Introduction to Abstract Algebra
"Rings First"

Bruno Benedetti
University of Miami

January 2022

Abstract

The main purpose of these notes is to understand what $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are, as well as their polynomial rings.

Contents

0 Preliminaries

In this section we briefly recall a few topics you probably know already. But even if you "know too much", it's good to agree on notation and terminology, so that we are all on the same page.

0.1 Natural numbers, induction, and Euclid's theorem

You are probably all familiar with the infinite set of natural numbers (also known as "nonnegative integers")

$$\mathbb{N} = \{0, 1, 2, 3, \ldots, n, n+1, \ldots\}$$

It is usually stated in textbooks that natural numbers "come from Nature". This is not entirely true: To accept them, three important abstraction steps are necessary. These steps are non-trivial, as throughout the history of mankind, not all populations have achieved or accepted them:

- the notion of *cardinality*, i.e. the realization that two finite sets in bijection with one another have something in common; whence the *names* of numbers. This is not universal: Even today, the Pirahã people in Amazonas, Brazil, have no names for numbers, and have no linguistic way of expressing exact quantity, not even "one". [1]
- the notion of *zero*, as the cardinality of an "empty set". The ancient Greeks had no symbol for zero, for example; Mayas did have a symbol for zero around the year 36 BC, using it as placeholder in their base-20 numerical system; arithmetic operations with zero were first introduced by the Indian mathematician Brahmagupta[2], around 650 AD.
- the existence of an *infinite set*, that is, a set that can be in bijection with a proper subset of itself. The bijection in this case is the *successor* map, that is, the map that adds one to each element; so an equivalent way of formulating this principle is, "the belief that every number has a successor". Once again, this intuition is not universal, and in logic there is a movement of logicians from around 1900, called *(strict) finitists*, who rejected it.

Given two natural numbers a and b, we say that "*a divides b*" (or equivalently that "*a is a divisor of b*", or equivalently that "*b is a multiple of a*") if there exists a natural number k such that

$$b = k \cdot a.$$

Prime numbers are the natural numbers with exactly two distinct divisors (so 1 is not prime!):

$$2, 3, 5, 7, 11, 13, \ldots$$

Given two natural numbers a and b, their *greatest common divisor*, denoted by $\gcd(a, b)$ is the largest integer that divides both a and b.

Example 1. If p is prime, then it has only 1 and p as divisors. Thus for any natural number n

$$\gcd(p, n) = \begin{cases} p & \text{if } p \text{ divides } n, \\ 1 & \text{otherwise.} \end{cases}$$

Induction is a standard technique to prove a statement for infinite subsets of $\mathbb{N}$. It is based on the fact that every natural number is obtained from 0 by adding 1 sufficiently many times. So if we show that a property P holds for zero and is maintained when we move from any number

[1] Frank et al., *Number as a cognitive technology: Evidence from Pirahã language and cognition*, Cognition 108 (2008), 819–824.

[2] Wallin, Nils-Bertil,"The History of Zero". YaleGlobal online, 19 November 2002

to its successor, then P holds also for one, for two, for three.... and eventually is shared by all natural numbers.

Formally, a proof by induction consists of two parts:

- ("Basis") We prove that the statement holds true for a specific integer n_0.
- ("Step") We prove that, if there exists a natural number n such that the statement holds for n, then the statement must hold also for $n+1$.

Once again, the validity of the statement for n_0 implies the validity for n_0+1, which in turn implies the validity for n_0+2, and so on: A domino effect, which eventually proves the statement for all integers $n \geq n_0$. If our basis was $n_0 = 0$, then we end up proving the statement for the whole of $\mathbb{N}$.

Example 2. Let us prove by induction that

$$\sum_{i=0}^{n+1} i = \binom{n+2}{2} \text{ for all } n \in \mathbb{N}.$$

- ("Basis") For $n = 0$, the formula above boils down to $0+1 = \binom{2}{2}$, which is correct. This brings good luck!
- ("Step") Let us assume that $\sum_{i=0}^{n+1} i = \binom{n+2}{2}$ holds true for some n. Then

$$\sum_{i=0}^{n+2} i = (n+2) + \sum_{i=0}^{n+1} i \overset{!}{=} (n+2) + \binom{n+2}{2} = \binom{n+2}{1} + \binom{n+2}{2} = \binom{n+3}{2}.$$

Non-Example 3. Here is a "fake proof" by induction that

$$\sum_{i=0}^{n+1} i = \frac{(2n+3)^2}{8} \text{ for all positive integers } n.$$

Let us assume that $\sum_{i=0}^{n+1} i = \frac{(2n+3)^2}{8}$ holds true for some n. Then

$$\sum_{i=0}^{n+2} i = (n+2) + \sum_{i=0}^{n+1} i \overset{!}{=} (n+2) + \frac{(2n+3)^2}{8} = \frac{4n^2+20n+25}{8} = \frac{(2n+5)^2}{8} = \frac{(2(n+2)+1)^2}{8}.$$

What did we do wrong? Induction consists of *two* parts, a step and a basis... The basis is not superfluous! We need a domino tile where our domino effect can start.

Remark 4. A common mistake is to memorize the induction step as follows, "Let us assume that the statement holds for all n; then let us prove it for $n+1$". This makes no sense: If we already know that the statement holds for all n, then we are done!, no need to think about $n+1$. That's not how induction works. What instead we are assuming is much less, namely, that the statement holds for *one* specific n; from there we want to be able to say the same thing about its successor, $n+1$.

We should pay special attention to whether our induction step is imposing extra conditions on n. If the induction step works only for $n \geq n_0$, then the basis for the induction should be its verification at n_0, and not at 0.

Example 5. Let us prove that $n^2 - 5n + 6 \geq 0$ for all integers n. Let's assume it for n; then

$$(n+1)^2 - 5(n+1) + 6 = (n^2+2n+1) - 5n - 5 + 6 = (n^2 - 5n + 6) + 2n - 4 \geq 2n - 4.$$

Now to conclude we would like to say that $2n - 4 \geq 0$. But this is true only when $n \geq 2$. So we are not done yet; it is incorrect for us to "make the domino tiling start" at $n = 0$ because as far as we know, the validity at 0 might not imply the validity at 1. So we proceed as follows:

- First, we ask ourselves whether the claim holds true for $n = 2$, which is the correct induction basis. Since $2^2 - 10 + 6 = 0$, the answer is "yes". Together with the induction step, this does prove

$$\text{“}n^2 - 5n + 6 \geq 0 \text{ for all integers } n \geq 2.\text{”}$$

- This is not what we were asked to do, but almost: we are left with only finitely many cases to consider!, namely, $n = 0$ and $n = 1$. We can check them by hand: for $n = 0$ we have $0 - 0 + 6 > 0$, for $n = 1$ we have $1^2 - 5 + 6 = 0$. This concludes the proof.

Non-Example 6. Here is a "fake proof" (by induction on the number n of students) that

"all students will receive the same grade in the final".

For $n = 1$, we are considering a class consisting of only one student, so the claim is clear. Now suppose we have proved the claim for classes with n students. Let C be any class with $n + 1$ students. Let x, y be any two students enrolled in the class, and let z be any other student. Consider $S \setminus \{x\}$: this is a set of n students, so we can apply the inductive assumption: Everybody in $S \setminus \{x\}$ is going to get the same grade. In particular, y and z will get the same grade. Analogously, by the inductive assumption everybody in $S \setminus \{y\}$ will get the same grade, so in particular x and z. Summing up, x, y, z will all get the same grade. But then *any* two students x, y will get the same grade. What is wrong here?

(Hint: If in a proof you pick three different elements from a set, then you are implicitly assuming that the set has at least three elements. Note that if we have a class of 3 students, and any pair of them gets the same grade, then indeed they all get the same grade...)

There is a variant of induction which sometimes is easier to use than the one above. It is called **strong induction**, or sometimes "complete' or "generalized induction". Essentially, it is just normal induction plus "bookkeeping", i.e. keeping track of everything you have proven before. It consists of two parts:

- ("Basis") Prove a statement for a specific integer n_0.
- ("Step") Prove that, if there is a natural number n such that the statement holds for *every* natural number k such that $n_0 \leq k \leq n$, then the statement must hold also for $n + 1$.

Let us use the shortening $P(n)$for "the property holds for n". It is easy to see how generalized induction works: First of all, $P(n_0)$ implies $P(n_0 + 1)$. Now that we know $P(n_0)$ and $P(n_0 + 1)$, we can infer $P(n_0 + 2)$. But then we know $P(n_0)$ $P(n_0 + 1)$, and $P(n_0 + 2)$, which together imply $P(n_0 + 3)$. And so on: Another domino effect, which results in a proof of the statement for all integers $n \geq n_0$. The difference with classical induction is that instead of proving

$$P(n) \Rightarrow P(n + 1),$$

where $P(n)$ stands for "the property holds for n", we keep track at each step of what we have proven already and show

$$[P(n) \text{ and } P(n - 1) \text{ and } P(n - 2) \ldots \text{ and } P_{n_0}] \Rightarrow P(n + 1).$$

Proposition 7. *Any integer $n \geq 2$ can be written as product of primes.*

Proof. The statement is true for $n = 2$, because "$2 = 2$" writes 2 as products of primes. Now let n be an integer, and suppose the claim has already been proven for any integer in $\{2, 3, \ldots, n-1\}$. If n is prime, then

$$n = n$$

is a valid way to write down n as a product of primes, and we are done. If n is composite, then

$$n = n_1 \cdot n_2,$$

with $2 \leq n_i < n$ (for $i = 1, 2$). By strong induction, both n_i can be written as product of primes. But then so can n. □

Corollary 8 (Euclid)**.** *There are infinitely many primes. More precisely, the number of primes not larger than n is at least* $\log_2 \log_2 n$.

Proof. Let $p_1, \ldots, p_r$ be the complete list of the first r primes. Consider the number $n \stackrel{\text{def}}{=} 1 + p_1 p_2 \cdots p_r$ and let p be any prime factor of n; the existence of one such p is guaranteed by Proposition 7. Were p belonging to $\{p_1, \ldots, p_r\}$, then

$$1 \;=\; p_1 \cdots p_r - n$$

would be a difference of two multiples of p; and thus 1 itself would be a multiple of p, a contradiction. So p does not belong to $\{p_1, \ldots, p_r\}$, and in particular, p is larger than all of $p_1, \ldots, p_r$. This already suffices to prove that primes are infinitely many (because for any r, given the first r primes, we showed how to produce yet another prime). But we can use it to say something more precise: Since p is prime and $p_r < p \leq n$, either p is the first prime after p_r (i.e. $p_{r+1} = p$), or there are further primes in between (i.e. $p_{r+1} < p$). Either way, we obtain

$$p_{r+1} \leq p \leq n = p_1 p_2 \cdots p_r + 1. \tag{1}$$

We claim that this implies the bound

$$p_{r+1} \leq 2^{2^r} \tag{2}$$

We verify the claim above by strong induction: the case $r = 0$ boils down to $2 \leq 2^1$, which is true. As for $r > 0$, applying 1 and strong induction we get

$$p_{r+1} \leq p_1 p_2 \cdots p_r + 1 \leq 2^{2^0} 2^{2^1} \cdots 2^{2^{r-1}} + 1 \leq 2^{2^r}.$$

So the claim is proven. But from the claim, the conclusion follows immediately: Let n be a positive integer. Let $r = \lfloor \log_2 \log_2 n \rfloor$. Then of course $r \leq \log_2 \log_2 n \leq r + 1$, or in other words,

$$2^{2^r} \leq n \leq 2^{2^{r+1}}.$$

But by inequality 2, the first $r + 1$ primes $p_1, \ldots, p_{r+1}$ are all $\leq 2^{2^r}$ by inequality 2. So in particular they are all all not larger than n. Thus the number of primes not larger than n is at least

$$r + 1 \geq \log_2 \log_2 n.$$

□

Remark 9. In some textbooks, the theorem above is often misquoted as follows: *given the set of the first k primes, their product plus one is a much larger prime.* This argument is incorrect:

$$1 + (2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13) = 30031$$

is not prime, for example, because it is divisible by 59.

The bound of Corollary 8 can be strengthened with the following beautiful argument, due to Erdös:

Theorem 10. *The number of primes not larger than n is at least* $\frac{1}{2} \log_2 n$.

Proof. Let $p_1, \ldots, p_r$ be the complete list of all primes that are not larger than n. Any natural number $k \leq n$ can be written as

$$k = p_1^{\epsilon_1} \cdots p_r^{\epsilon_r} \cdot m^2, \tag{3}$$

where the exponents ϵ_j are all either 0 or 1, and the integer m^2 is a perfect square. Clearly $m \leq \sqrt{n}$. Now, in the expression 3, not all the choices of ϵ_j in $\{0, 1\}$ and of integers $1 \leq m \leq \sqrt{n}$ will result in a product that is $\leq n$. Also, at the moment we do not know whether every $k \leq n$ can be determined *uniquely* by the choice of $\epsilon_1, \ldots, \epsilon_r$ and n. But certainly, if we count all such choices, we are overcounting all numbers from 1 to n. So

$$2^r \sqrt{n} \geq n.$$

Dividing by $\sqrt{n}$, we obtain

$$2^r \geq n^{\frac{1}{2}},$$

whence taking the logarithm in base 2 we get

$$r \geq \log_2(n^{\frac{1}{2}}) = \frac{1}{2} \log_2 n. \qquad \square$$

Remark 11. The bound can be improved further: in 1896, Hadamard and de-la-Vallee-Poussin independently proved that the number of primes $\leq n$ grows like $\frac{n}{\log_2 n}$; an elementary proof was fund in 1948 by Selberg and Erdös, again independently. This statement goes under the name of **PNT** (Prime Number Theorem).

We have arrived to the most important result of this Section:

Theorem 12 (Euclidean division)**.** *Let n, d be natural numbers, $d \neq 0$. There exist natural numbers (q, r) such that*

(1) $n = qd + r$,

(2) $0 \leq r < d$.

In addition, the pair (q, r) is uniquely determined by (n, d).

Notation alert. The number q is called "quotient of the division of n by d"; the number r is called "remainder of the division of n by d". Sometimes d is called "divisor". This explains why one chooses the letters q, r, d.

Proof. Let us prove existence first. The claim is clear for $n < d$ (by choosing $q \stackrel{\text{def}}{=} 0$ and $r \stackrel{\text{def}}{=} n$). The claim is also clear for $d = 1$ (by choosing $q \stackrel{\text{def}}{=} n$ and $r \stackrel{\text{def}}{=} 0$). So the interesting case is

$$n \geq d \geq 2.$$

We proceed by strong induction on the first component of the pair (n, d). Consider $n_1 = n - d$. Clearly $n_1 \in \mathbb{N}$ (we are in the case $n \geq d$) and $n_1 < n$. By strong induction, the existence part of the theorem holds for the pair (n_1, d): So we can find natural numbers (q_1, r_1) such that we can write

$$n_1 = q_1 d + r_1 \text{ and } 0 \leq r_1 < d.$$

But then

$$n = n_1 + d = (q_1 + 1)d + r_1, \text{ with } 0 \leq r_1 < d$$

is the desired "division of n by d".

As for uniqueness, say $n = qd + r = q'd + r'$, with $0 \leq r < d$ and $0 \leq r' < d$. We distinguish three cases:

- if $q = q'$, then $r = n - qd = n - q'd = r'$, so $(q, r) = (q', r')$ and we are done.
- if $q > q'$, then $q \geq q' + 1$. Multiplying by d we get $qd \geq q'd + d$. Hence

$$q'd + r' = n = qd + r \geq q'd + d + r,$$

  whence $r' \geq d + r$, a contradiction because $r \geq 0$ and $r' < d$.
- symmetrically, if $q' > q$ one gets $r \geq r' + d$, another contradiction because $r' \geq 0$ and $r < d$. □

Here is a famous result by Euclid, with a proof that uses strong induction three times. We will see a much simpler proof in the next section.

Lemma 13 (Euclid)**.** *Let a and b be natural numbers. If a prime number p divides ab, it divides either a or b.*

Proof. [3] We proceed by strong induction on the minimum of the pair $\{a, b\}$. Up to relabeling, we can assume $a \leq b$, so that a is the smallest of the pair. If $a = 0$, or $a = 1$, then the claim is clear. So we assume $a \geq 2$ and distinguish two cases: either a is prime, or not.

- Suppose a is prime. Let p be a prime that does not divide a but divides ab for some b. Via Theorem 12, write $p = aq + r$ with $0 \leq r < a$. Since $ab = pc$ for some integer c, we have that

$$ab = pc = (aq + r)c = acq + rc.$$

  This implies that $rc = a(b - cq)$, so the prime a divides rc. Since $r < a$, by strong induction the theorem holds for the pair $\{r, c\}$; that is, when a prime divides rc, it divides either r or c. But the prime a divides rc and does not divide r, because $r < a$. Hence, a divides c. Writing $c = ad$ for some integer d, we get

$$ab = pc = pad,$$

  and canceling a we get $b = pd$. So p divides b.
- Suppose a is not prime. Then $a = d_1 \cdot d_2$, with both $d_1, d_2 < a$. Let p be a prime that does not divide a, and divides ab for some b. Then p divides neither d_1 nor d_2 (or else it would divide a), but

$$p \text{ divides } d_1 d_2\, b.$$

  By strong induction (since $d_1 < a$) the statement of the theorem holds for the pair $\{d_1, d_2 b\}$: so either p divides d_1 (which is false), or p divides $d_2 b$. Hence,

$$p \text{ divides } d_2\, b.$$

  By strong induction (since $d_2 < a$) the statement of the theorem holds for the pair $\{d_2, b\}$: so either p divides d_2 (which is false), or p divides $d_2 b$. Hence, p divides b. □

Lemma 14. *Let $a_1, \ldots, a_n$ be natural numbers. If a prime number p divides their product, then p divides (at least) one of $\{a_1, \ldots, a_n\}$.*

Proof. The case $n = 2$ is Lemma 13. By induction, suppose p divides $a_1 \cdot \ldots \cdot a_n \cdot a_{n+1}$. Call $b \stackrel{\text{def}}{=} a_1 \cdot \ldots \cdot a_n$. Since p divides $b \cdot a_{n+1}$, by Lemma 13 either p divides a_{n+1}, or p divides b, in which case by inductive assumption p divides one of $\{a_1, \ldots, a_n\}$. □

[3] Proof due to Barry Cipra, math.stackexchange.com/questions/1581173/proof-of-euclids-lemma, 2015

Theorem 15 (Unique Factorization)**.** *Any integer $n \geq 2$ can be decomposed as product of weakly-increasing primes, and such decomposition is unique.*

Proof. We already know that n can be written as product of primes by Proposition 7; so up to reordering them in weakly-increasing order, the existence of such decomposition is clear. The hard part is to prove uniqueness. Suppose

$$p_1 \cdot p_2 \cdot \ \ldots \ \cdot p_r = n = q_1 \cdot q_2 \cdot \ \ldots \ \cdot q_s,$$

with p_i, q_j primes, listed so that

$$p_1 \leq p_2 \leq \ \ldots \ \leq p_r \quad \text{and} \quad q_1 \leq q_2 \leq \ \ldots \ \leq q_s.$$

Since p_1 divides the product of the q_j's, by Lemma 14 it must divide at least one of the q_j's. Because they are both primes, this actually means that p_1 is *equal* to one of the q_j's. Since q_1 is the smallest of the q_j's, this means that

$$p_1 \geq q_1.$$

Symmetrically, q_1 divides the product of the p_i's, so it must divide one of them by Lemma 14. By primality, q_1 is equal to one of the p_i's, so in particular

$$p_1 \leq q_1.$$

Now we cancel p_1 and q_1, and proceed recursively. Because

$$p_2 \cdot \ \ldots \ \cdot p_r = q_2 \cdot \ \ldots \ \cdot q_s,$$

we get that $p_2 = q_2$; and so on. It follows that $r = s$ and $p_i = q_i$ for each i. □

The unique factorization (Theorem 15) provides a way to find the gcd of two natural numbers a and b: We can decompose a and b into primes, and then collect together all common factors.

Corollary 16. *Suppose $a = p_1^{a_1} \cdot \ \ldots \ \cdot p_h^{a_h} \cdot p_{h+1}^{a_{h+1}} \cdot \ldots \cdot p_r^{a_r}$ and $b = p_1^{b_1} \cdot \ \ldots \ \cdot p_h^{b_h} \cdot q_{h+1}^{b_{h+1}} \cdot \ldots \cdot q_m^{b_m}$ are decompositions into distinct primes, so that $\{p_{h+1}, \ldots, p_r\} \cap \{q_{h+1}, \ldots, q_m\} = \emptyset$. (Here each a_i, b_i and c_i is a positive integer.) Then,*

$$\gcd(a, b) = p_1^{\min(a_1, b_1)} \cdot \ p_2^{\min(a_2, b_2)} \ \ \ldots \ \cdot p_h^{\min(a_h, b_h)}.$$

Example 17. Since $168 = 2^3 \cdot 3 \cdot 7$ and $60 = 2^2 \cdot 3 \cdot 5$, then

$$\gcd(168, 60) = 2^2 \cdot 3 = 12.$$

This method seems quick. The problem is that we have hidden the difficulty under the carpet: Given 168, how can you find its prime factors quickly? In general, factoring requires a huge amount of computational time. Many cryptography systems (e.g. RSA) that keep your emails and credit cards secure, are based on the fact that a product of two distinct primes uniquely determines the two primes, but it takes really long to figure them out from the product if you do not have extra information.

0.2 Functions and Quotients. The sets $\mathbb{Z}$ and $\mathbb{Q}$

Let X, Y be any two sets. Recall that their Cartesian product is defined by

$$X \times Y \stackrel{\text{def}}{=} \{(x, y) \text{ such that } x \in X, y \in Y\}.$$

Definition 18. A *function* $f : X \longrightarrow Y$ consists of two sets X, Y and a subset $F \subseteq X \times Y$, such that for each element $x \in X$ there is always exactly one element y of Y for which $(x, y) \in F$. Usually we denote this y by $f(x)$, and we say it is the *image of* x *(under* f*).* We also call X (resp. Y) the *domain* (resp. the *codomain*) of the function. The *image of the set* X is the set $\operatorname{Im} X \stackrel{\text{def}}{=} \{f(x) \text{ such that } x \in X\}$.

Functions are typically described by a formula that tells us how to find $f(x)$ given x. For example, given any set X, the *identity function on* X (usually denoted by id_X) is the function whose output is always identical to the input. In this case, our notation to "explain the function" is

$$\begin{aligned} id_X : X &\longrightarrow X \\ x &\longmapsto x. \end{aligned}$$

Sometimes one does not have an explicit formula, but there is still a clear general method to associate x with its image: for example,

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \mathbb{N} \\ x &\longmapsto \text{the } x\text{-th prime number.} \end{aligned}$$

In the worst case scenario, if we do not see a general pattern, we can always express f by specifying all its values:

$$\begin{aligned} f : \{0, 1, 2\} &\longrightarrow \{0, 1, 2\} \\ 0 &\longmapsto 1 \\ 1 &\longmapsto 0 \\ 2 &\longmapsto 2. \end{aligned}$$

Definition 19. A function $f : X \to Y$ is *injective* if for each $x \neq x'$ one has $f(x) \neq f(x')$.

We assume familiarity with logic and quantifiers ($\forall$, $\exists$) and logical equivalence (contrapositives, etc.) For example, it should be clear that an equivalent way to define injectivity is:

$$\forall x, x' \in X, \quad f(x) = f(x') \Rightarrow x = x'.$$

Injectivity depends not only on the "formula", but also on the domain involved. For example, the function "first letter of" is injective on the set { Alba, Bruno }, but not injective on the set { Alba, Alice, Bruno }.

Definition 20. A function $f : X \to Y$ is *surjective* if the image of X coincides with the codomain Y; or in other words, if for each $y \in Y$ there is some $x \in X$ (not necessarily unique) such that $y = f(x)$.

Surjectivity depends not only on the "formula" for f, but also on the domain and the codomain. For example, let E be the set of even natural numbers:

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \mathbb{N} \\ x &\longmapsto 2x \end{aligned} \text{ is not surjective,} \qquad \begin{aligned} f : \mathbb{N} &\longrightarrow E \\ x &\longmapsto 2x \end{aligned} \text{ is,} \qquad \begin{aligned} f : \mathbb{N} \setminus \{0\} &\longrightarrow E \\ x &\longmapsto 2x \end{aligned} \text{ is not.}$$

Definition 21. A function $f : X \to Y$ is *bijective* if it is both injective and surjective. That is, if for each $y \in Y$ there exists exactly one $x \in X$ such that $y = f(x)$.

Given a function $f : X \to Y$ and a function $g : Y \to Z$, their *composite* is the function

$$\begin{array}{rccc} g \circ f : & X & \longrightarrow & Z \\ & x & \longmapsto & g(\, f(x)\,). \end{array}$$

Proposition 22. *Let $f : X \to Y$ be a function between two non-empty sets.*
(1) *f is surjective $\iff$ there exists $g : Y \to X$ (called "right inverse") such that $f \circ g = id_Y$.*
(2) *f is injective $\iff$ there exists $g : Y \to X$ (called "left inverse") such that $g \circ f = id_X$.*
(3) *f is bijective $\iff$ there exists $g : Y \to X$ (called "inverse") such that $g \circ f = id_X$ and $f \circ g = id_Y$.*

Remark 23. Before starting with the proof, note that two functions are equal when they have same domain, same codomain, and they yield same outputs when given same inputs. So to verify an equality of functions like $g \circ f = id_Y$, both going from Y to Y, we'll need to check that $g \circ f(y) = id_Y(y)$ for all $y \in Y$.

Proof of Proposition 22.
(1), "⇒". Define

$$\begin{array}{rccc} g : Y & \longrightarrow & X \\ y & \longmapsto & \text{some } x \text{ such that } f(x) = y. \end{array}$$

(If there is more than one x such that $f(x) = y$, we simply choose one.) Then by construction, $f \circ g(y) = f(x) = y$ for all $y \in Y$. Hence $f \circ g = id_Y$.
(1), "⇐". For each $y \in Y$, we know that $id_Y(y) = f \circ g(y)$, so $y = f(g(y))$, which means $y \in \operatorname{Im} f$.
(2), "⇒". Choose a point x_0 of X. Define

$$\begin{array}{rccc} g : Y & \longrightarrow & X & \\ y & \longmapsto & \begin{cases} x_0, & \text{if } y \notin \operatorname{Im} f \\ \text{the unique } x \text{ such that } f(x) = y, & \text{if } y \in \operatorname{Im} f. \end{cases} \end{array}$$

Then for all x in X, $g \circ f(x) = g(f(x)) = x$. So $g \circ f = id_X$.
(2), "⇐". Suppose $f(x) = f(x')$. Applying g, and remembering that $g \circ f = id_X$, we get

$$x = id_X(x) = g \circ f(x) = g(f(x)) = g(f(x')) = g \circ f(x') = id_X(x') = x'.$$

(3), "⇒". This does not follow immediately from items (1) and (2), because a priori it could be that the two g's (right inverse and left inverse) are different. However, if f is bijective we can simply define

$$\begin{array}{rccc} g : Y & \longrightarrow & X \\ y & \longmapsto & \text{the unique } x \text{ such that } f(x) = y. \end{array}$$

and it is easy to see that it does the trick.
(3), "⇐". This follows from (1) and (2). (Why?) □

Definition 24. Let X be an arbitrary, non-empty set. An *equivalence relation on X* is a subset R of $X \times X$ that satisfies the following properties:
REL1: $(x, x) \in R$ for all x. ("reflexivity")
REL2: If $(x, y) \in R$, then $(y, x) \in R$. ("symmetry")
REL2: If $(x, y) \in R$ and $(y, z) \in R$, then $(x, z) \in R$. ("transitivity")

Definition 25. Let R be equivalence relation on a set X, instead of $(x, y) \in R$ we shall write $x \sim y$, and read it "x is in a relation with y". The *equivalence class* of an element x of X is

$$\overline{x} \stackrel{\text{def}}{=} \{y \in X \text{ such that } y \sim x\} \stackrel{\text{def}}{=} \{y \in X \text{ such that } (x, y) \in R\}.$$

Example 26. On any non-empty set X, one can always put two "extreme" equivalence relation: the first one is

$$R_0 = \{(x, y) \text{ such that } x = y\}.$$

Under this, any element of X is in a relation only with himself. So the equivalence classes are as small as possible: They contain one element each.

The other extreme is

$$R_1 = X \times X.$$

Under this, any element of X is in a relation with everyone! So there is only one giant equivalence class containing all elements.

Example 27. Let X be the set of students in your Algebra class. If we define

$$R = \{(x, y) \text{ such that } x, y \text{ are born in the same year}\}$$

this is an equivalence relation. You are in a relation with anybody who is born the same year as you. If you view the student names as files, you can think of the equivalence classes as folders, labeled by birthyear.

Example 28. Let $\mathbb{N}$ be the set of integers. let us define

$$R = \{(a, b) \text{ such that } a - b \text{ is even}\}.$$

This is an equivalence relation, because (REL1) $a - a$ is even, (REL2) if $a - b$ is even so is $b - a$, and (REL3) if $a - b$ and $b - c$ are even, so is their sum $a - c$. This equivalence relation is called *congruence mod 2*.

Non-Example 29. The empty relation $R = \emptyset$ is not an equivalence relation: It satisfies (REL2) and (REL3), but not (REL1).

Non-Example 30. Let $X = \mathbb{N}$. The relation

$$R = \{(a, b) \text{ such that } |a - b| < 5\}$$

is not an equivalence relation: It satisfies (REL1) and (REL2), but not (REL3).

For a real-life analogy, "being close to" is not a relation of equivalence: if it takes less than 5 minutes to go from a to b, and it takes also less than 5 minutes to go from b to c, not necessarily it takes less than 5 minutes to go from a to c! It could be that distances add up.

Non-Example 31. Let $X = \mathbb{N}$. The relation

$$R = \{(a, b) \text{ such that } a \leq b\}$$

is not an equivalence relation: It satisfies (REL1) and (REL3), but not (REL2).

Definition 32. Let R be an equivalence relation on a set X. The *quotient* $X/_\sim$ is the set of all equivalence classes in X. In other words,

$$X/_\sim = \{\overline{x} \text{ such that } x \in X\}.$$

By definition, two elements of X are equal in the quotient (i.e. $\overline{x} = \overline{x'}$ in ${}^{X}\!/_{\sim}$) if and only if they are in a relation with one another (i.e. $x \sim x'$). Very often in mathematics, we have to define **functions or operations on quotients**. To define a function $f : {}^{X}\!/_{\sim} \to Y$, the usual procedure is to first define a function $F : X \to Y$, and then to check "compatibility with the quotient", i.e. check that

$$x \sim x' \Longrightarrow F(x) = F(x').$$

In particular, we can define an internal operation

$$o : {}^{X}\!/_{\sim} \times {}^{X}\!/_{\sim} \longrightarrow {}^{X}\!/_{\sim}$$

by first defining an operation

$$O : X \times X \longrightarrow X,$$

and then by checking that O is compatible with the quotient, i.e.

$$x \sim x', y \sim y' \Longrightarrow O(x, y) \sim O(x', y').$$

Example 33 ($\mathbb{Z}$ as quotient)**.** The set $\mathbb{Z}$ can be defined as follows: on $X \stackrel{\text{def}}{=} \mathbb{N} \times \mathbb{N}$ we introduce the equivalence relation

$$(a, b) \sim (a', b') \stackrel{\text{def}}{\Longleftrightarrow} a + b' = a' + b.$$

Then $\mathbb{Z} \stackrel{\text{def}}{=} {}^{X}\!/_{\sim}$. By convention, we denote $\overline{(a, 0)}$ simply by 'a' and $\overline{(0, a)}$ simply by '$-a$'.

The hint is: you should think of $\overline{(a, b)}$ as what you have always written as '$a - b$'. On the set $\mathbb{Z}$, we can define addition componentwise

$$\overline{(a, b)} + \overline{(c, d)} = \overline{(a + c, b + d)}$$

and multiplication as

$$\overline{(a, b)} \cdot \overline{(c, d)} = \overline{(ac + bd, ad + bc)}.$$

Are these operations legitimate? if $X = \mathbb{N} \times \mathbb{N}$, it is easy to see that the componentwise addition from $X \times X$ to X is compatible with the quotient: if $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$, which means that $a + b' = a' + b$ and $c + d' = c' + d$, then $a + c + b' + d' = a' + c' + b + d$, which means that $(a + c, b + d) \sim (a' + c', b' + d')$. But what about multiplication?, if $(a, b) \sim (a', b')$ and $(c, d) \sim (c', d')$, is it true that $(ac + bd, ad + bc) \sim (a'c' + b'd', a'd' + b'c')$? To check this, we assume $a + b' = a' + b$ and $c + d' = c' + d$, and we want to show

$$ac + bd + a'd' + b'c' = ad + bc + a'c' + b'd'.$$

The trick is to add to both sides of the equality above the quantity $ac' + bd' + a'd + b'c$, and prove that on both sides you get as result $2(a' + b)(c' + d)$. We leave this as exercise.

Example 34 ($\mathbb{Q}$ as quotient)**.** The set $\mathbb{Q}$ can also be defined as follows: on $Y \stackrel{\text{def}}{=} \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ we introduce the equivalence relation

$$(a, b) \sim (a', b') \stackrel{\text{def}}{\Longleftrightarrow} ab' = a'b.$$

Then $\mathbb{Q} \stackrel{\text{def}}{=} {}^{Y}\!/_{\sim}$. By convention, we denote $\overline{(a, b)}$ by '$\frac{a}{b}$'.

In other words, two fractions $\frac{a}{b}, \frac{a'}{b'}$ are considered identical if $ab' = a'b$. For example, $\frac{1}{2}$, $\frac{-1}{-2}$, and $\frac{3}{6}$ are the same. So if you want each rational number to be represented by precisely one pair (a, b), perhaps you would prefer to write something like

$$\mathbb{Q} = \{0\} \cup \left\{ \frac{a}{b} \text{ such that } a, b \in \mathbb{Z},\ b > 0,\ a \neq 0,\ \text{and } \gcd(a, b) = 1 \right\}.$$

Addition is defined by

$$\frac{a}{b} + \frac{c}{d} \stackrel{\text{def}}{=} \frac{ad + cb}{bd},$$

and multiplication by

$$\frac{a}{b}\,\frac{c}{d} \stackrel{\text{def}}{=} \frac{ac}{bd}.$$

Note that both operations are compatible with the quotient. In other words, if $\frac{a}{b} = \frac{a'}{b'}$ and if $\frac{c}{d} = \frac{c'}{d'}$, then by definition $ab' = a'b$ and $cd' = c'd$. So

$$(ad + cb)b'd' = ab'dd' + cd'bb' = a'bdd' + c'dbb' = (a'd' + b'c')bd \qquad \text{and} \qquad acb'd' = a'c'bd,$$

which imply, respectively, that

$$\frac{a}{b} + \frac{c}{d} \stackrel{\text{def}}{=} \frac{ad + cb}{bd} = \frac{a'd' + c'b'}{b'd'} \stackrel{\text{def}}{=} \frac{a'}{b'} + \frac{c'}{d'} \qquad \text{and} \qquad \frac{a}{b}\,\frac{c}{d} \stackrel{\text{def}}{=} \frac{ac}{bd} = \frac{a'c'}{b'd'} \stackrel{\text{def}}{=} \frac{a'}{b'}\,\frac{c'}{d'}.$$

We all know that $\mathbb{Z}$ can be viewed as a subset of $\mathbb{Q}$, thanks to the identification

$$\begin{aligned} \iota : \quad \mathbb{Z} &\longrightarrow \mathbb{Q} \\ z &\longmapsto \tfrac{z}{1}. \end{aligned}$$

This map ι is injective, but not surjective, of course: elements like $\frac{1}{2}$ are not in the image.

Perhaps less known is that there is another map between $\mathbb{Z}$ and $\mathbb{Q}$ that *is* injective and surjective. This follows from these two facts, which we just sketch:

- A bijective map from $\mathbb{N}$ to a set A is simply an exhaustive list, without repetition, of the elements of A. For example, listing all the elements of $\mathbb{Z}$ as

  $$0, 1, -1, 2, -2, \ldots$$

  shows that there is a bijection from $\mathbb{N}$ to $\mathbb{Z}$: It is the function that sends i to the i-th element of the list.
- Given a fraction $\frac{a}{b}$ with $\gcd(a, b) = 1$, let us agree to call *value of* $\frac{a}{b}$ the natural number $|a| + |b|$. For example, the value of $\frac{3}{4}$ is 7. We can order all nonzero elements of $\mathbb{Q}$ by increasing value, with the following subroutine: If two fractions have same value k, we order them according to the absolute value of the numerator; if they are still tied, we give priority to the positive fraction. (The numerator of a fraction $\frac{a}{b}$ is a; remember that we are excluding the pairs (a, b) if a and b have a factor in common, like $(2, 2)$.) This way we immediately obtain a list of all nonzero fractions that starts as follows:

  $$\begin{array}{cccccc} \underset{\text{(value 2)}}{\frac{1}{1}, -\frac{1}{1},} & \underset{\text{(value 3)}}{\frac{1}{2}, -\frac{1}{2}, \frac{2}{1}, -\frac{2}{1},} & \underset{\text{(value 4)}}{\frac{1}{3}, -\frac{1}{3}, \frac{3}{1}, -\frac{3}{1},} & \underset{\text{(value 5)}}{\frac{1}{4}, -\frac{1}{4}, \frac{2}{3}, -\frac{2}{3}, \frac{3}{2}, -\frac{3}{2}, \frac{4}{1}, -\frac{4}{1},} & \ldots \\ & & & & \ldots \end{array}$$

Definition 35. A *countable* set is a set that is either finite or in bijection with $\mathbb{N}$.

What we have just proven is this fact:

Theorem 36 (Cantor)**.** *$\mathbb{Z}$ and $\mathbb{Q}$ are countable.*

0.3 The Euclidean Algorithm and Diophantine Equations

Thanks to Euclidean division (Theorem 12) we can write down an algorithm to find the gcd that is much quicker than factoring. In addition, this algorithm will allows us to find *all* integer solutions of a linear equation $ax + by = c$, with a, b, c integers. For example, we can determine all integers x, y such that $8x + 5y = 6$. Equations where all solutions are required to be integers are usually called *Diophantine*, after Diophantus of Alexandria, who introduced them 1800 years ago. Despite this long history, they hide plenty of open problems.

But let us not ramble too much. We were saying, the algorithm!

Theorem 37. *Let n, m be positive integers. Let $n = qm + r$, with $0 \leq r < m$. Then*

$$\gcd(n, m) = \gcd(m, r).$$

Proof. Set $d_1 \stackrel{\text{def}}{=} \gcd(n, m)$ and $d_2 \stackrel{\text{def}}{=} \gcd(m, r)$. Since d_1 divides both n and m, it divides also $r = n - qm$; so it's a common divisor of m and r. So $d_1 \leq d_2$. On the other hand, d_2 divides m and r, so it divides also $n = qm + r$. So it's a common divisor of n and m. So $d_2 \leq d_1$. □

Algorithm 38. INPUT: a, b positive integers, with $b < a$.

def $\gcd(a, b)$:
  Do the Euclidean division $a = qb + r$.
  if $r = 0$:
    return b.
  else:
    return $\gcd(b, r)$.

The algorithm is recursive. Termination is ensured by the fact that at each iteration, the remainder decreases. Eventually, it will get to zero: but if the remainder of the division of x by y is zero, it means that $\gcd(x, y) = y$.

Example 39. Let us compute $\gcd(168, 60)$ using the Euclidean algorithm.

$$\begin{cases} 168 &= 2 \cdot 60 + 48 \\ 60 &= 1 \cdot 48 + 12 \\ 48 &= 4 \cdot 12 + 0. \end{cases}$$

So the $\gcd(168, 60) = \gcd(60, 48) = \gcd(48, 12) = 12$. Note that we did not have to factor 168.

An important consequence of the Euclidean Algorithm is Bezout's identity, which we explain next. Recall that by $\mathbb{Z}$ we mean the infinite set of "integers"

$$\mathbb{Z} = \{0, 1, -1, 2, -2, \ldots, n, -n, \ldots\}.$$

Theorem 40 (Bezout's Identity)**.** *Let a, b, c be positive integers. The equation*

$$ax + by = c$$

admits an integer solution $(x_1, y_1) \in \mathbb{Z} \times \mathbb{Z}$ $\iff$ c is a multiple of $\gcd(a, b)$.

Proof.

"⇒" This is obvious: If an integer d divides both a and b, then it divides also $ax + by$. Apply this to $d \stackrel{\text{def}}{=} \gcd(a, b)$.

"⇐" First of all, it suffices to show the claim for $c = \gcd(a, b)$. In fact, if $ax+by = d$ has integer solutions (x_1, y_1), then for any integer k the equation $ax + by = dk$ has integer solutions too, namely, (kx_1, ky_1). (If $ax_1 + by_1 = d$, then $a(kx_1) + b(ky_1) = k(ax_1 + ay_1) = kd$.) So let us focus on the equation

$$ax + by = c \text{ with } c = \gcd(a, b), \text{ and } a \geq 0,\ b \geq 0.$$

Now, suppose the Euclidean algorithm stops at the i-th iteration. That is,

(Step 0) $a = q_1 b + r_1$, with $0 < r_1 < b$.

(Step 1) $b = q_2 r_1 + r_2$, with $0 < r_2 < r_1$.

(Step 2) $r_1 = q_3 r_2 + r_3$, with $0 < r_3 < r_2$.

$\vdots$

(Step i) $r_{i-1} = q_{i+1} r_i + r_{i+1}$, with $0 = r_{i+1} < r_i$.

In this case, the output is r_i. So we can express $r_i = \gcd(a, b)$ as integer combination of r_{i-2} and r_{i-1}. In turn, we can "get rid of" r_{i-1} by expressing it as an integer combination of r_{i-2} and r_{i-3}:

$$r_i = r_{i-2} - q_i r_{i-1} = r_{i-2} - q_i \left(r_{i-3} - q_{i-1} r_{i-2}\right) = \left(1 + q_i q_{i-1}\right) r_{i-2} - q_i r_{i-3}.$$

Next, we can get rid of r_{i-2} with the substitution $r_{i-2} = r_{i-4} - q_{i-2} r_{i-3}$, and so on. After $i - 1$ steps, we will be able to write

$$r_i = ax \pm by, \text{ for some integers } x, y. \qquad \square$$

Example 41. Let us find an integer solution of $168x + 60y = 24$. From Example 39, we see that

$$12 = 60 - 1 \cdot 48 = 60 - (168 - 2 \cdot 60) = 3 \cdot 60 - 1 \cdot 168.$$

This gives us the solution $x_1 = -1$, $y_1 = 3$ to $168x + 60y = 12$. Multiplying by 2 we get the solution

$$x = -2, y = 6$$

to the equation we wanted.

Remark 42. What about Diophantine equations of the type

$$3x - 7y = 1\ ?$$

Well, this is the same as studying $3x + 7y = 1$. In fact, if $3x + 7y = 1$ admits an integer solution (x_1, y_1), then obviously $3x - 7y = 1$ admits the solution $(x_1, -y_1)$.

More generally, it is one of the Exercises to extend Bezout's theorem to arbitrary integers a, b, c (not necessarily positive) as follows: "The equation $ax + by = c$ admits an integer solution $(x_1, y_1) \in \mathbb{Z} \times \mathbb{Z} \iff |c|$ is a multiple of $\gcd(|a|, |b|)$".

Non-Example 43. Let us find an integer solution of $168x - 60y = 20$. There are none, because

$$168x - 60y = 12(14x - 5y)$$

must be a multiple of 12, and 20 is not.

Bezout's theorem also allows us to give a four-line, beautiful proof of Lemma 13.

Second proof of Lemma 13. Suppose p divides ab but not a; we claim that p divides b. In fact, since $\gcd(a,p)=1$, by Theorem 40 there are integers x,y such that $ax+py=1$. Multiplying by b, we get

$$abx + pby = b. \tag{4}$$

But since p divides ab, there is an integer k such that $ab=pk$, so Equation 4 becomes

$$p(kx+by) = b. \qquad \square$$

Remark 44. The Euclidean algorithm yields precisely *one* solution to $ax+by=d$, where $d=\gcd(a,b)$. However, there are many more. For example, $168x+60y=24$ is solved by $x=-2, y=6$ but also by $x=3, y=-8$.

How can we find *all* other solutions? To answer this question, we need to study homogenous linear Diophantine equations. The most important part is not to get scared by these three adjectives! They are just abbreviations of concepts we know already. "Linear" means that the equations look like $ax+by=c$, with a,b,c integer coefficients; "homogeneous" means that the constant term c is zero; and "Diophantine" means that we are only interested in integer solutions.

Lemma 45. *Let a,b be positive integers such that $\gcd(a,b)=1$. Let*

$$\mathcal{H}_0(a,b) \stackrel{\text{def}}{=} \{(x,y) \in \mathbb{Z}\times\mathbb{Z} \text{ such that } ax+by=0\}$$

be the set of integer solutions of $ax+by=0$. Then

$$\mathcal{H}_0(a,b) = \{(bk,-ak) \text{ such that } k\in\mathbb{Z}\}.$$

Proof.

"$\supset$" This is easy: if $x=bk$ and $y=-ak$, then $ax+by=abk-abk=0$.

"$\subseteq$" Suppose $ax=-by$. Let $b=p_1^{a_1}\cdots p_k^{a_k}$ be a factorization of b. Together with a factorization of $-y$, this forms a factorization of ax. By the uniqueness of factorizations (Theorem 15), this coincides with the factorization of ax obtained by juxtaposing a factorization of a and a factorization of x. However, since $\gcd(a,b)=1$, for each $i\in\{1,\ldots,k\}$, p_i does not divide a. This means that the factor $p_i^{a_i}$ that appears in the factorization of ax comes completely from the factorization of x. So for each $i\in\{1,\ldots,k\}$, $p_i^{a_i}$ divides x. So x is of the form

$$x = bk \text{ for some } k\in\mathbb{Z}.$$

It follows that $abk+by=0$, whence canceling b we get $y=-ak$. $\square$

Theorem 46. *Let a,b,c be positive integers such that $\gcd(a,b)=1$. Let*

$$\mathcal{H}_c(a,b) \stackrel{\text{def}}{=} \{(x,y) \in \mathbb{Z}\times\mathbb{Z} \text{ such that } ax+by=c\}$$

be the set of integer solutions of $ax+by=c$. Then

$$\mathcal{H}_c(a,b) = \{(cx_1+bk, cy_1-ak) \text{ such that } k\in\mathbb{Z}\},$$

where (x_1,y_1) is some particular solution of $ax+by=1$ (for example, the one given by the Euclidean algorithm.)

Proof.
"$\supset$" Suppose $x_2 = cx_1 + bk$ and $y_2 = cy_1 - ak$, then

$$ax_2 + by_2 = acx_1 + abk + bcy_1 - abk = acx_1 + bcy_1 = c(ax_1 + by_1) = c.$$

"$\subseteq$" Suppose $(x_2, y_2) \in \mathcal{H}_c(a, b)$. This means that $ax_2 + by_2 = c$. But also $c(ax_1 + by_1) = c$. Subtracting, we get that

$$a(x_2 - cx_1) + b(y_2 - cy_1) = 0.$$

This means that $(x_2 - cx_1, y_2 - cy_1) \in \mathcal{H}_0(a, b)$. By Lemma 45,

$$(x_2 - cx_1, y_2 - cy_1) = (bk, -ak) \text{ for some } k \in \mathbb{Z}.$$

So for this k we have that $x_2 = cx_1 + bk$ and $y_2 = cy_1 - ak$. □

Remark 47. Theorem 46 completely solves the problem of finding all integer solutions for $ax + by = c$, for any triple of integers a, b, c. In fact:

- In view of the considerations of Remark 42, we can always reduce ourselves to the case where a, b, c are natural numbers.
- If c is not a multiple of $\gcd(a, b)$, say no more!, there are no solutions.
- If $c \neq 0$ is a multiple of $\gcd(a, b)$, the equation simplifies. After canceling the common factor $\gcd(a, b)$, we face an equation

$$a'x + b'y = c'$$

with $\gcd(a', b') = 1$. So now Theorem 46 applies.

Example 48. Let's find all integer solutions of

$$168x - 60y = 24. \tag{5}$$

By Remark 42, let us study the equation

$$168x + 60y = 24. \tag{6}$$

Since $\gcd(168, 24) = 12$, we can simplify the equation: dividing everything by 12, we get

$$14x + 5y = 2.$$

Using the Euclidean algorithm, we can find a solution of $14x + 5y = 1$, namely, $x_1 = -1$, $y_1 = 3$. Now Theorem 46 tells us that the set of solutions for Equation 6 is

$$\mathcal{H}_2(14, 5) = \{(2x_1 + 5k, 2y_1 - 14k) \text{ such that } k \in \mathbb{Z}\} = \{(-2 + 5k, 6 - 14k) \text{ such that } k \in \mathbb{Z}\}.$$

But then the set of solutions for our original Equation 5 is

$$\{(-2 + 5k, -6 + 14k) \text{ such that } k \in \mathbb{Z}\}.$$

Corollary 49. *, If a, b are positive integers and $\gcd(a, b) = 1$, one can find integers x, y such that $ax + by = 1$ and $0 \leq x < b$.*

Proof. Let x_1, y_1 be an integer solution of $ax + by = 1$ (for example, the one given by the Euclidean algorithm). If $0 \leq x_1 < b$ we are done. Otherwise, there are two cases:

- If x_1 is negative, let k be the smallest integer such that $x_1+bk \geq 0$. Then clearly $x_1+bk < b$, because if $x_1 + bk \geq k$ then already $x_1 + (k-1)b \geq 0$, contradicting the definition of k. By Theorem 46, also $(x_1 + bk, y_1 - ak)$ is an integer solution.
- If $x_1 \geq b$, write $x_1 = bq + r$, with $0 \leq r < b$. By Theorem 46, applied to $k = -q$, also $(r, y_1 + aq)$ is an integer solution. □

Corollary 50. *If a, b are positive integers and $\gcd(a,b) = 1$, one can find infinitely many positive integers x, y such that $ax - by = 1$.*

Proof. Let x_1, y_1 be an integer solution of $ax + by = 1$ (for example, the one given by the Euclidean algorithm). Then $(x_1, -y_1)$ is a solution of the equation $ax - by = 1$. Moreover, the set of solutions of $ax - by = 1$ is of the form

$$\{(x_1 + bk, -y_1 + ak) \text{ such that } k \in \mathbb{Z}\}.$$

But if k is large enough, both $x_1 + bk$ and $-y_1 + ak$ are positive. More precisely, for any integer

$$k > \max\left(\frac{|x_1|}{b}, \frac{|y_1|}{a}\right),$$

one has $bk > |x_1|$ and $ak > |y_1|$, so in particular

$$x_1 + bk > x_1 + |x_1| \geq 0 \quad \text{and} \quad -y_1 + ak > -y_1 + |y_1| \geq 0.$$ □

We conclude this chapter with a useful reminder of some linear algebra:

Proposition 51 (cf. Rotman[4])**.** *Given n positive integers $a_1, \ldots, a_n$,*

$$\gcd(a_1, \ldots, a_n) = 1 \iff \begin{array}{l} \exists \text{ an } n \times n \text{ matrix } A \text{ with entries in } \mathbb{Z} \text{ whose} \\ \text{first row is } a_1, \ldots, a_n, \text{ and whose determinant is } 1. \end{array}$$

Proof.

"⇐" If we expand the determinant across the first row, we obtain $1 = \det A$ as a linear combination of $a_1, \ldots, a_n$ (with coefficients in $\mathbb{Z}$). Set $d \stackrel{\text{def}}{=} \gcd(a_1, \ldots, a_n)$. Since d divides all of the a_i's, d must divide also any linear combination of them; so d divides $\det A = 1$. This means $d = 1$.

"⇐", *alternative proof*: Having determinant 1, the matrix A has an inverse B whose entries are all integers. Since $AB = I$, the scalar product of the first row of A with the first column $(b_1, \ldots, b_n)$ of B, yields 1. In particular, there exists an integer solution $(b_1, \ldots, b_n)$ to the equation

$$a_1x_1 + \ldots + a_nx_n = 1.$$

"⇒" By induction on n. For $n = 2$ this is essentially Bezout's theorem (Theorem 40): Since one can find some integers x_1, x_2 such that $a_1x_1 + a_2x_2 = 1$, the matrix

$$A = \begin{pmatrix} a_1 & a_2 \\ -x_2 & x_1 \end{pmatrix}$$

has determinant 1, as desired. For larger n, set

$$a \stackrel{\text{def}}{=} \gcd(a_1, \ldots, a_{n-1}) \qquad \text{and} \quad b_i \stackrel{\text{def}}{=} \frac{a_i}{a} \quad (\text{for } i = 1, \ldots, n-1).$$

[4] Rotman, *Introduction to the Theory of Groups*, Springer, 1995, Theorem VI.4, p. 488

By construction, $\gcd(b_1, \ldots, b_{n-1}) = 1$. So we can apply the inductive assumption and find an $(n-1) \times (n-1)$ matrix B with entries in $\mathbb{Z}$ whose first row is $b_1, \ldots, b_{n-1}$ and whose determinant is 1. Let C be the submatrix formed by the lower $n-2$ rows of B. Note that by definition

$$\det \begin{pmatrix} b_1 & \ldots & b_{n-1} \\ & C & \end{pmatrix} = \det B = 1.$$

Now, $\gcd(a, a_n) = \gcd(a_1, \ldots, a_n) = 1$. So by Theorem 40, we can find integers s and t such that

$$a_n t + as = 1.$$

With these integers and with the aforementioned $(n-2) \times (n-1)$ matrix C, let us create the new $n \times n$ matrix

$$A \stackrel{\text{def}}{=} \begin{pmatrix} ab_1 & \ldots & ab_{n-1} & a_n \\ & C & & 0 \\ -tb_1 & \ldots & -tb_{n-1} & s \end{pmatrix}.$$

Let us see that A is the desired matrix. First of all all entries of A are integers and its first row is indeed equal to $(a_1, \ldots, a_{n-1}, a_n)$. It remains to show that $\det A = 1$. To this end, let us expand down the last column:

$$\begin{aligned} \det A &= (-1)^{n+1} a_n \det \begin{pmatrix} & C & \\ -tb_1 & \ldots & -tb_{n-1} \end{pmatrix} + (-1)^{2n} s \cdot \det \begin{pmatrix} ab_1 & \ldots & ab_{n-1} \\ & C & \end{pmatrix} = \\ &= (-1)^{(n+1)} (-1)^{n-2} a_n \cdot \det \begin{pmatrix} -tb_1 & \ldots & -tb_{n-1} \\ & C & \end{pmatrix} + s \cdot a \cdot \det \begin{pmatrix} b_1 & \ldots & b_{n-1} \\ & C & \end{pmatrix} = \\ &= (-1)^{2n-1} a_n \cdot (-t) \cdot \det \begin{pmatrix} b_1 & \ldots & b_{n-1} \\ & C & \end{pmatrix} + s \cdot a \cdot \det \begin{pmatrix} b_1 & \ldots & b_{n-1} \\ & C & \end{pmatrix} = \\ &= (-1)^{2n} a_n \cdot t \cdot \det B + s \cdot a \cdot \det B = a_n t + as = 1. \quad \square \end{aligned}$$

Deeper thoughts 52. While Diophantine Equations of degree 1 are understood (via Theorem 46), the study of Diophantine equations of higher degree is extremely difficult. There are however two beautiful results worth mentioning:

Theorem 53 (Sum of Two Squares theorem)**.** *Let $c \geq 2$ be an integer. The equation $x^2 + y^2 = c$ has positive integer solutions $\iff$ in the prime decomposition of c, all the prime factors of the form $4m+3$ are raised to an even exponent.*

Theorem 54 (Sum of Two Cubes theorem, Broughan 2003[5])**.** *Let $c \geq 2$ be an integer. The equation $x^3 + y^3 = c$ has positive integer solutions $\iff$ there exists a divisor d of c with $\sqrt[3]{c} \leq d \leq \sqrt[3]{4c}$, such that the number $\sqrt{\frac{4c-d^3}{3d}}$ is an integer.*

Example 55. For example, 100 can be written as sum of two squares, because the primes 2 and 5 involved in its factorization are not of the form $4m+3$. To check if 100 can be written as sum of cubes, let us look for divisors d between $\sqrt[3]{100} \approx 4.64$ and $\sqrt[3]{400} \approx 7.37$. So $d \in \{5, 6, 7\}$. Since $\sqrt{\frac{400-5^3}{3 \cdot 5}}$, $\sqrt{\frac{400-6^3}{3 \cdot 6}}$ and $\sqrt{\frac{400-7^3}{3 \cdot 7}}$ are not integers, it follows that 100 is not the sum of two cubes.

In contrast, 91 has a divisor, 7, in between $\sqrt[3]{91} \approx 4.50$ and $\sqrt[3]{364} \approx 7.14$, for which $\sqrt{\frac{364-7^3}{3 \cdot 7}} = \sqrt{\frac{21}{21}} = 1$. Hence, 91 is the sum of two cubes. However, it is not the sum of two squares: $91 = 7 \cdot 13$ and 7 is a prime of the form $4m+3$ raised to an odd exponent.

[5] Kevin A. Broughan, *Characterizing the sum of two cubes*, J. Int. Seq. 6 (2003).

A final shout-out goes to one of the most famous statements in mathematics, "Fermat's last theorem", claimed by Pierre de Fermat in 1637:

$$\forall n \geq 3,\ x^n + y^n = z^n \text{ has no positive integer solutions.}$$

The first complete proof of this fact was published in 1995 by Andrew Wiles, 358 years after Fermat's conjecture, and uses extremely advanced algebraic geometry.

0.4 From $\mathbb{Q}$ to $\mathbb{R}$: The need for geometry

It turns out that $\mathbb{Q}$ is not enough to describe elementary geometry. For example, the diagonal of the square whose edge has length 1, cannot be measured exactly within $\mathbb{Q}$. The following proof goes all the way back to Pythagoras:

Lemma 56 (Pythagoras)**.** *Let p be any prime number. No fraction has p as square. In other words, $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$, and so on, are not in $\mathbb{Q}$.*

Proof. By contradiction, suppose we could write uniquely

$$p = \left(\frac{a}{b}\right)^2, \text{ with } a, b \in \mathbb{Z},\ b > 0,\ a \neq 0, \text{ and } \gcd(a, b) = 1.$$

Clearing denominators,

$$pb^2 = a^2.$$

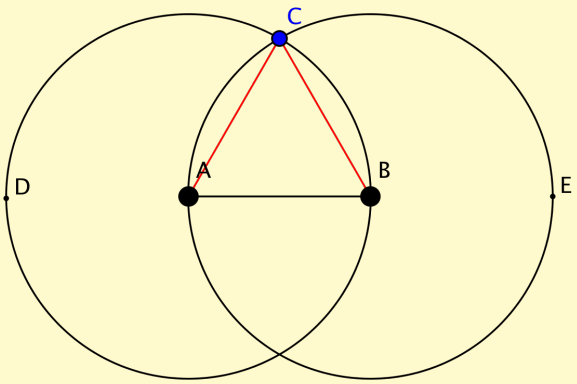
So p divides a^2. By Euclid's Lemma 13, this means that p divides a. So write $a = pk$, with k in $\mathbb{N}$. Plugging in, we get

$$pb^2 = (pk)^2 = p^2k^2,$$

or in other words $b^2 = pk^2$. But then p divides b^2 and by Lemma 13, p divides b. Hence p is a common factor of a and b. A contradiction, we assumed $\gcd(a, b) = 1$. □

Consider now the very first theorem of Euclid's *Elements*, perhaps the oldest global treatise of mathematics, published around 300 BC.

Proposition 1. *How to construct an equilateral triangle on a given segment.*



"It is required to construct an equilateral triangle on the segment AB. Describe the circle BCD with center A and radius AB. Again describe the circle ACE with center B and radius BA. Draw segments CA and CB from **the point C at which the circles cut one another** to the points A and B. Now, since the point A is the center of the circle CDB, therefore AC equals AB. Again, since the point B is the center of the circle CAE, therefore BC equals BA. But AC was proved equal to AB, so each of the segments AC and BC equals AB. Since things which equal the same thing also equal one another, AC also equals BC. Therefore the three segments AC, AB, and BC equal one another. Therefore the triangle ABC is equilateral, and it has been constructed on the given straight segment AB." □

We have highlighted a sentence in Euclid's original proof, from 2300 years ago. It is geometrically intuitive that the two circles should intersect somewhere, like Euclid claimed. But suppose for a moment that we lived in the plane $\mathbb{Q} \times \mathbb{Q}$. It would look like the usual Cartesian plane, except that we would only see the points whose coordinates are both rational. The definition of "circle with center A and radius r" as "the collection of points in $\mathbb{Q} \times \mathbb{Q}$ at distance r from A" would still make sense. But a rapid calculation with Pythagoras' theorem shows that if we place Cartesian coordinates with the origin in A, and if B has coordinates $(1, 0)$, say, then we should expect the two "rational circles" to intersect at the points

$$\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$$

which are not in $\mathbb{Q} \times \mathbb{Q}$ (cf. Lemma 56). So in the "rational plane" $\mathbb{Q} \times \mathbb{Q}$, already the first theorem of Euclid's book would be nonsense!, because the two "rational circles" would not intersect at all.

We point out that this "missing number" $\frac{\sqrt{3}}{2}$ can be easily expressed as "least upper bound" of the set

$$\left\{ x \in \mathbb{Q} \text{ such that } x^2 < \frac{3}{4} \right\},$$

which consists entirely of rational numbers. So the least upper bound of an infinite family of elements of $\mathbb{Q}$ need not be in $\mathbb{Q}$.

Another famous "limit" of doing geometry in $\mathbb{Q} \times \mathbb{Q}$ is the computation of curve lengths, like the perimeter of the unit circle. In his two essays *On the sphere and the Cylinder* and *Measurement of the Circle*, dating back to the third century b.C., Archimedes showed that in

any circle the ratio between perimeter and diameter is a number (called π) between $3 + \frac{10}{71}$ and $3 + \frac{1}{7}$. Archimedes obtained these bounds by comparing the perimeter of the inscribed regular polygon with 96 edges (which should be smaller than the circumference length) and the perimeter of the circumscribed regular polygon with 96 edges (which should be larger). The same method, applied to regular polygons with a higher and higher number of edges, lead to sharper and sharper bounds, allowing us to approximate this ratio π better and better. Implicit in Archimedes' reasoning, though, is the belief that such limit number π *must exist*, even though we can only compute rational approximations thereof. Quoting the Italian mathematician Giuseppe Peano[6]:

> "The Greek geometers, for what concerns the lengths of lines and the areas of surfaces (spheres, cylinders and so on), started from *postulates* instead of *definitions*. But the difference is only formal. The postulates that were stated by Archimedes in *On the sphere and the cylinder* are equivalent to the following definitions:
>
> - the length of a curvilinear plane convex arc is the common value of the least upper bound of the length of the polygonal inscribed arcs and the greatest lower bound of the circumscribed ones;
> - the area of a convex surface is the common value of the least upper bound of the length of the polygonal inscribed convex surfaces, and the greatest lower bound of the area of the circumscribed ones;
> - the length of a curvilinear arc is the least upper bound of the length of the polygonal inscribed arcs."

But as we saw above, there is no guarantee that such least upper bounds are in $\mathbb{Q}$. In conclusion, our rational plane is somewhat "incomplete": In order to do geometry, we need a larger set than $\mathbb{Q}$. This larger set, which you extensively studied in Calculus classes, is $\mathbb{R}$. Unfortunately, $\mathbb{R}$ can be formally defined only after taking a topology course. But below we sketch a construction that should give you an idea of what it is.

Definition 57. Let X be a set. A *sequence in* X is a function $a : \mathbb{N} \to X$ (not necessarily injective). For brevity, we denote the image $a(n)$ by a_n.

Definition 58. A *sequence in* $\mathbb{Q}$ is called

- *convergent*, if there exists $\ell \in \mathbb{Q}$ (called "limit" of the sequence) such that $\lim_{n\to\infty} a_n = \ell$; which means,

$$\forall k \in \mathbb{N}\ \exists M \in \mathbb{N} \text{ such that } \forall n \geq M \text{ we have } |a_n - \ell| < \frac{1}{k+1}.$$

- *Cauchy*, if

$$\forall k \in \mathbb{N}\ \exists M \in \mathbb{N} \text{ such that } \forall n, m \geq M \text{ we have } |a_n - a_m| < \frac{1}{k+1}.$$

Using the triangular inequality ($|a_n - a_m| \leq |a_n - \ell| + |a_m - \ell|$), it is easy to see that every convergent sequence is Cauchy. The converse is false: One of the many counterexamples is given by the increasing sequence

$$a_n = \sum_{i=0}^{n} \frac{1}{i!}.$$

[6] G. Peano, *Sulla definizione dell'area di una superficie*, Rendiconti dell'Accademia dei Lincei, 1890, 54–57; translated in A. Papadopoulos, *Metric Spaces Convexity and Nonpositive Curvature*, EMS 2005, p. 31.

To prove that this is Cauchy, note that for any three natural numbers $M \leq m < n$,

$$\begin{aligned} a_n - a_m &= \sum_{i=m+1}^{n} \frac{1}{i!} \leq \sum_{i=M+1}^{n} \frac{1}{i!} = \frac{1}{M!} \sum_{i=M+1}^{n} \frac{M!}{i!} = \frac{1}{M!} \sum_{i=M+1}^{n} \frac{1}{\prod_{h=M+1}^{i} h} \leq \\ &\leq \frac{1}{M!} \sum_{i=M+1}^{n} \frac{1}{\prod_{h=M+1}^{i} M} = \frac{1}{M!} \sum_{i=M+1}^{n} \frac{1}{M^{i-M}} \stackrel{!}{=} \frac{1}{M!} \sum_{j=1}^{n-M} \frac{1}{M^j} \leq \\ &\leq \frac{1}{M!} \sum_{j=1}^{\infty} \frac{1}{M^j} = \frac{1}{M!} \cdot \frac{\frac{1}{M+1}}{1 - \frac{1}{M+1}} = \frac{1}{M!} \cdot \frac{1}{M}. \end{aligned}$$

So this sequence satisfies the Cauchy definition by choosing $M = k$, for example. However, you might remember from Calculus (specifically, from Taylor series) that $\sum_{i=0}^{\infty} \frac{1}{i!} = e$; and intuitively, since one can prove that its limit e is not in $\mathbb{Q}$, the sequence (a_n) is not convergent in $\mathbb{Q}$.

(A cheap trick to generate counterexamples like this one, is the following: Pick a number that you already know to be irrational, like $\sqrt{2}$, and let a_i be the truncation at at the i-th digit of its decimal representation: $a_0 = 1$, $a_1 = 1.4$, $a_2 = 1.41$, $a_3 = 1.414$, and so on. Intuitively, a_n tends to $\sqrt{2}$, and for this reason the sequence will satisfy the Cauchy property, but $\sqrt{2}$ is not an element of $\mathbb{Q}$, so the sequence is not going to converge in $\mathbb{Q}$.)

The idea is now to artificially "add" to $\mathbb{Q}$ all the limits of all Cauchy sequences in $\mathbb{Q}$.

Definition 59. Let $\mathbb{R}$ be the set of all Cauchy sequences in $\mathbb{Q}$, with the following identification: We consider two sequences a_n and b_n identical if their difference $a_n - b_n$ converges to zero.

There is a natural way to view $\mathbb{Q}$ as subset of $\mathbb{R}$, via the map that associates to any rational number q the *constant* sequence, $(q, q, q, q, \ldots)$. However, there are much more elements in $\mathbb{R}$ than those coming from $\mathbb{Q}$. This was proven by Cantor using what is known today as *diagonal argument*. The starting point of this argument is that every real number x can be represented by means of a *decimal representation*:

$$x = a + \sum_{i=1}^{\infty} b_i \cdot 10^{-i}, \text{ with } a \in \mathbb{Z}, b_i \in \{0, 1, \ldots, 9\}.$$

This representation is not always unique. For example, 2.399999... is identical to 2.4 (because the difference tends to zero). However, this is the only thing that can go wrong: If we simply throw out all decimal representations that are eventually always nine, then every real number admits a unique decimal representation.

Theorem 60 (Cantor)**.** $\mathbb{R}$ *is not countable.*

Sketch of proof. It suffices to show that even the interval $(0, 1)$ is not countable. By contradiction, suppose we could list all elements of $(0, 1)$, as

$$x_1, x_2, \ldots, x_n, x_{n+1}, \ldots$$

By what we said above, every element x_i has a decimal representation, which consists of the integer 0 followed by a sequence of digits in $\{0, \ldots, 9\}$. (Remember we have thrown out representations that end with a nine periodic.)

Now construct an element y of $\mathbb{R}$ as follows: for the first decimal digit of y, choose either 0 or 1, making sure that your choice does not coincide with the first decimal digit of x_1. In particular, the number y we are writing down will be different from x_1: They differ in the first decimal place. For the second digit of y, again, choose 0 or 1, making sure not to agree with

the second decimal digit of x_2. In particular, $y \neq x_2$. And so on: For the i-th digit of y, choose either 0 or 1, disagreeing with the i-th digit of x_i. In the end, by construction you will have produced an element of $\mathbb{R}$ that is different from all x_i. A contradiction: The x_i were supposed to be a complete list of all elements of $\mathbb{R}$. □

In fact, Cantor proved something even stronger:

Theorem 61 (Cantor)**.** *Any countable subset of $\mathbb{R}$ has (Lebesgue) measure zero in $\mathbb{R}$. In particular, picking a random point in the real interval $[0,1]$, the probability that it is rational is zero.*

Sketch of proof. Let S be a countable subset of $\mathbb{R}$. Let $x_1, x_2, \ldots$ be a complete list (without repetitions) of all elements of S. For any $\varepsilon > 0$, let us construct a symmetric interval centered at the real number x_n and of radius $\frac{\varepsilon}{2^{n+1}}$,

$$A_n \stackrel{\text{def}}{=} \left(x_n - \frac{\varepsilon}{2^{n+1}}, x_n + \frac{\varepsilon}{2^{n+1}}\right).$$

These A_n's might not be disjoint, of course. Their (Lebesgue) measure obviously depends on n:

$$m(A_n) = \frac{2\varepsilon}{2^{n+1}} = \frac{\varepsilon}{2^n}.$$

But since S is contained in the union of all these intervals,

$$m(S) \leq m(\bigcup_n A_n) \leq \sum_{n=1}^{\infty} m(A_n) = \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon \cdot \sum_{n=1}^{\infty} \frac{1}{2^n} = \varepsilon \cdot 1 = \varepsilon.$$

Since ε can be chosen arbitrarily small, this implies that S has Lebesgue-measure zero. Now, probability is defined as the ratio of the measure of favorable cases and the measure of all possible cases. In particular, if we pick a random point in the segment $[0,1]$ of the real line, which has measure 1, the probability that we hit a rational is $\frac{0}{1}$. □

0.5 Modular Arithmetics and Divisibility Criteria

Fix an integer $m \geq 2$. Let a, b be two natural numbers. We say that "a is congruent to b modulo m", and write

$$a \equiv b \mod m,$$

if $a - b$ is multiple of m.

Example 62. 15 is congruent to 3 modulo 12. In fact, $15 - 3$ is a multiple of 12.

Congruence is an equivalence relation (check!). It behaves well with respect to products and sums:

Lemma 63. *If $a \equiv a' \mod m$ and $b \equiv b' \mod m$, then*

$$ab \equiv a'b' \mod m \quad \text{and} \quad a + b \equiv a' + b' \mod m.$$

Proof. By assumption, there are integers c, d such that $a - a' = cm$ and $b - b' = dm$. Then

$$ab - a'b' = ab - a'b + a'b - a'b' = b(a - a') + a'(b - b') = bcm + a'dm$$

is a multiple of m, so $ab \equiv a'b'$. Similarly,

$$(a + b) - (a' + b') = (a - a') + (b - b') = cm + dm$$

is a multiple of m, so $a + b \equiv a' + b'$. □

Note that every number a is congruent modulo 12 to some r such that $0 \leq r < 12$. In fact, we could simply perform the Euclidean division of a by 12: we get

$$a = q \cdot 12 + r, \text{ with } 0 \leq r < 12,$$

and by definition $a \equiv r \mod 12$.

Since congruence mod m is an equivalence relation on $\mathbb{Z}$, we want now to use the **quotient**, which we should call $\mathbb{Z}_m$, and perform operations on it. All this should sound familiar, because our clocks work modulo 12 and our calendar modulo 7. Twentyfive hours after ten pm, it is going to be eleven pm — nobody would call it "thirtyfive pm". And if today is Tuesday, you all know that in 28 days it will still be Tuesday.

Here is the formal definition.

Definition 64. Let $m \geq 2$ be an integer. Let $\mathbb{Z}_m$ be the quotient of $\mathbb{Z}$ by the equivalence relation "congruence mod m". As a set, $\mathbb{Z}_m$ has exactly m elements, which we should write as $\{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$, atlhough we will soon drop the overline bar from the notation. The *modular addition* and the *modular multiplication* are the two operations on $\mathbb{Z}_m$ defined as follows:

$$\begin{aligned} a \oplus b &\stackrel{\text{def}}{=} \text{the remainder of the division of } a+b \text{ by } m \\ a \odot b &\stackrel{\text{def}}{=} \text{the remainder of the division of } ab \text{ by } m. \end{aligned}$$

Note that the operations are well-defined: if $a \equiv a'$ and $b \equiv b'$, then $a+b \equiv a'+b'$ and $ab \equiv a'b'$ by Lemma 63. In plain words, if we drop the "overline bar" from the notation, $\overline{a} \oplus \overline{b}$ is the unique number in $0, \ldots, m-1$ congruent to $a+b$ mod m, and $a \odot b$ is the unique number in $0, \ldots, m-1$ congruent to ab mod m. (Later in the textbook, we will also drop the circle around the operation from the notation.)

Remark 65. An important difference between modular arithmetics and usual arithmetics, is that in modular arithmetics sums and products of positive integers can be zero. For example, in $\mathbb{Z}_6$ we have

$$\begin{aligned} 1 \oplus 5 &= 0. \\ 2 \odot 3 &= 0. \end{aligned}$$

Apart from this, modular operations behave in a very similar manner to usual arithmetic operations; later in the course, once we introduce quotients rings, we will understand why.

Lemma 66. *Both $\oplus$ and $\odot$ are associative: That is, for each a, b, c,*

$$(a \oplus b) \oplus c = a \oplus (b \oplus c) \quad \text{and } (a \odot b) \odot c = a \odot (b \odot c).$$

Proof. We prove it only for $\oplus$; the $\odot$ case is analogous and left as exercise. By definition, $a \oplus b$ is the remainder r_1 of the Euclidean division

$$(a+b) = q_1 m + r_1, \text{ with } 0 \leq r_1 < m.$$

So $(a \oplus b) \oplus c$ is the remainder r_2 of the Euclidean division

$$r_1 + c = q_2 m + r_2, \text{ with } 0 \leq r_2 < m.$$

So $r_2 = r_1 + c - q_2 m = (a + b - q_1 m) + c - q_2 m = (a+b+c) - (q_1+q_2)m$; hence,

$$(a+b+c) = (q_1+q_2)m + r_2,$$

and since $r_2 < m$, the expression above is a Euclidean division. In other words, r_2 is the remainder of the division of $a + b + c$ by $(q_1 + q_2)$.

On the other hand, $b \oplus c$ is by definition the remainder r_3 of the Euclidean division

$$(b + c) = q_3 m + r_3, \text{ with } 0 \leq r_3 < m.$$

So $a \oplus (b \oplus c)$ is the remainder r_4 of the Euclidean division

$$a + r_3 = q_4 m + r_4, \text{ with } 0 \leq r_4 < m.$$

But then $r_4 = a + r_3 - q_4 m = a + (b + c - q_3 m) - q_4 m = (a + b + c) - (q_3 + q_4)m$; and as above, we get that

$$(a + b + c) = (q_3 + q_4)m + r_4, \text{ with } 0 \leq r_4 < m$$

is also a Euclidean division. By the uniqueness of the Euclidean division, it follows that $r_2 = r_4$, so $(a \oplus b) \oplus c = a \oplus (b \oplus c)$. □

Lemma 67. *The operation $\odot$ distributes $\oplus$: That is, for each a, b, c,*

$$(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c).$$

Proof. As above, $a \oplus b \stackrel{\text{def}}{=} r_1$, where

$$(a + b) = q_1 m + r_1, \text{ with } 0 \leq r_1 < m.$$

In turn, $(a \oplus b) \odot c \stackrel{\text{def}}{=} r_2$, where

$$r_1 c = q_2 m + r_2, \text{ with } 0 \leq r_2 < m.$$

In particular, $r_2 = r_1 c - q_2 m = (a + b - q_1 m)c - q_2 m = (ac + bc) - (cq_1 + q_2)m$. Since $r_2 < m$, we see that r_2 is the remainder of the Euclidean division of $(ac + bc)$ by m.

Now let $r_3 = a \odot c$ and $r_4 = b \odot c$ be the remainders of the Euclidean division by m of ac and bc, respectively. Clearly, $r_3 + r_4$ will be the remainder of the Euclidean division by m of $(ac + bc)$. But then $r_3 + r_4 = r_2$. □

These properties make it possible to speed up calculations. Suppose we have to compute (abc) mod m. Instead of computing abc (which could be a large number) and then dividing it by m, here is what we could do: First we compute ab mod m, then we multiply this result (which is a number smaller than m) by c, and finally we take the remainder mod m.

Example 68. Let us compute the last two digits of $89 \cdot 67 \cdot 67 \cdot 3 \cdot 3$. The last two digits of any number n are the remainder of the division of n by 100. In other words, we have to find $89 \odot 67 \odot 67 \odot 3 \odot 3$ in $\mathbb{Z}_{100}$. Here is a quick strategy: We first compute $67 \odot 3$ in $\mathbb{Z}_{100}$. Since $67 \cdot 3 = 201 = 2 \cdot 100 + 1$, we have $67 \odot 3 = 1$. But then, in $\mathbb{Z}_{100}$

$$89 \odot 67 \odot 3 \odot 67 \odot 3 = 89 \odot 1 \odot 1 = 89.$$

Let us now see a simple way to use distributivity. Note that for any number a, we have $a \oplus (m - a) = 0$. A trick that sometimes pays off in modular arithmetics (especially if a is almost as large as m) is to perform all calculations using the "opposite" $m - a$. By the distributive property, for any b we have

$$(a \odot b) \oplus ((m - a) \odot b) = (a \oplus m - a) \odot b = 0 \odot b = 0.$$

Example 69. What are the last two digits of $7499 \cdot 8656$? The direct way to answer is to compute the remainder of the division of $7499 \odot 8656$ by 100, which (calculator in hand) is 44. However, it is much simpler to work in $\mathbb{Z}_{100}$. By the distributivity property, $99 \odot 56$ must be the number that added to $1 \odot 56$ is 0. But we can compute $1 \odot 56$ very easily: It is 56. Hence, $99 \odot 56$ must be $100 - 56 = 44$.

Similarly, using Newton's formula ($(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k\, y^{n-k}$) we have that

$$\begin{array}{rl} (m-a)^2 & = m^2 - 2ma + a^2 \\ (m-a)^3 & = m^3 - 3m^2a + 3ma^2 - a^3 \\ \vdots & = \quad \vdots \\ (m-a)^{2k} & = (\text{some multiple of } m) + a^{2k} \\ (m-a)^{2k+1} & = (\text{some multiple of } m) - a^{2k+1}. \end{array}$$

Therefore, if $0 \leq a < m$, in $\mathbb{Z}_m$ we have

$$(m-a)^h = \begin{cases} a^h & \text{if } h \text{ is even,} \\ m - a^h & \text{if } h \text{ is odd.} \end{cases} \tag{7}$$

Example 70. What is the last digit of 9^{17}? We should work in $\mathbb{Z}_{10}$. Since $9 + 1 = 10$, and $1^{17} = 1$, then 9^{17} should be $10 - 1^{17} = 9$.

Modular arithmetics can also be used to explain all divisibility criteria.

Proposition 71. *A number is divisible by 9 if and only if the sum of its digits is. Same for divisibility by 3.*

Proof. The point is that 10 is congruent to 1 modulo 9 (or modulo 3). We focus on divisibility by 9; the same argument applies also to 3. If a number x is written in base 10 with n digits

$$(b_{n-1}|b_{n-2}|\ldots|b_1|b_0),$$

it means that

$$x = 10^{n-1}b_{n-1} + 10^{n-2}b_{n-2} + \ldots + 10b_1 + b_0.$$

For example, $4235 = 10^3 \cdot 4 + 10^2 \cdot 2 + 10 \cdot 3 + 5$. Now 10 is congruent to 1 mod 9; hence, $10^{n-1}b_{n-1} + 10^{n-2}b_{n-2} + \ldots + 10b_1 + b_0$ is congruent to $1^{n-1}b_{n-1} + 1^{n-2}b_{n-2} + \ldots + 1 \cdot b_1 + b_0$, which is the sum of the digits of x. So if we call $s(x)$ the sum of the digits of x, we have just proven that

$$x \equiv s(x) \mod 9.$$

In particular, $x \equiv 0 \mod 9$ if and only if $s(x) \equiv 0 \mod 9$. □

Proposition 72. *A number is divisible by 2^k if and only if the number formed by reading only its last k digits is. Same for divisibility by 5^k.*

Proof. The point is that 2^k and 5^k divide 10^k, and in particular, they divide 10^h for all $h \geq k$. So if the digits of x are

$$(b_{n-1}|b_{n-2}|\ldots|b_1|b_0),$$

when we pass modulo 2^k we have

$$\begin{array}{rcll} x & = & 10^{n-1}b_{n-1} + 10^{n-2}b_{n-2} + \ldots + 10^k b_k & +10^{k-1}b_{k-1} + \ldots + 10b_1 + b_0 \equiv \\ & \equiv & 0 & +10^{k-1}b_{k-1} + \ldots + 10b_1 + b_0. \end{array}$$

Hence modulo 2^k, the number x is congruent to the number formed by its last k digits. The same is true modulo 5^k. □

Proposition 73. *A number is divisible by* 11 *if and only if the alternating sum of its digits is.*

Proof. The point is that 10 is the opposite of 1 in $\mathbb{Z}_{11}$. So modulo 11, by Equation 7, we have

$$10^h \equiv \begin{cases} 1 & \text{if } h \text{ is even,} \\ 10 & \text{if } h \text{ is odd.} \end{cases}$$

So if the digits of x are $(b_{n-1}|b_{n-2}|\dots|b_1|b_0)$, and if n is even, we have

$$\begin{aligned} x &= b_0 + 10b_1 + 10^2b_2 + 10^3b_3 + \dots + 10^{n-2}b_{n-2} + 10^{n-1}b_{n-1} + \equiv \\ &\equiv b_0 + 10b_1 + b_2 + 10b_3 + \dots + b_{n-2} + 10b_{n-1} \equiv \\ &\equiv b_0 - b_1 + b_2 - b_3 + \dots + b_{n-2} - b_{n-1}; \end{aligned}$$

whereas if n is odd we have

$$\begin{aligned} x &= b_0 + 10b_1 + 10^2b_2 + 10^3b_3 + \dots + 10^{n-2}b_{n-2} + 10^{n-1}b_{n-1} + \equiv \\ &\equiv b_0 + 10b_1 + b_2 + 10b_3 + \dots + 10b_{n-2} + b_{n-1} \equiv \\ &\equiv b_0 - b_1 + b_2 - b_3 + \dots - b_{n-2} + b_{n-1}. \end{aligned}$$

So either way, x is congruent to $\sum_{i=0}^{n-1}(-1)^i b_i$ modulo 11. In particular, x is congruent to 0 modulo 11 if and only if the alternating sum of its digits is. □

The same arguments can be adapted also to representations of numbers in some base $b \neq 10$. For example,

- Let $d > 1$ be any divisor of $b - 1$. A number is divisible by d if and only if the sum of its digits in base b is divisible by b. (The proof is analogous to that of Proposition 71.)
- Let $d > 1$ be any divisor of b. A number is divisible by d^k if and only if the number formed by reading only its last k digits in base b is divisible by d^k. (Proof: Like Proposition 72.)
- Let $d > 1$ be any divisor of $b + 1$. A number is divisible by d if and only if the alternating sum of its digits in base b is divisible by d. (Proof: Like Proposition 73.)

