Lecture notes of Algebra. Week 1

1. Introduction: various types of numbers; notation

- The *natural* numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.
  - They originate from counting, but note that (for us) they start from 0.
  - Can be added (and also multiplied), but not subtracted in general: we can say that $1-3$ *should be* the same as $2-4$, $3-5$, etc., but such a thing does not exist within the natural numbers.
  - *Solution:* we *invent* a symbol for it, namely -2 (the *opposite*).
- The *integer* numbers $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$ (or the *integers*).
  - They can be added and subtracted arbitrarily.
  - They can be arbitrarily multiplied, but not divided in general: even if we exclude dividing by 0, which could never make sense, we can say that $2:3$ should be the same as $4:6$, $6:9$, or even $(-2):(-3)$, etc., but such a thing does not exist within the integers.
  - *Solution:* we *invent* a symbol for it, namely $2/3$ (so we have introduced the *reciprocal,* which in case of 3 is $1/3$, and then $2/3$ also means $2 \cdot (1/3)$).
  - Differently from subtraction in $\mathbb{Z}$, we still need pairs of integers to represent the result of an arbitrary division, as the reciprocals alone, such as $1/2$, $1/(-3)$, etc., would not be enough.
- The *rational* numbers $\mathbb{Q}$.
  - They consist of *fractions* of integers, hence of the form a/b, with $a, b \in \mathbb{Z}$ and $b \neq 0$, with the understanding that a/b and c/d represent the same rational number exactly when $ad = bc$.

    (Note that the condition $ad = bc$ makes sense already in $\mathbb{Z}$.)
  - They can be arbitrarily added and subtracted, multiplied and divided, with the only exception that we cannot divide by zero.

    (There is no way to remedy this, as $0 \cdot b = 0$ whatever b is.)
- The *real* numbers $\mathbb{R}$.
  - These are harder to construct from $\mathbb{Q}$, we will not go into this.
  - They include things like π, and like $\sqrt{2}$, but not $\sqrt{-2}$, or $\sqrt{-1}$.
- The *complex* numbers $\mathbb{C}$.
  - Each complex number has the form $a + bi$, for unique $a, b \in \mathbb{R}$.
  - Operations are done 'normally' with the only extra rule that $i^2 = -1$.
  - More info and practice on $\mathbb{C}$ in the Calculus module.

Identifying $\mathbb{Z} \ni 2 = 2/1 \in \mathbb{Q}$ etc. we have $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

2. Divisibility and the greatest common divisor in the integers

2.1. Divisibility in the integers. What does it mean that an integer b divides an integer a? One interpretation is that a/b is an integer as well. There are at least two problems with this possible definition. Firstly, it does not tell you whether 0 divides 0 or not, because $0/0$ does not make sense. Secondly, for generalisations it is good to have a definition which only mentions integer numbers, while a/b might generally be a rational number.

Definition 1 (Divisibility in the integers). Let $a, b \in \mathbb{Z}$. We say that b divides a, and we write $b \mid a$, if there is $c \in \mathbb{Z}$ such that $a = b \cdot c$.

We can express divisibility in various equivalent ways. The expressions

- b divides a,
- b is a divisor of a,
- a is a multiple of b,
- a is divisible by b,

all have the same meaning, which we write symbolically as $b \mid a$.

Do not confuse the symbol $|$, a vertical bar, with the fraction sign $/$, as $b \mid a$ is not related with the fraction b/a, also written $\frac{a}{b}$, but rather with a/b.

In fact, the statement that $b \mid a$ is equivalent with the quotient a/b being an integer, but only for $b \neq 0$. We cannot divide 0 by 0 and so cannot write $0/0$, but $0 \mid 0$ is a true statement, because $0 = 0 \cdot 1$ (or $0 = 0 \cdot 3$, or $0 = 0 \cdot (-22)$, or even $0 = 0 \cdot 0$; uniqueness of c is not required in the above definition).

