}} \bmod p$. Therefore if if $\frac{p-1}{2}$ is even then $x = w^{\frac{p-1}{4}}$ is a squareroot of -1 modulo p. Vice versa, if $x^2 \equiv -1 \bmod p$ is solvable, say, with $x \equiv w^n \bmod p$, we conclude $\frac{p-1}{2} \equiv 2n \bmod p-1$, in particular, that $\frac{p-1}{2}$ must be even. □

We remark that Wilson's theorem gives us, for a prime number $p \equiv 1 \bmod 4$, a closed formula for a solution of $x^2 \equiv -1 \bmod p$, namely $x = \left(\frac{p-1}{2}\right)!$. Indeed,

$$\left(\frac{p-1}{2}\right)!^2 \equiv \left(\prod_{j=1}^{\frac{p-1}{2}} j\right)\left(\prod_{j=1}^{\frac{p-1}{2}} (p-j)\right) \equiv (p-1)! \equiv -1 \bmod p.$$

We leave it to the reader to find out where we used here that $p-1$ is divisible by 4. Note that this computation gives a second proof of the fact that $p \equiv 1 \bmod 4$ implies the solubility of the congruence

$x^2 \equiv -1 \bmod p$. Algorithmically it is, however, for big p not wise to compute a solution of $x^2 \equiv -1 \bmod p$ using this formula. If p is big this affords $(p-1)/2$ multiplications. It is better to proceed as in the first proof, namely to compute $w^{(p-1)/2}$ modulo p for some primitive root modulo p. At the first glance this seems also to afford $(p-1)/2$ multiplications. But there is a little important trick to reduce the computation of a power a^n to about $\log_2 n$ steps. This is best understood by an example: for computing a^{100} one proceeds as follows.

$$b = (a^2)^2, \quad c = ((b^2)^2)^2, \quad d = c^2, \quad a^{100} = d \cdot c \cdot b,$$

which makes 8 multiplications instead of 100. This method is sometimes called "divide and conquer". we discuss it in more detail in the section of remarks following this chapter.

Theorem (Thue)**.** *Let p be a prime. Then, for every r not divisible by p there exist numbers $0 < a, b < \sqrt{p}$ such that $b \equiv \pm ra \bmod p$. More generally, given integers $m > 0$ and $0 < A, B \leq m$, $AB > m$, then, for any r which is relatively prime to m there exist integers $0 < a < A$, $0 < b < B$ such that $b \equiv \pm ra \bmod m$.*

Proof. Consider the application which associates to each pair of integers (k, l) with $0 \leq k < A$, $0 \leq l < B$ the residue class modulo m of $kr + l$. Since there are $AB > m$ such pairs but only m residue classes modulo m, we conclude that there are two pairs $(k, l) \neq (k', l')$ such that

$$kr + l \equiv k'r + l' \bmod m.$$

Setting $a = |k - k'|$ and $b = |l - l'|$ we find $b \equiv \pm ar \bmod m$. It is clear that $|l - l'| < B$ and $|k - k'| < A$. Furthermore either $a \neq 0$ or $b \neq 0$ since (k, l) and (k', l') are different. But then also b respectively a is different from 0. Namely, if $b = 0$ then m would divide ra, and then also a (since r is relatively prime to m), which is only possible for $a = 0$. Vice versa $a = 0$ would imply that m divides b, whence $b = 0$. This proves Thue's theorem for general m. The special case for $m = p$ follows on taking $A = B = \lceil \sqrt{p} \rceil$, so that $AB > p$. □

As consequence of the two preceding theorems one obtains:

Theorem. *An odd prime p is a sum of two perfect squares if and only if $p \equiv 1 \bmod 4$.*

Proof. Indeed, assume $p = x^2 + y^2$ for two positive integers x and y. Clearly, x and y are smaller than $\sqrt{p}$. In particular, they are not divisible by p. But then we deduce from $x^2 \equiv -y^2 \bmod p$ that $(xy')^2 \equiv -1 \bmod p$, where y' is an inverse mod p of y. From the last but not least theorem we deduce $p \equiv 1 \bmod 4$.

Vice versa, if the latter is satisfied, we can solve $r^2 \equiv -1 \bmod p$. Choosing a and b for r as in Thue's Theorem we have $b^2 \equiv -a^2 \bmod p$. In other words p divides the number $n := a^2 + b^2$. But since a and b are smaller than $\sqrt{p}$ we have $n < 2p$. It follows $n = p$. □

The preceding theorem is an existence theorem, but it neither tells us how to find a decomposition of a prime as sum of two squares nor how many such decomposition's there are. It is not too hard to show that there is at most one solution $0 < x \leq y$ of $p = x^2 + y^2$. For small $p \equiv 1 \bmod 4$, for finding this solution, we can try all positive integers $x < \sqrt{p}$ until we find one such that $p - x^2$ is a perfect square. For large p this would not work since it needs to many steps. For this case we have the subsequent theorem, whose proof, however, would require methods from algebraic number theory and must therefore be skipped.