Example 74. Here are some conversions from base 10 to base 2:

$$\begin{array}{ll} 3_{(10)} = 11_{(2)} & 6_{(10)} = 110_{(2)} \\ 9_{(10)} = 1001_{(2)} & 12_{(10)} = 1100_{(2)} \\ 15_{(10)} = 1111_{(2)} & 18_{(10)} = 10010_{(2)} \\ 21_{(10)} = 10101_{(2)} & 24_{(10)} = 11000_{(2)} \\ 27_{(10)} = 11011_{(2)} & 30_{(10)} = 11110_{(2)} \end{array}$$

from which we see that the alternating sum of the digits in base 2 is again a multiple of 3.

*Divisibility by 7: A Recursive Algorithm

In $\mathbb{Z}_7$, the equation $3x = 1$ has a solution, namely, $x = 5$. This gives the following recursive criterion for testing divisibility by 7:

Proposition 75. *Let a, b be integers.*

$$10a + b \text{ is divisible by } 7 \iff a - 2b \text{ is divisible by } 7,$$

and the remainder of the left hand side (of the division by 7) is three times the remainder on the right.

Proof. Since in $\mathbb{Z}_7$ we have $10 = 3$, $1 = 3 \cdot 5$, and $5 = -2$, we can write

$$10a + b = 3a + b = 3(a + 5b) = 3(a - 2b). \qquad \square$$

The advantage of this criterion is that the integer $a - 2b$ is smaller than a, which is smaller than one tenth of $10a + b$. So we reduced the check for divisibility to a much smaller number. The criterion can be iterated, rapidly yielding a number with few digits, for which we are able to tell if it's a multiple of 7 or not.

Interestingly, there is nothing special about the prime 7; this recursive criterion works for all primes different than 2 and 5, and more generally, for all numbers with last digit $1, 3, 7$ or 9.

Proposition 76. *Let a, b be integers. Let $m \geq 2$ be any integer with last digit $1, 3, 7$ or 9. Let $v = m - x$, where x is a solution in $\mathbb{Z}_m$ of the equation $10x = 1$. Then*

$$10a + b \text{ is divisible by } m \iff a - vb \text{ is divisible by } m,$$

and the remainder of the left hand side (of the division by m) is 10 times the remainder on the right, (where the multiplication is performed in $\mathbb{Z}_m$).

Proof. The hard part is to check that a solution to the equation $10x = 1$ exists in $\mathbb{Z}_m$. By the divisibility criteria, m is neither a multiple of 2 nor a multiple of 5. Hence $\gcd(m, 10) = 1$. By Bezout's Theorem, and specifically by Corollary 49, we can find integers x, y such that $10x + my = 1$, with $x \in \{0, \ldots, m - 1\}$. Then in $\mathbb{Z}_m$ we have $10x = 1$ and

$$10a + b = 10(a + xb) = 10(a - vb). \qquad \square$$

Example 77. Let $m = 31$. Then $10x = 1$ in $\mathbb{Z}_{31}$ has the solution $x = 28$. So $v = 3$ and

$$10a + b \text{ is divisible by } 31 \iff a - 3b \text{ is divisible by } 31.$$

More generally, if m is of the form $10k + 1$, then obviously in $\mathbb{Z}_m$

$$10(m - k) = 10m - 10k = 10m - (m - 1) = 1,$$

so $x = m - k$ is a solution of $10x = 1$ in $\mathbb{Z}_m$ and $v = m - (m - k) = k$. So

$$10a + b \text{ is divisible by } 41 \iff a - 4b \text{ is divisible by } 41,$$

$$10a + b \text{ is divisible by } 51 \iff a - 5b \text{ is divisible by } 51,$$

and so on.

0.6 *Fermat's little theorem and decimal representation

In this section we worry about powers modulo a prime. This will help us understand the decimal representation of numbers. Have you ever noticed that

$$\tfrac{1}{7} = 0,\overline{142857}, \qquad \tfrac{2}{7} = 0,\overline{285714}, \qquad \tfrac{3}{7} = 0,\overline{428571},$$

$$\tfrac{4}{7} = 0,\overline{571428}, \qquad \tfrac{5}{7} = 0,\overline{714285}, \qquad \tfrac{6}{7} = 0,\overline{857142},$$

have periods of the same length? Also, the periods are formed by the same digits cyclically shifted! Why is that the case? Does this happen only for division by 7?

To explain this, we start with a simple fact.

Lemma 78. *Let n be any integer ≥ 2. Then,*

$$n \text{ is prime} \iff \text{for all integers } k \in \{1, \ldots, n-1\}, \binom{n}{k} \text{ is a multiple of } n.$$

Proof.
"⇒" By definition

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k(k-1) \cdots 2 \cdot 1}.$$

If n is prime, all of the factors below the fraction line are *not* multiples of n, because they are all smaller than n. So the prime factor n above the fraction line does not cancel out.

"⇐" If n is not prime, by Theorem 15 it has a prime divisor $p \in \{1, \ldots, n-1\}$. Say $n = pd$. The previous multiple of p is $n - p = p(d-1)$, so the numbers in $\{n-p+1, \ldots, n-1\}$ are not multiples of p. But then in the expression

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

a factor p is "canceled out" from the numerator. More precisely, the exponent of p in the unique factorization of n is larger than the exponent of p in the unique factorization of $\binom{n}{p}$. In particular, n cannot divide $\binom{n}{p}$. □

Lemma 79 (Freshman's dream)**.** *For any prime number p, for any integers x, y,*

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

Proof. By Newton's formula,

$$(x+y)^p = x^p + \binom{p}{1} x^{p-1}y + \ldots + \binom{p}{k} x^{p-k}y^k + \ldots + \binom{p}{1} xy^{p-1} + y^p.$$

But since p is prime, by Lemma 78 all binomials $\binom{p}{k}$ with $1 \leq k \leq p-1$ are multiples of p. □

Theorem 80 (Fermat's little theorem, 1640)**.** *If p is any prime number, then for any $a \in \mathbb{N}$ one has*

$$a^p \equiv a \mod p.$$

In particular, if $b \in \mathbb{N}$ is not a multiple of p, then $b^{p-1} \equiv 1 \mod p$.

Proof. By induction on a. For $a = 0$ the claim is clear. Suppose that $a^p \equiv a$ for some a; then using the Freshman's Dream (Lemma 79) we obtain

$$(a+1)^p \equiv a^p + 1^p \equiv a + 1.$$

This proves the first part of the claim. Now suppose b is not a multiple of p. By the first part we know that p divides $b^p - b$. But

$$b^p - b = b(b^{p-1} - 1)$$

and p does not divide b, so by Euclid's Lemma 13, p divides $(b^{p-1} - 1)$. Hence, $b^{p-1} \equiv 1 \mod p$. □

Remark 81. The "converse" of Fermat's little theorem is false: There are composite numbers c such that for any $a \in \mathbb{N}$,

$$a^c \equiv a \mod c.$$

(The smallest such number, found by Carmichael in 1910, is $c = 561 = 3 \cdot 11 \cdot 17$. It is known that there are infinitely many numbers c of this type, called "Carmichael numbers".) We will see that two slightly stronger versions of the converse are true (Theorem 152 and Proposition 154).

Fermat's little theorem tells us that the set of integers u such that $10^u - 1$ is a multiple of p, is never empty, unless 10 is a multiple of p. This justifies the following definition:

Definition 82. Let p be a prime number different than 2 or 5. The *multiplicative order of* 10 *mod* p is the smallest positive integer u such that p divides $(10^u - 1)$.

Lemma 83. *The multiplicative order of* 10 *mod* p *divides* $p - 1$.

Proof. Let u be the multiplicative order of 10 mod p. Write

$$(p-1) = qu + r \text{ with } 0 \leq r < u.$$

If $r = 0$ then u divides $(p-1)$ and we are done. If $r \neq 0$, we have

$$10^{p-1} = 10^{qu+r} = (10^u)^q \cdot 10^r.$$

Now 10^{p-1} is congruent to 1 modulo p, and 10^u is congruent to 1 modulo p. By lemma 63, it follows that 10^r is also congruent to 1 mod p. A contradiction, r would be a positive integer smaller than u such that p divides $10^r - 1$. □

Theorem 84 (Glaisher 1878 [7])**.** *Let* p *be a prime number different than* 2 *or* 5. *Let* k *be an integer that is not a multiple of* p. *Then the decimal expansion of* $\frac{k}{p}$ *is infinite periodic, and the number of digits in the period is the multiplicative order of* 10 *mod* p. *(In particular, it does not depend on* k.*)*

Proof. Without loss of generality, we can assume $k > 0$, otherwise we can replace it by $-k$. Moreover, we can reduce ourselves to the case $1 \leq k < p$: In fact, if $k \geq p$, we can replace k by the remainder r of the division

$$k = qp + r, \text{ with } 0 \leq r < p :$$

The decimal parts of $\frac{k}{p}$ and of $\frac{r}{p}$ are the same, because they differ by the integer q.

So, assume $1 \leq k < p$. Let u be the multiplicative order of 10 mod p. Let d be the positive integer such that

$$10^u - 1 = pd. \tag{8}$$

Recall from Calculus 2 the formula for Geometric Series of ratio $r < 1$:

$$a + ar + ar^2 + ar^3 + \ldots + a^n + a^{n+1} + \ldots = \frac{a}{1-r} \tag{9}$$

In particular, for $a = r = \frac{1}{10^u}$ we obtain:

$$\frac{1}{10^u} + \frac{1}{10^{2u}} + \frac{1}{10^{3u}} + \ldots = \frac{1}{10^u} \cdot \frac{1}{1 - \frac{1}{10^u}} = \frac{1}{10^u - 1}. \tag{10}$$

[7] Glaisher, J. W. L., "Periods of Reciprocals of Integers Prime to 10." Proc. Cambridge Philos. Soc. 3 (1878), 185-206

Putting together Equation 8 and Equation 10, we obtain

$$\frac{k}{p} = \frac{kd}{pd} = \frac{kd}{10^u - 1} = kd\left(\frac{1}{10^u} + \frac{1}{10^{2u}} + \frac{1}{10^{3u}} + \dots\right) \tag{11}$$

Since by assumption $1 \leq k < p$, equation 11 represents a number $\frac{k}{p}$ which is strictly between 0 and 1. Moreover, the number kd has at most u digits, because

$$kd < pd = 10^u - 1,$$

which is the number consisting of u nines. We can assume that kd has exactly u digits: If not, we can always make it a u-digit number by appending some initial zeroes. So Equation 11 shows us that the block of u digits representing kd repeats. This means that the number of digits of the period is $\leq u$.

So it remains to exclude that the period a w-digit number c, with $w < u$. Were this the case, we could write

$$\frac{k}{p} = c\left(\frac{1}{10^w} + \frac{1}{10^{2w}} + \frac{1}{10^{3w}} + \dots\right)$$

and by Equation 9, the right hand side would be equal to $\frac{c}{10^w - 1}$. So we would have

$$k(10^w - 1) = cp,$$

and since p does not divide k, by Euclid's Lemma 13 p divides $10^w - 1$: A contradiction with the minimality of u. □

Example 85. For $k = 2$ and $p = 7$, Fermat's little theorem tells us that 7 divides $10^6 - 1 = 999\,999$. It turns out that 7 does not divide 9, 99, 999, 9999, or 99999. In particular, the smallest u for which 7 divides $10^u - 1$, is $u = 6$. Hence the multiplicative order of 10 mod 7 is 6, and $d = \frac{999999}{7} = 142857$. Equation 11 becomes

$$\frac{2}{3} = \frac{285714}{999999} = 285714\left(\frac{1}{10^6} + \frac{1}{10^{12}} + \frac{1}{10^{18}} + \dots\right) = 0.\overline{285714}$$

So the number of decimal digits in the period is 6, which is exactly $p - 1$. If you try a different k, again with $p = 7$, you will always find decimal expansions with a 6-digit period.

Example 86. For $k = 1$ and $p = 11$, Fermat's little theorem tells us that 11 divides $10^{10} - 1 = 9\,999\,999\,999$. However, 11 divides already $10^2 - 1 = 99$. So $u = 2$ and $d = \frac{99}{11} = 9$. Equation 11 becomes

$$\frac{1}{11} = \frac{9}{99} = 9\left(\frac{1}{10^2} + \frac{1}{10^4} + \frac{1}{10^6} + \dots\right) = 0.\overline{09}$$

In this case the number of decimal digits in the period is 2, which divides $p - 1$.

Deeper thoughts 87. The primes p for which the multiplicative order of 10 mod p is exactly $(p-1)$, are called *full reptend primes*. Basically, the reciprocals $\frac{1}{p}$ of full reptend primes are the most annoying fractions to write down in decimal digits!, because they have the "longest period possible": By theorem 84, their period consists of exactly $p - 1$ digits. For example,

$$7, 17, 19, 23, 29, 47, 59, 61, 97$$

are full reptend primes, and

$$\frac{1}{7} = 0.\overline{142857}, \quad \frac{1}{17} = 0.\overline{0588235294117647}, \quad \frac{1}{19} = 0.\overline{052631578947368421}, \dots$$

Full-reptend primes seem quite frequent, but at the moment, it is not known whether they are infinitely many or not. Should there be infinitely many of them, it would mean that the decimal period of an integer can be arbitrarily large. The general belief is that reptend primes should be infinitely many; in fact, *Artin's conjecture* on primitive roots is that roughly 37.4% of primes are full reptend. In 1967 Hooley proved that Artin's conjecture is true if we believe another famous conjecture, the Generalized Riemann Hypothesis[8]. The Riemann hypothesis is one of the seven "millennium prize problems": Any correct solution to it results in a $1 million prize being awarded to the discoverer.

0.7 Exercises

0. Recalling that $\binom{n}{k} \stackrel{\text{def}}{=} \frac{n!}{k!(n-k)!}$, prove that for any integers $n \geq k \geq 1$ one has

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

1. Use induction on n to prove Newton's formula:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \, b^{n-k}.$$

   Hint: You may use exercise 0 and the following "reindexing trick":

$$\sum_{k=0}^{n} \binom{n}{k} a^{k+1} \, b^{n-k} = \sum_{k=1}^{n+1} \binom{n}{k-1} a^k \, b^{n+1-k}.$$

2. Let n be a positive integer. Prove that for any $k \in \{0, \ldots, n\}$, one has

$$\binom{n}{k} < 2^n.$$

3. Use induction to prove the *pidgeonhole principle*: any map from a set with $n+1$ objects to a set with n objects, is not injective.

4. Use induction to prove the *generalized pidgeonhole principle*: If more than kn objects are placed into n boxes, then at least one box must contain more than k objects. (The case $k = 1$ is the pidgeonhole principle.)

5. Prove that

$$1^3 + 2^3 + 3^3 + \ldots + n^3 = \frac{1}{4} n^2 (n+1)^2.$$

6. Prove that $n! > 3^n$ for n large.

7. Prove that for all $i \in \mathbb{N}$ and for all integers $n \geq 1$

$$\sum_{k=1}^{n} \binom{i+k-1}{i} = \binom{n+i}{i+1}.$$

8. Compute the $\gcd(528, 303)$ using the Euclidean algorithm.

[8] Hooley, Christopher (1967). "On Artin's conjecture". J. Reine Angew. Math. 225: 209–220.

9. Prove the "Euclidean division for $\mathbb{Z}$": Given two integers a, b, with $b \neq 0$, there exists a unique pair of integers (r, q) such that

$$a = bq + r, \text{ and } 0 \leq r < |b|.$$

10. Find all integer solutions of the equation $2x + 3y = 15$.

11. Prove a version of Bezout's theorem for $\mathbb{Z}$: Let a, b, c be integers. The equation $ax+by = c$ has integer solutions if and only if $\gcd(|a|, |b|)$ divides $|c|$.

12. Let a, b be positive integers such that $\gcd(a, b) = 1$. Let

$$\mathcal{H}_0^-(a, b) \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \text{ such that } ax - by = 0\}$$

be the set of integer solutions of $ax - by = 0$. Prove that

$$\mathcal{H}_0^-(a, b) = \{(bk, ak) \text{ such that } k \in \mathbb{Z}\}.$$

13. Find all integer solutions of $6x - 9y = 15$. Can you find all *positive* integer solutions?

14. Think of a number. Square it. Divide the result by 3. Why is the remainder always 0 or 1, but never 2? Justify your answer.

15. What is the last digit of 3^{2001}?

16. What are the last two digits of 913250946798^6?

17. Compute $12345678^{23456789} \bmod 3$.

18. Suppose that a number x can be written in the decimal representation as "*abcabc*", with a, b, c decimal digits. (For example, $x = 285285$.) Show that x is always a multiple of 13.

19. Find the remainder of the division of 11^{118} by 59. (Hint: use Fermat's little theorem.)

20. Write the number 73 on a piece of paper, fold it up, and give it to an unsuspecting friend. Ask your friend to write his/her birth year twice in a calculator. (E.g., I would write 19821982.) Then ask your friend if the number is divisible by any chance by 137; ask him/her to verify with the calculator. Then say, "please divide the result by your birthyear". Ask your friend to unwrap the paper: the calculator and the piece of paper will magically tell the same number, 73! Can you spoil the magic and explain the trick?

21. Under what conditions is a six-digit number whose decimal digits are *abcabc* divisible by $7, 9, 11$ and 13? (For example, 135135 is divisible by all of them.)

22. For any positive integers a and b, prove that $\gcd(a, b) \cdot \text{lcm}(a, b) = ab$.

1 C-Rings, Fields and Domains

A **commutative ring** or **C-ring** consists of a set A endowed with two operations $+$ and $\cdot$ that satisfy the following eight axioms:

(R0) The operations are *internal*. That is, for all x, y in A, the elements $x + y$ and $x \cdot y$ are both in A.

Additive properties:

(R1) The operation $+$ is *associative*. That is, for all x, y, z in A, $x + (y + z) = (x + y) + z$.

(R2) The operation $+$ is *commutative*. That is, for all x, y, z in A, $x + y = y + x$.

(R3) The operation $+$ has a unique *neutral element*. That is, there exists an element z in A such that for all x in A, $x + z = x$. From now on we denote such element by "0".

(R4) Every element has a unique *additive inverse*. That is, for all x in A there exists exactly one element y in A such that $x + y = 0$. From now on we denote such element by "$-x$".

Multiplicative properties:

(R5) The operation $\cdot$ is *associative*. That is, for all x, y, z in A, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

(R6) The operation $\cdot$ is *commutative*. That is, for all x, y, z in A, $x \cdot y = y \cdot x$.

Compatibility of $+$ and $\cdot$:

(R7) The operation $\cdot$ *distributes* $+$: for all x, y, z in A, $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

Notation. We write $a - b$ as a shortening of $a + (-b)$. Moreover, we usually write xy instead of $x \cdot y$. Note also that by associativity, it is not ambiguous to write $abcd$ instead of $a(b(cd))$ or of $(ab)(cd)$. In fact, no matter how you insert brackets, the result is always the same.

Remark 88. Some textbooks rephrase axiom (R3) as "The operation $+$ has a neutral element". Uniqueness is anyway necessary: Were there two distinct neutral elements z and w, we would have $z + w = z$ (because w is neutral) yet also $z + w = w$ (being z neutral), so $z = w$; a contradiction.
Similarly, some textbooks equivalently rephrase axiom (R4) as "Every element has an *additive inverse*". Also in this case, uniqueness is implicit: Were there two distinct elements y, y' in A such that $x + y = 0 = x + y'$, then we would have

$$y' = y' + 0 = y' + (x + y) = (y' + x) + y = 0 + y = y,$$

a contradiction.

Example 89. The empty set is not a commutative ring: In fact, by axiom (R4), any C-ring must contain at least one element, namely, the neutral element 0.
The set $\{0\}$, instead, *is* a C-ring. So the smallest C-ring has one element.

Example 90. Let m be a positive integer. The set $\mathbb{Z}_m = \{0, 1, \ldots, m\}$ is a C-ring, with the operations of addition and multiplication "modulo m". So for any positive integer m, there is a C-ring with exactly m elements.

Example 91. Let $\mathbb{Z}$ be the set of integers: $0, 1, -1, 2, -2, 3, -3, \ldots$ This infinite set is a C-ring with respect to the usual operations of addition and multiplication.
Instead the set $\mathbb{N}$ of natural numbers $0, 1, 2, 3, \ldots$ is not a C-ring, because (R4) fails: $\mathbb{N}$ does not contain the additive inverse of 3, for example.

Example 92. Let $2\mathbb{Z}$ be the set of EVEN integers: $0, 2, -2, 4, -4, \ldots$ This is a C-ring with respect to the usual addition and multiplication.
Instead, the set of ODD integers is not a C-ring. For example, (R3) fails: there is no neutral element with respect to addition. (Axiom $(R0)$ fails too: why?)

Example 93. The set of rational numbers

$$\mathbb{Q} \stackrel{\text{def}}{=} \left\{ \frac{a}{b} \text{ such that } a, b \in \mathbb{Z} \text{ and } b \neq 0 \right\}$$

is a C-ring with respect to the addition defined by

$$\frac{a}{b} + \frac{c}{d} \stackrel{\text{def}}{=} \frac{ad + cb}{bd}$$

and the multiplication defined by

$$\frac{a}{b} \frac{c}{d} \stackrel{\text{def}}{=} \frac{ac}{bd}.$$

Example 94. The set of real numbers $\mathbb{R}$ is a C-ring with respect to the usual addition and multiplication.

Example 95. Another example of C-ring is the set of complex numbers

$$\mathbb{C} \stackrel{\text{def}}{=} \{a + bi \text{ such that } a, b \in \mathbb{R}\}.$$

Recall that i is short for $\sqrt{-1}$, so $i^2 = -1$. Addition in $\mathbb{C}$ is defined as

$$(a + bi) + (c + di) \stackrel{\text{def}}{=} (a + c) + (b + d)i,$$

whereas the formula for multiplying is

$$(a + bi) \cdot (c + di) \stackrel{\text{def}}{=} (ac - bd) + (ad + bc)i.$$

Remark 96. In the definition of C-ring, we required only $+$ to have a neutral element, which we denoted by 0. However, the rings $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ have also a neutral element with respect to multiplication, namely, 1. In fact, $1x = x$ for all x. This leads us to the following definitions.

Definition 97. A **C-ring with 1** is a C-ring that satisfies the additional axiom

(R8) The operation $\cdot$ has a (necessarily unique[9]) *neutral element*. That is, there exists a unique element $z \neq 0$ in A such that for all x in A, $xz = x$. From now on we denote such neutral element by "1".

Example 98. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are C-ring with 1. For any positive integer m, $\mathbb{Z}_m$ is a C-ring with 1. Instead, $2\mathbb{Z}$ and $\{0\}$ are examples of C-rings "without 1".

[9] Were there two distinct neutral elements z and w, we would have $zw = z$ (because w is neutral) yet also $zw = w$ (being z neutral), so $z = w$; a contradiction.

1.1 Invertible elements and Fields

In the definition of C-ring with 1, we are requiring every element to have an additive inverse. In contrast, the presence of a multiplicative inverse is welcome, but not required. We will see very soon that there is a reason for this asymmetry, namely, that the element 0 cannot have a multiplicative inverse. However, it does make sense to study C-rings like $\mathbb{Q}$, where every element $\frac{a}{b}$ different from 0 has a multiplicative inverse, namely, $\frac{b}{a}$. These special C-rings are called "fields" and will be discussed below.

Definition 99. Let A be a C-ring with 1. An element x in A is called *invertible* if there exists an element y in A such that $xy = 1$.

For example, 1 is always invertible, because $1 \cdot 1 = 1$. Note that the definition of "invertible" makes sense only if there is an element 1 in the ring.

Proposition 100. *If it exists, the multiplicative inverse is also unique.*

Proof. Were there two distinct multiplicative inverses u and u' for a, we would have

$$u' = u'\,1 = u'(au) = (u'a)u = 1u = u. \qquad \square$$

Notation. From now on we denote the multiplicative inverse of x (whenever it exists!) by "x^{-1}". Recall that we decided to write $a-b$ as a shortening for $a+(-b)$; similarly, some authors introduce the notation $\frac{a}{b}$ as a shortening for ab^{-1}. We stress that the notation $\frac{a}{b}$ only makes sense because of the commutativity axiom (R6): Were the product not commutative, then we would need to distinguish ab^{-1} from $b^{-1}a$, so the notation "$\frac{a}{b}$" would be ambiguous. For these reasons, most textbooks prefer to use the notation ab^{-1} rather than $\frac{a}{b}$.

Definition 101. A **field** is a C-ring with 1 that satisfies another additional axiom:

(R9) Every element $a \neq 0$ of F is *invertible*.

Example 102. $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields; $\mathbb{Z}$ is not a field.

Is $\mathbb{Z}_m$ a field? It turns out that the answer depends on m.

Proposition 103. *In $\mathbb{Z}_m$ the invertible elements are those coprime with m.*

Proof. By definition of $\mathbb{Z}_m$, an element a has multiplicative inverse y if and only if $1 \equiv ay$ modulo m. This means that there is an integer x such that

$$1 - ay = mx.$$

In other words, a has a multiplicative inverse if and only if there exist two integers x, y that satisfy $1 = ay + mx$. But we proved in Theorem 40 that the equation $1 = ay + mx$ has integer solutions if and only if $\gcd(a, m) = 1$. $\square$

Corollary 104. *$\mathbb{Z}_m$ is a field $\iff$ m is a prime number.*

Proof.

"$\Leftarrow$" Let $a \neq 0$ be any element of $\mathbb{Z}_m$. Since m is prime, $\gcd(a, m) = 1$. So a is invertible by Proposition 103.

"$\Rightarrow$" By contradiction, suppose m has a divisor d such that $1 < d < m$. Then $\gcd(d, m) = d \neq 1$, so this particular d is not invertible by Proposition 103. $\square$

Corollary 105. *The smallest field is $\mathbb{Z}_2$.*

Proof. Since 2 is a prime number, $\mathbb{Z}_2$ is a field. On the other hand, a field must contain by definition at least two different elements, 0 and 1. □

Okay, we are good with definitions. Now we can see some general properties of C-rings.

Proposition 106 (Cancellation)**.** *Let A be a C-ring. Let $a, b, c \in A$. If $a + b = a + c$, then $b = c$.*

Proof. By axiom R4 we know that a has and additive inverse, that is, an element y such that $y + a = 0$. Let us add it: From $a + b = a + c$ we get that $y + (a + b) = y + (a + c)$. Applying associativity on both sides, we get that $(y + a) + b = (y + a) + c$. But $y + a = 0$, so

$$0 + b = 0 + c.$$

By definition of neutral element, the line above reads $a = c$. □

Remark 107. The Cancellation above is *with respect to* $+$. The analogous property with respect to the product does not always work. If $ab = ac$, we cannot conclude that $b = c$. (Why cannot we multiply both sides by a^{-1}?) In fact, even in $\mathbb{Z}$ there are easy counterexamples: $3 \cdot 0 = 7 \cdot 0$, but if we simplified we would get $3 = 7$, which is false. This is part of a more general phenomenon:

Proposition 108. *Let A be a C-ring. For all a in A, $a \cdot 0 = 0$.*

Proof. Being 0 neutral element, $0 = 0 + 0$, and $a \cdot 0 + 0 = a \cdot 0$. So

$$a \cdot 0 + 0 = a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0,$$

where in the last step we used distributivity. So by Cancellation, $0 = a \cdot 0$. □

Proposition 109. *Let A be a C-ring. For all a, b in A, $(-a)b = a(-b) = -(ab)$.*

Proof. Since the additive inverse is unique, to check that $(-a)b = -(ab)$ it suffices to prove that $(-a)b$ is an additive inverse of ab; that is, we need to show that $(-a)b + ab = 0$. This can be done using distributivity:

$$(-a)b + ab = (-a + a)b = 0b = 0,$$

where in the last step we applied Proposition 108. Similarly,

$$a(-b) + ab = a(-b + b) = a \cdot 0 = 0,$$

which shows that also $a(-b)$ is the additive inverse of ab. □

Proposition 110. *Let A be a C-ring. For all a, b in A, $(-a)(-b) = ab$.*

Proof. Obviously $-ab + ab = 0$. On the other hand, by Proposition 109, $-ab = (-a)b$; so

$$-ab + (-a)(-b) = (-a)b + (-a)(-b) = (-a)(b + -b) = (-a)0 = 0.$$

So both ab and $(-a)(-b)$ are the additive inverse of $-ab$. Hence, they must be equal. □

1.2 Zerodivisors and Domains

As we saw in Remark 107, cancellation with respect to addition works: If $a+b = a+c$, then $b = c$. In contrast, cancellation with respect to product works only in some C-rings. For example, it works in $\mathbb{Z}$. But in $\mathbb{Z}_{12}$ we have that $3 \cdot 1 = 3 \cdot 5$, yet we cannot simplify the 3. In this chapter, we focus on those C-rings in which cancellation with respect to product works.

Definition 111. Let A be a nonzero C-ring. An element $a \in A$ is called *zero-divisor* if there exists some element $b \neq 0$ such that $ab = 0$.

For example, in any nonzero C-ring, 0 is a zerodivisor, because we can pick any element $b \neq 0$ and by Proposition 108 we will have $0 \cdot b = 0$.

Proposition 112. *The sets of zero-divisors and of invertible elements have empty intersection.*

Proof. By contradiction, let a be an invertible zero-divisor. Let $b \neq 0$ be an element such that $ab = 0$. Multiplying by a^{-1} we obtain $b = 0$, a contradiction. □

Proposition 113. *Let A be a nonzero C-ring. The following are equivalent:*

(1) *The only zerodivisor is 0.*

(2) *For all a, b in A, if $ab = 0$, then either $a = 0$ or $b = 0$.*

(3) *For every $a \neq 0$, if $ab = ac$ then $b = c$.*

Proof.
(1) ⇒ (2). By contradiction, suppose there exist a, b in A such that $ab = 0$, but both $a \neq 0$ and $b \neq 0$. Then a and b are zerodivisors.
(2) ⇒ (3). If $ab = ac$, with $a \neq 0$, then $a(b - c) = 0$. So by the assumption, either $a = 0$ or $b - c = 0$. But $a \neq 0$, so $b - c = 0$.
(3) ⇒ (1). By contradiction, suppose there is a zerodivisor $a \neq 0$. So for some $b \neq 0$ we have $ab = 0$. So by Proposition 108, $ab = a \cdot 0$, which implies $b = 0$; a contradiction. □

Definition 114. Let A be a C-ring, $A \neq \{0\}$. If one of the three equivalent properties above is satisfied, A is called a *domain*.

Proposition 115. *$\mathbb{Z}_m$ is a domain $\iff$ m is prime.*

Proof.
"⇐" If m is a prime number and $ab = 0$ in $\mathbb{Z}_m$, then m divides ab. By Euclid's Lemma 13, m is either a factor of a or of b. In other words, either $a \equiv 0 \bmod m$ or $b \equiv 0 \bmod m$.
"⇒" By contradiction, suppose $m = ab$, for some $a, b > 1$. Then in $\mathbb{Z}_m$ we have $ab = 0$, with $a \neq 0$ and $b \neq 0$. So $\mathbb{Z}_m$ is not a domain, a contradiction. □

The previous proposition reveals that $\mathbb{Z}_m$ is a domain exactly whenever $\mathbb{Z}_m$ is a field. We will see in the next two results that this is not a coincidence.

Proposition 116. *Every field is a domain.*

Proof. If A is a field and $a \neq 0$, then a is invertible. So from $ab = ac$ we can multiply both sides by a^{-1} and get $b = c$. Via Proposition 113, this is equivalent to A being a domain. □

Alternative proof. By Proposition 112, every zero-divisor is also non-invertible. Since A is a field, the only non-invertible element is zero. Hence, the only zero-divisor is 0. □

The converse of Proposition 116 is false: $\mathbb{Z}$ and $2\mathbb{Z}$ are easy examples of **domains that are not fields**. In fact, $2\mathbb{Z}$ does not even contain 1, so *none* of its elements is invertible, because it does not even make sense to talk about invertibility... In $\mathbb{Z}$, there are two invertible elements, 1 and -1. The remaining integers $2, -2, 3, -3, \ldots$ are neither invertible nor zero-divisors. So there are more domains than fields; and there are way more non-invertible elements than zero-divisors.

However, both $\mathbb{Z}$ and $2\mathbb{Z}$ are infinite sets. If we stick to finite sets, then surprisingly there is no difference between domains and fields:

Theorem 117. *Every finite domain is a field.*

Proof. First we need to show that every finite domain A is a C-ring *with 1*. Let $0, x_1, x_2, \ldots, x_n$ be the elements of A, listed without repetitions. Consider the list

$$x_1x_1,\ x_1x_2,\ \ldots,\ x_1x_n.$$

These guys must be all distinct and nonzero, because if $x_1x_j = x_1x_k$ we would have $x_j = x_k$ (since A is a domain), and if $x_1x_j = 0$ we would have $x_j = 0$ (again, because A is a domain). Therefore, by the pidgeonhole principle (and here we use the finiteness assumption), our list $x_1x_1,\ x_1x_2,\ \ldots,\ x_1x_n$ is just a reshuffling of $x_1, \ldots, x_n$. In particular, x_1 must appear somewhere in that list; and more generally, each x_j must appear somewhere. So we have established that

(A) there exists an index e in $\{1, \ldots, n\}$ such that

$$x_1 = x_1x_e;$$

(B) more generally, for any integer j, there is an index k for which

$$x_j = x_1x_k.$$

Our goal is now to show that x_e "is the 1". In other words, we want to show that $x_ex_j = x_j$ for all j. In fact, fix j and consider x_ex_j. Using the facts (B) and (A) above, we have

$$x_j = x_1x_k = (x_1x_e)x_k = x_e(x_1x_k) = x_ex_j.$$

This holds for all j, so x_e is the desired neutral element! So 1 is in A. Next, to show that A is a field, consider an element $a \neq 0$. We need to show that a is invertible. Consider all the powers

$$a^2,\ a^3,\ a^4,\ \ldots, a^n,\ a^{n+1},\ \ldots$$

Since A is finite, these elements cannot be all distinct, so there exist integers $r < s$ such that $a^s = a^r$. In other words, $a^s - a^r = 0$. Since $r < s$ we can collect the common factor a^r, obtaining

$$a^r \cdot (a^{s-r} - 1) = 0. \tag{12}$$

But now A is a domain and $a \neq 0$. So it cannot be that $a \cdot a = 0$. So $a^2 \neq 0$. But again, if $a \neq 0$ and $a^2 \neq 0$, in a domain their product a^3 cannot be zero. So $a^3 \neq 0$. And so on. By induction, $a^r \neq 0$ for all r. So from equation 12, since A is a domain, we obtain

$$(a^{s-r} - 1) = 0.$$

Hence $a^{s-r} = 1$. So if we set $y \stackrel{\text{def}}{=} a^{s-r-1}$, we have found a y such that $ay = ya = a^{s-r} = 1$. $\square$

1.3 Nilpotent elements and reduced C-rings

The previous proof suggests us a new definition.

Definition 118. Let A be a nonzero C-ring. An element a of A is called *nilpotent* if there is a positive integer N such that $a^N = 0$.

Note that if $a^N = 0$, then also $a^{N+1} = a \cdot a^N = a \cdot 0 = 0$; so also $a^{N+2} = 0$, and so on.

Definition 119. A C-ring is called *reduced* if the only nilpotent element is 0.

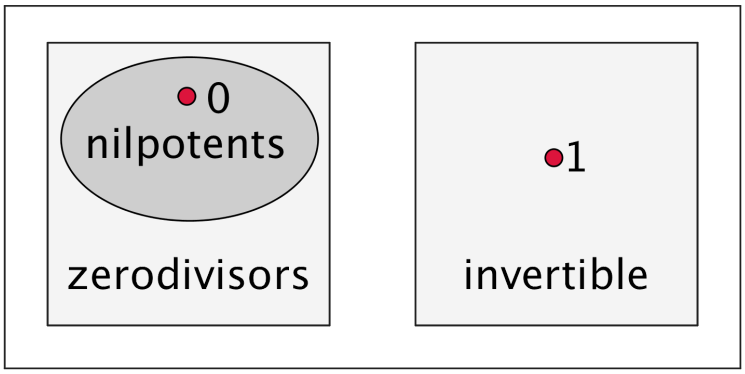
Proposition 120. *Every nilpotent element is a zero-divisor. (In particular, domains are reduced.)*

Proof. Let a be a nilpotent element and let r be the smallest positive integer for which $a^r = 0$. If $r = 1$, then $a = 0$, which is a zero-divisor. If $r \geq 2$, then $a^{r-1} \neq 0$ by definition of r (because $r - 1$ is smaller than r), and $a \cdot a^{r-1} = a^r = 0$. So either way, a is a zero-divisor. □

The converse is false: In $\mathbb{Z}_{10}$, the element 5 is a zerodivisor (because $2 \cdot 5 = 0$), but it is not nilpotent (because $5^n = 5$ for all positive integers n).

Corollary 121. *No nilpotent element is invertible. (In particular, fields are reduced.)*

Proof. Put together Propositions 112 and 120. □



1.4 *Gaussian Integers

Inside $\mathbb{C}$, consider the subset

$$\mathbb{Z}[i] \stackrel{\text{def}}{=} \{a + bi \text{ such that } a, b \in \mathbb{Z}\},$$

formed by the "complex numbers with integer coordinates". It is easy to see (compare the Exercises section) that $\mathbb{Z}[i]$ is a domain, with respect to the usual addition and multiplication of complex numbers. The neutral element of addition is of course $0 + 0i$, and the neutral element of multiplication is $1 + 0i$. In this section we discuss a natural, beautiful property of $\mathbb{Z}[i]$. We will see that this has some applications to number theory (cf. Theorem 304 and following ones).

Definition 122. The *algebraic norm* of a complex number $a + bi \in \mathbb{C}$ is the non-negative real number

$$\mathcal{N}(a + bi) \stackrel{\text{def}}{=} a^2 + b^2.$$

In case $a + bi \in \mathbb{Z}[i]$, the algebraic norm is of course a number in $\mathbb{N}$.

Remark 123. By Pythagoras' theorem, in the Gaussian plane, the algebraic norm of a point is the square of distance between the point and the origin.

Lemma 124. *The algebraic norm of the product is the product of the algebraic norms.*

Proof. For any a, b, c, d real numbers,

$$\begin{aligned}\mathcal{N}\left((a+bi)(c+di)\right) &= \mathcal{N}\left(ac-bd+(bc+ad)i\right) = (ac-bd)^2+(bc+ad)^2 = \\ &= a^2c^2 - \cancel{2abcd} + b^2d^2 + b^2c^2 + \cancel{2abcd} + a^2d^2 = \\ &= (a^2+b^2)(c^2+d^2) = \mathcal{N}(a+bi)\cdot\mathcal{N}(c+di). \quad \square\end{aligned}$$

Lemma 125. *The invertible elements in $\mathbb{Z}[i]$ are the elements of norm one; that is, ± 1 and $\pm i$.*

Proof. If $\mathcal{N}(a+bi) = 1$, then $(a+bi)(a-bi) = a^2+b^2 = \mathcal{N}(a+bi) = 1$. So the interesting implication is the converse. If $a+bi$ is invertible, then $1 = (a+bi)(c+di)$ for some c, d in $\mathbb{Z}$. Passing to the norms and using Lemma 124, we obtain

$$1 = \mathcal{N}\left((a+bi)(c+di)\right) = \mathcal{N}(a+bi)\cdot\mathcal{N}(c+di).$$

So $\mathcal{N}(a+bi)$ is an invertible element in $\mathbb{Z}$. Since it is non-negative, $\mathcal{N}(a+bi) = 1$. $\square$

Theorem 126 (Euclidean Division for $\mathbb{Z}[i]$)**.** *Let z, w be Gaussian integers, with $w \neq 0$. There exist Gaussian integers (q, r) such that*

(1) $z = qw + r$,

(2) $0 \leq \mathcal{N}(r) \leq \frac{\mathcal{N}(w)}{2} < \mathcal{N}(w)$.

However, the pair (q, r) is not necessarily unique.

Proof. Write $z = a+bi$ and $w = c+di$. Since $w \neq 0$, it is certainly invertible **inside** $\mathbb{C}$, though we don't know if the inverse of w is in $\mathbb{Z}[i]$ or not. So if we pass to $\mathbb{C}$, we have the identity

$$z\cdot w^{-1} = (a+bi)\cdot\frac{c-di}{c^2+d^2} = \frac{ac-bd}{c^2+d^2} + i\frac{bc-ad}{c^2+d^2}.$$

Now, let m be the closest integer to $\frac{ac-bd}{c^2+d^2}$. That is,

$$\left|\frac{ac-bd}{c^2+d^2} - m\right| \leq \frac{1}{2}. \tag{13}$$

Similarly, let n be the closest integer to $\frac{bc-ad}{c^2+d^2}$, so that

$$\left|\frac{bc-ad}{c^2+d^2} - n\right| \leq \frac{1}{2}. \tag{14}$$

Finally, set

$$q \stackrel{\text{def}}{=} m + ni. \tag{15}$$

From the previous equations, inside $\mathbb{C}$ we have the following identity:

$$(zw^{-1} - q) = \left(\frac{ac-bd}{c^2+d^2} + i\frac{bc-ad}{c^2+d^2}\right) - (m+ni) = \left(\frac{ac-bd}{c^2+d^2} - m\right) + i\left(\frac{bc-ad}{c^2+d^2} - n\right),$$

so from Inequalities 13 and 14 we get

$$\mathcal{N}(zw^{-1} - q) \leq \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}. \tag{16}$$

Let us define

$$r \stackrel{\text{def}}{=} z - qw.$$

This way, $z = qw + r$ is automatically satisfied; moreover, by Inequality 16 in $\mathbb{C}$ we have

$$\mathcal{N}(r) = \mathcal{N}(z - qw) = \mathcal{N}\left(w \cdot (zw^{-1} - q)\right) = \mathcal{N}(w) \cdot \mathcal{N}(zw^{-1} - q) \leq \mathcal{N}(w) \cdot \frac{1}{2}.$$

This shows existence. For a concrete example that shows how the pair (q, r) is not unique, start with the numbers $z = 2 + 3i$, $w = 1 + 2i$. Since

$$zw^{-1} = \frac{8}{5} - \frac{1}{5}i,$$

if we followed the proof above we would select the pair $(q, r) = (2, -i)$, for which indeed

$$\mathcal{N}(r) = 1 \leq \frac{\mathcal{N}(w)}{2} = 2.5.$$

However, another possible choice is the pair $(q', r') = (1, 1 + i)$, for which also

$$\mathcal{N}(r') = 2 \leq \frac{\mathcal{N}(w)}{2} = 2.5. \qquad \square$$

1.5 Exercises

1. Prove that in any C-ring A, $(a + b)^2 = a^2 + 2ab + b^2$ for all a, b. Did you use the commutativity property in your proof?

2. Let A be a C-ring. Prove that for all a, b, c in A, $a(b - c) = ab - ac$.

3. Let A be a C-ring with 1. Show that $(-1)(-1) = 1$. Show that for any a in A, $(-1)a = -a$.

4. Is $3\mathbb{Z}$, the set of integers multiple of 3, a C-ring?

5. Inside $\mathbb{C}$, consider the subset

$$\mathbb{Z}[i] \stackrel{\text{def}}{=} \{a + bi \text{ such that } a, b \in \mathbb{Z}\}.$$

   Prove that it is a C-ring and even a domain.

6. Prove that $\mathbb{Z}[i]$ is not a field.

7. Prove that the set of matrices

$$\left\{ \begin{pmatrix} a & x \\ 0 & a \end{pmatrix} \text{ such that } a, x \in \mathbb{Q} \right\}$$

   is a C-ring, with respect to the usual sum and (row-by-column) multiplication of matrices. Is it a domain? Is it a field?

8. Consider $A = \{f : \mathbb{R} \to \mathbb{R},\ f \text{ continuous }\}$. The *sum* of two functions f and g in A is defined as the function that maps x to the real number $f(x) + g(x)$. Similarly, the *product* of two functions f and g in A is the function that sends x to $f(x)g(x)$. Show that with these two operations, A is a C-ring.

9. Let X be any set. Let B be a C-ring with 1. Consider the set $A = \{f : X \to B\}$. Show that with sum and product defined as in the previous exercise, A is a C-ring with 1. Which function is the 0 of A? Which function is the 1 of A?

10. Let A be a C-ring. Suppose in A there is a nonzero nilpotent element. Is it true that in A there is a nonzero element a such that $a^2 = 0$?

11. Is there a nonzero element a in $\mathbb{Z}_{10}$ such that $a^2 = 0$? What about in $\mathbb{Z}_{20}$?

12. A positive integer m is called *squarefree* if either m is a prime or m is a product of distinct primes. Show that $\mathbb{Z}_m$ is reduced if and only if m is squarefree.

13. Let A be a C-ring. Prove that if a, b are nilpotent, $a + b$ is also nilpotent.

14. Let A be a C-ring with 1. Show that if a is nilpotent, then $1 - a$ is invertible.
Hint: what is $(1 - a)(1 + a + a^2 + \ldots + a^{N-1})$?

15. An element a in a C-ring A is called *idempotent* if $a^2 = a$. Prove that the only idempotent and nilpotent element is 0.

16. Prove that every idempotent element except 1 is a zerodivisor.

2 Polynomials

What is a monomial? What does "x" mean? When are two polynomials equal? In this section, we will try to answer these simple questions. Let us first recall a notion from calculus.

Definition 127. Let A be an arbitrary set. A *sequence in* A is a function $a : \mathbb{N} \to A$. The usual convention is to write a_i instead of $a(i)$. If 0 is an element of A, we say that a sequence is *eventually zero* if there exists an integer M such that $a_i = 0$ for all $i > M$. A common convention is to write down eventually zero sequences as finite vectors, by listing the images $a_0, a_1, \ldots, a_M$ and by forgetting the infinite sequence of zeroes that comes next.

Definition 128. Let A be a C-ring with 1. The sequences $f = (a_0, a_1, a_2, \ldots, a_n, a_{n+1}, \ldots)$ in A that are eventually zero are called *polynomials with coefficients in* A. If the a_i are not all zeroes, the *degree* of the polynomial f is

$$\deg f \stackrel{\text{def}}{=} \max\{k \text{ such that } a_k \neq 0\}.$$

The set of all polynomials with coefficients in A is denoted by $A[x]$. With the notation above, the set $A[x]$ can be written as

$$A[x] \stackrel{\text{def}}{=} \{f = (a_0, a_1, a_2, \ldots, a_n) \text{ such that } n \in \mathbb{N} \text{ and } a_i \in A\}.$$

The original ring A can be thought of as a subset of $A[x]$, as any element c of A can be identified with the "constant polynomial" $(c, 0, 0, \ldots)$ of $A[x]$. Every constant polynomial has degree zero, except for the zero polynomial, which does not have a degree.

Theorem 129. *$A[x]$ is a C-ring with 1, when endowed with the following operations:*

- *$(a_0, \ldots, a_n) + (b_0, \ldots, b_m)$ is the sequence $c_0, c_1, \ldots$ where $c_i \stackrel{\text{def}}{=} a_i + b_i$;*
- *$(a_0, \ldots, a_n) \cdot (b_0, \ldots, b_m)$ is the sequence $c_0, c_1, \ldots$ where $c_i \stackrel{\text{def}}{=} \sum_{k=0}^{i} a_k \cdot b_{i-k}$.*

The neutral element with respect to the sum is the sequence $(0, 0, 0, \ldots...)$; the neutral element with respect to the product is the sequence $(1, 0, 0, \ldots)$.

The proof is left as exercise. In view of the theorem, we adopt the notation $0 \stackrel{\text{def}}{=} (0, 0, \ldots)$ and $1 \stackrel{\text{def}}{=} (1, 0, 0, \ldots)$.

We are now ready to explain who "x" is:

Definition 130. We call x the sequence $(0, 1, 0, \ldots...)$. So for example

$$x^2 = x \cdot x = (0, 1, 0, 0 \ldots) \cdot (0, 1, 0, 0 \ldots) = (0, 0, 1, 0).$$

Similarly,

$$x^3 = x \cdot x^2 = (0, 1, 0, 0) \cdot (0, 0, 1, 0) = (0, 0, 0, 1).$$

By induction, one can verify that x^n is the sequence $c_0, c_1, \ldots$, where $c_i = 1$ if $i = n$ and $c_i = 0$ otherwise.

Remark 131. Note that

$$(a_0, a_1) = (a_0, 0) + (0, a_1) = a_0 \cdot (1, 0) + a_1 \cdot (0, 1) = a_0 \cdot 1 + a_1 \cdot x.$$

Similarly,

$$(a_0, a_1, a_2) = (a_0, 0, 0) + (0, a_1, 0) + (0, 0, a_2) = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2.$$

More generally,

$$(a_0, \ldots a_n) = \sum_{k=0}^{n} a_k \cdot x^k.$$

Notation. We can now write down polynomials the way you are used to. In fact, in view of the identity above, we will write down the polynomial $(a_0, \dots a_n)$ as

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n.$$

Definition 132. Let f be a polynomial in $A[x]$. Say $f = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$ with the notation above. Let $b \in A$. We call *evaluation of f at b* the element

$$f(b) \stackrel{\text{def}}{=} a_0 + a_1 b + a_2 b^2 + \ldots + a_n b^n.$$

Notationally, you can think of $f(b)$ as the result of "plugging in b for x".

This way, every polynomial f in $A[x]$ naturally induces a function from A to A; namely, the function $\tilde{f}$ that sends b to $f(b)$.

Remark 133. With our definition, two polynomials are equal if they have the same coefficients in the same positions. For example, the two polynomials of $\mathbb{Z}_3[x]$

$$f = x + 2 \qquad \text{and} \qquad g = x^3 + 2$$

are different: the polynomial f corresponds to the sequence (2,1), whereas g corresponds to the sequence (2,0,0,1). However, the two induced functions

$$\begin{array}{rcl} \tilde{f} : \mathbb{Z}_3 & \to & \mathbb{Z}_3 \\ b & \mapsto & b + 2 \end{array} \qquad \text{and} \qquad \begin{array}{rcl} \tilde{g} : \mathbb{Z}_3 & \to & \mathbb{Z}_3 \\ b & \mapsto & b^3 + 2. \end{array}$$

are *equal*, because they yield same outputs if they are given same inputs! In fact, by Fermat's Little Theorem (Theorem 80), one has $b^3 \equiv b \mod 3$ for all $b \in \mathbb{N}$.

2.1 Degree of a polynomial

Definition 134. The degree of a polynomial f is the maximum index k such that the coefficient of x^k is not zero. If the degree of f is n, we will often refer to $a_n x^n$ as the *leading term* of f, and to a_n as the *leading coefficient* of f. A polynomial is called *monic* if its leading coefficient is 1.

Let us see how the degree behaves with respect of sum and product.

Lemma 135. *Let A be a C-ring with 1. Let f, g be nonzero polynomials in $A[x]$. Then either $f + g = 0$, or*

$$\deg(f + g) \leq \max\{\deg f, \deg g\}.$$

If in addition $\deg f \neq \deg g$, then

$$\deg(f + g) = \max\{\deg f, \deg g\}.$$

Proof. Write $f = (a_0, \dots, a_n)$ with $a_n \neq 0$, and write $g = (b_0, \dots, b_m)$ with $b_m \neq 0$. Then $n = \deg f$ and $m = \deg g$. Now:

- If $n < m$, then $f + g$ is the polynomial $(a_0 + b_0, \dots, a_n + b_n, b_{n+1}, \dots, b_m)$, of degree m.
- If $n > m$, then $f + g$ is the polynomial $(a_0 + b_0, \dots, a_m + b_m, a_{m+1}, \dots, a_n)$, of degree n.
- If $n = m$, then $f + g$ is the polynomial $(a_0 + b_0, \dots, a_m + b_m)$. In this case we cannot be sure that the degree is n, because it could be that $a_m = -b_m$, so that $a_m + b_m = 0$ and the degree is then smaller than m. For this reason, we only claim $\deg(f + g) \leq \max\{\deg f, \deg g\}$. □

Lemma 136. *Let A be a C-ring with 1. Let f, g be non-zero polynomials in $A[x]$. Then either $fg = 0$, or*

$$\deg(f \cdot g) \leq \deg f + \deg g.$$

If in addition A is a domain, then

$$\deg(f \cdot g) = \deg f + \deg g.$$

Proof. As in the previous proof, write $f = (a_0, \ldots, a_n)$ with $a_n \neq 0$, and write $g = (b_0, \ldots, b_m)$ with $b_m \neq 0$. Then

$$f \cdot g = (a_0 b_0,\ a_0 b_1 + a_1 b_0,\ \ldots\ ,\ a_n b_m).$$

If A is a domain, from $a_n \neq 0$ and $b_m \neq 0$ it follows that $a_n \cdot b_m \neq 0$, whence $\deg(f \cdot g) = n + m$. If instead A is not a domain, it could be that $a_n \cdot b_m = 0$, in which case the degree is lower, or it could even be that $fg = 0$, in which case fg does not have a degree. □

Example 137. In $\mathbb{Z}_4[x]$, consider the degree-five polynomial $f = 2x^5 + 1$. Then

$$f \cdot f = 4x^{10} + 4x^5 + 1 = 0x^{10} + 0x^5 + 1 = 1.$$

So $\deg(f \cdot f) = 0$, which is lower than $\deg f + \deg f$, and in fact it is even lower than $\deg f$! Of course, this happened because $\mathbb{Z}_4$ is not a domain, and inside such C-ring we have $2 \cdot 2 = 0$.

The proof of the next Lemma is left as exercise:

Lemma 138. *Let A be a C-ring with 1. Let f, g be polynomials in $A[x]$. Suppose that $f \neq 0$ and that the leading coefficient of g is not a zero-divisor. Then $\deg(f \cdot g) = \deg f + \deg g$. In particular, $\deg g \leq \deg(f \cdot g)$.*

Theorem 139. *Let A be a C-ring with 1. Then the following are equivalent:*

(1) *A is a domain.*

(2) *$A[x]$ is a domain.*

Moreover, if A is a domain, then $\{$invertible elements of $A[x]\} = \{$invertible elements of $A\}$.

Proof.

(1) ⇒ (2). To show $A[x]$ is a domain, take polynomials $f, g \neq 0$. Since $f \neq 0$ it will have a leading term $a_n x^n$ with $a_n \neq 0$. Similarly $g \neq 0$ will have a leading term $b_m x^m$ with $b_m \neq 0$. But then the product $f \cdot g$ cannot be zero, because it has leading term $a_n \cdot b_m$, which is not zero because of the assumption that A is a domain.

(2) ⇒ (1). Note that $A \subseteq A[x]$, because we can view the elements of A as the polynomials of $A[x]$ of degree 0. Now let $f, g \neq 0$ in A. Were $fg = 0$ in A, then we would have $fg = 0$ in $A[x]$, a contradiction. So A is a domain.

Last claim, "$\supset$": If $ab = 1$ in A, then since $A \subseteq A[x]$, the equality $ab = 1$ also holds in $A[x]$.

Last claim, "$\subseteq$": Suppose $fg = 1$ in $A[x]$. Since A is a domain, by Lemma 136 we have that $\deg f + \deg g = \deg(fg) = \deg 1 = 0$. This means that $\deg f = \deg g = 0$, so both f, g are constant polynomials; in other words, both f and g are in A. □

Remark 140. For the last claim of the previous theorem, it is important to assume that A is a domain. Let us have another look at the polynomial $f = 2x^5 + 1$ of Example 137: It is invertible in $\mathbb{Z}_4[x]$, because $f \cdot f = 1$. However, f does not have degree zero.

Remark 141. One may wonder if the second part of Theorem 139 is valid only for domains. The answer is negative. We will see in Corollary 280 that if m is a product of distinct primes, then in $\mathbb{Z}_m[x]$ (which is not a domain!) the invertible elements all have degree zero.

2.2 Euclidean division

Sometimes one uses the adjective "monic" to indicate the polynomials that have leading coefficient 1. One of the most important aspects of monic polynomials, is that we can always divide them by one another, as the next Theorem explains.

Theorem 142. *Let A be a C-ring with 1. Let f, g in $A[x]$ be polynomials. Suppose that the leading coefficient of g is invertible. Then, there exist q, r in $A[x]$ such that:*

I.) $f = q \cdot g + r$;

II.) *either* $r = 0$, *or* $\deg r < \deg g$.

Moreover, the pair (q, r) is uniquely determined by the pair (f, g).

Proof. Existence. There are three cases: either $\deg f \geq \deg g$, or $\deg f < \deg g$, or $\deg f$ is undefined (because $f = 0$). The second and third case are obvious: Set $q \stackrel{\text{def}}{=} 0$, $r \stackrel{\text{def}}{=} f$ and we are done. So let us assume from now on that $\deg f \geq \deg g$. Also, if $\deg g = 0$, then g is constant, so the leading coefficient of g is g itself. By the assumption, g is invertible. The claim is then easy to show: set $q \stackrel{\text{def}}{=} f \cdot g^{-1}$ and $r \stackrel{\text{def}}{=} 0$.

Hence, we can assume from now on that $\deg f \geq \deg g \geq 1$. We proceed by strong induction on the degree of f. Set $n \stackrel{\text{def}}{=} \deg f$, $m \stackrel{\text{def}}{=} \deg g$ and write

$$f = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \quad \text{and} \quad g = b_0 + b_1 x + b_2 x^2 + \ldots + b_m x^m.$$

Since we know that b_m is invertible, consider

$$f' \stackrel{\text{def}}{=} f - a_n \cdot (b_m)^{-1} \cdot x^{n-m} \cdot g.$$

This f' is the difference of two polynomials of same degree n. The second polynomial is constructed in such a way that its leading coefficient is a_n, the same as that of f. As a result, when we take the difference, the two leading coefficients "cancel out". So $\deg f' < n$. So by strong induction, the theorem holds for the pair (f', g): Namely, there exist polynomials q', r' in in $A[x]$ such that:

I.) $f' = q' \cdot g + r'$;

II.) either $r' = 0$, or $\deg r' < \deg g$.

It follows that

$$\begin{aligned} f &= f' + a_n \cdot (b_m)^{-1} \cdot x^{n-m} \cdot g = (q' \cdot g + r') + a_n \cdot (b_m)^{-1} \cdot x^{n-m} \cdot g = \\ &= (q' + a_n \cdot (b_m)^{-1} \cdot x^{n-m}) \cdot g + r'. \end{aligned}$$

If we set $r \stackrel{\text{def}}{=} r'$, $q \stackrel{\text{def}}{=} (q' + a_n \cdot (b_m)^{-1} \cdot x^{n-m})$, we obtain that

$$f = qg + r' \text{ with either } r' = 0 \text{ or } \deg r' < \deg g.$$

Uniqueness. Suppose that

$$\begin{aligned} f &= q_1 \cdot g + r_1 && \text{with either } r_1 = 0 \text{ or } \deg r_1 < \deg g, \text{ and also} \\ f &= q_2 \cdot g + r_2, && \text{with either } r_2 = 0 \text{ or } \deg r_2 < \deg g. \end{aligned}$$

We want to show $q_1 = q_2$ and $r_1 = r_2$. From the lines above we get $q_1 \cdot g + r_1 = q_2 \cdot g + r_2$, so

$$(q_1 - q_2) \cdot g = r_2 - r_1.$$

If $q_1 - q_2 = 0$, then $r_2 - r_1 = 0$, and we are done. So suppose $q_1 - q_2 \neq 0$. We will show that this leads to a contradiction. First of all, the leading coefficient of g is invertible, so it is not a zero-divisor (Proposition 112). Hence if $q_1 - q_2 \neq 0$ we can apply Lemma 138 and conclude

$$\deg g \;\leq\; \deg\left((q_1 - q_2) \cdot g\right) \;=\; \deg(r_2 - r_1).$$

This contradicts Lemma 135, because each r_i is either 0, or of degree smaller than $\deg g$. $\square$

But how does one concretely perform one such division? The right algorithm is simply derived from the previous theorem, as follows. Suppose you have to divide f by g.

1. Initialize $q \stackrel{\text{def}}{=} 0$.
2. If the leading term of g is not invertible, return an error message and stop.
3. If $\deg g = 0$, output $(f \cdot g^{-1}, 0)$ and stop.
4. If $\deg f < \deg g$, output $(0, f)$ and stop.
5. While $\deg f \geq \deg g \geq 1$:
   - divide the leading term of f by the leading term of g. Let m_1 be the resulting monomial. Add m_1 to q.
   - Write $m_1 g$ under f, and perform the subtraction $f - m_1 g$. By construction, this kills the leading term of f. Replace f by the lower-degree polynomial $f - m_1 g$, and go back to step 4.
6. Output $(q, f - qg)$.

Below is an example of how I graphically compute the division in $\mathbb{Q}[x]$ of $f = 12x^3 + 4x^2 - 6$ by $g = 3x - 2$. If you are used to a different graphical representation, that is also fine, of course. The remainder I obtain in the end is $r = -\frac{2}{3}$ (last line to the left); the quotient is $q = 4x^2 + 4x + \frac{8}{3}$ (last line to the right). Every new iteration of the algorithm produces a new horizontal bar on the left, and a new monomial (of smaller and smaller degree) composing q on the right. I like this notation because it is consistent with the way I was taught to graphically represent long divisions of integers, when I was in elementary school. Obviously, if you were taught long division of integers in a different way, I encourage you to keep track of the long division of polynomials in a way analogous to that one.

$12x^3 + 4x^2 + 0x - 6$	$\overline{3x} - 2$
$12x^3 - 8x^2$	$4x^2 + 4x + {}^{8}/_{3}$
$// + 12x^2 + 0x$	
$12x^2 - 8x$	
$// + \overline{8x} - 6$	
$8x - {}^{16}/_{3}$	
$// - {}^{2}/_{3}$	

The Euclidean division has several spectacular consequences. Let us start with one.

Definition 143. Let A be a C-ring with 1. Let f be a nonzero polynomial in $A[x]$. An element a of A is called a *root* of f if $f(a) = 0$. (That is, if "plugging in $x = a$ we get an expression that is equal to zero".)

For example, 3 is a root of $x^2 - 5x + 6$, because $3^2 - 5 \cdot 3 + 6 = 0$.

Theorem 144 (Ruffini)**.** *Let A be a C-ring with 1. Let f be a polynomial in $A[x]$. Then for any $a \in A$, we can write*

$$f = (x - a) \cdot q + f(a).$$

In particular,

$$a \text{ is a root of some polynomial } g \iff (x - a) \text{ divides } g.$$

Proof. The leading coefficient of $x - a$ is 1, so we can apply the Euclidean division to f and $g = x - a$: there exist polynomials q, r such that

$$f = (x - a) \cdot q + r$$

and either $r = 0$, or $\deg r < \deg(x - a) = 1$. In both cases, r must be a constant. In particular, r remains unchanged if we plug in $x = a$. So let's plug in $x = a$: we get

$$f(a) = 0 \cdot q(a) + r = r.$$

So the remainder of the Euclidean division of f by $x - a$ is exactly $f(a)$. In particular, $f(a) = 0$ if and only if f is a multiple of $x - a$. □

Theorem 145. *Let A be a domain. If $a_1, \dots a_n$ are distinct roots of some nonzero polynomial $f \in A[x]$, then $\deg f \geq n$ and*

$$f = g(x - a_1)(x - a_2) \cdot (x - a_n)$$

for some nonzero polynomial $g \in A[x]$ of degree $\deg f - n$.

Proof. By induction on n. The case $n = 1$ is given by Ruffini's theorem 144: if a_1 is a root of f, then $f = q \cdot (x - a_1)$, with $q \neq 0$ (otherwise $f = 0$). So by Lemma 136 we have $\deg f = \deg q + 1 \geq 1$. Setting $g \stackrel{\text{def}}{=} q$ we are done.

Now suppose we have already proven the theorem for every nonzero polynomial with $n - 1$ distinct roots. Let f be a polynomial with n distinct roots, $a_1, \dots, a_n$. By Ruffini's theorem applied to a_n, we have

$$f = q \cdot (x - a_n),$$

and since we are in a domain, $\deg f = \deg q + 1$. Now, if we plug in $x = a_1$, which is a root of f, we get

$$0 = q(a_1) \cdot (a_1 - a_n).$$

But by assumptions $a_1 - a_n \neq 0$ and A is a domain: hence, $q(a_1) = 0$. The same applies also to $a_2, a_3, \dots, a_{n-1}$: we get

$$0 = f(a_i) = q(a_i) \cdot (a_i - a_n),$$

which implies $q(a_i) = 0$. In conclusion, q has $n - 1$ distinct roots. By the inductive assumption, $\deg q \geq n - 1$ and

$$q = g(x - a_1) \cdots (x - a_{n-1})g$$

for some $g \in A[x]$. But then $\deg f = \deg q + 1 \geq (n - 1) + 1 = n$ and

$$f = q(x - a_n) = g(x - a_1) \cdots (x - a_{n-1})(x - a_n).$$ □

Non-Example 146. Consider the polynomial $x^2 - 4$ in $\mathbb{Z}_{12}$. It has degree two, but four different roots: $2, 4, 8, 10$. Note that $(x - 2)(x - 10) = x^2 - 4$, but also $(x - 4)(x - 8) = x^2 - 4$.

The period in $\mathbb{Z}_m$ and two converses of Fermat

We will now provide two "partial converses" of Fermat's Little Theorem.

Definition 147. For any element a of $\mathbb{Z}_m$, let us call *period* of a (and denote by $\pi(a)$) the smallest positive integer k for which $a^k = 1$ in $\mathbb{Z}_m$. If there is no integer k for which $a^k = 1$ in $\mathbb{Z}_m$, we set $\pi(a) = \infty$.

For example, the period of 3 in $\mathbb{Z}_6$ is infinite. If instead m is a prime number, then in view of Fermat's little theorem, only 0 in $\mathbb{Z}_m$ has infinite period, because for any $b \neq 0$ we know that $b^{m-1} = 1$. This does not mean, however, that the period of every $b \neq 0$ in $\mathbb{Z}_m$ is $m - 1$. For example, in $\mathbb{Z}_7$ we have $2 \neq 0$, $2^2 \neq 0$, and $2^3 = 0$; so $\pi(2) = 3$. Note that every nonzero element in $\mathbb{Z}_7$ has period $1, 2, 3$ or 6.

Lemma 148. *Let $s, m \in \mathbb{N}, m \geq 2$. Then $a^s = 1$ in $\mathbb{Z}_m \iff \pi(a)$ divides s.*

Proof. In $\mathbb{N}$, write the Euclidean division

$$s = q \cdot \pi(a) + r.$$

Using the two sides of the expression above as exponents, we have that in $\mathbb{Z}_m$

$$a^s = \left(a^{\pi(a)}\right)^q \cdot a^r = 1 \cdot a^r.$$

So if $\pi(a)$ divides s, clearly $r = 0$ and $a^s = a^0 = 1$. Conversely, if $a^s = 1$, then $a^r = 1$: and were $r > 0$ we would get a contradiction, since $\pi(a)$ was supposed to be the smallest positive integer k for which $a^k = 1$. □

Remark 149. In view of Lemma 148, Fermat's little theorem can be rephrased as follows:
"if m is prime, for every $a \neq 0$ in $\mathbb{Z}_m$ one has that $\pi(a)$ divides $m - 1$".

Lemma 150. *Let $m \in \mathbb{N}, m \geq 2$. Let a, b be in $\mathbb{Z}_m$. If $\gcd(\pi(a), \pi(b)) = 1$, then*

$$\pi(ab) = \pi(a) \cdot \pi(b).$$

Proof. Let $c = \pi(a)$ and $d = \pi(b)$. Since

$$(ab)^{cd} = a^{cd} \cdot b^{cd} = (a^c)^d \cdot (b^d)^c = 1^d \cdot 1^c = 1,$$

by Lemma 148, $\pi(ab)$ must divide cd. Note that any prime dividing c does not divide d, and the other way around, since $\gcd(c, d) = 1$. So write

$$c = p_1^{e_1} \cdots p_k^{e_k} \text{ and } d = p_{k+1}^{e_{k+1}} \cdots p_h^{e_h},$$

where the p_i's are distinct primes (not necessarily in increasing order). If $\pi(ab)$ divides cd, then by the Fundamental Theorem of Arithmetics

$$\pi(ab) = p_1^{f_1} \cdots p_k^{f_k} p_{k+1}^{f_{k+1}} \cdots p_h^{f_h}, \text{ with } f_i \leq e_i \text{ for all } i.$$

So if $\pi(ab)$ is strictly smaller than cd, some f_i is strictly smaller than the corresponding e_i. Up to swapping the labels of c and d, we can assume that this $f_i < e_i$ happens for $i = 1$. This means that

$$\pi(ab) \text{ divides } p_1^{e_1 - 1} p_2^{e_2} \cdots p_k^{e_k} p_{k+1}^{e_{k+1}} \cdots p_h^{e_h} = \frac{c}{p_1} d.$$

But then

$$1 = (ab)^{\frac{c}{p_1} d} = a^{\frac{c}{p_1} d} \cdot b^{\frac{c}{p_1} d} = a^{\frac{c}{p_1} d} (b^d)^{\frac{c}{p_1}} = a^{\frac{c}{p_1} d}.$$

Hence, by Lemma 148, the period of a must divide $\frac{c}{p_1} d$. A contradiction, p_1 appears with larger exponent in $c = \pi(a)$ than in $\frac{c}{p_1} d$. □

Lemma 151. *Let A be any C-ring with 1. Let t, n be two positive integers. If t divides n, then the polynomial $x^t - 1$ divides the polynomial $x^n - 1$.*

Proof. The trick is to introduce a new variable $y \stackrel{\text{def}}{=} x^t$. Then for any $\ell \in \mathbb{N}$ it is clear by Ruffini's theorem that the polynomial $y - 1$ divides the polynomial $y^\ell - 1$ in $A[y]$. In particular, this is true for $\ell = \frac{n}{t}$; so there exists a polynomial $G \in A[y]$ such that

$$y^\ell - 1 = (y - 1)G.$$

Now plug back in x^t for y: we obtain that there exists a polynomial $g \in A[x]$, of degree $t \deg G$, such that

$$x^n - 1 = (x^t - 1)g.$$

□

Theorem 152 (Euler–Gauss, Lucas–Lehmer)**.** *For any integer $m \geq 2$, the following are equivalent:*

(1) *m is prime.*

(2) *Some $a \in \mathbb{Z}_m$ with $\gcd(a, m) = 1$ has period equal to $m - 1$.*

Proof. The theorem is obvious for $m = 2$, since the period of 1 is 1. So let's assume $m \geq 3$.

(1) $\Rightarrow$ (2). (This direction is called "Euler–Gauss' theorem" or "**cyclicity of the multiplicative group of $\mathbb{Z}_m$ for m prime**".) By Fermat's Little Theorem, all nonzero elements of $\mathbb{Z}_m$ are roots of the polynomial $x^{m-1} - 1$. Since $\mathbb{Z}_m$ is a domain (because m is prime), by Theorem 145 this polynomial factors as

$$x^{m-1} - 1 = (x - 1)(x - 2) \cdots (x - m + 1)$$

and every factor of $x^{m-1} - 1$ has as many distinct roots as its degree. Now look at the prime decomposition of $m - 1$,

$$m - 1 = p_1^{e_1} \cdots p_k^{e_k}.$$

Fix an $i \in \{1, \ldots, k\}$, and let

$$t_i \stackrel{\text{def}}{=} (p_i)^{e_i}, \quad s_i \stackrel{\text{def}}{=} (p_i)^{e_i - 1} = \frac{t_i}{p_i}.$$

Since t_i divides $m - 1$, by Lemma 151 $x^{t_i} - 1$ divides $x^{m-1} - 1$, so any root of $x^{t_i} - 1$ is also a root of $x^{m-1} - 1$. This implies that $x^{t_i} - 1$ is a polynomial whose roots are all distinct. We claim that $x^{t_i} - 1$ has exactly t_i roots (that is, as many as its degree). In fact, suppose it had k roots $b_1, \ldots, b_k$, with $k < t_i$. Then by Theorem 145

$$x^{t_i} - 1 = (x - b_1) \cdots (x - b_k) g \tag{17}$$

for some polynomial g of degree $t_i - k$ that has no root. At the same time

$$x^{m-1} - 1 = h(x^{t_i} - 1) = gh(x - b_1) \cdots (x - b_k) \tag{18}$$

for some polynomial $h \in \mathbb{Z}_m[x]$ of degree $m - 1 - t_i$. Reasoning as in the proof of Theorem 145, all the $m - 1 - k$ roots of $x^{m-1} - 1$ that are *not* in the set $\{b_1, \ldots, b_k\}$ would then be distinct roots of gh; and since g has no roots, they would all be roots of h. But they are too many:

$$m - 1 - k > m - 1 - t_i = \deg h,$$

a contradiction. So the claim is proven, and in $\mathbb{Z}_m$ there are t_i distinct solutions for the equation $x^{t_i} = 1$. On the other hand, if we set $s_i \stackrel{\text{def}}{=} p_i^{e_i - 1}$, then the polynomial $x^{s_i} - 1$ has at most s_i roots by Theorem 145, and all of them are also roots of $x^{t_i} - 1$, which is a multiple of $x^{s_i} - 1$ by Lemma 151. Since $s_i < t_i$, for sure there will be an element a_i of $\mathbb{Z}_m$ that satisfies simultaneously $a_i^{t_i} = 1$ and $a_i^{s_i} \neq 1$.

We claim that the period of a_i is exactly t_i. In fact, were $\pi(a_i) < t_i$, then by Lemma 148 $\pi(a_i)$ would divide t_i. But since $t_i = (p_i)^{e_i}$, this means that $\pi(a_i)$ divides $s_i = (p_i)^{e_i - 1}$. Which implies that $a_i^{s_i} = 1$, a contradiction with how we chose a_i. So our second claim is proven and we are ready to conclude: if $m - 1 = p_1^{e_1} \cdots p_k^{e_k}$ is the prime decomposition of $m - 1$, for each prime factor p_i we can find an element $a_i \in \mathbb{Z}_m$ such that $\pi(a_i) = p_i^{e_i}$. Because $p_i^{e_i}$ and $p_j^{e_j}$ are coprime if $i \neq j$, via Lemma 150 that the period of the product $a_1 \cdots a_s$ is the product of the periods of the a_i's. In other words, the element $a \stackrel{\text{def}}{=} a_1 a_2 \cdots a_s$ has period exactly $m - 1$. Since m is prime, obviously $\gcd(a, m) = 1$.

(2) $\Rightarrow$ (1). (This direction is known as "Lucas–Lehmer's criterion".) View a as an element of $\mathbb{N}$, and consider the subset

$$S \stackrel{\text{def}}{=} \{1 = a^0, a^1, a^2, \ldots, a^{m-2}\} \subseteq \mathbb{N}.$$

Let r_i be the remainder of the Euclidean division of a^i by m. By definition of "period", we know that $r_i \neq 0$ for all i. Since $\gcd(a, m) = 1$, and the prime factors of a^i are the same of a, then also $\gcd(a^i, m) = 1$; and thus $\gcd(r_i, m) = 1$. We claim that the elements

$$\{1 = r_0, r_1, \ldots, r_{m-2}\} \subseteq \{1, 2, \ldots, m-1\}$$

are all distinct. In fact, suppose by contradiction that $r_i = r_j$ for some $0 \leq i < j \leq m-1$. This means that

$$a^i - q_i m = a^j - q_j m,$$

which can be rewritten as

$$m(q_j - q_i) = a^i(a^{j-i} - 1).$$

But m and a^i are coprime. So in $\mathbb{Z}$, by the Unique Factorization Theorem, m must divide $(a^{j-i} - 1)$. So in $\mathbb{Z}_m$ we have $a^{j-i} = 1$; a contradiction, because $j - i \leq j \leq m-2$, whereas the period of a is by assumption equal to $m-1$. So the claim is proven. In particular, by the pigeonhole principle,

$$\{1 = r_0, r_1, \ldots, r_{m-2}\} = \{1, 2, \ldots, m-1\}.$$

Now let d be any proper divisor of m, so that obviously $\gcd(d, m) = d$. Since d belongs to the right hand side above, there will be an $i \in \{0, \ldots, m-2\}$ such that $r_i = d$. That is, the remainder of dividing a^i by m is d. But then we have

$$d = \gcd(d, m) = \gcd(r_i, m) = \gcd(a^i, m) = 1. \qquad \square$$

Corollary 153. *Let p be a prime number. In $\mathbb{Z}_p$, the product of two non-squares is a square.*

Proof. By theorem 152, there exists a (nonzero) element $a \in \mathbb{Z}_p$ such that every nonzero element b of Z_p can be written as $b = a^s$, for a suitable non-negative integer s. Clearly, s is even if and only if b is a perfect square. But if $b = a^s$ and $c = a^t$ with s, t both odd, their product is $bc = a^{s+t}$, with $s + t$ even. $\square$

Proposition 154 (Agrawal – Kayal – Saxena, 2002)**.** *Let a, m be positive integers, $m \geq 2$, with $\gcd(a, m) = 1$. Then*

$$m \text{ is prime} \iff \text{in } \mathbb{Z}_m[x], \quad (x + a)^m = x^m + a.$$

Proof. See the Exercises. $\square$

2.3 Complex conjugation

There is a special operation on polynomials of $\mathbb{R}[x]$ and $\mathbb{C}[x]$ that is worth studying.

Definition 155. The *conjugate* of a complex number $z = a + ia'$, with $a, a' \in \mathbb{R}$, is $\overline{z} \stackrel{\text{def}}{=} a - ia'$. More generally, the *conjugate* of a polynomial

$$f = z_0 + z_1 x + \ldots + z_n x^n \in \mathbb{C}[x]$$

is the polynomial

$$\overline{f} \stackrel{\text{def}}{=} \overline{z_0} + \overline{z_1} x + \ldots + \overline{z_n} x^n \in \mathbb{C}[x].$$

Lemma 156. *Let $f \in \mathbb{C}[x]$.*

(1) $\overline{\overline{f}} = f$;

(2) $\overline{f} = f$ *if and only if* $f \in \mathbb{R}[x]$;

(3) $\overline{f+g} = \overline{f} + \overline{g}$;

(4) $\overline{f \cdot g} = \overline{f} \cdot \overline{g}$;

(5) $f \cdot \overline{f} \in \mathbb{R}[x]$.

Proof. The first two items are obvious. The fifth follows from the fourth and the second, since

$$\overline{f \cdot \overline{f}} = \overline{f} \cdot \overline{\overline{f}} = \overline{f} \cdot f = f \cdot \overline{f}.$$

Thus it suffices to prove the third and fourth item. Suppose

$$f = \sum_{j=1}^{n} (a_j + ia'_j)x^j \quad \text{and } g = \sum_{j=1}^{m} (b_j + ib'_j)x^j,$$

where the a_j, a'_j, b_j, b'_j are real numbers. Then it is easy to check that

$$\overline{f+g} = \overline{f} + \overline{g} = \sum_j (c_j - ic'_j)x^j,$$

where each $c_j = a_j + b_j$ and each $c'_j = a'_j + b'_j$. Moreover,

$$f \cdot g = \sum_j (d_j + id'_j)x^j,$$

where each $d_j = \sum_{k=0}^{j}(a_k b_{j-k} - a'_k b'_{j-k})$ and each $d'_j = \sum_{k=0}^{j} a'_k b_{j-k} + a_k b'_{j-k}$. Thus

$$\overline{f \cdot g} = \sum_j (d_j - id'_j)x^j.$$

On the other hand,

$$\overline{f}\,\overline{g} = \sum_j z_j x^j,$$

where

$$\begin{aligned} z_j &= \textstyle\sum_{k=0}^{j}(a_k - ia'_k) \cdot (b_{j-k} - ib'_{j-k}) = \sum_{k=0}^{j}(a_k b_{j-k} - a'_k b'_{j-k}) - i(a'_k b_{j-k} + a_k b'_{j-k}) = \\ &= \textstyle\sum_{k=0}^{j}(a_k b_{j-k} - a'_k b'_{j-k}) - i\sum_{k=0}^{j}(a_k b_{j-k} - a'_k b'_{j-k}) \overset{\text{def}}{=} d_j - id'_j. \quad \square \end{aligned}$$

Proposition 157 (Complex Conjugate root theorem)**.** *If $\alpha \in \mathbb{C}$ is the root of a polynomial $f \in \mathbb{C}[x]$, then $\overline{\alpha}$ is a root of $\overline{f}$.*

Proof. Suppose $f = z_0 + z_1 x + \ldots + z_n x^n$. Since α is a root,

$$0 = z_0 + z_1\alpha + \ldots + z_n\alpha^n.$$

So if we pass to the conjugate and repeatedly apply Lemma 156, we get that

$$\overline{0} = \overline{z_0 + z_1\alpha + \ldots + z_n\alpha^n} = \overline{z_0} + \overline{z_1} \cdot \overline{\alpha} + \ldots + \overline{z_n} \cdot \overline{\alpha}^n = \overline{f}(\overline{\alpha}).$$

But 0 is real, so $\overline{0} = 0$. $\square$

Corollary 158. *Let $f \in \mathbb{R}[x]$ and let $\alpha \in \mathbb{C}$. Then α is a root of f if and only if $\overline{\alpha}$ is.*

Proof. Straightforward from Proposition 157 and the fact that $\overline{f} = f$, because $f \in \mathbb{R}[x]$. □

Example 159. Since i is a root of the polynomial $x^3 + x \in \mathbb{R}[x]$, then also its conjugate $-i$ must be.

2.4 Symmetric Polynomials

We have shown how to construct polynomial in one variable. But how to construct polynomials in several variables, like $2xy+z^2$? The answer we are going to give in this course is: by induction. If A is a C-ring with 1, so is $B = A[x]$. Hence it makes sense to start out with B and create a new polynomial ring out of it, with a new variable y:

$$A[x, y] \stackrel{\text{def}}{=} B[y] = A[x][y].$$

For example, $xy - y^2 + 3 + x^3$ is an element of C: it can be viewed as a polynomial in y of degree-two, with coefficients $(3 + x^3, x, -1)$.