Hence the integer b divides the integer a when the equation

$$a = bc$$

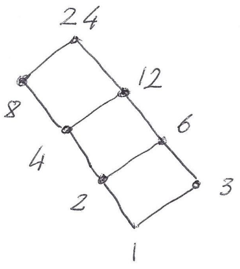
has a solution c in the integers. For example, 2 divides 6 because $6 = 2 \cdot 3$, and so the above equation, for $a = 6$ and $b = 2$, admits the solution $c = 3$. Note that $6 = 2 \cdot 3 = 3 \cdot 2$, hence the same equation tells us that 3 divides 6. In fact, the divisors of a nonzero integer a come in pairs: we can match any divisor b of a to the divisor a/b (which may possibly equal b, in case $a = b^2$).

For $a \in \mathbb{Z}$ we set

$$D(a) = \{x \in \mathbb{Z} : x \mid a\},$$

the set of divisors of a. Note that if $b \in D(a)$ then $-b \in D(a)$ as well, and also $D(-a) = D(a)$ for every a.

Example. $D(24) = \{\pm1, \pm2, \pm3, \pm4, \pm6, \pm8\pm12, \pm24\}$. The divisors are found more systematically starting from the factorisation of 24 into prime factors, $24 = 2\cdot2\cdot2\cdot3 = 2^3\cdot3$, and are best arranged in a *Hasse diagram*,



where the ascending lines indicate that the number below divides the one above. For simplicity we have omitted the ± signs, hence omitting the negative divisors, because after all $b \mid -a$ if and only if $b \mid a$, etc. Of course this representation works best if the number has only two prime divisors.

Example. Every integer b divides 0, because $0 = b \cdot 0$. Hence $D(0) = \mathbb{Z}$.

Note that if we had to place 0 somewhere in a (partial) Hasse diagram of $\mathbb{Z}$, we would have to place it 'higher' than any other integer: with respect to divisibility 0 is sort of 'the largest integer,' despite being the smallest in absolute value.

Example. However, if 0 divides a, then $a = 0 \cdot c = 0$ for some c, and so $a = 0$. This means that the only integer $a \in \mathbb{Z}$ such that $0 \in D(a)$ is $a = 0$ itself.

For each a we have $a \mid a$ (*reflexivity*), because $a = a \cdot 1$. Divisibility also enjoys *transitivity*: if $a \mid b$ and $b \mid c$, then $a \mid c$.

Symmetry does not generally hold, as $b \mid a$ does not imply $a \mid b$. In fact, $b \mid a$ and $a \mid b$ can only hold simultaneously when $a = \pm b$. Because this fact is important for the sequel, here is a proof.

Proof. Since the assertions $\pm b \mid \pm a$, with all possible choices of signs, are equivalent to each other, we may assume that $a, b \geq 0$, and so we only need to prove that under this condition $b \mid a$ and $a \mid b$ together imply $a = b$. First, if $a = 0$ then $0 \mid b$ implies $b = 0$, and conversely. Hence we may actually assume $a, b > 0$. Now $b \mid a$ means $b = ac$ for some integer c, which must be positive as well, hence $c \geq 1$, and so $b = ac \geq a \cdot 1 = a$. Similarly we find $a \leq b$, and we conclude $a = b$ as desired. □

2.2. Division with remainder in the integers.

Theorem 2 (Division with remainder in the integers). *Given two integers a, b, with $b > 0$, there exist unique $q, r \in \mathbb{Z}$ such that $a = b \cdot q + r$, with $0 \leq r < b$.*

We will not give a formal proof of Theorem 2, but the division which Theorem 2 describes is the ordinary division with remainder which you know from school, except that here we allow a to be negative.

Example. Dividing $a = -13$ by $b = 5$ gives quotient $q = -3$ and remainder 2, because $-13 = 5 \cdot (-3) + 2$, and $0 \leq 2 < 5$.

Theorem 2 actually remains true also for $b < 0$, provided that we replace the second condition with $0 \leq r < |b|$. Of course there is no way to make the theorem work for $b = 0$.

The following variant of integer division allows negative remainders, and aims at making the remainders as small as possible *in absolute value*. This may be convenient in certain situations in order to have simpler calculations.