Theorem (Cornacchia)**.** *Let be a prime, $p \equiv 1 \bmod 4$, and let x a solution of $x^2 \equiv -1 \bmod p$ with $p/2 < x < p$. Denote by $\{r_n\}$ the sequence of numbers such that $r_0 = p$, $r_1 = x$ und $r_n = r_{n-2} \% r_{n-1}$[7] $(n \geq 2)$. Let l be the smallest index such that $r_l < \sqrt{p}$. Then $p - r_l^2$ is a perfect square.*

Note that an x as in the theorem always exist and can also be easily computed. Namly, choose a primitive root w mod p and compute a positive x such that $x \equiv w^{\frac{p-1}{4}} \bmod p$. Then $x^2 \equiv -1 \bmod 4$. If $x < p/2$ we replace x by $p - x$ so that then $p/2 < x < p$. As already explained above computing powers by even large exponents is no problem. This leads then to the following algorithm.

[7] For two integers a and $b \neq 0$ we use $a\%b$ for the remainder of a after division by b

Algorithm: Find a representation of a given prime p as sum of two perfect squares.

```
import math

def find_squares( p):
    """
    Return the solution (x,y) of 0<x<=y of
        p=x^2+y^2 if it exists,
    otherwise throw an exception. Input
        must be prime p = +1 mod 4.

    EXAMPLE
        >>> find_squares(7829)
        (50, 73)
        >>> find_squares(100049)
        (215, 232)
        >>> find_squares(1000037)
        (134, 991)

    """
    assert p%4 == 1, 'Error: %d must be a
        prime = 1 mod 4' % p
    # Find a solution of x^2=-1 mod p
    # using Wilson's theorem
    x = 1
    for j in range(2,(p-1)/2):
        x *= j
        x = x%p

    # Modify x if necessary so that p > x >
        p/2
    if 2*x < p:
```

```
        x = p - x

    # compute r_l as in the preceding
        theorem
    a=p; b=x
    while b*b > p:
        r=a%b; a=b; b=r

    a = int( math.sqrt(p-b*b))
    return (a,b) if a < b else (b,a)
```

4. Remarks

We end this chapter by additional material for those readers with some basic knowledge of abstract algebra.

4.1. Axiomatic characterization of the integers. In Section 1 we mentioned that it is not difficult to characterize the integers axiomatically. In fact, this can be done as follows. Let R be an *ordered ring without zero-divisors*. In other words, R is a set equipped with two binary operations "+" and "$\cdot$" which fulfill the axioms of a unitary commutative ring without zero-divisors, and there exists a subset $R_{\geq 0}$ of R which is closed under addition and multiplication, and which, for any $a \neq 0$ in R, contains either a or $-a$. That R is unitary means that there exists a multiplicative neutral element. This is unique and henceforth denoted by 1_R. One defines $a \leq b$ if $b - a$ is in $R_{\geq 0}$, and this relation defines then a *total order* on R. The integers are an example of such an ordered ring. However, there are also other examples, like for example the rational or real numbers. We assume now in addition that every subset in $R_{\geq 0}$ possesses a smallest element. We can then prove:

Theorem. *The induction axiom holds in R, i.e. $R_{\geq 0}$ is the only subset of $R_{\geq 0}$ which contains 0 and with every element a also $a + 1_R$.*

Proof. Indeed, let A be such a subset. If A was different from $R_{\geq 0}$ then $B := R_{\geq 0} \setminus A$ possesses a smallest element a_0. Clearly $a_0 > 0$. Furthermore $a_0 < 1$ (since otherwise $a_0 - 1$ would be in B). But there are no elements $0 < b < 1$ in R since for any such element b we would have $b^2 < b$ (as follows from $(1-b)b \in R_{\geq 0}$ and $(1-b)b \neq 0$ since R does not have any divisors of zero), contradicting the fact that the set of all $0 < b < 1$ would posses a minimal element if non-empty. Therefore B must be empty and $A = \mathbb{R}_{\geq 0}$. □

We can now prove that R is nothing else than the ring of integers, up to a possible different naming of its elements. More precisely, we shall prove the following theorem.