More generally, we define inductively

$$A[x_1, \ldots, x_n] \stackrel{\text{def}}{=} A[x_1 \ldots x_{n-1}][x_n].$$

A priori, this definition depends on the order given to the variables; but one can verify that $A[x, y]$ is "the same as" $A[y, x]$. (We will have a more precise notion of "same as" later on, after we introduce the notion of isomorphism.)

Definition 160. Let d be a positive integer. A polynomial $m \in A[x_1, \ldots, x_n]$ is called a *monomial of degree d* if it is of the form

$$\alpha\, x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \ldots \cdot x_n^{\alpha_n}, \text{ with } \alpha \in A, \alpha_i \in \mathbb{N}, \text{ and } \sum_{i=1}^{n} \alpha_i = d.$$

A *homogeneous polynomial of degree d* is a polynomial that is sum of monomials of the same degree d. A *homogeneous polynomial* is a polynomial that is homogeneous of degree d, for some d.

Example 161. The polynomial $x_1^2 x_3 x_4 - x_3^3 x_4$ is homogeneous (of degree 4).

Definition 162. A polynomial $p \in A[x_1, \ldots, x_n]$ is called *symmetric* if it remains unchanged when we swap the labels of any two of the variables $x_1, \ldots, x_n$.

Example 163. The polynomial $x_1 x_2 + x_1^2$ is homogeneous, but not symmetric: If we swap x_1 with x_2, it changes to $x_1 x_2 + x_2^2$. Instead, the sum $S = x_1 + \ldots x_n$ and the product $P = x_1 \cdot \ldots \cdot x_n$ are symmetric and homogeneous (of degree 1 and n, respectively). Their sum $S + P$ is of course symmetric, but not homogeneous.

Definition 164. The *elementary symmetric polynomials* in n variables $x_1, \ldots, x_n$, are the $n+1$ polynomials defined by

$$\begin{array}{ll} e_0 & \stackrel{\text{def}}{=} 1 \\ e_1 & \stackrel{\text{def}}{=} \sum_{1\leq j\leq n} x_j = x_1 + x_2 + \ldots + x_n \\ e_2 & \stackrel{\text{def}}{=} \sum_{1\leq j<k\leq n} x_j x_k \\ e_3 & \stackrel{\text{def}}{=} \sum_{1\leq j<k<l\leq n} x_j x_k x_l \\ \vdots & \qquad\vdots \\ e_i & \stackrel{\text{def}}{=} \sum_{1\leq j_1<j_2<\ldots<j_i\leq n} x_{j_1} x_{j_2} \cdots x_{j_i} \\ \vdots & \qquad\vdots \\ e_n & \stackrel{\text{def}}{=} x_1 x_2 \cdots x_n. \end{array}$$

Note that each e_i is homogeneous of degree i.

Example 165. For $n = 4$ we are talking about the five polynomials

$$\begin{array}{ll} e_0 & \stackrel{\text{def}}{=} 1 \\ e_1 & \stackrel{\text{def}}{=} x_1 + x_2 + x_3 + x_4 \\ e_2 & \stackrel{\text{def}}{=} x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 \\ e_3 & \stackrel{\text{def}}{=} x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 \\ e_4 & \stackrel{\text{def}}{=} x_1x_2x_3x_4. \end{array}$$

Remark 166. Let q_1, q_2, q_3, q_4 be four elements of $\mathbb{Q}$. Let us consider the product

$$p \stackrel{\text{def}}{=} (x + q_1)(x + q_2)(x + q_3)(x + q_4),$$

Then p is a polynomial in $\mathbb{Q}[x]$ of degree four. Who are its coefficients?

$$\begin{array}{rcl} p & = & x^4 + (q_1 + q_2 + q_3 + q_4)x^3 + (q_1q_2 + q_1q_3 + q_1q_4 + q_2q_3 + q_2q_4 + q_3q_4)x^2 + \\ & & +(q_1q_2q_3 + q_1q_2q_4 + q_1q_3q_4 + q_2q_3q_4)x + q_1q_2q_3q_4 = \\ & = & x^4 + e_1(q_1, q_2, q_3, q_4)x^3 + e_2(q_1, q_2, q_3, q_4)x^2 + e_3(q_1, q_2, q_3, q_4)x + e_4(q_1, q_2, q_3, q_4). \end{array}$$

More generally, if $q_1, \ldots, q_n$ are elements of A and

$$p \stackrel{\text{def}}{=} (x + q_1) \cdot (x + q_2) \cdots (x + q_n),$$

then the coefficient of x^i in p can be obtained by taking the polynomial $e_{n-i}(x_1, \ldots, x_n)$ and plugging in q_i for x_i $(i = 1, \ldots, n)$. These identities go under the name of *Viete's formulas*, and were first noticed by François Viete in the 16th century.

Definition 167. Let A be any C-ring with 1. The *lexicographic order* in $A[x_1, \ldots, x_n]$ is defined as follows: among two terms

$$A = \alpha x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \quad \text{and} \quad B = \beta x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n},$$

we say

$$A > B \stackrel{\text{def}}{\Longleftrightarrow} \left\{ \begin{array}{ll} \text{either} & a_1 > b_1, \\ \text{or} & a_1 = b_1 \text{ but } a_2 > b_2, \\ \text{or} & a_1 = b_1,\ a_2 = b_2 \text{ but } a_3 > b_3, \\ \vdots & \qquad\vdots \\ \text{or} & a_1 = b_1,\ a_2 = b_2,\ \ldots,\ a_{n-1} = b_{n-1} \text{ but } a_n > b_n. \end{array} \right.$$

Given a polynomial p, the largest of its terms is called *leading term* of p, and denoted by $LT(p)$.

Example 168. Among the terms x^3y, x^3z, x^2z^3, y^2 and z^4, we have

$$x^3y > x^3z > x^2y^3 > y^2 > z^4.$$

Example 169. The leading term of the symmetric polynomial e_1 is obviously x_1. More generally, in the symmetric polynomial e_i, the leading term (= the largest one with the lexicographic order) is

$$x_1 \cdot x_2 \cdots x_{i-1} \cdot x_i.$$

Remark 170. The lexicographic order is compatible with the product: if B, C, D are nonzero monomials in $A[x_1, \ldots, x_n]$, and $B < C$, then $BD < CD$. In particular, the leading term of a product is the product of the leading terms.

The lexicographic order can be viewed as a mathematical translation of the usual alphabetic order used in dictionaries. It first appeared in Gauss' 1816 proof of the following theorem, which we present following the exposition by Cox–Little–O'Shea[10]:

Theorem 171 (Fundamental Theorem of Symmetric Polynomials)**.** *Any symmetric polynomial in n variables can be written uniquely as a polynomial expression in terms of $e_1, \ldots, e_n$.*

Proof. Existence.
We proceed by strong induction on the leading term. If the leading term is 1, then the polynomial is e_0. (Or if the leading term is x_1, then it is easy to see that the polynomial is e_1.)

Let p be a symmetric polynomial and let $A \stackrel{\text{def}}{=} LT(p)$. Let us assume by strong induction that any symmetric polynomial with leading term $B < A$ *can* be written as a polynomial expression of $e_1, \ldots, e_n$. Let us write

$$A = \alpha x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}.$$

Because A is the leading term of a *symmetric* polynomial, then automatically $a_1 \geq a_2 \geq \ldots \geq a_n$. (In fact, if for example $a_1 < a_2$, then the term $B = \alpha x_1^{a_2} x_2^{a_1} x_3^{a_3} \cdots x_n^{a_n}$, which is larger than A, would also appear in p, because p is symmetric; but then B would be a term larger than the leading term, a contradiction.) So it makes sense to consider the following polynomial:

$$h \stackrel{\text{def}}{=} \alpha e_1^{a_1 - a_2} \cdot e_2^{a_2 - a_3} \cdots e_{n-1}^{a_{n-1} - a_n} \cdot e_n^{a_n}.$$

Who is the leading term of h? By example 169, the leading term of e_i is $x_1 \cdots x_i$. Since the leading term of the product is the product of the leading terms, the leading term of h is

$$\begin{aligned} LT(h) &= \alpha \cdot LT(e_1)^{a_1 - a_2} \cdot LT(e_2)^{a_2 - a_3} \cdots LT(e_{n-1})^{a_{n-1} - a_n} \cdot LT(e_n)^{a_n} = \\ &= \alpha \cdot (x_1)^{a_1 - a_2} \cdot (x_1 x_2)^{a_2 - a_3} \cdots (x_1 \cdots x_{n-1})^{a_{n-1} - a_n} \cdot (x_1 \cdots x_n)^{a_n} = \\ &= \alpha \cdot x_1^{a_1} \cdot x_2^{a_2} \cdot x_3^{a_3} \cdots x_{n-1}^{a_{n-1}} \cdot x_n^{a_n} \;=\; A. \end{aligned}$$

So h has by construction the same leading term of p. Yippie-kay-yay! Then consider $p - h$: This is again a symmetric polynomial (because p, h are) and its leading term is smaller than $LT(p)$, which was "killed" in the subtraction. Hence, by strong induction, $p - h$ can be written as a polynomial expression in terms of $e_1, \ldots, e_n$. But h is by construction a polynomial in $e_1, \ldots, e_n$. Therefore $p = (p - h) + h$ can be written a polynomial expression in terms of $e_1, \ldots, e_n$.

Uniqueness.
By contradiction, suppose there are two different polynomials G_1, G_2 in $A[y_1, \ldots, y_n]$ such that when we plug in e_i for y_i, we obtain

$$G_1(e_1, \ldots, e_n) = G_2(e_1, \ldots, e_n).$$

[10] D. Cox, J. Little, D. O'Shea, *Ideals, Varieties, and Algorithms*, Springer Verlag, 2007; Chapter 7.

Set $G \stackrel{\text{def}}{=} G_1 - G_2$. By the assumption, G is a nonzero polynomial such that $G(e_1, \dots, e_n) = 0$. This G will be a sum of expressions of the form

$$M_{(a_1,\dots,a_n)} = \alpha_{(a_1,\dots,a_n)} \cdot y_1^{a_1} y_2^{a_2} \cdots y_n^{a_n}. \tag{19}$$

Here $\alpha_{(a_1,\dots,a_n)}$ is just a coefficient in A; we used this complicated notation just to remember that the various terms come with different coefficients. So formally the coefficient of each monomial "depends" on the n-tuple $(a_1, \dots, a_n)$ we are considering. Now plug in e_i for y_i. As we do it, we use the convention of passing to lowercase notation: so for us $g \stackrel{\text{def}}{=} G(e_1, \dots, e_n)$ will be a sum of expressions of the form

$$m_{(a_1,\dots,a_n)} = \alpha_{(a_1,\dots,a_n)} \cdot e_1^{a_1} e_2^{a_2} \cdots e_n^{a_n}. \tag{20}$$

Now that we are in $A[x_1, \dots, x_n]$, the leading term of the monomial in Equation 20, again exploiting example 169, is

$$\begin{aligned} LT\left(m_{(a_1,\dots,a_n)}\right) &= \alpha_{(a_1,\dots,a_n)} \cdot (x_1)^{a_1} (x_1 x_2)^{a_2} \cdots (x_1 \cdots x_n)^{a_n} = \\ &= \alpha_{(a_1,\dots,a_n)} \cdot (x_1)^{a_1+a_2+\dots+a_n} \cdot (x_2)^{a_2+\dots+a_n} \cdots (x_n)^{a_n}. \end{aligned}$$

Since the map from $\mathbb{N}$ to $\mathbb{N}$ defined by

$$(a_1, \dots, a_n) \longmapsto (a_1 + a_2 + \dots + a_n, a_2 + \dots + a_n, \dots, a_n)$$

is injective, the leading terms of the various monomials $m_{(a_1,\dots,a_n)}$, as $(a_1, \dots, a_n)$ varies, are all different. But then, no cancellation can occur, so g cannot be zero. This contradicts the assumption that $g = G(e_1, \dots, e_n) = 0$. □

The previous proof gives you also an **algorithm** to find a polynomial representation in $e_1, \dots e_n$ of any symmetric polynomial in n variables. The idea is to iteratively "kill the leading term", and iterate until you get the zero polynomial. Let us illustrate it with a couple of examples.

Example 172. The polynomial $f = x^4 + y^4 + z^4$ is symmetric in $\mathbb{Z}[x, y, z]$ and homogeneous. The elementary symmetric polynomials we want to use to express f are

$$e_1 = x + y + z, \qquad e_2 = xy + yz + xz, \qquad \text{and } e_3 = xyz,$$

whose leading terms are

$$LT(e_1) = x, \qquad LT(e_2) = xy, \qquad \text{and } LT(e_3) = xyz.$$

The leading term of f is x^4. We want to write it as product of powers of the three leading terms above, i.e. we look for natural numbers a, b, c such that

$$x^4 = x^a \cdot (xy)^b \cdot (xyz)^c.$$

This leads to the system

$$\begin{cases} 4 = a + b + c, \\ 0 = b + c, \\ 0 = c, \end{cases}$$

which immediately tells us $c = 0$, $b = 0$, $a = 4$. In other words, x^4 is simply a power of x, which is the leading term of e_1. So we subtract:

$$\begin{aligned} f_1 \stackrel{\text{def}}{=} f - (e_1)^4 = \; & \underline{-4x^3y} - 4x^3z - 6x^2y^2 - 12x^2yz - 6x^2z^2 - 4xy^3 - 12xy^2z - 12xyz^2 \\ & -4xz^3 - 4y^3z - 6y^2z^2 - 4yz^3 \end{aligned}$$

The leading term of f_1, underlined above, is $-4x^3y$. Again, we look for natural numbers a, b, c such that

$$x^3y = x^a \cdot (xy)^b \cdot (xyz)^c.$$

This time we get the system

$$\begin{cases} 3 = a + b + c, \\ 1 = b + c, \\ 0 = c, \end{cases}$$

whence $c = 0$, $b = 1$, $a = 2$. So to kill the leading term of f_1, we should subtract $(e_1)^2e_2$, with the right coefficient:

$$f_2 \stackrel{\text{def}}{=} f_1 + 4(e_1)^2e_2 = \underline{2x^2y^2} + 8x^2yz + 2x^2z^2 + 8xy^2z + 8xyz^2 + 2y^2z^2.$$

The leading term of f_2 is $2x^2y^2$. Again, we look for natural numbers a, b, c such that

$$x^2y^2 = x^a \cdot (xy)^b \cdot (xyz)^c,$$

and solving the system we get $a = 0$, $b = 2$, $c = 0$. So to kill the leading term of f_2, we should subtract $2(e_2)^2$:

$$f_3 \stackrel{\text{def}}{=} f_2 - 2(e_2)^2 = \underline{4x^2yz} + 4xy^2z + 4xyz^2.$$

The leading term is $4x^2yz$. We look for a, b, c such that $x^2y^2 = x^a \cdot (xy)^b \cdot (xyz)^c$ and we obtain $a = 1, b = 0, c = 1$. So to kill the leading term of f_3, we subtract $4e_1e_3$:

$$f_4 \stackrel{\text{def}}{=} f_3 - e_1e_3 = 0.$$

Finally we obtained 0, so we stop. (The algorithm always reaches zero after a finite number of steps, because the leading term keeps getting smaller at every step.) The conclusion is that

$$\begin{aligned} f &= (e_1)^4 + f_1 = (e_1)^4 - 4(e_1)^2e_2 + f_2 = (e_1)^4 - 4(e_1)^2e_2 + 2(e_2)^2 + f_3 = \\ &= (e_1)^4 - 4(e_1)^2e_2 + 2(e_2)^2 + e_1e_3, \end{aligned}$$

which is the desired polynomial expression.

Example 173. Consider the symmetric polynomial $g = x^4 + y^4 + z^4 + xyz$ in $\mathbb{Z}[x, y, z]$. Rather than applying the algorithm directly, we notice that $g = f + h$, where $f = x^4 + y^4 + z^4$ and $h = xyz$. Since we already showed in the previous example that

$$f = (e_1)^4 - 4(e_1)^2e_2 + 2(e_2)^2 + e_1e_3,$$

we conclude that

$$g = (e_1)^4 - 4(e_1)^2e_2 + 2(e_2)^2 + e_1e_3 + e_3.$$

2.5 Exercises

1. Find a nilpotent polynomial f of degree one, with coefficients in some C-ring A.
2. Prove Lemma 138.
3. Prove Theorem 129.
4. Is the converse of Theorem 145 true? Can there be polynomials f of degree one with no roots? What about degree two?

5. Find a gcd of $x^3 - x + 1$ and x^4 in $\mathbb{Z}_2[x]$. Call it $d(x)$. Find polynomials a, b such that

$$d(x) = (x^3 - x + 1)a(x) + x^4 b(x).$$

6. Find a gcd of $x^2 + 1$ and $x^3 + 1$ in $\mathbb{Z}_3[x]$. What if I asked, "in $\mathbb{Z}_2[x]$"?

7. How many polynomials of degree four in $\mathbb{Z}_3[x]$ are there?

8. Write down $x^4 + y^4 + x^2y^2$ as polynomial expression of the elementary symmetric polynomials in $\mathbb{Z}[x, y]$.

9. Prove the following stronger version of Fermat's little theorem 80: Let a, m be any integers such that $\gcd(a, m) = 1$. Then m is prime if and only if in $\mathbb{Z}_m[x]$,

$$(x + a)^m = x^m + a.$$

3 Ideals, Homomorphisms, PIDs and Noetherian domains

3.1 Subrings

Definition 174. Let A be a C-ring with operations $+, \cdot$. A subset $B \subseteq A$ is called *C-subring* of A if B is a C-ring with the same operations of A.

With this definition, it would be very long to verify that a certain set is a subring. It turns out however that many properties are automatically satisfied; for example, associativity always holds, because if $x(yz) = (xy)z$ for all x, y, z *in* A, then in particular $x(yz) = (xy)z$ for all x, y, z *in* B. Similarly for commutativity and distributivity. So essentially, to check if a subset B of A is a subring, we only need to verify:

– that the operations are internal to B (that is, for each $x, y \in B$, both $x + y$ and $x \cdot y$ are in B);
– that the neutral element of the sum 0 is in B;
– that for any element x in B, its additive inverse $-x$ is still in B.

None of the axioms above is superfluous. However, there is a way to simplify the test even further.

Proposition 175. *A subset $B \subseteq A$ is a* subring *of A if and only if B satisfies the following conditions:*

(SR1) *For each $x, y \in B$, the difference $x - y$ is in B;*
(SR2) *For each $x, y \in B$, the product $x \cdot y$ is in B.*

Proof. Let x, y be elements of B. If B is a subring, also $-y \in B$, so $x - y \stackrel{\text{def}}{=} x + (-y)$ is also in B. Hence, (SR1) and (SR2) both hold.

Conversely, suppose (SR1) and (SR2) hold. Let x be an element of B. Applying (SR1) to $y = x$, we get that the difference $0 = x - x$ is in B. So B contains 0. But then if y is any element of B, applying again (SR1) to $x = 0$ we get that $0 - y$ is in B too. So B contains all additive inverses. It remains to show that the sum is internal with respect to B. This is now easy because if x, y are in B, then we know that $-y$ is in B, so by (SR1) $x + y = x - (-y)$ is also in B. □

Example 176. $2\mathbb{Z}$ is a subring of $\mathbb{Z}$, because the difference and the product of two even numbers is even. $\mathbb{Z}$ is a subring of $\mathbb{Q}$, by identifying every z with $\frac{z}{1}$; in fact, the difference and the product of two fractions with denominator 1 is again a fraction with denominator 1. Analogously, $\mathbb{Q}$ is a subring of $\mathbb{R}$ and $\mathbb{R}$ is a subring of $\mathbb{C}$.

Non-Example 177. $\mathbb{Z}_2$ is not a subring of $\mathbb{Z}$, because the operations are different (in $\mathbb{Z}_2$ one has $1 + 1 = 0$). The set of odd integers is not a subring of $\mathbb{Z}$: it it satisfies (SR2) but not (SR1). (Why?) Finally, consider

$$A = \{q\sqrt{2} \text{ such that } q \in \mathbb{Q}\}.$$

This is not a subring of $\mathbb{R}$: it satisfies (SR1) but not (SR2), as you can see picking $x = y \stackrel{\text{def}}{=} \sqrt{2}$.

The notion of "subring" allows us to compare rings that are contained in one another. However, there is a more general way to compare rings that are connected by a special function, namely, a function that respects the ring structure.

3.2 Homomorphisms

Definition 178. Let A, B be two C-rings. A function $\varphi : A \to B$ is called a *(ring) homomorphism* if for each x, y in A, it satisfies:
(RH1) $\varphi(x+y) = \varphi(x) + \varphi(y)$;
(RH2) $\varphi(x \cdot y) = \varphi(x) \cdot \varphi(y)$.

Remark 179. Let k be a positive integer. If φ is any homomorphism, note that in particular

$$\varphi(kx) = \varphi\left(x + x + \ldots x\right) = \varphi(x) + \varphi(x) + \ldots + \varphi(x) = k\varphi(x), \text{ and}$$

$$\varphi(x^k) = \varphi\left(x \cdot x \cdot \ldots x\right) = \varphi(x) \cdot \varphi(x) \cdot \ldots \cdot \varphi(x) = [\varphi(x)]^k, \text{ for all } x.$$

Example 180 (Inclusion)**.** Let B be a subring of A. Then the inclusion $\iota : B \to A$, mapping any b in B to itself, is a homomorphism. It is injective, but unless $B = A$, it is not surjective. Note: only in case $B = A$ this is called the "identity map"! Remember that two maps are the same if *domain and codomain are the same*, and same inputs are mapped to same outputs.

Example 181 (Zero map)**.** Let A, B be arbitrary C-rings. Then the zero map $\mathbf{0}$: $B \to A$ is a homomorphism, called the *zero map* or sometimes the *trivial homomorphism*. Unless $B = 0$, it is not injective; and unless $A = 0$, it is not surjective either.

Example 182 (Complex Conjugation)**.** The map conj : $\mathbb{C} \to \mathbb{C}$ that sends any complex number $z = a + ib$ to its conjugate $\overline{z} = a - ib$ is a homomorphism by Lemma 156. It is both injective and surjective. In fact, conj $\circ$ conj is the identity homomorphism.

Example 183 (Polynomial Evaluation)**.** Let a be any element of a C-ring A with 1. The *evaluation homomorphism* from $A[x]$ to A is the map that plugs in $x = a$ in any polynomial of $A[x]$. Formally, it is the map $\varphi_a : A[x] \to A$ that maps $f \mapsto f(a)$.

For example, if $a = 7$ and $A = \mathbb{Z}$, $\varphi_7 : \mathbb{Z}[x] \to \mathbb{Z}$ is the map that "plugs in $x = 7$". So $\varphi_7(x^2 + x + 1) = 55$. It is a homomorphism because it doesn't matter whether we first sum/multiply two polynomials and then plug in $x = 7$, or whether we first plug in and then sum/multiply. In fact, $\varphi_7(x) \cdot \varphi_7(x) + \varphi_7(x) + \varphi_7(1) = 7 \cdot 7 + 7 + 1 = 55$.

The evaluation is surjective, because for any $c \in A$, if we view c as a constant polynomial $c \in A[x]$, then $\varphi_a(c) = c$. (Plugging in $x = a$ does not change anything, if the polynomial does not contain any x!) However, the evaluation is not injective. For example, $\varphi_a(x - a) = 0 = \varphi_a(0)$.

Example 184 (Multiplication by idempotent)**.** Given any C-ring A, let r be an element such that $r^2 = r$. (As a concrete example, you may choose as A the C-ring formed by the 2×2 diagonal matrices with entries in $\mathbb{R}$, and $r \overset{\text{def}}{=} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.) Then the *multiplication by* r,

$$\begin{array}{rccc} \mu_r : & A & \longrightarrow & A \\ & a & \longmapsto & ra, \end{array}$$

is an injective homomorphism: one has in fact

$$\mu_r(a) + \mu_r(b) = ra + rb = r(a+b) = \mu_r(a+b) \quad \text{and} \quad \mu_r(a)\mu_r(b) = rarb = r^2ab \overset{!}{=} rab = \mu_r(ab),$$

where you see the importance of assuming $r^2 = r$.

Non-Example 185 (Multiplication)**.** Let $f : \mathbb{Z} \to 2\mathbb{Z}$ be the map defined by $f(x) = 2x$. This map is both injective and surjective, but it is not a homomorphism! It is true that $f(a+b) = f(a) + f(b)$, but $f(ab) \neq f(a)f(b)$.

Example 186 (Remainder of division by n)**.** Let n be an integer ≥ 2. Consider the map $\varphi_n : \mathbb{N} \to \mathbb{Z}_n$ that maps any integer $z \geq 0$ to the remainder of the division of z by n. We can extend it to a homomorphism $\varphi_n : \mathbb{Z} \to \mathbb{Z}_n$ as follows: if $z < 0$, then $-z \in \mathbb{N}$, so we set $\varphi_n(z) = -\varphi_n(-z)$. This homomorphism is surjective, but not injective. (Why?)

Example 187 (Projection from $\mathbb{Z}_m$ to $\mathbb{Z}_d$)**.** Let m be an integer ≥ 2. Let d be a divisor of m. Define a function $\pi : \mathbb{Z}_m \longrightarrow \mathbb{Z}_d$ by $\pi(\overline{z}) \stackrel{\text{def}}{=} \varphi_d(z)$, where φ_d is the map from the example above. This function is well-defined, because if $\overline{z} = \overline{z'}$ in $\mathbb{Z}_m$, then $z - z'$ is a multiple of m, so in particular it is a multiple of d; hence, $\varphi_d(z) = \varphi_d(z')$. It is easy to see that π is a homomorphism that is surjective, but not injective, unless $d = m$.

Definition 188. A bijective homomorphism is called an *isomorphism*. We say that two C-rings A and B are *isomorphic*, and we write

$$A \cong B,$$

if there exists some isomorphism from A to B.

One should reiterate that if A and B are isomorphic, then *some* homomorphism from A to B is bijective, but not necessarily all of them. For example, the zero map is not (unless $B = A = (0)$.)

Lemma 189. *The composition of a homomorphism $\varphi : A \to B$ and a homomorphism $\psi : B \to C$ is a homomorphism from A to C, denoted by $\psi \circ \varphi$.*
Moreover, the inverse of any isomorphism is again an isomorphism.

Proof. Let x, y be elements of A.

$$\psi \circ \varphi(x + y) \stackrel{\text{def}}{=} \psi\left(\varphi(x) + \varphi(y)\right) = \psi(\varphi(x)) + \psi(\varphi(y)) = \psi \circ \varphi(x) + \psi \circ \varphi(y).$$

Similarly for the product. As for the final claim, suppose $\varphi : A \to B$ is a bijective homomorphism. Consider $\varphi^{-1} : B \to A$. For any b, b' in B,

$$\varphi\left(\varphi^{-1}(b) + \varphi^{-1}(b')\right) = \varphi\left(\varphi^{-1}(b)\right) + \varphi\left(\varphi^{-1}(b')\right) = b + b',$$

so applying φ^{-1} to both sides, we get $\varphi^{-1}(b) + \varphi^{-1}(b') = \varphi^{-1}(b + b')$. Similarly,

$$\varphi\left(\varphi^{-1}(b) \cdot \varphi^{-1}(b')\right) = \varphi\varphi^{-1}(b) \cdot \varphi\varphi^{-1}(b') = b \cdot b',$$

so $\varphi^{-1}(b) \cdot \varphi^{-1}(b') = \varphi^{-1}(b \cdot b')$. □

Proposition 190. *For any homomorphism $\varphi : A \to B$, we have:*

- $\varphi(0) = 0$.
- *for any x in A, $\varphi(-x) = -\varphi(x)$*
- *for any x, y in A, $\varphi(x - y) = \varphi(x) - \varphi(y)$.*
- *In case A, B are C-rings with 1, and $\varphi(1) = 1$, if $x \in A$ is invertible then so is $\varphi(x)$, and $\varphi(x)^{-1} = \varphi(x^{-1})$.*

Proof. For any $x \in A$ we have $\varphi(x) = \varphi(x + 0) = \varphi(x) + \varphi(0)$, so by cancellation $0 = \varphi(0)$. This shows the first item. Now for any $x \in A$, $\varphi(-x) + \varphi(x) = \varphi(-x + x) = \varphi(0) = 0$; so $\varphi(-x) = -\varphi(x)$. In particular $\varphi(x - y) \stackrel{\text{def}}{=} \varphi(x + (-y)) = \varphi(x) + (-\varphi(y)) \stackrel{\text{def}}{=} \varphi(x) - \varphi(y)$. Finally, for any x we have $\varphi(x^{-1}) \cdot \varphi(x) = \varphi(x^{-1} \cdot x) = \varphi(1)$. Since $\varphi(1) = 1$ by assumption, we conclude. □

Note that we are not claiming $\varphi(1) = 1$, because this is not always the case: Compare Example 181, in which 1 is mapped to 0, and Example 184, in which 1 is mapped to some r such that $r^2 = 1$. In case φ is the zero map from $\mathbb{Q}$ to $\mathbb{Q}$, even if x is invertible $\varphi(x)$ is not invertible, of course.

3.3 Ideals

Definition 191. Let $\varphi : A \to B$ be a homomorphism. The *kernel of* φ is the set

$$\ker \varphi \stackrel{\text{def}}{=} \{a \in A \text{ such that } \varphi(a) = 0\},$$

or in other words, the pre-image of 0. The *image of* φ is the set

$$\operatorname{Im} \varphi \stackrel{\text{def}}{=} \{\varphi(a) \text{ such that } a \in A\}.$$

Proposition 192. *With the notation above,* $\ker \varphi$ *is a subring of* A *and* $\operatorname{Im} \varphi$ *is a subring of* B*.*

Proof. By Proposition 175, it suffices to check (SR1) and (SR2). Let x, y be in $\ker \varphi$. Then by Proposition 190

$$\varphi(x - y) = \varphi(x) - \varphi(y) = 0 - 0 = 0$$
$$\varphi(x \cdot y) = \varphi(x) \cdot \varphi(y) = 0 \cdot 0 = 0.$$

So both $x - y$ and $x \cdot y$ are in $\ker \varphi$.

Similarly, if $\varphi(x)$ and $\varphi(y)$ are elements of $\operatorname{Im} \varphi$, then $\varphi(x) - \varphi(y)$ and $\varphi(x) \cdot \varphi(y)$ are also elements of $\operatorname{Im} \varphi$, because $\varphi(x) - \varphi(y) = \varphi(x - y)$ and $\varphi(x) \cdot \varphi(y) = \varphi(x \cdot y)$. □

It is obvious from the definition that $\varphi : A \to B$ is surjective if and only if $\operatorname{Im} \varphi = B$. A less obvious criterion holds for injectivity.

Proposition 193. *A homomorphism* $\varphi : A \to B$ *is injective if and only if* $\ker \varphi = \{0\}$*.*

Proof. Only if: By Proposition 190 we know that $\varphi(0) = \{0\}$. So if φ is injective, there cannot be another element $x \neq 0$ such that $\varphi(x) = 0$. This shows $\ker \varphi = \{0\}$.

If: Suppose $\varphi(x) = \varphi(y)$. Then $\varphi(x - y) = \varphi(x) - \varphi(y) = 0$, so $x - y$ belongs to $\ker \varphi$. But if $\ker \varphi = \{0\}$, this means that $x - y = 0$. So $x = y$. □

The kernel and the image are very important subrings. The kernel, however, has a richer structure. If x is in $\ker \varphi$, and if y is in $\ker \varphi$, we have seen that xy is, too. If you remember the proof (look up), this was because $0 \cdot 0 = 0$. But the truth is that $0 \cdot$ *anything* is still zero! So even if x is in $\ker \varphi$ and y is **not** in $\ker \varphi$, we still have that $x \cdot y$ is in $\ker \varphi$, because

$$\varphi(x \cdot y) = \varphi(x) \cdot \varphi(y) = 0 \cdot \varphi(y) = 0.$$

This inspires the following definition.

Definition 194. A subset $I \subseteq A$ is called an *ideal* of A if it satisfies the following condition:
(I1) For each $x, y \in I$, the difference $x - y$ is in I;
(I2) For each $x \in I$, for each $a \in A$, the product $a \cdot x$ is in I.

A comparison with Proposition 175 immediately reveals that *all ideals are subrings*. In fact, condition (I1) is identical to (SR1); yet condition (I2) is stronger than (SR2). We want ax to be in I not only if $a \in I$, but even if a is an element of A outside I.

Non-Example 195. $\mathbb{Z}$ is a subring of $\mathbb{Q}$, but it is not an ideal of $\mathbb{Q}$. (For example, $3 \in \mathbb{Z}$, $\frac{1}{2} \in \mathbb{Q}$, yet $3 \cdot \frac{1}{2} \notin \mathbb{Z}$). This shows that not every subring is an ideal.

Example 196. For any $\varphi : A \to B$ homomorphism, the kernel $\ker \varphi$ is always an ideal of A. In contrast, the image is not always an ideal of B: for a counterexample, take any subring $B \subseteq A$ that is not an ideal (like $\mathbb{Z} \subseteq \mathbb{Q}$) and look at the inclusion map $\iota : B \to A$.

Example 197. Let $n \in \mathbb{N}$. Then

$$(n) \stackrel{\text{def}}{=} \{\text{multiples of } n\}$$

is an ideal of $\mathbb{Z}$. (The difference of two multiples of n is a multiple of n; moreover, a multiple of n times any number, yields a multiple of n.)

Proposition 198. *All ideals of $\mathbb{Z}$ are of the form (n), for some $n \in \mathbb{N}$.*

Proof. Let I be an ideal of $\mathbb{Z}$. If $I = (0)$, we are done. Otherwise, I contains some nonzero element a. By (I1), I must contain $-a$. Since one of a and $-a$ is positive, we can safely say that I contains some *positive* integer. Let n be the **smallest** positive integer contained in I. Our goal is to show that $I = (n)$:

- $I \supset (n)$ is easy: Since n is in I, all multiples of n must be in I too by the condition (I2).
- To show $I \subseteq (n)$, it suffices to show that any positive element of I is a multiple of n. So, choose $x > 0$ in I. Let us divide it by n: we obtain

  $$x = q \cdot n + r, \qquad \text{with } 0 \leq r < n.$$

  If we manage to prove that $r = 0$, then x is a multiple of n and we are done. By contradiction, suppose $r \neq 0$. Recall that n was by definition the smallest positive integer in I; since $r < n$, it follows that $r \notin I$. Yet $qn \in I$ by (I2), so $r = x - qn$ is the difference of two elements of I and it is not in I: A contradiction with (I1). □

Example 199. Every C-ring A has two "obvious" ideals, namely, (0) and A itself. Interestingly, some C-rings have only these two ideals! In fact, we will see now that having only two ideals is a property that characterizes fields.

Lemma 200 ("Explosion")**.** *Let A be a C-ring with 1. If an ideal I contains an invertible element, then $I = A$. (We'll say, the ideal I "explodes" to the whole C-ring.)*

Proof. If $u \in I$ is invertible, then $u \cdot u^{-1} \in I$ by (I2). So $1 \in I$. But any $a \in I$ can be written as $a = a \cdot 1$. So since 1 is in I, $a \cdot 1$ is in I by (I2). □

Proposition 201. *Let A be a C-ring with 1. The following are equivalent:*

(1) *A has only two ideals.*

(2) *A is a field.*

Proof.

(2) ⇒ (1). Let I be an ideal of A different than (0). Then I contains an element $u \neq 0$, which is necessarily invertible because A is a field. So by the Explosion Lemma (Lemma 200), we must have $I = A$.

(1) ⇒ (2). Let g be any nonzero element of A. We want to show that g is invertible. Consider the set

$$(g) \stackrel{\text{def}}{=} \{ \text{ multiples of } g \ \} = \{ag \text{ such that } a \in A\}.$$

Let us check that this is an ideal. Let $a_1 g, a_2 g$ be two elements of (g), with a_1, a_2 in A. Then the difference $a_1 g - a_2 g = (a_1 - a_2)g$ is still in (g), because $a_1 - a_2$ is in A. Similarly, if $a \in A$, then $a \cdot (a_1 g) = (a a_1) \cdot g$, which is in (g) because $a a_1$ is in A. So (I1) and (I2) are satisfied.

Since $g \neq 0$, the ideal (g) is not (0). But by the assumption, A has only two ideals, which necessarily are (0) and A by Example 199. So $(g) = A$, and in particular, $1 \in (g)$. This means that there exists $a \in A$ such that $ag = 1$. Hence g is invertible. By the arbitrariety of g, A is a field. □

3.4 Generators and Principal Ideals

Let A be any C-ring. Fix an element g of A. We have seen in the proof of Proposition 201 (and also in Proposition 198, for $A = \mathbb{Z}$) that we can always form the ideal

$$(g) \stackrel{\text{def}}{=} \{ \text{ multiples of } g\} = \{a \cdot g \text{ such that } a \in A\}.$$

The ideals of this form are called *principal*. The element g is called a *generator* of (g). Of course, if A contains 1, then $1 \cdot g \in (g)$.

Example 202. In $\mathbb{Z}$, the ideal (2) is the set of all even numbers.

Proposition 203. *If A is a C-ring with 1, the ideal (g) is the smallest ideal of A containing g.*

Proof. Let J be any ideal of A containing x. By (I2), J must contain all multiples of x. So, $J \supset (x)$. □

Remark 204. Given a principal ideal, its generator is not uniquely determined. For example, even numbers are the ideal generated by 2, but also the ideal generated by -2. The next Lemma, however, shows that the generator of a principal ideal is unique *up to an invertible factor*, if the C-ring is a domain.

Lemma 205. *Let a, b be elements of a C-ring A with 1. If A is a domain, the following are equivalent:*

(1) $(a) = (b)$

(2) $a = b \cdot u$, *with u invertible.*

Proof.

(2) ⇒ (1). If $a = b \cdot u$, then any multiple of a is also a multiple of b, so $(a) \subseteq (b)$. But since u is invertible, multiplying by u^{-1} we get $a \cdot u^{-1} = b$; so any multiple of b is also a multiple of a, hence $(b) \subseteq (a)$. (For this implication, we did not use that A is a domain.)

(1) ⇒ (2). Suppose $(a) = (b)$. Since $b \in (a)$, we know that $b = a \cdot y$, for some y in A. Symmetrically, since $a \in (b)$, we know that $a = b \cdot x$, for some x in A. Substituting $a \cdot y$ for b in the latter equation, we obtain

$$a = a \cdot y \cdot x, \text{ for some } x, y \text{ in } A.$$

Now we distinguish two cases. If $a = 0$, then $b = 0 \cdot y = 0$ too; so in this case we can write $a = b \cdot 1$, and 1 is invertible. If instead $a \neq 0$, since A is a domain we can cancel the factor a and conclude $1 = yx$. So both x and y is invertible, and setting $u = x$, we are done. □

Non-Example 206. Let $A = \{f : \mathbb{R} \longrightarrow \mathbb{R} \text{ continuous}\}$. Please check that A is a C-ring with point-wise sum and multiplication, and that the invertible elements in A are the continuous functions that never vanish. Now draw for me the three continuous functions

$$a : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} -x & \text{if } x < 0 \\ 0 & \text{if } 0 \leq x \leq 1 \\ x - 1 & \text{if } x > 1 \end{cases}$$

$$b : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} x & \text{if } x < 0 \\ 0 & \text{if } 0 \leq x \leq 1 \\ x - 1 & \text{if } x > 1. \end{cases}$$

$$c : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} -1 & \text{if } x < 0 \\ 2x - 1 & \text{if } 0 \leq x \leq 1 \\ +1 & \text{if } x > 1. \end{cases}$$

All three functions vanish somewhere, so none of them is invertible. It is easy to check that $a = bc$ and $b = ac$, so

$$(a) = (b).$$

Now let u be any continuous function such that $a = bu$. This u must be a continuous function that yields -1 on $x \leq 0$, and $+1$ on $x \geq 1$. By the Intermediate Value Theorem of Calculus, this u must vanish in some point $x \in (0,1)$. So there is no *invertible* u in A such that $a = bu$. With a completely identical reasoning, one shows that there is no invertible v in A such that $b = av$. Of course, there is no conflict with Lemma 205, because A is not a domain.

Inspired by this, we give the following generalization.

Definition 207. Let A be any C-ring. Let n be a positive integer. Let $g_1, \ldots, g_n$ be elements of A (not necessarily distinct). Set

$$(g_1, \ldots, g_n) \stackrel{\text{def}}{=} \left\{ \begin{array}{l} \text{linear combinations of the } g_i \\ \text{with coefficients in } A \end{array} \right\} \stackrel{\text{def}}{=} \{a_1 g_1 + \ldots + a_n g_n \text{ such that } a_i \in A\}.$$

The set $\{g_1, \ldots, g_n\}$ is called *a set of generators* of the ideal $(g_1, \ldots, g_n)$.

Proposition 208. *If A is a C-ring with 1, then $(g_1, \ldots, g_n)$ is the smallest ideal of A containing $g_1, \ldots, g_n$.*

Proof. It is an ideal because (I1) the sum of two linear combinations is still a linear combination, and (I2) if we multiply a linear combination of the x_i with coefficients in A by an element of A, we still have a linear combination (with rescaled coefficients). That $(g_1, \ldots, g_n)$ contains each g_i follows from choosing $a_i = 1$ and $a_j = 0$ for all $j \neq i$. Moreover, any ideal J of A that contains $g_1, \ldots, g_n$ must contain their linear combinations; so J contains $(g_1, \ldots, g_n)$. □

Remark 209. The *order* of the generators is not important. For example,

$$(x_1, x_2) = (x_2, x_1),$$

because any element of the set of the left (say, $ax_1 + bx_2$, with a, b in A) can be expressed, using commutativity, as an element of the set on the right (in our case, $bx_2 + ax_1$). For this reason, we speak of a *set* of generators, and not of an *ordered list* of generators.

That said, the *set* of generators is not uniquely determined by the ideal – not even if we are in a domain. For example, consider in $\mathbb{Z}$ the ideal

$$(4,6) = \{4m + 6n \text{ such that } m, n \in \mathbb{Z}\}.$$

Using Diophantine equations, it is easy to see that this is just the set $2\mathbb{Z}$. (Any number of the form $4m + 6n = 2(2m + 3m)$ is obviously even. As for the other inclusion, let $2k$ be an arbitrary even number; since $\gcd(2,3) = 1$, the equation $k = 2x + 3y$ has solutions in $\mathbb{Z}$. But then $2k = 4x + 6y$ has solutions.) Finally, note that $(4,6) = 2\mathbb{Z} = (2)$, yet $(4) \subsetneq (4,6)$ and $(6) \subsetneq (4,6)$. Hence, $(4,6)$ is an inclusion–minimal set of generators, in the sense that none of them is in the ideal generated by the others. This shows that two inclusion–minimal sets of generators might have different cardinalities. (This is a striking difference with the theory of vector spaces.)

Remark 210. Given the ideal $I = (g_1, g_2, \ldots, g_n)$, sometimes one says that the g_i's *are generators* for I. Remember that generating is teamwork! The element g_1 alone does *not* generate I. For this reason, one has to be careful with the wording. For example, if $I = (4,6) \subseteq \mathbb{Z}$, it would be incorrect to say that "4 is a generator of I", because it seems that we are implying $(4) = (4,6)$, which is false. What we mean when we say "4 and 6 are generators of I" is "the set $\{4,6\}$ generates I".

Note that principal ideals are ideals that have a set of generators consisting of only one element. This does not mean that *every* set of generators has only one element. For example, in $\mathbb{Z}$ the ideal $2\mathbb{Z}$ of even numbers is generated by $(4, 6)$, but also by (2). In fact, we have seen in Proposition 198 that *all ideals of* $\mathbb{Z}$ *are principal.*

In the next section, we will investigate this particular property further.

3.5 PIDs

$\mathbb{Z}$ is an example of a domain in which all ideals are principal. This combination of properties is particularly fertile and interesting.

Definition 211. A C-ring A is called *PID* (Principal Ideal Domain) if
(PID1) A is a domain, and
(PID2) every ideal of A is principal.

Example 212. Proposition 198 can be now re-worded as: "$\mathbb{Z}$ is a PID".

Proposition 213. *Every field is a PID. In particular,* $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ *are PID; and* $\mathbb{Z}_m$ *is PID if and only if* m *is prime.*

Proof. Let A be a field. By Proposition 116 A is a domain. By Proposition 201 A has only two ideals, (0) and $A = (1)$. Both are principal. (When m is not a prime, $\mathbb{Z}_m$ is not a PID because of the "D" in "PID".) □

Proposition 214. *Every PID is a C-ring with* 1.

Proof. Since A is an ideal of itself, it must be principal; so $A = (g)$ for some $g \in A$. Now $g \in (g)$, so by definition there exists an element u in A such that $g = u \cdot g$. We claim that this u is the neutral element of the product. In fact, for any $x \in A = (g)$, there exists an $a \in A$ such that $x = g \cdot a$, which implies

$$u \cdot x = u \cdot (g \cdot a) = (u \cdot g) \cdot a = g \cdot a = x.$$ □

Remark 215. Any proper ideal of a PID ring, viewed as ring of its own, is not a PID, because by Lemma 200 it cannot contain 1. In particular, $2\mathbb{Z}$ is an example of a domain that is not a PID. Here is a cleverer example:

Proposition 216. $\mathbb{Z}[x]$ *is a domain that is not a PID; the ideal* $(x, 2)$ *is not principal.*

Proof. Note first that $(x, 2)$ is the set of all polynomials with integer coefficient whose constant term is even. This includes as proper subset the polynomials with *all* coefficients even; so

$$(2) \subsetneq (x, 2) \subsetneq \mathbb{Z}[x].$$

By contradiction, suppose $(x, 2) = (f)$ for some polynomial f in $\mathbb{Z}[x]$. Since $2 \in (x, 2) = (f)$,

$$2 = f \cdot g \text{ for some polynomial } g \in \mathbb{Z}[x].$$

Since we are in a domain, by Lemma 136 we have that $0 = \deg 2 = \deg f + \deg g$, which implies that $\deg f = 0 = \deg g$. So f is an integer that divides 2. So either $f = \pm 1$ and $(x, 2) = (1) = \mathbb{Z}[x]$ by Lemma 200, a contradiction; or $f = \pm 2$ and $(2) = (x, 2)$, another contradiction. □

What about $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$, or the $\mathbb{Z}_p[x]$ with p prime? Are they PID? (The $\mathbb{Z}_m[x]$ with m composite are certainly not PID, because they are not even domains, cf. Theorem 139 and Proposition 115). The answer is surprisingly beautiful.

Theorem 217. *Let A be a C-ring. The following are equivalent:*

(1) *A is a field.*

(2) *$A[x]$ is a PID.*

Before we start proving anything, please read again the reason why $\mathbb{Z}$ is a PID (Proposition 198) and why $\mathbb{Z}[x]$ is not (Proposition 216). You will see where we copied our proof from!

Proof of Theorem 217. Let us start with two forewords. First, we already know A domain $\Leftrightarrow$ $A[x]$ domain (Theorem 139.) Second, since A is a field, the leading coefficient of any nonzero polynomial in $A[x]$ is invertible. So if $g \neq 0$, we can always divide f by g if we want to: Compare the assumptions of Theorem 142.

(1) $\Rightarrow$ (2). Let I be an ideal of $A[x]$. If $I = (0)$, then I is principal and we are done. If I contains a nonzero polynomial c of degree zero, then $c \in A$, which is a field. So c is invertible. But then by Lemma 200, I explodes, so $I = A[x] = (1)$. In particular, I is again principal.
So the interesting case is if I contains 0 and then only polynomials of positive degree. In this case, the set of polynomials in I of positive degree is non-empty. Let g be a **smallest-degree** (nonzero) polynomial inside I. In other words: Choose g such that for any polynomial $f \neq 0$ of I, we have $\deg f \geq \deg g$. We claim that

$$I = (g).$$

The containment $\supset$ is obvious, because $g \in I$. So let us try to prove

$$I \subseteq (g).$$

Pick any polynomial f in I. Let us divide it by g: We can!, because the leading coefficient of g is nonzero and thus invertible. So we have

$$f = q \cdot g + r, \qquad \text{with either } r = 0 \text{ or } \deg r < \deg g.$$

If we manage to prove $r = 0$, then f is a multiple of g and we are done. So by contradiction, assume $\deg r > 0$. Recall that g was by definition the smallest-degree nonzero polynomial in I; since $\deg r < \deg g$, it follows that $r \notin I$. Yet $q \cdot g \in I$ by (I2), so $r = f - q \cdot g$ is the difference of two polynomials in I and it is not in I: A contradiction.

(2) $\Rightarrow$ (1). For any $a \neq 0$ in A, we are going to show that

$$a \text{ is invertible} \iff \text{the ideal } (x, a) \subseteq A[x] \text{ is principal.}$$

From this coimplication the claim follows immediately, because if $A[x]$ is a PID, all ideals are principal and so any $a \neq 0$ must be invertible.
Let us prove the claim. The direction "$\Rightarrow$" is trivial because if a is invertible, by Lemma 200 $(x, a) = A[x] = (1)$. So let us talk about "$\Leftarrow$". Let f be a polynomial such that $(x, a) = (f)$. In particular, $a \in (f)$ and $x \in (f)$.; so we can find polynomials g, h such that

$$\begin{cases} a &= f \cdot g \\ x &= f \cdot h. \end{cases} \tag{21}$$

Being in a domain, we can pass to the degrees and apply Lemma 136: we have

$$\begin{cases} 0 &= \deg f + \deg g \\ 1 &= \deg f + \deg h. \end{cases}$$

This implies $\deg f = \deg g = 0$ and $\deg h = 1$. So we know that $f \in A$, that $g \in A$, and that h is of the form

$$h = h_0 + h_1 x, \qquad \text{for some } h_0, h_1 \in A.$$

Let us plug this back into our System 21: We have

$$\begin{cases} a &= f \cdot g \\ x &= f \cdot (h_0 + h_1 x) = (f h_0) + (f h_1) x. \end{cases} \tag{22}$$

Now, the latter equation is an equality of polynomials, and two polynomials are the same if and only if they have the same coefficients. So our system becomes

$$\begin{cases} a &= f \cdot g \\ 0 &= f \cdot h_0 \\ 1 &= f \cdot h_1. \end{cases} \tag{23}$$

In particular, f is invertible. (As a bonus we get that $h_1 = f^{-1}$ and $h_0 = 0$). So $(x, a) = (f) = A[x]$ by Lemma 200. In particular, $1 \in (x, a)$, so we can find polynomials g', h' such that

$$1 = x \cdot g' + a \cdot h'.$$

By evaluating at $x = 0$, we get

$$1 = a \cdot h'(0),$$

which tells us that a is invertible. □

Corollary 218. *$\mathbb{R}[x, y]$, $\mathbb{Q}[x, y]$, $\mathbb{C}[x, y]$ etc. are not PID.*

Proof. Let $A = \mathbb{R}[x]$. Then A is not a field, because clearly the polynomial x is not invertible. (For any polynomial f, the product $f \cdot x$ has degree at least one... so it cannot be a constant.) Since $\mathbb{R}[x, y] \overset{\text{def}}{=} A[y]$, it follows from Theorem 217 that $A[y]$ is not a PID. In fact, Theorem 217 tells us also how to construct an ideal that is not principal: in $A[y]$ we should take an ideal (a, y), with a not invertible in A. Since we saw that in A the element $a = x$ is not invertible, we conclude that the ideal $(x, y) \subseteq \mathbb{R}[x, y]$ is not principal.

The same reasoning works also for $\mathbb{C}[x, y]$, $\mathbb{Q}[x, y]$, $\mathbb{Z}_p[x, y]$. Also, by induction, we get that $\mathbb{R}[x_1, \ldots, x_n]$ etc. is not PID. □

3.6 Noetherian domains

It is natural to introduce a weakening of the Principal Ideal Domains definition, namely, we could study domains in which every ideal is finitely generated. (The definition we present here seems different; later on we will prove that it is equivalent.) This weaker property is important because it turns out that domains like $\mathbb{C}[x, y]$, $\mathbb{Q}[x, y]$, $\mathbb{R}[x, y]$ and $\mathbb{Z}_p[x, y]$ are Noetherian, even if they are not PID. These rings are crucial to do Algebraic Geometry.

Definition 219. Let A be a C-ring. A is called *Noetherian* if every ascending chain of ideals of A eventually stabilizes. That is, if whenever

$$I_0 \subseteq I_1 \subseteq \ldots \subseteq I_n \subseteq I_{n+1} \subseteq \ldots$$

there exist m such that $I_m = I_{m+1} = I_{m+2} = \ldots$.

Example 220. All fields are Noetherian. In fact, by Proposition 201 any field A has only two ideals, (0) and A, so any ascending chain of ideals looks like $(0) \subsetneq A$.

Example 221. $\mathbb{Z}$ is Noetherian. In fact, $(m) \subsetneq (d)$ in $\mathbb{Z}$ if and only if d divides m strictly. So every ascending chain of ideals eventually stops, basically because any number m has only finitely many divisors. Note that in $\mathbb{Z}$ there are infinite descending chains of ideals.

Example 222. If $\mathbb{K}$ is a field, then $\mathbb{K}[x]$ is Noetherian. In fact, we have seen in Theorem 217 that ideals in $\mathbb{K}[x]$ are principal. Now, $(f) \subsetneq (g)$ in $\mathbb{K}[x]$ if and only if $f = g \cdot h$, with h not invertible. This implies that $h \notin \mathbb{K}$, because $\mathbb{K}$ is a field and every nonzero element of a field is invertible. So $\deg h \geq 1$. Since we are in a domain, $\deg f = \deg g + \deg h$, hence

$$\deg g = \deg f - \deg h \leq \deg f - 1.$$

So every ascending chain of ideals eventually stops, basically because every polynomial f has finite degree.

Non-Example 223. The set of real-valued functions $A \stackrel{\text{def}}{=} \{f : \mathbb{R} \to \mathbb{R}\}$ is a C-ring (though not a domain) with respect to pointwise sum and multiplication. Consider the ideals

$$I_n \stackrel{\text{def}}{=} \{f : \mathbb{R} \to \mathbb{R} \text{ such that } f(x) = 0 \text{ for each } x \in [n, \infty)\}.$$

Clearly $I_n \subsetneq I_{n+1}$, because if f is zero on $[n, \infty)$, in particular it is zero on $[n+1, \infty)$ (but the converse is false, e.g. the function that is identically zero on $[n+1, \infty)$ and identically 1 on $(-\infty, n+1)$ is in I_{n+1} but not in I_n). Since this ascending chain never stabilizes, A is not Noetherian.

Non-Example 224 (The domain $\mathbb{Q}[x_1, \ldots, x_\infty]$)**.** There is an obvious identification of $\mathbb{Q}$ as the subring of $\mathbb{Q}[x]$ formed by the constant polynomials. More generally, there's an obvious identification of $\mathbb{Q}[x_1, \ldots, x_n]$ as the subring of $\mathbb{Q}[x_1, \ldots, x_{n+1}]$ formed by the polynomials "not containing the variable x_{n+1}". Now, consider the set

$$\mathbb{Q}[x_1, \ldots, x_\infty] \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} \mathbb{Q}[x_1, \ldots, x_n].$$

No need to be scared: Any element f of this set $\mathbb{Q}[x_1, \ldots, x_\infty]$ belongs to some $\mathbb{Q}[x_1, \ldots, x_n]$, so such f is a normal polynomial with finitely many variables. Moreover, for any two elements f, g of $\mathbb{Q}[x_1, \ldots, x_\infty]$, there will be some integer m such that f, g both belong to $\mathbb{Q}[x_1, \ldots, x_m]$; so it makes sense to speak of $f + g$ and fg inside this $\mathbb{Q}[x_1, \ldots, x_m]$; but since $\mathbb{Q}[x_1, \ldots, x_m]$ is a subring of $\mathbb{Q}[x_1, \ldots, x_n]$ for all $n \geq m$, the definitions of "$f + g$" and "fg" do not depend on the m chosen. This way we can naturally define a sum operation and a product operation on $\mathbb{Q}[x_1, \ldots, x_\infty]$.

It is easy to see that $\mathbb{Q}[x_1, \ldots, x_\infty]$ is a C-ring and even a **domain** with these operations. The verification is left to you, but here is a guideline: if f, g, h are three polynomials in $\mathbb{Q}[x_1, \ldots, x_\infty]$, then there exists some m such that f, g, h all belong to $\mathbb{Q}[x_1, \ldots, x_m]$; and since $f(gh) = (fg)h$ in $\mathbb{Q}[x_1, \ldots, x_m]$, the same equality holds in $\mathbb{Q}[x_1, \ldots, x_\infty]$, which shows the associativity of the product. Another example: if $f \cdot g = 0$ in $\mathbb{Q}[x_1, \ldots, x_\infty]$, then $f \cdot g = 0$ in some $\mathbb{Q}[x_1, \ldots, x_m]$, which is a domain; so either $f = 0$ or $g = 0$ in $\mathbb{Q}[x_1, \ldots, x_m]$; but the same equalities hold in $\mathbb{Q}[x_1, \ldots, x_\infty]$, proving that $\mathbb{Q}[x_1, \ldots, x_\infty]$ is a domain. However, the infinite chain of ideals

$$(0) \subsetneq (x_1) \subsetneq (x_1, x_2) \subsetneq \ldots \subsetneq (x_1, \ldots, x_n) \subsetneq (x_1, \ldots, x_n, x_{n+1}) \subsetneq \ldots$$

never stabilizes: so $\mathbb{Q}[x_1, \ldots, x_\infty]$ is **not Noetherian**.

Proposition 225. *Let A be a C-ring. Then*

$$A \text{ is Noetherian} \iff \text{Every ideal of } A \text{ is finitely generated.}$$

Proof.
"$\Leftarrow$" Let $I_0 \subseteq I_1 \subseteq \ldots \subseteq I_n \subseteq I_{n+1} \subseteq \ldots$ be an ascending chain of ideals. Consider

$$U \stackrel{\text{def}}{=} \bigcup_{j \in \mathbb{N}} I_j.$$

Because the ideals are nested, we claim that U is an ideal. In fact, if $x, y \in U$, then $x \in I_k$ for some k and $y \in I_h$ for some h; so if we set $m = \max(k, h)$, both x and y are in the ideal I_m. So $x - y$ and xy are both in I_m. But $I_m \subseteq U$; so $x - y$ and xy are in U.
Since all ideals are finitely generated, there exists $a_1, \ldots, a_k$ in A such that $U = (a_1, \ldots, a_k)$. Since these k elements are in $U = \bigcup_{j \in \mathbb{N}} I_j$, for each i there exists m_i such that $a_i \in I_{m_i}$. If m is the largest of these m_i, all of $a_1, \ldots, a_k$ are in I_m. But then $U \subseteq I_m$, because any $x \in U$ is generated by $a_1, \ldots, a_k$ and therefore is in I_m. It follows that

$$I_m = I_{m+1} = I_{m+2} = \ldots = U.$$

"$\Rightarrow$" Suppose that an ideal I is not finitely generated. Pick g_1 in I. Set

$$I_1 \stackrel{\text{def}}{=} (g_1) \subseteq I.$$

Since I is not finitely generated, it cannot be that $I_1 = I$; so we can find g_2 that is in I but not in g_1. Again,

$$I_2 \stackrel{\text{def}}{=} (g_1, g_2) \subsetneq I,$$

so we can find $g_3 \in I \setminus I_2$. And so on. This creates an infinite ascending chain of (finitely generated) ideals. □

Corollary 226. *If A is a PID, then A is Noetherian.*

Proof. Principal ideals are finitely generated. □

How about the converse? This question is easy to dismiss:

Example 227. For any m, the C-ring $\mathbb{Z}_m$ is Noetherian, because it is finite. In particular, some Noetherian C-rings are not domains.

A more interesting question is whether all Noetherian *domains* are PID. The answer is still no: a counterexample is given, for example, by $\mathbb{Z}[x]$. The reason why $\mathbb{Z}[x]$ is Noetherian is the following wonderful result of Hilbert, called "Hilbert's Basis Theorem" or – with the original German expression — "Hilberts Basissatz". David Hilbert published it in 1888, when he was 26. This publication almost put an end to an entire branch of mathematics, called "classical invariant theory". Hilbert sent the article containing his theorem to the leading math journal of the time, *Mathematische Annalen*. The journal sent it to two referees, the main expert of invariant theory, Paul Gordan, and the famous geometer Felix Klein. Gordan which suggested rejection, famously claiming "*das ist nicht Mathematik; das ist Theologie*" ("this is not math; it's theology"). Luckily Klein pushed for the publication of the article, which appeared in 1890. Thanks also to this result, in 1895 Hilbert obtained a Professorship at the University of Göttingen.

Theorem 228 (Hilbert's Basis Theorem)**.** *Let A be a C-ring with* 1*.*

$$A \text{ is Noetherian} \iff A[x] \text{ is Noetherian.}$$

Proof.

"$\Leftarrow$" Let J be any ideal of A. By Proposition 225, we want to show that J is finitely generated. Consider the ideal of polynomials whose constant term is in J:

$$I \stackrel{\text{def}}{=} \{\, g \in A[x] \text{ such that } g(0) \in J \,\}.$$

Let us verify that I is an ideal: If $g(0) \in J$ and $h(0) \in J$, certainly $(h-g)(0) = h(0) - g(0)$ is in J, because J is an ideal. Similarly, if $g(0) \in J$ and f is an arbitrary polynomial of $A[x]$, then $f \cdot g(0) = f(0) \cdot g(0)$ is in J, again because J is an ideal.
Since $A[x]$ is Noetherian, there are polynomials $g_1, \ldots, g_m$ such that $I = (g_1, \ldots, g_m)$. Consider the constant terms of these polynomials,

$$c_i \stackrel{\text{def}}{=} g_i(0) \text{ for all } i \in \{1, \ldots, m\}.$$

We claim that J is finitely generated in A by $c_1, \ldots, c_m$. In fact:

- Because each g_i is in I, each c_i is in J. So $(c_1, \ldots, c_m) \subseteq J$.
- Let $a \in J$. The polynomial $g = x + a$ has the property that $g(0) = a$. Hence $g \in I$. Since $I = (g_1, \ldots, g_m)$, we can find polynomials $f_1, \ldots, f_m$ in $A[x]$ such that

  $$g = f_1 \cdot g_1 + f_2 \cdot g_2 \ldots + f_m \cdot g_m.$$

  Evaluating at $x = 0$, we obtain

  $$a = f_1(0) \cdot c_1 + f_2(0) \cdot c_2 + \ldots f_m(0) \cdot c_m,$$

  which proves that $a \in (c_1, \ldots, c_m)$. So $J \subseteq (c_1, \ldots, c_m)$.

"$\Rightarrow$" By contradiction, let I be an ideal of $A[x]$ that is not finitely generated. Clearly $I \neq (0)$. Let f_0 be a minimum-degree polynomial in I. Since I is not finitely generated, $(f_0) \subsetneq I$; so let f_1 be a polynomial of smallest degree among those in $I \setminus (f_0)$. Again, $(f_0, f_1) \subsetneq I$, so we can choose f_2 of minimal degree in $I \setminus (f_0, f_1)$, and so on. Note that the integers

$$\deg f_0,\ \deg f_1,\ \deg f_2, \ldots$$

form a non-decreasing sequence. Let ℓ_n be the leading coefficient of f_n. Consider the following ascending chain of ideals of A:

$$(\ell_0) \subseteq (\ell_0, \ell_1) \subseteq (\ell_0, \ell_1, \ell_2) \subseteq \ldots$$

Since A is Noetherian, the chain eventually stabilizes. So there is an integer N such that $(\ell_0, \ell_1, \ldots, \ell_N)$ contains also ℓ_{N+1}. This means that we can find elements $a_i \in A$ such that

$$\ell_{N+1} = a_0 \cdot \ell_0 + a_1 \cdot \ell_1 + \ldots + a_N \ell_N.$$

Now consider the polynomial

$$g \stackrel{\text{def}}{=} a_0 \cdot (x^{\deg f_{N+1} - \deg f_0}) \cdot f_0 + a_1 \cdot (x^{\deg f_{N+1} - \deg f_1}) \cdot f_1 + \ldots + a_N \cdot (x^{\deg f_{N+1} - \deg f_N}) \cdot f_N.$$

The summands above are polynomials of the same degree, namely, $\deg f_{N+1}$. The leading coefficient of g is by construction $a_0 \cdot \ell_0 + a_1 \cdot \ell_1 + \ldots + a_N \ell_N$, which equals ℓ_{N+1}. In other

words, we constructed a polynomial $g \in (f_0, f_1, \ldots, f_N)$ with exactly the same degree and the same leading coefficient of f_{N+1}. As a consequence, $f_{N+1} - g$ has lower degree, because the leading coefficients cancel out. Since f_{N+1} was chosen of minimal degree among the polynomials outside $(f_0, f_1, \ldots, f_N)$, it follows that $f_{N+1} - g$ must be in $(f_0, f_1, \ldots, f_N)$. But by definition, $g \in (f_0, f_1, \ldots, f_N)$, yet the sum $(f_{N+1} - g) + g = f_{N+1}$ is not in the ideal $(f_0, f_1, \ldots, f_N)$: A contradiction with the definition of ideal. □

Remark 229. A more elegant way to show "⇐" will arise from the content of the next chapter, once we introduce quotients. In fact, one can easily prove that the quotient of any Noetherian C-ring is again Noetherian, and then conclude by observing that A is isomorphic to the quotient of $A[x]$ by the ideal (x).

Corollary 230. *The domains $\mathbb{Z}[x]$, $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$ are all Noetherian.*

Corollary 231. $\mathbb{Z}[x, y] \stackrel{\text{def}}{=} \mathbb{Z}[x][y]$ *is again Noetherian. Same for $\mathbb{Q}[x, y]$, $\mathbb{R}[x, y]$, $\mathbb{C}[x, y]$.*

Corollary 232. *If A is Noetherian, then $A[x_1, \ldots, x_n]$ is Noetherian for any n.*

Remark 233. $\mathbb{Q}[x, y]$, $\mathbb{R}[x, y]$, $\mathbb{C}[x, y]$, $\mathbb{C}[x_1, \ldots, x_n]$ are examples of C-rings that are Noetherian but not PID. One may wonder if there is an upper bound on the number of generators of any ideal in the C-ring. The answer is no. For example, in $\mathbb{C}[x, y]$ one can prove that the ideal

$$J_n \stackrel{\text{def}}{=} \left(x^n,\ x^{n-1}y, \ldots,\ xy^{n-1}, y^n\right),$$

above defined by $n + 1$ generators, cannot be generated with n generators or less.

3.7 Exercises

1. Prove that the subring of a domain is always a domain.
2. Is the subring of a PID always a PID? (Hint: think of subrings of $\mathbb{Q}[x]$.)
3. Let A be a C-ring and I an ideal of A. Let

   $$\operatorname{rad} I = \{a \in A \text{ such that } a^n \in I \text{ for some natural number } n\}.$$

   Prove that $\operatorname{rad} I$ is an ideal. This is called the **radical** of I. Show that $\operatorname{rad} \operatorname{rad} I = \operatorname{rad} I$.
4. Let A be a C-ring; let N be the set of all nilpotent elements of A. Is N an ideal of A? What about the set of all zero-divisors?
5. Let $f : A \to B$ be a nonzero homomorphism between C-rings with 1. Show that if B is a domain, then $f(1) = 1$.
6. Let $f : A \to B$ be a homomorphism between fields. Show that either f is injective, or f is trivial.
7. How many ideals with less than 10 elements does $\mathbb{Z}$ have?
8. Show that a domain A is a field if and only if A has finitely many ideals.
9. Show that if A and B are Noetherian C-rings, then so is their Cartesian product

   $$A \times B \stackrel{\text{def}}{=} \{(a, b) \text{ such that } a \in A, b \in B\}.$$

4 Quotient C-rings; prime, maximal, and radical ideals

4.1 Quotient C-rings

Let A be a C-ring. Given any ideal I of A, we can define a relation $\overset{\text{I}}{\sim}$ on A as follows:

$$a \overset{\text{I}}{\sim} b \overset{\text{def}}{\iff} a - b \in I.$$

This is an equivalence relation, because (REL1) $a - a$ is in I, (REL2) if $a - b$ is in I so is $b - a$, and (REL3) if $a - b$ is in I and $b - c$ is in I, so is their sum $a - c$. Hence, we can form the quotient ${}^{A}\!/_{\overset{\text{I}}{\sim}}$. To make the notation lighter, we omit the tilde, calling the quotient ${}^{A}\!/_{I}$. Just to reiterate: **By definition, two elements a, b of A are equal in the quotient ${}^{A}\!/_{I}$ (i.e. $\overline{a} = \overline{b}$) if and only if $a - b \in I$.**

So far, ${}^{A}\!/_{I}$ has been defined as a set. But indeed, it turns out to be a C-ring.

Theorem 234. *The quotient ${}^{A}\!/_{I}$ is a C-ring when equipped with the following operations:*

$$\begin{aligned} \overline{a} \oplus \overline{b} &\overset{\text{def}}{=} \overline{a + b}, \text{ and} \\ \overline{a} \odot \overline{b} &\overset{\text{def}}{=} \overline{a \cdot b}. \end{aligned}$$

Moreover, if A is a C-ring with 1, so is ${}^{A}\!/_{I}$.

Proof. First of all we need to verify that these two operations are well-defined. Suppose that $\overline{a} = \overline{a'}$ and $\overline{b} = \overline{b'}$. Is it true that $\overline{a + b} = \overline{a' + b'}$? If not, $\oplus$ would have two different results, and so it would not be a legal operation. Similarly, we have to check whether $\overline{a \cdot b} = \overline{a' \cdot b'}$. Let us do it! Remember $\overline{a} = \overline{a'}$ means that $a - a' \in I$, and $\overline{b} = \overline{b'}$ means that $b - b' \in I$. Now

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

is the sum of two elements in I, so it is also in I; hence, we verified that $\overline{a + b} = \overline{a' + b'}$. Verifying $\overline{a \cdot b} = \overline{a' \cdot b'}$ is a bit more complicated. The trick is to write

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

which is the sum of two elements in I (by property I2), so it is in I.

Now that we know the operations $\oplus$ and $\odot$ make sense, we have eight axioms to check.

(R0) The operations are *internal*. In fact, since $a + b$ and $a \cdot b$ are both in A, indeed $\overline{a + b}$ and $\overline{a \cdot b}$ are elements of ${}^{A}\!/_{I}$.

(R1) The operation $\oplus$ is *associative*:

$$\overline{a} \oplus (\overline{b} \oplus \overline{c}) \overset{\text{def}}{=} \overline{a} \oplus \overline{b + c} \overset{\text{def}}{=} \overline{a + (b + c)} = \overline{(a + b) + c)} = \overline{a + b} \oplus \overline{c} = (\overline{a} \oplus \overline{b}) \oplus \overline{c}.$$

(R2) The operation $\oplus$ is *commutative*:

$$\overline{a} \oplus \overline{b} \overset{\text{def}}{=} \overline{a + b} = \overline{b + a} = \overline{b} \oplus \overline{a}.$$

(R3) The neutral element of $\oplus$ is $\overline{0}$:

$$\overline{a} \oplus \overline{0} \overset{\text{def}}{=} \overline{a + 0} = \overline{a}.$$

(R4) Every element $\overline{a}$ has an *additive inverse*, which is $\overline{-a}$. In fact,

$$\overline{a} \oplus \overline{-a} \stackrel{\text{def}}{=} \overline{a + (-a)} = \overline{0}.$$

(R5) The operation $\odot$ is *associative*:

$$\overline{a} \odot (\overline{b} \odot \overline{c}) \stackrel{\text{def}}{=} \overline{a} \odot \overline{b \cdot c} \stackrel{\text{def}}{=} \overline{a \cdot (b \cdot c)} = \overline{(a \cdot b) \cdot c)} = \overline{a \cdot b} \odot \overline{c} = (\overline{a} \odot \overline{b}) \odot \overline{c}.$$

(R6) The operation $\odot$ is *commutative*:

$$\overline{a} \odot \overline{b} \stackrel{\text{def}}{=} \overline{a \cdot b} = \overline{b \cdot a} = \overline{b} \odot \overline{a}.$$

(R7) We have *distributivity*:

$$\overline{a} \odot (\overline{b} \oplus \overline{c}) \stackrel{\text{def}}{=} \overline{a} \odot \overline{b + c} \stackrel{\text{def}}{=} \overline{a \cdot (b + c)} = \overline{(a \cdot b) + (a \cdot c)} = \overline{(a \cdot b)} \oplus \overline{(a \cdot c)} = (\overline{a} \odot \overline{b}) \oplus (\overline{a} \odot \overline{c}).$$

Finally, if A is a C-ring with 1, then the neutral element of ${}^{A}\!/_{I}$ is simply $\overline{1}$. In fact,

$$\overline{a} \odot \overline{1} \stackrel{\text{def}}{=} \overline{a \cdot 1} = \overline{a}.$$ □

Example 235. Let $A = \mathbb{Z}$ and $I = (2)$. We can form the equivalence relation

$$a \overset{\text{I}}{\sim} b \overset{\text{def}}{\iff} a - b \in (2).$$

In fact, we have already seen this equivalence relation in Example 28: This is congruence mod 2, and the quotient ${}^{Z}\!/_{(2)}$ is what we called $\mathbb{Z}_2$. More generally, ${}^{\mathbb{Z}}\!/_{(m)} \stackrel{\text{def}}{=} \mathbb{Z}_m$.

4.2 Homomorphism theorems.

We start by discussing a function that naturally arises whenever there is a quotient, namely, the *projection*

$$\begin{array}{rccc} p: & A & \to & {}^{A}\!/_{I} \\ & a & \mapsto & \overline{a}. \end{array}$$

Proposition 236. *Let A be a C-ring and let I be any ideal of A. The projection $p : A \to {}^{A}\!/_{I}$ is a surjective homomorphism with kernel equal to I.*

Proof. By definition,

$$\ker p = \{a \in A \text{ such that } p(a) = \overline{0}\} = \{a \in A \text{ such that } \overline{a} = \overline{0}\}.$$

Because of how we defined ${}^{A}\!/_{I}$, the projection is surjective. Because of how we defined $\oplus$ and $\odot$, the projection is a homomorphism:

$$\begin{aligned} p(a) \oplus p(b) &= \overline{a} \oplus \overline{b} \stackrel{\text{def}}{=} \overline{a + b} = p(a + b), \text{ and} \\ p(a) \odot p(b) &= \overline{a} \odot \overline{b} \stackrel{\text{def}}{=} \overline{a \cdot b} = p(a \cdot b). \end{aligned}$$

Finally, recall that two elements of A end up in the same equivalence class in ${}^{A}\!/_{I}$ if and only if their difference is in I; in particular, $\overline{a} = \overline{0}$ if and only if $a \in I$. Hence $\ker p = I$. □

Corollary 237. $\{$*ideals*$\}$ =$\{$*kernels of homomorphisms*$\}$.

Proof. We already know that the kernel of any homomorphism is an ideal. Conversely, any ideal I of some C-ring A is the kernel of the homomorphism $p : A \to {}^{A}/_{I}$. □

Theorem 238 (First Homomorphism Theorem)**.** *Let $f : A \to B$ be any homomorphism between two C-rings. Then, there exists a (unique) homomorphism*

$$g : {}^{A}/_{\ker f} \to B \qquad \textit{such that}$$

1) *g is injective;*
2) $\operatorname{Im} g = \operatorname{Im} f$*;*
3) $f = g \circ p$*.*

Proof. Let us start from the end and force property number 3) by defining

$$g(\overline{a}) \overset{\text{def}}{=} f(a) \quad \text{for all } a.$$

Is this a good definition? If a, a' are distinct elements of A such that $\overline{a} = \overline{a'}$, is it true that $f(a) = f(a')$? Well, by definition of quotient, $\overline{a} = \overline{a'}$ if and only if $a - a' \in \ker f$. So

$$\overline{a} = \overline{a'} \text{ in } {}^{A}/_{\ker f} \overset{\text{def}}{\iff} a - a' \in \ker f \iff f(a - a') = 0 \iff f(a) = f(a') \overset{\text{def}}{\iff} g(\overline{a}) = g(\overline{a'}).$$

The stream of implications from left to right tells us that g is a well-defined function; the converse implications, from right to left, tell us that g is injective. It remains to see that $\operatorname{Im} g = \operatorname{Im} f$. Let $b \in B$:

$$b \in \operatorname{Im} g \iff \exists a \in A \text{ such that } g(\overline{a}) = b \overset{\text{def}}{\iff} \exists a \in A \text{ such that } f(a) = b \iff b \in \operatorname{Im} f. \qquad \square$$

Corollary 239. *If $f : A \to B$ is a surjective ring homomorphism, then ${}^{A}/_{\ker f} \cong B$.*

Proof. Applying the previous theorem to the case $\operatorname{Im} f = B$, we know that there is an injective homomorphism $g : {}^{A}/_{\ker f} \to B$ such that $\operatorname{Im} g = B$. So g is an isomorphism. □

Corollary 240. *For any C-ring A,*

$${}^{A}/_{(0)} \cong A.$$

Proof. Apply Corollary 239 to the case when f is the identity map $id : A \to A$. Clearly this map is both surjective and injective, so $\ker f = (0)$. □

Proposition 241 (Complex numbers as quotient)**.** ${}^{\mathbb{R}[x]}/_{(x^2+1)} \cong \mathbb{C}$.

Proof. Let us consider the function

$$\begin{array}{rccc} \varphi : & \mathbb{R}[x] & \to & \mathbb{C} \\ & f & \mapsto & f(i). \end{array}$$

This map is a homomorphism: In fact, it is the restriction to $\mathbb{R}[x]$ of the "polynomial evaluation at i", which is the map from $\mathbb{C}[x]$ to $\mathbb{C}$ defined in Example 183. The map φ is also surjective, because by definition any element z of $\mathbb{C}$ can be written as

$$z = a + b \cdot i, \text{ with } a, b \in \mathbb{R}.$$

Hence the polynomial $a + b \cdot x$ is mapped to z under φ. So $z \in \operatorname{Im} \varphi$, and by the genericity of z, the map φ is surjective. By the first isomorphism theorem (Corollary 239),

$${}^{\mathbb{R}[x]}/_{\ker \varphi} \cong \mathbb{C}.$$

At this point we are almost done: All we need to prove is $\ker \varphi = (x^2 + 1)$.

- $\ker\varphi \supset (x^2+1)$, because if $g = h(x^2+1)$ for some polynomial h, when we plug in $x = i$ we get $g(i) = h(i)\cdot(-1+1) = 0$. So $\varphi(g) = 0$.
- $\ker\varphi \subseteq (x^2+1)$: Suppose $\varphi(g) = 0$ for some polynomial $g \in \mathbb{R}[x]$. This means that $g(i) = 0$, so i is a root of g. Since g has all real coefficients, by Corollary 158 also the conjugate $-i$ is a root. So g is a multiple of $(x-i)(x+i)$, which equals x^2+1. □

Proposition 242 (Gaussian integers)**.** *If* $\mathbb{Z}[i] \stackrel{\text{def}}{=} \{a + b\cdot i \text{ with } a, b \in \mathbb{Z}\}$*, then*

$$\mathbb{Z}[x]/(x^2+1) \cong \mathbb{Z}[i].$$

Proof. Analogous to the previous one. □

Proposition 243. *Let A be a C-ring with 1. Fix any $a \in A$. Then*

$$A[x]/(x-a) \cong A.$$

Proof. The tactic is to find a surjective homomorphism $\varphi : A[x] \to A$ such that $\ker\varphi = (x-a)$, from which we would conclude using the first isomorphism theorem (Corollary 239). The natural choice is to define

$$\begin{array}{rccc} \varphi_a : & A[x] & \to & A \\ & f & \mapsto & f(a), \end{array}$$

the evaluation at a (or in plain words, the map that plugs in $x = a$). Recall that it is surjective because for any $b \in A$, we can view b as a constant polynomial in $A[x]$, and plugging in $x = a$ will leave b unchanged; so $\varphi_a(b) = b$. It remains to prove $\ker\varphi_a = (x-a)$.

- $\ker\varphi_a \supset (x-a)$, because if $g = h\cdot(x-a)$, when we plug in $x = a$ we get $g(a) = h(a)\cdot(a-a) = 0$. So $\varphi_a(g) = 0$.
- $\ker\varphi_a \subseteq (x-a)$: Let $g \in \ker\varphi_a$. This means $g(a) = 0$. But by Ruffini's theorem (Thm. 144), if $g(a) = 0$ then g is a multiple of $x-a$.

□

Proposition 244. *Let A be a C-ring with 1. Fix any $a, b \in A$. Then*

$$A[x]/(x-a, b) \cong A/(b).$$

Proof. Let φ_a the map defined in the previous proof,

$$\begin{array}{rccc} \varphi_a : & A[x] & \to & A \\ & f & \mapsto & f(a). \end{array}$$

Let p_b be the projection to the quotient

$$\begin{array}{rccc} p_b : & A & \to & A/(b) \\ & y & \mapsto & \overline{y}. \end{array}$$

Consider the composition $\gamma \stackrel{\text{def}}{=} p_b \circ \varphi_a$. This is a homomorphism (cf. Lemma 189) from $A[x]$ to $A/(b)$. It is surjective, because both p_b and φ_a are, and the composition of two surjective maps is surjective. It remains to show that $\ker\gamma = (X-a, b)$.

- $\ker\gamma \supset (x-a, b)$: If $f = g\cdot b + h\cdot(x-a)$, when we plug in $x = a$ we get

  $$f(a) = g(a)\cdot b,$$

  which is a multiple of b. So when we apply p_b to that, we get zero. So $p_b(\varphi_a(f)) = 0$, which means that $f \in \ker\gamma$.

- $\ker\gamma \subseteq (x-a,b)$: Let $f \in \ker\gamma$. This means

$$0 = \gamma(f) \stackrel{\text{def}}{=} p_b(\,\varphi_a(f)\,) \stackrel{\text{def}}{=} p_b(\,f(a)\,).$$

By definition of p_b, the fact that $0 = p_b(f(a))$ means that $f(a) \in (b)$. In other words, $f(a)$ is a multiple of b. Now we can apply Ruffini's theorem (Thm. 144) to the polynomial f:

$$f = q \cdot (x-a) + f(a).$$

This expresses f as a multiple of $(x-a)$ plus a multiple of b. So $f \in (x-a,b)$. □

4.3 Operations with ideals

Let us start with a simple observation.

Lemma 245. *Let A be a C-ring. Let I, J be ideals of A. The* intersection

$$I \cap J$$

is also an ideal. It is the largest ideal contained in both I and J.

Proof. Let x, y be in $I \cap J$. Then x, y is in I, which is an ideal; so both $x-y$ and xy are in I. Analogously for J. So $x-y$ and xy are in $I \cap J$. This shows that $I \cap J$ is an ideal. Now suppose H is another ideal such that $H \subseteq I$ and $H \subseteq J$. Then $H \subseteq I \cap J$. □

Careful: while the *intersection* is always an ideal, the *union* need not be!