Theorem 3 (Variant of division with remainder in the integers). *Given two integers a, b, with $b > 0$, there exist unique $q, r \in \mathbb{Z}$ such that $a = b \cdot q + r$, with $-b < 2r \leq b$.*

The condition on the remainder is equivalent to $-b/2 < r \leq b/2$, but we have preferred the other formulation because it takes place in the integers rather than in the rationals.

Example. When $a = 28$ and $b = 5$, using this variant of division we will get $28 = 5 \cdot 6 - 2$ instead of $28 = 5 \cdot 5 + 3$, so the remainder will be $r = -2$ (which is smaller than 3 in absolute value).

When $a = 21$ and $b = 6$, we will have $21 = 6 \cdot 3 + 3$ or $21 = 6 \cdot 4 - 3$, but the condition on the remainder chooses the positive remainder $r = 3$ in this case.

The theorem on division with remainder (in either version, Theorem 2 or Theorem 3) yields the following useful characterization of divisibility.

Corollary 4. *Let $a, b \in \mathbb{Z}$ with $b \neq 0$. The following assertions are equivalent:*

(1) *b divides a;*
(2) *the remainder of the division of a by b is zero.*

Proof. If b divides a then $a = b \cdot c = b \cdot c + 0$ for some c. Because of uniqueness, the remainder must be zero.

Conversely, if $r = 0$, then $a = b \cdot q + r = b \cdot q$, and hence b divides a. □

2.3. The greatest common divisor. Recall the definition of greatest common divisor which is usually learnt in school. Let a, b be positive integers. An integer d is called the greatest common divisor of a and b if

(1) d divides a and b, and
(2) if c is any integer which divides both a and b, then $c \leq d$.

This definition is not good for us because it does not generalise properly to other contexts (polynomials and further). Also, with this definition there would be no greatest common divisor in the special case $a = b = 0$, because the divisors of 0 are *all the integers*, and they do not have a maximum.

A better definition replaces 'max in the natural roder given by $\leq$' with 'max with respect to divisibility'.

Definition 5 (Greatest common divisor). Let $a, b \in \mathbb{Z}$. An integer d is called *a greatest common divisor* of a and b if

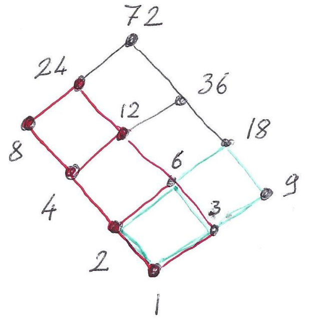
(1) d divides a and b, and
(2) if c is any integer which divides both a and b, then $c \mid d$.

A greatest common divisor of a and b (or *a GCD* in short) is denoted with $\gcd(a, b)$, or more simply with (a, b) (which we will preferably use).

Example. Let $a = 24$ and $b = 18$, and compute their GCD as learnt in school. First we need to factorise a and b as products of prime numbers, and in this case we have

$$24 = 2 \cdot 2 \cdot 2 \cdot 3 = 2^3 \cdot 3 \qquad\qquad 18 = 2 \cdot 3 \cdot 3 = 2 \cdot 3^2$$

The school's rule (which we will recall later together with primes and unique factorisation) tells us that $(24, 18) = 2 \cdot 3 = 6$, and also that the least common multiple (see definition later) of 24 and 18 is $2^3 \cdot 3^2 = 72$. In such a simple case we can actually arrange all divisors of 72 in the following Hasse diagram,



where in red and green we see the Hasse diagrams of the divisors of 24 and 18, respectively. Note that the common divisors of 24 and 18 are precisely the divisors of 6, so $D(24) \cap D(18) = \{\pm 1, \pm 2, \pm 3, \pm 6\}$. Hence 6 is the greatest common divisor of 24 and 18 in the school's sense of being the larger numerically, but also in the stronger sense of our definition.

The school's rule used in the above example to find the greatest common divisor of two (or more) integers is something like *the GCD equals the product of all common prime factors, taken with the lower (or lowest) exponent.* This is correct but is not practical, because factorisation into products of primes is a hard computational problem when the numbers are large, as we will see (while computing a greatest common divisor is much faster, using Euclid's algorithm, see below). Besides, it is theoretically more convenient to deduce the unique factorisation of integers from the existence of the greatest common divisor and its properties (which we will do in a later section), rather than the other way around.