Theorem. *There exists one and only one isomorphism of rings of $\mathbb{Z}$ with R which maps $\mathbb{Z}_{\geq 0}$ onto $R_{\geq 0}$.*

Proof. Any isomorphism maps 1 to 1_R, and then any integer $n = n \cdot 1$ to $n \cdot 1_R$ (where, for negative n we mean by $n \cdot 1_R$ the element additive inverse of 1_r added $|n|$-many times to itself). Let vice versa f denote the map from $\mathbb{Z}$ to R which takes n to $n \cdot 1_R$. It is clear from the definition that this f is a homomorphism of rings. Note that 1_R is in $R_{\geq 0}$ (if -1_R was in $R_{\geq 0}$ then $1_R = (-1_R)(-1_R)$ is in $R_{\geq 0}$, a contradiction. Therefore any n-fold sum of 1_R is in $R_{\geq 0}$. In particular, f takes $\mathbb{Z}_{\geq 0}$ into $R_{\geq 0}$. The map f is injective. If $n \cdot 1_R = 0$ and $n \geq 2$, then $-1_R = (n-1) \cdot 1_r$ is in $\mathbb{R}_{\geq 0}$, a contradiction.

Finally, f is surjective since its image contains 0 and with every element a also $a + 1$, hence it contains $R_{\geq 0}$, and then consequently all of R. □

4.2. p-adic numbers. As we saw in the section on algebraic congruences mod p^n, given a polynomial $f(x)$ in one variable with integer coefficients and a number $0 \leq t_0 < p$ such that $f(t_0) \equiv 0 \bmod p$ and $f'(t_0) \not\equiv 0 \bmod p$, there exist one and only one sequence of numbers $0 \leq t_j < p$ such that $y_{n+1} := \sum_{\nu=0}^{n} t_\nu p^\nu$ satisfies $f(y_{n+1}) \equiv 0 \bmod p^{n+1}$ for all $n \geq 0$. The y_{n+1} look like the partial sums of some infinite p-adic expansion of some object, and we wondered what this object might be.

The answer to this can indeed be given. Namely, for a rational number r set $|r|_p = p^{-s}$, where s is the unique integer such that p does not occur in the factorization of r/p^s. We also set $|0|_p = 0$. The function $|\cdot|_p$ shares the same properties as the usual absolute value $|\cdot|_\infty$ on the set of rational numbers. Namely, we have $|r|_p = 0$ if and only if $r = 1$, we have $|rs|_p = |r|_p \cdot |s|_p$, and finally $|r + s|_p \leq |r|_p + |s|_p$ (in fact we even have even the stronger *ultrametric property* $|r + s|_p \leq \max(|r|_p, |s|_p)$). Using these properties one sees that the Cauchy sequences $\mathcal{C}_p$ of rational numbers with respect to the *valuation* $|\cdot|_p$ form under term-wise addition and multiplication a ring, and that the subset $\mathcal{N}_p$ of rational sequences converging to zero with respect to $|\cdot|_p$ form a ideal in $\mathcal{C}_p$. The quotient ring

$$\mathbb{Q}_p := \mathcal{C}_p/\mathcal{N}_p$$

turns out to be a field, the *field of p-adic numbers*. The map which associates to a rational number r the constant sequence with value r defines an embedding of fields (so that one identifies $\mathbb{Q}$ with its image under this embedding). The valuation $|\cdot|_p$ can be uniquely extended to all of $\mathbb{Q}_p$ so that the three properties of a valuation (and the ultrametric property) are still satisfied. The field $\mathbb{Q}_p$ is then complete with respect to $|\cdot|_p$, i.e. every Cauchy sequence of $\mathbb{Q}_p$ converges. The field $\mathbb{Q}$ is dense in $\mathbb{Q}_p$. One sets

$$\mathbb{Z}_p := \big\{x \in \mathbb{Q}_p : |x|_p \leq 1\big\}.$$

Using the ultrametric property it is easy to verify that $\mathbb{Z}_p$ is a ring, the ring of integers of $\mathbb{Q}_p$. Note that $\mathbb{Z}_p$ contains the ring $\mathbb{Z}$.

Coming back to our sequence of the y_n the congruences $y_m \equiv y_n \bmod p^n$ translate into $|y_m - y_n|_p \leq p^{-n}$, and hence our sequence is a Cauchy sequence and converges, say, towards y. In other words,

$$y = \lim_n y_{n+1} = \lim_n \sum_{\nu=0}^{n} t_\nu p^\nu,$$

and as in real analysis it is common in p-adic analysis too to denote this limit by $\sum_{\nu=0}^{\infty} t_\nu p^\nu$. Moreover, $f(y_n) \equiv 0 \bmod p^n$ translates to $|f(y_n)|_p \leq p^{-n}$, i.e. $f(y_n)$ converges to 0. Finally, it is not hard to show that polynomials are continuous functions (with respect to $|\cdot|_p$),

so that

$$f(y) = f(\lim_n y_n) = \lim_n f(y_n) = 0.$$

We can therefore state:

Theorem. *Let f be a polynomial in one variable with integral coefficients, and p a prime number. Assume $f(x) \equiv 0 \bmod p$ and $f'(x) \not\equiv 0 \bmod p$ Then there exists exactly one solution y in $\mathbb{Z}_p$ of $f(y) = 0$ with $|y - x|_p < 1$.*