Non-Example 246. In $\mathbb{Z}$, consider the set $W = (2) \cup (3)$. It does not contain 5, because 5 is neither a multiple of 2 nor of 3. So $2, 3 \in W$ but $2+3 \notin W$. Hence, W is not an ideal.

So is there a smallest ideal containing both I and J, somehow balancing Lemma 245? The answer is positive.

Lemma 247. *Let A be a C-ring. Let I, J be ideals of A. The* sum

$$I + J \stackrel{\text{def}}{=} \{i + j \text{ such that } i \in I, j \in J\}$$

is also an ideal. It is the smallest ideal containing both I and J.

Proof. Let $x = i + j$ and $x' = i' + j'$ be elements of $I + J$, with i, i' in I and j, j' in J. The difference can be written as

$$x - y = (i - i') + (j - j'),$$

which is the sum of an element in I and of an element in J. Furthermore, for any a in the C-ring A

$$ax = a(i+j) = ai + aj$$

which is the sum of an element of I and an element of J. So both $x-y$ and ax are in $I+J$. So $I+J$ is an ideal. As for the second claim, choose any ideal H containing I and J. Then for any $i \in I$ and for any $j \in J$, we have that i, j are in H; so H must contain also their sum $i+j$. □

Proposition 248. *If $I = (g_1, \ldots, g_m)$ and $J = (h_1, \ldots, h_n)$, then $I+J = (g_1, \ldots, g_m, h_1, \ldots, h_n)$.*

Proof.

"$\subseteq$". Let $i+j$ be a generic element of $I+J$. By the assumption, we can write

$$i = a_1 g_1 + \ldots + a_m g_m \qquad \text{and} \qquad j = b_1 h_1 + \ldots + b_n h_n,$$

for suitable coefficients a_k, b_k in A. So the sum $i+j$ is in $(g_1, \ldots, g_m, h_1, \ldots, h_n)$.

"$\supseteq$" If $x = [a_1 g_1 + \ldots + a_m g_m] + b_1 h_1 + \ldots + b_n h_n$ is the generic element of $I+J$, then the part of the sum between brackets is in I, while the remaining part is in J. □

***Product Ideals**

There is a third operation, commonly used in commutative algebra.

Lemma 249. *Let A be a C-ring. Let I, J be ideals of A. The* product

$$I \cdot J \stackrel{\text{def}}{=} \left\{ \sum_{k=1}^{n} i_k \cdot j_k \text{ such that } i_k \in I,\, j_k \in J,\, n \in \mathbb{N} \right\}$$

is also an ideal. It is the smallest ideal containing all 'mixed' products of one element from I and one from J.

Proof. Let x, y be in $I \cdot J$. Then we can write

$$x = \sum_{k=1}^{n} i_k j_k \qquad \text{and } y = \sum_{h=1}^{m} a_h \cdot b_h, \qquad \text{with } i_k, a_h \in I,\, j_k, b_h \in J.$$

So if we relabel $i_{n+h} \stackrel{\text{def}}{=} -a_h$ and $j_{n+h} \stackrel{\text{def}}{=} -b_h$ for all $h \in \{1, 2, \ldots, m\}$, we have that

$$x = \sum_{k=1}^{n} i_k j_k \qquad \text{and } -y = \sum_{h=1}^{m} i_{n+h} \cdot j_{n+h} = \sum_{k=n+1}^{n+m} i_k \cdot j_k, \qquad \text{with } i_k \in I, \text{ and } j_k \in J.$$

It follows that

$$x - y = \sum_{k=1}^{n+m} i_k j_k$$

is an element of $I \cdot J$. As for $x \cdot y$, there is no need for such tedious computations: By the property (I2) of ideals, $x \in I$. For the same reason, $y \in J$. So $x \cdot y$ is by definition in $I \cdot J$.

As for the final claim, any ideal H that contains all 'mixed' products of one element from I and one from J, must contain also their sum. So H contains $I \cdot J$. □

Proposition 250. *Let A be a C-ring. Let I, J be ideals of A.*

$$(I + J) \cdot (I \cap J) \subseteq I \cdot J.$$

Proof. Let $x = i + j$ be a generic element of $I + J$. Let z be an element of $I \cap J$. Then writing

$$x \cdot z = i \cdot z + z \cdot j$$

we see that $x \cdot z$ is an element of $I \cdot J$. Because the generic element s of $(I + J) \cdot (I \cap J)$ is the sum of finitely many elements like $x \cdot z$, and $I \cdot J$ contains each summand, it must contain the sum s. □

Note that it follows from the definition of product ideals that

$$I \cdot J \subseteq I \cap J.$$

The containment is usually strict: For example, in $\mathbb{Z}$, if $I = J = (2)$, then $I \cdot I = (4)$, which is strictly contained in $I \cap I = I = (2)$.

However, in some cases one has equality.

Lemma 251. *For any ideal I in a C-ring A with 1,*

$$I \cdot A = I = A \cdot I.$$

Proof. Clearly $I \cdot A \subseteq I \cap A = I$. To show that $I \subseteq I \cdot A$, simply write down any element i of I as $i \cdot 1$. Symmetrically, if we write $i = 1 \cdot i$ we see that any element i of I belongs to $A \cdot I$. □

Two ideals I, J in a C-ring A are called *coprime* if $I + J = A$.

Corollary 252. *Let A be a C-ring with 1. Let I, J be ideals of A. If I, J are coprime, then*

$$I \cdot J = I \cap J.$$

Proof. All we need to show is the inclusion $I \cap J \subseteq I \cdot J$. Since $I + J = A$, Proposition 250 tells us that

$$A \cdot (I \cap J) \subseteq I \cdot J,$$

but $A \cdot (I \cap J)$ is simply $(I \cap J)$. □

Remark 253. The previous implication cannot be reversed. In $A = \mathbb{R}[x, y]$ the ideals (x) and (y) are not coprime, because $(x)+(y) = (x, y)$ is a proper ideal. (It is the ideal of all polynomials 'whose constant term is zero'.) However, it is true that

$$(x) \cap (y) = (xy) = (x) \cdot (y).$$

4.4 Prime ideals and quotients

Definition 254. Let A be a C-ring with 1. An ideal $P \subseteq A$ is called *prime* if it satisfies the following properties:
(P1): P is *proper*, that is, $P \neq A$.
(P2): If $a \cdot b \in P$, then either $a \in P$ or $b \in P$.

Note that (P2) is the converse of (I2), the second property in the definition of "ideal". Hence, (P2) could be equivalently replaced by

"any product of two elements is in P if and only if at least one of them is in P".

The next Proposition explains why we call them "prime" ideals.

Proposition 255. *(n) is a prime ideal of $\mathbb{Z} \iff$ either $n = 0$ or n is a prime number.*

Proof. Up to replacing n by $-n$, we can assume $n \geq 0$.
"⇒" Prime ideals must be proper, so $n \neq 1$. If $n = 0$ there is nothing to show. So, suppose $n > 1$. If n is not prime, then $n = a \cdot b$, with $1 < a < n$. Thus $a \notin (n)$. Since $1 < b < n$, also $b \notin (n)$. But $a \cdot b = n \in (n)$. So by definition (n) is not a prime ideal.
"⇐" Since $n \neq 1$, the ideal (n) is proper. Now suppose $a \cdot b \in (n)$. If $n = 0$, this means $ab = 0$, but since $\mathbb{Z}$ is domain, either $a = 0$ or $b = 0$. So (0) is prime. If instead n is a prime number and n divides ab, by Euclid's Lemma 13 n must divide either a or b. In other words, either $a \in (n)$ or $b \in (n)$, so again (n) is prime. □

Theorem 256. *Let A be any C-ring with 1. Let I be any ideal of A. Then*

$$I \text{ is prime} \iff {}^{A}\!/_{I} \text{ is a domain.}$$

Proof.
"⇐" Clearly I is proper, otherwise ${}^{A}\!/_{I} = (0)$ would not be a domain. (Check again Definition 114: The C-ring (0) is excluded from the definition.) Say $ab \in I$. Then $\overline{ab} = \overline{0}$ in ${}^{A}\!/_{I}$. But $\overline{ab} = \overline{a} \odot \overline{b}$ and ${}^{A}\!/_{I}$ is a domain; it follows that either $\overline{a} = \overline{0}$ or $\overline{b} = \overline{0}$.

"⇒" Suppose $\overline{a} \odot \overline{b} = \overline{0}$ in A/I. This means that $\overline{ab} = \overline{0}$ in A/I. So $ab \in I$. Since I is prime, either $a \in I$ (which implies $\overline{a} = \overline{0}$) or $b \in I$ (which implies $\overline{b} = \overline{0}$). □

Corollary 257. *In any domain D, the ideal (0) is prime.*

Proof. Since $D/(0) \cong D$ is a domain, by Theorem 256 the ideal (0) is prime. □

4.5 Maximal ideals and quotients

Definition 258. Let A be a C-ring with 1. An ideal $M \subseteq A$ is called *maximal* if it satisfies the following properties:
(M1): M is *proper*, that is, $M \neq A$.
(M2): If $M \subsetneq J \subseteq A$ for some J ideal, then $J = A$.

Property (M2) can be conveniently rephrased as "No other proper ideal of A contains M." This explains the terminology "maximal".

Non-Example 259. (9) is not maximal in $\mathbb{Z}$. In fact,

$$(9) \subsetneq (3) \subsetneq \mathbb{Z}.$$

Example 260. (3) is maximal in $\mathbb{Z}$. (M1) is obvious; let us check (M2). Suppose that

$$(3) \subsetneq J \text{ for some ideal } J.$$

We know that $J = (n)$ for some $n > 0$, because all ideals of $\mathbb{Z}$ are of this form (Prop. 198). Since $(3) \subseteq (n)$, we have that $3 \in (n)$, so n must divide 3. But 3 is prime, so it has only itself and 1 as divisors. So there are two cases:

- either $n = 3$, a contradiction with $(3) \subsetneq J$;
- or $n = 1$, in which case $J = \mathbb{Z}$.

Proposition 261. *Let $n \in \mathbb{N}$. An ideal (n) is maximal in $\mathbb{Z}$ if and only if n is a prime number.*

Proof. Exercise. □

Non-Example 262. In $\mathbb{Z}[x]$, the ideal $(x) \stackrel{\text{def}}{=} \{ax \text{ such that } a \in \mathbb{Z}[x]\}$ is not maximal, because

$$(x) \subsetneq (x, 2) \subsetneq \mathbb{Z}[x].$$

(See also Remark 266 below.)

Example 263. In $\mathbb{Q}[x]$, the ideal $(x) \stackrel{\text{def}}{=} \{bx \text{ such that } b \in \mathbb{Q}[x]\}$ is maximal. To see this, suppose that

$$(x) \subsetneq J \text{ for some ideal } J.$$

Let f be any polynomial in J that is not in (x). If we perform the Euclidean division

$$f = q \cdot x + r,$$

the remainder r is a constant that cannot be zero (otherwise f would be a multiple of x) and hence it is invertible, because $\mathbb{Q}$ is a field. Yet $r = f - qx$ is the difference of two elements of J. Which means that J contains r, an invertible element. So $J = \mathbb{Q}[x]$ by the Explosion Lemma 200.

Theorem 264. *Let A be any C-ring with 1. Let I be any ideal of A. Then*

$$I \text{ is maximal} \iff {}^{A}/_{I} \text{ is a field.}$$

Proof.

"$\Leftarrow$" Clearly I is proper, otherwise ${}^{A}/_{I} = (0)$ would not be a field. (A field must have at least two elements, 0 and 1.) So, suppose that $I \subsetneq J$ for some ideal J; we want to show that $J = A$. Choose any element $b \in J$ that is not in I. Then $\overline{b} \neq \overline{0}$ in ${}^{A}/_{I}$. Since the latter is a field, $\overline{b}$ is invertible. This means that there is an $a \in A$ such that

$$\overline{a} \odot \overline{b} = \overline{1} \text{ in } {}^{A}/_{I}.$$

By definition of quotient, this means that $ab - 1 \in I$. So we can write

$$1 = a \cdot b + i, \text{ for some } i \in I.$$

But $a \in J$, so $ab \in J$; and $I \subseteq J$, so also $i \in J$. It follows that $1 \in J$. So $J = A$, by Lemma 200.

"$\Rightarrow$" Let $\overline{a} \neq 0$ in ${}^{A}/_{I}$; we want to show that $\overline{a}$ is invertible. Since it is nonzero in the quotient, $a \notin I$. So consider the ideal

$$J = I + (a).$$

This is certainly larger than I, because it contains a which is not in I. Now we play the card that I was maximal to conclude that $J = A$. In particular, $1 \in J = I + (a)$! So there exists an element $i \in I$ and an element $b \in A$ such that

$$1 = i + a \cdot b.$$

But then in ${}^{A}/_{I}$ we get

$$\overline{1} = \overline{0} + \overline{a} \odot \overline{b}.$$

□

Corollary 265. *Any maximal ideal is prime.*

Proof. Let I be an ideal of a C-ring A with 1. By Theorems 256 & 264 and Proposition 116,

$$I \text{ maximal} \iff {}^{A}/_{I} \text{ field} \implies {}^{A}/_{I} \text{ domain} \iff I \text{ prime.}$$

□

Remark 266. Not every prime ideal is maximal. For example, consider the ideal (x) in $\mathbb{Z}[x]$. By Proposition 243, we know that

$${}^{\mathbb{Z}[x]}/_{(x)} \cong \mathbb{Z},$$

which is a domain but not a field. By Theorems 256 & 264, it follows that (x) is prime but not maximal.

Remark 267. We have seen in Theorem 117 that every *finite* domain is a field. So if we know that ${}^{A}/_{I}$ is finite, the only "$\Longrightarrow$" in the proof of Corollary 265 becomes "$\Longleftrightarrow$", and we can say that I is maximal if and only if I is prime.

For example, this is the case when $A = \mathbb{Z}$. Since all ideals of $\mathbb{Z}$ are of the form (n), all quotients of $\mathbb{Z}$ are finite; so prime and maximal ideals in $\mathbb{Z}$ are the same.

We will see next another important class of rings where prime and maximal ideals are basically the same (with the exception of (0).)

Proposition 268. *Any non-zero prime ideal in a PID ring is maximal.*

Proof. If I is a nonzero ideal in a PID C-ring A, then $I = (i)$ with $i \neq 0$. Let J be any ideal such that $I \subsetneq J$. Again, since A is a PID, we can write $J = (j)$ for some j in A. We want to show that $J = A$. Since $(i) \subseteq (j)$, the element i is a multiple of j. So we can write

$$i = j \cdot a \text{ for some } a \in A. \tag{24}$$

Now we use that I is prime. Since $j \cdot a = i$ belongs to I, either j belongs to I (which is impossible!, we wanted $I \subsetneq J$) or a belongs to I. So a has to be in $I = (i)$: which means,

$$a = b \cdot i \text{ for some } b \in A. \tag{25}$$

Putting together Equations 24 and 25, we get

$$i = j \cdot b \cdot i.$$

Since A is a domain and $i \neq 0$, we can cancel the i and obtain

$$1 = j \cdot b.$$

So j is invertible. Hence $J = A$ by the Explosion Lemma (Lemma 200). □

As a corollary, we obtain that in $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$,"prime" and "maximal" ideals are the same. So **who are they**? Here is what we know so far:

- In $\mathbb{C}[x]$, the ideals of the form $(x - a)$, with $a \in \mathbb{C}$, are certainly maximal, because by Proposition 243,
$$\mathbb{C}[x]/(x-a) \cong \mathbb{C},$$
which is a field. But are there more maximal ideals?
- In $\mathbb{R}[x]$ there are more types of maximal ideals:
  1. all ideals of the form $(x-r)$, with $r \in \mathbb{R}$, are again maximal, because $\mathbb{R}[x]/(x-r) \cong \mathbb{R}$ is a field;
  2. yet also $x^2 + 1$ is maximal, because by Proposition 241 also $\mathbb{R}[x]/(x^2+1) \cong \mathbb{C}$ is a field!

4.6 Radical Ideals and quotients

Definition 269. An ideal J of a C-ring A is called *radical* if the following implication holds for any x in A and for any positive integer n:

$$x^n \in J \Rightarrow x \in J.$$

Proposition 270. *The radical ideals of $\mathbb{Z}$ are (0), (1), and those generated by products of distinct primes.*

Proof. Let $m = p_1 p_2 \cdots p_s$, with $p_i < p_{i+1}$ for all i. Then if x^n is a multiple of m, it means that all primes p_i appear in the factorization of x^n, and thus of x. So x is a multiple of m.
Conversely, suppose $m = p_1^{a_1} p_2^{a_2} \cdots p_s^{a_s}$, with $p_i < p_{i+1}$ for all i, and suppose that $a_i \geq 2$ for some i. Then consider $x \stackrel{\text{def}}{=} \frac{m}{p_i}$. Clearly x^2 is a multiple of m, but x is not. □

Recall that a C-ring is 'reduced' if the only nilpotent element is 0 (Definition 119).

Theorem 271. *Let A be any C-ring with 1. Let I be any ideal of A. Then*

$$I \text{ is radical} \iff {}^{A}\!/_{I} \text{ is reduced.}$$

Proof. Exercise. □

Proposition 272. *Prime ideals are radical, and so are the intersections of prime ideals.*

Proof. Suppose that there is an element x and an integer n such that $x^n \in P$ but $x^{n-1} \notin P$. In particular, $x \in P$. Then $x^n = x \cdot x^{n-1}$ shows that P is not prime.
As for the second part: Since prime ideals are radical, it suffices to show that the intersection of radical ideals is radical. But if $x^n \in I \cap J$, with both I and J radical, then $x \in I$ and $x \in J$. □

Definition 273. Let I be any ideal in a C-ring A. The *radical of I* is the set

$$\operatorname{rad}(I) \stackrel{\text{def}}{=} \{x \in A \text{ such that } \exists n \in \mathbb{N} \setminus \{0\} \text{ for which } x^n \in I\}.$$

Theorem 274. *In any C-ring A, for any ideal I of A, $\operatorname{rad}(I)$ is the smallest radical ideal containing I.*
In particular, $I = \operatorname{rad}(I)$ if and only if I is radical.

Proof. That $\operatorname{rad}(I) \supseteq I$ is obvious: if x is in I, then x^n is in I for all n. Let us see that $\operatorname{rad}(I)$ is an ideal. Pick a in A, and x, y in $\operatorname{rad}(I)$. So $x^m \in I$ and y^n in I for some positive integers m, n. Clearly then for any a in A, $(ax)^m = a^m x^m$ is in I. Moreover, by Newton's formula,

$$(x-y)^{m+n} = \sum_{k=0}^{m+n} \binom{m+n}{k} x^k (-y)^{m+n-k} \in I,$$

because either $k \geq m$ (in which case x^k is in I) or $k < m$, and thus $n + m - k > n$ (in which case $(-y)^{m+n-k}$ is in I). So both ax and $x - y$ are in $\operatorname{rad}(I)$, which proves that $\operatorname{rad}(I)$ is an ideal. Moreover, $\operatorname{rad}(I)$ is radical, because if x^m is in $\operatorname{rad}(I)$, it means that x^{mn} is in I for some n, and so by definition x is in $\operatorname{rad}(I)$. Finally, let R be any radical ideal containing I. For any x in $\operatorname{rad}(I)$, if n is an integer for which $x^n \in I \subseteq R$, we have that x is in R: so $\operatorname{rad}(I) \subseteq R$. We are left with the second claim, which is now trivial: If $I = \operatorname{rad}(I)$ then I is equal to a radical ideal, hence radical. Conversely, if I is radical, then obviously I is the smallest radical ideal containing I. □

Theorem 275. *Let A be a C-ring with 1. For any ideal $I \subseteq A$, the ideal $\operatorname{rad}(I)$ coincides with the intersection of all prime ideals containing I.*

Proof. Let R be the intersection of all prime ideals contaning I. By Proposition 272, R is a radical ideal containing I, so by Theorem 274 $\operatorname{rad} I \subseteq R$.

The converse is more interesting. We prove the contrapositive. Suppose $x \notin \operatorname{rad} I$. Then I contains *none* of the elements of

$$S \stackrel{\text{def}}{=} \{1, x, x^2, \ldots, x^n, x^{n+1}, \ldots, \}$$

Let Z be an ideal that contains I, is disjoint from all elements above, and is inclusion-maximal with respect to these two properties.[11]

[11] One may wonder why I am not choosing Z to be simply the complement of S. The reason is that the complement of S might not be an ideal. For example, if $A = \mathbb{Z}$ and $S \stackrel{\text{def}}{=} \{1, 3, 9, \ldots, 3^n, 3^{n+1}, \ldots, \}$, then the complement of S contains both 2 and 5, but not their difference. However, any largest possible ideal inside the complement of S is a prime ideal.

We claim that Z **is prime**. Let us prove the claim. By contradiction suppose $ab \in Z$, with $a, b \notin Z$. Then the ideal $Z + (a)$ is larger than Z; so by inclusion-maximality of Z, the ideals $Z + (a)$ and $Z + (b)$ are no longer disjoint from the list S. In other words,

$$x^m = z_1 + a_1 a \qquad \text{and} \qquad x^n = z_2 + a_2 b,$$

for some positive integers m, n, for some z_1, z_2 in Z and for some a_1, a_2 in A. But then

$$x^{m+n} = (z_1 + a_1 a)(z_2 + a_2 b) = z_1 z_2 + z_1(a_2 b) + z_2(a_1 a) + ab(a_1 a_2) \in Z,$$

because all four summands are in Z; a contradiction.

Okay, so Z is prime. But by construction, $x \notin Z$. But this means that x is not contained in one of the prime ideals that contain I. So x is not contained in the intersection of the prime ideals containing I. So $x \notin R$. □

Corollary 276. *Radical ideals are the same as intersections of prime ideals. Moreover, in any C-ring with* 1*, the set of all nilpotent elements coincides with the intersection of all prime ideals.*

Proof. It's easy to see that the intersection of arbitrarily many prime ideals is radical. The converse is given by the previous Theorem: Any radical ideal is intersection of (possibly infinitely many) prime ideals. The second claim is the previous Theorem applied to $I = (0)$. □

We now show that in Noetherian rings, every ideal is the intersection of finitely many primes.

Lemma 277. *Let A be a C-ring with* 1*. Let R be a radical ideal of A. If ab is in R, then*

$$R \quad = \quad \mathrm{rad}(R + (a)) \cap \mathrm{rad}(R + (b)).$$

Proof. The containment $\subseteq$ is obvious. As for $\supseteq$: suppose that x is in the right hand side. Then there are positive integers m, n such that x^m is in $R + (a)$, and x^n is in $R + (b)$. In other words,

$$x^m = r + ay \;\text{ and }\; x^n = s + bz, \qquad \text{for some } r, s \in R \text{ and some } y, z \in A.$$

But then

$$x^{m+n} = x^m \cdot x^n = (r + ay)(s + bz) = rs + rbz + say + abz$$

is in R, because all four summands are in R! (We used the assumption that ab is in R.) And because R is radical, $x^{m+n} \in R$ implies $x \in R$. □

Theorem 278. *Any radical ideal in a Noetherian C-ring with* 1 *is the intersection of* finitely many *prime ideals.*

Proof. Let A be a Noetherian C-ring with 1 and by contradiction, let R_0 be a radical ideal of A that is not the intersection of finitely many prime ideals. Clearly R_0 is not prime (otherwise $R_0 = R_0$ would be a way to write it as intersection of one prime ideal...). So there are elements a, b outside R_0 such that $ab \in R_0$. By Lemma 277 we can write R_0 as intersection of two radical ideals, $R \stackrel{\text{def}}{=} \mathrm{rad}(R_0 + (a))$ and $S \stackrel{\text{def}}{=} \mathrm{rad}(R_0 + (b))$. Were *both* R and S writable as intersection of finitely many prime ideals, so would R_0. So, at least one of R, S is not: call it R_1. Then, R_1 is strictly larger than R_0, and again is not the intersection of finitely many prime ideals. Now repeat the previous reasoning with R_1 instead of R_0. Eventually this contradicts the assumption that A is Noetherian, because we obtain an infinite ascending chain of radical ideals

$$R_0 \subsetneq R_1 \subsetneq R_2 \subsetneq \dots$$

□

We conclude this section with a beautiful theorem (although sadly, often ignored by textbooks) that fully solves the problem of **what polynomials are invertible in** $A[x]$.

Theorem 279. *Let A be a C-ring with 1. Let $f = a_0 + a_1x + \ldots + a_nx^n \in A[x]$. Then*

$$f \text{ is invertible} \iff a_0 \text{ is invertible and all other } a_i\text{'s are nilpotent.}$$

Proof. "$\Longleftarrow$". Clearly if a_i is nilpotent, so is a_ix^i. We claim that if u is invertible and b is nilpotent, then $u+b$ is invertible. To show the claim, pick an n for which $b^n = 0$ and apply the identity

$$1 - c^n = (1-c)(1 + c + c^2 + c^3 \ldots + c^{n-1}) \tag{26}$$

to $c \stackrel{\text{def}}{=} -bu^{-1}$. Since $c^n = -b^nu^{-n} = 0$, Equation 26 tells us that $1-c$ is invertible. But then so is $(1-c)u = u - cu = u + b$. Hence the claim is proven. The conclusion now follows because a_0 is invertible, so $a_0 + a_1x$ is invertible, so $(a_0 + a_1x) + a_2x^2$ is invertible, and so on.
"$\Longrightarrow$". That a_0 is invertible follows immediately from the fact that the constant term of a product of polynomials is the product of the constant terms. So if $fg = 1$ in $A[x]$, and b_0 is the constant term of g, then $a_0b_0 = 1$. Now, let P be any prime ideal of A. Then $B \stackrel{\text{def}}{=} {}^{A}\!/_{P}$ is a domain by Theorem 256. If $fg = 1$ in $A[x]$, then $\overline{fg} = \overline{1}$ in $B[x]$. But being B a domain, by Theorem 139 this means that $\overline{f}$ has degree 0. So for all $i > 0$, we have $\overline{a_i} = \overline{0}$. In other words, for all $i > 0$, the coefficient a_i of f is in P. But this holds for any prime P. So each a_i belongs to the intersection of all primes of $A[x]$. By Corollary 276, this means that each a_i is nilpotent. □

Corollary 280. *For a C-ring A, the following are equivalent:*

- *A is reduced.*
- *$\{$invertible elements of $A[x]\} = \{$invertible elements of $A\}$.*

Proof. Exercise. □

Since $\mathbb{Z}_m$ is reduced if and only if m is a product of distinct primes, we conclude:

Corollary 281. *If m is a product of distinct primes,*

$$\{\text{invertible elements of } \mathbb{Z}_m[x]\} = \{\text{invertible elements of } \mathbb{Z}_m\}.$$

Otherwise, "$\supsetneq$" holds.

4.7 *Domains and quotients: The field of fractions

In this section we show how to solve the following problem: If you are given a domain A, can you find a smallest field $F(A)$ of which A is a subring? The answer is positive; the main idea is to mimic the construction of $\mathbb{Q}$ from $\mathbb{Z}$. Essentially, an element $\frac{a}{b}$ of $\mathbb{Q}$ is given by a pairs of integers (a, b) with $b \neq 0$; but remember that we identify two fractions

$$\frac{a}{b} = \frac{c}{d} \qquad \text{if and only if} \qquad ad = bc.$$

Definition 282. Let A be a domain. On the set

$$X_A \stackrel{\text{def}}{=} \{(a, b) \text{ such that } a, b \in A, b \neq 0\}$$

we define the relation

$$(a, b) \sim (c, d) \stackrel{\text{def}}{\iff} ad = bc.$$

Let verify that this is indeed an equivalence relation:
(REL1) $(a,b) \sim (a,b)$ because $ab = ba$.
(REL2) If $(a,b) \sim (c,d)$, then $ad = bc$, so by commutativity $cb = da$. Hence $(c,d) \sim (a,b)$.
(REL3) Suppose $(a,b) \sim (c,d)$ and $(c,d) \sim (e,f)$. Then $ad = bc$ and $cf = de$. It follows that

$$(ad)f = (bc)f = b(cf) = b(de).$$

But we are in a commutative domain and $d \neq 0$, we can simplify the d and get $af = be$. Which tells us $(a,b) \sim (e,f)$.
So $\sim$ is indeed an equivalence relation.

Definition 283. With the notation above, we call $F(A)$ the quotient $X_A/_\sim$. On this set we define two operations as follows:

$$\begin{array}{rcl} \overline{(a,b)} \oplus \overline{(c,d)} & \stackrel{\text{def}}{=} & \overline{(ad+bc,\, bd)}, \text{ and} \\ \overline{(a,b)} \odot \overline{(c,d)} & \stackrel{\text{def}}{=} & \overline{(ac,\, bd)}. \end{array}$$

We should check that this is a good definition. Suppose that $(a,b) \sim (a',b')$ and $(c,d) \sim (c',d')$; so we know that $ab' = a'b$, and $cd' = c'd$. Is it true that $(ad+bc,\, bd) \sim (a'd' + b'c',\, b'd')$? In other words, is it true that $(ad+bc)b'd' = bd(a'd' + b'c')$? Let us check:

$$(ad+bc)b'd' = (ab')dd' + bb'(cd') = (a'b)dd' + bb'(c'd) = a'bdd' + bb'c'd = bd(a'd' + b'c').$$

Similarly, is it true that $(ac, bd) \sim (a'c', b'd')$, or in other words, is it true that $(ac)(b'd') = (bd)(a'c')$? Indeed,

$$(ac)(b'd') = (ab')(cd') = (a'b)(c'd) = (bd)(a'c').$$

So the operations are well defined. To understand what is going on, it helps greatly to imagine a pair (a,b) as a fraction $\frac{a}{b}$: This way it is clear that the operations introduced above are just the usual way to add and multiply fractions. From now on, we adopt the notation

$$\frac{a}{b} \stackrel{\text{def}}{=} \overline{(a,b)}.$$

We omit the proof of the following theorem, which can be done as (long!) exercise:

Theorem 284. *With the operations above, $F(A)$ is a* field*, where:*
- *The neutral element of the sum is $\frac{0}{b}$, where b is any nonzero element in A.*
- *The additive inverse of $\frac{a}{b}$ is $\frac{-a}{b}$.*
- *The neutral element of the product is $\frac{b}{b}$, where b is any nonzero element in A.*
- *The multiplicative inverse of an element $\frac{a}{b}$, with $a \neq 0$, is $\frac{b}{a}$.*

The field $F(A)$ is called the *field of fractions* of A.

Example 285. The field of fractions of $\mathbb{Z}$ is $\mathbb{Q}$.

If A is a C-ring with 1, there is a natural inclusion map

$$\begin{array}{rccc} \iota: & A & \to & F(A) \\ & a & \mapsto & \frac{a}{1}. \end{array}$$

For example, we can identify every element z of $\mathbb{Z}$ with a fraction with denominator 1.

If A is a C-ring without 1, it's no problem! We can still define the same map with a trick, by setting

$$\begin{array}{rccc} \iota: & A & \to & F(A) \\ & a & \mapsto & \frac{a^2}{a}. \end{array}$$

We still have to show that $F(A)$ is the *smallest* field containing A. In other words, if a field $\mathbb{K}$ contains A, we want to prove that $\mathbb{K}$ contains $F(A)$ too. The formal way to write this down is using injective maps, instead of inclusions.

Theorem 286. *Let A be a domain. Let $\mathbb{K}$ be a field such that $j : A \to \mathbb{K}$ is an injective (ring) homomorphism. Then there exists a (unique) injective homomorphism $h : F(A) \to \mathbb{K}$ such that*

$$j = h \circ \iota,$$

where $\iota : A \to F(A)$ is the natural inclusion defined above.

Proof. Note that if $b \neq 0$ in A, then $j(b) \neq 0$ in $\mathbb{K}$, because j is injective. But $\mathbb{K}$ is a field, so the element $j(b)$ has a multiplicative inverse, $[j(b)]^{-1}$. So, let us start by defining

$$h(\frac{a}{b}) \stackrel{\text{def}}{=} j(a) \cdot [j(b)]^{-1}.$$

We should check that h is well-defined. Suppose $\frac{a}{b} = \frac{c}{d}$, i.e. suppose $ad = bc$. Is it true that $j(a) \cdot [j(b)]^{-1} = j(c) \cdot [j(d)]^{-1}$? This is the same as asking whether $j(a)j(d) = j(b)j(c)$. Indeed, since j is a homomorphism,

$$j(a)j(d) = j(ad) = j(bc) = j(b)j(c).$$

So h is a well-defined function. The fact that $j = h \circ \iota$ is because

$$h \circ \iota(a) = h(\frac{a^2}{a}) = j(a^2)[j(a)]^{-1} = j(a).$$

So it only remains to show that h is a homomorphism, namely, that for any a, b, c, d with $bd \neq 0$,

$$h(\frac{a}{b}) \cdot h(\frac{c}{d}) = h(\frac{ac}{bd}) \text{ and} \tag{27}$$

$$h(\frac{a}{b}) + h(\frac{c}{d}) = h(\frac{ad+bc}{bd}). \tag{28}$$

To prove Equation 27, by definition of h, we have to prove that

$$j(a) \cdot [j(b)]^{-1} \cdot j(c) \cdot [j(d)]^{-1} = j(ac) \cdot [j(bd)]^{-1},$$

or equivalently, that

$$j(a) \cdot j(c) \cdot j(bd) = j(ac) \cdot j(b) \cdot j(d).$$

But this is true simply because j is a homomorphism, so both sides are equal to $j(abcd)$.

Finally, to prove Equation 28, we have to show that

$$j(a) \cdot [j(b)]^{-1} + j(c) \cdot [j(d)]^{-1} \quad = j(ad+bc) \cdot [j(bd)]^{-1}.$$

or equivalently

$$\left(j(a) \cdot [j(b)]^{-1} + j(c) \cdot [j(d)]^{-1}\right) \cdot j(bd) = j(ad+bc).$$

Now, since $j(bd) = j(b)j(d)$, it follows that $[j(b)]^{-1} \cdot j(bd) = j(d)$ and $[j(d)]^{-1} \cdot j(bd) = j(b)$. So the expression above simplifies to

$$j(a) \cdot j(d) + j(c) \cdot j(b) = j(ad+bc),$$

which is true because j is a homomorphism. □

4.8 Exercises

1. Prove that ${}^{\mathbb{Z}_6}/_{(2)}$ is a domain.
2. Prove that ${}^{\mathbb{Z}_{p^2}}/_{(p)}$ is a field, for any prime p.
3. Show that ${}^{\mathbb{Z}[x,y]}/_{(y)}$ is isomorphic to $\mathbb{Z}[x]$.
4. Let a be an element of a C-ring A. Let $B = {}^{A}/_{(a)}$. If you now consider the ring $A[x]$, this also has an ideal (a), formed by all polynomials that are multiple of a. Is it true that
$$ {}^{A[x]}/_{(a)} = B[x]? $$
5. Inside $\mathbb{Q}$, consider the subring
$$ \mathbb{Z}\left[{}^{1}/_{2}\right] \stackrel{\text{def}}{=} \left\{ \frac{a}{2^n} \text{ such that } a \in \mathbb{Z},\, p \in \mathbb{N} \right\}. $$
Prove that it is isomorphic to ${}^{\mathbb{Z}[x]}/_{(2x-1)}$.
6. In the quotient ring $A = {}^{\mathbb{Q}[x]}/_{(x^2-1)}$ consider the ideal I generated by the element $\overline{x^3 - 7x + 6}$. Is I proper? Is the quotient ${}^{A}/_{I}$ a field?
7. Consider the ideal $J = (x+1, x-1)$ in $\mathbb{Z}[x]$. Is it proper? Is it prime? Is it maximal?
8. If every proper ideal of A is prime, show that A is a field.
9. Prove that (5) is not prime in $\mathbb{Z}[i]$.
10. The *Jacobson radical* $J(A)$ of a C-ring A, is the intersection of all maximal ideals of A. Show that $J(\mathbb{Z}) = (0)$.
11. Let I, J be ideal of a C-ring A. Let P be a prime ideal of A. Assuming $I \cdot J \subseteq P$, prove that either $I \subseteq P$ or $J \subseteq P$.
12. Show that the quotient ${}^{\mathbb{Z}_5[x]}/_{(x^3+2x^2+3)}$ is a field.
13. Which elements are invertible in $\mathbb{Z}_4[x,y]$? Which elements are zero-divisors?
14. Is the subring of a Noetherian ring always Noetherian? (Hint: some domains are not Noetherian; fields are Noetherian...)
15. What is the field of fractions of $\mathbb{R}[x]$?

5 Irreducible elements and unique factorization domains

We want to study the elements that cannot be decomposed as products. Here the hardest task is coming up with a good definition: In fact, in a C-ring with 1, any element a can be written as $1 \cdot a$, and also as $(-1) \cdot a$. Somewhat in analogy with the usual definition of "prime number", which excludes 0 and ± 1, we would also like to avoid trivial decompositions like:

$$0 = 0 \cdot 0 = 1 \cdot 0 ... \text{ and } 1 = 1 \cdot 1 = (-1) \cdot (-1) ...$$

So long story short, here is the definition we want:

Definition 287. Let A be a C-ring with 1. An element $p \in A$ is called *irreducible* if it satisfies the following properties:
(IR0) $p \neq 0$;
(IR1) p is not invertible;
(IR2) if $p = a \cdot b$, then at least one of a, b is invertible.

Remark 288. In the definition above, we could equivalently replace the "at least one" by "exactly one". In fact, were a, b both invertible, then $p = a \cdot b$ would also be invertible — so p would not be irreducible because of (IR1).

Example 289. In $\mathbb{Z}$, the irreducible elements are precisely the prime numbers.

Example 290. Not all C-rings have irreducible elements. Consider for example the C-ring $A = \{f : \mathbb{R} \to \mathbb{R}\}$. For any function f, let us call the *zero set of* f

$$Z(f) \stackrel{\text{def}}{=} \{x \in \mathbb{R} \text{ such that } f(x) = 0.\}$$

Clearly, a function f is invertible if and only if f never vanishes; which happens if and only if $Z(f) = \emptyset$. Now, consider any non-invertible function f, different than the zero function. When is f irreducible? To answer this question, consider the function h that is identically 0 on $Z(f)$ and identically 1 elsewhere. Then

$$f = f \cdot h$$

and neither f nor h are invertible!, because they both vanish somewhere. So f is not irreducible. In conclusion, A is a C-ring that has *no* irreducible elements.

Once the definition is setup, let us let the theorems flow.

Proposition 291. *Let A be a domain with 1. Let p be any nonzero element of A. Then*

$$(p) \text{ is a prime ideal } \implies p \text{ is irreducible.}$$

Proof. Since (p) is prime, it is proper, so p is not invertible. We also know $p \neq 0$ by assumption. So all we need to show is (IR2). Suppose $p = a \cdot b$. Since $a \cdot b \in (p)$ and (p) is a prime ideal, one of a, b must belong to (p).

- If $a \in (p)$, then $a = p \cdot c$ for some $c \in A$. So we have $p = a \cdot b = p \cdot c \cdot b$. Since we are in a domain, we get $1 = c \cdot b$, so b is invertible.
- If $b \in (p)$, then $b = d \cdot p$ for some $d \in A$. So we have $p = a \cdot b = a \cdot d \cdot p$. Since we are in a domain, we get we get $1 = a \cdot d$, so a is invertible.

Hence, one of a, b must be invertible. □

Corollary 292. *If A is a domain, then all monic degree-one polynomials in $A[x]$ are irreducible.*

Proof. For any polynomial of the form $x - a$, with $a \in A$, we know by Proposition 243 that

$$A \cong {}^{A[x]}\!/_{(x-a)}.$$

Since A is a domain, by Theorem 256 the ideal $(x - a)$ is prime, so by Proposition 291 the polynomial $x - a$ is irreducible. □

Remark 293. In Corollary 292, all three assumptions are needed. For example, in the domain $A = \mathbb{Z}$, the degree-one (non-monic!) polynomial $2x$ and the monic polynomial x^2 (not of degree one) are both reducible. Also, if A is not a domain, not necessarily all monic polynomials of degree one in $A[x]$ are irreducible. For example, in $A = \mathbb{Z}_{12}$, we have

$$x = (4x + 3)(3x + 4),$$

and by Theorem 279 neither $4x+3$ nor $3x+4$ are invertible (because 3 and 4 are not nilpotent). So in $\mathbb{Z}_{12}[x]$, the polynomial x is reducible. In fact, one can prove the following result (thanks to Aldo Conca for clarifying the direction '(iii) ⇒ (ii)'):

Theorem 294. *Let $n \in \mathbb{N}$. Let $A = {}^{\mathbb{Z}}\!/_{(n)}$. The following are equivalent:*

(i) *x is irreducible in $A[x]$;*
(ii) *all monic degree-one polynomials are irreducible in $A[x]$;*
(iii) *either $n = 0$ or n is a prime power.*

Proof. "(ii) ⇒ (i)" is obvious.
"(iii) ⇒ (ii)". If $n = 0$, then $A = \mathbb{Z}$ is a domain, and by Corollary 292 we conclude. If $n = p^k$ for some prime p and some positive integer k, suppose by contradiction that in $A[x]$ we have

$$x + a = g \cdot h$$

for some a in A and some $g, h \in A[x]$ not invertible. Now, "pass mod p", i.e. apply to this equality the homomorphism from $\mathbb{Z}_n$ to $\mathbb{Z}_p$ that sends $\overline{z}$ to $\tilde{z}$ for any $z \in \mathbb{Z}$, cf. Example 187. We get that in $\mathbb{Z}_p[x]$

$$x + \tilde{a} = \tilde{g} \cdot \tilde{h}.$$

But $\mathbb{Z}_p[x]$ is a domain! Thus $1 = \deg \tilde{g} + \deg \tilde{h}$. Without loss, we can assume $\deg \tilde{g} = 0$ and $\deg \tilde{h} = 1$. So $g = c$ for some $c \in \mathbb{Z}_p$, $c \neq 0$. Going back to $A = \mathbb{Z}_n$, g is of the form

$$g = pg' + c,$$

for some $g' \in A[x]$. So any coefficient of g is a multiple of p, and thus nilpotent in $\mathbb{Z}_n$, except for the constant term of g, which is $pd + c$, where d is the constant term of g'. Since $(pd)^k = 0$ in $\mathbb{Z}_n$, pd is nilpotent, and since $1 \leq c \leq p - 1$, c is coprime with p and thus invertible in $\mathbb{Z}_n$; and as we saw in the proof of Theorem 279, the sum of an invertible and a nilpotent is always invertible. So in conclusion, by Theorem 279, g is invertible in $\mathbb{Z}_n$. A contradiction.
"(i) ⇒ (iii)". By contradiction, write $n = p^h q^k s$, where p, q are distinct primes, h, k are positive integers, and s is some integer that is not a multiple of p or q. Let $a = p^h$, and $b = q^k s$. By Bezout's theorem, there are integers c, d such that $ad + bc = 1$ in $\mathbb{Z}$. But then in $\mathbb{Z}_n$ we have $\overline{ab} = \overline{n} = 0$ and $\overline{ad + bc} = 1$, so

$$(\overline{a}x + \overline{c - acd})(\overline{b}x + \overline{d - bcd}) = \overline{ab}x^2 + (\overline{ad - abcd + bc - abcd})x + \overline{cd - bc^2d - acd^2 + abc^2d^2} = x.$$

It remains to prove that the two factors on the left are not invertible. By Theorem 279, it suffices to prove that $\overline{a}$ and $\overline{b}$ are not nilpotent. But this is obvious from the way we chose a and b: the prime q divides n but not a, so no power of a will be a multiple of n. Similarly, since p divides n but not b, no power of b will be a multiple of n. □

Here is a partial converse of of Proposition 291:

Proposition 295. *Let A be a PID. Let p be a nonzero element of A. Then*

$$p \text{ is irreducible } \iff (p) \text{ is a prime ideal } \iff (p) \text{ is a maximal ideal.}$$

Proof. We already established that the two implication towards the left are true for any domain with 1 (whether PID or not). So let us prove that in any PID ring, if p irreducible, then (p) is maximal. Say $(p) \subsetneq J$ for some ideal J. Since A is a PID, $J = (a)$ for some a. Since $(p) \subseteq (a)$, we have that

$$p = a \cdot b \text{ for some } a, b \in A.$$

By the definition of irreducible, one of a, b must be invertible. If b is invertible, then $p \cdot b^{-1} = a$, so $a \in (p)$ and $(p) = (a) = J$; a contradiction. Hence, a is invertible, which implies $J = A$. □

The actual converse of Proposition 291 is false. Too see this, we have to 'take a walk on the wild side' and study non-PID C-rings like the following subset of $\mathbb{C}$:

$$\mathbb{Z}[\sqrt{-5}] \stackrel{\text{def}}{=} \{a + bi\sqrt{5} \text{ such that } a, b \in \mathbb{Z}\}.$$

It is easy to see that $\mathbb{Z}[\sqrt{-5}]$ is a subring of $\mathbb{C}$. Since $\mathbb{C}$ is a domain, so is $\mathbb{Z}[\sqrt{-5}]$. Note also that $\mathbb{Z}$ is a subring of $\mathbb{Z}[\sqrt{-5}]$, because any integer z can be written as $z + 0i\sqrt{5}$. The C-ring $\mathbb{Z}[\sqrt{-5}]$ contains a few "surprises":

Proposition 296. *In $\mathbb{Z}[\sqrt{-5}]$ the ideals (2) and (3) are not prime.*

Proof. The key is that 6 can be factored in two different ways:

$$2 \cdot 3 = 6 = (1 + i\sqrt{5}) \cdot (1 - i\sqrt{5}).$$

So the elements $1 + i\sqrt{5}$ and $1 - i\sqrt{5}$ are not in (2), because 1 is not a multiple of 2. Yet their product is 6, which is in (2). So (2) is not prime. Same for (3). □

Lemma 297. *Let $p \neq 5$ be any prime number. Then*

$$p \text{ is reducible in } \mathbb{Z}[\sqrt{-5}] \iff p = A^2 + 5B^2 \text{ for some integers } A, B.$$

In particular, $p = 2$ and $p = 3$ are irreducible.

Proof. Let us start with the obvious remark that 2 and 3 cannot be written as $A^2 + 5B^2$ for any integers a, b, because if $|A| \geq 1$ and $|B| \geq 1$,

$$A^2 + 5B^2 \geq 1^2 + 5 \cdot 1^2 = 6. \tag{29}$$

"⇒" Let p be a prime number, $p \neq 5$, that *cannot* be written as $A^2 + 5B^2$ for any integers A, B. We want to show that p is irreducible. Write $p = (a + bi\sqrt{5}) \cdot (c + di\sqrt{5})$. Without loss of generality we can assume $|a| \leq |c|$: In fact, if not we can swap the two factors. Moreover, we can assume $G \stackrel{\text{def}}{=} \gcd(a, b) = 1$, because if $G > 1$ we can "collect G and move it into the second factor", as follows:

$$p = (\frac{a}{G} + \frac{b}{G} i\sqrt{5}) \cdot (Gc + Gdi\sqrt{5});$$

and now if we set $a' \stackrel{\text{def}}{=} \frac{a}{G}$, $b' = \frac{b}{G}$, and $c' = Gc$, indeed $\gcd(a', b') = 1$ and $|a'| \leq |c'|$.

Now we are ready to consider the equation

$$p = (a + bi\sqrt{5}) \cdot (c + di\sqrt{5}) = (ac - 5bd) + (ad + bc)i\sqrt{5}.$$

By equating the coefficients, we immediately obtain

$$\begin{cases} p = & ac - 5bd \\ 0 = & ad + bc. \end{cases} \tag{30}$$

We distinguish three cases, according to whether a and b are zero or not.

- If $a = 0$, then System 30 becomes

$$\begin{cases} p = & -5bd \\ 0 = & bc, \end{cases}$$

  which is impossible because p is prime (and not 5), so it cannot be a multiple of 5.
- If $a \neq 0$ and $b \neq 0$, since we are in a domain $ab \neq 0$. Moreover,

$$ad + bc = 0 \implies bc = -ad \implies \frac{bc}{ab} = -\frac{ad}{ab} \implies \frac{c}{a} = -\frac{d}{b}.$$

  Set $k \stackrel{\text{def}}{=} \frac{c}{a} = -\frac{d}{b}$. Then by definition

$$\begin{cases} c = & ak \\ d = & -bk. \end{cases} \tag{31}$$

  Note that a priori $k \in \mathbb{Q}$. But if $k = \frac{u}{v}$ with $\gcd(u, v) = 1$, then ak can be an integer only if a is a multiple of v; similarly $-bk$ can be an integer only if b is a multiple of v; but we assumed at the beginning that $G = \gcd(a, b) = 1$, so $v = 1$. So, in fact, k is an integer.
  Plugging the values of System 31 into the first equation of System 30, we get

$$p = a \cdot ak - 5b \cdot bk = k(a^2 + 5b^2). \tag{32}$$

  This tells us that $k > 0$. Since k is an integer, $k \geq 1$. On the other hand, Equation 32 tells us that k divides a prime p, so either $k = 1$ or $k = p$. Either way we get a contradiction: If $k = 1$ Equation 32 reads $p = a^2 + 5b^2$, which contradicts how p was chosen. If instead $k = p$, Equation 32 reads $1 = a^2 + 5b^2$, which is clearly impossible by Inequality 29.
- If $a \neq 0$ and $b = 0$, then System 30 becomes

$$\begin{cases} p = & ac \\ 0 = & ad. \end{cases}$$

  But then $ad = 0$ implies that $d = 0$, because we are in a domain. Moreover, p is prime: So from $p = ac$ and $|a| \leq |c|$ it follows that a is ± 1. So in this case the factor $(a + bi\sqrt{5})$ is invertible.

In conclusion, p is irreducible in $\mathbb{Z}[\sqrt{-5}]$.

"$\Leftarrow$" The idea is to write

$$p = (A + Bi\sqrt{5}) \cdot (A - Bi\sqrt{5}). \tag{33}$$

The only thing to verify is that neither of the factors above is invertible. We prove a more general fact. Suppose that for some nonzero integers a, b, c, d, we have

$$1 = (a + bi\sqrt{5}) \cdot (c + di\sqrt{5}).$$

We claim that

$$a = c = \pm 1 \text{ and } b = d = 0.$$

From the claim the conclusion follows immediately, because were for example $(A - Bi\sqrt{5})$ invertible, by setting $a \stackrel{\text{def}}{=} A$ and $b \stackrel{\text{def}}{=} -B$ we would obtain that $A = \pm 1$ and $B = 0$; which plugged back into Equation 33 would tell us $p = 1$, a contradiction.
So, let us prove the claim. By equating the coefficients in the identity above, we immediately obtain

$$\begin{cases} 1 = & ac - 5bd \\ 0 = & ad + bc. \end{cases} \tag{34}$$

Reasoning exactly as in the case "⇒" above, we can assume that $\gcd(a, b) = 1$; we distinguish the three cases "$a = 0$" (which leads to a contradiction), "$a \neq 0$ and $b \neq 0$" (which eventually leads to a contradiction with Inequality 29), and "$a \neq 0$ and $b = 0$". In this third case System 34 becomes

$$\begin{cases} 1 = & ac \\ 0 = & ad \end{cases}$$

which tells us that $d = 0$ and a is invertible in $\mathbb{Z}$. So $a = \pm 1$. □

Example 298. The primes $2, 3, 7, 11, 13, 17, 23$ are all irreducible in $\mathbb{Z}[\sqrt{-5}]$, because they are not 5 and they cannot be written as $a^2 + 5b^2$. (One can verify this either by hand, or by using Theorem 301 below.) Instead, 5 is reducible as

$$5 = (0 + i\sqrt{5})(0 - i\sqrt{5})$$

and $29 = 3^2 + 5 \cdot 2^2$ is reducible as

$$29 = (3 + 2i\sqrt{5}) \cdot (3 - 2i\sqrt{5}).$$

Deeper thoughts 299. One can find infinitely many other examples of irreducible elements p such that (p) is not prime, using the following two deep results in number theory:

Theorem 300 (Dirichlet)**.** *Let $m < n$ be positive integers. If $\gcd(m, n) = 1$, there are infinitely many primes congruent to m modulo n.*

Theorem 301 (Lagrange)**.** *Let p be a prime.*

- *p can be written as $p = a^2 + 5b^2$ for some positive integers a, b if and only if either $p \equiv 1 \mod 20$, or $p \equiv 9 \mod 20$.*
- *$2p$ can be written as $2p = a^2 + 5b^2$ for some positive integers a, b if and only if either $p \equiv 3 \mod 20$, or $p \equiv 7 \mod 20$.*

In particular, Theorem 301 paired with Lemma 297 tells us that for any prime $p \neq 5$

$$p \text{ is irreducible in } \mathbb{Z}[\sqrt{-5}] \iff p \text{ is congruent to one of } 3, 7, 11, 13, 17, 19 \mod 20.$$

Now, let p be a prime congruent to 3 or 7 mod 20. By Dirichlet's theorem, there are infinitely many such examples. By what we said above, p is irreducible in $Z[\sqrt{-5}]$. However, again by Lagrange, there are integers a, b such that $2p = a^2 + 5b^2$; hence,

$$2 \cdot p = (a + bi\sqrt{5}) \cdot (a - bi\sqrt{5}).$$

Now, a is not a multiple of p; otherwise from $5b^2 = 2p - a^2$ we would get that also b is a multiple of p; this would imply that $2p$ is a multiple of p^2; so $p = 2$, a contradiction with the fact that p is congruent to 3 or 7 mod 20. But since a is not a multiple of p, neither $a + bi\sqrt{5}$ nor $a - bi\sqrt{5}$ are in the ideal (p). Yet their product is. So the ideal (p) is not prime.

*More on Gaussian Integers

In view of Lemma 297, one may wonder which elements of $\mathbb{Z}[i]$ are irreducible. Here we show that the *primes* in $\mathbb{Z}$ that are irreducible in $\mathbb{Z}[i]$ are exactly the primes of the form $4k+1$. Note: It is not true that all irreducible elements of $\mathbb{Z}[i]$ are in $\mathbb{Z}$.

Example 302. The number $1+i$ is irreducible in $\mathbb{Z}[i]$, because it has norm 2. In fact, by Lemma 125, if $zw = 1+i$ then $\mathcal{N}(z)\cdot\mathcal{N}(w) = 2$, which implies that one of z, w must have norm one and thus be invertible.

Example 303. The prime number 2 is not irreducible in $\mathbb{Z}[i]$, because $2 = (1+i)(1-i)$, and none of these two factors is invertible by Lemma 125. Note that 2 has norm 4.

Let us warm up with an easy consequence of Theorem 126:

Theorem 304. *$\mathbb{Z}[i]$ is a PID.*

Proof. The proof is very similar to the direction (1) $\Rightarrow$ (2) in the proof of Theorem 217. Let I be an ideal of $\mathbb{Z}[i]$. If $I = (0)$, then I is principal and we are done. If I contains nonzero elements, let w be a **smallest-norm** nonzero element of I. We claim that

$$I = (w).$$

The containment $\supset$ is obvious, because $z \in I$. To prove

$$I \subseteq (g),$$

pick any Gaussian integer z in I. Let us divide it by w, according to Theorem 126. Since

$$z = q\cdot w + r \qquad \text{with } \mathcal{N}(r) < \mathcal{N}(w),$$

we have that $r = z - qw$ is in I, but has smaller norm than w: A contradiction. □

We can now prove an analog of Lemma 297.

Lemma 305. *Let $p \neq 2$ be any prime number. Then*

$$p \text{ is reducible in } \mathbb{Z}[i] \iff p = a^2 + b^2 \text{ for some positive integers } a, b.$$

In particular, $p = 3$ and $p = 7$ are irreducible.

Proof. Left as exercise. (*Hint*: Try to mimic the proof of Lemma 297; pay particular attention to the case $a = 0$, which is slightly different; the conclusion in this case is that one of the two factors is invertible.) □

The next theorem was first announced by Fermat in 1654; the first complete proof was given by Euler 100 years later. We reproduce Zagier's amazing proof from 1990[12].

Definition 306. Let S be any finite set. An *involution* $i : S \to S$ is a function such that $i \circ i$ is the identity. In other words, $i(i(s)) = s$ for all s. A *fixed point* for i is an element $s \in S$ such that $i(s) = s$.

Lemma 307. *The number of fixed points of any involution $i : S \to S$ has the same parity of the cardinality of S (that is, either they are both even, or they are both odd).*

[12] D. Zagier, *A one-sentence proof that every prime $p \equiv 1 \mod 4$ is a sum of two squares*, Amer. Math. Monthly 97 (1990), no. 2, 144.

Proof. Any element x that is not a fixed point can be paired with its image $i(x)$. □

Theorem 308 (Fermat 1654; Euler, 1755; Zagier, 1990)**.** *Let $p \neq 2$ be a prime number. Then*

$$p \equiv 1 \mod 4 \iff p = a^2 + b^2 \text{ for some positive integers } a, b.$$

Proof.

"⇐" This is easy: Suppose $p = a^2 + b^2$ for some integers a, b. Since p is odd, one of a, b must be even and the other odd. Without loss of generality write $a = 2k$ and $b = 2\ell + 1$; then

$$a^2 + b^2 = (2k)^2 + (2\ell + 1)^2 = 4(k^2 + \ell^2 + \ell) + 1.$$

"⇒" If $p = 4k + 1$ is prime, consider the set

$$S = \{(x, y, z) \text{ positive integers such that } x^2 + 4yz = p\}.$$

This set is finite and admits two different involutions:

- the 'easy' involution $(x, y, z) \longmapsto (x, z, y)$, whose set of fixed points $F \subseteq S$ is

  $$F = \{(x, y, y) \text{ such that } x^2 + (2y)^2 = p\}.$$

  This F corresponds to all possible representations of p as sums of a (necessarily odd) square and an even square. As we saw in the direction "⇐", this simply describes all possible representations of p as sums of two squares.
- A more complicated involution, defined by

  $$(x, y, z) \longmapsto \begin{cases} (x + 2z, z, y - x - z) & \text{if } x < y - z, \\ (2y - x, y, x - y + z) & \text{if } y - z < x < 2y, \\ (x - 2y, x - y + z, y) & \text{if } x > 2y. \end{cases}$$

  This function is well-defined because if $x^2 + 4yz = p = 4k + 1$, then x is odd, so the case $x = 2y$ is void; and similarly $x = y - z$ is impossible because we would have $p = (y - z)^2 + 4yz = (y + z)^2$. We leave it as exercise to verify (by cases) that it is indeed an involution.

We claim that the second involution has exactly one fixed point. In fact, if $x < y - z$ then the first component increases. If $x > 2y$, then the first component decreases. So the only fixed points can be in the region $y - z < x < 2y$. Setting $(x, y, z) = (2y - x, y, x - y + z)$ yields $x = y$ (and no conditions on z). But then the equation $x^2 + 4xz = p$ tells us that x divides the prime p, so either $x = p$ (which is impossible, $p^2 + 4pz$ is larger than p) or $x = 1$. In the latter case $y = x = 1$ and $4k + 1 = p = 1 + 4z$, so $z = k$. Hence, the only fixed point is $(1, 1, k)$.

Conclusion: from the second involution and from Lemma 307, we know that the cardinality of S is odd. But then from the first involution, we conclude that the number of representations of p as a sum of two squares is odd. And what is odd, cannot be zero. □

Corollary 309. *Let $p \neq 2$ be any prime number. Then*

$$p \text{ is irreducible in } \mathbb{Z}[i] \iff p \equiv 3 \mod 4.$$

Proof. Odd primes modulo 4 are either congruent to 1, or congruent to 3. The claim follows immediately by putting together Lemma 305 and Theorem 308. □

Deeper thoughts 310. In a letter to Blaise Pascal dated September 25, 1654, Fermat announced also the following statements, which were later proven by Euler (and can be viewed as analogs of Lagrange's theorem 301): For any prime p different than 2,

- p can be written as $p = a^2 + 2b^2$ for some positive integers a, b if and only if either $p \equiv 1$ or $p \equiv 3 \mod 8$.
- p can be written as $p = a^2 + 3b^2$ for some positive integers a, b if and only if $p \equiv 1 \mod 3$.

One may wonder if there are "Zagier-style" proof, using a clever involution, for these statements as well. There is active research on this! In 2010 Christian Elsholtz has found a combinatorial proof for the case $p = a^2+2b^2$. Check out his webpage at TU Graz for updates! and in particular www.math.tugraz.at/~elsholtz/WWW/papers/zagierenglish9thjuly2002.ps.

5.1 UFDs

In this section we study domains where factorization into irreducible elements exists and is unique. The main examples we have in mind are $\mathbb{Z}$ (which is a PID) and polynomial rings in many variables, like $\mathbb{R}[x, y]$ (which are not PIDs). We call these rings UFDs, which is short for Unique Factorization Domains. The main result is that in these rings, irreducible elements generate prime ideals, although not necessarily maximal.

We also show that every PID is a UFDs. (If your UFD is a PID, then any irreducible element does generate a maximal ideal, cf. Proposition 291).

Definition 311. A C-ring A with 1 is called a *UFD* if it satisfies the following conditions:

(D) A is a domain.

(F) If $a \in A$ is not invertible and not zero, then $a = p_1 \cdot p_2 \cdot \ldots \cdot p_r$ with p_i irreducible.

(U) If $a = p_1 \cdot p_2 \cdot \ldots \cdot p_r = q_1 \cdot q_2 \cdot \ldots \cdot q_s$ with all p_i, q_j irreducible, then $r = s$ and, up to reordering, $(p_i) = (q_i)$ for all i.

Example 312. $\mathbb{Z}$ is a UFD.

Non-Example 313. $\mathbb{Z}[\sqrt{-5}]$ is **not** a UFD. We saw in Proposition 296 that elements like 6 have two different factorizations.

Lemma 314. *Let A be a PID. If $a \neq 0$, and a is not invertible, then some irreducible element p divides a.*

Proof. If a is irreducible, then choosing $p \stackrel{\text{def}}{=} a$ we are done. Otherwise, we have

$$a = a_1 \cdot b_1, \text{ with } a_1, b_1 \text{ not invertible.}$$

In particular, $(a) \subsetneq (a_1)$. Of course, none of a_1, b_1 is zero, since their product a is not zero. If one of a_1, b_1 is irreducible, then a has an irreducible factor and we are done. If not, then in particular a_1 is reducible, that is,

$$a_1 = a_2 \cdot b_2, \text{ with } a_2, b_2 \text{ not invertible.}$$

In particular, $(a) \subsetneq (a_1) \subsetneq (a_2)$.

By repeating the reasoning above, either we find an irreducible factor at some point, or we create an infinite ascending chain of ideals

$$(a) \subsetneq (a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \ldots \subsetneq (a_n) \subsetneq (a_{n+1}) \subsetneq \ldots$$

that never stabilizes, which would contradict Corollary 226 (every PID is Noetherian!). □

Theorem 315. *A PID $\implies$ A UFD.*

Proof. Let us check the properties of UFDs one by one.

(D) Clearly A is a domain.

(F) By contradiction, let $a \neq 0$ be a non-invertible element that *cannot* be factored into a product of irreducibles. By Lemma 314, a has an irreducible factor p_1. In other words

$$a = p_1 a_1 \text{ for some } a_1 \in A.$$

Clearly a_1 is nonzero (because a is nonzero) and not invertible (otherwise we would have $(a) = (p_1)$, so a would be irreducible and $a = a$ would be a decomposition of a into irreducibles; a contradiction.) So by Lemma 314, a_1 has an irreducible factor p_2, and we can write

$$a = p_1(p_2 a_2) \text{ for some } a_2 \in A.$$

Clearly a_2 is not zero and not invertible (otherwise $a = p_1 p_2$ would decompose a into irreducibles). So reapplying Lemma 314, a_2 has an irreducible factor p_3, and

$$a = p_1 p_2 p_3 a_3 \text{ for some } a_3 \in A.$$

And so on. Since irreducible elements are not invertible, from $a_i = p_{i+1} a_{i+1}$ it follows that $(a_i) \subsetneq (a_{i+1})$ for each i. But then we have obtained an infinite ascending chain

$$(a) \subsetneq (a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \ldots \subsetneq (a_n) \subsetneq (a_{n+1}) \subsetneq \ldots$$

A contradiction with Corollary 226.

(U) Now suppose that $a = p_1 \cdot p_2 \cdot \ldots \cdot p_r = q_1 \cdot q_2 \cdot \ldots \cdot q_s$ with all p_i, q_j irreducible. Since q_1 is irreducible, (q_1) is maximal and in particular prime by Proposition 295. Since $p_1 \cdot p_2 \cdot \ldots \cdot p_r \in (q_1)$, then one of the p_i must belong to (q_1). Up to relabeling, assume it is p_1. Now we know that $p_1 \in (q_1)$, so $p_1 = q_1 \cdot \alpha$; but p_1 is irreducible, so α must be invertible. Hence, $(p_1) = (q_1)$. Since we are in a domain, we can cancel out q_1 from $p_1 \cdot p_2 \cdot \ldots \cdot p_r = q_1 \cdot q_2 \cdot \ldots \cdot q_s$; we obtain

$$p_2 \cdot \ldots \cdot p_r = q_2' \cdot \ldots \cdot q_s,$$

where we replaced q_2 by $q_2' \stackrel{\text{def}}{=} \alpha q_2$ (a replacement that does not affect the ideal generated.) Now we repeat the previous reasoning. Since (q_2) is prime, and $p_2 \cdot \ldots \cdot p_r \in (q_2)$, one of the p_i's (say, p_2) belongs to (q_2). Then $p_2 = \beta q_2$, with β necessarily invertible because p_2 is irreducible. So $(p_2) = (q_2)$, and so on. □

Corollary 316. *The polynomial rings in one variable over a field, like $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$ and $\mathbb{Z}_p[x]$ with p prime, are all UFDs.*

What about $\mathbb{Z}[x]$ or $\mathbb{R}[x, y]$? The answer will come in the next Section. For the moment, let us prove a fact relating irreducible elements and prime/maximal ideals.

Proposition 317. *Let A be a UFD. Let p be a (nonzero) element of A. Then*

$$p \text{ is irreducible} \implies (p) \text{ is a prime ideal.}$$

Proof. Say $a \cdot b \in (p)$. Then

$$a \cdot b = p \cdot c \quad \text{for some } c \in A.$$

Whether c is invertible or not, note that p is a factor of $a \cdot b$. Now, another possible factorization of $a \cdot b$ is obtained by juxtaposing the factors of a and the factors of b. Yet in UFDs, factorization is unique. It follows that p must appear either among the factors of a, or among the factors of b. (Or both.) So either $a \in (p)$, or $b \in (p)$. □

Can we strengthen the conclusion to "(p) is maximal"? It turns out that the answer is no; compare Remark 340.

Deeper thoughts 318. Every PID is a UFD; also, every PID is a Noetherian domain. One might wonder if there is any implication between the UFD property and the property of being a Noetherian domain. With some effort one can see that the answer is negative:

- the C-ring $\mathbb{Z}[\sqrt{-5}]$ of Example 313 is a Noetherian domain, but not a UFD;
- the "polynomial ring with infinitely many variables" $\mathbb{R}[x_1, x_2, \dots, x_n, x_{n+1}, \dots]$ is an example of C-ring that is not Noetherian. An infinite ascending chain of ideals that never stabilizes is given by the ideals $I_n \stackrel{\text{def}}{=} (x_1, x_2, \dots x_n)$. However, one can prove that this C-ring is a UFD.
- the C-ring $A = \{f : \mathbb{R} \to \mathbb{R}\}$ is neither Noetherian, nor UFD (because we saw in Example 290 that in A there are no irreducible elements).

Greatest Common Divisor in UFDs

Definition 319. We call two irreducible elements p, q *associate*, if they are the same up to multiplication by an invertible element; or in other words, if $(p) = (q)$. If $a \in A$, we say that an irreducible element p of A is a *factor* of a if an element associate to p appears in the factorization of a. We say that an irreducible element p of A has *multiplicity* k in a, if k is the largest number of elements associate to p that appear in the factorization of a.

Example 320. If $a = p^3q^4$ and $b = p'qq'r$, with p, p', q, r irreducible, and with $(p) = (p')$, $(q) = (q')$. Then

- p has multiplicity 3 in a and 1 in b,
- q has multiplicity 4 in a and 2 in b,
- r is not a factor of a, but it has multiplicity 1 in b.

Lemma 321. *In any UFD ring, for any two nonzero elements a, b it makes sense to define (up to invertible factors) their greatest common divisor* $\gcd(a, b)$. *One has*

$$\gcd\left(\frac{a}{\gcd(a,b)}, \frac{b}{\gcd(a,b)}\right) = 1.$$

Proof. If we look at the factorizations of a and of b, we can now collect all irreducible elements that are factors of both a and b with the minimum of the two multiplicities, mimicking Corollary 16. By construction, $\frac{a}{\gcd(a,b)}$ is the product of the factors of a that are not factors of b; symmetrically, $\frac{b}{\gcd(a,b)}$ is the product of the factors of b that are not factors of a. Hence $\frac{a}{\gcd(a,b)}$ and $\frac{b}{\gcd(a,b)}$ have no common factor. □

Example 322. In $\mathbb{Z}[x, y]$ consider $a = 24y^2 - 12x^2y^2$ and $b = 8x^2 - 16$. The respective decompositions into irreducibles are

$$a = 2^2 \cdot 3 \cdot y^2 \cdot (2 - x^2) \quad \text{and} \quad b = 2^3(x^2 - 2).$$

Obviously $x^2 - 2$ and $2 - x^2$ are associate. Hence, $\gcd(a, b) = 2^2(X^2 - 2) = 4x^2 - 8$. Or, if you prefer, $\gcd(a, b) = 8 - 4x^2$ is also correct.

Remark 323. In a PID, the ideal (a, b) is always generated by $\gcd(a, b)$. (Proof: exercise!) But in an arbitrary UFD, this might not be true: Sometimes the ideal (a, b) could be not principal! In general, one can only say

$$(a, b) \subseteq (\gcd(a, b)),$$

but unless we are in a PID, there is no guarantee that $\gcd(a,b)$ belongs to the ideal (a,b). For example, we will see in the next Section that $\mathbb{Z}[x]$ and $\mathbb{Z}[x,y]$ are both UFD; in $\mathbb{Z}[x,y]$ one has $\gcd(2x,2y)=2$, but $2 \notin (2x,2y)$ for degree reasons. Or if you look back at Proposition 216, in $\mathbb{Z}[x]$ one has $\gcd(X,2)=1$ but $1 \notin (X,2)$.

5.2 *Gauss' theorem

This section is dedicated to the proof of the following theorem:

Theorem 324 (Gauss)**.** *Let A be a C-ring with* 1*.*

$$A \text{ is a UFD} \iff A[x] \text{ is a UFD.}$$

Prerequisite for the proof is the section on Fields of Fractions. Make sure you remember it before going on! While you review it, we can walk through one direction of the theorem, which is actually much easier than the other one.

Lemma 325. *Let A be a UFD ring. Let f be a degree-zero polynomial of $A[x]$.*

$$f \text{ is irreducible in } A[x] \iff f \text{ is irreducible in } A.$$

Proof. Obviously f is nonzero in $A[x]$ if and only if it is nonzero in A. The rest follows from the definition of 'irreducible' and the fact that the invertible elements of $A[x]$ are precisely the invertible elements of A (by Theorem 139.) □

Proposition 326. *If $A[x]$ is a UFD, then A is a UFD.*

Proof. Let us check the three axioms of UFDs.

(D) A is a domain by Theorem 139.

(F) Let $a \neq 0$ be a non-invertible element of A. Let us view a as a degree-zero polynomial in $A[x]$; being $A[x]$ a UFD, we know that a will factor as

$$a = p_1 \cdot \ldots \cdot p_s$$

with p_i irreducible *in* $A[x]$. But since $A[x]$ is a domain, necessarily $\deg p_i = 0$ for all i; so these p_i are all in A. By Lemma 325, all p_i are irreducible in A.

(U) Suppose

$$a = p_1 \cdot \ldots \cdot p_s = q_1 \cdot \ldots \cdot q_r$$

are two different factorizations of a into irreducible elements in A. By Lemma 325, we can view this as two different factorizations of a into irreducible elements in $A[x]$. Being $A[x]$ UFD, we conclude that $r = s$ and, up to reordering, $(p_i) = (q_i)$ as ideals of $A[x]$. This means that for each i there is an invertible polynomial $u_i \in A[x]$ such that $p_i = u_i \cdot q_i$. But the invertible elements of $A[x]$ are just the invertible elements of A, by Theorem 139. So $u_i \in A$ and $(p_i) = (q_i)$ also as ideals of A. □

So the hard part of Gauss' theorem is to show that if A is UFD, so is $A[x]$. We need a few Lemmas first.

Definition 327. Let A be a UFD. Let f be a nonzero polynomial in $A[x]$. The *content* of f is the gcd of all nonzero coefficients of f. We call *primitive* the polynomials whose content is invertible in A. Recall that *monic* polynomials are those with leading coefficient 1.

Obviously, monic implies primitive, because if one of the coefficients is 1, their greatest common divisor must be 1.

Example 328. Let $A = \mathbb{Z}$. Let $f = x^2 + 6x + 9$. Then f is monic and primitive, although not irreducible (it is the square of $x + 3$.)

Example 329. Let $A = \mathbb{Z}$. Let $f = 3x^2 + 6x - 9$. Then $\text{content}(f) = 3$, which is not invertible in $\mathbb{Z}$. So f is not primitive. In contrast, $f + 1$ is primitive, but not monic.

Example 330. Let $A = \mathbb{R}[x]$. Let $g = 3x^2y - 9x - 3y$ in $\mathbb{R}[x, y] = A[y]$. Then $\text{content}(g) = 3$, which is invertible in A; so g is primitive. In contrast, $g + 3y = 3x^2y - 9x$ is not primitive, because its content is $3x$, which is not invertible in A.

Being a gcd, the content is defined up to an invertible coefficient. So we might as well say, primitive polynomials are those whose content is (equivalent to) 1. The next lemma shows that "primitive" is in some sense a weaker notion than "irreducible", at least for polynomials of degree 1 or higher.

Lemma 331. *If A is a UFD, any irreducible polynomial of $A[x]$ of positive degree is primitive.*

Proof. By contradiction, choose f irreducible polynomial with $\deg f \geq 1$ and not primitive. Write

$$f = \text{content}(f) \cdot f', \text{ with } f' \text{ primitive.}$$

Since f is irreducible and $\text{content}(f)$ is not invertible, the other factor f' must be invertible. Yet $\deg f' = \deg f \geq 1$. But $A[x]$ is a domain by Theorem 139, so for any nonzero polynomial g one has $\deg(f' \cdot g) = \deg f' + \deg g \geq 1$. So f' cannot be invertible: A contradiction. □

The converse is "very false": It is easy to construct primitive polynomials that are not irreducible, as the next Lemma shows.

Lemma 332 (Gauss)**.** *Let A be a UFD. Let $f, g \in A[x]$. If f and g are primitive, so is $f \cdot g$.*

Proof. Let $f = a_0 + a_1x + \ldots + a_mx^m$ and let $g = b_0 + b_1x + \ldots + b_nx^n$. We proceed by contradiction: Suppose that some irreducible element p divides all coefficients of $f \cdot g$. Since f is primitive, p does not divide all coefficients of f. Let a_h be the coefficient with smallest index h such that p does *not* divide a_h. Similarly, since g is primitive, let b_j be the coefficient with smallest index j such that p does *not* divide b_j. Since we are in a UFD and p is irreducible, p does not divide $a_h \cdot b_j$ (otherwise it would be a factor of one of them).

Now consider the coefficient in $f \cdot g$ of x^{h+j}. By the way we defined products of polynomials, this coefficient is

$$c_{h+j} \stackrel{\text{def}}{=} \sum_{k=0}^{h+j} a_k \cdot b_{h+j-k}.$$

But in the sum above, all summands except $a_h \cdot b_j$ are themselves multiples of p. In fact, if k is smaller than h, then by definition of h, a_k is a multiple of p. If instead k is larger than h, then $h + j - k < j$; so by definition of j, the coefficient b_{h+j-k} is a multiple of p.

But then c_{h+j} is the sum of a multiple of p and a non-multiple of p; so it cannot be a multiple of p. This contradicts the assumption that every coefficient of $f \cdot g$ is a multiple of p. □

Lemma 333 (Gauss)**.** *Let A be a UFD. Let $f, g \in A[x]$ be nonzero polynomials. Then*

$$\text{content}(f \cdot g) = \text{content}(f) \cdot \text{content}(g).$$

Proof. Let $c_f \stackrel{\text{def}}{=} \text{content}(f)$ and $c_g \stackrel{\text{def}}{=} \text{content}(g)$. Then we can write $f = c_f \cdot f'$ with f' primitive, and $g = c_g \cdot g'$ with g' primitive. Obviously

$$f \cdot g = (c_f \cdot f') \cdot (c_g \cdot g') = c_f c_g \cdot (f' \cdot g'),$$

and by Lemma 332, no irreducible elements divides all coefficients of $f' \cdot g'$.

So if p is any irreducible factor of all coefficients of $f \cdot g$, necessarily it divides $c_f c_g$. Conversely, if p divides one of c_f or c_g, it divides all coefficients of $f \cdot g$, so it must divide $\text{content}(f \cdot g)$. □

For the next Lemmas, all due to Gauss, it helps to think of the inspiring case $A = \mathbb{Z}$, $F(A) = \mathbb{Q}$.

Lemma 334. *Let A be a UFD. Let f be a primitive polynomial in $A[x]$ of positive degree. Let $\mathbb{K} = F(A)$ be the field of fractions of A. Assume f is reducible over $\mathbb{K}[x]$ as*

$$f = G \cdot H \quad \textit{for some } G, H \textit{ in } \mathbb{K}[x] \textit{ of positive degree.}$$

Then f is also reducible over $A[x]$ as

$$f = g \cdot h \quad \textit{for some } g, h \textit{ in } A[x] \textit{ of positive degree and primitive.}$$

Moreover,
(1) *if f is monic, g and h can be chosen to be monic;*
(2) *if G and H are irreducible, so are g and h.*

Proof. Let us write down G as

$$G = \frac{b_0}{a_0} + \frac{b_1}{a_1}x + \ldots + \frac{b_m}{a_m}x^m, \quad \text{with } \gcd(a_i, b_i) = 1.$$

Since we are in a UFD, it makes sense to take the least common multiple of the a_i's. Let us call it a. Now we can clear denominators:

$$\tilde{g} \stackrel{\text{def}}{=} a \cdot G = b_0 \cdot \frac{a}{a_0} + b_1 \frac{a}{a_1}x + \ldots + b_m \frac{a}{a_m}x^m \in A[x],$$

because a is a multiple of each a_i. Now that we know it is in $A[X]$, we can collect out the content of $\tilde{g}$, and write

$$\tilde{g} = \text{content}(\tilde{g}) \cdot g', \text{ with } g' \text{ primitive.}$$

Similarly, we can clear denominators for H: if we call c the least common multiple of all denominators of H, then

$$\tilde{h} \stackrel{\text{def}}{=} c \cdot H \in A[x],$$

and we can factor out the content of this new polynomial $\tilde{h}$ as $\tilde{h} = \text{content}(\tilde{h}) \cdot h'$. Now, consider the polynomial $ac \cdot f$. Putting together what we said, we have

$$ac \cdot f = ac \cdot GH = (aG) \cdot (cH) = \tilde{g} \cdot \tilde{h} = \text{content}(\tilde{g}) \cdot g' \cdot \text{content}(\tilde{h}) \cdot h' = \text{content}(\tilde{g} \cdot \tilde{h}) \cdot g' \cdot h',$$

where in the last step we applied Lemma 333. But in the equation above, f is primitive by assumption, and $g' \cdot h'$ is primitive by Lemma 333. It follows that $ac = \text{content}(\tilde{g} \cdot \tilde{h}) \cdot u$, for some invertible element $u \in A$. But then we can cancel out this common factor: we obtain

$$f = ug' \cdot h'$$

where both ug' and h' are primitive polynomials in $A[x]$. Note that $\deg(ug') = \deg G = m$ and $\deg h' = \deg H \stackrel{\text{def}}{=} n$. This suffices to prove the initial part of the claim, by setting $g \stackrel{\text{def}}{=} ug'$ and $h \stackrel{\text{def}}{=} h'$.

For item (1), we make one more modification, by multiplying each polynomial by the leading coefficient of the other. In fact, if v is the leading coefficient of g and w is leading coefficient of h, then the leading coefficient of f is obviously $v \cdot w$. So if $v \cdot w = 1$, we can set

$$g_0 \stackrel{\text{def}}{=} w \cdot g \qquad \text{and } h_0 \stackrel{\text{def}}{=} v \cdot h.$$

The leading term of g_0 is now $w \cdot v = 1$, and the leading term of h_0 is $v \cdot w = 1$. This shows (1).

To show (2), observe that by construction $g = \alpha G$, for some α in $\mathbb{K}$ that is nonzero. So certainly α is invertible in $\mathbb{K}$, because $\mathbb{K}$ is a field. Were g reducible in $A[x]$ as $g = g_1 \cdot g_2$, then G would be reducible too in $\mathbb{K}[x]$ as $G = (\alpha^{-1} g_1) \cdot g_2$. Similarly, h and H are connected by a nonzero element of A, which is invertible in $\mathbb{K}$: Were h reducible, so would be H. □

Lemma 335. *Let A be a UFD. Let f be a primitive polynomial in $A[x]$ of positive degree. Then f decomposes in a unique way (up to multiplication with invertible elements) as a product of irreducible, primitive polynomials in $A[x]$ of positive degree.*

Proof. Let us pass to $\mathbb{K} = F(A)$, the field of fractions of A. Using the natural injection of A into its field of fractions ($a \mapsto \frac{a}{1}$), we can view f as a polynomial in $\mathbb{K}[x]$. Yet $\mathbb{K}[X]$ is a PID by Theorem 217, and so in particular a UFD by Theorem 315. So f splits as

$$f = P_1 \cdot P_2 \cdot \ldots \cdot P_s, \qquad \text{for some } P_i \text{ irreducible in } \mathbb{K}[x].$$

Note that in $\mathbb{K}[x]$ every nonzero polynomial of degree zero is invertible, because $\mathbb{K}$ is a field; so all irreducible polynomials of $\mathbb{K}[x]$ have positive degree. So we can apply Lemma 334, obtaining

$$f = p_1 \cdot p_2 \cdot \ldots \cdot p_s, \qquad \text{for some } p_i \text{ primitive and irreducible in } A[x].$$

This shows the existence of the decomposition we wanted. It remains to show uniqueness. For this assume that

$$f = q_1 \cdot q_2 \cdot \ldots \cdot q_r, \qquad \text{for some } q_i \text{ primitive and irreducible in } A[x].$$

By Lemma 334, the q_i's are irreducible in $\mathbb{K}[x]$, which is a UFD; so we know that $r = s$ and up to reshuffling them, $(p_i) = (q_i)$ as ideals of $\mathbb{K}[x]$. This means that for each i in $\{1, 2, \ldots, s\}$, we can find an invertible element c_i in $\mathbb{K}$ such that

$$p_i = c_i \cdot q_i \text{ for all } i.$$

Note that "invertible in $\mathbb{K}$" does not imply "invertible in A"!, so we're not done yet. But since $\mathbb{K}$ is the field of fractions of A, any invertible element of $\mathbb{K}$ is of the form $c_i = \frac{b_i}{a_i}$, with $a_i \neq 0$ and $b_i \neq 0$. So for each i in $\{1, 2, \ldots, s\}$, we can find two nonzero elements a_i and b_i such that

$$a_i \cdot p_i = b_i \cdot q_i.$$

So far we have not used the assumption that the q_i's are primitive. We do so now. From the equation above, passing to the content and recalling that p_i and q_i are primitive, we get

$$a_i = b_i \cdot u_i \text{ for some } u_i \text{ invertible in } A.$$

Hence $b_i \cdot q_i = a \cdot p_i = b_i \cdot u_i \cdot p_i$. Canceling b_i (we are in a domain), we get that $q_i = u_i \cdot p_i$, with u_i invertible in A. In other words, $(p_i) = (q_i)$ also as ideals of $A[x]$. □

Theorem 336 (Gauss)**.** *If A is a UFD, so is $A[x]$.*

Proof. Let us check the three axioms of UFDs.