Note that if d is a greatest common divisor of a and b, then $-d$ is also a greatest common divisor. But there cannot be other greatest common divisors of a and b. In fact, if d and d' are both greatest common divisors of a and b, then Definition 5 implies $d \mid d'$ and $d' \mid d$, and we know that $d' = \pm d$ follows. Hence the GCD is not unique, but almost (and we will accept and use the traditional expression 'the GCD' in place of the more correct 'a GCD').

3. The Euclidean algorithm in the integers

To prove that the GCD of any two integers actually exist we will describe an actual algorithm to compute it, called the *Euclidean algorithm.*

3.1. The Euclidean algorithm. Before describing the Euclidean algorithm we start with a numerical example.

Example. We compute the GCD of $a = 78$ and $b = 33$, using the Euclidean algorithm which we will explain below. That consists of a sequence of divisions, in this case

$$\begin{aligned} 78 &= 33 \cdot 2 + 12 \\ 33 &= 12 \cdot 2 + 9 \\ 12 &= 9 \cdot 1 + 3 \\ 9 &= 3 \cdot 3 + 0 \end{aligned}$$

So we have started with dividing one of the given numbers by the other, with remainder: $a = bq + r$. Then we have discarded the first number a, let b and the remainder r take the previous roles of a and b, and we have divided again. This continues until one of the divisions had remainder zero. The remainder of the previous division (hence the last nonzero remainder, if we like) is the greatest common divisor of a and b, so in this case $(78, 33) = 3$. Correct, because $78 = 2 \cdot 3 \cdot 13$ and $33 = 3 \cdot 11$.

It is convenient to think of the list of remainders as including the two original numbers, hence $78, 33, 12, 9, 3, 0$ in this case. Each of them (starting from the third) is determined as the remainder of dividing the previous two in that order. Hence division with remainder gives us a recursive way to construct this sequence (where the first two are given to get started), and the sequence ends as soon as it reaches zero (which it eventually will because it is decreasing).

Now we present a general description of the Euclidean algorithm, starting with a lemma which explains the crucial reason why the algorithm works. We introduce the notation $D(a, b) = D(a) \cap D(b)$ for the set of common divisors of two integers a, b. Then it is easy to see that a greatest common divisor of a and b is any number d such that $D(a,b) = D(d)$. For example, the GCD of $a = 0$ and $b = 0$ is 0, because $D(0,0) = \mathbb{Z} = D(0)$. More generally, the GCD of any a and 0 is a, because $D(a,0) = D(a) \cap D(0) = D(a) \cap \mathbb{Z} = D(a)$.

In the general case, the existence of the GCD can be proved as follows. Because $D(-a) = D(a)$, we can restrict ourselves to the case $a, b \geq 0$. Because $D(a,b) = D(b,a)$, we may also assume $a \geq b$. The crucial step is the following lemma.

Lemma 6. *If $a = bq + c$, then $D(a,b) = D(b,c)$.*

Proof. We have to show that every element of $D(a,b)$ is also an element of $D(b,c)$, and the converse, namely, that every element of $D(b,c)$ is also an element of $D(a,b)$.

If $d \in D(a,b)$, that is, if d is any common divisor of a and b, then d also divides bq, and hence it also divides the difference $a - bq = c$. Because d divided b in the first place, we conclude that it divides both b and c, which means $d \in D(b,c)$.