(D) $A[x]$ is a domain by Theorem 139.

(F) Let $f \in A[x]$ be a nonzero polynomial that is not invertible. If $\deg f = 0$, then f is in A, which is a UFD, so we can decompose it into irreducible elements. If $\deg f \geq 1$, we write it as

$$f = \text{content}(f) \cdot f', \qquad \text{with } f' \text{primitive.}$$

Then by Lemma 335 we know how to decompose f' into irreducible primitive polynomials. Also, we know how to decompose content(f) into irreducible degree-zero polynomials, because content$(f) \in A$. Putting the two decompositions together, we obtain one for f.

(U) Suppose a can be written in two ways as products of irreducible elements. Let us separate the irreducible factors of degree zero from the irreducible factors that have positive degree. We can write

$$P \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_r \quad = a = \quad Q \cdot q_1 \cdot q_2 \cdot \ldots \cdot q_s,$$

where P, Q are in A and p_i, q_j are irreducible polynomials in $A[x]$ of positive degree. (Note: P, Q are not necessarily irreducible: they are the product of the irreducible factors of degree zero.) By Lemma 331, all the p_i's and q_j's are primitive. So both P and Q are the content of a. Remember that the content is defined up to an invertible element, so we cannot claim that $Q = P$, but we can say that there is an invertible element u such that

$$Q = P \cdot u.$$

Now, we already know that A is a UFD, so the factorizations of P and of Q into degree-zero irreducible polynomials are going to be the same (up to reordering the factors, and up to multiplication by invertible elements).
So all we need to prove is that the factors p_i and q_j of positive degree are the same. From $Q = P \cdot u$ we get

$$P \cdot p_1 \cdot p_2 \cdot \ldots \cdot p_r = a = Q \cdot q_1 \cdot q_2 \cdot \ldots \cdot q_s = P \cdot u \cdot q_1 \cdot q_2 \cdot \ldots \cdot q_s.$$

Since we are in a domain, we can cancel P:

$$p_1 \cdot p_2 \cdot \ldots \cdot p_r = u \cdot q_1 \cdot q_2 \cdot \ldots \cdot q_s.$$

But now the polynomial on the left is primitive: So we are looking at two different factorizations of a primitive polynomial into primitive irreducible polynomials. Applying Lemma 335, we conclude that $r = s$ and up to reordering them, $(p_i) = (q_i)$ for all i. In other words, also the irreducible factors of positive degree are the same. □

Corollary 337. *$\mathbb{Z}[x]$ is a UFD.*

Corollary 338. *$\mathbb{K}[x, y]$ is a UFD for any field $\mathbb{K}$.*

Proof. Let $A = \mathbb{K}[x]$. Then $\mathbb{K}[x, y] = A[y]$. We already know that A is a PID by Theorem 217; hence A is a UFD by Theorem 315. So $A[y]$ is UFD by Theorem 336. □

Corollary 339. *For any positive integer n, for any UFD A, the C-ring $A[x_1, \ldots, x_n]$ is a UFD.*

Proof. By induction on n. The case $n = 1$ is Theorem 336. Set $A_n \stackrel{\text{def}}{=} A[x_1, \ldots, x_n]$. Since

$$A_{n+1} \stackrel{\text{def}}{=} A[x_1, \ldots, x_n, x_{n+1}] = A_n[x_{n+1}],$$

and by inductive assumption A_n is UFD, via Theorem 336 we conclude that A_{n+1} is UFD. □

Remark 340. $\mathbb{Z}[x]$, $\mathbb{K}[x,y]$ and $A[x_1,\dots,x_n]$ are examples of C-rings that are UFD but not PID. We saw that:

- in PID C-rings, if p is irreducible then the ideal (p) is prime and maximal (Prop. 295).
- in UFD C-rings, if p is irreducible then the ideal (p) is prime (Proposition 317).
- in domains, if p is irreducible we cannot conclude anything on (p) (cf. Proposition 296.)

This triggers a natural question: If we are in a C-ring that is a UFD but not a PID, is it still true that prime ideals of the type (p) are necessarily maximal? The answer is negative. In fact, in $\mathbb{Z}[x]$ we already know that (2) is not maximal, because $(2) \subsetneq (x,2)$. However, 2 is irreducible in $\mathbb{Z}[x]$, so (2) is prime in $\mathbb{Z}[x]$ by Proposition 317. (Alternatively, one can prove that (2) is prime but not maximal by showing that

$$\mathbb{Z}[x]/_{(2)} \cong \mathbb{Z}_2[x],$$

which is a domain but not a field.)

Similarly, let A be any UFD, and let $n \geq 2$. Then one can show that the ideal $(x_2,\dots,x_n)$ is prime in $A[x_1,\dots,x_n]$ but not maximal, for example because the quotient

$$A[x_1,\dots,x_n]/_{(x_2,\dots,x_n)} \cong A[x_1]$$

is a domain but not a field.

5.3 Irreducibility in $\mathbb{Z}[x]$

A natural problem is to classify which polynomials of $\mathbb{Z}[x]$ and $\mathbb{Q}[x]$ are irreducible. This would allow us to quickly figure out the minimum polynomials of all elements of $\mathbb{Q}$. Unfortunately, no complete classification is known (although in some sense we do know that "most polynomials are irreducible": see Lindsay Childs, A Concrete Introduction to Higher Algebra, pp. 563–567. The first thing to say is that understanding irreducible polynomials in $\mathbb{Z}[x]$ and in $\mathbb{Q}[x]$ are basically equivalent problems. Recall that a polynomial in $\mathbb{Z}[x]$ is "primitive" if it is not a multiple of an integer $c > 1$ (or equivalently, if the greatest common divisor of all its coefficients is 1).

Theorem 341. *Let $f \in \mathbb{Z}[x]$ be a primitive polynomial. Then*

$$f \text{ is irreducible in } \mathbb{Z}[x] \iff f \text{ is irreducible in } \mathbb{Q}[x].$$

Proof. This is Lemma 334 applied to the case $A = \mathbb{Z}$ and $\mathbb{K} = F(A) = \mathbb{Q}$. □

Remark 342. The "primitive" assumption is necessary for both directions: For example, 3 is invertible in $\mathbb{Q}$ but not in $\mathbb{Z}$, so

- $f = 3$ is irreducible in $\mathbb{Z}[x]$, but not in $\mathbb{Q}[x]$;
- $f = 3x+6 = 3(x+2)$ is irreducible in $\mathbb{Q}[x]$ (because $x+2$ is irreducible and 3 is invertible), but not in $\mathbb{Z}[x]$.

Even if no characterization exists, we have several useful criteria that often allow to conclude irreducibility over $\mathbb{Z}[x]$. Below we collect a few. Let us start with an easy one.

Lemma 343. *Let A be a C-ring with 1. Let $f \in A[x]$ be a monic polynomial of degree 2 or 3.*

$$f \text{ has no roots} \iff f \text{ is irreducible.}$$

Proof. "⇐" If f has a root a, by Ruffini's theorem (Theorem 144) there is a polynomial h such that $f = h \cdot (x-a)$. By Lemma 138, $\deg h = \deg f - 1 \geq 1$. Moreover, the leading coefficient of h is 1 as well. In particular, h is not invertible, because for any polynomial g again by Lemma 138 we have $\deg gh = \deg g + \deg h \geq \deg g + 1 \geq 1$.

"⇒" Suppose $f = g \cdot h$, with g, h not invertible. Since f is monic, the leading coefficients of g and h have product 1, so they are both invertible. Without of generality (i.e. up to multiplying each of them by the leading coefficient of the other), we can assume that both g and h are monic. Were $\deg g = 0$, then g would coincide with its leading coefficient 1 and thus be invertible: a contradiction. So $\deg g \geq 1$ and similarly $\deg h \geq 1$. Also, $\deg g + \deg h = \deg f$ by Lemma 138. But if two positive integers add up to 2 (or 3), one of them must be 1. Say that g has degree 1; then it is of the form $x - b$. This implies that b is a root of f. □

Remark 344. If f is not monic, the direction "⇒" of the previous theorem no longer holds. For example, in $\mathbb{Z}[x]$ the primitive polynomial $(2x-1)\cdot(3x-1)$ is reducible and has no root. However, if we know that A is a field, then having no root implies irreducibility for all polynomials of degree 2 or 3.

Remark 345. If A is not a domain, monic polynomials of degree-1 in $A[x]$ need not be irreducible. For example, inside $\mathbb{Z}_6[x]$ we have

$$x + 1 = (4x + 1)(3x + 1)$$

and neither $4x + 1$ nor $3x + 1$ is invertible, by Corollary 281. So $x + 1$ is reducible in $\mathbb{Z}_6[x]$.

Proposition 346 (Modulo-m Criterion)**.** *Let f be a primitive polynomial in $\mathbb{Z}[x]$ and let F be its leading coefficient. If there exists a integer m such that*

- $\gcd(m, F) = 1$ *and*
- $\overline{f}$ *is irreducible in* $\mathbb{Z}_m[x]$,

then f is irreducible in $\mathbb{Z}[x]$.

Proof. Suppose that $f = g \cdot h$ is a reduction in $\mathbb{Z}[x]$. Let G and H be the leading coefficients of g and h, respectively. The idea is to show that $\overline{f} = \overline{g} \cdot \overline{h}$ is a reduction in $\mathbb{Z}_m[x]$. Clearly $\overline{g} \neq 0$ and $\overline{h} \neq 0$, otherwise f would have all coefficients multiple of m, which contradicts the assumption that f is primitive. The delicate point is to verify that $\overline{g}$ and $\overline{h}$ are not invertible in $\mathbb{Z}_m[x]$. For this we need the assumption that $\gcd(m, F) = 1$: Since $F = G \cdot H$, also G and H are coprime with m. And any polynomial of positive degree whose leading coefficient is coprime with m, cannot be invertible in $\mathbb{Z}_m[x]$ (compare Lemma 138, or use Corollary 281.) So neither $\overline{g}$ nor $\overline{h}$ are invertible, and $\overline{f}$ is reducible. □

Example 347. Consider $f = x^2 + 3x + 1$. In $\mathbb{Z}_3[x]$ we have $\overline{f} = x^2 + 1$, which has no root and therefore is irreducible in $\mathbb{Z}[x]$ by Lemma 343.

Example 348. Consider $f = 8x^3 - 6x - 1$. In $\mathbb{Z}_5[x]$ we have $\overline{f} = 3x^3 + 4x - 1$, which has no root and therefore is irreducible in $\mathbb{Z}[x]$ by Lemma 343. Interestingly, one of the roots of f is the cosine of 20°. In fact, the well-known trigonometric identities

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \quad \text{and} \quad \sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$$

yield formulas for $\cos(2\alpha)$ and for $\sin(2\alpha)$. But since (by the left identity above)

$$\cos(\alpha + 2\alpha) = \cos\alpha\cos(2\alpha) - \sin\alpha\sin(2\alpha),$$

plugging in the formulas for $\cos(2\alpha)$ and for $\sin(2\alpha)$, after some patient calculations, we obtain the so-called "triple–angle formula"

$$\cos(3\alpha) = 4\cos^3(\alpha) - 3\cos\alpha.$$

In particular, say that $\alpha = 20^\circ$ and $x = \cos\alpha$. Then by the triple-angle formula

$$\frac{1}{2} = \cos(60^\circ) = \cos(3\alpha) = 4\cos^3(\alpha) - 3\cos\alpha = 4x^3 - 3x,$$

whence clearing denominators we get $8x^3 - 6x - 1 = 0$.

Example 349. Consider $f = x^5 - x^3 + 6x^2 - 9$. In $\mathbb{Z}_2[x]$ we have $\overline{f} = x^5 + x^3 + 1$. This degree-5 polynomial has no root in $\mathbb{Z}_2$, because $0^5 + 0^3 + 1 = 1$ and $1^5 + 1^3 + 1 = 1$. Still, $\overline{f}$ could be reducible in $\mathbb{Z}_2[x]$ as product of a degree-2 polynomial g and a degree-3 polynomial h. Since there are finitely many options for the coefficients, we can simply check all cases to exclude this, and conclude that f is irreducible. Let us quickly see how to do it. First of all, a "trick": If u, v are the leading coefficients of g and h, respectively, we can assume without loss that $u = v = 1$. (If not, replace g by the polynomial $g' \stackrel{\text{def}}{=} vg$ and h by $h' \stackrel{\text{def}}{=} uh$.) So, suppose we can write

$$x^5 + x^3 + 1 = (x^2 + ax + b) \cdot (x^3 + cx^2 + dx + e).$$

By equating coefficients, this reduces to

$$\begin{cases} 0 &= a + c \\ 1 &= d + ac + b \\ 0 &= e + ad + bc \\ 0 &= ae + bd \\ 1 &= be. \end{cases} \tag{35}$$

Now, the last line tells us that $b \neq 0$ and $e \neq 0$. Since we are in $\mathbb{Z}_2$, this means $b = e = 1$. The second last line tells us that $a + d = 0$, which (in $\mathbb{Z}_2$!) means $d = a$. Similarly, the first line tells us $c = a$. But then the second and the third line read

$$\begin{cases} 1 &= a + a^2 + 1 \\ 0 &= 1 + a^2 + a, \end{cases} \tag{36}$$

a contradiction. Hence, $x^5 + x^3 + 1$ is irreducible in $\mathbb{Z}_2[x]$ and $f = x^5 - x^3 + 6x^2 - 9$ is irreducible in $\mathbb{Z}[x]$, by the mod-2 criterion.

Non-Example 350. The converse of the Irreducibility-mod-m criterion does not hold. For example, for any pair of distinct primes a, b, we will see in Corollary 404 that the polynomial

$$f = x^4 - 2(a+b)x^2 + (a-b)^2$$

is irreducible over $\mathbb{Z}$. However, one can prove (see the exercises) that f is reducible modulo any prime p. For example, in $\mathbb{Z}_2[x]$ we have

$$\overline{f} = x^4 + (a-b)^2 = x^4 - (a-b)^2 = (x^2 + a - b)(x^2 - a + b),$$

while in $\mathbb{Z}_3[x]$ we have

$$\overline{f} = x^4 + (a+b)x^2 + (a-b)^2.$$

Remark 351. The Irreducibility-mod-m criterion might not work if m has a factor in common with the leading coefficient of f. For example, let $g = 2x + 1$ and $h = x^2 + 2$. Consider the product

$$f \stackrel{\text{def}}{=} g \cdot h = 2x^3 + x^2 + 4x + 2.$$

By construction, f is reducible in $\mathbb{Z}[x]$. Of course, the polynomials g and h are not invertible in $\mathbb{Z}[x]$, because they have positive degree and $\mathbb{Z}$ is a domain. But if we pass mod 4, we get

$$\overline{f} = \overline{g} \cdot \overline{h} = 2x^3 + x^2 + 2$$

and now the polynomial $\overline{g} = (2x+1)$ *is* invertible in $\mathbb{Z}_4[x]$, because $(2x+1)(2x+1) = 1$ in $\mathbb{Z}_4[x]$. (Compare also Corollary 281.) One can prove that $\overline{h}$, instead, is irreducible[13]. But then $\overline{f}$ mod 4 is irreducible, even if f itself was not.

Counterexamples modulo a prime p can be generated as follows. Choose an integer n and set $g_n \stackrel{\text{def}}{=} px^n + 1$. Let ℓ be any polynomial in $\mathbb{Z}[x]$ whose class is irreducible in $\mathbb{Z}_p$. (For example $x^2 + x + 1$ for $p = 2$.) Then the polynomial $f_n \stackrel{\text{def}}{=} g_n \cdot \ell$ is reducible by construction. However, in $\mathbb{Z}_p$ $\overline{g_n} = \overline{1}$, so $\overline{f_n} = \overline{\ell}$ is irreducible.

The next result is perhaps the most famous criterion of irreducibility.

Theorem 352 (Eisenstein's criterion, 1850)**.** *Let $f = a_0 + a_1 x + \ldots + a_n x^n$ be a primitive polynomial in $\mathbb{Z}[x]$. If there exists a prime p such that*

- *p does not divide a_n,*
- *p divides each a_i, for $i < n$, and*
- *p^2 does not divide a_0,*

then f is irreducible over $\mathbb{Z}[x]$.

Proof. Say $f = g \cdot h$ with g, h in $\mathbb{Z}[x]$. Let G and H be the leading coefficients of g and h, respectively. If we pass modulo p, in $\mathbb{Z}_p[x]$ we have

$$\overline{f} = \overline{g} \cdot \overline{h}.$$

By the assumptions, the only coefficient of f that "survives" the quotient is a_n, so

$$\overline{f} = \overline{a_m} \cdot x^n. \tag{37}$$

But we are in $\mathbb{Z}_p[x]$, which is a UFD, and each factor x is irreducible (while $\overline{a_m}$ is invertible). So the factorization of $\overline{f}$ given by Equation 37 must be unique. This means that $\overline{g}$ and $\overline{h}$ are of the form

$$\overline{g} = \overline{G} \cdot x^a \quad \text{and } \overline{h} = \overline{H} \cdot x^b,$$

with $a + b = n$. In particular, the constant term of $\overline{g}$ and that of $\overline{h}$ are both zero. This means that in $\mathbb{Z}$, the constant terms of g and h are both multiples of p. But then the constant term a_0 of f, which is the product of the constant terms of g and h, must be a multiple of p^2: A contradiction. □

Example 353. The polynomial $x^5 - 5x^2 + 25x + 15$ is irreducible, because 5 divides all coefficients except the leading term; yet the constant term is not a multiple of $5^2 = 25$.

Example 354. The polynomials of the type $x^n - p$, with p prime, are always irreducible in $\mathbb{Z}[x]$. (Same for polynomials of the type $x^n - A$, where A is any product of distinct primes.) This shows that in $\mathbb{Z}[x]$ and $\mathbb{Q}[x]$ there are irreducible polynomials of any degree. This is going to be a striking difference with $\mathbb{C}[x]$ and with $\mathbb{R}[x]$: Compare Corollary 428 below.

[13] For details, see e.g. J. von zur Gathen, S. Hartlieb, *Factoring modular polynomials*, J. Symbolic Comput. 26 (1998), 583–606.

Non-Example 355. The converse of Eisenstein's criterion does not work. For example, the polynomial $x^2 - 4x + 11$ is irreducible in $\mathbb{Z}[x]$ (and even in $\mathbb{R}[x]$, since its roots are complex), but there is no way to infer it from Eisenstein's criterion. Same for the degree-4 polynomials constructed in Non-example 350 (because their constant term is a perfect square, so if a prime p divides it, p^2 divides it too.)

Remark 356. In the proof of Eisenstein's criterion, at some point we used that $\mathbb{Z}_p[x]$ is a UFD to claim that $\overline{a_m} \cdot x^m$ factors into low-degree polynomials in only one way, as $x^m = (\alpha \cdot x^a) \cdot (\beta \cdot x^b)$, with $a + b = m$. It is easy to see that the UFD assumption, however, is not necessary. This way Eisenstein's criterion can be generalized as follows:

Lemma 357. *Let D be a domain. Let n be a positive integer. The only way x^n can factor inside $D[x]$ is as $x^n = \alpha x^\ell \cdot \beta x^m$, with $\ell + m = n$ and $\alpha\beta = 1$.*

Proof. Suppose $x^n = g \cdot h$ in $D[x]$. Without loss, we can assume that the leading coefficients of g and h are 1. Let us write

$$g = a_0 + a_1 x + \ldots + a_{\ell-1} x^{\ell-1} + x^\ell \qquad \text{and} \qquad h = b_0 + b_1 x + \ldots + b_{m-1} x^{m-1} + x^m.$$

Let i be the smallest integer such that $a_i \neq 0$. Clearly, $i \leq \ell$. Similarly, let j be the smallest integer such that $b_j \neq 0$. Clearly, $j \leq m$. Now $a_i b_j x^{i+j}$ is the lowest-degree nonzero term of $g \cdot h$. But by assumption $g \cdot h$ consists only of one nonzero term, namely, x^m. So it must be

$$a_i b_j \, x^{i+j} = x^n = x^{\ell+m},$$

which forces $i = \ell$ and $j = m$. Hence g and h are monomials. □

Proposition 358 (Eisenstein for domains)**.** *Let A be a domain. Let $f = a_0 + a_1 x + \ldots + a_n x^n$ be a polynomial in $A[x]$. If there exists a prime ideal P such that*

- *a_n does not belong to P,*
- *a_i belongs to P for all $i < n$, and*
- *a_0 does not belong to P^2 (i.e., a_0 cannot be written as product of two elements of P),*

then f cannot be written as product of polynomials of lower degree. In particular, if f is not a multiple of a constant polynomial, then f is irreducible.

Proof. Left as exercise.
(Hint: if you set $D \stackrel{\text{def}}{=} {}^{A}\!/_{P}$, then D is a domain, and $\overline{f} = \overline{a_n} \cdot x^n$.) □

Example 359. Consider the polynomial in two variables $f = 3x^3 + 2xy - y \in \mathbb{Z}[x, y]$. Let us apply Eisenstein's criterion for domains with $A = \mathbb{Z}[x]$ and $P = (y)$. Since $y \in P$, $y \notin P^2$, and $3 \notin P$, it follows that f is irreducible.

Example 360. Let $D \stackrel{\text{def}}{=} \mathbb{Z}[\sqrt{-5}]$. This D is a domain, but not a UFD. It is not difficult to see that the ideal $P \stackrel{\text{def}}{=} (2, 1 + \sqrt{-5})$ is prime and even maximal in D, since

$${}^{D}\!/_{P} \cong \mathbb{Z}_2.$$

Using Eisenstein's criterion for domains, the polynomial $f = (1 + \sqrt{5}) + 2x^3 + x^7$ is irreducible.

We conclude mentioning two modern criteria of irreducibility, one using ideas from combinatorics, and one from calculus. (There are many more we are not citing, see e.g. Győry–Hajdu–Tijdeman, *Irreducibility criteria of Schur-type and Pólya-type*, Monatsh Math 163 (2011), 415–443.)

Lemma 361. *If $a_1, \ldots, a_m$ are distinct integers, $m \geq 5$, then $(x-a_1)(x-a_2)\cdots(x-a_m) \pm 1$ has m distinct real roots.*

Proof. Let $p(x) \stackrel{\text{def}}{=} (x-a_1)(x-a_2)\cdots(x-a_m)$. Let $D \stackrel{\text{def}}{=} \{z+\frac{1}{2} \,:\, z \in \mathbb{Z}\}$ be the set of "halves of odd integers". We claim that for any x in D, $|p(x)| > 1$. To see this, let us start with the most critical case, namely when $m = 5$ and the integers are consecutive. For example say $a_1 = 1$, $a_2 = 2$, ..., $a_5 = 5$. In this case it is easy to see that $\min_{x\in D} |p(x)|$ is achieved at both $x = 2.5$ and $x = 3.5$: One has

$$p(2.5) = |(-1.5)(-0.5)(0.5)(1.5)(2.5)| = \frac{45}{32} > 1.$$

It is clear that if we spread the five numbers a_i further apart, this value $\min_{x\in D} |p(x)|$ will grow. Or also, it will grow if we consider more than five integers. So we proved that for any x in D, $|p(x)| \geq \frac{45}{32} > 1$. In particular, this means that if x is chosen from D, then $p(x) \pm 1$ has the same sign as $p(x)$. As x ranges over D, the sign of $p(x)$ (and thus of $p(x) \pm 1$) changes as follows:

- positive for x large; specifically, for $x \geq a_m + 0.5$, because all factors are positive;
- negative if $a_{m-1} + 0.5 \leq x \leq a_m - 0.5$, because exactly one of the factors is negative;
- positive if $a_{m-2} + 0.5 \leq x \leq a_{m-1} - 0.5$, because exactly two of the factors are negative;
- and so on.

The last sign depends on the parity of m. This shouldn't surprise you: You know from calculus that $\lim_{x\to-\infty} p(x)$ is $-\infty$ if m is odd, and $+\infty$ if m is even.

If m is odd, we conclude using the intermediate value theorem as follows:

- $p(x) - 1$ is negative in the m values $a_1, a_2, \ldots, a_m$ and positive in the $\frac{m+1}{2}$ values $a_1 + 0.5$, $a_3 + 0.5$, ..., $a_m + 0.5$. The intermediate value theorem grants the existence of m roots.
- $p(x) + 1$ is positive in the m values $a_1, a_2, \ldots, a_m$ and negative in the $\frac{m+1}{2}$ values $a_1 - 0.5$, $a_2 + 0.5$, ..., $a_{m-1} + 0.5$. The intermediate value theorem grants then m roots.

If m is even, symmetrically, we conclude as follows:

- $p(x) - 1$ is negative in the m values $a_1, a_2, \ldots, a_m$ and positive in the $\frac{m}{2}$ values $a_1 - 0.5$, $a_2 + 0.5$, ..., $a_m + 0.5$. The intermediate value theorem grants m roots.
- $p(x) + 1$ is positive in the m values $a_1, a_2, \ldots, a_m$ and negative in the $\frac{m}{2}$ values $a_1 + 0.5$, $a_3 + 0.5$, ..., $a_{m-1} + 0.5$. The intermediate value theorem grants m roots. □

Theorem 362. *If $a_1, \ldots, a_m$ are distinct integers, $m > 1$, $f(x) = (x-a_1)(x-a_2)\cdots(x-a_m)-1$ is irreducible in $\mathbb{Q}[x]$ and in $\mathbb{Z}[x]$. If in addition m is odd, then also $f + 2$ is irreducible.*

Proof. By contradiction, suppose $f = gh$, with g, h monic polynomials in $\mathbb{Z}[x]$ of degree less than $m = \deg f$. For each $i \in \{a_1, \ldots, a_m\}$, we have

$$g(i) \cdot h(i) = f(i) = -1.$$

But $g(i)$ and $h(i)$ are integers: so one is 1, and the other is -1. In particular $g(i) = -h(i)$ for all $i \in \{a_1, \ldots, a_m\}$. But then the polynomial $g + h$ has m roots and degree less than m. This implies that $g + h$ is the zero polynomial. So $g = -h$ and $f = -g^2$, a contradiction: the leading coefficient of $-g^2$ is -1. This shows that f is irreducible.

As for $f + 2$, the argument is similar: since $g(i) \cdot h(i) = f(i) + 2 = 1$, either they are both 1 or both -1, so $g - h$ is the zero polynomial. So $f = g^2$, a contradiction with $\deg f$ odd. □

Remark 363. The assumption on m odd for the irreducibility of $f + 2$, cannot be removed in general: For example, the polynomial $f(x) = (x-1)(x-2)(x-3)(x-4) + 1$ is reducible, it's the square of $x^2 - 5x + 5$.

Corollary 364. *For any $m \geq 2$, there exist degree-m polynomials in $\mathbb{Z}[x]$ that are irreducible, but have m real roots.*

Proof. For any $m \geq 5$, choose $f(x) = (x-1)(x-2)\cdots(x-m) - 1$. The existence of the m real roots follows from Lemma 361. Irreducibility follows from Theorem 362, applied to $a_i = i$.

As for $2 \leq m \leq 5$: consider for example

$$g(x) = (x-1)(x-5)\cdots(x-4m+3) - 1.$$

Irreducibility follows from Theorem 362, applied to $a_i = 4i - 3$. For the existence of the m real roots, we mimic the proof of Lemma 361, this time using the set $D' \stackrel{\text{def}}{=} \{3, 7, 9, 11\}$. We obtain that $\min_{x \in D'} |p(x)| = 4$, so again $p(x) \pm 1$ has the same sign of $p(x)$ if x is in D'. Reasoning on sign changes as in the proof of Lemma 361, we conclude. □

Our next gem is my favorite criterion: **Cohn's criterion**, which says that if a prime number p is expressed in base 10 as $p = a_0 + 10a_1 + \ldots + 10^m a^m$, with digits $0 \leq a_i \leq 9$, then the polynomial $f = a_0 + a_1 x + \ldots + a_m x^m$ is irreducible in $\mathbb{Z}[x]$. For example, since 911 is prime, the polynomial $9x^2 + x + 1$ is irreducible. It turns out that there is nothing special about base 10: The theorem works in all bases, although to prove it, we need to assume a fact that we shall prove later, the "fundamental theorem of algebra", namely, that every monic polynomial of degree d in $\mathbb{C}[x]$ factors as product of d monic polynomials of degree one.

Theorem 365 (Cohn; Brillhart–Filaseta–Odlyzko; Ram Murty, 2002 [14])**.** *Let $f = a_0 + a_1 x + \ldots + a_m x^m$ be a polynomial in $\mathbb{Z}[x]$. Define*

$$H \stackrel{\text{def}}{=} \max\left\{ \left|\frac{a_0}{a_m}\right|, \left|\frac{a_1}{a_m}\right|, \ldots, \left|\frac{a_{m-1}}{a_m}\right| \right\}.$$

Suppose that $f(n)$ is prime for some *integer $n \geq H + 2$. Then f is irreducible in $\mathbb{Z}[x]$.*

Proof. Note that $H > 0$, otherwise $f = a_m x^m$ and in particular $f(n) = a_m n^m$, which is never prime; a contradiction. If we view f in $\mathbb{C}[x]$, we claim that for any complex root α of f

$$|\alpha| < H + 1. \tag{38}$$

If $|\alpha| \leq 1$, this is obvious, because by construction $H > 0$. So the interesting case is $|\alpha| > 1$. Since α is a root of f, we know that

$$-a_m \alpha^m = a_0 + a_1\alpha + a_2\alpha^2 + \ldots + a_{m-1}\alpha^{m-1}.$$

Passing to the norm, we obtain

$$|\alpha|^m = \frac{|a_0 + a_1\alpha + a_2\alpha^2 + \ldots + a_{m-1}\alpha^{m-1}|}{|a_m|} \leq \frac{|a_0|}{|a_m|} + \frac{|a_1\alpha|}{|a_m|} + \frac{|a_2\alpha^2|}{|a_m|} + \ldots + \frac{|a_{m-1}\alpha^{m-1}|}{|a_m|},$$

where in the second step we used the so-called "triangular inequality" of the complex norm ($|x + y| \leq |x| + |y|$ for all x, y). But using the definition of H, it is easy to see that in turn

$$\frac{|a_0|}{|a_m|} + \frac{|a_1\alpha|}{|a_m|} + \frac{|a_2\alpha^2|}{|a_m|} + \ldots + \frac{|a_{m-1}\alpha^{m-1}|}{|a_m|} \;\leq\; H \cdot \left(1 + |\alpha| + |\alpha|^2 + \ldots + |\alpha|^{m-1}\right) = H\,\frac{|\alpha|^m - 1}{|\alpha| - 1}.$$

Now since $|\alpha| > 1$ we can multiply the inequality above by $(|\alpha| - 1)$, and obtain

$$|\alpha|^{m+1} - |\alpha|^m \leq H \cdot (|\alpha|^m - 1) < H \cdot |\alpha|^m.$$

[14] M. Ram Murty, *Prime Numbers and Irreducible Polynomials*, Amer. Math. Monthly 109 (2002), pp. 452–458.

Dividing by $|\alpha|^m$, we get $|\alpha| - 1 < H$, which is precisely Inequality 38.

Now we are ready to complete the proof of the theorem. Suppose by contradiction that $f = g \cdot h$, with neither g nor h invertible. Since $f(n)$ is a prime number, either $g(n)$ or $h(n)$ must be ± 1. Without loss of generality, we can assume $g(n) = 1$. Clearly, any complex root of g is also a complex root of f. Let $G \in \mathbb{Z}$ be the leading coefficient of g. Then (for the fundamental theorem of algebra, which we will fully prove in the next chapter) g factors inside $\mathbb{C}[x]$ as

$$g = G \cdot (x - \beta_1)(x - \beta_2) \cdot \ldots \cdot (x - \beta_k),$$

and for what we said above, all β_i's are also roots of f. What happens if we plug in $x = n$? In view of Inequality 38, we have $|\beta_i| < H + 1$ for all i; so in particular $n - |\beta_i| > n - (H + 1)$ for all i. Using again the triangular inequality (in the form $|x - y| \geq |x| - |y|$ for all x, y) we have

$$|n - \beta_i| \geq n - |\beta_i| > n - (H + 1) \quad \text{for all } i. \tag{39}$$

But then

$$1 = |1| = |g(n)| = |G| \cdot |n - \beta_1| \cdot |n - \beta_2| \cdot \ldots \cdot |n - \beta_k| > |G| \cdot (n - H - 1)^k.$$

Yet recall that $|G| \geq 1$, because G is a nonzero integer. Moreover, $(n - H - 1) \geq 1$, because by assumption $n \geq H + 2$. So we have obtained

$$1 > |G| \cdot (n - H - 1)^k \geq 1 \cdot 1^k = 1,$$

a contradiction. □

Corollary 366. *Let $f = a_0 + a_1 x + \ldots + a_{m-1} x^{m-1} + x^m$ be a monic polynomial in $\mathbb{Z}[x]$. If $f(n)$ is prime for* some *integer $n \geq 2 + \max_i |a_i|$, then f is irreducible.*

Example 367. The polynomial $x^4 + 6x^2 + 1$ is irreducible, because $\max\{1, 6, 1\} = 6$, and $f(6 + 2) = f(8) = 4481$, which is already a prime number.

Corollary 368 (Cohn's theorem, weak version)**.** *Let $b \geq 2$ be an integer. If a prime number p is expressed in base b as $p = a_0 + ba_1 + \ldots + b^m a^m$, with digits $0 \leq a_i \leq b - 2$, $a_m \neq 0$, then the polynomial $f = a_0 + a_1 x + \ldots + a_m x^m$ is irreducible.*

Proof. The a_i's are all non-negative; so setting

$$H \stackrel{\text{def}}{=} \max\left\{ \left|\frac{a_0}{a_m}\right|, \left|\frac{a_1}{a_m}\right|, \ldots, \left|\frac{a_{m-1}}{a_m}\right| \right\} = \max\left\{ \frac{a_0}{a_m}, \frac{a_1}{a_m}, \ldots, \left|\frac{a_{m-1}}{a_m}\right| \right\},$$

the assumptions $a_i \leq b - 2$ and $a_m \geq 1$ clearly imply that $H \leq b - 2$. Thus $f(b)$ is prime for an integer $b \geq H + 2$. By Theorem 365, we conclude. □

Note that for $b = 10$ this look like what we promised, but not quite: we can say that $8x^2 + 8x + 7$ is irreducible since 887 is prime, but we cannot use the corollary above to claim that $9x^2 + x + 1$ is irreducible from the fact that 911 is prime, because we are in base 10 and one of the digits of 911 is a nine. A natural attempt to fix this is to try to replace the "$n \geq H + 2$" in the assumptions of Theorem 365, with a weaker "$n \geq H + 1$" assumption. However, this attempt fails: The polynomial

$$g(x) = x^3 - 9x^2 + x - 9$$

is clearly reducible, has all coefficients of absolute value ≤ 9, yet $g(10) = 101$ is a prime number. So we have to follow another path, 'tightening the screws' of the bound on the roots norm a bit more (and using an extra non-negativity assumption on the coefficients a_m and a_{m-1}):

Lemma 369 (Ram Murty). *Let H be a positive integer. Let $f = a_0 + a_1 x + \ldots + a_m x^m$ be a polynomial in $\mathbb{Z}[x]$, with $a_m \geq 1$, $a_{m-1} \geq 0$, and $|a_i| \leq H$ for all $0 \leq i \leq m-2$. Let α be any root of f in $\mathbb{C}$. Then either the real part of α is ≤ 0, or*

$$|\alpha| < \frac{1}{2}(1 + \sqrt{1+4H}). \tag{40}$$

Proof. If $|\alpha| \leq 1.5$, then Equation (40) is trivial, since $\frac{1}{2}(1+\sqrt{1+4H}) \geq \frac{1}{2}(1+\sqrt{5}) > 1.5$. So we can assume that $|\alpha| > 1.5$. By contradiction, suppose the real part of α is positive and $|\alpha| \geq \frac{1}{2}(1+\sqrt{1+4H})$. In particular, by elementary calculus we have

$$1 - \frac{H}{|\alpha|^2 - |\alpha|} \geq 0. \tag{41}$$

Now, recall that for any two real numbers x, y, one has $|x+y| \geq |x| - |y|$, because of the triangle inequality $|x| \leq |x+y| + |-y|$. So since $|\alpha| > 1.5$,

$$\begin{aligned} \left|\frac{f(\alpha)}{\alpha^m}\right| &= \left|a_n + \frac{a_{m-1}}{\alpha} + \frac{a_{m-2}}{\alpha^2} + \ldots + \frac{a_0}{\alpha^m}\right| \geq \left|a_m + \frac{a_{m-1}}{\alpha}\right| - \left|\frac{a_{m-2}}{\alpha^2} + \ldots + \frac{a_0}{\alpha^n}\right| \geq \\ &\geq \left|a_m + \frac{a_{m-1}}{\alpha}\right| - H\left(\frac{1}{|\alpha|^2} + \frac{1}{|\alpha|^3} + \ldots + \frac{1}{|\alpha|^m}\right) > \\ &> \left|a_m + \frac{a_{m-1}}{\alpha}\right| - H\left(\sum_{i=2}^{\infty} \frac{1}{|\alpha|^i}\right) = \left|a_m + \frac{a_{m-1}}{\alpha}\right| - H\frac{1}{|\alpha|^2 - |\alpha|} \end{aligned} \tag{42}$$

where the last step is a geometric series. However, since we are in the case $Re(\alpha) > 0$, and since by assumption $a_m \geq 1$ and $a_{m-1} \geq 0$, we have

$$\left|a_m + \frac{a_{m-1}}{\alpha}\right| \geq Re(a_m + \frac{a_{m-1}}{\alpha}) \geq 1. \tag{43}$$

So putting the three inequalities (41), (42), and (43) together, we conclude that

$$\left|\frac{f(\alpha)}{\alpha^m}\right| > \left|a_m + \frac{a_{m-1}}{\alpha}\right| - H\frac{1}{|\alpha|^2 - |\alpha|} \geq 1 - H\frac{1}{|\alpha|^2 - |\alpha|} \geq 0,$$

which contradicts the fact that α is a root of f. □

Theorem 370 (Cohn's theorem). *Let $b \geq 3$ be an integer. If a prime number p is expressed in base b as $p = a_0 + ba_1 + b^2 a_2 + \ldots + b_m a^m$, with digits $0 \leq a_i \leq b-1$, then the polynomial $f = a_0 + a_1 x + \ldots + a_m x^m$ is irreducible.*

Proof. Suppose $f(x) = g(x)h(x)$, with g, h of positive degree. Plugging in $x = b$, we get

$$p = g(b)h(b),$$

which tells us that (up to relabeling) $g(b) = \pm 1$ and $h(b) = \pm p$. Again, assuming the Fundamental Theorem of Algebra (which we will prove later), we have that

$$g(x) = c(x - \alpha_1) \cdots (x - \alpha_k),$$

where $\alpha_1, \ldots, \alpha_k$ are some of the roots of f, and c is a nonzero integer. But for every root α of f, we know by the previous Lemma that either the real part of α is ≤ 0, in which case clearly $|b - \alpha| \geq b$, or Equation 40 holds, in which case we also have that

$$|\alpha| < \frac{1}{2}(1 + \sqrt{1+4H}) \leq \frac{1}{2}(1 + \sqrt{1+4(b-1)}) \leq b - 1.$$

So either way,

$$|b - \alpha| > 1.$$

But this implies

$$|g(b)| = |c| \cdot |b - \alpha_1| \cdots |b - \alpha_k| > |c| 1^k \geq 1,$$

contradicting the fact that $g(b) = \pm 1$. □

Remark 371. The theorem above holds also for $b = 2$, though extra work is needed. See M. Ram Murty, *Prime Numbers and Irreducible Polynomials*, Amer. Math. Monthly 109 (2002), pp. 452–458. The "converse" of Cohn's theorem ('if the polynomial $f = a_0 + a_1 x + \ldots + a_m x^m$ is irreducible, then the number $a_0 + 10a_1 + \ldots + 10^m a^m$ is prime') is easily shown to be false: For example, $x^3 + 7$ is irreducible by Eisenstein's criterion, yet $1007 = 19 \cdot 53$ is not prime.

Deeper thoughts 372. An open problem in number theory is Bunyakovsky's conjecture[15]:

Conjecture 373 (Bunyakovsky, 1857)**.** *Let $f \in \mathbb{Z}[x]$. $f(k)$ is prime for infinitely many positive integers $n \iff f$ is irreducible and all the numbers $\{f(k) : k \in \mathbb{N}\}$ have no factor common to all of them.*

It is easy to prove "⇒": the set $\{0, 1, \ldots, H+1\}$ is obviously finite. If $f(k)$ takes prime values for infinitely many positive inputs k, then the set $A \stackrel{\text{def}}{=} \{k \geq H + 2 \text{ such that } f(k) \text{ is prime}\}$ is also infinite. In particular, it is non-empty, so by Theorem 365 f is irreducible. Were there a unique prime p such that for all $k \in A$, $f(k) = p$, then the polynomial $f(x) - p$ would have infinitely many roots, a contradiction. So, there are k, k' in A such that $f(k)$ and $f(k')$ are *different* prime numbers. In particular they are coprime, so the gcd of all $f(i)$'s is 1.

As for "⇐", we have no clue. We only know that it is true for $\deg f = 1$, using Dirichlet's theorem 300. The polynomial $x^2 + 1$ is already an open case: Are there infinitely many primes of the form $n^2 + 1$, like 2 and 5? We do not know.

As a final remark: The condition of no-common-factor in Bunyakovski's conjecture is necessary: There are irreducible polynomials f such that only one of the $f(k)$'s is prime. For example, the polynomial

$$f = x^2 + x + 2$$

is irreducible over $\mathbb{Z}$, $\mathbb{Q}$ and even $\mathbb{R}$, because it has complex roots. Yet by Fermat's little theorem, f has the property that if we plug in any nonnegative integer k, the output is always even. So $f(k)$ is prime if and only if $f(k) = 2$. This has only one solution, namely, $k = 0$.

5.4 Exercises

1. Let $f(x) = 2x^7 + 25x^3 + 10x^2 - 30$. Prove that $\mathbb{Q}[x]/(f)$ is a field.

2. For what integer values of b is the polynomial $2x^2 + b$ irreducible in $\mathbb{Z}[x]$?

3. Prove Proposition 358.

4. Let p be a prime. Factor the polynomial $f = x^{p-1} - 1$ in $\mathbb{Z}_p[x]$. (Hint: How many roots can it have at most? How many roots does it have at least?)

5. Is $x^3 + 2$ irreducible in $\mathbb{Z}_7[x]$?

6. Show that $x^4 - x^2 + 2$ is reducible in $\mathbb{Z}_4[x]$. (Hint: Don't forget to prove that the factors are not invertible. Can a monic polynomial in $\mathbb{Z}_4[x]$ be invertible?)

[15] M. Ram Murty, *Prime Numbers and Irreducible Polynomials*, Amer. Math. Monthly 109 (2002), pp. 452–458.

7. Show that $x^4 + 1$ is reducible in $\mathbb{Z}_3[x]$.

8. Is $x^4 + x^3 + x + 2$ irreducible in $\mathbb{Z}_3[x]$?

9. Is $x^4 + 4x^3 + 16x^2 + 34$ irreducible in $\mathbb{Z}$?

10. Show that the only irreducible quadratic polynomials of $\mathbb{Z}_3[x]$ are $x^2 + 1$, $x^2 + x - 1$, $x^2 - x - 1$.

11. Show that in $\mathbb{Z}_2$ all polynomials of the form $x^4 + ax^2 + b$ are reducible.

12. Let p be an odd prime. For any a in $\mathbb{Z}_p$, show that in $\mathbb{Z}_p$ there is an element s such that $a = 2s$. Use this to write $x^4 + ax^2 + b$ in three different ways as

$$x^4 + ax^2 + b = (x^2 + s)^2 - (s^2 - b^2) = (x^2 + b)^2 - (2b - 2s)x^2 = (x^2 - b)^2 - (-2b - 2s)x^2.$$

13. Use Corollary 153 to show that in $\mathbb{Z}_p$, if $2b-2s$ and $(-2b-2s)$ are not squares, then s^2-b^2 must be a square. Thus at least one of $2b - 2s$, $(-2b - 2s)$, and $s^2 - b^2$ is a square in $\mathbb{Z}_p$. Conclude using the previous two exercises that $x^4 + ax^2 + b$ is always reducible mod p.

6 Field extensions

Let us go back to fields. Rather than studying subfields, we are interested in "extensions": that is, fields that contain the one we started with. A first legitimate question is whether we can always construct a field that contains a given one. The answer to this question is "yes", and the reason can be sketched in one line: Given a field $\mathbb{K}$, construct $A = \mathbb{K}[x]$ and then pass to the field of fractions of A; this is a field that contains $\mathbb{K}$ as subring.

But is every field a field extension of some other field? The next theorem answers this question with a "yes, except the $\mathbb{Z}_p$'s and $\mathbb{Q}$".

Theorem 374. *Every field $\mathbb{F}$ is an extension either of of $\mathbb{Q}$ or of some $\mathbb{Z}_p$, with p prime. Moreover, if $\mathbb{F}$ is finite, then it must be an extension of $\mathbb{Z}_p$, for some p prime.*

Proof. Since $1_F \in \mathbb{F}$, then also $1_F + 1_F \in \mathbb{F}$; and by induction,

$$m \cdot 1_F = 1_F + 1_F + \ldots + 1_F \in \mathbb{F}.$$

Let us consider the C-ring homomorphism

$$\begin{array}{rccc} \varphi: & \mathbb{Z} & \longrightarrow & \mathbb{F} \\ & m & \longmapsto & m \cdot 1_F. \end{array}$$

There are two cases to consider:

- If $\ker \varphi \neq \{0\}$, since $\mathbb{Z}$ is a PID (Proposition198) there exists a positive integer m such that $\ker \varphi = (m)$. Now by the First Isomorphism Theorem,

  $$\mathbb{Z}_m = {}^{\mathbb{Z}}\!/_{(m)} = {}^{\mathbb{Z}}\!/_{\ker \varphi} \cong \operatorname{Im} \varphi \subseteq \mathbb{F}.$$

  So $\mathbb{F}$ contains a subring isomorphic to $\mathbb{Z}_m$. Now $\mathbb{F}$ is a field, so in particular a domain: Any subring of it must be a domain too! So by Proposition 115 the number m must be prime. Hence, $\mathbb{F}$ is an extension of $\mathbb{Z}_p$ for some p prime.
- If instead $\ker \varphi = \{0\}$, by the First Isomorphism Theorem (Theorem 238) we have that

  $$\mathbb{Z} \;=\; {}^{\mathbb{Z}}\!/_{(0)} = {}^{\mathbb{Z}}\!/_{\ker \varphi} \cong \operatorname{Im} \varphi \subseteq \mathbb{F}$$

  which in particular implies that $\mathbb{F}$ is infinite. (So for $\mathbb{F}$ finite, we must be in the case above.) Now since $\mathbb{F}$ is a field, for any a, b in $\mathbb{Z}$, with $b \neq 0$, $\mathbb{F}$ contains also the element

  $$\varphi(a) \cdot (\varphi(b))^{-1}.$$

  So we can extend φ to a map

  $$\begin{array}{rccc} \tilde{\varphi}: & \mathbb{Q} & \longrightarrow & \mathbb{F} \\ & \frac{a}{b} & \longmapsto & \varphi(a) \cdot (\varphi(b))^{-1}. \end{array}$$

  This map is injective, because

  $$\tilde{\varphi}(\frac{a}{b}) = \tilde{\varphi}(\frac{a'}{b'}) \iff \varphi(a)(\varphi(b))^{-1} = \varphi(a')(\varphi(b'))^{-1} \iff \varphi(ab') = \varphi(a'b) \iff ab' = a'b,$$

  where the last step follows by the injectivity of φ (we are in the case $\ker \varphi = 0$); so

  $$\tilde{\varphi}(\frac{a}{b}) = \tilde{\varphi}(\frac{a'}{b'}) \iff \frac{a}{b} = \frac{a'}{b'}.$$

  This means that $\mathbb{Q}$ is isomorphic to a subring $\operatorname{Im} \tilde{\varphi}$ of $\mathbb{F}$. □

Remark 375. Not all extension of $\mathbb{Z}_p$ are finite fields: See Corollary 499.

6.1 Algebraic and transcendental numbers

From now on, we will write "Let $\mathbb{K} \subseteq \mathbb{F}$ be fields" as short for "Let $\mathbb{F}$ be a field, and let $\mathbb{K}$ be a field that is a subring of $\mathbb{F}$".

Let $\mathbb{K} \subseteq \mathbb{F}$ be fields, and let $\alpha \in \mathbb{F}$ be an element of the larger field. Consider the Polynomial Evaluation homomorphism

$$\begin{array}{rccc} \varphi_\alpha : & \mathbb{K}[x] & \longrightarrow & \mathbb{F} \\ & f & \longmapsto & f(\alpha) \end{array}$$

which takes a polynomial and "plugs in α for x" (cf. Example 183). The kernel of this homomorphism is an ideal of $\mathbb{K}[x]$, which by Theorem 217 is a PID. Hence, $\ker \varphi_\alpha$ is a principal ideal: There is a polynomial p such that

$$\ker \varphi_\alpha = (p).$$

Definition 376. Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let $\alpha \in \mathbb{F}$. We say that

- α is *transcendental over* $\mathbb{K}$, if $\ker \varphi_\alpha = (0)$; in other words, if the only polynomial of $\mathbb{K}[x]$ that vanishes in α is the zero polynomial.
- α is *algebraic over* $\mathbb{K}$, if it is not transcendental; or in other words, if some non-zero polynomial of $\mathbb{K}[x]$ vanishes in α.

Definition 377. Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let $\alpha \in \mathbb{F}$. If α is algebraic over $\mathbb{K}$, we call *minimum polynomial of α (over $\mathbb{K}$)* the *unique monic* polynomial $p \in \mathbb{K}[x]$ such that

$$\ker \varphi_\alpha = (p),$$

and if $\deg p = d$, we say that "α has degree d over $\mathbb{K}$".

Note that the "degree of α over $\mathbb{K}$" is always an integer ≥ 1, because the minimum polynomial cannot have degree zero. (Nonzero constant polynomials do not vanish.)

Remark 378. Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let $\alpha \in \mathbb{F}$. Suppose we see a (monic) polynomial f of $\mathbb{K}[x]$ that vanishes in α. Then we cannot conclude that f is the minimum polynomial of α; what we know for sure though, is that f is a *multiple* of the minimum polynomial of α. In fact,

$$f \text{ vanishes on } \alpha \iff f \text{ is a multiple of } p.$$

Example 379. Consider the fields $\mathbb{Q} \subseteq \mathbb{R}$. Clearly $\alpha = \frac{2}{7}$ is algebraic over $\mathbb{Q}$: It suffices to consider the degree-1 monic polynomial $x - \frac{2}{7}$, which vanishes in $\frac{2}{7}$ and belongs to $\mathbb{Q}[x]$. Let p be the mimimum polynomial. We claim that $p = x - \frac{2}{7}$. In fact,

$$x - \frac{2}{7} \in \ker \varphi_{\frac{2}{7}} \overset{\text{def}}{=} (p).$$

On the other hand, by Ruffini's theorem (Theorem 144), any polynomial p that vanishes in $\frac{2}{7}$ is a multiple of $x - \frac{2}{7}$. So

$$(x - \frac{2}{7}) = (p).$$

Since they are both monic, $x - \frac{2}{7} = p$. So $\alpha = \frac{2}{7}$ is algebraic over $\mathbb{Q}$ of degree 1.

In fact, a more general fact holds:

Lemma 380. *Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Then $\{$elements of degree 1 over $\mathbb{K}\} = \mathbb{K}$.*

Proof.

"$\supset$" Any element α of $\mathbb{K}$ is algebraic over $\mathbb{K}$, because the polynomial $x - \alpha$ vanishes in it. To show that α is algebraic of degree one, let us show that $x - \alpha$ is the minimum polynomial. This is a consequence of Ruffini's theorem (Theorem 144): Any polynomial that vanishes on α is a multiple of $x - \alpha$, and thus belongs to the ideal $(x - \alpha)$.

"$\subseteq$" If an element α of $\mathbb{F}$ has degree 1 over $\mathbb{K}$, then its minimum polynomial must be of the form $x - \alpha$. But all coefficients of the minimum polynomial are by definition in $\mathbb{K}$; so $\alpha \in \mathbb{K}$. □

Here is a useful criterion to understand whether a polynomial that vanishes on α is indeed the minimum polynomial of α, or just a multiple of it.

Proposition 381. *Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let $\alpha \in \mathbb{F}$. Let f be a monic, nonzero polynomial that vanishes in α. The following are equivalent:*

(1) *f is the minimum polynomial of α;*

(2) *f has the same degree of the minimum polynomial of α;*

(3) *all polynomials of degree smaller than f do not vanish on α;*

(4) *f is irreducible;*

(5) *the quotient C-ring $\mathbb{K}[x]/(f)$ is a field.*

Proof.
(1) $\Rightarrow$ (2). Obvious.
(2) $\Rightarrow$ (3). Any nonzero polynomial g that vanishes on α is a multiple of the minimum polynomial p; since $\mathbb{K}[x]$ is a domain, it follows that $\deg g \geq \deg p = \deg f$.
(3) $\Rightarrow$ (4). Since it vanishes in α, f is not an invertible constant. By contradiction, suppose

$$f = a \cdot b \text{ with neither } a \text{ nor } b \text{ invertible.}$$

In particular, $\deg a \geq 1$ and $\deg b \geq 1$. Since we are in a domain, by Lemma 136

$$\deg a = \deg f - \deg b \leq \deg f - 1 \text{ and } \deg b = \deg f - \deg a \leq \deg f - 1.$$

So having smaller degree than f, by the assumption both a and b do *not* vanish on α. Yet

$$0 = f(\alpha) = a(\alpha) \cdot b(\alpha),$$

a contradiction with the fact that $\mathbb{F}[x]$ is a domain.
(4) $\Rightarrow$ (1). Let p be the minimum polynomial of α. Since f vanishes on α, by definition $f \in (p)$. So $f = p \cdot g$ for some polynomial g. Since f is irreducible and p is not invertible, g must be invertible. So $(f) = (p)$. But they are both monic, so $f = p$.
(4) $\Leftrightarrow$ (5). This is an immediate consequence of Proposition 295 and the fact that $\mathbb{K}[x]$ is a PID.

□

This justifies the name "minimum polynomial": It is the smallest degree polynomial of $\mathbb{K}[x]$ that vanishes on α.

Example 382. Consider the fields $\mathbb{R} \subseteq \mathbb{C}$. We claim that $i \stackrel{\text{def}}{=} \sqrt{-1}$ is algebraic over $\mathbb{R}$ of degree 2. In fact, the polynomial x^2+1 vanishes on i and has degree 2. This already tells us that i is algebraic over $\mathbb{R}$ of degree ≤ 2. Moreover, the quotient

$$\mathbb{R}[x]/(x^2+1) \equiv \mathbb{C}$$

is a field, so by Proposition 381 x^2+1 is the minimum polynomial of i.

Example 383. The real numbers $\sqrt{2}$, $\sqrt[3]{2}$, $\sqrt[4]{2}$... are not in $\mathbb{Q}$, but they are algebraic over $\mathbb{Q}$. Consider the polynomial $x^n - 2$. For any $n \geq 2$, this polynomial is irreducible by Eisenstein's criterion (cf. Example 354.) Via Proposition 381, this tells us that $\sqrt[n]{2}$ has degree n.

But what about $\sqrt{2}+\sqrt{3}$? With some effort, one can show that $x^4 - 10x^2 + 1$ vanishes in $\sqrt{2}+\sqrt{3}$. Can you prove that this polynomial is irreducible? We will see a new method in the next section. (If you wish to see the solution, go to Example 405.)

Deeper thoughts 384. In general, it hard to prove that a certain number of $\mathbb{R}$ is transcendental over $\mathbb{Q}$. Here are a few famous results:

- In 1844, Joseph Liouville discovered the first such number, $\sum_{n=1}^{\infty} \frac{1}{10^{n!}} \approx 0.110001\ldots$
- In 1873, Hermite proved that e is transcendental over $\mathbb{Q}$. Recall from Calculus (specifically, from the Taylor expansion of the exponential e^x) that $e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.71828\ldots$.
- In 1882, Ferdinand von Lindemann proved that π is transcendental over $\mathbb{Q}$. Recall from Calculus (specifically, from the Taylor series of $f(x) = 4\arctan x$ evaluated at $x = 1$) that $\pi = \sum_{n=0}^{\infty} (-1)^n \frac{4}{2n+1} \approx 3.14159\ldots$.
- It follows from Lindemann's method that (A) $\ln a$ with $a \notin \{0, 1\}$ is transcendental over $\mathbb{Q}$ as long as a is algebraic over $\mathbb{Q}$; (B) $\sin b$, $\cos b$ and $\tan b$ are all transcendental over $\mathbb{Q}$, as long as b is algebraic over $\mathbb{Q}$ and different from 0. In particular, $\ln 2, \ln 3, \ldots$, $\sin 2, \sin 3, \ldots$, $\cos 2, \cos 3, \ldots$, and ratios like $\frac{\ln 2}{\ln 3}$, are all transcendental over $\mathbb{Q}$.
- In 1900, in his list of 23 Problems for the new century, David Hilbert thought of a new possible way to generate transcendental numbers. He posed the question: *Is it true that if a, b are algebraic numbers over $\mathbb{Q}$, $a \notin \{0, 1\}$ and $b \notin \mathbb{Q}$, then a^b is always transcendental over $\mathbb{Q}$?* For example, is $2^{\sqrt{2}}$ transcendental over $\mathbb{Q}$? Hilbert admitted that he wasn't expecting to witness an answer to this problem during his lifetime; but in 1930, Rodion Kuzmin proved that indeed $2^{\sqrt{2}}$ is transcendental over $\mathbb{Q}$. In 1934, Aleksandr Gelfond and Theodor Schneider independently extended Kuzmin's method and were able to answer Hilbert's question with a full "yes". Their work is now known as Gelfond–Schneider theorem. In particular, $2^{\sqrt{2}}, 2^{\sqrt{3}}, 2^{\sqrt[3]{2}}, 2^i$, and $\sqrt{2}^{\sqrt{2}}$, are all transcendental over $\mathbb{Q}$.
  (Careful: if only one of a, b is algebraic over $\mathbb{Q}$, it could be that a^b is still algebraic. For example, if $a = 2^{\sqrt{3}}$ and $b = \sqrt{3}$, then $a^b = 2^3 = 8$. If $a = 3$ and $b = \frac{\ln 2}{\ln 3}$, $a^b = 2$. Obviously, if a if transcendental over $\mathbb{Q}$ and n is an integer, then a^n is still transcendental over $\mathbb{Q}$; so $\pi^2, \pi^3, \ldots$ are all transcendental over $\mathbb{Q}$.)
- There's still plenty of things to discover on transcendental numbers. For example, we know from Lindemann's work that e^π is transcendental. But as of today, we have no clue whether π^e is transcendental. Same for $e \cdot \pi$, $e + \pi$, e^e, or π^π. It is conjectured that they all are transcendental. See also Corollary 401.
- A case where we know even less is Euler-Mascheroni's constant, introduced by Euler in 1735, and defined as

$$\gamma \stackrel{\text{def}}{=} \lim_{n\to\infty} \left(-\ln n + \sum_{k=1}^{n} \frac{1}{k} \right) \approx 0.5772\ldots$$

This number γ is conjectured to be transcendental, but we don't even know whether it is rational or not!

As you can see, if a real number is expressed as a limit of a series, it is typically hard to decide whether such number is transcendental. There is a remarkable twist though. This does not mean that transcendental numbers are rare. In fact, the opposite is true: Most real numbers are transcendental over $\mathbb{Q}$! It's the algebraic ones who are rare:

Theorem 385 (Cantor)**.** *The set of numbers algebraic over* $\mathbb{Q}$ *is countable. In particular, it has measure* 0 *in* $\mathbb{R}$*.*

Sketch of proof. Every polynomial in $\mathbb{Q}[x]$ is determined by its coefficients, which are finitely many elements in $\mathbb{Q}$; and $\mathbb{Q}$ is countable. Hence, there are countably many polynomials in $\mathbb{Q}[x]$. (Here we used the fact the finite union of countable sets is countable, which is not hard to prove.) Each polynomial of $\mathbb{Q}[x]$ has finitely many roots; which implies that all possible roots of all possible polynomials in $\mathbb{Q}[x]$ are countably many. Hence, the set of numbers algebraic over $\mathbb{Q}$ is countable. By Theorem 61, every countable subset of $\mathbb{R}$ has measure zero in $\mathbb{R}$. □

In particular, if we choose a random point in the segment $[0, 1]$, what is the probability that the point (has a coordinate that) is algebraic over $\mathbb{Q}$? By definition, the probability should be a ratio of the measures of favorable cases versus all cases, which yields a fraction $\frac{0}{1}$. So the probability to hit an algebraic number is 0. Which means that with probability 1, we have picked a transcendental number!

6.2 Vector spaces, field extensions, and degree

In this section, given $\mathbb{K} \subseteq \mathbb{F}$ fields, we recall some linear algebra to define $[\mathbb{F} : \mathbb{K}]$, as the dimension over $\mathbb{K}$ of $\mathbb{F}$ as $\mathbb{K}$-vector space. Moreover, for any $\alpha \in \mathbb{F}$, we introduce the two C-rings $\mathbb{K}(\alpha)$ and $\mathbb{K}[\alpha]$. One of them is by construction a field, the other – a priori – is not; but we will see that sometimes they coincide. In fact, we will see that they coincide precisely when α is algebraic.

Definition 386 (Vector spaces)**.** Let $\mathbb{K}$ be a field. A $\mathbb{K}$-*vector space* is a set V together with an operation $(v, w) \mapsto v + w$ from $V \times V$ to V (called "*sum*") and an operation $(\lambda, v) \mapsto \lambda v$ from $\mathbb{K} \times V$ to V (called "*scalar multiplication*") that satisfy the following axioms:

(VS1) $+$ is associative. That is, for all u, v, w in V, $(u + v) + w = u + (v + w)$.

(VS2) $+$ is commutative. That is, for all v, w in V, $v + w = w + v$.

(VS3) $+$ has a (unique) neutral element. That is, $\exists! z$ in V such that for all v in V, $v + z = v$. We shall call this z "0".

(VS4) Every element has a unique additive inverse. That is, for all v in V $\exists! y$ in V such that $v + y = 0$. We shall call this y "$-v$".

(VS5) For all v in V and for all λ, μ in $\mathbb{K}$, $\lambda(\mu v) = (\lambda\mu)v$.

(VS6) For all v in V and for all λ, μ in $\mathbb{K}$, $(\lambda + \mu)v = \lambda v + \mu v$.

(VS7) For all v, w in V and for all λ in $\mathbb{K}$, $\lambda(v + w) = \lambda v + \lambda w$.

(VS8) For all v in V, $1v = v$.

Example 387. If $\mathbb{K} \subseteq \mathbb{F}$ are fields, then the larger field $\mathbb{F}$ has a natural structure of vector space over the smaller field. In fact, there is an obvious scalar multiplication

$$\begin{aligned} \mathbb{K} \times \mathbb{F} &\longrightarrow \mathbb{F} \\ (\lambda, \alpha) &\mapsto \lambda\alpha. \end{aligned}$$

Because this external operation on $\mathbb{K}$ is "stolen" from the internal multiplication in $\mathbb{F}$, it is easy to see that it satisfies distributivity and the other axioms of vector spaces.

Recall the following terms of the theory of vector spaces:

Definition 388. Let $\mathbb{K}$ be a field. Let V be a $\mathbb{K}$-vector space. Let $W \subset V$. We say that W is:

- *linearly independent* (LI), if for any finite subset $\{w_1, \ldots, w_n\}$ of W the following implication holds:

$$\text{if } a_1 w_1 + \ldots + a_n w_n = 0 \text{ for some } a_1, \ldots, a_n \in \mathbb{K}, \quad \text{then } a_1 = a_2 = \ldots = a_n = 0;$$

- *a set of generators (for V)*, in which case we write $\operatorname{span}(W) = V$, if for every v in V one can find finitely many scalars $a_1, \ldots, a_n$ of $\mathbb{K}$ and vectors $w_1, \ldots, w_n$ of W such that

$$v = a_1 w_1 + \ldots a_n w_n.$$

- *a basis*, if it is a LI set of generators.

More generally, if $W \subset V$, we denote by $\operatorname{span}(W)$ the smallest vector subspace of V containing W. Note that if $W = \{w_1, \ldots, w_n\}$ is finite, then $\operatorname{span}(W) = \{a_1 w_1 + \ldots + a_n w_n \ : \ a_i \in \mathbb{K}\}$. Recall also the following fundamental theorem of linear algebra:

Lemma 389 (Steinitz Exchange Lemma)**.** *Let $\mathbb{K}$ be a field. Let V be a $\mathbb{K}$-vector space. Let $U = \{u_1, \ldots, u_m\}$ be a LI set in V. Let $W = \{w_1, \ldots, w_n\}$ be a set of generators for V. Then for all $k \in \{0, \ldots, m\}$, one has $k \leq n$, and up to relabeling the elements of W, one has*

$$V = \operatorname{span}(u_1, \ldots, u_k, w_{k+1}, \ldots, w_n).$$

In particular (choosing $k = m$), one has $m \leq n$.

Proof. By induction. The case $k = 0$ is clear. Suppose the claim holds for some $k < m$. Then

$$u_{k+1} \in V = \operatorname{span}(u_1, \ldots, u_k, w_{k+1}, \ldots, w_n),$$

so we can find elements $\mu_1, \ldots, \mu_n$ in $\mathbb{K}$ such that

$$u_{k+1} = \sum_{j=1}^{k} \mu_j u_j \ + \ \sum_{j=k+1}^{n} \mu_j w_j. \tag{44}$$

But since the u_j's are linearly independent, at least one of $\{\mu_{k+1}, \ldots, \mu_n\}$ must be nonzero. This already implies $k + 1 \leq n$. Up to reordering, suppose $\mu_{k+1} \neq 0$. Then

$$w_{k+1} \text{ is in } \operatorname{span}(u_1, \ldots, u_k, u_{k+1}, w_{k+2}, \ldots, w_n).$$

This implies, together with Equation 44, that

$$\operatorname{span}(u_1, \ldots, u_k, u_{k+1}, w_{k+2}, \ldots, w_n) = \operatorname{span}(u_1, \ldots, u_k, w_{k+1}, w_{k+2}, \ldots, w_n) = V. \qquad \square$$

Theorem 390. *Let $\mathbb{K}$ be a field. Let V be a $\mathbb{K}$-vector space. Any two bases of V have the same cardinality, which we shall call "$\dim_{\mathbb{K}} V$".*

Proof. If V has a finite set of generators, the claim is an immediate consequence of the Steinitz Exchange Lemma. Now suppose that V has a LI set $(u_i)_{i\in I}$ and a set of generators $(w_j)_{j\in J}$, with J infinite. We claim that also in this case, the cardinality of I is smaller than or equal to the cardinality of J. From the claim, the conclusion follows. So let us prove the claim by contradiction. Suppose the cardinality of I is larger. Every LI set can be expanded to a maximal LI set, which is therefore a basis; in the expansion the cardinality can only increase. So without loss, we can assume that the $(u_i)_{i\in I}$ are a basis for V. Thus every element of the other set of generators can be written as

$$w_j = \sum_{i\in E_j} a_{ij} u_i, \text{ for some finite subsets } E_j \text{ of } I. \tag{45}$$

But since J is infinite, the cardinality of the union of the E_j's is the same as the cardinality of J, and thus smaller than the cardinality of I. Choose an element $i_0 \in I$ that does not belong to any of the subsets E_j. Since u_{i_0} is in $V = \text{span}(w_j)_{j\in J}$, we can write u_{i_0} as linear combination of the w_j; and using Equation 45, in turn we can express each w_j as a linear combination of all the u_i's, except u_{i_0}. So in conclusion, we can write u_{i_0} as linear combination of the other u_i's, a contradiction with linear independence of the u_i's. □

Definition 391. Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. We define the "degree of $\mathbb{F}$ over $\mathbb{K}$" as

$$[\mathbb{F} : \mathbb{K}] \stackrel{\text{def}}{=} \dim_{\mathbb{K}} \mathbb{F}.$$

Example 392. Consider the fields $\mathbb{R} \subseteq \mathbb{C}$. Since by definition $\mathbb{C} = \{a + bi \text{ such that } a, b \in \mathbb{R}\}$, it is clear that

$$[\mathbb{C} : \mathbb{R}] = \dim_{\mathbb{R}} \mathbb{C} = 2.$$

Example 393. The dimension is not always finite. Consider the fields $\mathbb{Q} \subseteq \mathbb{R}$. Choose any positive integer n. Choose any element of $\mathbb{R}$ transcendental over $\mathbb{Q}$, for example π. Then the field $\mathbb{R}$ contains the n elements $1, \pi, \pi^2, \pi^3, \ldots, \pi^{n-1}$. Now let

$$a_0 + a_1\pi + a_2\pi^2 + a_3\pi^3 + \ldots + a_{n-1}\pi^{n-1} = 0$$

be any linear combination of these elements with coefficients in $\mathbb{Q}$. Then the polynomial $a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1}x^{n-1}$, by construction, vanishes when you plug in π for x. But since π is transcendental, the only possible polynomial that vanishes on π is the zero polynomial. So all a_i must be zero. Hence, the n elements $1, \pi, \pi^2, \pi^3, \ldots, \pi^{n-1}$ are linearly independent over $\mathbb{Q}$. In particular $\dim_{\mathbb{Q}} \mathbb{R} \geq n$. But since n was chosen arbitrarily, this proves that

$$[\mathbb{R} : \mathbb{Q}] = +\infty.$$

For the next theorem we adopt the convention that

- for any integer n, the products $n \cdot \infty$ and $\infty \cdot n$ both equal ∞, and
- ∞ times ∞ also equals ∞.

Theorem 394 (Lagrange)**.** *Let $\mathbb{K} \subseteq \mathbb{G} \subseteq \mathbb{F}$ be fields. Then*

$$[\mathbb{F} : \mathbb{K}] = [\mathbb{F} : \mathbb{G}] \cdot [\mathbb{G} : \mathbb{K}].$$

Proof. Let $v_1, \ldots, v_n$ be a basis of $\mathbb{F}$ over $\mathbb{G}$. Let $w_1, \ldots, w_m$ be a basis of $\mathbb{G}$ over $\mathbb{K}$. The idea is to show that the

$$\{w_i \cdot v_j \text{ such that } i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}\}$$

are a basis of $\mathbb{F}$ over $\mathbb{K}$. (The same argument extends also to the infinite-dimensional case; we leave this to the reader.)

First let us show that they are generators. Let $x \in \mathbb{F}$. Since the v_i generate $\mathbb{F}$ over $\mathbb{G}$, we can write

$$x = \sum_{j=1}^{n} g_j v_j \text{ for some } g_j \in \mathbb{G}. \tag{46}$$

In turn, because the w_i are generators, each g_j can be written as

$$g_j = \sum_{i=1}^{m} k_{i,j} w_i \text{ for some } k_{i,j} \in \mathbb{G}.$$

Plugging this back into Equation 46, we obtain

$$x = \sum_{j=1}^{n} \sum_{i=1}^{m} k_{i,j} \cdot w_i v_j \text{ for some } k_{i,j} \in \mathbb{G}.$$

It remains to show that the $w_i v_j$'s are independent over $\mathbb{K}$. So, take any linear combination of the type

$$0 = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{i,j} \cdot w_i v_j, \qquad \text{with } a_{i,j} \in \mathbb{K}. \tag{47}$$

Now for each j set $g_j \stackrel{\text{def}}{=} \sum_i a_{i,j} \cdot w_i$. Since these coefficients $a_{i,j}$'s are in $\mathbb{K} \subseteq \mathbb{G}$, and the w_i are also in $\mathbb{G}$, it follows that each g_j is in $\mathbb{G}$. And since Equation 47 can be rewritten as

$$0 = \sum_{j=1}^{n} g_j \cdot v_j,$$

it yields a linear combination of the v_j with coefficients in $\mathbb{G}$. Since the v_j's are linear independent over $\mathbb{G}$, it follows that $g_j = 0$ for all j. Recalling the definition of g_j, this means that for all j we have

$$\sum_i a_{i,j} \cdot w_i = 0,$$

which yields now a linear combination of the w_i with coefficients in $\mathbb{K}$. But the w_i are also linearly independent. Hence, for all i and for all j we can conclude that $a_{i,j} = 0$. □

Now we are ready to connect this with our study of field extensions.

Definition 395. Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$. Set

$$\mathbb{K}(\alpha_1, \ldots, \alpha_n) \stackrel{\text{def}}{=} \text{ the smallest subfield of } \mathbb{F} \text{ that contains } \mathbb{K} \text{ and all the } \alpha_i\text{'s.}$$

Definition 396. Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let $\alpha \in \mathbb{F}$. Let $\varphi_\alpha : \mathbb{K}[x] \to \mathbb{F}$ be the C-ring homomorphism that "plugs in $x = \alpha$". Set

$$\mathbb{K}[\alpha] \stackrel{\text{def}}{=} \operatorname{Im} \varphi_\alpha = \{a_0 + a_1\alpha + \ldots + a_n\alpha^n \text{ such that } n \in \mathbb{N}, a_i \in \mathbb{K}\}.$$

Note the following facts:

- by definition $\mathbb{K}(\alpha)$ is a field, whereas $\mathbb{K}[\alpha]$ is a C-ring;
- $\mathbb{K}[\alpha]$ contains $\mathbb{K}$ (set $n = 0$ in the definition) and also α (set $n = 1$, $a_0 = 0$ and $a_1 = 1$ in the definition); it also contains all *positive* powers of α (but α might not be invertible...);
- any field that contains elements $a_0, \ldots, a_n$ and α contains also $a_0 + a_1\alpha + \ldots + a_n\alpha^n$;
- in particular, the smallest field that contains $\mathbb{K}$ and α contains also the C-ring $\mathbb{K}[\alpha]$, so
$$\mathbb{K}[\alpha] \subseteq \mathbb{K}(\alpha).$$

The next result clarifies precisely when this inclusion is strict and when it is an equality.

Theorem 397. *Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let $\alpha \in \mathbb{F}$.*

(1) *α is transcendental over $\mathbb{K}$ $\iff$ $\mathbb{K}[\alpha] \cong \mathbb{K}[x]$ $\iff$ $\mathbb{K}[\alpha]$ is not a field.*

(2) *α is algebraic over $\mathbb{K}$ $\iff$ $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$ $\iff$ $\mathbb{K}[\alpha]$ is a field.*

Proof. Consider $\varphi_\alpha : \mathbb{K}[x] \to \mathbb{F}$, the C-ring homomorphism that plugs in $x = \alpha$. By the first isomorphism theorem,
$$\mathbb{K}[\alpha] \stackrel{\text{def}}{=} \operatorname{Im}\varphi_\alpha \cong {}^{\mathbb{K}[x]}\!/_{\ker\varphi_\alpha}.$$
Now:
$$\alpha \text{ is transcendental} \iff \ker\varphi_\alpha = (0) \iff {}^{\mathbb{K}[x]}\!/_{\ker\varphi_\alpha} = {}^{\mathbb{K}[x]}\!/_{(0)} = \mathbb{K}[x].$$
Since x is not invertible in $\mathbb{K}[x]$, it follows that if α is transcendental, $\mathbb{K}[\alpha]$ is not a field. In particular, $\mathbb{K}[\alpha] \neq \mathbb{K}(\alpha)$.

If instead α is algebraic, then $\ker\varphi_\alpha = (p)$, where p is the minimum polynomial. By Proposition 381, this is equivalent to $\mathbb{K}[\alpha] \cong {}^{\mathbb{K}[x]}\!/_{\ker\varphi_\alpha} = {}^{\mathbb{K}[x]}\!/_{(p)}$ being a field. Also, $\mathbb{K}[\alpha] \subseteq \mathbb{K}(\alpha)$. But by the definition of $\mathbb{K}(\alpha)$, we must have $\mathbb{K}[\alpha] = K(\alpha)$. □

Theorem 398. *Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let α be an element of $\mathbb{F}$.*

(1) *If α is algebraic over $\mathbb{K}$, then*
$$[\mathbb{K}(\alpha) : \mathbb{K}] = \deg p,$$
where p is the minimum polynomial of α.

(2) *If α is transcendental over $\mathbb{K}$, then $[\mathbb{K}(\alpha) : \mathbb{K}] = +\infty$.*

Proof.
(1) Set $d \stackrel{\text{def}}{=} \deg p$ for brevity. The idea is to show that
$$1,\ \alpha,\ \alpha^2,\ \ldots,\ \alpha^{d-1}$$
form a basis for $\mathbb{K}(\alpha)$ over $\mathbb{K}$.

First let us show that they are generators. Since α is algebraic, $\mathbb{K}(\alpha) = \mathbb{K}[\alpha] \stackrel{\text{def}}{=} \operatorname{Im}\varphi_\alpha$, by Theorem 397. So it suffices to show that *any* element of $\operatorname{Im}\varphi_\alpha$ can be written as combination of $\{1, \alpha, \alpha^2, \ldots, \alpha^{d-1}\}$ with coefficients in $\mathbb{K}$. Now, choose any $f \in \mathbb{K}[x]$. Since p is monic, we can divide f by p; the remainder is either 0, or a polynomial of degree less than d. Either way, we can write
$$f = q \cdot p + (r_0 + r_1 \cdot x + r_2 \cdot x^2 + \ldots + r_{d-1} \cdot x^{d-1}), \text{ for some } r_i \in \mathbb{K}.$$

The r_i may or may not be zero; of course, they depend on f. Now let us plug in α for x. Since $p(\alpha) = 0$ (by definition of minimum polynomial), we get

$$f(\alpha) = 0 + r_0 + r_1 \cdot \alpha + r_2 \cdot \alpha^2 + \ldots + r_{d-1} \cdot \alpha^{d-1}.$$

Hence $\operatorname{Im} \varphi_\alpha$ is generated by $\{1, \alpha, \alpha^2, \ldots, \alpha^{d-1}\}$.
Next, let us show linear independence. This resembles the game we played in Example 393. Suppose

$$a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3 + \ldots + a_{d-1}\alpha^{d-1} = 0$$

is any linear combination of the powers of α with coefficients in $\mathbb{Q}$. Then the polynomial

$$g \stackrel{\text{def}}{=} a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1}$$

vanishes when you plug in α for x. In other words, this polynomial g belongs to $\ker \varphi_\alpha$, which (by definition of minimum polynomial) is the ideal (p). So g is a multiple of p. If $g \neq 0$ we immediately get a contradiction: the degree of g is at most $d-1$, while the degree of p is d. So we must have $g = 0$. This implies that all a_i's are zero. Hence, the powers of α are linearly independent.

(2) The idea is to show, exactly as in Example 393, that for any positive integer n, the elements $1, \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{n-1}$ are linearly independent over $\mathbb{Q}$. This proves that $\dim_{\mathbb{K}} \mathbb{K}(\alpha) \geq n$, for any n.

□

Corollary 399. *Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. If $[\mathbb{F} : \mathbb{K}]$ is finite, then every element of $\mathbb{F}$ is algebraic over $\mathbb{K}$, and its degree is a divisor of $[\mathbb{F} : \mathbb{K}]$.*

Proof. Let d be the degree of α. By Theorem 398, $d = [\mathbb{K}(\alpha) : \mathbb{K}]$. By Theorem 394,

$$[\mathbb{F} : \mathbb{K}] = [\mathbb{F} : \mathbb{K}(\alpha)] \cdot [\mathbb{K}(\alpha) : \mathbb{K}] = [\mathbb{F} : \mathbb{K}(\alpha)] \cdot d.$$ □

Remark 400. The converse of the previous corollary does not work: It could be that every element of $\mathbb{F}$ is algebraic over $\mathbb{K}$, yet $[\mathbb{F} : \mathbb{K}] = +\infty$. An example will be given by $\overline{\mathbb{Q}}$ over $\mathbb{Q}$, cf. Proposition 413 below.

Corollary 401. *At least one of $\pi + e$ and πe is transcendental over $\mathbb{Q}$.*

Proof. Recall that if a, b are real numbers, by $\mathbb{Q}(a, b)$ we denote the smallest subfield of $\mathbb{R}$ containing $\mathbb{Q}$, a, and b. By contradiction, suppose that $\pi + e$ and πe are both algebraic. Then $\mathbb{K} \stackrel{\text{def}}{=} Q(\pi + e, \pi e)$ has finite degree over $\mathbb{Q}$. Set $k \stackrel{\text{def}}{=} [\mathbb{K} : \mathbb{Q}]$. Now, there is a degree-2 polynomial

$$f = x^2 - (\pi + e)x + \pi e$$

that belongs to $\mathbb{K}[x]$ and has π and e as roots. So $[\mathbb{Q}(\pi, e) : \mathbb{K}] \leq 2$. By Theorem 398,

$$[\mathbb{Q}(\pi, e) : \mathbb{Q}] = [\mathbb{Q}(\pi, e) : \mathbb{K}] \cdot [\mathbb{K} : \mathbb{Q}] \leq 2 \cdot k.$$

In particular, by Corollary 399, all elements of $\mathbb{Q}(\pi, e)$ should be algebraic over $\mathbb{Q}$. A contradiction, π is an element of $\mathbb{Q}(\pi, e)$ that is transcendental. □

Let us see an application of Theorems 398 and 394 to the problem of figuring out the degree of an element of $\mathbb{R}$ that is algebraic over $\mathbb{Q}$.

Lemma 402. *For any a, b positive integers, set $\mathbb{K}_a \stackrel{\text{def}}{=} \mathbb{Q}(\sqrt{a})$: then*

$$\mathbb{Q}(\sqrt{a}+\sqrt{b}) = \mathbb{Q}(\sqrt{a},\sqrt{b}) = \mathbb{K}_a(\sqrt{b}).$$

Proof. The first equality is proven as follows:

"$\subseteq$" Any field that contains both $\sqrt{a}$ and $\sqrt{b}$, must also contain $\sqrt{a}+\sqrt{b}$; so it contains $\mathbb{Q}(\sqrt{a}+\sqrt{b})$.

"$\supset$". Let $\mathbb{F}$ be a field that contains $x \stackrel{\text{def}}{=} \sqrt{a}+\sqrt{b}$. If $a=b$, then $\mathbb{F}$ contains also $\frac{x}{2} = \sqrt{a} = \sqrt{b}$ and we are done. If $a \neq b$, then $\mathbb{F}$ contains also

$$y \stackrel{\text{def}}{=} (\sqrt{a}+\sqrt{b})^3 = (a+3b)\sqrt{a} + (b+3a)\sqrt{b}.$$

In particular, $\mathbb{F}$ must also contain

$$y - (a+3b)x = (b+3a-a-3b)\sqrt{q} = 2(a-b)\sqrt{b}.$$

Since $a \neq b$, the element $2(a-b)$ is invertible in $\mathbb{Q}$; so $\mathbb{F}$ contains also $\sqrt{b}$. Then $\mathbb{F}$ contains also $x - \sqrt{b} = \sqrt{a}+\sqrt{b}-\sqrt{b} = \sqrt{a}$.

As for the second equality:

"$\subseteq$" $\mathbb{K}_a(\sqrt{b})$ is a field that contains $\mathbb{Q}$, $\sqrt{a}$, and $\sqrt{b}$.

"$\supset$" The polynomial $x^2 - b$ vanishes in $\sqrt{b}$. Hence, $\sqrt{b}$ is algebraic over $\mathbb{K}_a$ of degree ≤ 2. If it is algebraic of degree 1, then the conclusion follows immediately. If it is algebraic of degree 2, by Theorem 397 $\mathbb{K}_a(\sqrt{b}) = \mathbb{K}_a[\sqrt{b}]$, and the generic element of $\mathbb{K}_a(\sqrt{b})$ can then be written as

$$k + c\sqrt{b}, \quad \text{with } k \in \mathbb{K}.$$

But obviously $\mathbb{K}_a \stackrel{\text{def}}{=} \mathbb{Q}(\sqrt{a}) \subseteq \mathbb{Q}(\sqrt{a},\sqrt{b})$, so $k + c\sqrt{b} \in \mathbb{Q}(\sqrt{a},\sqrt{b})$. □

Proposition 403. *Let a, b be two positive integers. If none of $\sqrt{a}, \sqrt{b}, \sqrt{\frac{a}{b}}$ is in $\mathbb{Q}$, then*

$$[\mathbb{Q}(\sqrt{a}+\sqrt{b}) : \mathbb{Q}] = 4.$$

Proof. By Lemma 402 and Theorem 394,

$$[\mathbb{Q}(\sqrt{a}+\sqrt{b}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{a},\sqrt{b}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{a},\sqrt{b}) : \mathbb{Q}(\sqrt{a})] \cdot [\mathbb{Q}(\sqrt{a}) : \mathbb{Q}].$$

Now $\mathbb{Q}(\sqrt{a}) : \mathbb{Q}] = 2$, because the degree-2 polynomial $x^2 - a$ vanishes in $\sqrt{a}$ (which proves $\mathbb{Q}(\sqrt{a}) : \mathbb{Q}] \leq 2$) and $\sqrt{a} \notin \mathbb{Q}$ (which proves $\mathbb{Q}(\sqrt{a}) : \mathbb{Q}] > 1$, by Lemma 380).

So it remains to show that $[\mathbb{Q}(\sqrt{a},\sqrt{b}) : \mathbb{Q}(\sqrt{a})] = 2$. By Lemma 402, if $\mathbb{K}_a \stackrel{\text{def}}{=} \mathbb{Q}(\sqrt{a})$, then $\mathbb{Q}(\sqrt{a},\sqrt{b}) = \mathbb{K}_a(\sqrt{b})$. So we have to prove $[\mathbb{K}_a(\sqrt{b}) : \mathbb{K}_a] = 2$. Again, it is easy to see that $x^2 - b$ is a degree-2 polynomial with coefficients in $\mathbb{K}_a$ that vanishes on $\sqrt{b}$: hence, $[\mathbb{K}(\sqrt{b}) : \mathbb{K}] \leq 2$. So all we need to do is to exclude that $[\mathbb{K}(\sqrt{b}) : \mathbb{K}] = 1$. Using Lemma 380, if we show that that $\sqrt{b} \notin \mathbb{K}_a$ we are done. Also, $\sqrt{b}$ is algebraic over $\mathbb{K}_a$, so by Theorem 397 we can write

$$\mathbb{K}_a = \mathbb{Q}(\sqrt{a}) = \mathbb{Q}[\sqrt{a}] = \{c + d\sqrt{a} \text{ such that } c, d \in \mathbb{Q}\}.$$

Let us proceed by contradiction. Suppose there are c, d in $\mathbb{Q}$ such that $\sqrt{b} = c + d\sqrt{a}$. We distinguish three cases:

- if $c \neq 0$ and $d \neq 0$, then $b = c + ad^2 + 2cd\sqrt{a}$, whence we get $\sqrt{a} \in \mathbb{Q}$: A contradiction.
- if $c = 0$ and $d \neq 0$, then $\sqrt{\frac{a}{b}} = \frac{1}{d} \in \mathbb{Q}$, a contradiction.
- if $d = 0$, then $\sqrt{b} = c \in \mathbb{Q}$, a contradiction. □

Corollary 404. *Let a, b be two positive integers. If none of $\sqrt{a}, \sqrt{b}, \sqrt{\frac{a}{b}}$ is in $\mathbb{Q}$, then the polynomial*

$$f_{a,b} = x^4 - 2(a+b)x^2 + (a-b)^2$$

is irreducible. (In particular, if a, b are distinct primes then $f_{a,b}$ is irreducible.)

Proof. If $x = \sqrt{a} + \sqrt{b}$, then $x^2 = a + b + 2\sqrt{ab}$, so

$$x^2 - (a+b) = 2\sqrt{ab}.$$

Squaring again we get

$$x^4 + (a+b)^2 - 2(a+b)x^2 = 4ab,$$

so we see that

$$x^4 - 2(a+b)x^2 + (a-b)^2 = 0.$$

This means that $x = \sqrt{a} + \sqrt{b}$ is a root of $f_{a,b}$. So $f_{a,b}$ is a multiple of the minimum polynomial p of $\sqrt{a} + \sqrt{b}$. On the other hand, p has degree 4 by Theorem 398 and Proposition 403. Since $f_{a,b}$ is monic and has also degree 4, it follows that $f_{a,b} = p$. □

Example 405. $[\mathbb{Q}(\sqrt{2} + \sqrt{3}) : \mathbb{Q}] = 4$. Like in the proof of Corollary 404, one can verify that the minimum polynomial is

$$f_{2,3} = x^4 - 10x^2 + 1.$$

Non-Example 406. $[\mathbb{Q}(\sqrt{3} + \sqrt{12}) : \mathbb{Q}] = 2$, because $\sqrt{12} = 2\sqrt{3}$, so $\mathbb{Q}(\sqrt{3} + \sqrt{12}) = \mathbb{Q}(\sqrt{3})$. If you compute the polynomial $f_{3,12}$ of Corollary 404, you obtain

$$f_{3,12} = x^4 - 30x^2 + 81,$$

which vanishes in $x = \sqrt{3} + \sqrt{12}$, but is not irreducible. In fact,

$$x^4 - 30x^2 + 81 = (x^2 - 3) \cdot (x^2 - 27).$$

6.3 Algebraic closure and Nullstellensatz

Algebraically closed fields are basically fields $\mathbb{K}$ that contain the roots of all the polynomials with coefficients in $\mathbb{K}$. There is a natural example, namely, the field $\mathbb{C}$; but we are not allowed to say it out loud, because we are going to prove that $\mathbb{C}$ is algebraically closed only in the next section. So for the moment, think of this as an abstract theory; later we'll discover that all we manage to prove, is going to be true for $\mathbb{C}$.

Definition 407. A field $\mathbb{K}$ is called *algebraically closed* if every element algebraic over $\mathbb{K}$ has degree one (that is, it is in $\mathbb{K}$). Equivalently, $\mathbb{K}$ is algebraically closed if the following implication holds: For every field $\mathbb{F}$ containing $\mathbb{K}$, if every element of $\mathbb{F}$ is algebraic over $\mathbb{K}$, then $\mathbb{F} = \mathbb{K}$.

The most important question we wish to address is: How to "expand" a field that does not have this property, to a larger field that does? To this end, we introduce an operation, called "algebraic closure". A priori, given a field $\mathbb{K}$, it just outputs a larger set $\overline{\mathbb{K}}$; but we will prove later that this set $\overline{\mathbb{K}}$ is indeed an algebraically closed field.

Definition 408. Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. The *algebraic closure of* $\mathbb{K}$ is the set

$$\overline{\mathbb{K}} \stackrel{\text{def}}{=} \{a \in \mathbb{F} \text{ such that } a \text{ is algebraic over } \mathbb{K}.\}$$

Lemma 409. *With the notation above, $\overline{\mathbb{K}}$ is a field.*

Proof. Let α, β be elements of $\overline{\mathbb{K}}$. We want to show that

1. $\alpha - \beta \in \overline{\mathbb{K}}$;
2. $\alpha \cdot \beta \in \overline{\mathbb{K}}$;
3. unless $\alpha = 0$, then there exists $\alpha^{-1} \in \overline{\mathbb{K}}$.

For all three items, the trick is to consider $\mathbb{K}(\alpha, \beta)$. Since β is algebraic over $\mathbb{K}$, there is a polynomial in $\mathbb{K}[x]$ that vanishes in β; the same polynomial also belongs to $\mathbb{K}(\alpha)[x]$ and shows that β is algebraic over $\mathbb{K}(\alpha)$. So $[\mathbb{K}(\alpha, \beta) : \mathbb{K}(\alpha)]$ is finite. Also $[\mathbb{K}(\alpha) : \mathbb{K}]$ is finite: It is the degree of the minimum polynomial of α. But then by Theorem 394,

$$[\mathbb{K}(\alpha, \beta) : \mathbb{K}] = [\mathbb{K}(\alpha, \beta) : \mathbb{K}(\alpha)] \cdot [\mathbb{K}(\alpha) : \mathbb{K}]$$

is finite. Hence by Corollary 399 every element of $\mathbb{K}(\alpha, \beta)$ is algebraic over $\mathbb{K}$. In particular, $\alpha - \beta$, $\alpha \cdot \beta$ and α^{-1} are algebraic over $\mathbb{K}$. So they belong to $\overline{\mathbb{K}}$. □

Lemma 410. *With the notation above, $\overline{\mathbb{K}}$ is algebraically closed.*

Proof. Say $\mathbb{K} \subseteq \overline{\mathbb{K}} \subseteq \mathbb{F}$. Let $\alpha \in \mathbb{F}$ be an element algebraic over $\overline{\mathbb{K}}$. We want to show that α belongs to $\overline{\mathbb{K}}$; or in other words, that α is also algebraic over $\mathbb{K}$.

By the assumption, we know that there is a nonzero polynomial

$$h_0 + h_1 x + \ldots + h_n x^n, \text{ with } h_i \in \overline{\mathbb{K}}$$

that vanishes in α. The same polynomial proves that α is algebraic over $\mathbb{K}(h_0, h_1, \ldots, h_n)$, of degree $\leq n$. Moreover, each $h_i \in \overline{\mathbb{K}}$, so each h_i is algebraic over $\mathbb{K}$; hence, the field $\mathbb{K}(h_0, h_1, \ldots, h_n)$ has finite degree over $\mathbb{K}$. Therefore, by Theorem 394,

$$[\mathbb{K}(\alpha, h_0, h_1, \ldots, h_n) : \mathbb{K}] = [\mathbb{K}(\alpha, h_0, h_1, \ldots, h_n) : \mathbb{K}(h_0, h_1, \ldots, h_n)] \cdot [\mathbb{K}(h_0, h_1, \ldots, h_n) : \mathbb{K}].$$

So $[\mathbb{K}(\alpha, h_0, h_1, \ldots, h_n) : \mathbb{K}]$ is finite. So every element of it (and in particular α) is algebraic over $\mathbb{K}$. □

Proposition 411. *The algebraic closure of $\mathbb{K}$ is the smallest algebraically closed field that contains $\mathbb{K}$.*

Proof. By Lemmas 409 and 410, $\overline{\mathbb{K}}$ is an algebraically closed field. If $\mathbb{F}$ is any algebraically closed field that contains $\mathbb{K}$, $\mathbb{F}$ must contain any element algebraic over $\mathbb{K}$. So $\mathbb{F}$ contains $\overline{\mathbb{K}}$. □

Example 412. The algebraic closure of $\mathbb{Q}$ is a new field we have not encountered before. It contains $\mathbb{Q}$ but it is different than $\mathbb{R}$. For example, $\overline{\mathbb{Q}}$ contains i, which is not in $\mathbb{R}$; and $\mathbb{R}$ contains π, which is not algebraic over $\mathbb{Q}$. In Theorem 385 we have seen that $\overline{\mathbb{Q}} \cap \mathbb{R}$ is a measure zero subset of $\mathbb{R}$. In any case, clearly $\overline{\mathbb{Q}}$ is a subset of $\overline{\mathbb{R}}$. (We will show later on that $\overline{\mathbb{R}} = \mathbb{C}$.)

Proposition 413. *Both $[\overline{\mathbb{R}} : \overline{\mathbb{Q}}]$ and $[\overline{\mathbb{Q}} : \mathbb{Q}]$ are infinite.*

Proof. Suppose by contradiction that $[\overline{\mathbb{Q}} : \mathbb{Q}] = d$. By Example 354, elements like $\sqrt[d+1]{2}$ are algebraic over $\mathbb{Q}$ of degree $d + 1$, a contradiction with Corollary 399.

Suppose now that $[\overline{\mathbb{R}} : \overline{\mathbb{Q}}] = d$. Since $\pi \in \mathbb{R} \subseteq \overline{\mathbb{R}}$, this means that there exists a polynomial of degree d that that vanishes in π, of the form

$$h_0 + h_1 x + \ldots + h_n x^n, \text{ with } h_i \in \overline{\mathbb{Q}}.$$

Now we can use the same trick of the proof of Lemma 410. The polynomial above shows that π is algebraic over $\mathbb{Q}(h_0, h_1, \ldots, h_n)$, which has finite degree over $\mathbb{Q}$. So by Theorem 394,

$$[\mathbb{Q}(\pi, h_0, h_1, \ldots, h_n) : \mathbb{Q}] = [\mathbb{Q}(\pi, h_0, h_1, \ldots, h_n) : \mathbb{Q}(h_0, h_1, \ldots, h_n)] \cdot [\mathbb{Q}(h_0, h_1, \ldots, h_n) : \mathbb{Q}]$$

is finite. So by Corollary 399, π is algebraic over $\mathbb{Q}$, a contradiction. □

Remark 414. We will see in Corollary 499 that for any prime number p, the algebraic closure $\overline{\mathbb{Z}_p}$ is a field with infinitely many elements. In particular, no finite field is algebraically closed.

Theorem 415. *Let $\mathbb{K}$ be a field. The following are equivalent:*

1. *$\mathbb{K}$ is algebraically closed.*
2. *$\overline{\mathbb{K}} = \mathbb{K}$.*
3. *Any non-constant polynomial f of $\mathbb{K}[x]$ can be written as a product of degree-one polynomials in $\mathbb{K}[x]$.*
4. *Any non-constant polynomial f of $\mathbb{K}[x]$ has at least one root in $\mathbb{K}$.*
5. *The irreducible monic polynomials of $\mathbb{K}[x]$ are those of the form $x - a$, with $a \in \mathbb{K}$.*
6. *The maximal ideals of $\mathbb{K}[x]$ are all of the form $(x - a)$, with $a \in \mathbb{K}$.*

Proof.

(1) $\Rightarrow$ (4). (This is known as "Cauchy–Kronecker–Steinitz theorem"). Let g be any irreducible factor of f. Since $\mathbb{K}$ is a field, and irreducible elements cannot be invertible, necessarily $\deg g \geq 1$. Inside $\mathbb{K}[x]$ the ideal (g) is maximal (by Proposition 295), so by Theorem 264 the quotient $\mathbb{G} \stackrel{\text{def}}{=} \mathbb{K}[x]/(g)$ is a field. If $g = g_0 + g_1 x + \ldots + g_n x^n$, then inside the quotient $\mathbb{G}$ the element $\alpha \stackrel{\text{def}}{=} \overline{x}$ satisfies the equation

$$\overline{0} = \overline{g_0} + \overline{g_1}\alpha + \ldots + \overline{g_n}\alpha^n.$$

Now, g has positive degree, so if in $\mathbb{K}[x]$ we divide any degree-zero polynomial c by g, we obtain 0 as result and c itself as remainder. So the map from $\mathbb{K}$ to $\mathbb{G}$ that sends c to $\overline{c}$ is injective. In other words, we can identify $\mathbb{K}$ with a subset of $\mathbb{G}$. So $\mathbb{G}$ is an extension of $\mathbb{K}$. So if $g_i \in \mathbb{K}$, with slight abuse of notation we will simply write g_i instead of $\overline{g_i} \in \mathbb{K}[x]/(g)$. If $g = g_0 + g_1 x + \ldots + g_n x^n$, then by construction the element $\alpha \stackrel{\text{def}}{=} \overline{x}$ of $\mathbb{G}$ satisfies $0 = g_0 + g_1\alpha + \ldots + g_n\alpha^n$. So α is algebraic over $\mathbb{K}$. But then by the assumption α is already in $\mathbb{K}$! And since α is a root of g, and f is a multiple of g, it follows that α is a root of f.

(4) $\Rightarrow$ (3). If $n = 1$ the claim is obvious. Otherwise, since f has a root α_1, by Ruffini's theorem (Theorem 144) we can write $f = (x - \alpha_1)f_1$, for some polynomial $f_1 \in \mathbb{K}[x]$ of positive degree. If $\deg f_1 = 1$ we stop; otherwise, since f_1 has a root in $\mathbb{K}[x]$, we can write it as $f_1 = (x - \alpha_2)f_2$ for some polynomial $f_1 \in \mathbb{K}[x]$ of positive degree. And so on.

(3) $\Rightarrow$ (2). Obviously $\overline{\mathbb{K}} \supseteq \mathbb{K}$. For the other inclusion, if α is algebraic over $\mathbb{K}$, then it is a root of some monic polynomial $f \in \mathbb{K}[x]$. But by assumption f can be written as

$$f = (x - a_1)(x - a_2)\cdots(x - a_n), \text{ for some } a_i \in \mathbb{K}.$$

Plugging in $x = \alpha$, we obtain

$$0 = (\alpha - a_1)(\alpha - a_2)\cdots(\alpha - a_n),$$

and since $\mathbb{K}[x]$ is a domain, $\alpha = a_i$ for some i. Which implies $\alpha \in \mathbb{K}$.

(2) $\Rightarrow$ (1). By Lemma 410, $\overline{\mathbb{K}}$ is algebraically closed.

(4) ⇒ (6). Every ideal of the form $(x-a)$ is maximal by Theorem 264, since $\mathbb{K}[x]/(x-a) \cong \mathbb{K}$ (Proposition 243). Conversely, let J be a maximal ideal of $\mathbb{K}[x]$. Since $\mathbb{K}[x]$ is a PID (Theorem 217), then $J = (f)$ for some f. Since J is proper, f is not a constant, so by assumption it has a root α. By Ruffini's theorem (Theorem 144), f is a multiple of $(x-\alpha)$, which implies that $J \subseteq (x-\alpha)$. Being J maximal, it has to be $J = (x-\alpha)$.

(6) ⇔ (5). Since $\mathbb{K}[x]$ is a PID, by Proposition 295 the maximal ideals of $\mathbb{K}[x]$ are precisely the ideals generated by irreducible polynomials.

(5) ⇒ (4). By the assumption, any irreducible factor g of f is of the form $(x-a)$. Being a multiple of $x-a$, the polynomial f must vanish in a. □

It is natural to ask whether we can also characterize the irreducible polynomials with more than one variable, if $\mathbb{K}$ is algebraically closed. A first, negative answer is that Theorem 415 does not extend to $\mathbb{K}[x_1, \ldots, x_n]$ with $n > 1$. For example, for any $n \geq 3$, inside $\mathbb{K}[x_1, \ldots, x_n]$ the following degree-d polynomial (called "Fermat polynomial") is irreducible:

$$F_d \stackrel{\text{def}}{=} x_1^d + x_2^d + \ldots + x_n^d.$$

Such irreducible polynomial generates a prime ideal by Proposition 317 (because $\mathbb{K}[x_1, \ldots, x_n]$ is a UFD, cf. Gauss' theorem 336). However, there should be no surprise if this prime ideal is not maximal, because $\mathbb{K}[x_1, \ldots, x_n]$ is not a PID when $n > 1$. In fact,

$$(x_1^d + x_2^d + \ldots + x_n^d) \subsetneq (x_1, \ldots, x_n)$$

and the latter ideal is maximal.

So a legitimate question is to ask if, instead of the irreducible polynomials, we can characterize the maximal ideals, in a similar fashion to Theorem 415. The answer this time is positive and represents a cornerstone of classical algebraic geometry. We omit the proof for brevity.

Theorem 416 (Hilbert's Nullstellensatz)**.** *Let $\mathbb{K}$ be a field and let n be any positive integer. $\mathbb{K}$ is algebraically closed $\iff$ every maximal ideal of $\mathbb{K}[x_1, \ldots, x_n]$ is of the form*

$$(x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$$

with $a_1, \ldots, a_n \in \mathbb{K}$.

In the very special case when $n = 1$, the Nullstellensatz boils down to the case (1) ⇔ (6) of Theorem 415 above.

The derivative trick

It is easy to check whether a polynomial has all roots distinct, thanks to the "derivative trick". In the following subsection, we adopt the notation $f[i]$ to denote the coefficient of x^i in f.

Definition 417. Let A be a C-ring with 1. Let $f = a_0 + a_1 x + \ldots + a_n x^n \in A[x]$. The *derivative (polynomial) of f* by

$$D(f) \stackrel{\text{def}}{=} a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}.$$

In other words, $D(f)$ is defined by the identity

$$D(f)[i] = (i+1) \cdot f[i+1]. \tag{48}$$

Note that this is an entirely algebraic concept, we never mentioned limits or epsilons. Of course, there are some analogies with the derivatives you learned in calculus:

Lemma 418. *Let A be a C-ring. For all polynomials f, g in $A[x]$, and for all a in A, one has:*
- $D(f+g) = D(f) + D(g)$;
- $D(af) = a \cdot D(f)$;
- $D(f \cdot g) = D(f) \cdot g + f \cdot D(g)$.

Proof. The first two are left as exercise. We prove the third, which is more difficult. Since

$$D(f)g[i] = \sum_{j=0}^{i} D(f)[j]g[i-j] = \sum_{j=0}^{i}(j+1)f[j+1]g[i-j] = (i+1)f[i+1]g[0] + \sum_{j=0}^{i-1}(j+1)f[j+1]g[i-j],$$

using a reindexing trick (i.e. setting $h \stackrel{\text{def}}{=} j+1$, so that $j = h-1$) we get

$$D(f)g[i] = (i+1)f[i+1]g[0] + \sum_{h=1}^{i} hf[h]g[i+1-h]. \tag{49}$$

Similarly

$$fD(g)[i] = \sum_{j=0}^{i} f[j]D(g)[i-j] = \sum_{j=0}^{i} f[j](i-j+1)g[i-j+1]$$

and with a reindexing trick (this time with $h = j$), we obtain

$$fD(g)[i] = \sum_{h=0}^{i}(i+1-h)f[h]g[i+1-h] = (i+1)f[0]g[[i+1] + \sum_{h=1}^{i}(i+1-h)f[h]g[i+1-h]. \tag{50}$$

Now we can sum equations 49 and 50:

$$D(f)g[i] + fD(g)[i] = (i+1)f[i+1]g[0] + (i+1)f[0]g[[i+1] + \sum_{h=1}^{i}(\cancel{h} + i + 1 - \cancel{h})f[h]g[i+1-h],$$

from which it is easy to see that

$$D(f)g[i] + fD(g)[i] = \sum_{h=0}^{i+1}(i+1)f[h]g[i+1-h] = D(fg)[i]. \qquad \square$$

Now let α be an element of some C-ring B. Let A be any subring of B, not necessarily containing α, and let $f \in A[x]$. If there is a polynomial $g \in B[x]$ such that $f = (x-\alpha)^2 \cdot g$, then also $D(f)$ is a multiple of $(x-\alpha)$, because

$$D(f) = 2(x-\alpha)g + (x-\alpha)^2 D(g) = (x-\alpha)\left(2g + (x-\alpha)D(g)\right).$$

In case B is a UFD C-ring, we can even compute in $B[x]$ the $\gcd(f, D(f))$, and it will also be a multiple of $(x-\alpha)$. By contrapositive, this leads to the following fact:

Proposition 419. *Let $\mathbb{K}$ be a field. Let $f \in \mathbb{K}[x]$. If $\gcd(f, D(f)) = 1$ in $\mathbb{K}[x]$, then (in any field extension $\mathbb{F}$ of $\mathbb{K}$) any two roots of f are distinct.*

Proof. The greatest common divisor in $\mathbb{K}[x]$ can be found by means of the Euclidean algorithm. The same is true in $\mathbb{F}[x]$: and the calculations are exactly the same. Thus if $\gcd(f, D(f)) = 1$ in $\mathbb{K}[x]$, we also have $\gcd(f, D(f)) = 1$ in $\mathbb{K}[x]$. Had f a multiple root α in $\mathbb{F}$, then $f = (x-\alpha)^2 \cdot g$ for some $g \in \mathbb{F}[x]$, whence we would conclude that $1 = \gcd(f, D(f))$ is a multiple of $(x-\alpha)$, a contradiction. $\square$

6.4 The fundamental theorem of algebra

In this section, we are going to prove that $\mathbb{C}$ is algebraically closed, and so in particular it is the algebraic closure of $\mathbb{R}$. This follow from the fact that every non-constant polynomial in $\mathbb{C}[x]$ has at least one root. Despite several attempts by famous intellectuals (D'Alembert, 1746) and by some of the best mathematicians of all times (Euler, 1749; Lagrange, 1772; Laplace, 1795; and the 22-year old Gauss, 1799) the first entirely correct and complete proof of this fact was apparently given in 1813 by an amateur mathematician, Jean-Robert Argand, who worked as bookkeeper in Paris. Two years later Gauss provided a new proof, which is essentially the one we display here.

This theorem remains a beautiful example of interactions between mathematical fields. It is not difficult to derive the theorem in a Topology course or in a Complex Analysis course. For example, it follows easily from Liouville's theorem in Complex Analysis. Roughly speaking, Liouville's theorem says that all bounded complex-differentiable functions that can defined on the whole complex plane, are actually constant. Now given a monic polynomial p of degree n, if it is has no root, then the function $f(z) = \frac{1}{p(z)}$ would be defined on the whole plane. Being the reciprocal of a polynomial, it is not difficult to see that f is complex-differentiable and bounded. Hence, the function $f(z) = \frac{1}{p(z)}$ must be constant, which means that the polynomial p we started with had degree zero.

The original proof by Argand follows this strategy. The proof we sketch follows instead an approach outlined by Gauss in 1815, and later modernized by Emil Artin and Van der Waerden. For details, we recommend the article by J. Shipman, *Improving the Fundamental Theorem of Algebra*, The Mathematical Intelligencer 29 (2007), 9–14. Let us warm up with a few easy facts.

Lemma 420. *Every polynomial $p \in \mathbb{R}[x]$ of odd degree has a real root.*

Proof. Up to dividing p by its leading coefficient, we may assume that the leading term of p is x^n, for some n odd. Since

$$\lim_{x\to\infty} p(x) = \lim_{x\to\infty} x^n = +\infty \quad \text{and} \quad \lim_{x\to-\infty} p(x) = \lim_{x\to-\infty} x^n = -\infty,$$

by the intermediate value theorem of calculus it follows that there exists $x_0 \in \mathbb{R}$ such that $p(x_0) = 0$. □

Lemma 421. *$\mathbb{C}$ contains "all square roots": That is, For any z in $\mathbb{C}$, there exists $\delta \in \mathbb{C}$ such that $\delta^2 = z$.*

Proof. Let $z = a + ib$. If $b = 0$ then $z \in \mathbb{R}$ and the claim is obvious, so let us assume $b \neq 0$. We look for an element $\delta = x + iy$ such that

$$a + ib = (x + iy)(x + iy) = x^2 - y^2 + i(2xy).$$

In other words, given real numbers a and b, we need to solve over $\mathbb{R}$ the system

$$\begin{cases} x^2 - y^2 &= a \\ 2xy &= b. \end{cases}$$

Substituting $y = \frac{b}{2x}$ into the first equation, we obtain

$$x^2 - \frac{b^2}{4x^2} = a,$$

or in other words,

$$4x^4 - b^2 = 4ax^2.$$

Setting $t \stackrel{\text{def}}{=} x^2$ and imposing $t > 0$, this gives rise to a quadratic equation,

$$4t^2 - 4at - b^2 = 0,$$

which certainly has a positive solution, namely, $t = \frac{2a+\sqrt{4a^2+4b^2}}{4}$. (It is positive because $\sqrt{4a^2+4b^2} \geq \sqrt{4a^2} = |2a|$.) Once we found t, we immediately derive $x = \sqrt{t}$ and $y = \frac{b}{2x}$. □

Lemma 422. *Every monic polynomial $q \in \mathbb{C}[x]$ of degree two has a complex root.*

Proof. A generic degree-two monic polynomial of $\mathbb{C}[x]$ can be written as

$$q = x^2 + (2a + 2bi)x + (c + di), \text{ with } a, b, c, d \text{ in } \mathbb{R}.$$

By Lemma 421, we can find a complex number δ such that $\delta^2 = (a^2 - b^2 - c) + (2ab - d)i$. We claim that q factors as

$$q = (x + a + bi + \delta) \cdot (x + a + bi - \delta).$$

In fact,

$$\begin{aligned}(x + a + bi + \delta)(x + a + bi - \delta) &= (x + a + bi)^2 - \delta^2 = \\ &= (x^2 + a^2 - b^2 + 2ax + 2bix + 2abi) - (a^2 - b^2 - c + 2abi - di) = \\ &= x^2 + 2ax + 2bix + c + di = q.\end{aligned}$$

□

Lemma 423. *Let $\mathbb{K} \subseteq \mathbb{F}$ be fields. Let n be a positive integer. Let $p \in \mathbb{K}[x]$ be a degree-n polynomial such that in $\mathbb{F}[x]$ p splits as product of degree-one factors, i.e.*

$$p = (x - \phi_1)(x - \phi_2) \dots (x - \phi_n), \text{ for some } \phi_1, \dots, \phi_n \in \mathbb{F}.$$

For any symmetric *polynomial f in $\mathbb{F}[x_1, \dots, x_n]$, the evaluation $f(\phi_1, \dots, \phi_n)$ belongs to $\mathbb{K}$.*

Proof. Let us recall the theory of elementary symmetric polynomials, cf. Remark 166. Since

$$a_0 + a_1 x + \dots + a_{n-1}x^{n-1} + x^n = (x - \phi_1)(x - \phi_2) \dots (x - \phi_n),$$

we see that the product of the ϕ_i's is $(-1)^n a_0$; and more generally,

$$e_{n-k}(\phi_1, \dots \phi_n) = \pm a_k. \tag{51}$$

But both a_k and $-a_k$ are in $\mathbb{K}$, because $p \in \mathbb{K}[x]$. So Equation 51 tells us that if we evaluate the elementary symmetric polynomials at $x_j = \phi_j$ for all j, we obtain an element of $\mathbb{K}$. But by Theorem 171, every symmetric polynomial f in the variables $x_1, \dots, x_n$ can be written as a polynomial expression (with coefficients in $\mathbb{K}$!) in the elementary symmetric polynomials $e_j(x_1, \dots, x_n)$. So when we plug in ϕ_1 for x_1, ϕ_2 for x_2, ..., and ϕ_n for x_n, the resulting quantity $f(\phi_1, \dots, \phi_n)$ is a polynomial expression of elements of a field $\mathbb{K}$. So it is in $\mathbb{K}$. □

Example 424. Say $n = 3$, $p = x^3 + 2x + 3$, and $f = x^4 + y^4 + z^4$. Suppose that in some field $\mathbb{F} \supset \mathbb{Q}$ (for example $\mathbb{F} = \mathbb{Q}(\sqrt{-11})$ or $\mathbb{F} = \mathbb{C}$ would do) p splits as

$$x^3 + 2x + 3 = (x - \phi_1)(x - \phi_2)(x - \phi_3).$$

This is an equality of polynomials, so coefficientwise it means that

$$\begin{cases} 3 &= -\phi_1\phi_2\phi_3 \\ 2 &= \phi_1\phi_2 + \phi_1\phi_3 + \phi_1\phi_3 \\ 0 &= \phi_1 + \phi_2 + \phi_3 \\ 1 &= 1. \end{cases}$$

Now, we saw in Example 172 that the symmetric polynomial f can be written as

$$f = (e_1)^4 - 4(e_1)^2 e_2 + 2(e_2)^2 + e_1 e_3,$$

where $e_1 = x + y + z$, $e_2 = xy + xz + yz$, and $e_3 = xyz$. But then if we plug in $x = \phi_1$, $y = \phi_2$ and $z = \phi_3$, the system above tells us that $e_1(\phi_1, \phi_2, \phi_3) = 0$, $e_2(\phi_1, \phi_2, \phi_3) = 2$, and $e_3(\phi_1, \phi_2, \phi_3) = -3$; so we obtain

$$f(\phi_1, \phi_2, \phi_3) = 0^4 - 4 \cdot 0^2 \cdot 2 + 2 \cdot 2^2 + 0 \cdot (-3) = 8,$$

which is in $\mathbb{Q}$, as promised.

Theorem 425 (Gauss)**.** *Every non-constant polynomial $p \in \mathbb{R}[x]$ has a complex root.*

Proof. Let $n \stackrel{\text{def}}{=} \deg p$. Up to dividing p by its leading coefficient, we may assume that the leading term of p is x^n. So write

$$p = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + x^n.$$

We proceed by induction on *the largest natural number k such that 2^k divides n.* (I am not kidding, it is really a proof that was written by a genius.) If $k = 0$ it means that n is odd, so using Lemma 420 we can conclude immediately that p has a real root.

If $k > 0$, then $n = 2^k \cdot m$ with m odd. Let $\mathbb{F}$ be a field extending $\mathbb{R}$ (for example, the algebraic closure of $\mathbb{R}$) in which p splits as

$$p = (x - \phi_1)(x - \phi_2) \ldots (x - \phi_n), \text{ with } \phi_i \in \mathbb{F}.$$

Now, choose an arbitrary number $t \in \mathbb{R}$. Let us define a polynomial

$$q_t(x) \stackrel{\text{def}}{=} \prod_{1 \le i < j \le n} (x - \phi_i - \phi_j - t\phi_i\phi_j) \in \mathbb{F}[x].$$

This q_t is a polynomial of one variable, in $\mathbb{F}[x]$; but we make two surprising claims:
(1) q_t is actually in $\mathbb{R}[x]$;
(2) the degree of q_t is 2^{k-1} times an *odd* number.

Alright, so let us prove the claims:
(1) If we set

$$Q_t \stackrel{\text{def}}{=} \prod_{1 \le i < j \le n} (x - x_i - x_j - t x_i x_j) \in \mathbb{R}[x_1, \ldots x_n, x],$$

clearly q_t is obtained by evaluating Q_t at at $x_1 = \phi_1$, $x_2 = \phi_2$, ..., $x_n = \phi_n$. Note that Q_t is not symmetric, but if we set $B_n \stackrel{\text{def}}{=} \mathbb{R}[x_1, \ldots x_n]$ and view $\mathbb{R}[x_1, \ldots x_n, x]$ as $B_n[x]$, the *coefficients* of Q_t are symmetric polynomials in $\mathbb{R}[x_1, \ldots, x_n]$. But then by Lemma 423, the coefficients of $q_t(x)$ are all real numbers! So we discovered that $q_t \in \mathbb{R}[x]$.
(2) Recall that $n = 2^k m$ with m odd, and we are in the case $k > 0$, so n is even. Let

$$m' \stackrel{\text{def}}{=} m(n-1).$$

Clearly, m' is also odd. Why do we care? Because the degree of q_t is exactly

$$\deg q_t = \binom{n}{2} = \frac{n(n-1)}{2} = \frac{2^k m(n-1)}{2} = 2^{k-1} m'.$$

So miraculously, we can apply induction! We can conclude that the polynomial q_t has at least one complex root. In other words, there are two distinct integers i, j in $\{1, \ldots, n\}$ (depending on t) such that

$$\phi_i + \phi_j + t\phi_i\phi_j \text{ is a complex number.} \tag{52}$$

Everything we said so far holds for a generic real number t. In correspondence with any given t, we found a couple $\{i, j\}$ of elements in $\{1, \ldots, n\}$ that satisfies Equation 52. But since there are $\binom{n}{2}$ pairs from a size-n set, and since there are obviously more than $\binom{n}{2}$ real numbers, by the pidgeonhole principle we can find **two** distinct real numbers t and s such that, for the **same** pair (i, j),

$$\phi_i + \phi_j + t\phi_i\phi_j \text{ and } \phi_i + \phi_j + s\phi_i\phi_j \text{ are both in } \mathbb{C}.$$

So if we subtract these two quantities and divide the difference by $t-s$, the result, which happens to be $\phi_i\phi_j$, is still in $\mathbb{C}$. But then also $\phi_i + \phi_j$ is in $\mathbb{C}$, because

$$\phi_i + \phi_j = (\phi_i + \phi_j + t\phi_i\phi_j) - t(\phi_i\phi_j).$$

Time to conclude: ϕ_i and ϕ_j are roots of a monic degree-2 polynomial with coefficients in $\mathbb{C}$, namely,

$$x^2 - (\phi_i + \phi_j)x + (\phi_i\phi_j).$$

Hence, by Lemma 422, one of ϕ_i and ϕ_j must be in $\mathbb{C}$. So p has at least one root in $\mathbb{C}$. □

Theorem 426 (Fundamental Theorem of Algebra, Argand 1813)**.** $\mathbb{C}$ *is algebraically closed. Moreover, every non-constant polynomial* $f \in \mathbb{R}[x]$ *can be factored over* $\mathbb{R}[x]$ *into polynomials of degree one or two.*

Proof. Recursively, it suffices to show that any non-constant polynomial $f \in \mathbb{C}[x]$ has a root in $\mathbb{C}$. By Lemma 156, $f \cdot \overline{f}$ is a polynomial in $\mathbb{R}[x]$. By Theorem 425, $f \cdot \overline{f}$ has a complex root α. So by Ruffini's theorem (Theorem 144),

$$f \cdot \overline{f} = (x - \alpha) \cdot g \text{ for some } g \in \mathbb{R}[x].$$

But $\mathbb{R}[x]$ is a UFD, and $(x - \alpha)$ is irreducible; so $(x - \alpha)$ must be a factor either of f, or of $\overline{f}$; in the latter case, $(x - \overline{\alpha})$ is a factor of f by Proposition 157. So either way, f has a root in $\mathbb{C}$.

As for the second part of the claim: Suppose now that $f \in \mathbb{R}[x]$, $f \neq 0$. Up to dividing f by its nonzero leading coefficient, we may assume that f is monic. By the first part of the claim, which we have just proved, f will have n complex roots, $r_1, \ldots, r_n$. If these roots are all real, then f factors in $\mathbb{R}[x]$ as

$$f = (x - r_1)(x - r_2)\cdots(x - r_n)$$

and we are done. Otherwise, we know by Proposition 157 that the non-real roots of f come in pairs: Whenever α is a root, so is $\overline{\alpha}$. So, suppose that there are r real roots and $2s$ complex non-real roots, so that

$$f = (x - r_1)\cdots(x - r_k) \cdot (x - \alpha_1) \cdot (x - \overline{\alpha_1})\cdots(x - \alpha_s) \cdot (x - \overline{\alpha_s}).$$

The product of the two binomials $(x-\alpha)(x-\overline{\alpha})$ is in $\mathbb{R}[x]$, by Lemma 156 (or by straightforward computation). So if we group together the binomials corresponding to conjugate roots, we get that f splits as product of r degree-one polynomials and s degree-two polynomials. □

Corollary 427. $\mathbb{C}$ *is the algebraic closure of* $\mathbb{R}$.

Proof. By Theorem 426, $\mathbb{C}$ is algebraically closed. But any element of $\mathbb{C}$ is of the form $x = a+ib$, so it is algebraic over $\mathbb{R}$ because it is a root of the polynomial

$$(x - a - ib)(x - a + ib) = x^2 - 2ax + (a^2 + b^2).$$

□

Corollary 428. *The irreducible polynomials in $\mathbb{C}[x]$ are the degree-one polynomials. The irreducible polynomials in $\mathbb{R}[x]$ are the degree-one polynomials, plus the degree-two polynomials $x^2 + bx + c$ with $\Delta = b^2 - 4c < 0$.*

Proof. In $\mathbb{R}[x]$, polynomials of degree 2 without roots are irreducible. Theorem 426 tells us that no polynomial of higher degree is irreducible. □

Remark 429. Notice the stark difference with $\mathbb{Q}[x]$, where there are irreducible polynomials of arbitrarily high degree, cf. Example 354.

Corollary 430. *In $\mathbb{C}[x]$, every maximal ideal is of the form $(x - a)$, with $a \in \mathbb{C}$. In $\mathbb{R}[x]$, the maximal ideals are the ideals of the form $(x - a)$, with $a \in \mathbb{R}$, plus the ideals of the form $(x^2 + bx + c)$, with $\Delta = b^2 - 4c < 0$.*

6.5 *$\mathbb{Q}(\sqrt{p})$ and Fibonacci numbers

How can you tell if a given number is Fibonacci? Do you need to reconstruct them all? In this section, we learn a new way to describe Fibonacci numbers. It involves studying the field $\mathbb{Q}(\sqrt{p})$, where p is a prime number of the form $4k + 1$. This subsection is mostly copied from the article *The quadratic field $\mathbb{Q}(\sqrt{5})$ and a certain Diophantine equation*, by D. A. Lind.

Definition 431. An element algebraic over $\mathbb{Q}$ is called an *algebraic integer* if the coefficients of its minimum polynomial are all in $\mathbb{Z}$. When both α and α^{-1} are algebraic integers, α is called a *unit*.

Lemma 432. *The algebraic integers of degree one are precisely the integers.*

Proof. They are the roots of monic polynomials $x - z$, with $z \in \mathbb{Z}$. □

Lemma 433. *Let p be a prime number of the form $4k+1$. The algebraic integers in $\mathbb{Q}[\sqrt{p}]$ are precisely*

$$\left\{ \frac{a + b\sqrt{p}}{2} \text{ such that } a, b \in \mathbb{Z}, \text{ and } a \equiv b \mod 2 \right\}.$$

Proof. "$\subseteq$" Since $\mathbb{Q}(\sqrt{p}) = \mathbb{Q}[\sqrt{p}]$, any element of $\mathbb{Q}(\sqrt{p})$ can be written as

$$u = \frac{a + b\sqrt{p}}{c},$$

with a, b, c integers such that $\gcd(a, b, c) = 1$ and $c \in \mathbb{N}$. When is one such u an algebraic integer? We distinguish two cases:

- If $b = 0$, then $u = \frac{a}{c}$ is in $\mathbb{Q}$. By Lemma 432, if u is an algebraic integer, then $u \in \mathbb{Z}$ and a is a multiple of c. So without loss we can assume $c = 2$ and $a = 2u$. Then indeed u is of the form
$$u = \frac{a + 0\sqrt{p}}{2}$$
with a even.

- If $b \neq 0$, then $u = \frac{a+b\sqrt{p}}{c}$ is algebraic over $\mathbb{Q}$ of degree 2, and its minimum polynomial is

$$f = x^2 - \frac{2a}{c}x + \frac{a^2 - pb^2}{c^2}.$$

  If $c = 1$, then setting $a' = 2a$ and $b' = 2b$ we can certainly write u in the form $u = \frac{a'+b'\sqrt{p}}{2}$, with $a' - b'$ even. So let us focus on the case $c \geq 2$. Clearly $f \in \mathbb{Z}[x]$ if and only if both $\frac{2a}{c}$ and $\frac{a^2-pb^2}{c^2}$ are integers. But then $\frac{4a^2}{c^2}$ and $\frac{4a^2-4pb^2}{c^2}$ are also integers, and so is their difference $\frac{4pb^2}{c^2}$. So for any prime factor q of c, we know that q^2 divides $4a^2$ and that q^2 divides $4pb^2$. So q must be 2: The other option (q divides both a^2 and b^2) would contradict $\gcd(a, b, c) = 1$. So c is a power of 2. Were $c = q^r = 2^r$ with $r \geq 2$, then c^2 would be a multiple of 16, and we would get another contradiction: In order for $\frac{4a^2}{c^2}$ and $\frac{4pb^2}{c^2}$ to be integers, both a and b would need to be even to cancel out the factor 16 at the denominator, and this again would contradict $\gcd(a, b, c) = 1$. (Here we used that p is prime – or better, that it is a squarefree integer, i.e. a product of distinct primes.) So c is a power of 2, but the exponent can only be 1: in other words, we found out that $c = 2$. Moreover, we point out that a, b are both odd. (Were they both even, then $\gcd(a, b, c)$ would not be 1; and were only one of a, b even, then $\frac{a^2-pb^2}{4}$ would not be an integer.) In particular, $a \equiv b \mod 2$.

"$\supseteq$" The case $b = 0$, using Lemma 432, is easy and left as exercise. So, suppose $u = \frac{a+b\sqrt{p}}{2}$ with $b \neq 0$. To show that its minimum polynomial f is in $\mathbb{Z}[x]$, all we need to verify is that $\frac{a^2-pb^2}{4}$ is an integer. By the assumption, in $\mathbb{Z}_4$ we have $\overline{p} = \overline{1}$. Moreover, since a and b are congruent mod 2, their squares are congruent mod 4 (because in $\mathbb{Z}_4$ we have $\overline{(n+2)}^2 = \overline{n^2 + 4n + 4} = \overline{n^2}$ for all $n \in \mathbb{N}$). So in $\mathbb{Z}_4$ one has

$$\overline{a^2 - pb^2} = \overline{a}^2 - \overline{p} \cdot \overline{b}^2 = \overline{a}^2 - \overline{1} \cdot \overline{a}^2 = \overline{0}. \qquad \square$$

Because of Lemma 433, we can define a "norm function" N on the algebraic integers in $\mathbb{Q}[\sqrt{p}]$, as follows:

$$\begin{array}{rccc} N: & \{ \text{ algebraic integers of } \mathbb{Q}[\sqrt{p}]\} & \longrightarrow & \mathbb{Z} \\ & \frac{a+b\sqrt{p}}{2} & \longmapsto & \frac{a^2-pb^2}{4}. \end{array}$$

Note that the norm can be negative: For example, for $p = 5$, the element $\alpha = \frac{1+\sqrt{5}}{2}$, which will play a crucial role below, has norm -1.

Lemma 434. *The function N above satisfies $N(xy) = N(x)N(y)$ for all x, y; in particular, the units are precisely the algebraic integers with norm ± 1.*

Proof. If $x = \frac{a+b\sqrt{p}}{2}$ and $y = \frac{c+d\sqrt{p}}{2}$, with $a \equiv b$ and $c \equiv d \mod 2$, then xy can be written as

$$xy = \frac{(ac + pbd) + (ad + bc)\sqrt{p}}{4} = \frac{1}{2}\left(\frac{ac + pbd}{2} + \frac{ad + bc}{2}\sqrt{p}\right).$$

Note that in $\mathbb{Z}_2$ $\overline{a} = \overline{b}$ and $\overline{c} = \overline{d}$, so

- $\overline{ad + bc} = \overline{ad} + \overline{ad} = \overline{0}$.
- $\overline{ac + pbd} = \overline{ac + pac} = \overline{(p+1)ac} = \overline{0}$, because $p + 1$ is even.

So if we set $a' \stackrel{\text{def}}{=} \frac{ac+pbd}{2}$ and $b' \stackrel{\text{def}}{=} \frac{ad+bc}{2}$, we have that a', b' are both integers! We claim that $a' \equiv b' \bmod 2$. To see this, it suffices to show that $2a' + 2b'$ is always a multiple of 4. Indeed, write $b = a + 2t$ and $d = c + 2z$, for $t, z \in \mathbb{Z}$. In $\mathbb{Z}_4$, since $\overline{p} = \overline{1}$, we have

$$\begin{aligned}\overline{2a' + 2b'} &= \overline{ac + pbd + ad + bc} = \overline{ac + (a+2t)(c+2z) + a(c+2z) + (a+2t)c} = \\ &= \overline{ac + ac + 2ct + 2az + 4tz + ac + 2az + ac + 2ct} = \overline{0}.\end{aligned}$$

So $a' \equiv b' \bmod 2$ and xy is again an algebraic integer, by Lemma 433. Thus it makes sense to compute

$$\begin{aligned}N(xy) &= N(\tfrac{a'+b'\sqrt{p}}{2}) = \tfrac{(a')^2 - p(b')^2}{4} = \tfrac{(ac+pbd)^2 - p(ad+bc)^2}{16} = \\ &= \tfrac{a^2c^2 + p^2b^2d^2 + \cancel{2pabcd} - pa^2d^2 - pb^2c^2 - \cancel{2pabcd}}{16} = \\ &= \tfrac{a^2c^2 - pa^2d^2 - pb^2c^2 + p^2b^2d^2}{16} = \tfrac{a^2 - 5b^2}{4}\,\tfrac{c^2 - pd^2}{4} = N(x)N(y).\end{aligned}$$

But then of course if $xy = 1$, it must be $N(x)N(y) = 1$, and since both are integers, this means that $N(x)$ is either 1 or -1. Conversely, suppose $N(x) = \pm 1$. Then if $x = \frac{a+b\sqrt{p}}{2}$, set $y = \frac{a-b\sqrt{p}}{2}$: we have

$$xy = \frac{a + b\sqrt{p}}{2} \cdot \frac{a - b\sqrt{p}}{2} = \frac{a^2 - pb^2}{4} = N(x) = \pm 1$$

by assumption, so the inverse of x is either y or $-y$. In both cases, x is a unit. □

Let us focus on the case $a = 1$, $b = 1$. The element $\frac{1+\sqrt{p}}{2}$ has norm $\frac{1-p}{4}$, so in general it is not a unit; but it is a unit for $p = 5$.

Theorem 435. *Let $\alpha \stackrel{\text{def}}{=} \frac{1+\sqrt{5}}{2}$. The units in $\mathbb{Q}[\sqrt{5}]$ are precisely*

$$\{\pm\alpha^z \text{ such that } z \in \mathbb{Z}\}.$$

Proof. The fact that $N(\alpha) = -1$ together with Lemma 434 immediately implies that any element of the form α^z is a unit (because it has norm ± 1). So the hard part is to show the opposite implication. We first claim that there is no unit in the real interval $(1, \alpha)$. We prove the claim by contradiction. Let $u = \frac{a+b\sqrt{5}}{2}$ be one unit between 1 and α. Since $N(u) = \pm 1$, we obtain

$$-u < -1 \leq N(u) \leq 1 < u.$$

Dividing by u, which is positive, and adding one to the result, we obtain

$$-1 < \frac{N(u)}{u} < 1.$$

Add this inequality "componentwise" to the inequality $1 < u < \alpha$: We obtain

$$0 < \frac{N(u)}{u} + u < 1 + \alpha.$$

But $\frac{N(u)}{u} + u = \frac{a-b\sqrt{5}}{2} + \frac{a+b\sqrt{5}}{2} = a$! So the equation above tells us that the integer a is either 1 or 2. In both cases, however, we have a contradiction, because there is no integer b such that

$$1 < \frac{1 + b\sqrt{5}}{2} < \alpha \quad \text{or } 1 < \frac{2 + b\sqrt{5}}{2} < \alpha.$$

So the claim is proven. Now suppose that u is an arbitrary unit in $\mathbb{Q}[\sqrt{5}]$. Up to replacing it with $-u$, which is also a unit by Lemma 434, we can assume $u > 0$. If u is of the form α^n, there

is nothing to show. If not, since the function $f(n) = \alpha^n$ tends to infinity, clearly we can find an integer n such that

$$\alpha^n < u < \alpha^{n+1}.$$

But since α^{-n} is also a unit, and the product of two units is a unit by Lemma 434, we get that $v \stackrel{\text{def}}{=} \alpha^{-n}u$ is also a unit; and in addition, it satisfies

$$1 < v < \alpha.$$

A contradiction with the claim we proved above. □

Definition 436 (Fibonacci, Lucas)**.** The Fibonacci numbers are defined recursively by

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

whereas the Lucas numbers are defined recursively by

$$L_{n+2} = L_{n+1} + L_n, \qquad L_0 = 2, L_1 = 1.$$

Lemma 437. *Let $\alpha \stackrel{\text{def}}{=} \frac{1+\sqrt{5}}{2}$. For all $n \in \mathbb{N}$, one has*

$$\alpha^n = \frac{L_n + F_n\sqrt{5}}{2}.$$

Proof. By induction on n. When $n = 0$ the equation above is $1 = 1$. When $n = 1$, the equation above is $\alpha = \frac{1+1\sqrt{5}}{2}$, also true. Before moving to the induction step, note that

$$\alpha^2 = \frac{1+5+2\sqrt{5}}{4} = \frac{2+2\sqrt{5}+4}{4} = \frac{2+2\sqrt{5}}{4} + 1 = \alpha^1 + \alpha^0,$$

whence multiplying by α^n we get

$$\alpha^{n+2} = \alpha^{n+1} + \alpha^n.$$

This shows us how to prove the induction step cleverly:

$$\begin{aligned}\alpha^{n+2} &= \alpha^{n+1} + \alpha^n = \tfrac{L_{n+1}+F_{n+1}\sqrt{5}}{2} + \tfrac{L_n+F_n\sqrt{5}}{2} = \tfrac{(L_{n+1}+L_n)+(F_{n+1}+F_n)\sqrt{5}}{2} = \\ &= \tfrac{L_{n+2}+F_{n+2}\sqrt{5}}{2}.\end{aligned}$$

□

Theorem 438. *Either F_n and L_n are both even, or they are both odd. Moreover, the units in $\mathbb{Q}[\sqrt{5}]$ are precisely*

$$\left\{ \pm\frac{L_n + F_n\sqrt{5}}{2} \text{ such that } n \in \mathbb{N} \right\}$$

and their inverses.

Proof. The second claim is obtained by putting together Lemma 437 and Theorem 435; all we did was to rewrite the set "$\{\pm\alpha^z$ such that $z \in \mathbb{Z}\}$" as "$\{\pm\alpha^n$ such that $n \in \mathbb{N}\}$ and their inverses". The first claim is implied by the second, because if $\frac{L_n+F_n\sqrt{5}}{2}$ is a unit, then by Lemma 434 one has $L_n \equiv F_n \bmod 2$. □

Theorem 439 (Fibonacci numbers as rational solutions)**.** *All rational solutions of the equations*

$$x^2 - 5y^2 = 4$$

are given by $x = L_{2n}$, $y = F_{2n}$. Similarly, all rational solutions of the equations

$$x^2 - 5y^2 = -4$$

are given by $x = L_{2n+1}$, $y = F_{2n+1}$.

Proof. By theorem 438, $\frac{L_k+F_k\sqrt{5}}{2}$ is a unit, so by Lemma 434 its norm is ± 1. But we can be more precise: using Lemma 437 and the fact that $N(\alpha) = \frac{1^2-5\cdot 1^2}{4} = -1$, we obtain

$$\frac{(L_{2n})^2 - 5(F_{2n})^2}{4} = N(\frac{L_{2n} + F_{2n}\sqrt{5}}{2}) = N(\alpha^{2n}) = [N(\alpha)]^{2n} = [-1]^{2n} = 1, \text{ and}$$

$$\frac{(L_{2n+1})^2 - 5(F_{2n+1})^2}{4} = N(\frac{L_{2n+1} + F_{2n+1}\sqrt{5}}{2}) = N(\alpha^{2n+1}) = [-1]^{2n+1} = -1.$$

Conversely,

- if $x^2 - 5y^2 = 4$, then $x \equiv y \bmod 2$ and

  $$N(\frac{x + y\sqrt{5}}{2}) = \frac{x^2 - 5y^2}{4} = \frac{4}{4} = 1.$$

  So $(\frac{x+y\sqrt{5}}{2})$ is a unit by Lemma 434. So by Theorem 438, $\frac{x+y\sqrt{5}}{2} = \alpha^k$ for some k. Since $N(\frac{x+y\sqrt{5}}{2}) = 1$, the integer k must be even, so $k = 2n$ for some n. But then by Lemma 437 we have

  $$x + y\sqrt{5} = 2\alpha^{2n} = L_{2n} + F_{2n}\sqrt{5}.$$

  Since $[\mathbb{Q}(\sqrt{5} : \mathbb{Q}] = 2$, the elements 1 and $\sqrt{5}$ are a basis for $\mathbb{Q}(\sqrt{5})$ over $\mathbb{Q}$, which means that every element of $\mathbb{Q}(\sqrt{5})$ can be *uniquely* represented as $a + b\sqrt{5}$, for a, b in $\mathbb{Q}$. So it must be $x = L_{2n}$ and $y = F_{2n}$.
- Similarly, if $x^2 - 5y^2 = -4$, then $x \equiv y \bmod 2$ and

  $$N(\frac{x + y\sqrt{5}}{2}) = \frac{x^2 - 5y^2}{4} = \frac{-4}{4} = -1.$$

  So by Theorem 438, we have $\frac{x+y\sqrt{5}}{2} = \alpha^k$ for some k odd. If $k = 2n + 1$, then by Lemma 437 we have

  $$x + y\sqrt{5} = 2\alpha^{2n+1} = L_{2n+1} + F_{2n+1}\sqrt{5},$$

  so as above we conclude $x = L_{2n+1}$ and $y = F_{2n+1}$. □

Corollary 440 (Testing Fibonacci)**.** *A number $m \in \mathbb{N}$ is Fibonacci $\iff$ either $5m^2 + 4$ or $5m^2 - 4$ is a perfect square.*

Proof. By Theorem 439, $5m^2 + 4$ is a perfect square if and only if m is a Fibonacci number of even index, while $5m^2 - 4$ is a perfect square if and only if m is a Fibonacci number of odd index. □

6.6 Exercises

1. Let p be a prime number. Is there a field $\mathbb{K}$ such that

   $$\mathbb{Q} \subsetneq \mathbb{K} \subsetneq \mathbb{Q}[\sqrt[p]{2}]\,?$$

2. Prove that $\mathbb{Q}(\sqrt{3} + i) = \mathbb{Q}(\sqrt{3}, i)$ and find the minimum polynomial of $\sqrt{3} + i$ over $\mathbb{Q}$.
3. Find the minimum polynomial of $\sqrt[3]{7} - 1$ over $\mathbb{Q}$.
4. Find the minimum polynomial of $\sqrt[3]{3} - \sqrt[3]{9}$ over $\mathbb{Q}$.

5. We know (by von Lindemann's theorem) that π is transcendental over $\mathbb{Q}$. Could it be that π^3 is algebraic over $\mathbb{Q}$?

6. Prove that $[\overline{\mathbb{Q}}(\pi) : \overline{\mathbb{Q}}]$ is infinite.

7. Prove that $x^3 + x - 1$ is irreducible in $\mathbb{Z}_5[x]$. If u is a root, how many elements does $\mathbb{Z}_5(u)$ have?

8. Prove that there are $\frac{p(p-1)}{2}$ monic irreducible quadratic polynomials in $\mathbb{Z}_p[x]$.

9. (hard) Prove that there are $\frac{p(p-1)}{3}$ monic irreducible cubic polynomials in $\mathbb{Z}_p[x]$.

7 Groups

7.1 Groups, subgroups, homomorphisms, and Lagrange

Definition 441. A **group** consists of a set G endowed with an operation $(x, y) \mapsto xy$ that satisfies the following axioms:

(G0) The operation is *internal*. That is, for all x, y in G, the element xy is in A.

(G1) The operation is *associative*. That is, for all x, y, z in G, $x(yz) = (xy)z$. In view of this, we usually leave out brackets and simply write xyz.

(G2) The operation has a (necessarily unique[16]) *neutral element*. That is, there is an element e in G such that for all x in G, $xe = x = ex$.

(G3) Every element has a (necessarily unique[17]) *inverse*. That is, for all x in G there exists an element y in G such that $xy = e = yx$. We denote such y by x^{-1}.

Remark 442. Instead of ab^{-1} one might be tempted to write $\frac{a}{b}$. Don't do it! Our operation might not be commutative. So if you write $\frac{a}{b}$, it's not clear whether you mean ab^{-1} or $b^{-1}a$.

Example 443 (Bijections and Permutations)**.** Given an arbitrary set A, the set

$$G = \{f : A \longrightarrow A \text{ bijective}\}$$

is a group with respect to composition; the neutral element is the identity function. A case of particular interest is when A is finite; in this case, the bijections are called *permutations*. We'll see them in more details soon.

Example 444 (Fields)**.** $\mathbb{R}$ is a group with respect to addition, with 0 as neutral element. Moreover, $\mathbb{R}^* \stackrel{\text{def}}{=} \mathbb{R} \setminus \{0\}$ is a group with respect to multiplication!, with neutral element 1. More generally, any C-ring A is a group with respect to the operation of addition. The neutral element e is what we called "0". If A is a field, and we set $A^* \stackrel{\text{def}}{=} A \setminus \{0\}$, this is a group with respect to the operation of multiplication (and with neutral element what we called "1"). If A is not a field, then * is not a group, since some elements do not have multiplicative inverse.

Example 445. For any positive integer m, $\mathbb{Z}_m$ is a group with respect to addition modulo m. Of course, $\mathbb{Z}_m^* \stackrel{\text{def}}{=} \mathbb{Z}_m \setminus \{0\}$ is a group with respect to multiplication (modulo m) if and only if m is prime. If m is *not* prime, however, it is still true that *the integers in* $\{1, \ldots, m-1\}$ *coprime with* m form a group with respect to multiplication modulo m. This is usually called *multiplicative group of integers modulo* m, denoted by $(U_m, \cdot)$, and well-studied in the literature.

Example 446. Let G, H be groups. Then the cartesian product

$$G \times H \stackrel{\text{def}}{=} \{(g, h) \text{ such that } g \in G, h \in H\}$$

is a group with respect to the "entrywise" operation

$$(g_1, h_1)(g_2, h_2) \stackrel{\text{def}}{=} (g_1 g_2, h_1 h_2).$$

In fact, the neutral element is just the pair (e_G, e_H) of the respective neutral elements; and the inverse of the pair (g, h) is simply the pair (g^{-1}, h^{-1}).

[16] Were there two neutral elements e_1 and e_2, we would have $e_1 e_2 = e_1$ (because e_2 is neutral) yet also $e_1 e_2 = e_2$ (because e_1 is neutral), so $e_1 = e_2$.

[17] Were there two inverses y_1 and y_2 for the same element x, we would have $y_1 x y_2 = e y_2 = y_2$ (because y_1 is inverse) yet also $y_1 x y_2 = y_1 e = y_1$ (because y_2 is inverse), so $y_1 = y_2$.

Example 447 (Quaternions)**.** Consider on $\mathbb{R}^4$ minus the origin the following operation, introduced by Hamilton in 1843:

$$\begin{aligned}&(a,b,c,d)\otimes(a',b',c',d')\stackrel{\text{def}}{=}\\&\quad\stackrel{\text{def}}{=}(aa'-bb'-cc'-dd',\ \ ab'+a'b+cd'-c'd,\ \ ac'+a'c-bd'+b'd,\ \ ad'+a'd+bc'-b'c).\end{aligned}$$

With respect to these operations, $\mathbb{R}^4\setminus\{(0,0,0,0)\}$ becomes a non-Abelian group, called the *quaternions*. It is easy to see that the neutral element is $(1,0,0,0)$ and the inverse of an element (a,b,c,d) is

$$(a,b,c,d)^{-1}=\frac{1}{a^2+b^2+c^2+d^2}\,(a,-b,-c,-d).$$

Proposition 448 (Cancellation)**.** *Let G be a group. For any $a,b,c\in G$, if $ab=ac$ then $b=c$.*

Proof. Multiply by a^{-1}. □

Proposition 449 (Inverse of product)**.** *Let G be a group. For any $a,b\in G$, one has*

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

Proof. Since the inverse is unique, we only need to check that $(ab)(b^{-1}a^{-1})=e=(b^{-1}a^{-1})(ab)$. This follows from associativity: for example,

$$(ab)(b^{-1}a^{-1})=(a(bb^{-1}))a^{-1}=(ae)a^{-1}=aa^{-1}=e.$$ □

Definition 450. Let G be a group. A *subgroup* of G is a subset $H\subseteq G$ that is a group with respect to the same operation.

Proposition 451. *Let G be a group. $H\subseteq G$ is a subgroup $\iff$ H satisfies*

(SG1) for each a,b in H, the element ab^{-1} is in H.

Proof.
"$\Rightarrow$" This is easy: if $a,b\in H$ group, then b^{-1} is in H, so ab^{-1} is in H.
"$\Leftarrow$" Applying (SG1) to $b=a$ we get that the neutral element $e=aa^{-1}$ is in H. But then for each b in H we get that eb^{-1} is in H, again by (SG1). So H contains the inverse of any of its elements. Finally, we should check that the operation is internal. Let x,y be arbitrary elements of H. We have just proven that $y^{-1}\in H$. Applying (SG1) to $a=x$ and $b=y^{-1}$, we get that $x((y^{-1})^{-1})$ is in H. In other words, $xy\in H$. □

Example 452. If $(B,+,\cdot)$ is a subring of $(A,+,\cdot)$, then $(B,+)$ is also a subgroup of $(A,+)$. For example, $(\mathbb{Z},+)$ is a subgroup of $(\mathbb{Q},+)$, which in turn is a subgroup of $(\mathbb{Q}[\sqrt{2}],+)$, which is a subgroup of $(\mathbb{R},+)$, which is a subgroup of $(\mathbb{C},+)$.

Example 453. If $\mathbb{K}\subseteq\mathbb{F}$ is a field extension, then $(\mathbb{K}^*,\cdot)$ is a subgroup of $(\mathbb{F}^*,\cdot)$ (the star here denotes that the zero element has been removed from the set). For example, $(\mathbb{Q}^*,\cdot)$ is a subgroup of $(\mathbb{R}^*,\cdot)$, which in turn is a subgroup of $(\mathbb{C}^*,\cdot)$.

Example 454. If $(H_j)_{j\in I}$ is a family of subgroups of a group G, then their intersection is also a subgroup of G. This can easily be checked via Proposition 451, as follows: Let $a,b\in\bigcap_{j\in I}I_j$. Then for each j, both a,b belong to the subgroup I_j. Hence ab^{-1} is in I_j. Since this holds for all j, ab^{-1} is in $\bigcap_{j\in I}I_j$.

Group homomorphisms

Definition 455. A function $f : G \longrightarrow H$ between two groups is called *group homomorphism* if

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

A bijective homomorphism is called *isomorphism*. Two groups G, H are called *isomorphic* if there exist an isomorphism between them.

Proposition 456. *Let $f : G \longrightarrow H$ be a group homomorphism. Then $f(e_G) = e_H$. Moreover, for any a in G, the inverse of $f(a)$ is $f(a^{-1})$.*

Proof. By the cancellation property (Proposition 448), from

$$f(e_G)f(e_G) = f(e_G) = e_H f(e_G)$$

one gets $f(e_G) = e_H$. But then similarly from

$$f(a)f(a^{-1}) = f(e_G) = e_H = f(a)[f(a)]^{-1}$$

one gets $f(a^{-1}) = [f(a)]^{-1}$. □

Proposition 457. *For any $f : G \longrightarrow H$ group homomorphism,* $\operatorname{Im} f$ *is a subgroup of H and*

$$\ker f \stackrel{\text{def}}{=} \{x \in G \text{ such that } f(x) = e_H\}$$

is a subgroup of G. Moreover, f is injective if and only if $\ker f = \{e_G\}$.

Proof. Left as exercise. □

Example 458. Any C-ring homomorphism is in particular a group homomorphism, if we focus on the operation of addition.

Example 459. Let $f : \mathbb{Z} \to 2\mathbb{Z}$ be the map defined by $f(x) = 2x$. This map is a bijective group homomorphism, although it is not a C-ring homomorphism (cf. Example 185).

Example 460. The map

$$\begin{array}{rccc} f: & (\mathbb{R}^*, \cdot) & \longrightarrow & (\mathbb{R}, +) \\ & x & \longmapsto & \log(x^2) \end{array}$$

is a surjective homomorphism: $f(xy) = \log(xy)^2 = 2\log x + 2\log y = f(x) + f(y)$. Since the neutral element of $(\mathbb{R}, +)$ is 0,

$$\ker f = \{x \text{ such that } \log x^2 = 0\} = \{x \text{ such that } x^2 = 1\} = \{-1, +1\}.$$

Example 461 (Invertible matrices and SL_n)**.** Let $\mathbb{K}$ be any field. Let

$$GL_n(\mathbb{K}) \stackrel{\text{def}}{=} \{n \times n \text{ invertible matrices with entries in } \mathbb{K}\}.$$

This is a group with respect to row-by-column multiplication. The neutral element e is the identity matrix, with ones on the diagonal and zeroes elsewhere. Now consider the map

$$\begin{array}{rccc} f: & (GL_n(\mathbb{K}), \times) & \longrightarrow & (\mathbb{K}^*, \cdot) \\ & A & \longmapsto & \det A. \end{array}$$

Binet's formula (namely, $\det(AB) = \det A \det B$) implies that f is a homomorphism. Hence its kernel, called "special linear group", is a subgroup of $GL_n(\mathbb{K})$. It can be described as

$$SL_n(\mathbb{K}) \stackrel{\text{def}}{=} \{n \times n \text{ matrices with entries in } \mathbb{K} \text{ and determinant } 1\}.$$

Example 462 (Orthogonal matrices and rotations)**.** The *orthogonal group* is defined as

$$O_n(\mathbb{K}) \stackrel{\text{def}}{=} \{n \times n \text{ matrices A with entries in } \mathbb{K} \text{ such that } A\,A^T = I\}.$$

(Since $\det A = \det A^T$, Binet's formula tells us that

$$1 = \det I = \det(A\,A^T) = \det A \, \det A^T = (\det A)^2,$$

so matrices in $O_n(\mathbb{K})$ have determinant ± 1 and $O_n(\mathbb{K})$ could be viewed as a subgroup of $GL_n(\mathbb{K})$.) Consider the map

$$\begin{array}{rccc} f: & (O_n(\mathbb{K}), \times) & \longrightarrow & (\mathbb{K}^*, \cdot) \\ & A & \longmapsto & \det A. \end{array}$$

The kernel of this map is called the *rotation group* $SO_n(\mathbb{K})$. By definition,

$$SO_n(\mathbb{K}) \stackrel{\text{def}}{=} O_n(\mathbb{K}) \cap SL_n(\mathbb{K}).$$

Lagrange's theorem

Lemma 463. *Let a, b be two distinct elements of a finite group G. Let H be a subgroup of G. Then the sets $aH \stackrel{\text{def}}{=} \{ah \text{ such that } h \in H\}$ and $bH \stackrel{\text{def}}{=} \{bh \text{ such that } h \in H\}$, called* "left cosets"*, have the same cardinality, which is equal to the number of elements of H. The same is true for the sets $Ha \stackrel{\text{def}}{=} \{ha \text{ such that } h \in H\}$ and $Hb \stackrel{\text{def}}{=} \{hb \text{ such that } h \in H\}$, called* "right cosets"*.*

Proof. Fix a, b in G. The two functions

$$\begin{array}{rccc} \psi: & aH & \longrightarrow & bH \\ & x & \longmapsto & ba^{-1}x \end{array} \quad \text{and} \quad \begin{array}{rccc} \phi: & bH & \longrightarrow & aH \\ & x & \longmapsto & ab^{-1}x \end{array}$$

are well-defined and inverse of one another. In particular, since G finite, we infer that aH and bH have the same cardinality. This is true for any a, b in G; so in particular, it is true if we choose $b = e$. But $eH \stackrel{\text{def}}{=} \{eh \text{ such that } h \in H\} = H$. The proof for Ha and Hb is completely analogous, and left as exercise. □

Theorem 464 (Lagrange)**.** *Let G be a group with g elements. If H is a subgroup of G with h elements, then h divides g. In fact, $\frac{g}{h}$ counts the number of left cosets (and also the number of right cosets).*

Proof. If $H = G$ there is nothing to show. Otherwise, pick an element a_1 not in H. Clearly,

$$a_1 = a_1 e \in a_1 H.$$

If $G = H \cup a_1 H$ stop; otherwise, pick an element a_2 not in $H \cup a_1 H$. Clearly $a_2 \in a_2 H$. And so on. We claim that $H, a_1 H, a_2 H \ldots$ are all disjoint. Let us prove the claim by contradiction. Set $a_0 \stackrel{\text{def}}{=} e$. Suppose

$$x \in a_i H \cap a_j H, \text{ with } i < j.$$

This means that there exist elements $h_i, h_j \in H$ such that $a_i h_i = x = a_j h_j$. If we set $h \stackrel{\text{def}}{=} h_i h_j^{-1}$, we have that

$$a_j = x h_j^{-1} = a_i h_i h_j^{-1} = a_i h \in a_i H.$$

A contradiction, a_j was chosen outside $H \cup \ldots \cup a_{j-1} H$.

So the claim is proven. Since G is finite, the inductive process described above eventually ends, and at some point we will be able to write

$$G = H \cup a_1 H \cup \ldots \cup a_t H.$$

But by Lemma 463, these $t+1$ disjoint cosets all have the same cardinality, namely, h. So $n = (t+1)h$. This shows that $\frac{g}{h}$ counts the number of left cosets. With a symmetric argument, one can partition G into its right cosets (all of identical cardinality) and conclude that $\frac{g}{h}$ is also the number of right cosets. □

7.2 Permutations

Let n be any positive integer. Let $[n] \stackrel{\text{def}}{=} \{1, \ldots, n\}$. Let

$$\mathcal{S}_n \stackrel{\text{def}}{=} \{\sigma : [n] \longrightarrow [n] \text{ bijective}\}.$$

The elements of $\mathcal{S}_n$ are called *permutations*. There are three types of notation to write down the same permutation:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{pmatrix}, \quad \sigma = (12)(456), \quad \text{and } \sigma = (12)(45)(56).$$

The second notation writes σ as product of disjoint cycles; the third, as product of non-disjoint flips. We will explain them in a few minutes. The first notation is called *two-line notation* and it is the most intuitive: the rule is, σ maps each elements of the first row into the element of the second row immediately below. (In this case the first row is ordered, but it does not matter: What matters is that below each i sits $\sigma(i)$.) For example, $\sigma(3) = 3$. To compose two functions, we write them on top of one another, remembering that when we write $\tau \circ \sigma(1)$ the first permutation to be applied to 1 is σ, so σ should be on top. The two-line notation of $\tau \circ \sigma$ is then obtained by looking at only the first and the last row, ignoring all intermediate ones. For example, if

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 4 & 3 & 5 & 6 \end{pmatrix} \qquad \text{and} \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{pmatrix}$$

then

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \\ 2 & 1 & 4 & 5 & 6 & 3. \end{pmatrix} \qquad \text{and} \qquad \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 4 & 3 & 5 & 6 \\ 2 & 1 & 5 & 3 & 6 & 4. \end{pmatrix}$$

En passant, notice that $\tau \circ \sigma$ and $\sigma \circ \tau$ are different, so $\mathcal{S}_6$ is not an Abelian group.

Definition 465. Let $2 \leq k \leq n$ be integers. A *cycle (of length k)* in a permutation $\sigma \in \mathcal{S}_n$ is a k-tuple

$$(a_1, a_2, \ldots, a_k),$$

such that $\sigma(a_k) = a_1$ and $\sigma(a_i) = a_{i+1}$ for all $i \in \{1, \ldots, k-1\}$. Cycles of length two are called *flips* (or *transpositions*).
Any cycle g (of length k) is naturally associated to a permutation $\gamma \in \mathcal{S}_n$, as follows: $\gamma = \sigma$ on the elements of the cycle, and $\gamma = id$ otherwise.

Example 466. In the permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{pmatrix},$$

there is a cycle of length 3, namely, $g = (4, 5, 6)$. Its associated permutation is

$$\gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 5 & 6 & 4 \end{pmatrix}.$$

Theorem 467. *Every permutation different than the identity can be written as product of disjoint cycles, in a unique way.*

Proof. Let a_1 be the smallest integer such that $\sigma(a_1) \neq a_1$. Let t_1 be the smallest integer such that $\sigma^{t_1}(a_1) = a_1$. Then the first cycle is

$$\left(a_1, \sigma(a_1), \sigma^2(a_1), \ldots, \sigma^{t_1-1}(a_1)\right).$$

Now let a_2 be the smallest integer that does not belong to the cycle above, and satisfies $\sigma(a_2) \neq a_2$. Let t_2 be the smallest integer such that $\sigma^{t_2}(a_2) = a_2$. The second cycle is

$$\left(a_2, \sigma(a_2), \sigma^2(a_2), \ldots, \sigma^{t_2-1}b(a_2)\right).$$

And so on. We sketch the algorithm with the help of an example. Suppose

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 4 & 6 & 5 & 7 & 9 & 2 & 1 \end{pmatrix}.$$

To find the first cycle, we start with 1 and apply iteratively σ, until we get back to 1. So

$$\sigma(1) = 3, \quad \sigma(3) = 4, \quad \sigma(4) = 6, \quad \sigma(6) = 7, \quad \sigma(7) = 9, \quad \sigma(9) = 1.$$

So the first cycle is $(1, 3, 4, 6, 7, 9)$. Now let us consider the smallest integer not contained in this cycle, and apply σ repeatedly, until we get back to such integer. In our case, we re-start with 2:

$$\sigma(2) = 8, \quad \sigma(8) = 2.$$

So $(2, 8)$ is the second cycle. By construction, it is disjoint from the first cycle, because σ is injective. Now the smallest integer that belongs to neither of the previous cycles is 5. Since $\sigma(5) = 5$, we are done. Our final result is

$$\sigma = (1, 3, 4, 6, 7, 9)(2, 8).$$

Now, technically what we found is just a *list* of disjoint cycles. But if we interpreted every cycle as its associated permutation in $\mathcal{S}_n$, the list can actually be interpreted a product of permutations. More precisely, if $\gamma, \gamma_1, \gamma_2$ are the permutations of $\mathcal{S}_n$ associated respectively to σ, to $(1, 3, 4, 6, 7, 9)$, and to $(2, 8)$, then it is clear that

$$\gamma = \gamma_1 \circ \gamma_2.$$

For this reason, we speak of "product of cycles". Note that disjoint cycles commute:

$$\gamma = \gamma_1 \circ \gamma_2 = \gamma_2 \circ \gamma_1.$$

To complete our "proof by example", we claim that up to commuting the disjoint cycles, this decomposition is unique. This is easy: Let

$$\gamma = \eta_1 \circ \ldots \circ \eta_k$$

be another decomposition into disjoint cycles. Without loss of generality, suppose 1 appears in η_1. Since $\sigma(1) = 3, \sigma(3) = 4$, etc., it is clear that η_1 must be the cycle $(1, 3, 4, 6, 7, 9)$. Similarly, suppose η_2 is the cycle containing 2: Then $\eta_2 = (2, 8)$. Since $\eta(5)$ must be 5, we conclude that $\gamma_i = \eta_i$ for all i. □

Lemma 468. *Every cycle of length k can be written as product of $k-1$ non-disjoint flips (not necessarily in a unique way).*

Proof. If $k \geq 2$, we claim that

$$(a_1, a_2, \ldots, a_k) = (a_1, a_2)(a_2, a_3) \cdots (a_{k-2}, a_{k-1})(a_{k-1}, a_k).$$

By this we mean that if γ is the permutation of $\mathcal{S}_n$ associated to $(a_1, \ldots, a_k)$, and γ_i is the permutation of $\mathcal{S}_n$ associated to (a_i, a_{i+1}), then

$$\gamma = \gamma_1 \circ \gamma_2 \circ \ldots \circ \gamma_{k-1}.$$

As a warm up, let us check this first for the element a_1. By definition, $\gamma(a_1) = a_2$. On the other hand, γ_i swaps a_i with a_{i+1}, so it has no effect on a_1 if $i \geq 2$. Formally,

$$\gamma_i(a_1) = \begin{cases} a_1 & \text{if } i \geq 2 \\ a_2 & \text{if } i = 1. \end{cases}$$

So

$$\gamma_1 \circ \gamma_2 \circ \ldots \circ \gamma_{k-1}(a_1) = \gamma_1 \circ \gamma_2 \circ \ldots \circ \gamma_{k-2}(a_1) = \ldots = \gamma_1(a_1) = a_2,$$

as desired. Now let us check the effect on the generic element a_j, with $j < k$. Clearly $\gamma(a_j) = a_{j+1}$, with the exception of a_k, for which $\gamma(a_k) = a_1$. On the other hand,

$$\gamma_i(a_j) = \begin{cases} a_j & \text{if } i \geq j+1 \\ a_{j+1} & \text{if } i = j \\ a_{j-1} & \text{if } i = j-1 \\ a_j & \text{if } i \leq j-2. \end{cases}$$

So if $j < k$, we have

$$\gamma_1 \circ \ldots \circ \gamma_{k-1}(a_j) = \ldots = \gamma_1 \circ \ldots \circ \gamma_j(a_j) = \gamma_1 \circ \ldots \circ \gamma_{j-1}(a_{j+1} = \ldots = \gamma_1(a_{j+1}) = a_{j+1}.$$

For a_k instead we have

$$\gamma_1 \circ \ldots \circ \gamma_{k-1}(a_k) = \gamma_1 \circ \ldots \circ \gamma_{k-2}(a_{k-1}) = \gamma_1 \circ \ldots \circ \gamma_{k-3}(a_{k-2}) = \ldots = \gamma_1(a_2) = a_1. \quad \square$$

Example 469. Let us verify that $(1, 3, 4, 6, 7, 9) = (1, 3)(3, 4)(4, 6)(6, 7)(7, 9)$. In fact, the right hand side is given by the first and the last row of the matrix

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 3 & 4 & 5 & 6 & 9 & 8 & 7 \\ 1 & 2 & 3 & 4 & 5 & 7 & 9 & 8 & 6 \\ 1 & 2 & 3 & 6 & 5 & 7 & 9 & 8 & 4 \\ 1 & 2 & 4 & 6 & 5 & 7 & 9 & 8 & 3 \\ 3 & 2 & 4 & 6 & 5 & 7 & 9 & 8 & 1 \end{pmatrix}$$

and the left hand side is precisely

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 2 & 4 & 6 & 5 & 7 & 9 & 8 & 1 \end{pmatrix}$$

Definition 470. A permutation $\sigma \in \mathcal{S}_n$ is called *even* if it can be written as the product of an even number of flips, and *odd* if it can be written as the product of an odd number of flips.