The converse is similar. If if d is a common divisor of b and c, then d also divides bq, and hence it also divides the sum $bq + c = a$. Therefore, d divides both a and b, which means $d \in D(a,b)$. □

We use Lemma 6 as follows. Because we have seen the case $b = 0$, suppose $b > 0$. Dividing a by b we find

$$a = bq_1 + r_1, \qquad 0 \leq r_1 < b.$$

According to Lemma 6 we have $D(a,b) = D(b,r_1)$, and so for the sake of computing our GCD we may replace a with b and b with r_1, with the advantage that r_1 is smaller than b. Assuming $r_1 > 0$ we divide b by r_1, with remainder $r_2 < r_1$, and continue in the same way by dividing and replacing, until one of the divisions has remainder zero. That is bound to happen because the successive remainders $r_1, r_2, r_3, \ldots$ are strictly decreasing nonnegative integers. In conclusion, we have performed a sequence of divisions

$$\begin{aligned}
a &= bq_1 + r_1, & 0 &< r_1 < b,\\
b &= r_1q_2 + r_2, & 0 &< r_2 < r_1,\\
r_1 &= r_2q_3 + r_3, & 0 &< r_3 < r_2,\\
&\vdots & &\vdots\\
r_{i-2} &= r_{i-1}q_i + r_i, & 0 &< r_i < r_{i-1},\\
r_{i-1} &= r_iq_{i+1}, & &
\end{aligned}$$

where the remainder r_{i+1} equals zero. Then the last nonzero remainder, r_i is the desired GCD of a and b, because

$$D(a,b) = D(b,r_1) = D(r_1,r_2) = \cdots = D(r_{i-1},r_i) = D(r_i,0) = D(r_i).$$

For example, let us compute the GCD of 18 and 14. We have $18 = 14 \cdot 1 + 4$, and so $D(18, 14) = D(14, 4)$. Again, $14 = 4 \cdot 3 + 2$, and so $D(14, 4) = D(4, 2)$. And again, $4 = 2 \cdot 2 + 0$, and so $D(4, 2) = D(2, 0) = D(2)$. Hence 2 is the desired GCD. Of course this was also easy to do with the school method but, for example, finding the GCD of 1987 and 2203 by the school method would not be so easy. This method is called *the Euclidean algorithm (of successive divisions).*

3.2. Examples and further comments on the Euclidean algorithm.

Example. Compute the GCD of 59 and 22:

$$\begin{aligned} 59 &= 22 \cdot 2 + 15 \\ 22 &= 15 \cdot 1 + 7 \\ 15 &= 7 \cdot 2 + 1 \end{aligned}$$

Hence $(59, 22) = 1$. Note that when the Euclidean algorithm reaches a remainder 1 there is no need to do or write down the last division $7 = 1 \cdot 7 + 0$, as it is obvious that its remainder will be zero. When two integers have greatest common divisor 1 as in this case we say that they are *relatively prime,* or that they are *coprime.* Do not confuse being coprime with being prime: 59 is actually a prime but 22 is not. Two integers may be coprime without either being prime, for example 4 and 15.

Example. Compute the GCD of 34 and 21:

$$\begin{aligned} 34 &= 21 \cdot 1 + 13 \\ 21 &= 13 \cdot 1 + 8 \\ 13 &= 8 \cdot 1 + 5 \\ 8 &= 5 \cdot 1 + 3 \\ 5 &= 3 \cdot 1 + 2 \\ 3 &= 2 \cdot 1 + 1 \end{aligned}$$

Hence $(34, 21) = 1$. In this case the Euclidean algorithm has been as slow as it can possibly be, because all quotients happened to be 1. This will occur exactly when the starting numbers a and b are consecutive Fibonacci numbers. The *Fibonacci numbers* $F_0, F_1, F_2, \ldots$ are defined by the recurrence relation

$$F_n = F_{n-1} + F_{n-2} \qquad (n \geq 2; \quad F_0 = 1, F_1 = 1),$$

and so their sequence begins with $0, 1, 1, 2, 3, 5, 13, 21, 34, 55, 89, 144, \ldots$

The following example illustrates how computing a greatest common divisor by means of the Euclidean algorithm can be much faster than by the school method of factorising the two numbers.

Example. We use the Euclidean algorithm to compute the greatest common divisor of 391 and 299. The Euclidean algorithm reads

$$\begin{aligned} 391 &= 299 \cdot 1 + 92 \\ 299 &= 92 \cdot 3 + 23 \\ 92 &= 23 \cdot 4. \end{aligned}$$

Hence $(391, 299) = 23$.