Example 471. A k-cycle is an even permutation if k is odd, and an odd permutation if k is even. In fact, any k-cycle is the product of $k-1$ flips.

Theorem 472. *Every permutation of $\mathcal{S}_n$ is either even or odd, but not both.*

Proof. The identity can be written as $(12)(12)$, so it is even. Any other element of $\mathcal{S}_n$ is a product of disjoint cycles by Theorem 467, and each cycle in turn decomposes into (not necessarily disjoint) flips as in Lemma 468. This proves that every permutation is either even or odd. It remains to prove that a product of an odd number of flips cannot be rewritten as a product of an even number of flips. To see this, consider the polynomial of

$$p_n \stackrel{\text{def}}{=} \prod_{1 \le i < j \le n} (x_i - x_j).$$

Now consider any flip (i, j), with $i < j$. This naturally induces a map from $\mathbb{R}[x_1, \ldots, x_n]$ to itself, as follows: given a polynomial, we replace every occurrence of x_i with x_j, and the other way around. We claim that this map sends p_n to $-p_n$. In fact, the factor $(x_i - x_j)$ is replaced by $(x_j - x_i)$; also, for any a such that $i < a < j$, *both* factors $(x_i - x_a)$ and $(x_a - x_j)$ change sign, and sign changes in an even number cancel out; and finally, all other factors stay the same. For example, consider in $\mathcal{S}_4$ the flip $(2, 4)$. The associated map sends the polynomial

$$p_4 \stackrel{\text{def}}{=} (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4)$$

to the polynomial

$$(x_1 - x_4)(x_1 - x_3)(x_1 - x_2)(x_4 - x_3)(x_4 - x_2)(x_3 - x_2) = -p_4.$$

More generally, any permutation σ induces a map from $\mathbb{R}[x_1, \ldots, x_n]$ to itself, namely, the map that for all i replaces x_i with $x_{\sigma(i)}$. Since every permutation is a composition of flips, it is clear that any odd permutation will map p_n to $-p_n$, whereas any even permutation will map p_n to itself. Hence no permutation can be simultaneously even and odd. □

Corollary 473. *For each $n \ge 3$, every even permutation of $\mathcal{S}_n$ can be written as product of 3-cycles.*

Proof. By Theorem 472 any even permutationcan be written as product of an even number $2s$ of flips. The trick is to simply to pair them. We claim in fact that the product of any two flips is either the identity, or a 3-cycle, or a product of two 3-cycles. In fact,

- if $a = c$ and $b = d$, then $(a, b)(a, b) = id$;
- if $a = c$ but $b \neq d$, then $(a, b)(a, d) = (a, d, b)$;
- if a, b, c, d are all distinct, then $(a, b)(c, d) = (a, b)(a, c)(a, c)(c, d) = (a, b, c)(c, d, a)$. □

Proposition 474. *For $n \ge 2$, the set*

$$A_n \stackrel{\text{def}}{=} \{\textit{even permutations}\}$$

is a subgroup of $\mathcal{S}_n$ with exactly $\frac{n!}{2}$ elements.

Proof. Consider the following function between sets

$$\begin{array}{rccc} \psi: & A_n & \longrightarrow & (\mathcal{S}_n \setminus A_n) \\ & \sigma & \longmapsto & \sigma \circ (1,2). \end{array}$$

This function is well-defined because composing with a single flip γ changes parity. Moreover, it is invertible, the inverse being

$$\begin{array}{rccc} \phi: & (\mathcal{S}_n \setminus A_n) & \longrightarrow & A_n \\ & \tau & \longmapsto & \tau \circ (1,2). \end{array}$$

This proves that A_n and its complement within $\mathcal{S}_n$ have the same number of elements. On the other hand, $\mathcal{S}_n$ has $n!$ elements, because to write down a bijection $\sigma : [n] \to [n]$ we have n choices for $\sigma(1)$, $n-1$ choices for $\sigma(2)$, and so on. Hence, A_n has $\frac{n!}{2}$ elements.

Now let σ and τ be even permutations. Write $\sigma = \gamma_1 \cdots \gamma_{2s}$ and $\tau = \mu_1 \cdots \gamma_{2t}$ for some flips γ_i, μ_j. Applying the same flip twice is the same as leaving things unchanged, so $\gamma_i^{-1} = \gamma_i$. But then

$$\sigma^{-1} \circ \tau = (\gamma_1 \cdots \gamma_2 s)^{-1} \mu_1 \cdots \mu_{2t} = (\gamma_{2s}^{-1} \cdots \gamma_1^{-1}) \mu_1 \cdots \mu_{2t} = \gamma_{2s} \cdots \gamma_1 \mu_1 \cdots \mu_{2t}$$

is the product of an even number of flips. By Proposition 451, A_n is a subgroup. □

It turns out that A_n has a richer property than just being a subgroup. In fact, for each $\sigma, \tau \in \mathcal{S}_n$, if τ is even, then the permutation $\sigma\tau\sigma^{-1}$ is also even. This property is called *normality* and it is crucial to form quotients.

7.3 Normal subgroups, quotients, and simple groups

Definition 475 (Normal)**.** A subgroup H of G is called *normal* if for each $g \in G$, for each $h \in H$, $ghg^{-1} \in H$.

Example 476. If the elements g and h commute, then $ghg^{-1} = hgg^{-1} = h$. So certainly in groups where the operation is commutative, like $(\mathbb{Z}, +)$ or $(\mathbb{Q}^*, \cdot)$, every subgroup is normal.

Example 477. Given any group G, the subgroups $H = \{0\}$ and $H = G$ are always normal.

Non-Example 478. Consider in $G = GL_2(\mathbb{R})$ the subgroup $H = UT_2(\mathbb{R})$ of upper triangular matrices, with nonzero determinant. Then H is not normal. In fact, choosing

$$h = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } g = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{one has } ghg^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \notin UT_2(\mathbb{R}).$$

Lemma 479. *Let H be a subgroup of a group G. The following are equivalent:*
(1) *H is normal.*
(2) *For each g in G, for any $h \in H$, there is a k in H such that $gh = kg$.*
(3) *For each g in G, the sets $gH \stackrel{\text{def}}{=} \{gh \text{ such that } h \in H\}$ and $Hg \stackrel{\text{def}}{=} \{hg \text{ such that } g \in G\}$ coincide.*
(4) *For each $a, b \in G$, $ab^{-1} \in H$ if and only if $a^{-1}b \in H$.*

Proof.
(1) ⇒ (2): We know that $ghg^{-1} \in H$. Setting $k \stackrel{\text{def}}{=} ghg^{-1}$, we have $kg = gh$.
(2) ⇒ (3): $gH \subseteq Hg$, because any element of the form gh can also be rewritten in the form kg for some $k \in H$. Symmetrically, $Hg \subseteq gH$. So $gH = Hg$.

(3) $\Rightarrow$ (4): Suppose $ab^{-1} \in H$. Set $h \stackrel{\text{def}}{=} ab^{-1}$. Then $a = ab^{-1}b = hb \in Hb$. Since by assumption $Hb \subseteq bH$, it follows that $a = bk$ for some $k \in H$. So $a^{-1}b = (bk)^{-1}b = k^{-1}b^{-1}b = k^{-1}$ is in H. The converse implication is similar: if $k = b^{-1}a \in H$, then $a = bk \in bH \subseteq Hb$, so we can find $h \in H$ such that $a = hb$. Hence, $ab^{-1} = h \in H$.

(4) $\Rightarrow$ (1): For every h in H and for every G in G, we want to show that ghg^{-1} is in H. In other words, if we set $a \stackrel{\text{def}}{=} gh$ and $b \stackrel{\text{def}}{=} g$, we want to show that ab^{-1} is in H. But by the assumption, this is equivalent to proving that $a^{-1}b \in H$. But $a^{-1}b = h^{-1}g^{-1}g = h^{-1}$. □

Proposition 480. *If a subgroup H of G contains half of the elements of G, then H is normal.*

Proof. Let x be an element of G that is not in H. The left coset xH is disjoint from H and has the same number of elements of H (cf. Lemma 463), so xH is simply the complement of H. The same applies to Hx. But then $xH = Hx$, so by Lemma 479 H is normal. □

Corollary 481. *The set*

$$A_n \stackrel{\text{def}}{=} \{\textit{even permutations}\}$$

is a normal subgroup of $\mathcal{S}_n$.

Proof. This follows straightforwardly either from the definition, or from the fact that A_n has half the elements of $\mathcal{S}_n$ (cf. Proposition 474.) □

Quotients

Definition 482. Let H be a **normal** subgroup of G. Let $\sim$ be the relation of equivalence

$$a \sim b \stackrel{\text{def}}{\Longleftrightarrow} a^{-1}b \in H \text{ (or equivalently, } \stackrel{\text{def}}{\Longleftrightarrow} ab^{-1} \in H).$$

Then the quotient

$$G/H \stackrel{\text{def}}{=} \{\overline{g} \text{ such that } g \in G\},$$

is the set of all classes of equivalences.

Proposition 483. *The classes of equivalence of $\sim$ are the left (or equivalently, since H is normal, "right") cosets with respect to H. In particular, if G has g elements and H has h elements, then G/H has $\frac{g}{h}$ elements.*

Proof. For any a, b in G,

$$a \sim b \stackrel{\text{def}}{\Longleftrightarrow} \exists h \in H \text{ such that } a^{-1}b = h \Longleftrightarrow \exists h \in H \text{ such that } b = ah \Longleftrightarrow b \in aH.$$

So the set of elements in a relation with a is precisely aH. Hence, the elements of G/H are the (left) cosets. By theorem 464, $\frac{g}{h}$ counts precisely the left cosets, so we conclude. □

Proposition 484. *Let H be a* normal *subgroup of G. Then G/H is a group with respect to the operation*

$$\overline{a}\,\overline{b} \stackrel{\text{def}}{=} \overline{ab}$$

which makes the projection

$$p : \begin{array}{ccc} G & \to & G/H \\ g & \mapsto & \overline{g} \end{array}$$

a surjective group homomorphism.

Proof. The fact that H is normal is crucial to verify that the operation is well defined. In fact, if $x' \sim x$ and $y' \sim y$, then $h_x \overset{\text{def}}{=} x'x^{-1}$ and $h_y \overset{\text{def}}{=} y'y^{-1}$ both belong to H. Now

$$x'y'(xy)^{-1} = x'y'y^{-1}x^{-1} = x'h_yx^{-1}.$$

Since H is normal, we can write $x'h_y = kx'$, for some $k \in H$. So we can continue the chain of equalities with

$$x'h_yx^{-1} = kx'x^{-1} = kh_x,$$

which is an element of H. Hence, $x'y'(xy)^{-1} \in H$, which means that $x'y' \sim xy$. This shows that the operation is well defined.

The rest is easy: The neutral element is $\overline{e}$, where e is the neutral element of G, and the inverse of $\overline{x}$ is $p(x^{-1}) = \overline{x^{-1}}$. □

Theorem 485 (First Homomorphism Theorem for Groups, Noether 1927)**.** *Let $f : A \to B$ be any homomorphism between two groups. Then, there exist a (unique) homomorphism*

$$g : {}^{A}\!/_{\ker f} \to B \qquad \textit{such that}$$

1) *g is injective;*
2) $\operatorname{Im} g = \operatorname{Im} f$*;*
3) $f = g \circ p$.

Proof. Let us start from the end and force property number 3) by defining

$$g(\overline{a}) \overset{\text{def}}{=} f(a) \quad \text{for all } a.$$

Is this a good definition? If a, a' are distinct elements of A such that $\overline{a} = \overline{a'}$, is it true that $f(a) = f(a')$? By definition of quotient, $\overline{a'} = \overline{a}$ if and only if $a'a^{-1} \in \ker f$, if and only if $f(a'a^{-1}) = e_B$. Multiplying both sides by $f(a)$, this is the same as saying, $f(a'a^{-1})f(a) = f(a)$, or in other words, $f(a') = f(a)$, cf. Proposition 456. Summing up,

$$\overline{a'} = \overline{a} \text{ in } {}^{A}\!/_{\ker f} \iff f(a) = f(a') \overset{\text{def}}{\iff} g(\overline{a}) = g(\overline{a'}).$$

The stream of implications from left to right tells us that g is a well-defined function; the converse implications, from right to left, tell us that g is injective. It remains to see that $\operatorname{Im} g = \operatorname{Im} f$. Let $b \in B$:

$$b \in \operatorname{Im} g \iff \exists a \in A \text{ such that } g(\overline{a}) = b \overset{\text{def}}{\iff} \exists a \in A \text{ such that } f(a) = b \iff b \in \operatorname{Im} f. \qquad \square$$

Corollary 486. *If there is a surjective homomorphism $f : G \longrightarrow H$, then H is isomorphic to ${}^{G}\!/_{\ker f}$.*

Example 487. In Example 461 we introduced the kernel $SL_n(\mathbb{K})$ of the surjective homomorphism

$$\begin{array}{rccc} f: & (GL_n(\mathbb{K}), \times) & \longrightarrow & (\mathbb{K}^*, \cdot) \\ & A & \longmapsto & \det A. \end{array}$$

By Corollary 486,

$$\frac{GL_n(\mathbb{K})}{SL_n(\mathbb{K})} \text{ is isomorphic to } \mathbb{K}^*.$$

Example 488. Consider the surjective homomorphism

$$\begin{array}{rccc} \pi_1: & G \times H & \longrightarrow & G \\ & (g,h) & \longmapsto & g. \end{array}$$

Since $\ker \pi_1 = \{(g,h) \text{ such that } g = e_G\} = \{e_G\} \times H$, by Corollary 486 we have that

$$\frac{G \times H}{\{e_G\} \times H} \text{ is isomorphic to } G.$$

We conclude this Chapter with a useful proposition on normal subgroups.

Proposition 489. *Let H, K be two normal subgroups of a group G. If $H \cap K = (e)$, then the smallest subgroup of G containing both H and K is isomorphic to $H \times K$.*

Proof. Let us start with an observation. For any h in H and for any k in K, by the normality of K we have that $hkh^{-1} \in K$; similarly, by the normality of H, we have that $kh^{-1}k^{-1} \in H$. In particular,

$$(hkh^{-1})k^{-1} = h(kh^{-1}k^{-1}) \in H \cap K.$$

But then by assumption $hkh^{-1}k^{-1} = e$, which means $hk = kh$. So the elements of H always commute with the elements of K. Now let us set

$$\begin{array}{rccc} \varphi: & H \times K & \longrightarrow & G \\ & (h,k) & \longmapsto & hk. \end{array}$$

Is it a homomorphism? Indeed, since h' commutes with k,

$$\varphi(h,k) \cdot \varphi(h',k') = hk \cdot h'k' = hh' \cdot kk' = \varphi(hh', kk') = \varphi((h,k) \cdot (h',k')).$$

Let us check that φ is injective. Assume $hk = e$. Then

$$h^{-1} = h^{-1}e = h^{-1}hk = k.$$

But then since $h^{-1} \in H$, we have that $k = h^{-1} \in H \cap K$, which implies $k = e$, so $h = e$. Hence φ is injective. So $H \times K$ is isomorphic to $\operatorname{Im}\varphi = \{hk \text{ such that } h \in H, k \in K\}$.

To conclude our proof, it remains to show that $\operatorname{Im}\varphi$ is the smallest subgroup containing H and K. Indeed $\operatorname{Im}\varphi$ contains any element h of H, which can be written as $h = he$. Symmetrically, it contains any element of K, by writing it as $k = ek$. Finally, any subgroup containing H and K must contain the products of their elements; so it must contain $\operatorname{Im}\varphi$. □

Simple Groups

A_n has a very interesting property: When $n \geq 5$, it has no normal subgroup other than the trivial ones. Below we prove it only for $n = 5$.

Definition 490. We say that a group G is *simple* if its only normal subgroups are $\{e\}$ and G itself.

Proposition 491. *If G has a prime number of elements, then G is simple.*

Proof. Let p be the number of elements of G. Let H be any subgroup of G (normal or not). By Lagrange's theorem 464), the cardinality of H must divide the prime number p, so it is either 1 or p. Hence either $H = \{e\}$ or $H = G$ □

The converse of the previous proposition is false in general.

Proposition 492. *A_5, which has 60 elements, is simple.*

Proof. Let N be a normal subgroup of A_5 different than $\{e\}$. We want to show that $N = A_5$. The proof is in two parts:
(A) We show that if N contains one 3-cycle, then it contains all 3-cycles.
(B) We show that N must contain a 3-cycle.
From this the conclusion follows using Corollary 473: since N contains all 3-cycles, any permutation of A_5 can be written as product of elements of N and is therefore in N. So, let us prove the claims.

(A) Suppose N contains the 3-cycle (a', b', c'). Let u', v' be the other two elements of [5]. Now let (a, b, c) be an arbitrary 3-cycle, and let u, v be the other two elements of [5]. Consider the two permutations

$$\sigma_0 = \begin{pmatrix} a' & b' & c' & u' & v' \\ a & b & c & u & v \end{pmatrix} \quad \text{and} \quad \sigma_1 = \begin{pmatrix} a' & b' & c' & u' & v' \\ a & b & c & v & u \end{pmatrix}.$$

Since they differ by a flip (namely, $\sigma_0 = \sigma_1 \circ (u, v)$), one of σ_0 and σ_1 must be even: Call it σ. Then, it is easy to see that

$$\sigma \circ (a', b', c') \circ \sigma^{-1} = (a, b, c).$$

But then since N is normal and (a', b', c') belongs to N, so does (a, b, c).

(B) By contradiction, suppose N does not contain any 3-cycles. Then a generic nontrivial element of $N \subseteq A_5$ can be
- either a product of two disjoint flip, $(a, b)(c, d)$,
- or a 5-cycle (v, w, x, y, z).

We will show that both cases leads to a contradiction.
- Suppose $(a, b)(c, d) \in N$. Let f be the element of [5] different than a, b, c and d. Since N is normal, it must contain also

$$(c, f, d)\,[(a, b)(c, d)]\,(c, f, d)^{-1} = (a, b)(c, f).$$

But then N contains

$$[(a, b)(c, d)]\,[(a, b)(c, f)] = (c, f, d),$$

which is a 3-cycle: A contradiction.
- Suppose $(v, w, x, y, z) \in N$. Then by normality N contains also

$$[(v, w)(x, y)]\,(v, w, x, y, z)\,[(v, w)(x, y)]^{-1} = (v, y, x, z, w).$$

So N contains the product

$$(v, w, x, y, z)(v, y, x, z, w) = (v, z, c),$$

which is a 3-cycle: A contradiction. □

In fact, the proof of the above Proposition can be "expanded" to prove an interesting fact about $\mathcal{S}_5$:

Proposition 493. *The only normal subgroups of $\mathcal{S}_5$ are $\{e\}$, A_5 and $\mathcal{S}_5$.*

Proof. Let N be a normal subgroup of $\mathcal{S}_5$ different than $\{e\}$. We want to show that N contains A_5. From this the claim follows because A_5 has half of the elements of $\mathcal{S}_5$ (so by Lagrange's Theorem 464 there is no subgroup strictly between A_5 and $\mathcal{S}_5$).

From the proof of Proposition 492 we already know that if N contains even flips like a 3-cycle $(a', b'c')$, or a product $(a, b)(c, d)$ of two 2-cycles, or a 5-cycle like (v, w, x, y, z), then N contains a *all* 3-cycles and so by Corollary 473 N contains A_5. But a priori, a generic element of N could also be

- a single flip (a, b), or
- a 4-cycle (v, w, x, y).

Let us show that even in these two cases, N contains A_5.

- Suppose $(a, b) \in N$. Then for all $c \in [5]$ different than a and b we have
$$(a, c) = (b, c)\,(a, b)\,(b, c)$$
so by normality $(a, c) \in N$. Reapplying the same reasoning, for any $d \neq a$ we get that
$$(c, d) = (a, d)\,(a, c)\,(a, d)$$
so by normality $(c, d) \in N$. This means that N contains all flips. So $N = \mathcal{S}_5$.
- Suppose $(v, w, x, y) \in N$. Then by normality N contains also
$$(v, x, y, w) = (v, w)\,(v, w, x, y)\,(v, w).$$
But since N contains (v, w, x, y) and (v, x, y, w), it contains also the 3-cycle
$$(v, y, x) = (v, w, x, y)\,(v, x, y, w),$$
so N contains all 3-cycles and again by Corollary 473 N contains A_5. □

Here is another property of $\mathcal{S}_5$ that will be useful later on:

Proposition 494. *If a subgroup of $\mathcal{S}_5$ (normal or not) contains both a 5-cycle and a flip, it is the whole $\mathcal{S}_5$.*

Proof. Without loss of generality, suppose S is the subgroup generated by $(1, 2, 3, 4, 5)$ and by a flip $(1, k)$. Then it contains the product of cycles
$$(1, 2, \ldots, 5)\,(1, k)\,(1, 2, \ldots, 5)^{-1} = (2, k + 1).$$
Similarly, it also contains
$$(1, 2, \ldots, 5)\,(2, k + 1)\,(1, 2, \ldots, 5)^{-1} = (3, k + 2).$$
Iterating, S contains all flips $(a, a + k - 1)$, where the sum is taken mod 5. But then S contains also $(k, 2k - 1)$ and the product
$$(k, 2k - 1)\,(1, k)\,(k, 2k - 1) = (1, 2k - 1).$$
So starting from $(1, k)$ in S, we found out that $(1, 2k - 1)$ is in S. We can reapply the whole reasoning from the beginning then, starting from $(1, 2k - 1)$ instead of $(1, k)$. We would obtain that all flips $(a, a + 2k - 1)$ are in S, and that $(1, 2(2k - 1) - 1)$ is in S. And so on. But the map
$$\begin{array}{rccl} \varphi: & \{1, \ldots 5\} & \longrightarrow & \{1, \ldots, 5\} \\ & x & \longmapsto & 2x - 1 \mod 5 \end{array}$$
is bijective. Written as permutation, it would be the 4-cycle $(3, 5, 4, 2)$. In particular, just by reiterating the argument above, we obtain that S contains all flips $(1, b)$, for any $b \geq 2$, as well as all flips of the type $(a, a + b - 1)$. So S contains all flips. By Theorem 472, we conclude. □

Remark 495. While it is true that $(1,2)$ and $(1,2,\ldots,n)$ generate $\mathcal{S}_n$ for all n, to claim that *any* 2-cycle and any n-cycle generate $\mathcal{S}_n$ it is necessary to assume that n is prime. For example, $(1,3)$ and $(1,2,3,4)$ do not generate $\mathcal{S}_4$. More generally, if d is a divisor of n, then $(1,d+1)$ and $(1,2,\ldots,n)$ do not generate $\mathcal{S}_n$. In the exercises you are asked to prove that $(1,\ldots,n)$ and $(n-1,n)$ generate $\mathcal{S}_n$.

*Finite fields, and their algebraic closure

We already know by Lehmer's theorem 152 that if $\mathbb{K} = \mathbb{Z}_p$, then the "multiplicative group" $\mathbb{K} \setminus \{0\}$ is cyclic. We now show that this result is also true for finite fields. (For this result, Lagrange's theorem for groups is crucial.) As a consequence, we prove that the algebraic closure of any $\mathbb{Z}_p$ is infinite (Corollary 499).

Theorem 496 (Primitive root theorem)**.** *If $\mathbb{K}$ is any finite field, the "multiplicative group" $\mathbb{K} \setminus \{0\}$ is cyclic.*

Proof. Let $\mathbb{K}$ be a finite field. By Theorem 374, $\mathbb{K}$ is an extension of $\mathbb{Z}_p$, for some p prime. Let m be the cardinality of $\mathbb{K} \setminus \{0\}$, which is a group with respect to multiplication. Let λ be the maximum of the periods of the elements in $\mathbb{K} \setminus \{0\}$, or in other words, the exponent of $\mathbb{K} \setminus \{0\}$. Then from Proposition 513 we know that λ divides m and $a^\lambda = 1$ for all $a \in \mathbb{K}$. So the polynomial $x^\lambda - 1$ has at least m roots in $\mathbb{K}$. But being $\mathbb{K}[x]$ a domain, $x^\lambda - 1$ has also at most λ roots in $\mathbb{K}$, and $\lambda \leq m$. It follows that $\lambda = m$. Hence, $\mathbb{K} \setminus \{0\}$ is cyclic. □

Theorem 497. *There exists a field with q elements if and only if q is a prime power.*

Proof. (From Lindsay Childs, A concrete Introduction to Higher Algebra, page 501) "⇒": Let $\mathbb{K}$ be a finite field. By Theorem 374, $\mathbb{K}$ is an extension of $\mathbb{Z}_p$, for some p prime. Also, by the Primitive Root Theorem, there is an element α of $\mathbb{K}$ such that every nonzero element of $\mathbb{K}$ can be written uniquely as α^s for some non-negative integer s. Now, define a map $\phi : \mathbb{Z}_p[x] \to \mathbb{K}$ by sending $\overline{1}$ to $1_\mathbb{K}$ and x to α. Basically, ϕ is the composition of the inclusion $\mathbb{Z}_p[x] \to \mathbb{K}[x]$ with the evaluation at α. Clearly ϕ is a surjective homomorphism, because α is a primitive root. Since $\mathbb{Z}_p[x]$ is a PID, there is some polynomial $g \in \mathbb{Z}_p[x]$ for which $\ker \phi = (g)$. Thus by the first isomorphism theorem of C-rings,

$$\mathbb{K} \cong {}^{\mathbb{Z}_p[x]}\!/_{(g)},$$

and g is some polynomial that must be irreducible, otherwise the quotient would not be a field. In particular, $\deg g \geq 1$. But then it is easy to see that ${}^{\mathbb{Z}_p[x]}\!/_{(g)}$ has exactly $p^{\deg g}$ elements.

"⇐": Let $q \stackrel{\text{def}}{=} p^n$ for some prime p and some positive integer n. Consider the polynomial

$$h(x) \stackrel{\text{def}}{=} x^q - x.$$

In the algebraic closure of $\mathbb{Z}_p$, this h decomposes as product of (monic) linear factors. We know that all roots of h are distinct because of the "derivative trick": in fact, since in $\overline{\mathbb{Z}_p}$ one has $q = 0$, we have

$$D(h) = qx^{q-1} - 1 = -1,$$

so by Proposition 419 h has no multiple roots. Now let S be the subset of $\overline{\mathbb{Z}_p}$ contanining all roots of h, or more formally,

$$S \stackrel{\text{def}}{=} \{\alpha \in \overline{\mathbb{Z}_p} \,:\, \alpha^q = \alpha\}.$$

By what we said above, S has exactly q elements. It remains to see that S is a field. This is easy: if α and β are in S, then

- so is $\alpha+\beta$, because $(\alpha+\beta)^q = \alpha^q + \beta^q = \alpha+\beta$, where the first equality follows by the same argument of the Freshman's dream;
- so is $\alpha\cdot\beta$, because $(\alpha\cdot\beta)^q = \alpha^q\cdot\beta^q = \alpha\cdot\beta$;
- so is $-\alpha$, because when $p=2$ $-\alpha=\alpha$, and when p is odd, q is also odd, so $(-\alpha)^q = (-1)^q\cdot\alpha^q = -1\cdot\alpha = -\alpha$;
- so is α^{-1}, because $(\alpha^{-1})^q = (\alpha^q)^{-1} = \alpha^{-1}$. □

Corollary 498. *For any positive integer d and any prime p there is an irreducible polynomial in $\mathbb{Z}_p[x]$ of degree d.*

Proof. By Theorem 497, there exists a field $\mathbb{F}$ with p^d elements. The direction "⇒" of Theorem 497 also tells us that $\mathbb{F}$ is isomorphic to a quotient of $\mathbb{Z}_p[x]$ by some irreducible polynomial of degree d. □

Corollary 499. *The algebraic closure of $\mathbb{Z}_p$, with p prime, is a field extension of $\mathbb{Z}_p$ that has infinitely many elements.*

Proof. By Lemma 410, $\overline{\mathbb{Z}_p}$ is a field. The fact that $\overline{\mathbb{Z}_p}$ is an extension of $\mathbb{Z}_p$ follows from Theorem 374 and the fact that $1+1+\ldots+1$ (p times) is equal to zero in $\overline{\mathbb{Z}_p}$. By the previous Corollary, there exists an irreducible polynomial in $\mathbb{Z}_p[x]$ of degree d. Any root r_d of such polynomial is algebraic of degree d over $\mathbb{Z}_p$. But then already the subset $\{r_d \text{ such that } d\in\mathbb{N}, d\neq 0\}$ of $\overline{\mathbb{Z}_p}$ is infinite, because if $d\neq d'$ it is obvious that $r_d \neq r_{d'}$. □

7.4 Generators, periods, and cyclic groups

Definition 500. Let a be an element of a group G. Let $z\in\mathbb{Z}$. We define

$$a^z \stackrel{\text{def}}{=} \begin{cases} aa\cdots a \text{ (z times)}, & \text{if } z>0, \\ e & \text{if } z=0, \\ a^{-1}\cdots a^{-1} \text{ (-z times)}, & \text{if } z<0. \end{cases}$$

It is clear that the inverse of a^n is $(a^{-1})^n$, which by definition is a^{-n}.

Proposition 501 (Power properties)**.** *Let G be a group. Let a be any element of G. For any integers z, w one has*

$$a^z a^w = a^{z+w} \qquad \text{and} \qquad (a^z)^w = a^{zw}.$$

Proof. Left as exercise. (Hint: Do the case $w>0$ and $z>0$ first. Then do all other cases.) □

Definition 502 (period)**.** Let G be a group. Let a be an element of G. The *period of a* is defined as

$$\pi(a) \stackrel{\text{def}}{=} \begin{cases} +\infty & \text{if all powers of } a \text{ are distinct}, \\ t & \text{if } t \text{ is the smallest positive integer for which } a^t = e. \end{cases}$$

Remark 503. The two cases above cover all possibilities: If it is not true that all powers of a are distinct, then $a^z = a^w$ for some $z\neq w$, whence using cancellation (Proposition 448) we get that $a^{z-w} = e = a^{w-z}$. This means that the set of integers $n>0$ for which $a^n = e$ is non empty.

Lemma 504. *Let a be an element of a group G. Let t be a positive integer. Then*

$$a^t = e \iff t \text{ is a multiple of } \pi(a).$$

Proof.

"⇐" Easy: if $t = m\pi(a)$, then $a^t = (a^{\pi(a)})^m = e^m = e$.
"⇒" Let us perform a Euclidean division

$$t = q\pi(a) + r \text{ with } 0 \leq r < \pi(a).$$

If $r = 0$ then t is a multiple of $\pi(a)$ and we are done. If $r > 0$, we have

$$e = a^t = a^{q\pi(a)+r} = (a^{\pi(a)})^q a^r = e^q a^r = a^r,$$

a contradiction with the definition of period: r is smaller than $\pi(a)$. □

Proposition 505. *Let G be a group with n elements. If an element a has period m, then a^k has period $\frac{m}{\gcd(m,k)}$.*

Proof. Set $m' \stackrel{\text{def}}{=} \frac{m}{\gcd(d,k)}$ and $k' \stackrel{\text{def}}{=} \frac{k}{\gcd(m,k)}$; clearly, $\gcd(m', k') = 1$ and $m'k = m'k' \gcd(m,k) = mk'$. Since a has period m, we know that

$$e = a^m = a^{m' \gcd(m,k)}.$$

So by the Power Properties (Proposition 501) we have

$$(a^k)^{m'} = a^{km'} = a^{k'm} = (a^m)^{k'} = e^{k'} = e.$$

Now let t be any integer such that

$$(a^k)^t = e.$$

We want to prove that t is a multiple of m'. From $a^{kt} = e$, using Lemma 504 we see that kt must be a multiple of $\pi(a) = m$. Write this as

$$kt = mq \text{ for some } q \in \mathbb{N}. \tag{53}$$

Dividing by the common factor $\gcd(m,k)$, Equation 53 becomes

$$k't = m'q.$$

Since k' has no factor in common with m', by the Unique Factorization Theorem t must be a multiple of m'. □

We now wonder how the period behaves with respect to the product.

Proposition 506. *Let G be a group with n elements. Let a, b in G. If $\gcd(\pi(a), \pi(b)) = 1$ and $ab = ba$, then $\pi(ab) = \pi(a)\pi(b)$.*

Proof. Exercise. (See the proof of Lemma 150.) □

Remark 507. The assumption "$ab = ba$" cannot be removed. Inside S^3, let $a = (1,3)$ and $b = (1,2)$. Then a, b do not commute. One can see that $ab = (123)$. Thus

$$\pi(a) = \pi(b) = 2, \quad \text{but } \pi(ab) = 3 \neq \pi(a)\pi(b).$$

Definition 508. Let $a_1, \ldots, a_n$ be elements of a group G. We denote by

$$< a_1, \ldots, a_n > \stackrel{\text{def}}{=} \text{ the smallest subgroup of } G \text{ containing all the } a_i\text{'s.}$$

The definition makes sense, because given two subgroups of G containing the a_i's, their intersection is still a subgroup of G containing the a_i's. So there is indeed a unique inclusion-minimal subgroup containing the a_i's. We have the following alternative description:

Lemma 509. *For any group G and for any element $a \in G$,*

$$< a >= \{a^z \text{ such that } z \in \mathbb{Z}\} = \{e, a, a^2, a^3, \dots, a^{\pi(a)-1}\}$$

is a subgroup of G with exactly $\pi(a)$ elements.

Proof. For the second equality, $\supseteq$ is obvious; $\subseteq$ follows from the fact that if $z = q \cdot \pi(a) + r$ (Euclidean division), then $a^z = (a^{\pi(a)})^q \cdot a^r = a^r$. As for the first equality: if a^z and a^w be two elements of $\{a^z \text{ such that } z \in \mathbb{Z}\}$, so is

$$a^z (a^w)^{-1} = a^z a^{-w} = a^{z-w}.$$

Thus by Proposition 451, $\{a^z \text{ such that } z \in \mathbb{Z}\}$ is a subgroup. It contains $a^1 = a$. Also, any subgroup of G contaning a contains all its powers. Hence, $\{a^z \text{ such that } z \in \mathbb{Z}\}$ is the smallest subgroup containing a, which is what we denoted by $< a >$. □

Remark 510. The proof above crucially uses that there is only one generator. In general,

$$< a, b > \supseteq \{a^z\, b^t \text{ such that } z, t \in \mathbb{Z}\}$$

but the set on the left may be larger, because it contains also elements of the type aba or $ab^{-1}a^5a^{-3}b^7$, which we do not know how to rearrange. Of course, this problem would be solved if we knew in advance that a and b *commute* (that is, $ab = ba$): Then we could rewrite $aba = a^2b$ and $ab^{-1}a^5b^{-7}a^{-3}b^5 = a^3b^{-3}$. We leave it as Exercise to prove the following Proposition.

Proposition 511. *Let G be a group G. Let $a_1, \dots, a_n \in G$ be elements any two of which commute, i.e. $a_i a_j = a_j a_i$ for all i, j. Then*

$$< a_1, \dots, a_n >= \{a_1^{z_1}\, a_2^{z_2} \cdots a_n^{z_n} \text{ such that } z_i \in \mathbb{Z}\}.$$

Proposition 512. *Let G be a group with n elements. Then for all $a \in G$, one has $a^n = e$.*

Proof. For each a, the subgroup $< a >$ has cardinality $\pi(a)$ by Lemma 509. By Lagrange's theorem 464 the integer $\pi(a)$ divides n. But then by Lemma 504 one has $a^n = e$. □

In fact, one can slightly improve on the previous result, as follows.

Proposition 513. *Let G be a group with n elements and with largest period λ. Then λ divides n. Moreover, for all $a \in G$ that commute with a largest period element, one has $a^\lambda = e$.*

Proof. (From Lindsay Childs, A concrete introduction to Higher Algebra, page 389) Let g be the element such that $\pi(g) = \lambda$. Since the subgroup $< g >$ has cardinality λ by Lemma 509, by Lagrange's theorem 464 λ divides n. Now let a be an arbitrary element of G that commutes with g. Let $m = \pi(a)$. We know that $m \leq \lambda$. Suppose by contradiction that m does not divide λ. Then there exists some prime p such that p^r is the highest power dividing m, p^s is the highest power of p that divides λ, and $r > s$. Since a has period m, by Proposition 505 the element

$$d \stackrel{\text{def}}{=} a^{\frac{m}{p^r}} \text{ has period } \frac{m}{\gcd(m, \frac{m}{p^r})} = \frac{m}{\frac{m}{p^r}} = p^r.$$

On the other hand, since $\pi(g) = \lambda$, by Proposition 505

$$c \stackrel{\text{def}}{=} g^{p^s} \text{ has period } \frac{\lambda}{\gcd(\lambda, p^s)} = \frac{\lambda}{p^s}.$$

But by definition of r and s, the two integers p^r and $\frac{\lambda}{p^s}$ are coprime. So by Proposition 506 the element cd has order $p^r \cdot \frac{\lambda}{p^s} = p^{r-s} \cdot \lambda$. But this is larger than λ, a contradiction. □

Remark 514. The assumption "that commute with a largest period element" cannot be omitted. For example, in $\mathcal{S}_3$ the largest period is 3, but the smallest integer t such that $a^t = e$ for all $a \in \mathcal{S}_3$ is clearly $t = 6$. Usually in the literature the *exponent* of a group G with n elements is defined to be the smallest integer t such that $a^t = e$ for all $a \in G$. By the previous Proposition, for Abelian groups the exponent is simply the largest period of an element.

How do homomorphisms behave with respect to periods?

Lemma 515. *For any homomorphism $f : G \longrightarrow H$ between groups, $f(e_G) = e_H$.*

Proof. Since $e_G e_G = e_G$, applying f on both sides and using that f is a homomorphism we get

$$f(e_G)f(e_G) = f(e_G) = f(e_G)e_H.$$

So by Cancellation (Proposition 448)

$$f(e_G) = e_H.$$ □

Proposition 516. *For any homomorphism $f : G \longrightarrow H$ between groups, and for each $x \in G$, the period of $f(x)$ divides the period of x.*

Proof. Let t be the period of x. From $x^t = e_G$, we get

$$(f(x))^t = f(x^t) = f(e_G).$$

But by Lemma 515 we know that $f(e_G) = e_L$. So

$$(f(x))^t = e_L.$$

By Lemma 504, this means that the period of $f(x)$ divides t. □

Corollary 517. *If $\gcd(m, n) = 1$, the only homomorphism between $\mathbb{Z}_m$ and $\mathbb{Z}_n$ is the zero homomorphism.*

Proof. Let $f : \mathbb{Z}_m \longrightarrow \mathbb{Z}_n$ be an arbitrary homomorphism and set $y \stackrel{\text{def}}{=} f(\overline{1})$. By Proposition 516, the period of y divides $\pi(1) = m$. By Proposition 512, the period of y divides also n. So, since $\gcd(m, n) = 1$, the period of y must be 1. So $y = 0$. □

Cyclic and finitely generated groups

Definition 518 (Cyclic, Finitely Generated)**.** A group G is called *cyclic* is there is an element $a \in G$ such that $G =< a >$.
More generally, a group G is called *finitely generated* if there are elements $a_1, \ldots a_n$ in G such that $G =< a_1, \ldots, a_n >$.

Example 519. $(\mathbb{Z}, +)$ is cyclic. In fact, $\mathbb{Z} =< 1 >=< -1 >$. Note that with the additive notation, $(-1)^n$ is just $-n$.

Proposition 520. *Every subgroup of $(\mathbb{Z},+)$ is cyclic, of the form $< n >$ for some $n \in \mathbb{N}$. Thus any quotient of $(\mathbb{Z},+)$ is isomorphic either to $(\mathbb{Z},+)$ or to $(\mathbb{Z}_n,+)$, with $n > 0$.*

Proof. Let S be a subgroup of $(\mathbb{Z},+)$. If $S = \{0\}$ then $S =< 0 >$ and the quotient $\mathbb{Z}/S$ is $\mathbb{Z}$ itself. Otherwise, let n be the smallest positive integer in S. Clearly

$$< n > \subseteq S.$$

To see that equality holds, take a generic x is in S. Then the Euclidean division

$$x = q \cdot n + r$$

tells us that $r = x - qn$ is in S, because in $(\mathbb{Z},+)$ this is how we denote the operation $x \cdot (n^q)^{-1}$. Thus if $r > 0$ we have a contradiction with how n was chosen. So $r = 0$, which means that $x = qn$. With the additive notation, this means that x is in $< n >$. Thus $S =< n >$ and the quotient $\mathbb{Z}/S$ is $\mathbb{Z}_n$. □

Non-Example 521. $(\mathbb{Q}^*,\cdot)$ is not cyclic. By contradiction, were two coprime integers (a,b), with $b > 0$, for which

$$\mathbb{Q} =< \frac{a}{b} >= \left\{ \left(\frac{a}{b}\right)^z \text{ such that } z \in \mathbb{Z} \right\}, \tag{54}$$

Let p be a prime number that is neither a factor of a nor of b. Obviously $\frac{1}{p}$ belongs to $\mathbb{Q}$. So by Equation 54 there is some integer z such that

$$\frac{1}{p} = \left(\frac{a}{b}\right)^z.$$

If $z > 0$, clearing denominators we get that $b^z = pa^z$, a contradiction: p does not divide b.
If $z < 0$, clearing denominators we get that $a^{-z} = pb^{-z}$, a contradiction: p does not divide a.
(With a similar idea one can show that $(\mathbb{Q}^*,\cdot)$ is not even finitely generated; see the Exercises at the end of the section.)

Example 522. Any finite group is finitely generated: We can use all its elements as generators. However, $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ is an example of a finite group that is not cyclic. This can be verified directly (G has four elements, so we can check $< a >$ for all $a \in G$) or using Theorem 550 below.

Proposition 523. *Every cyclic group is isomorphic either to $\mathbb{Z}$ or to some $\mathbb{Z}_m$, with $m \in \mathbb{N}$.*

Proof. If $G =< a >$, the map

$$\begin{array}{rccc} f: & (\mathbb{Z},+) & \longrightarrow & G \\ & k & \longmapsto & a^k \end{array}$$

is a surjective group homomorphism. Thus by the First Isomorphism Theorem (Corollary 486), G is isomorphic to the quotient $\mathbb{Z}/\ker f$. Via Proposition 520 we conclude.

□

Example 524. In $\mathbb{Z}_6$, which has six elements, all elements must have finite period (a submultiple of six). Let's check: The element 0 is the identity, so it has period 1. The element 3 has period 2, because $3+3 = 0$. The elements 2 and 4 have period 3, because $2+2+2 = 0$ and $4+4+4 = 0$. The elements 1 and 5 have period 6.

Example 525. In $\mathbb{Z}$, the element 0 is the identity and therefore has period 1. Any other element has infinite period.

Remark 526. There are infinite groups in which all elements have finite period. For example, the polynomial C-ring $\mathbb{Z}_m[x]$ is a group with respect to addition. Any element summed to itself m times yields zero, which is the neutral element. Hence, the period of every element of $\mathbb{Z}_m[x]$ divides m.

The following, beautiful proof is due to McKay[18]:

Theorem 527 (Cauchy)**.** *If the cardinality of a finite group G is a multiple of some prime number p, then some element has period p.*

Proof by McKay. We are going to show something much stronger, namely, that the number of period-p elements is of the form $kp-1$ for some positive integer k. In particular, this number is at least $p-1$, which is positive. Say G has n elements. Look at the set

$$S \stackrel{\text{def}}{=} \{(x_1, x_2, \ldots, x_p) \text{ such that } x_1 \cdot x_2 \cdot \ldots \cdot x_p = e\}.$$

Once we freely choose $x_1, \ldots, x_{p-1}$ in $\{1, \ldots, n\}$, there is a unique x_p such that $x_1 \cdot x_2 \cdot \ldots \cdot x_p = e$; so the set S has cardinality n^{p-1}. Now partition S into equivalence classes, as follows. If all components of a p-tuple are equal, its equivalence class shall consist of only 1 element. If instead we have a $(x_1, \ldots, x_p)$ with $x_i \neq x_j$ for some i, j, then the equivalence class of $(x_1, \ldots, x_p)$ shall consist of the p (different!) elements

$$(x_1, x_2, \ldots, x_{p-1}, x_p),\ (x_2, \ldots, x_{p-1}, x_p, x_1),\ \ldots, (x_p, x_1, x_2, \ldots, x_{p-1}).$$

So let a be the number of equivalence classes with only one member, and b be the number of equivalence classes with p members. Since

$$n^{p-1} = a \cdot 1 + b \cdot p,$$

and n is a multiple of p, we obtain that a is also a multiple of p. But by definition, a counts exactly the elements x such that $(x, x, \ldots, x)$ is in S, which means that $x^p = e$. So we can conclude that the equation $x^p = e$ has a number of solutions in G that is a multiple of p. Write this number as kp. One of the solutions is the neutral element, because $e^p = e$. So there are exactly $kp - 1$ elements y different than the identity, such that $y^p = e$. The period of any such y must divide p, which is a prime number. So the period of any such y is exactly p. □

7.5 *The Second Isomorphism Theorem and Sylow theory

We wish to improve on Cauchy's theorem 527. This section follows the textbook *Algebra* by Goldhaber–Ehrlich.

Definition 528 (Normalizer; Centralizer)**.** Let S be any subset of a group G.

- the *normalizer of S in G* is $N_G(S) \stackrel{\text{def}}{=} \{x \in G \text{ such that } xS = Sx\}$.
- the *centralizer of S in G* is $C_G(S) \stackrel{\text{def}}{=} \{x \in G \text{ such that } xs = sx \text{ for all } s \in S\}$.

The centralizer of G in G is sometimes called "the *center* of G".

Proposition 529. *Let S be any subset of a group G.*

(i) *$C_G(S)$ and $N_G(S)$ are both subgroups of G.*
(ii) *$C_G(S) \subseteq N_G(S)$.*
(iii) *If S consists of a single element s, then normalizer and centralizer of S coincide.*
(iv) *If S is a subgroup of G, then S is a normal subgroup of $N_G(S)$.*

[18] James H. McKay, *Another Proof of Cauchy's Theorem*, American Mathematical Monthly 66 (1959), page 119.

(v) *If S is a subgroup of G, then $N_G(S) = G$ if and only if S is normal in G.*

Proof. Exercise. □

The notion of "normalizer" is useful to give us a second Isomorphism Theorem:

Theorem 530 (Second Isomorphism Theorem)**.** *Let H, K be subgroups of a group G, such that $H \subseteq N_G(K)$. Then:*

- *$HK = \{hk$ such that $h \in H$, $k \in K\}$ is the smallest subgroup containing H and K;*
- *K is a normal subgroup of HK;*
- *$H \cap K$ is a normal subgroup of H, and*

$$HK/K \cong H/H \cap K.$$

Proof.
- For all $h, h' \in H$ and $k, k' \in K$, the elements $hk^{-1}h^{-1}$ and $h'k'h^{-1}$ are in K, because $H \subseteq N_G(K)$; hence $(h'k')(hk)^{-1} = h'k'k^{-1}h^{-1} = \left(h'k'h^{-1}\right)\left(hk^{-1}h^{-1}\right) \in HK$. Thus HK is a subgroup. Clearly $HK \supseteq H$ because any $h \in H$ can be written as $h = he$, with $e \in K$. Symmetrically, $HK \supseteq K$. Any subgroup containing H and K must also contain all the products hk.
- For all h in H and all k, k' in K, since $h^{-1} \in H \subseteq N_G(K)$ we can write $kk'k^{-1}h^{-1}$ as $h^{-1}k''$, for some $k'' \in K$. Thus $(hk)k'(hk)^{-1} = hkk'k^{-1}h^{-1} = hh^{-1}k'' = k'' \in K$.
- For all x in $H \cap K$ and $h \in H$, clearly hxh^{-1} is in H. Moreover, since x is in K and $H \subseteq N_G(K)$, hxh^{-1} is in H. So hxh^{-1} is in $H \cap K$. This shows normality. Now define

$$\begin{array}{rccc} \varphi: & H & \longrightarrow & HK/K \\ & h & \longmapsto & \overline{he}. \end{array}$$

This is clearly a homomorphism. The kernel of this map is

$$\ker\varphi = \{h \in H \text{ such that } he \in K\} = H \cap K.$$

To conclude with the First Homomorphism Theorem, it remains to show that φ is surjective. The generic element of HK/K is of the form $\overline{hk}$, for some $h \in H$ and some $k \in K$. Yet $(hk)^{-1}(he) = k^{-1}h^{-1}h = k \in K$, which proves that $\varphi(h) = \overline{hk}$. □

Remark 531. The assumption $H \subseteq N_G(K)$ in the previous theorem cannot be omitted. For example, inside $\mathcal{S}_3$, consider the subgroups $H = \{e, (12)\}$ and $K = \{e, (23)\}$. Neither of them is contained in the normalizer of the other. The smallest subgroup containing both is $\mathcal{S}_3$ itself, which is larger than the sets HK and KH, none of which is a subgroup. Note also that the assumption $H \subseteq N_G(K)$ would be automatically verified in case K is normal.

We are now going to partition a group G according to a particular relation of equivalence called **conjugation**.

Definition 532. Given two elements a, b of a group G, we say that *a is conjugate to b* if $a = g^{-1}bg$ for some g in G. It is easy to see that this is a relation of equivalence; so we will also say "a and b are conjugates".

Example 533. Let G be a group. The equivalence class of a center element $b \in C_G(G)$ with respect to conjugation has only one element, namely, b itself. In fact, any a in G is conjugate to b if and only if for some g one has $a = g^{-1}bg$; but since $bg = gb$, this implies $a = b$.

Definition 534. Let G be a group. The *index* of a subgroup H of G is the cardinality of the set of left (or right) cosets of H. By Lagrange 's theorem 464, if G is finite $[G : H] = \frac{|G|}{|H|}$.

The next Lemma generalizes Example 533, because if s is in the center of G, then it commutes with all elements of G, so $N_G(\{s\}) = C_G(\{s\}) = G$.

Lemma 535. *Let G be a finite group.*

(i) *For any element s of G, the cardinality of the equivalence class of s with respect to conjugation is $[G : C_G(\{s\})]$.*

(ii) *For any subset S of G, the cardinality of the subsets of G conjugate to S is $[G : N_G(S)]$.*

Proof. The first claim follows from the second one by taking $S = \{s\}$, since centralizer and normalizer of a one-element set are the same. So let us prove the second claim. Set $N \stackrel{\text{def}}{=} N_G(S)$. For any x, y in G, we have $x^{-1}Sx = y^{-1}Sy$ if and only if $Sxy^{-1} = xy^{-1}S$, if and only if xy^{-1} is in N. So $x^{-1}Sx = y^{-1}Sy$ if and only if the right cosets Nx and Ny are equal. Hence, the map that sends $x^{-1}Sx$ to Nx is a one-to-one mapping between the class of subsets conjugate to S, and the right cosets of N. □

Theorem 536 (First Sylow Theorem, 1872)**.** *If the cardinality of a finite group G is a multiple of p^m, where p is some prime number and m is some positive integer, then G has a (not necessarily cyclic, unique, or normal) subgroup with p^m elements.*

Proof. We proceed by induction on the cardinality of G, the claim being clear if $|G| \in \{1, 2, 3\}$. Suppose the theorem holds for all groups with less than n elements, and let us consider a group G with n elements. There are two cases:

- If there exists a subgroup $H \subsetneq G$ such that p does not divide $[G : H]$, then p^m divides H, so by inductive assumption, H has a subgroup S with p^m elements. This S is also a subgroup of G, of course.
- If not, then for all subgroups $H \subsetneq G$ one has that p divides $[G : H]$. Now partition G according to the equivalence relation of conjugation. Let $A \subseteq G$ be a complete set of representatives, without repetitions, of the conjugate classes of G. By Example 533, A contains one copy of the center $C_G(G)$. Let $B \stackrel{\text{def}}{=} A \setminus C_G(G)$. By Lemma 535, part (i),

$$|G| \;=\; |C_G(G)| + \sum_{s \in B} [G : C_G(\{s\})]. \tag{55}$$

But since the centralizers involved in the sum above are proper subgroups, by assumption p divides $[G : C_G(\{s\})]$ for all s in B. So Equation 55 tells us that $|C_G(G)|$ is a multiple of p. By Cauchy's theorem 527, $C_G(G)$ has a (necessarily cyclic) subgroup H with p elements[19]. Since any two elements in $|C_G(G)|$ commute, this H is normal. Thus we may consider the quotient group G/H. If $|G| = p^m k$, obviously

$$[G : H] = \frac{|G|}{|H|} = \frac{p^m k}{p} = p^{m-1}k < |G|.$$

Using the inductive assumption if $m > 1$, or common sense if $m = 1$, we conclude that

$$G/H \text{ has a subgroup } S' \text{ of size } p^{m-1}.$$

Now by the First Isomorphism Theorem, $S' = S/H$ for some subgroup S of G. But then $|S| = p \cdot p^{m-1} = p^m$. □

[19] Cauchy's theorem is much easier to prove for Abelian groups, and here we are invoking it for $|C_G(G)|$, which is indeed Abelian. See the next Section and Corollary 571 for more.

Definition 537. Let G be a finite group. Let p be a prime. If m is the largest integer such that p^m divides $|G|$, any subgroup of G with p^m elements is called a *Sylow p-subgroup of G*.

The First Sylow Theorem guarantees the existence of a Sylow p-subgroup for any p. (If p does not divide $|G|$, the corresponding subgroup is just the identity.) Is it unique? Sometimes!

Example 538. S^3, which has six elements, has one subgroup with three elements, A_3, which is therefore "the" Sylow 3-subgroup. In fact, the two elements of period 3 within S^3 are the cycles (123) and (132), which are inverses of one another and thus belong to the same subgroup. The other nontrivial elements of S^3, namely, (12), (13) and (23), have period 2; so each of them generates a distinct size-two subgroup. So S^3 has three different Sylow 2-subgroups!

A natural follow-up question is, how many Sylow p-subgroups are there in a given group? In the same 1872 paper where he proved the "First Sylow Theorem", Ludwig Sylow, who was a high school teacher in Fredrikshald (now Halden), Norway, provided an interesting answer. To explain this answer we need some technicalities, like a slightly stronger form of Lemma 535.

Definition 539. Given two elements a, b of a group G, and given a subgroup P of G, we say that *a is P-conjugate to b* if $a = g^{-1}bg$ for some g in P. It is easy to see that this is a relation of equivalence; so we will also say "a and b are P-conjugates".

Example 540. If a is P-conjugate to some b of the center of G, then $a = g^{-1}bg = g^{-1}gb = b$. But the same is true if a is P-conjugate to some b in $C_G(P)$, which contains $C_G(G)$. So also the equivalence class of elements of $C_G(P)$ with respect to P-conjugation have only one element.

Lemma 541. *Let G be a finite group. Let P be a subgroup.*
For any subset S of G, the cardinality of the subsets of G P-conjugate to S is $[P : (P \cap N_G(S))]$.

Proof. Set $N' \stackrel{\text{def}}{=} P \cap N_G(S)$. For any x, y in P, we have $x^{-1}Sx = y^{-1}Sy$ if and only if $Sxy^{-1} = xy^{-1}S$, if and only if xy^{-1} is in N'. So $x^{-1}Sx = y^{-1}Sy$ if and only if the right cosets $N'x$ and $N'y$ are equal. Hence, the map that sends $x^{-1}Sx$ to $N'x$ is a one-to-one mapping between the class of subsets conjugate to S, and the right cosets of N'. □

Lemma 542. *Let G be a finite group. Let p be a prime that divides $|G|$. Let P be any subgroup of G such that $|P|$ is a power of p. If S is a Sylow p-subgroup, then $P \cap N_G(S) = P \cap S$.*

Proof. We have seen in Proposition 529 that S is a normal subgroup of $N_G(S)$, so the inclusion $P \cap S \subseteq P \cap N_G(S)$ is trivial. Set $Q \stackrel{\text{def}}{=} P \cap N_G(S)$. Then by the Second Isomorphism Theorem 530, since $Q \subseteq N_G(S)$ we have that QS is a subgroup, S is normal in QS, $Q \cap S = P \cap S$ is normal in Q, and

$$QS/S \cong Q/P \cap S. \tag{56}$$

Since we want to show that $Q/P \cap S$ is trivial, Equation (56) tells us that it suffices to show that QS/S is trivial; or in other words, that $QS = S$. Indeed, from Equation (56) passing to the cardinalities we immediately obtain

$$\frac{|QS|}{|S|} = \frac{|Q|}{|P \cap S|}.$$

This implies that the cardinality of QS is a power of p: In fact, by Lagrange's theorem, since by assumption the cardinality of P is some power of p, also the cardinality of its subgroups Q and $P \cap S$ must be a power of p; and by definition of Sylow p-subgroup, $|S|$ is also a power of p. But then so is $|QS|$. Hence QS is a subgroup of G containing S, and its cardinality is a power of p. By definition of Sylow p-subgroup, we must have $QS = S$, as desired. □

Theorem 543 (Second and Third Sylow Theorem, 1872)**.** *Let G be a group, p a prime, S a Sylow p-subgroup of G. Then:*

(i) *Every subgroup of G whose cardinality is a power of p is contained in a conjugate of S.*
(ii) *All Sylow p-subgroups are conjugate.*
(iii) *There are exactly $[G : N_G(S)]$ distinct Sylow p-subgroups of G.*
(iv) $[G : N_G(S)] \equiv 1 \mod p$.

Proof. Note first that by Lemma 535, the class $\mathfrak{C}$ of all subgroups of G that are conjugate to S has exactly $[G : N_G(S)]$ elements.

(i) Let P be an arbitrary subgroup of G such that $|P|$ is a power of p. Let us use the equivalence relation of P-conjugation to partition $\mathfrak{C}$ into nonempty classes of P-conjugate subgroups:

$$\mathfrak{C} = \mathfrak{C}_1 \cup \mathfrak{C}_2 \cup \ldots \mathfrak{C}_k, \quad k \leq [G : N_G(S)].$$

Now for each $i = 1, \ldots, k$, let S_i be a subgroup from the class $\mathfrak{C}_i$. By Lemma 541, the number of subgroups in $\mathfrak{C}_i$ is equal to $[P : P \cap N_G(S_i)]$. But by Lemma 542, this quantity equals $[P : P \cap S_i]$. Hence

$$[G : N_G(S)] = [P : P \cap S_1] + [P : P \cap S_2] + \ldots + [P : P \cap S_k]. \tag{57}$$

Since $|P|$ is a power of p, every summand on the right is a power of p, possibly $p^0 = 1$. On the other hand, by definition of Sylow p-subgroups, $[G : S]$ cannot be a multiple of p; otherwise if $|S| = p^a$, by the First Sylow Theorem there would be a subgroup of size p^{a+1} in G containing S, a contradiction. But since $S \subsetneq N_G(S)$ (cf. Proposition 529), also $[G : N_G(S)]$ cannot be a multiple of p (or else a group with size a power of p would strictly contain $N_G(S)$ and thus S). So in Equation (57) at least one of the summands must be 1. That is, there exists an $i \in \{1, \ldots, k\}$ such that $[P : P \cap S_i] = 1$. This happens if and only if $P = P \cap S_i$, if and only if $P \subseteq S_i$.

(ii) Let S and S' be two Sylow p-subgroups. By part (i), S' is contained and thus equal to a conjugate of S.

(iii) Let S be a Sylow p-subgroup. By part (ii), any other Sylow p-subgroup is conjugate to S. But the converse is also true, namely, any subgroup conjugate to S is clearly a Sylow p-subgroup. Thus the number of Sylow p-subgroups is identical to the number of subgroups of G conjugate to S, which is $[G : N_G(S)]$.

(iv) Let S be a Sylow p-subgroup. In the identity (57), for any P we can choose the S_i so that exactly one of them (say, S_1) is equal to S. In the particular case $P = S$, this tells us that $[P : P \cap S_1] = [S : S \cap S] = 1$. In contrast, $[P : P \cap S_i] = [S : S \cap S_i] > 1$ for $i > 1$. This implies that $[S : S \cap S_i]$ is a multiple of p, since the cardinality of S is a power of p. So the identity (57) for $P = S$ boils down to

$$[G : N_G(S)] = 1 + ([S : S \cap S_2] + \ldots + [S : S \cap S_k]), \tag{58}$$

where the part between curve brackets is a multiple of p. □

Corollary 544. *If S is a Sylow p-subgroup of G, then S is normal in G if and only if S the only Sylow p-subgroup of G.*

Proposition 545. *Let $p < q$ be two primes. If p does not divide $q - 1$, every group with pq elements is isomorphic to $\mathbb{Z}_p \times \mathbb{Z}_q$.*

Proof. Let n_p and n_q be the number of Sylow p-subgroups and of Sylow q-subgroups, respectively. By the previous theorem, $n_p \equiv 1 \mod p$ and $n_q \equiv 1 \mod q$. Now, if H is a Sylow p-subgroup,

$[G : H]$ and $[G : N_G(H)]$ cannot be multiples of p, so they are either 1 or q. But $n_p = [G : N_G(H)] = q$ leads to $(q \equiv 1 \mod p)$, which contradicts the assumption. So $n_p = 1$. By the Corollary above, this tells us that H is normal in G, with $|H| = p$. On the other hand, if K is a Sylow q-subgroup, $[G : K]$ and $[G : N_G(K)]$ cannot be multiples of q, so they are either 1 or p. But $n_q = [G : N_G(K)] = p$ means that $(p \equiv 1 \mod q)$, which is impossible since $p < q$. So $n_q = 1$. Thus K is normal in G, with $|K| = q$. By Lagrange's theorem, $H \cap K$ must be $\{e\}$. Thus $HK \cong H \times K \cong \mathbb{Z}_p \times \mathbb{Z}_q$. □

Remark 546. The previous Proposition is an "if and only if": For any pair of primes $p < q$, if p divides $q - 1$, one can see that there exists an integer n such that $n \neq 1 \mod p$, yet $nq = 1 \mod q$; and one can prove that the group

$$G = \langle x, y \text{ such that } x^p = 1 = y^q, yx = x^n y \rangle$$

has exactly pq elements and is not Abelian. However, G and $\mathbb{Z}_p \times \mathbb{Z}_q$ are the only two groups with pq elements.

7.6 Abelian Groups, Chinese remainders, and the totient function

Definition 547 (Abelian)**.** A group G is called *abelian* if its operation satisfies the commutative property, that is, for each x, y in G one has $xy = yx$.

Proposition 548. *Every cyclic group is abelian. The converse is false.*

Proof. If $G =< a >$, for any z, w integers one has

$$a^z a^w = a^{z+w} = a^{w+z} = a^w a^z.$$

As for the converse, we have seen that $(\mathbb{Q}*, \cdot)$ is not cyclic in Example 521. □

Remark 549. A finitely generated group might not be abelian. In fact, many finite groups are not abelian, like the permutation group $\mathcal{S}_3 = \{f : \{1, 2, 3\} \longrightarrow \{1, 2, 3\} \text{ bijective}\}$. We will discuss this example in the next Section.

It is easy to see that products of abelian groups are abelian. Products of cyclic groups are sometimes cyclic, and sometimes not:

Theorem 550. *Let A, B be two cyclic groups with a and b elements, respectively. Then:*
- *If* $\gcd(a, b) = 1$*, then $A \times B$ is cyclic.*
- *If* $\gcd(a, b) = d > 1$*, then* every *element of $A \times B$ has period $\leq \frac{ab}{d}$, so $A \times B$ is not cyclic.*

Proof. By definition of product group, $(x, y)^t = (x^t, y^t)$. So using Lemma 504,

$$(x, y)^t = (e_A, e_B) \Longleftrightarrow \begin{cases} x^t = e_A \\ y^t = e_B \end{cases} \Longleftrightarrow \begin{cases} t \text{ is a multiple of } \pi(x) \\ t \text{ is a multiple of } \pi(y). \end{cases} \tag{59}$$

- If $\gcd(a, b) = 1$, choose (x, y) in $A \times B$ such that $A =< x >$ and $B =< y >$. Clearly $\pi(x) = a$ and $\pi(y) = b$. Since they have no common factor, any number that is multiple of both a, b must also be a multiple of ab. Hence the coimplication 59 becomes

  $$(x, y)^t = (e_A, e_B) \Longleftrightarrow t \text{ is a multiple of } ab.$$

  In particular, the smallest t for which $(x, y)^t = (e_A, e_B)$ is ab.

- If $\gcd(a,b) = d > 1$, set $a' \stackrel{\text{def}}{=} \frac{a}{d}$ and $b' \stackrel{\text{def}}{=} \frac{b}{d}$. For any (x,y) in $A \times B$, we know that $x^a = e_A$ and $y^b = e_B$ by definition of period. In particular, $x^{ab'} = e_A$, because ab' is a multiple of a; and $y^{ab'} = e_B$, because $ab' = a'db' = a'b$ is a multiple of b. Hence,

$$(x,y)^{ab'} = (e_A, e_B).$$

  So the period of (x,y) is at most ab', which is smaller than ab. Since this holds for all (x,y), no element of $A \times B$ has period ab. Hence $A \times B$ is not cyclic. □

Corollary 551. *$\mathbb{Z}_a \times \mathbb{Z}_b$ is isomorphic to $\mathbb{Z}_{ab}$ $\iff$ $\gcd(a,b) = 1$.*

Proof. Straightforward from Lemma 550 and Proposition 523. □

In fact, it is very easy to write down the isomorphism explicitly. We do this in the slightly more general case of n cyclic factors, instead of two.

Lemma 552. *Let $m_1, \ldots, m_n$ be positive integers such that $\gcd(m_i, m_j) = 1$ for all $i \neq j$. Set $m \stackrel{\text{def}}{=} m_1 m_2 \cdots m_n$. For any integer x,*

$$m \text{ divides } x \iff \text{each } m_i \text{ divides } x.$$

Proof. The direction '$\Rightarrow$' is trivial, so let us focus on '$\Leftarrow$'. Let p be a prime that divides m. By Euclid's Lemma, p divides *at least one* of the m_i's; but by the pairwise-coprime assumption, p divides *at most one* of the m_i's. So it divides exactly one of the m_i's. So if

$$m = p_1^{a_1} p_2^{a_2} p_k^{a_k}$$

then each m_i is either of the form $m_i = p_j^{a_j}$, for some j, or (up to reordering the p_j's) of the form

$$m_i = p_j^{a_j} p_{j+1}^{a_{j+1}} \cdots p_{j+h}^{a_{j+h}} \text{ for some } j \in \{1, \ldots, k\},\ h \in \{1, \ldots, k-j\}.$$

So if an integer x is a multiple of all m_i's, it means that the exponent of each p_j in the factorization of x is a_j or larger. So m divides x. □

Theorem 553 (Chinese Remainder Theorem)**.** *Let $m_1, \ldots, m_n$ be positive integers such that $\gcd(m_i, m_j) = 1$ for all $i \neq j$. Set $m \stackrel{\text{def}}{=} m_1 m_2 \cdots m_n$. For any z in $\mathbb{Z}$, let us denote by $[z]$ its equivalence class modulo m, and by $[z]_i$ its equivalence class modulo m_i. The following trivial projection is a C-ring isomorphism:*

$$\begin{array}{rccccc} \pi: & \mathbb{Z}_m & \longrightarrow & \mathbb{Z}_{m_1} \times & \ldots \times & \mathbb{Z}_{m_n} \\ & [z] & \longmapsto & ([z]_1, & \ldots, & [z]_n). \end{array}$$

Moreover, $\gcd([z], m) = 1$ if and only if $\gcd([z]_i, m_i) = 1$ for all i.

Proof. The finite sets $\mathbb{Z}_m$ and $\mathbb{Z}_{m_1} \times \ldots \times \mathbb{Z}_{m_n}$ have the same number of elements, so any injective map between them is automatically surjective. But our projection π is well-defined and injective, because

$$[z] = [y] \stackrel{\text{def}}{\iff} m \text{ divides } z - y \stackrel{!}{\iff} \text{each } m_i \text{ divides } z - y \stackrel{\text{def}}{\iff} [z]_i = [y]_i \text{ for all } i.$$

(The $\Leftarrow$ direction of the $\stackrel{!}{\iff}$ above is given by the previous Lemma.) So the map π is automatically also surjective. It is easy to see that the map π is a C-ring homomorphism.

For the second claim, recall that $[z]_i$ is obtained by taking the remainder of the Euclidean division of $[z]$ by m_i, so for each i we have

$$[z] = q_i m_i + [z]_i.$$

If for some i we have $\gcd([z]_i, m_i) > 1$, then some prime p divides both $[z]_i$ and m_i; so p obviously divides $[z]$ and m; so $\gcd([z], m) > 1$. Conversely, if $\gcd([z], m) > 1$, then some prime p divides both $[z]$ and m. But then by Euclid's Lemma p divides at least one of the m_i's. Without loss, suppose it divides m_1. Then by the expression above, p will divide also $[z]_1 = [z] - q_1 m_1$, so $\gcd)([z]_1, m_1) > 1$. □

Corollary 554 (also cited as 'Chinese Remainder Theorem'; Sun-Tzu, 3rd Century AD)**.** *Let $m_1, \ldots, m_n$ be positive integers such that $\gcd(m_i, m_j) = 1$ for all $i \neq j$. Set $m \stackrel{\text{def}}{=} m_1 m_2 \cdots m_n$. For any $a_1, \ldots, a_n$ in $\mathbb{N}$, the system*

$$\begin{cases} x \equiv a_1 \mod m_1 \\ x \equiv a_2 \mod m_2 \\ \quad \vdots \\ x \equiv a_n \mod m_n \end{cases} \tag{60}$$

admits a unique solution x_0 in $\{0, \ldots, m-1\}$. Moreover, any further integer solution is congruent to such x_0 modulo m.

Proof. By Theorem 553, for any $a_1, \ldots, a_n$ in $\mathbb{N}$ there is a unique element $[z]$ in $\mathbb{Z}_m$ such that

$$\pi([z]) = ([a_1]_1, \ \ldots, \ [a_n]_n).$$ □

Example 555. "There are certain things whose number is unknown. If we count them by threes, we have two left over; by fives, we have three left over; and by sevens, two are left over. How many things are there?" (Sun-Tzu). To re-phrase with our lingo: Is there a number at the same time congruent to 2 mod 3, congruent to 3 mod 5, and congruent to 2 mod 7? There certainly is, because $3, 5, 7$ are coprime. And once you find any such number $[z]$ in $\mathbb{Z}_{105}$, clearly this corresponds to infinitely many in $\mathbb{Z}$, by repeatedly adding 105. So how do you concretely find the first solution? Well, I didn't tell you. Maybe you can find Sun-tzu's trick. Good luck!

The Chinese remainder theorem has another important consequence:

Definition 556. For any integer $m \geq 2$, the *totient function* $\phi(m)$ counts the positive integers coprime with m (or equivalently, the invertible elements in the C-ring $\mathbb{Z}_m$).

Lemma 557. *If $m_1, m_2 \geq 2$ are coprime integers, then $\phi(m_1 m_2) = \phi(m_1) \cdot \phi(m_2)$.*

Proof. Set $m \stackrel{\text{def}}{=} m_1 m_2$. The second part of Theorem 553 tells us that the natural projection π restricts to a bijection

$$\{ \text{ invertible in } \mathbb{Z}_m \} \ \cong \ \{ \text{ invertible in } \mathbb{Z}_{m_1} \} \times \{ \text{ invertible in } \mathbb{Z}_{m_2} \}.$$

The sizes of these three sets are precisely $\phi(m)$, $\phi(m_1)$, and $\phi(m_2)$. □

Lemma 558. *If $m = p^a$ is a prime power, then $\phi(m) = p^a(1 - \frac{1}{p})$.*

Proof. This one's easy. If you write down the numbers from 1 to p^a, those *not* coprime with p^a are simply the multiples of p, and there are p^{a-1} such multiples. So the remaining $p^a - p^{a-1}$ numbers are those coprime with p^a. We just rewrote $p^a - p^{a-1}$ as $p^a(1 - \frac{1}{p})$. □

Theorem 559 (Euler)**.** *For any integer $n \geq 2$,*

$$\phi(n) \;=\; n \cdot \prod_{p \text{ prime divisor of } n} \left(1 - \frac{1}{p}\right).$$

Proof. Suppose n factors as

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}.$$

Setting $m_i \stackrel{\text{def}}{=} p_i^{a_i}$ for all $i = 1, 2, \ldots, k$, and applying Lemmas 557 and 558, we obtain

$$\phi(n) \;= \phi(m_1) \cdots \phi(m_k) = p_1^{a_1}\left(1 - \frac{1}{p_1}\right) \cdots p_k^{a_k}\left(1 - \frac{1}{p_k}\right) = n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_k}\right). \quad \square$$

Example 560. Since $200 = 2^3 \cdot 5^2$, the invertible elements in $\mathbb{Z}_{200}$ are exactly

$$200\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{5}\right) = 200 \cdot \frac{1}{2} \cdot \frac{4}{5} = 80.$$

Instead, since $210 = 2 \cdot 3 \cdot 5 \cdot 7$, the invertible elements in $\mathbb{Z}_{210}$ are

$$210\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{7}\right) = 210 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} = 48.$$

Corollary 561. *For any integer n, $\sqrt{\frac{n}{2}} \leq \phi(n) \leq n - 1$.*

Proof. The upper bound is obviously attained when n is a prime number, and all elements of $\mathbb{Z}_n$ except 0 are invertible. To show the lower bound[20], note that by Lemmas 557 and 558 the quantity $\phi(n)$ is the product of factors of the type $p^a(1 - \frac{1}{p})$, which can be rewritten as $p^{a-1}(p-1)$. We need to bound each one of these factors from below. We claim that *apart from the exceptional case when $p = 2$ and $a = 1$,*

$$p^{a-1}(p-1) \geq (\sqrt{p})^a. \tag{61}$$

In fact, for $a \geq 2$, inequality 61 is obvious: since $a - 1 \geq \frac{a}{2}$, we have $p^{a-1}(p-1) \geq p^{a-1} \geq (\sqrt{p})^a$. If instead $a = 1$, inequality 61 simplifies to $p - 1 \geq \sqrt{p}$. Using elementary calculus, you can check that $p-1)^2 > p$ if and only if $p \in (-\infty, \frac{3-\sqrt{5}}{2}) \cup (\frac{3+\sqrt{5}}{2}, \infty)$. So for $p = 2$ the inequality is false, but it becomes true for every other prime $p \geq 3 > \frac{3+\sqrt{5}}{2}$.

Next, we claim that *unless n is twice an odd number*, the stronger bound $\phi(n) \geq \sqrt{n}$ holds. From the claim the conclusion follows immediately, because if $n = 2n'$ with n' odd, in particular n' is not twice an odd number, so we already know that $\phi(n') \geq \sqrt{n'}$: hence

$$\phi(n) = \phi(2)\phi(n') = \phi(n') \geq \sqrt{n'} = \sqrt{\frac{n}{2}}.$$

So let us prove the claim. Here is the trick: We do know that every factor of the type $p^{a-1}(p-1)$ in $\phi(n)$ will be at least $(\sqrt{p})^a$. In fact, the "exceptional case" where $p = 2$ and $a = 1$ in the factorization of n, can only occur if n is twice an odd number! So we conclude

$$\phi(n) = p_1^{a_1-1}(p_1 - 1) \cdots p_k^{a_k-1}(p_k - 1) \geq \sqrt{p_1}^{a_1} \cdots \sqrt{p_k}^{a_k} = \sqrt{p_1^{a_1} \cdots p_k^{a_k}} = \sqrt{n}. \quad \square$$

[20] Proof by F. Nicolas, *A simple, polynomial-time algorithm for the matrix torsion problem*, arXiv:0806.2068.

Deeper thoughts 562. With much more effort, the lower bound above can be improved a lot: There are lower bounds proportional to $\frac{n}{\log\log n}$. Intuitively, in proportion to n, the lower totients occur when n is a product of distinct primes, and the lowest occur if you take the product of the first t distinct primes. (Compare Example 560). So if p_k is the k-th prime in increasing order, figuring out the precise lower bound for $\phi(n)$ with respect to n has to do with the growth of the expression

$$\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{4}{5}\cdot\frac{6}{7}\cdots\frac{p_t-1}{p_t},$$

and with advanced tools, this can be estimated.
There are many open problems on the totient function. In 1922, Carmichael conjectured that there is no number m with an *exclusive* totient: that is, a number m such that for all $n \neq m$ one has $\phi(n) \neq \phi(m)$. In 1932, Lehmer conjectured that for no composite number n, $\phi(n)$ divides $n-1$. Since for prime numbers $\phi(p)$ does divide $p-1$ (they are equal!), Lehmer's conjecture can be rephrased as "$\phi(n)$ divides $n-1$ if and only if p is prime".

We conclude with the main motivation for studying the totient, namely, the following extension of Fermat's Little Theorem (Theorem 80):

Theorem 563 ("Gauss' Theorem")**.** *If* $\gcd(a,m)=1$*, then* $a^{\phi(m)}=1$*.*

Proof. Since the set of invertible integers of $\mathbb{Z}_m$ has $\phi(m)$ elements, and it forms a multiplicative group with 1 as neutral element, by Proposition 512 $a^{\phi(m)}=1$ for all a in such group. □

For example, $\phi(10)=4$. You can check that for each $a\in\{1,3,7,9\}$, one has $a^4=1$ in $\mathbb{Z}_{10}$.

Remark 564. The Cartesian product of sets is "associative" and "commutative" up to isomorphism. By this we mean that there is an obvious isomorphism

$$\begin{array}{rccc} f: & G\times H & \longrightarrow & H\times G \\ & (g,h) & \longmapsto & (h,g) \end{array}$$

and there is another obvious isomorphism

$$\begin{array}{rccc} f: & G\times(H\times K) & \longrightarrow & (G\times H)\times K \\ & (g,(h,k)) & \longmapsto & ((g,h),k) \end{array}$$

For this reason, when we write down Cartesian products of several groups, we usually omit brackets: We will simply write $G\times H\times K$ instead of $(G\times H)\times K$ or $G\times(H\times K)$.

Example 565. The two 60-element groups $\mathbb{Z}_2\times\mathbb{Z}_{30}$ and $\mathbb{Z}_6\times\mathbb{Z}_{10}$ are not cyclic. Using Corollary 551 (and the notation of Remark 564) we can break them further:

$$\begin{array}{ll} \mathbb{Z}_2\times\mathbb{Z}_{30} & \cong \mathbb{Z}_2\times(\mathbb{Z}_3\times\mathbb{Z}_{10})\cong\mathbb{Z}_2\times\mathbb{Z}_3\times\mathbb{Z}_2\times\mathbb{Z}_5, \\ \mathbb{Z}_6\times\mathbb{Z}_{10} & \cong(\mathbb{Z}_2\times\mathbb{Z}_3)\times(\mathbb{Z}_2\times\mathbb{Z}_5). \end{array}$$

Hence, they are both isomorphic to $\mathbb{Z}_2\times\mathbb{Z}_2\times\mathbb{Z}_3\times\mathbb{Z}_5$.

In contrast, the groups $\mathbb{Z}_3\times\mathbb{Z}_{20}$, $\mathbb{Z}_4\times\mathbb{Z}_{15}$, and $\mathbb{Z}_5\times\mathbb{Z}_{12}$, are all cyclic and isomorphic to $\mathbb{Z}_{60}$. So in some sense they are all the same group. We usually prefer to express this group by separating the different prime powers, which can be done by Corollary 551, as above:

$$\mathbb{Z}_3\times\mathbb{Z}_{20}\cong\mathbb{Z}_3\times\mathbb{Z}_4\times\mathbb{Z}_5.$$

Note that you are not allowed to split $\mathbb{Z}_4$ further, because $\mathbb{Z}_4$ is *not* $\mathbb{Z}_2\times\mathbb{Z}_2$, by Corollary 551.

A remarkable structural result for finite Abelian groups, is that they are all of this type – namely, products of cyclic groups. To see this, we make an observation using some elementary Linear Algebra:

Lemma 566. *Let $G =< g_1, \ldots, g_n >$ be an Abelian group. Let $a_1, \ldots, a_n$ be any n integers with $\gcd(a_1, \ldots, a_n) = 1$. Then there is another set of n generators for G that includes the element $g_1^{a_1} g_2^{a_2} \cdots g_n^{a_n}$.*

Proof. By Proposition 51, there is an integer matrix A of determinant 1 whose first row is $a_1, \ldots, a_n$. Let $a_{i,j}$ be the element in the i-th row, j-th column of A. Define

$$x_i \stackrel{\text{def}}{=} g_1^{a_{i,1}} \cdot g_2^{a_{i,2}} \cdots g_n^{a_{i,n}}. \tag{62}$$

Since $g_1, \ldots, g_n$ are in G, every x_i is obviously in G. By construction,

$$x_1 = g_1^{a_{1,1}} \cdot g_2^{a_{1,2}} \cdots g_n^{a_{1,n}} = g_1^{a_1} g_2^{a_2} \cdots g_n^{a_n}.$$

So all we need to show is that $G =< x_1, \ldots, x_n >$.

To see this, consider the matrix $B \stackrel{\text{def}}{=} A^{-1}$. Since A has determinant 1, B has also integer entries (by the adjoint formula for calculating the inverse matrix). Let $b_{i,j}$ be the element in the i-th row, j-th column of B. Using Equation 62, we get

$$\begin{aligned}
&x_1^{b_{j,1}} \cdot (x_2)^{b_{j,2}} \cdots x_n^{b_{j,n}} = \\
&= \left(g_1^{b_{j,1}a_{1,1}} \cdots g_n^{b_{j,1}a_{1,n}}\right) \cdot \left(g_1^{b_{j,2}a_{2,1}} \cdots g_n^{b_{j,2}a_{2,n}}\right) \cdots \left(g_1^{b_{j,n}a_{n,1}} \cdots g_n^{b_{j,n}a_{n,n}}\right) = \\
&= (g_1)^{\sum_{i=1}^n b_{j,i}a_{i,1}} \cdot (g_2)^{\sum_{i=1}^n b_{j,i}a_{i,2}} \cdot \ldots \cdot (g_n)^{\sum_{i=1}^n b_{j,i}a_{i,n}} = \\
&\stackrel{!}{=} (g_1)^0 \cdot (g_2)^0 \cdot \cdots \cdot (g_{j-1})^0 \cdot (g_j)^1 \cdot (g_{j+1})^0 \cdots (g_n)^0 = g_j,
\end{aligned}$$

where the marked equality is due to the fact that $\sum_{i=1}^n b_{j,i}a_{i,k}$ is the (j,k)-entry of the matrix BA, which is the identity matrix; but the (j,k)-entry of the identity matrix is always 0, unless $k = j$, in which case it is equal to 1. So in conclusion, for every $j \in \{1, \ldots, n\}$, we have

$$g_j = x_1^{b_{j,1}} \cdots x_n^{b_{j,n}}. \tag{63}$$

Equation 63 tells us that every g_j is in $< x_1, \ldots, x_n >$. So

$$G = < g_1, \ldots, g_n > \subseteq < x_1, \ldots, x_n > \subseteq G,$$

which implies $G =< x_1, \ldots, x_n >$. □

Theorem 567 (Smith 1861, Kronecker 1870)**.** *Every finite Abelian group is isomorphic to the product of cyclic groups.*

Proof by E. Schenkman. The idea is to proceed by induction on the *smallest number of generators* n of G. If $n = 1$, then G is cyclic, and we are done. Suppose now $n \geq 2$. Let g_1 be an element of smallest period among those elements that form a generating set of n elements for G. So to recap our assumptions:

- there are elements $g_2, \ldots, g_n$ so that $G =< g_1, \ldots, g_n >$;
- any subset $X \subset G$ with less than n elements cannot be a generating set for G;
- no element x with $\pi(x) < \pi(g_1)$ can be part of a size-n generating set for G.