In particular, we discover that $391 = 17 \cdot 23$ and $299 = 13 \cdot 23$ (the complete factorisations of these two numbers into prime factors), without previously factorising either number. Note that factorising 391 directly, for example, would have taken a while, because the standard procedure would be:

- checking if 391 is divisible by 2: it is not (easy, check if last digit is even or odd);
- checking if it is divisible by 3: it is not (a shortcut is checking if the sum of its digits is divisible by 3);
- checking if it is divisible by 5: it is not (because its last digit is not 0 or 5);
- checking if it is divisible by 7: it is not (not so easy, just perform the division);
- checking if it is divisible by 11: it is not (a shortcut is taking the alternating sum of the digits (that is, with alternating signs, here $3 - 9 + 1$) is divisible by 11);
- checking if it is divisible by 13: it is not (not so easy, just perform the division);
- checking if it is divisible by 17, and finally finding that it is.

Example. Compute the GCD of 2203 and 1987:

$$\begin{aligned} 2203 &= 1987 \cdot 1 + 216 \\ 1987 &= 216 \cdot 9 + 43 \\ 216 &= 43 \cdot 5 + 1 \end{aligned}$$

Hence $(2203, 1987) = 1$, so these two numbers are coprime (or relatively prime). In this case both 2203 and 1987 are actually prime numbers, and so finding their GCD by factorising them would have taken even longer than in the previous example.

In particular, because $\sqrt{1987}$ is about 44.5, we would have spent a long time trying and dividing it by the primes $2, 3, 5, 7, 11, \ldots, 43$ before discovering that it actually is a prime. At that point it would be enough to check that 1987 does not divide 2203 in order to conclude that $(2203, 1987) = 1$. However, this procedure would have taken a long time (most of it to factorise 1987).

By contrast, the Euclidean algorithm was very fast to give us $(2203, 1987) = 1$, but does not tell us that they are prime. More generally, when the Euclidean algorithm on a and b concludes with $(a, b) = 1$, it gives us no clue about the factorisations of a and b.

The successive divisions in the Euclidean algorithm can also be done in the variant where the remainders are taken as small as possible in absolute value. When possible this may make the Euclidean algorithm conclude a little faster.

Example. We compute the GCD of 29 and 18, on the left in the standard way, and on the right taking negative remainders when convenient:

$$\begin{aligned}
29 &= 18 \cdot 1 + 11 &\qquad 29 &= 18 \cdot 2 - 7\\
18 &= 11 \cdot 1 + 7 &\qquad 18 &= 7 \cdot 3 - 3\\
11 &= 7 \cdot 1 + 4 &\qquad 7 &= 3 \cdot 2 + 1\\
7 &= 4 \cdot 1 + 3 & &\\
4 &= 3 \cdot 1 + 1 & &
\end{aligned}$$

We see that taking negative remainders whenever those are smaller than the positive ones in absolute value made the algorithm run faster (in fewer steps), and this is generally a good strategy. However, even if when we choose to take negative remainders only sometimes, that may still make our algorithm run a bit faster than with just standard divisions. Here are a couple of examples:

$$\begin{aligned}
29 &= 18 \cdot 2 - 7 &\qquad 29 &= 18 \cdot 1 + 11\\
18 &= 7 \cdot 2 + 4 &\qquad 18 &= 11 \cdot 2 - 4\\
7 &= 4 \cdot 2 - 1 &\qquad 11 &= 4 \cdot 3 - 1
\end{aligned}$$

Example. Compute the GCD of 34 and 21 as in a previous example (the one on Fibonacci numbers), but allowing negative remainders:

$$\begin{aligned}
34 &= 21 \cdot 2 - 8\\
21 &= 8 \cdot 3 - 3\\
8 &= 3 \cdot 3 - 1
\end{aligned}$$

Hence $(34, 21) = 1$. This has taken half as many divisions as doing it in the normal way as before. In fact, note that the successive remainders, in absolute value, have been $34, 21, 8, 3, 1$, so every other Fibonacci number (apart from the start) rather than going through all of them as we did before, 34,21,13,8,5,3,2,1.