Set $H \stackrel{\text{def}}{=} < g_1 >$ and $K \stackrel{\text{def}}{=} < g_2, \ldots, g_n >$. Being generated by less than n elements, K is by inductive assumption a product of cyclic groups. H is clearly cyclic. What we want to show is that $G \cong H \times K$. But via Proposition 489, all we need to show is that $H \cap K = (e)$. (The normality of H, K is automatic, because G is Abelian; the smallest subgroup containing both H and K is $< g_1, \ldots, g_n >= G$.) So let us prove that $H \cap K \subseteq (e)$, the other inclusion being obvious. By contradiction, suppose there exists a $z \neq e$ inside $H \cap K$. Since $z \in H =< g_1 >$, we can write $z = g_1^{a_1}$ for some $a_1 \in \{1, \ldots, k-1\}$. Also, $z \in K$, so $z = g_2^{a_2} g_3^{a_3} \cdot g_n^{a_n}$ for some $a_2, \ldots, a_n \in \mathbb{N}$. Set $d \stackrel{\text{def}}{=} \gcd(a_1, a_2, \ldots, a_n)$. Because d is the greatest common divisor of the a_i's, the integers $-\frac{a_1}{d}$, $\frac{a_2}{d}$, $\ldots$, $\frac{a_n}{d}$ have gcd equal to 1. By Lemma 566, we can find another generating set of size n for G that includes the element

$$x \stackrel{\text{def}}{=} g_1^{-\frac{a_1}{d}} g_2^{\frac{a_2}{d}} \cdot \cdots \cdot g_n^{\frac{a_n}{d}}.$$

By construction,

$$x^d = g_1^{-a_1} g_2^{a_2} \cdot \cdots \cdot g_n^{a_n} = (g_1^{a_1})^{-1} \cdot (g_2^{a_2} \cdot \cdots \cdot g_n^{a_n}) = z^{-1} \cdot z = e.$$

So $\pi(x)$ divides d. In turn, d divides a_1, which was smaller than k. This implies that

$$\pi(x) \leq d \leq a_1 < k = \pi(g_1),$$

a contradiction with how g_1 was chosen. □

Corollary 568. *For every finite Abelian group G, there exist natural numbers h, $m_1, \ldots, m_h$ and prime numbers $p_1 \leq \ldots \leq p_h$ (not necessarily distinct), such that G can be decomposed as*

$$G \cong \mathbb{Z}_{p_1^{m_1}} \times \mathbb{Z}_{p_2^{m_2}} \times \ldots \times \mathbb{Z}_{p_h^{m_h}}.$$

Proof. Using Corollary 551 (and the notation of Remark 564), we have seen that it is possible to break up every cyclic group until it is the product of cyclic groups *whose sizes are prime powers* (not necessarily distinct). Compare Example 565. □

Deeper thoughts 569. One may wonder if the multiplicative group of invertible elements in Z_m is cyclic when m is not prime. The answer is not trivial: Gauss proved that it is cyclic whenever m is either 2, or 4, or the power of an *odd* prime, or twice a prime power of an odd prime. We may check that in particular $U_8 = \{1, 3, 5, 7\}$ is not cyclic: The square of every element is 1. So $(U_8, \cdot)$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Exercise for you: Show that $(U_{12}, \cdot)$ is also isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Deeper thoughts 570. The multiplicative group of invertible elements in $\mathbb{Z}_p$ is cyclic, but our proof leaves no insight on *who the generator is*. Finding the "primitive root", that is, the generator is in general a very difficult problem. More generally, for U_m (whether m is prime or not) it is hard to find a set of generators of smallest cardinality. A table of smallest generating sets of U_m for $m \in \{2, 3, \ldots, 128\}$ can be found at the link
https://en.wikipedia.org/wiki/Multiplicative_group_of_integers_modulo_n

Corollary 571 (Converse of Lagrange for Abelian groups)**.** *Let G be an Abelian group with n elements. If d is a divisor of n, then G has a subgroup with exactly d elements.*

Proof. By Corollary 568,

$$G \cong \mathbb{Z}_{p_1^{m_1}} \times \mathbb{Z}_{p_2^{m_2}} \times \ldots \times \mathbb{Z}_{p_h^{m_h}}, \text{ with } n = p_1^{m_1} p_2^{m_2} \cdots p_h^{m_h}.$$

Since d divides n, by the Unique Decomposition theorem of integers it will decompose as

$$d = p_1^{d_1} p_2^{d_2} \cdots p_h^{d_h}, \quad \text{with } d_i \leq m_i \text{ for each } i.$$

Now, in $\mathbb{Z}_{p_1^{m_1}}$ there is clearly an element of period $p_1^{d_1}$, namely, the element

$$a_1 \stackrel{\text{def}}{=} p_1^{m_1 - d_1}.$$

(In fact, by construction $p_1^{d_1} \cdot a_1 = p_1^{d_1} p_1^{m_1 - d_1} = p_1^{m_1} \equiv 0$.) In particular, the subgroup

$$A_1 \stackrel{\text{def}}{=} < p_1^{m_1 - d_1} >$$

has exactly $p_1^{d_1}$ elements. Similarly, inside $\mathbb{Z}_{p_i^{m_i}}$ the subgroup $A_i \stackrel{\text{def}}{=} < p_i^{m_i - d_i} >$ has exactly $p_i^{d_i}$ elements. It follows that the subgroup we are looking for is

$$H \stackrel{\text{def}}{=} A_1 \times A_2 \times \ldots \times A_h. \qquad \square$$

Theorems 567 allow us to classify all possible finite Abelian groups. For example, Example 565 classifies all Abelian groups with 60 elements. However, not all groups are Abelian: there might be additional 60-order group that are not Abelian (so using the Structure Theorem as radar, they are invisible). In fact, there is a famous non-Abelian group with 60 elements, called A_5, that we will study in Section 7.2.

7.7 *Fundamental theorem of finitely generated Abelian groups

In this section we prove a slightly stronger version of Theorem 567, namely, that all *finitely generated* Abelian groups (even the infinite ones) are isomorphic to products of cyclic groups. It is important to stress up front that not all Abelian groups are finitely generated: For example, $\mathbb{Q}$ is not. Also, every uncountable Abelian group cannot be finitely generated. (See the Exercises.) Hence, this theorem will concern only *some* of the Abelian groups.

Throughout the present Section, we only consider Abelian groups. For this reason, we decided to make a change in the notation: we shall denote the operation in the group G by $+$, instead of $\cdot$. Consistently, if $a \in G$ and $n \in \mathbb{N}$, if we need to express the result of operating a with itself n times, we shall use na instead of a^n. Also, we will denote the inverse of a by $-a$, instead of a^{-1}, and the neutral element by 0 instead of e. For example, the definition of "period" of a becomes the following:

$$\pi(a) \stackrel{\text{def}}{=} \begin{cases} +\infty & \text{if all multiples of } a \text{ are distinct,} \\ t & \text{if } t \text{ is the smallest positive integer for which } ta = 0. \end{cases}$$

Similarly, Proposition 511 tells us that for Abelian groups,

$$< a_1, \ldots, a_n > = \{z_1 a_1 + z_2 a_2 + \cdots z_n a_n \text{ such that } z_i \in \mathbb{Z}\}.$$

Changing notation at this point in the book may seem a bit lunatic. But to highlight how notation can simplify life considerably, let us reprove Lemma 566 with the new notation. Given an Abelian group $G =< g_1, \ldots, g_n >$ and n elements $a_1, \ldots, a_n$ whose gcd is 1, we want to show that there is another set of n generators for G that includes the element $\sum_{i=1}^n a_i g_i$.

Rewritten proof of Lemma 566. By Proposition 51, there is an integer matrix A of determinant 1 whose first row is $a_1, \ldots, a_n$. Having determinant 1, the matrix A^{-1} has also integer entries

(by the adjoint formula for calculating the inverse matrix). Now, let Γ be the column vector with entries $g_1, \ldots, g_n$. Define formally

$$X = A\Gamma,$$

and call $x_1, \ldots, x_n$ the entries of X. Since $x_i = \sum_{j=1}^{n} a_{i,j} g_j$, each x_i is in $< g_1, \ldots, g_n >$. Also, the first row of A is $(a_1, \ldots, a_n)$, so by construction $x_1 = \sum_{i=1}^{n} a_i g_i$. Our goal is thus to show that G is generated also by $\{x_1, \ldots, x_n\}$. Let us call $b_{j,k}$ the (j,k)-entry of A^{-1}, which is also an integer matrix because $\det A = 1$. Since

$$A^{-1}X = \Gamma,$$

we have that

$$\sum_{k=1}^{n} b_{j,k} x_k = g_j. \tag{64}$$

Equation 65 implies that g_j is in $< x_1, \ldots, x_n >$. Since this holds for all the generators g_j, it follows that every element of G is in $< x_1, \ldots, x_n >$. □

Isn't it easier? This new notation basically allows us to use Linear Algebra. Let us now recall what is perhaps the most important algorithm in Linear Algebra, namely, **Gauss' reduction**.

For fixed n, there are three types of so-called "elementary $n \times n$ matrices":

- $E_{i,j}$ (obtained from the identity by swapping rows i, j), which has determinant -1;
- $E_{i,j}(z)$ (obtained from the identity by replacing row i with "row i plus z times row j"), which has determinant 1;
- $E_i(z)$ (obtained from the identity by multiplying row i by the scalar z), of determinant z.

For example, for $n = 2$, we have:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_{1,2}(4) = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}, \quad \text{and } E_2(7) = \begin{pmatrix} 1 & 0 \\ 0 & 7 \end{pmatrix}.$$

These elementary matrices help us perform the analogous operations on a generic matrix A. In fact,

- $E_{i,j}A$ is the matrix obtained from A by swapping rows i, j;
- $E_{i,j}(z)A$ is obtained from A by replacing row i with "row i plus z times row j";
- $E_i(z)A$ is obtained from A by multiplying row i by z.

For example,

$$E_{1,2}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}; \; E_{1,2}(4)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+4c & b+4d \\ c & d \end{pmatrix}; \; E_2(7)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 7c & 7d \end{pmatrix}.$$

Analogously for columns:

- $AE_{i,j}$ is obtained from A by swapping columns i, j;
- $AE_{i,j}(z)$ is obtained from A by replacing column i with "column i plus z times column j";
- $AE_i(z)$ is obtained from A by multiplying column i by z.

For example,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} E_{1,2} = \begin{pmatrix} b & a \\ d & c \end{pmatrix}; \; \begin{pmatrix} a & b \\ c & d \end{pmatrix} E_{1,2}(4) = \begin{pmatrix} a+4b & b \\ c+4d & d \end{pmatrix}; \; \begin{pmatrix} a & b \\ c & d \end{pmatrix} E_2(7) = \begin{pmatrix} a & 7b \\ c & 7d \end{pmatrix}.$$

Now, you may remember that the inverse of $E_{i,j}$ is $E_{i,j}$ itself (swapping back!); the inverse of $E_{i,j}(z)$ is $E_{i,j}(-z)$ (resubtracting back what we had added); and the inverse of $E_i(z)$ is $E_i(\frac{1}{z})$

(scaling back). These three inverses certainly exist in the world of matrices with *real* coefficients, or, more generally, of matrices with coefficients in a field, like $\mathbb{Q}$ or $\mathbb{Z}_2$. But if we want to focus on matrices with *integer* coefficients, then $E_i(\frac{1}{z})$ does not exist, unless $z = \pm 1$ (because its determinant is z). So in order to operate within the realm of integer matrices, we are not going to allow elementary matrices of the third kind, except for $E_i(-1)$, whose inverse is itself.

The result you have seen in the Linear Algebra course is basically the following:

Definition 572 (Diagonal-positive)**.** We say that an $m \times n$ matrix with entries in $\mathbb{Z}$ (or $\mathbb{Q}$, or $\mathbb{R}$) is *diagonal-positive* if there exists an integer $t \in \{1, \ldots, \min(m, n)\}$ such that $a_{11}, \ldots, a_{tt}$ are all positive, while all other entries of A are zero.

The notation "diagonal-positive" is specific to this book; but a very similar notion, called "Smith normal form", is standard in the literature. Google it!

Theorem 573 (Gauss)**.** *Given any nonzero matrix A (square or rectangular) with entries in a field, one can find a diagonal-positive matrix D* with all its nonzero entries equal to 1*, such that*

$$D = UAV,$$

where U, V are products of elementary matrices.

Our next goal is to show that if the entries of A are in $\mathbb{Z}$, it is still possible to find a diagonal-positive matrix D such that

$$D = UAV,$$

where U, V are products of elementary matrices. However, we will no longer have the guarantee that the positive entries of D are equal to 1. Intuitively, the reason is that when we reach a matrix like $D = \begin{pmatrix} 1 & 0 \\ 0 & 7 \end{pmatrix}$, we no longer have the possibility of multiplying the second row by $\frac{1}{7}$, because we decided to "ban" elementary matrices of the third kind.

Theorem 574 (Diagonalization of integral matrices)**.** *Let A be any nonzero matrix (square or rectangular) with entries in $\mathbb{Z}$. Then we can find square matrices U, V, with determinant ± 1, such that the matrix $D \stackrel{\text{def}}{=} UAV$ is diagonal-positive.*

Proof. Let us try to set up a reduction algorithm. Even though we cannot rescale rows or columns, we are still allowed to (1) exchange rows or columns in A (because it corresponds to multiplying A to the left or to the right by elementary matrices of the first kind); (2) replace a row/column of A by that row/column plus z times another row/column (because it corresponds to multiplying A by elementary matrices of the second kind); (3) change all signs in a single row/column of A. Here is the right strategy:

- Find the integer b with the smallest absolute value. (It exists, because A is a non-zero matrix.) With operations of row exchange or column exchange, we can bring it to the top-left position: Row 1, column 1.
- Now we want to "zero out" all other entries in the first column. Let's start with the element in row 2, column 1. Call it a. If $a = 0$ we do nothing. If $a \neq 0$, we perform the Euclidean division for integers (see the Exercises in Chapter 1): we have

  $$a = bq + r \text{ with } 0 < r \leq |b|.$$

  Now let us replace row 2 by "row 2 minus q times row 1". This way, the element a gets replaced by $a - bq$, also known as r. If $r = 0$ we are happy: we zeroed out the entry in

row 2, column 1, as planned. If not, by definition r has smaller absolute value than b. So we can swap row 1 and row 2, taking this r into the top-left position; and we restart the algorithm. Steps of this type can only occur a finite number of times, because at each step the element in the top-right corner has smaller and smaller absolute value. Note that this step modifies only row 1 and row 2 of the matrix, without touching the other rows.

- Similarly, we zero out the element in row i, column 1, for fixed $i = 3, 4, 5....$, in this order. This modifies row i and possibly row 1, if you have to swap it with row i; but it does not modify any other row. Concretely, this means that the work you do to zero out row 7 does not spoil the work you did before, to zero out row 4.
- Now the first column has a nonzero integer d at the top, and all zeroes below. Next, we want to "zero out" all other entries in the first row. Let's start with the element in row 1, column 2. Call it c. We perform the Euclidean division

$$c = dq + r \text{ with } 0 < r \leq |d|$$

and we replace column 2 by "column 2 minus q times column 1". As before, if $r = 0$ we have reached our goal. If $r > 0$, we swap columns 1 and 2, thereby replacing the integer d with an integer r of smaller absolute value. Again, this can happen only a finite number of times.

- Note: This last swap may indeed ruin the work we had done before to zero out the first column. If this is the case, we need to start over again. However, each time we "restart", the absolute value of the top-left element decreases strictly. Thus this unpleasant situation of "my work has been ruined!, I need to restart" happens only finitely many times. After which, you do manage to clear out the first column *and* the element in row 1, column 2.
- Similarly, we zero out the element in row 1, column j, for all $j \geq 3$.
- At this point the first row and the first column have all zeroes, except for the element at the top left corner. Now we induct: "ignore" the first row and the first column, and proceed to find, in the resulting submatrix, an integer b with the smallest absolute value. Place it with swaps in row 2, column 2; then zero out the second row and the second column; and so on.
- This way we obtain a matrix that has nonzero entries only on the main diagonal (and all consecutive). To make it diagonal-positive, we change the sign of each row that has a negative element. □

Example 575. Let us diagonalize the integer matrix

$$A = \begin{pmatrix} 18 & 6 & -6 & 0 \\ 7 & 2 & 4 & 0 \\ 27 & 8 & 8 & 2 \end{pmatrix}.$$

In red we have highlighted (one occurrence of) the smallest positive integer. We start the reduction by placing the red element in the top-left corner, via swaps:

$$\begin{pmatrix} 18 & 6 & -6 & 0 \\ 7 & 2 & 4 & 0 \\ 27 & 8 & 8 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 7 & 2 & 4 & 0 \\ 18 & 6 & -6 & 0 \\ 27 & 8 & 8 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 7 & 4 & 0 \\ 6 & 18 & -6 & 0 \\ 8 & 27 & 8 & 2 \end{pmatrix}.$$

Next, we zero out the first row by means of Euclidean divisions. Whenever we obtain a nonzero remainder, we highlight it in red, and take it to the top-left-corner.

$$\begin{pmatrix} 2 & 7 & 4 & 0 \\ 6 & 18 & -6 & 0 \\ 8 & 27 & 8 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 1 & 4 & 0 \\ 6 & 0 & -6 & 0 \\ 8 & 3 & 8 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 4 & 0 \\ 0 & 6 & -6 & 0 \\ 3 & 8 & 8 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 4 & 0 \\ 0 & 6 & -6 & 0 \\ 3 & 2 & 8 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & -6 & 0 \\ 3 & 2 & -4 & 2 \end{pmatrix}$$

Now we zero-out the first column, using Euclidean divisions:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & -6 & 0 \\ 3 & 2 & -4 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & -6 & 0 \\ 0 & 2 & -4 & 2 \end{pmatrix}.$$

Now we ignore the first row and first column, and seek a smallest positive integer in the remaining rows/columns. We then take it via swaps to position $(2,2)$, and then zero out the second row.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & -6 & 0 \\ 0 & 2 & -4 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & -4 & 2 \\ 0 & 6 & -6 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 6 & 6 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 6 & 6 & 0 \end{pmatrix}.$$

Finally, we zero out the second column, and we are done!

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 6 & 6 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \end{pmatrix} \stackrel{\text{def}}{=} D.$$

Explicitly,

$$D = UAV,$$

where

$$U \stackrel{\text{def}}{=} E_{3,2}(-3)E_{2,3}E_{3,1}(-3)E_{1,2} \quad \text{and} \quad V \stackrel{\text{def}}{=} E_{1,2}E_{2,1}(-3)E_{1,2}E_{2,1}(-2)E_{3,1}(-4)E_{3,2}(2)E_{4,2}(-1).$$

Remark 576. The algorithm above relies on the fact that in $\mathbb{Z}$ it is possible to perform a Euclidean division. However, it is possible to modify the algorithm slightly to make it work for any matrix with entries in a PID ring R. We will see this in Theorem 693. Moreover, the result can be strengthened into "$d_{i,i}$ divides $d_{j,j}$ for every $j > i$".

Now we are ready to go back and prove that finitely generated Abelian groups are products of cyclic groups.

Definition 577. We say that some elements $a_1, \ldots, a_n$ of an Abelian group G form a *basis* for G if the presentation homomorphism

$$\begin{array}{rccc} \varphi: & \mathbb{Z}^n & \longrightarrow & G \\ & (z_1, \ldots, z_n) & \longmapsto & z_1a_1 + z_2a_2 + \ldots + z_na_n \end{array}$$

is bijective. Equivalenty, $a_1, \ldots, a_n$ form a basis of G if every element of G can be written in a unique way as combination of the a_i's with coefficients in $\mathbb{Z}$. We also say that a group G is *free of rank* n if it has a basis consisting of n elements.

Remark 578. $G =< a_1, \ldots, a_n >$ if and only if the map φ above is surjective. So every basis is a special set of generators; namely, a set of generators for which the presentation homomorphism is injective.

Example 579. Consider $\mathbb{Z}[i]$ with respect to addition. This is a free Abelian group: the presentation homomorphism

$$\begin{array}{rccc} \varphi: & \mathbb{Z}^2 & \longrightarrow & \mathbb{Z}[i], \\ & (a,b) & \longmapsto & a + ib \end{array}$$

is bijective.

Non-Example 580. $\mathbb{Z}_2$ is not free. In fact, we know that $\mathbb{Z}_2 =< \overline{1} >$, so the (only possible) presentation homomorphism

$$\begin{array}{rccc} \varphi: & \mathbb{Z} & \longrightarrow & \mathbb{Z}_2 \\ & 1 & \longmapsto & \overline{1} \end{array}$$

is surjective; but it is not injective. In the literature, one usually says that "$\mathbb{Z}_2$ has a system of generators, but it does not have a basis".

Remark 581. By definition, every free group of rank n is isomorphic to $\mathbb{Z}^n$. In particular, all free groups are infinite.

Non-Example 582. $\mathbb{Q} \times \mathbb{Z}$ is infinite, but still not free. By contradiction, suppose one could find a surjective presentation homomorphism

$$\varphi: \quad \mathbb{Z}^n \quad \longrightarrow \quad \mathbb{Q} \times \mathbb{Z}.$$

Then the composition $\pi_1 \circ \varphi$ with the projection onto the first coordinate would yield a surjective homomorphism from $\mathbb{Z}^n$ to $\mathbb{Q}$, which would contradict the fact that $\mathbb{Q}$ is not finitely generated.

The next theorem is the reason why we spent time in proving diagonalization of integral matrices. It says that all subgroups of $\mathbb{Z}^n$ are free! And in fact, it gives us even more information:

Lemma 583. *All subgroups of $\mathbb{Z}^n$ are finitely generated, by at most n generators.*

Proof. By induction. The case $n = 1$ is Proposition 520. For $n \geq 2$, let H be a subgroup of $\mathbb{Z}^{n+1}$. Let $\pi : \mathbb{Z}^{n+1} \to \mathbb{Z}$ be the projection onto the last component. Let $K = \pi(H)$. This is a subgroup of $\mathbb{Z}$, so by Proposition 520, it is generated by some elements k (namely, the smallest positive integer in K). Since $k \in \pi(H)$, we can find integers $a_1, \ldots, a_n$ such that $(a_1, a_2, \ldots, a_n, k)$ is in H. Now consider a generic element $(h_1, \ldots, h_n, h_{n+1})$ in H. Since h_{n+1} is a multiple of k, let q be the integer such that $h_{n+1} = qk$. Then we can write

$$(h_1, \ldots, h_n, h_{n+1}) = q(a_1, \ldots, a_n, k) + (h_1 - qa_1, \ldots, h_n - qa_n, 0). \tag{65}$$

But for fixed $a_1, \ldots, a_n$ (and k), the set

$$S \stackrel{\text{def}}{=} \{(h_1 - qa_1, \ldots, h_n - qa_n, 0) \text{ such that } h_1, \ldots, h_n, q \in \mathbb{Z}\}$$

is a subgroup of $\mathbb{Z}^n$, so it is finitely generated by at most n elements. Equation (66) tells us that $H = K + S$ is finitely generated by at most $n + 1$ elements. □

Remark 584. For arbitrary (non-Abelian groups), it is not true that subgroups of finitely generated groups are finitely generated. For example, let $G =< x, y >$ and let S be the subgroup generated by all elements of the form $y^n x y^{-n}$, with $n \geq 1$. Then there is no finite set of generators for S.

Theorem 585. *Let H be a nonzero subgroup of $\mathbb{Z}^n$. Then there are positive integers $d_1, d_2, \ldots, d_t$ and there is a basis $\{e_1, e_2, \ldots e_n\}$ of $\mathbb{Z}^n$ such that $\{d_1 e_1, d_2 e_2, \ldots, d_t e_t\}$ are a basis of H.*

Proof. Let $\{b_1, \ldots, b_n\}$ be any basis of $\mathbb{Z}^n$. For example, one could choose $b_1 = (1, 0, \ldots, 0)$, $b_2 = (0, 1, 0, \ldots, 0)$, and so on, until $b_n = (0, \ldots, 0, 1)$. Let $\{g_1, \ldots, g_r\}$ be a set of $r \leq n$ generators for H (not necessarily a basis), which exists by Lemma 583. Since $H \subseteq \mathbb{Z}^n$, every element of H can be written as linear combination of the b_i's. In particular, each g_j can be written as a linear combination of the b_i's, with coefficients in $\mathbb{Z}$. In other words,

$$(g_1, g_2, \ldots, g_r) = (b_1, b_2, \ldots, b_n)A, \text{ for some } A \in M_{n,r}(\mathbb{Z}). \tag{66}$$

Now we diagonalize A according to Theorem 574. Let $d_1, \ldots, d_t$ be the nonzero diagonal elements in the diagonal-positive matrix $D = UAV$. Obviously,

$$U^{-1}D = U^{-1}UAV = AV. \tag{67}$$

Now, define

$$(e_1, \ldots, e_n) \stackrel{\text{def}}{=} (b_1, \ldots, b_n)U^{-1}.$$

It is an exercise to see that because U^{-1} is invertible, and because the vectors $\{b_1, \ldots, b_n\}$ form a basis of $\mathbb{Z}^n$, then also $\{e_1, \ldots, e_n\}$ are a basis of $\mathbb{Z}^n$. But then if we take Equation 67 and multiply to the right by V, plugging in Equation 68 we see that

$$(g_1, g_2, \ldots, g_r)V = (b_1, b_2, \ldots, b_n)AV = (b_1, b_2, \ldots, b_n)U^{-1}D = (e_1, e_2, \ldots, e_n)D. \tag{68}$$

But by the shape of D, this tells us that

$$(g_1, g_2, \ldots, g_r)V = (d_1e_1, d_2e_2, \ldots, d_te_t, 0, \ldots, 0) \tag{69}$$

Now we make three final claims:

- Equation 70 tells us that the d_ie_i are all in H, because each d_ie_i can be obtained multiplying $(g_1|g_2|\ldots|g_r)$ by the i-th column of V. (So each d_ie_i is a combination with coefficients in $\mathbb{Z}$ of $g_1, \ldots, g_r$).
- Equation 70 tells us also that

  $$(g_1|g_2|\ldots|g_r) = (d_1e_1, d_2e_2, \ldots, d_te_t, 0, \ldots, 0)V^{-1},$$

  so every g_j can be written as linear combination of the d_ie_i's. This implies that

  $$H =< d_1e_1, \ldots, d_te_t > .$$

- Consider now the presentation homomorphism

  $$\begin{array}{rccc} \varphi: & \mathbb{Z}^t & \longrightarrow & H \\ & (z_1, \ldots, z_t) & \longmapsto & z_1(d_1e_1) + \ldots + z_t(d_te_t). \end{array}$$

  Is this injective? Suppose $0 = z_1d_1e_1 + \ldots + z_td_te_t$. Of course, this could be rewritten as

  $$0 = z_1d_1e_1 + \ldots + z_td_te_t + 0e_{t+1} + \ldots + 0e_n.$$

  But $e_1, \ldots, e_n$ are a basis. Hence,

  $$z_1d_1 = \ldots = z_td_t = 0.$$

  Finally we can use the fact that the d_i's are all positive, to conclude that $z_1 = \ldots = z_t = 0$. Hence, the presentation homomorphism is injective and the d_ie_i are a basis for H. □

Theorem 586. *Every finitely generated Abelian group is isomorphic to the product of cyclic groups. More precisely, if G is any Abelian group generated by n elements, then there exist natural numbers $s, t \geq 0$, with $s + t \leq n$, such that G is isomorphic to*

$$\mathbb{Z}^s \times \mathbb{Z}_{d_1} \times \ldots \times \mathbb{Z}_{d_t}$$

for suitable integers $d_1, \ldots, d_t$ greater than 1.

Proof. Let $G =< g_1, \ldots, g_n >$. Consider the presentation homomorphism

$$\begin{array}{rccc} \varphi: & \mathbb{Z}^n & \longrightarrow & G \\ & (z_1, \ldots, z_n) & \longmapsto & z_1 g_1 + \ldots z_n g_n. \end{array}$$

Set $H \stackrel{\text{def}}{=} \ker \varphi$. By the first isomorphism theorem,

$$\mathbb{Z}^n / H \cong G.$$

On the other hand, H is a subgroup of $\mathbb{Z}^n$, so we can apply Theorem 585. We have that

$$\mathbb{Z}^n / H = \mathbb{Z}^s \times \mathbb{Z}_{d_1} \times \ldots \times \mathbb{Z}_{d_t},$$

with $s + t = n$. However, it is possible that some d_i's are equal to 1, which result in irrelevant factors of the type $\mathbb{Z}/_{< 1 >} = \{0\}$ in the decomposition. If we discard these factors, then we obtain a decomposition

$$\mathbb{Z}^s \times \mathbb{Z}_{d_1} \times \ldots \times \mathbb{Z}_{d_{t'}}$$

with $s + t' \leq n$. □

We conclude with a uniqueness result.

Definition 587. Let G be an Abelian group. The *torsion subgroup* of G is

$$T(G) \stackrel{\text{def}}{=} \{a \in G \text{ such that } \pi(a) \text{ is finite } \}.$$

It is a subgroup by Proposition 451: In fact, suppose $\pi(x) = m$ and $\pi(y) = n$. Then both $mnx = 0$ and $mny = 0$, so $mn(x - y) = 0$. This proves that $\pi(x - y)$ is finite.

Lemma 588. *Let G, H be Abelian groups.*
(1) *If G is finite, $T(G) = G$.*
(2) *If G is free, $T(G) = (0)$.*
(3) *$T(G \times H) \cong T(G) \times T(H)$.*
(4) *Any isomorphism from G to H restricts to an isomorphism from $T(G)$ to $T(H)$.*

Proof.
(1) Obvious.
(2) Follows from (4) and the fact that $T(\mathbb{Z}^n) = (0)$.
(3) Let $(g, h) \in T(G) \times T(H)$. Then for some natural numbers n, m, we have $mg = 0 = nh$. Then clearly
$$mn(g, h) = (mng, mnh) = (0, 0),$$
so $(g, h) \in T(G \times H)$. This proves $T(G) \times T(H) \subseteq T(G \times H)$; the other inclusion is similar.
(4) Let $f : G \to H$ be an isomorphism. If $nx = 0$, then $nf(x) = f(nx) = f(0) = 0$. This shows that the image under f of any element of $T(G)$ belongs to $T(H)$. Now let f' be the restriction of f to $T(G)$. Automatically then f' is an injective homomorphism, because so is f. It remains to check that f' is surjective. Let $h \in T(H) \subseteq H$, and suppose that $nh = 0$. Because f was surjective, there is a g in G such that $f(g) = h$. But then
$$f(gn) = nf(g) = nh = 0;$$
but f was also injective, so $gn = 0$, which proves that $g \in T(G)$.

□

Lemma 589. *Let K, L be finite Abelian groups. If $\mathbb{Z}^t \times K \cong \mathbb{Z}^u \times L$, then $t = u$ and $K \cong L$.*

Proof. Applying Lemma 588, part (1), (2), (3), in this order, we get that

$$K = T(K) \cong (0) \times T(K) \cong T(\mathbb{Z}^t) \times T(K) = T(\mathbb{Z}^t \times K).$$

Similarly, we obtain

$$L = T(L) \cong (0) \times T(L) \cong T(\mathbb{Z}^u) \times T(L) = T(\mathbb{Z}^u \times L).$$

But by assumption $\mathbb{Z}^t \times K \cong \mathbb{Z}^u \times L$, so by Lemma 588, part (4), we have $T(\mathbb{Z}^t \times K) \cong T(\mathbb{Z}^u \times L)$. Putting everything together, we get

$$K \cong T(\mathbb{Z}^t \times K) \cong T(\mathbb{Z}^u \times L) \cong L.$$

Now, choose an isomorphism $\eta : L \to K$. Then it is easy to see that

$$\begin{array}{rccc} \delta: & \mathbb{Z}^t \times L & \longrightarrow & \mathbb{Z}^t \times K \\ & (z, \ell) & \longmapsto & (z, \eta(l)) \end{array}$$

is also an isomorphism. Let us compose δ with the isomorphism between $\mathbb{Z}^t \times K$ and $\mathbb{Z}^u \times L$ that exists by assumption. We obtain an isomorphism $\varphi : \mathbb{Z}^t \times L \longrightarrow \mathbb{Z}^u \times L$. By Lemma 588, part (4), this isomorphism restricts to an isomorphism of $(0) \times L$ to $(0) \times L$. Now consider

$$\begin{array}{rccc} \iota: & \mathbb{Z}^t & \longrightarrow & \mathbb{Z}^t \times L \\ & z & \longmapsto & (z, 0) \end{array} \quad \text{and} \quad \begin{array}{rccc} \pi: & \mathbb{Z}^u \times L & \longrightarrow & \mathbb{Z}^u \\ & (z, \ell) & \longmapsto & z. \end{array}$$

We claim that $\pi \circ \varphi \circ \iota : \mathbb{Z}^t \to \mathbb{Z}^u$ is injective. In fact, suppose $\pi \circ \varphi \circ \iota(z) = 0$. Then $0 = \pi(\varphi(z, 0))$. This means that $\varphi(z, 0)$ is of the form $(0, \ell)$ for some $\ell \in L$. But because φ restricts to an isomorphism of $(0) \times L$ to $(0) \times L$, then $(0, \ell) = \varphi(0, \ell')$ for some $\ell' \in L$. Since φ is injective, this means that

$$(z, 0) = (0, \ell')$$

which implies $z = 0$. So $\pi \circ \varphi \circ \iota$ is injective. By the first isomorphism theorem, this means that $\mathbb{Z}^t$ is isomorphic to a subgroup of $\mathbb{Z}^u$. But by Theorem 585, any subgroup of $\mathbb{Z}^u$ has a basis consisting of $\leq u$ elements. In particular, $t \leq u$.

Symmetrically, by using φ^{-1} instead of φ, one proves that $u \leq t$. Hence, $u = t$. □

Theorem 590 (Smith 1861, Kronecker 1870)**.** *For every finitely generated Abelian group G, there exist natural numbers t, h, $m_1, \ldots, m_h$ and prime numbers $p_1 \leq \ldots \leq p_h$ (not necessarily distinct), such that G can be decomposed as*

$$G \cong \mathbb{Z}^t \times \mathbb{Z}_{p_1^{m_1}} \times \mathbb{Z}_{p_2^{m_2}} \times \ldots \times \mathbb{Z}_{p_h^{m_h}}.$$

Moreover, such decomposition is unique.

Proof. The existence is determined by applying recursively Corollary 551 to the finite cyclic factors given by Theorem 567. Let us prove uniqueness. Suppose

$$\mathbb{Z}^t \times \mathbb{Z}_{p_1^{m_1}} \times \mathbb{Z}_{p_2^{m_2}} \times \ldots \times \mathbb{Z}_{p_h^{m_h}} \cong G \cong \mathbb{Z}^u \times \mathbb{Z}_{q_1^{n_1}} \times \mathbb{Z}_{q_2^{n_2}} \times \ldots \times \mathbb{Z}_{q_h^{n_h}}$$

with the q_i primes in weakly increasing order, not necessarily distinct.

Let us apply Lemma 589 to $K \stackrel{\text{def}}{=} \mathbb{Z}_{q_1} \times \ldots \times \mathbb{Z}_{q_h}$ and $L \stackrel{\text{def}}{=} \mathbb{Z}_{r_1} \times \ldots \times \mathbb{Z}_{r_k}$. We immediately get that $t = u$. Moreover, Lemma 589 tells us that $K \cong L$, so in particular their cardinalities are the same. This implies that

$$p_1^{m_1} p_2^{m_2} \cdots p_h^{m_h} = q_1^{n_1} q_2^{n_2} \cdots q_h^{n_h}.$$

But the p_i and the q_i are prime numbers, so by the Unique Decomposition of Integers, we know that the same primes must appear with same exponent on both sides of the equality above. The conclusion follows from the fact that no prime-power cyclic group can be decomposed further, as explained in Corollary 551. □

7.8 Solvable groups

Definition 591 (Solvable)**.** A group G is called *solvable* if there is a chain of subgroups

$$\{e\} = G_0 \subseteq G_1 \subseteq \ldots \subseteq G_n = G$$

such that for all i, G_i is normal in G_{i+1}, and the quotient ${}^{G_{i+1}}\!/_{G_i}$ is Abelian.

Example 592. Every Abelian group is solvable. In fact, the chain

$$\{e\} = G_0 \subseteq G_1 = G$$

does the trick.

Remark 593. In view of Theorem 590, an equivalent definition of solvable groups is obtained by replacing "the quotient ${}^{G_{i+1}}\!/_{G_i}$ is Abelian" with "the quotient ${}^{G_{i+1}}\!/_{G_i}$ is cyclic". In fact, once we decompose an Abelian group G as

$$G \cong \mathbb{Z}^t \times \mathbb{Z}_{p_1^{m_1}} \times \mathbb{Z}_{p_2^{m_2}} \times \ldots \times \mathbb{Z}_{p_h^{m_h}},$$

it is clear that

$$G_1 \stackrel{\text{def}}{=} \mathbb{Z}_{p_1^{m_1}} \times \mathbb{Z}_{p_2^{m_2}} \times \ldots \times \mathbb{Z}_{p_h^{m_h}}$$

is normal in G and the quotient ${}^{G}\!/_{G_1}$ is cyclic. At the same time,

$$G_2 \stackrel{\text{def}}{=} \mathbb{Z}_{p_2^{m_2}} \times \ldots \times \mathbb{Z}_{p_h^{m_h}}$$

is normal in G_1 and the quotient ${}^{G_1}\!/_{G_2}$ is cyclic, and so on.

Proposition 594. *Subgroups and quotients of solvable groups are solvable. Conversely, if N is a normal subgroup of a group G such that N and ${}^{G}\!/_{N}$ are solvable, then G is solvable.*

Sketch of proof. We give only the outline; verifying the details is left as exercise. Let G be a solvable group. Let H be a subgroup of a solvable group G. If $\{e\} = G_0 \subseteq G_1 \subseteq \ldots \subseteq G_n = G$ is the chain that shows the solvability of G, then

$$\{e\} = G_0 \cap H \subseteq G_1 \cap H \subseteq \ldots \subseteq G_n \cap H = H$$

proves the solvability of H, because ${}^{G_{i+1} \cap H}\!/_{G_i \cap H}$ is a subgroup of ${}^{G_{i+1}}\!/_{G_i}$, which is Abelian by assumption; and subgroups of Abelian groups are Abelian.

Similarly, let N be a normal subgroup of a solvable group G. Let $\pi : G \to {}^{G}\!/_{N}$ be the projection and let $\overline{G_i} \stackrel{\text{def}}{=} \pi(G_i)$. If $\{e\} = G_0 \subseteq G_1 \subseteq \ldots \subseteq G_n = G$ is the chain that shows the

solvability of G, then one can see that $\overline{G_i}$ is normal in $\overline{G_{i+1}}$, and any two elements commute in the quotient $\overline{G_{i+1}}/\overline{G_i}$. So

$$\{e\} = \overline{G_0} \subseteq \overline{G_1} \subseteq \ldots \subseteq \overline{G_n} = G/N$$

shows the solvability of G/N.

For the converse, the idea is the following: let $\{e\} = H_0 \subseteq H_1 \subseteq \ldots \subseteq H_m = G/N$ be a chain that shows the solvability of G/N. Each H_i can be written as

$$H_i = G_i/N$$

for a suitable subgroup G_i of G containing N, and one can show that H_i is normal in H_{i+1} if and only if G_i is normal in G_{i+1}. Now let $\{e\} = N_0 \subseteq N_1 \subseteq \ldots \subseteq N_n = N$ be a chain that shows the solvability of N. We can extend it as follows

$$\{e\} = N_0 \subseteq N_1 \subseteq \ldots \subseteq N_n = N \stackrel{!}{=} G_0 \subseteq G_1 \subseteq \ldots \subseteq G_m = G$$

to obtain a chain that shows the solvability of G. □

Example 595. A_5 is not solvable. In fact, its only normal proper subgroup is $\{e\}$, but the quotient is A_5, which is not Abelian.

Example 596. $\mathcal{S}_5$ is not solvable either. In fact, its only normal subgroup other than $\{e\}$ is A_5; the quotient $\mathcal{S}_5/A_5$ is the two-element group $\mathbb{Z}_2$, so it is indeed Abelian; but then we are stuck, because A_5 is simple.

Example 597. $\mathcal{S}_d$ is not solvable for any $d \geq 5$. In fact, $\mathcal{S}_5$ is (isomorphic to) a subgroup of $\mathcal{S}_d$, because any permutation of $\{1, \ldots, 5\}$ can be identified with a permutation of $\{1, \ldots, d\}$ that fixes the elements $6, 7, \ldots, d$. Were $\mathcal{S}_d$ solvable, by Proposition 594 $\mathcal{S}_5$ would be solvable too.

Deeper thoughts 598. A very deep result of Feit and Thompson (from 1963) states that if G has an odd number of elements, then G is solvable. In particular, this immediately implies that all non-Abelian simple groups have an even number of elements. However, there are Abelian groups with an odd number of elements, namely, the cyclic groups $\mathbb{Z}_p$ for p prime. Moreover, a theorem by Burnside from 1904 says that that the number of elements of a finite simple group is either a prime, or a number divisible by at least three distinct primes (like A_5, which has $2^2 \cdot 3 \cdot 5$ elements). The converse of Burnside's theorem does not work: for example, 30, 42, and 60 are the three smallest numbers that are divisible by three distinct primes; but by inspection, there is no simple group with 30 elements, nor any with 42. So the smallest non-Abelian simple group is A_5.

These studies lead to one of the most remarkable achievements of modern mathematics, namely, the complete classification of all *finite simple groups*. This huge project, initiated by Gorenstein and finally completed by Aschbacher–Smith in 2004, involves thousands and thousands of pages of proofs. The conclusion is that finite simple groups are:

- the cyclic groups $\mathbb{Z}_p$, with p primes (the only Abelian ones in the list);
- the groups A_n, for $n \geq 5$;
- the so-called "groups of Lie type", which form also an infinite class;
- plus 27 exceptional groups, called "sporadic groups".

Deeper thoughts 599. The name "solvable" comes from the so-called **Galois theory**. We all know that any degree-two equation $ax^2 + bx + c = 0$ can be solved by means of a "formula involving radicals":

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a.}$$

There are more complicated formula involving radicals, published by the Italian mathematician Cardano in 1545, to solve any equation of degree three or four. But what about the generic equation of degree n? This problem was eventually settled two centuries later, thanks to the work by Ruffini (1799), Abel (1824), and most notably Galois. Galois proved that a polynomial equation $p(x) = 0$ of degree d is solvable by radicals if and only if a certain group associated to p is solvable. He also showed that this group is a subgroup of $\mathcal{S}_d$, and sometimes it coincides with $\mathcal{S}_d$, as it happens in the case of the polynomial

$$p = x^5 - 10x^4 + 2.$$

But since $\mathcal{S}_5$ is not solvable (cf. Example 596), there cannot be any formula involving radicals to solve the quintic equation $x^5 - 10x^4 + 2 = 0$.

Note how unexpected yet philosophically powerful this result is: It does not say, "we have not found a formula yet". It says that "no formula can exist"!

Galois' result were hastily written in letter to his friend Auguste Chevalier on May 29, 1832, the day before he was shot in a duel over a love affair. He was 20.

Remark 600. From Calculus we know that the polynomial $p(x) = x^5 - 10x^4 + 2 = 0$ has at least one real root α by the intermediate value theorem, since

$$\lim_{x \to -\infty} p(x) = -\infty \quad \text{and} \quad \lim_{x \to +\infty} p(x) = +\infty.$$

We also know that such root α cannot be a rational number: In fact, by Eisenstein's criterion (Thm. 352, cf. also Theorem 341), $x^5 - 10x^4 + 2$ is irreducible over $\mathbb{Q}[x]$, so in particular it cannot have rational roots. However, the situation is not so mysterious as it seems: There are plenty of approximation methods (based on the intermediate value theorem) that allow us to compute the roots of p with arbitrary precision. It turns out that p has three real roots and two complex ones:

$$x_1 \approx 0.68063, \quad x_2 \approx -0.65817, \quad x_3 \approx 9.9998, \quad \text{and} \quad x_{4,5} \approx -0.01113 \pm 0.66809 \cdot i.$$

So even if a *formula by radicals* does not exists, we are still able to *compute* the solutions with arbitrary degree of precision!

If you are interested, stay tuned, Chapter 8 is about this.

7.9 Exercises

1. In $GL_2(\mathbb{R})$, what is the period of the element $a = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$?
2. is the multiplicative group of nonzero real numbers isomorphic to the additive group of all real numbers?
3. Show that if G is abelian, then

$$< a_1, \ldots, a_n >= \{(a_1)^{z_1}(a_2)^{z_2} \cdots (a_n)^{z_n} \text{ such that } z_i \in \mathbb{Z}\}.$$

4. Prove that $\mathbb{Q}$ is not finitely generated. (Hint: Use the previous exercise.)

5. Let x, y be two group elements such that $x^{2018} = y^{2019}$ and $xyx = yxy$. Prove that $x = y = e$.
6. Prove Proposition 511.
7. Let M be a matrix of integer coefficients. Prove that M^{-1} exists and has integer coefficients if and only if $\det M = \pm 1$.
8. Let $\{z_1, \ldots, z_n\}$ be any basis of $\mathbb{Z}^n$. Let M be an $n \times n$ matrix of integer coefficients, with $\det M = \pm 1$. Show that $\{z_1 M, \ldots, z_n M\}$ are again a basis of $\mathbb{Z}^n$.
9. Prove that every Abelian group that is finitely generated has countably many elements.
10. Let G be a group where $x^2 = e$ for all $x \in G$. Prove that G is Abelian.
11. Let G be a group with $2p$ elements, p prime. Prove that if every element of G has period 1 or 2, then G is Abelian. Use this to show that G contains a subgroup with p elements. (Hint: what can the period of an element $x \neq e$ be?)
12. For any groups G, H, K, prove that $G \times (H \times K)$ is isomorphic to $(G \times H) \times K$.
13. Represent the following as product of disjoint cycles:

$$(1267)(34562)(68) \qquad (123456)(1357)(163) \qquad (14)(15)(16)(17)$$

14. Prove that if a subgroup of $\mathcal{S}_n$ contains $(1, \ldots, n)$ and $(n-1, n)$, then it is the whole $\mathcal{S}_n$.

8 Galois Theory

This chapter is largely based on the beautiful book by John and Margaret Maxfield, Abstract Algebra and Solutions by Radicals, Dover, New York, 1992.

Given a degree-two polynomial in $\mathbb{R}[x]$

$$f(x) = ax^2 + bx + c$$

you have probably seen in high school that its "complex roots" are

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{70}$$

Can we find analogous formulas for degree three? What about higher degree? The answer to these questions is provided by Galois theory, and requires a careful, philosophical reflection on what a *formula* really is, and how we can obtain one.

Formula 71 comes from a clever change of variables that allows us to reduce ourselves to the case $a = 1$ and $b = 0$, which we already solved in Lemma 421. The trick is simply to define a new variable

$$y \stackrel{\text{def}}{=} x + \frac{b}{2a}.$$

Then obviously $x = y - \frac{b}{2a}$ and

$$f(x) = a\left(y - \frac{b}{2a}\right)^2 + b\left(y - \frac{b}{2a}\right) + c = \left(ay^2 - by + \frac{b^2}{4a}\right) + \left(by - \frac{b^2}{2a}\right) + c = ay^2 - \frac{b^2 - 4ac}{4a}.$$

So

$$f(x) = 0 \iff ay^2 = \frac{b^2 - 4ac}{4a} \iff y^2 = \frac{b^2 - 4ac}{(2a)^2}.$$

So we can extract the square root, and once we find y, we can of course recover x by remembering that $x = -\frac{b}{2a} + y$. This is precisely how Formula 71 is obtained.

The trick above shows also the following interesting fact. If we define the "discriminant of f" as $\Delta(f) \stackrel{\text{def}}{=} b^2 - 4ac$, there are three cases:

- if $\Delta > 0$ then f has two distinct real solutions,
- if $\Delta = 0$ then $f = y^2 = \left(x + \frac{b}{2a}\right)^2$ has only one real solution,
- if $\Delta < 0$ then f has two distinct complex solutions that are conjugate.

Remark 601. Formula 71 contains a minor confusion between algebra and calculus, which is better to clarify right away. Calculus focuses on functions defined on subintervals of $\mathbb{R}$, like for example the bijection

$$\begin{array}{rcl} g : \mathbb{R}_{\geq 0} & \longrightarrow & \mathbb{R}_{\geq 0} \\ x & \longmapsto & x^2 \end{array}.$$

The inverse of such function is called "square root". By definition, the square root is thus defined only on nonnegative numbers, and it returns always a nonnegative number. This explains why, when $\Delta > 0$, Formula 71 reads "$\pm\sqrt{\Delta}$": There are two solutions here! The square root function selects only one of them, but we shouldn't forget about the other one. However, when $\Delta < 0$, the expression "$\pm\sqrt{\Delta}$" does not make sense from the point of view of Calculus. What we mean in this case is "*any* of the two solutions of the equation $z^2 = \Delta$ (cf. Lemma 421)".

Remark 602. With the convention above, we can find the complex roots of any degree-two polynomial *with complex coefficients*. Given $f(x) = ax^2 + bx + c \in \mathbb{C}[x]$, the change of variable $y \stackrel{\text{def}}{=} x + \frac{b}{2a}$ reduces it to the polynomial $a(y^2 - \frac{b^2-4ac}{4a^2})$; so after dividing by a, we want to study the roots of the polynomial $y^2 - c'$, with $c' \stackrel{\text{def}}{=} \frac{b^2-4ac}{(2a)^2}$. And we know how to find the root of the latter using Lemma 421.

8.1 Degree-three equations

The solution to equations of degree three was published in 1545 by the Italian mathematician Girolamo Cardano, exploiting previous work by Niccoló Tartaglia and Scipione Del Ferro. Cardano omitted the details of the most important case (the so-called "casus irreducibilis", see below), which was clarified in 1572 by Rafael Bombelli.

As a warm up, let us start by solving the equation $x^3 = 1$ over $\mathbb{C}$.

Lemma 603. *Let* $\xi \stackrel{\text{def}}{=} -\frac{1}{2} + \frac{\sqrt{3}}{2}i$. *Then* $\xi^0 = 1$, $\xi^1 = \xi$, $\xi^2 = \overline{\xi}$ *are the three solution of the equation* $z^3 = 1$ *in* $\mathbb{C}$.

Proof. Since $\xi^2 = \left(\frac{1}{4} - \frac{3}{4}\right) - 2i\frac{1}{2}\frac{\sqrt{3}}{2} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i = \overline{\xi}$, we have $\xi^3 = \xi \cdot \overline{\xi} = \frac{1}{4} + \frac{3}{4} = 1$. □

Corollary 604. *Let* $r \in \mathbb{R}$, $r \neq 0$. *With the notation above, the three complex solutions of* $z^3 = r$ *are* $\sqrt[3]{r}$, $\sqrt[3]{r} \cdot \xi$, *and* $\sqrt[3]{r} \cdot \overline{\xi}$; *of these three, only the first one is in* $\mathbb{R}$.

Suppose now we have a degree-three polynomial equation with real coefficients,

$$0 = ax^3 + bx^2 + cx + d.$$

Analogously to the degree-two case, let us start by reducing ourselves to the case where $a = 1$ and $b = 0$. For this we define the new variable

$$y \stackrel{\text{def}}{=} x + \frac{b}{3a}.$$

Then obviously $x = y - \frac{b}{3a}$ and

$$\begin{aligned} 0 &= a\left(y - \tfrac{b}{3a}\right)^3 + b\left(y - \tfrac{b}{3a}\right)^2 + c\left(y - \tfrac{b}{3a}\right) + d = \\ &= ay^3 - by^2 + \tfrac{b^2}{3a}y - \tfrac{b^3}{27a^2} + by^2 - \tfrac{2b^2}{3a}y + \tfrac{b^3}{9a^2} + cy - \tfrac{bc}{3a} + d = \\ &= a\left(y^3 + \tfrac{3ac-b^2}{3a^2}y + \tfrac{2b^3-9abc+27a^2d}{27a^3}\right). \end{aligned}$$

So setting $p \stackrel{\text{def}}{=} \frac{3ac-b^2}{3a^2}$ and $q \stackrel{\text{def}}{=} \frac{2b^3-9abc+27a^2d}{27a^3}$, we reduced ourselves to the problem of solving a degree-three equation of the type

$$0 = y^3 + py + q. \tag{71}$$

At this point, it seems we are stuck. But here Cardano comes up with a new, brilliant change of variables. The idea is to introduce two new variables u, v such that

$$\begin{cases} u + v = y \\ 3uv = -p. \end{cases} \tag{72}$$

Using these new variables, Equation 72 becomes

$$0 = (u+v)^3 - 3uv(u+v) + q = u^3 + v^3 + q.$$

So the *sum* of u^3 and v^3 is $-q$. But the *product* of u^3 and v^3 can also be computed: in fact,

$$u^3 \cdot v^3 = (uv)^3 = \left(-\frac{p}{3}\right)^3 = -\frac{p^3}{27}.$$

Knowing their sum and their product, we can guess who u^3 and v^3 are! Namely, they are the two solutions of the quadratic equation

$$z^2 + qz - \frac{p^3}{27} = 0,$$

whose discriminant is

$$\Delta = q^2 + \frac{4}{27}p^3.$$

So, solving for u and v, we can eventually find $y = u + v$!

Using this idea, Cardano and Bombelli came up with the following formula for $y = u + v$:

$$y_k = \xi^k \cdot \sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}} + \xi^{2k} \cdot \sqrt[3]{\frac{-q - \sqrt{\Delta}}{2}}, \quad k = 0, 1, 2, \tag{73}$$

where the two cubic radicals are defined to be any two cubic roots with product $\frac{-p}{3}$.

How many of the roots are real? This depends on the Δ above. First of all, we know that $y^3 + py + q$ has at least one real root by the intermediate value theorem (since $\lim_{y\to-\infty} y^3 + py + q = -\infty$ and $\lim_{y\to+\infty} y^3 + py + q = +\infty$). Also, the number of non-real roots must be even, because of Corollary 158. So the number of complex, non-real roots is either 0 or 2. Let's see how Δ reflects the difference. Let y_0, y_1, y_2 be the three roots (real or complex) of $y^3 + py + q$. This means that

$$y^3 + py + q = (y - y_0)(y - y_1)(y - y_2).$$

By comparing coefficients, one gets the system of equations

$$\begin{cases} 0 = y_0 + y_1 + y_2 \\ p = y_0y_1 + y_0y_2 + y_1y_2 \\ q = -y_0y_1y_2. \end{cases} \tag{74}$$

From this system, after long and tedious calculations, one gets the identity

$$[(y_0 - y_1)(y_0 - y_2)(y_1 - y_2)]^2 \stackrel{!}{=} -4p^3 - 27q^2 = -27\Delta. \tag{75}$$

So we have the following cases:

Case I. $\Delta = 0$. By Equation 76, this happens if and only if two of the roots coincide. But then this excludes the case of two complex conjugate roots (because they would be distinct; and since the third root must be real, the third root would also be different from the other two roots.) We can in fact be more concrete. We can directly factor $y^3 + py + q$, as

$$y^3 + py + q \stackrel{!}{=} y^3 - \frac{27q^2}{4p^2}y - \frac{27q^3}{4p^3} = \left(y - \frac{3p}{q}\right)\left(y + \frac{3q}{2p}\right)^2,$$

where for the marked step we used that $\frac{27q^2}{4p^3} = -1$ (since $\Delta = q^2 + \frac{4}{27}p^3 = 0$.) So summing up, in this case we have 3 real roots, (at least) two of which coincide.

Case II. $\Delta > 0$. In this case y_0 is real, whereas y_1 and y_2 are complex conjugate, non-real roots. To see this we argue by contradiction: Were all three roots real, then $(y-y_0)(y-y_1)(y-y_2)$ would be a real number, so its square would be nonnegative, so by Equation 76 $\Delta \leq 0$.

Case III. $\Delta < 0$. This case is usually called *casus irreducibilis* in the literature. In this case all three roots are real. To prove this, suppose by contradiction that $y_0 = c$ is real, whereas $y_1 = a + ib$, and $y_2 = a - ib$ are complex conjugate, with $a, b, c \in \mathbb{R}$, $b \neq 0$. Plugging in into Equation 76, we get

$$[(c - a - ib)(c - a + ib)(2ib)]^2 = -27\Delta,$$

or in other words,

$$[(c - a)^2 + b^2]^2(-4b^2) = -27\Delta,$$

which implies that $\Delta > 0$, a contradiction.

8.2 Degree-four equations

Also the solution to equations of degree four was published in 1545 by the Italian mathematician Girolamo Cardano; the trick is to reduce it to equations of degree three. Below is a solution due to Descartes. Writing down the full formula is too long, but we give the exact procedure to find all solutions.

We should start from an equation of the type

$$0 = ax^4 + bx^3 + cx^2 + dx + e.$$

By now you should be used to the next step. We leave it to you to prove that we can reduce ourselves to the case $a = 1$ and $b = 0$, by performing the change of variable

$$y \stackrel{\text{def}}{=} x + \frac{b}{4a}.$$

So there is no loss of generality if we focus on equations of the type

$$0 = y^4 + py^2 + qy + r.$$

Now comes Descartes' idea. We want to write the right hand side as product of two degree-two polynomials, which we can assume to be monic. If we manage to do this, then we reduced ourselves to a previously solved problem, because we know how to find roots of degree-two polynomials in $\mathbb{C}[x]$ (cf. Remark 602). So let's do it! We seek four complex numbers s, t, u, v such that

$$y^4 + py^2 + qy + r = (y^2 + sy + t) \cdot (y^2 + uy + v).$$

Equating the coefficients, we obtain the following system:

$$\begin{cases} 0 &= s + u \\ p &= t + v + su \\ q &= sv + tu \\ r &= tv. \end{cases} \tag{76}$$

From the first equation we can get rid of s, obtaining

$$\begin{cases} s &= -u \\ p + u^2 &= t + v \\ q &= u(t - v) \\ r &= tv. \end{cases}$$

Now, consider the quantity $u^2(p + u^2)^2 - q^2$. We have

$$u^2(p + u^2)^2 - q^2 = u^2(t + v)^2 - u^2(t - v)^2 = u^2 \cdot [(t + v)^2 - (t - v)^2] = u^2 \cdot 4tv = 4ru^2.$$

But the equation $u^2(p + u^2)^2 - q^2 = 4ru^2$ is *cubic in* u^2! If we do a change of variable $w = u^2$, it becomes

$$w^3 + 2p \cdot w^2 + (p^2 - 4r) \cdot w - q^2 = 0.$$

Which after the usual change of variable $z = w + \frac{2p}{3}$ becomes

$$z^3 - \frac{p^2 + 4r}{3} z - \frac{2p^3 - 72pr + 27q^2}{27} = 0.$$

Once we find z (by solving the cubic equation), we can figure out $w = z - \frac{2p}{3}$ and $u = \sqrt{z - \frac{2p}{3}}$. And once u is known, by System 77 we have

$$\begin{cases} s & = -u \\ (t+v) & = u^2 + p \\ (t-v) & = \frac{q}{u} \\ tv & = r, \end{cases}$$

and from the first three equations we can easily derive

$$\begin{cases} s & = -u \\ t & = \frac{1}{2}\left(u^2 + p + \frac{q}{u}\right) \\ v & = \frac{1}{2}\left(u^2 + p - \frac{q}{u}\right). \end{cases}$$

Now the formula for degree-two equations allows us to derive the two roots of $(y^2 + sy + t)$ and the two roots of $(y^2 + uy + v)$. From each of these four solutions, if we subtract $\frac{b}{4a}$, we obtain a root for the original quartic, $ax^4 + bx^3 + cx^2 + dx + e = 0$.

8.3 Degree-n equations with only two nonzero terms

Here we show how to solve equations of the type $tx^n + ux^m = 0$, with $t, u \neq 0$ complex numbers. Without loss of generality, we can assume:

- $m \neq n$ (if not, $x = 0$ is the only solution);
- $m < n$ (if not, swap m and n);
- $m = 0$ (if not, divide by x^m);
- $t = 1$ (if not, divide by t).

So without loss, we can focus on equations of the type

$$x^n = z, \qquad \text{with } z \neq 0.$$

To tackle this, let us write z in the so-called "polar form":

Lemma 605 (Polar form of z)**.** *Let $z \in \mathbb{C}$, $z \neq 0$. There is a unique way to write z as*

$$z = r(\cos x + i \sin x), \qquad \textit{with } r, x \in \mathbb{R}, r > 0, x \in [0, 2\pi).$$

Proof. [21] To check that this is always possible: if $z = a + ib$, with a, b in $\mathbb{R}$, choose

$$r \stackrel{\text{def}}{=} \sqrt{a^2 + b^2} \qquad \text{and} \qquad x \stackrel{\text{def}}{=} \begin{cases} \arccos \frac{a}{r} & \text{if } b \geq 0, \\ -\arccos \frac{a}{r} & \text{if } b < 0, \end{cases}$$

Either way, $\cos x = \frac{a}{r}$. Moreover,

$$\sin^2 x = 1 - \cos^2 x = 1 - \frac{a^2}{r^2} = \frac{a^2 + b^2}{r^2} - \frac{a^2}{r^2} = \frac{b^2}{r^2}.$$

This tells us that $|\sin x| = \frac{|b|}{r}$. To check that the absolute value is not needed, we go back to the definition of x. When b is positive, then $x \stackrel{\text{def}}{=} \arccos \frac{a}{r}$, so $x \in [0, \pi]$, which means that $\sin x$ is also positive. When b is negative, then $x \stackrel{\text{def}}{=} -\arccos \frac{a}{r}$, so $x \in [-\pi, 0]$, which means that $\sin x$

[21] There is a wrong formula that systematically appears in textbooks, namely, "if $z = a + ib$ set $r \stackrel{\text{def}}{=} \sqrt{a^2 + b^2}$ and $x \stackrel{\text{def}}{=} \arctan \frac{b}{a}$". This works only for x positive. For example, when $z = (-1, 1)$ one has $\arctan(-1) = -\frac{1}{4}\pi$, whereas the correct value for x (see the formula we give) is $x = \frac{3}{4}\pi$.

is negative. In both cases, $\sin x$ and $\frac{b}{r}$ have the same sign. So they are equal. This proves that with our choice of r and x, indeed $a = r\cos x$ and $b = r\sin x$.

To check uniqueness, suppose

$$r(\cos x + i\sin x) = r'(\cos x' + i\sin x').$$

Passing to the norms, we get $r^2 = r'^2$, and since they are both positive, $r = r'$. So we can divide out by r:

$$\cos x + i\sin x = \cos x' + i\sin x'.$$

By equating the real part, we get $\cos x = \cos x'$, so $x' = \pm x + 2k\pi$ for some integer k. But by equating the imaginary part we get $\sin x = \sin x'$; so the case $x' = -x+2k\pi$ should be discarded. Hence, $x' = x + 2k\pi$. Since x and x' are in the interval $[0, 2\pi)$, necessarily $x = x'$. □

Lemma 606 (de Moivre's formula)**.** *For any nonnegative integer n, for any real number x,*

$$(\cos x + i\sin x)^n = \cos nx + i\sin nx. \tag{77}$$

Proof. By induction. The cases $n = 0$ and $n = 1$ are obvious. The induction step follows from the trigonometric identities

$$\begin{aligned}\cos(x+y) &= \cos x\cos y - \sin x\sin y\\ \sin(x+y) &= \sin x\cos y + \cos y\sin x,\end{aligned}$$

applied to $y = nx$. More specifically,

$$\begin{aligned}(\cos x + i\sin x)^{n+1} &= (\cos x + i\sin x)(\cos x + i\sin x)^n = (\cos x + i\sin x)(\cos nx + i\sin nx) =\\ &= (\cos x\cos nx - \sin x\sin nx) + i(\sin x\cos nx + \cos x\sin x) \overset{!}{=}\\ &= \cos(x+nx) + i\sin(x+nx) = \cos(n+1)x + i\sin(n+1)x.\end{aligned}$$ □

Proposition 607. *If $z = r(\cos x + i\sin x)$ in polar form, then the complex numbers*

$$\sqrt[n]{r}\left(\cos\frac{x+2k\pi}{n} + i\sin\frac{x+2k\pi}{n}\right), \quad k = 0,\dots,n-1, \tag{78}$$

are the n roots of the polynomial $x^n - z$.

Proof. We claim that the list above does not contain repetitions as k varies from 0 to $n-1$. We argue by contradiction. Suppose there are two integers $h < k$ in $\{0,\dots,n-1\}$ such that

$$\sqrt[n]{r}\left(\cos\frac{x+2h\pi}{n} + i\sin\frac{x+2h\pi}{n}\right) = \sqrt[n]{r}\left(\cos\frac{x+2k\pi}{n} + i\sin\frac{x+2k\pi}{n}\right).$$

If two complex numbers are equal, their real (resp. their imaginary) parts coincide, so

$$\cos\frac{x+2h\pi}{n} = \cos\frac{x+2k\pi}{n} \qquad \text{and, respectively,} \qquad \sin\frac{x+2h\pi}{n} = \sin\frac{x+2k\pi}{n}.$$

So, for some integer ℓ,

$$\frac{x+2h\pi}{n} + 2\ell\pi = \frac{x+2k\pi}{n},$$

or equivalently $h + \ell n = k$. But h, k are in $\{0,\dots,n-1\}$, so $\ell = 0$ and $h = k$, a contradiction.

Hence Equation 79 lists n different complex numbers. These numbers are all roots of $x^n - z$, because of de Moivre's formula 78:

$$\left(\sqrt[n]{r}\left(\cos\frac{x+2k\pi}{n} + i\sin\frac{x+2k\pi}{n}\right)\right)^n = r(\cos(x+2k\pi)+i\sin(x+2k\pi)) = r(\cos x+i\sin x) = z.$$

Since a degree-n polynomial in $\mathbb{C}[x]$ (which is a domain) cannot have more than n roots, we conclude that Equation 79 is listing all of the roots of $x^n - z$. □

Corollary 608. *If r is a positive real number, the n complex roots of $x^n - r$ are*

$$\sqrt[n]{r} \cdot \xi^k, \qquad k = 0, \ldots, n-1, \qquad \text{where } \xi \stackrel{\text{def}}{=} \cos\frac{2\pi}{n} + i \sin\frac{2\pi}{n}.$$

Proof. In polar form $r = r \cdot (\cos 0 + i \sin 0)$, so if we plug in $x = 0$ into Equation 79, we get that the roots are $\sqrt[n]{r}\left(\cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}\right)$ (for $0 \le k \le n-1$), which by de Moivre's formula 78 is the same as $\sqrt[n]{r}\left(\cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}\right)^k$. □

Definition 609. Set $\xi = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$ as above. The n complex numbers

$$\xi^k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \qquad k = 0, \ldots, n-1,$$

are called the "n-th roots of unity". By Corollary 608, they are the roots of $x^n - 1$.

Some of these roots are more significant then others.

Definition 610. An "n-th roots of unity" u is called *primitive* if there is no positive integer $k < n$ such that u is a k-th root of unity.

For example, for any $n \ge 1$, the ξ^1 of Definition 609 is always primitive. In contrast, if $n = ab$ is the product of two integers larger than one, then ξ^a is not primitive, because

$$(\xi^a)^b = (\xi^1)^{ab} = (\xi^1)^n = 1.$$

8.4 The Galois group

In order to understand the roots of polynomials in $\mathbb{K}[x]$, with $\mathbb{K}$ a field, we take a more abstract approach. The idea is to give some kind of "algebraic structure" to the set of all roots.

Definition 611. Let f be a polynomial with coefficients in a field $\mathbb{K}$. The *splitting field (over $\mathbb{K}$) of f* is the smallest field (or more precisely, the smallest subfield of $\overline{\mathbb{K}}$) containing $\mathbb{K}$ and all roots of f.

Example 612. The splitting field of the polynomial $x^2 + 1$ over $\mathbb{R}$ is $\mathbb{C}$. The splitting field of x^2+1 over $\mathbb{Q}$, is $\mathbb{Q}(i)$. The splitting field of x^2+1 over $\mathbb{Z}_2$, is $\mathbb{Z}_2$ itself, because $x^2+1 = x^2-1 = (x+1)(x-1)$.

Non-Example 613. There's no polynomial in $\mathbb{Q}[x]$ with splitting field $Q(\sqrt[3]{2})$. (In fact, the minimum polynomial over $\mathbb{Q}$ of $\sqrt[3]{2}$ is $x^3 - 2$, which has other two complex roots).

Definition 614. Let $\mathbb{F}$ be a field containing $\mathbb{K}$. An *automorphism of $\mathbb{F}$ that fixes $\mathbb{K}$* is any bijective ring homomorphism $\varphi : \mathbb{F} \longrightarrow \mathbb{F}$ that maps any element of $\mathbb{K}$ to itself.

Definition 615 (Galois Group)**.** Let $\mathbb{K}$ be a field. Let f be a polynomial in $\mathbb{K}[x]$. The *Galois group of f (over $\mathbb{K}$)* is the group of all automorphisms of the splitting field of f that fix $\mathbb{K}$, with respect to the operation of composition.

Notational remark. It is obvious that the Galois group is really a group: Composing two automorphism that fix $\mathbb{K}$ one gets an automorphism that fixes $\mathbb{K}$ as well. Also, if f and g are two polynomials in $\mathbb{K}[x]$ with same splitting field, it is clear from the definition that the Galois group of f and g over $\mathbb{K}$ are the same. For this reason we usually omit f from the notation, and denote the Galois group of f over $\mathbb{K}$ by "Gal($\mathbb{F}/\mathbb{K}$)", where $\mathbb{F}$ is the splitting field of f.

Example 616. Consider $f = x^2+1$ in $\mathbb{R}[x]$. Its splitting field is $\mathbb{R}[i] = \mathbb{C}$. The conjugation from $\mathbb{C}$ to $\mathbb{C}$ is an automorphism that fixes $\mathbb{R}$, because the conjugate of any real number r is r itself. So $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ contains at least two elements: the identity and the conjugation. Does it contain any other map? To answer this, let us compute the value of any function φ in $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ on the generic element of $\mathbb{C}$, which can be written as $a + bi$ with $a, b \in \mathbb{R}$. Since φ is a homomorphism that fixes $\mathbb{R}$,

$$\varphi(a + bi) = \varphi(a) + \varphi(b)\varphi(i) = a + b\varphi(i).$$

Hence the value of φ on *any* complex number is completely determined by $\varphi(i)$. On the other hand, we know that $i^2 + 1 = 0$. So

$$\varphi(i)^2 + 1 = \varphi(i)\varphi(i) + \varphi(1) = \varphi(i^2 + 1) = \varphi(0) = 0.$$

So $\varphi(i)$ also satisfies the equation $x^2 + 1 = 0$! Which means that either $\varphi(i) = i$ (in which case φ is the identity) or $\varphi(i) = -i$ (in which case φ is the conjugate homomorphism). Being a two-element group, $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ is $\mathbb{Z}_2$.

This example immediately generalizes as follows:

Theorem 617. *Let f be polynomial of $\mathbb{K}[x]$, with $\mathbb{K}$ field. Let $\mathbb{F}$ be the splitting field of f.*

1. *Any function in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$ is completely determined by its value on the roots of f.*
2. *Any function in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$ maps roots to roots.*
   In particular, if f has n distinct roots, the Galois group is a subgroup of $\mathcal{S}_n$.
3. *If $f = g \cdot h$ with both g, h in $\mathbb{K}[x]$, then any function $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$ maps roots of g to roots of g, and roots of h to roots of h.*
   In particular, if g has ℓ roots, h has $n - \ell$ roots, and all these roots are distinct, then the Galois group of gh is a subgroup of $\mathcal{S}_\ell \times \mathcal{S}_{n-\ell}$.

Proof. Let $f = a_0 + a_1 x + \ldots + a_n x^n$, with a_i in $\mathbb{K}$. Let $\alpha_1, \ldots, \alpha_n$ be the roots of f in $\mathbb{F}$, listed any way you want, possibly with repetitions. Let φ be any function in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$.

1. Set $\mathbb{K}' \stackrel{\text{def}}{=} \mathbb{K}(\alpha_1, \ldots, \alpha_{n-1})$ and $\alpha \stackrel{\text{def}}{=} \alpha_n$. So $\mathbb{F} = \mathbb{K}(\alpha_1, \ldots, \alpha_n) = \mathbb{K}'(\alpha)$, with α algebraic over $\mathbb{K}'$. Let d be the degree of the minimum polynomial of α over $\mathbb{K}'[x]$. By Theorem 398 we have that $d = [\mathbb{K}'(\alpha) : \mathbb{K}']$, and any element w of $\mathbb{F} = \mathbb{K}'(\alpha)$ can be uniquely written as

   $$w = \sum_{i=0}^{d-1} \lambda_i \cdot \alpha^i, \qquad \text{with } \lambda_i \in \mathbb{K}'.$$

   Applying φ and using the fact that it is a homomorphism, we get that

   $$\varphi(w) = \sum_{i=0}^{d-1} \varphi(\lambda_i) \cdot (\varphi(\alpha))^i.$$

   Hence the value of $\varphi(w)$ is completely determined, once we know $\varphi(\alpha)$ and the restriction of φ to $\mathbb{K}' = \mathbb{K}(\alpha_1, \ldots, \alpha_{n-1})$. Repeating this argument on $\mathbb{K}'$, we obtain that the value of φ on w is completely determined once we know $\varphi(\alpha_n)$, $\varphi(\alpha_{n-1})$, and the restriction of φ to $\mathbb{K}(\alpha_1, \ldots, \alpha_{n-2})$. And so on. Iterating n times, we reach the conclusion that $\varphi(w)$ is determined once we know $\varphi(\alpha_n)$, $\varphi(\alpha_{n-1})$, …, $\varphi(\alpha_1)$, and the restriction of φ to $\mathbb{K}$ (which is the identity function.)

(2) Since φ fixes $\mathbb{K}$, for all i we have $\varphi(a_i) = a_i$. Since φ is a homomorphism,

$$f(\varphi(\alpha_1)) = a_0 + a_1\varphi(\alpha_1) + \ldots + a_n(\varphi(\alpha_1))^n = \varphi(a_0 + a_1\alpha_1 + \ldots + a_n\alpha_1^n) = \varphi(0) = 0.$$

Hence, φ maps the root α_1 to some other root (possibly, α_1 itself).
So if f has n distinct roots, we can represent φ by a permutation in $\mathcal{S}_n$. But by item (1) above, this representation is unique. This yields an injective function from $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$ to $\mathcal{S}_n$. (Note: We are not claiming that *all* permutations of $\mathcal{S}_n$ arise this way; this is false, see e.g. Example 619, or compare the statement of item (3) with the following count: There are $5! = 120$ permutations in $\mathcal{S}_5$, but only $2! \cdot 3! = 12$ in $\mathcal{S}_2 \times \mathcal{S}_3$...)

(3) As above, from $g(\alpha_1) = 0$, one deduces that $g(\varphi(a_1)) = 0$; and same for h. Then f can be represented by a pair of permutations: one in $\mathcal{S}_\ell$ (keeping track of how the roots of g are shuffled) and one in $\mathcal{S}_{n-\ell}$ (keeping track of how the roots of h are shuffled). □

Corollary 618. *The Galois group of any polynomial is finite.*

Example 619. Consider the polynomial $f = (x^2 - 3)(x^2 - 2)$ over $\mathbb{Q}$. The smallest field containing $\mathbb{Q}$ and all roots of f is $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2} + \sqrt{3})$, which is an extension of degree four over $\mathbb{Q}$. The generic element of it can be written as $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{2}\sqrt{3}$. But for any automorphism φ that fixes $\mathbb{Q}$,

$$\begin{aligned}\varphi(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{2}\sqrt{3}) &= \varphi(a) + \varphi(b)\varphi(\sqrt{2}) + \varphi(c)\varphi(\sqrt{3}) + \varphi(d)\varphi(\sqrt{2})\varphi(\sqrt{3}) = \\ &= a + b\varphi(\sqrt{2}) + c\varphi(\sqrt{3}) + d\varphi(\sqrt{2})\varphi(\sqrt{3}),\end{aligned}$$

so φ is completely determined once we know $\varphi(\sqrt{2})$ and $\varphi(\sqrt{3})$, in agreement with what Theorem 617 says. On the other hand, by Theorem 617, φ maps roots to roots; and the roots are $\pm\sqrt{2}$ and $\pm\sqrt{3}$. Also, $f = g \cdot h$, with $g = x^2 - 2$ and $h = x^2 - 3$. So φ cannot send $\sqrt{2}$ to $\sqrt{3}$, because it must also send roots of g to roots of g, and roots of h to roots of h. Long story short, there are four possible options:

$$\begin{array}{lccc} \varphi_0: & \mathbb{Q}(\sqrt{2},\sqrt{3}) & \to \mathbb{Q}(\sqrt{2},\sqrt{3}) & \qquad \varphi_1: \quad \mathbb{Q}(\sqrt{2},\sqrt{3}) \to \mathbb{Q}(\sqrt{2},\sqrt{3}) \\ & \sqrt{2} & \mapsto +\sqrt{2} & \qquad\qquad\quad \sqrt{2} \mapsto +\sqrt{2} \\ & \sqrt{3} & \mapsto +\sqrt{3} & \qquad\qquad\quad \sqrt{3} \mapsto -\sqrt{3} \\ \varphi_2: & \mathbb{Q}(\sqrt{2},\sqrt{3}) & \to \mathbb{Q}(\sqrt{2},\sqrt{3}) & \qquad \varphi_3: \quad \mathbb{Q}(\sqrt{2},\sqrt{3}) \to \mathbb{Q}(\sqrt{2},\sqrt{3}) \\ & \sqrt{2} & \mapsto -\sqrt{2} & \qquad\qquad\quad \sqrt{2} \mapsto -\sqrt{2} \\ & \sqrt{3} & \mapsto +\sqrt{3} & \qquad\qquad\quad \sqrt{3} \mapsto -\sqrt{3} \end{array}$$

Note that $\varphi_3 = \varphi_1 \circ \varphi_2 = \varphi_2 \circ \varphi_1$. So the set $\mathrm{Aut}(\mathbb{F}/\mathbb{K})$, with respect to the operation of composition, forms a group isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. (The explicit group isomorphism is $\psi(\varphi_0) = (0,0)$, $\psi(\varphi_1) = (1,0)$, $\psi(\varphi_2) = (0,1)$, $\psi(\varphi_3) = (1,1)$.)

Proposition 620. *Let $p \geq 3$ be a prime number. The Galois group of $x^p - 1 = (x-1)(x^{p-1} + x^{p-2} \ldots + x + 1)$ over $\mathbb{Q}$ is cyclic and isomorphic to $(\mathbb{Z}_{p-1}, +)$.*

Proof. Recall, that by Corollary 608 the "p-th roots of unity" are the p distinct complex numbers

$$1,\ \xi,\ \xi^2,\ \ldots,\ \xi^{p-1}, \quad \text{where } \xi \stackrel{\text{def}}{=} \left(\cos\frac{2\pi}{p} + i\sin\frac{2\pi}{p}\right).$$

This means that over $\mathbb{Q}(\xi)$, the polynomial $x^p - 1$ splits as

$$x^p - 1 \ = \ (x-1)(x-\xi)(x-\xi^2)\cdots(x-\xi^{p-1}).$$

So $\mathbb{Q}(\xi)$ is the splitting field of $x^p - 1$. Hence any φ in $\mathrm{Gal}(\mathbb{Q}(\xi)/\mathbb{Q})$ is completely determined by its value on ξ. On the other hand, we know that $\xi^p - 1 = 0$. So

$$\varphi(\xi)^p - 1 = \varphi(\xi^p - 1) = \varphi(0) = 0.$$

This means that φ must be one of the other roots. Define φ_i to be the unique map that sends ξ to ξ^i. Then it is easy to see that for each i,

$$\varphi_i = (\varphi_1 \circ \varphi_1 \circ \ldots \circ \varphi_1) \qquad (i \text{ times.})$$

So the Galois group is cyclic, generated by φ_1. The isomorphism with $(\mathbb{Z}_{p-1}, +)$ is easily obtained by mapping the function φ_i to the element $i \in \mathbb{Z}_{p-1}$. □

Remark 621. The Galois group of $x^p - 1$ is the same as that of $x^{p-1} + x^{p-2} \ldots + x + 1$; the latter is irreducible over $\mathbb{Q}$.

Corollary 622. *Let p, q be distinct prime numbers. The Galois group of the (reducible) polynomial $(x^{p-1} + x^{p-2} \ldots + x + 1)(x^{q-1} + x^{q-2} \ldots + q + 1)$ is $\mathbb{Z}_{p-1} \times \mathbb{Z}_{q-1}$.*

So far, we have never said how many elements the Galois group should have. Theorem 617 (together with Lagrange theorem 464) immediately tells us that for polynomials with n distinct roots, the cardinality of the Galois group divides $n!$, which is the cardinality of $\mathcal{S}_n$. But maybe we can say more:

- Are there cases in which the Galois group really is $\mathcal{S}_n$? (For this f has to be irreducible, otherwise Theorem 617, item (3), gives a contradiction.)
- If f is irreducible, is it true that the Galois group always has at least n elements, like in the examples above? (Again, f has to be irreducible, otherwise $x^3 - 1$ is a counterexample by Proposition 620. But as we saw in Remark 621, $x^3 - 1 = (x-1)(x^2 + x + 1)$, and $\mathbb{Z}_2$ is actually the Galois group of $x^2 + x + 1$.)

We are now going to show that the cardinality of the Galois group is exactly $[\mathbb{F} : \mathbb{K}]$. For this we need to add an extra assumption on the field $\mathbb{K}$:

Definition 623. A field $\mathbb{K}$ has *characteristic zero* if the ring homomorphism from $\mathbb{Z}$ to $\mathbb{K}$ that sends 1 to 1 is injective.

For example, $\mathbb{Q}$ and all of its extensions have characteristic zero; $\mathbb{Z}_p$ and its extensions do not. This assumption is crucial to avoid pathologies like the one below.

Non-Example 624. Let p be any prime number. Set $\tau \stackrel{\text{def}}{=} \pi^p$. Consider the two (non-algebraic) extension $\mathbb{F} \stackrel{\text{def}}{=} \mathbb{Z}_p(\pi)$ and $\mathbb{G} \stackrel{\text{def}}{=} \mathbb{Z}_p(\tau)$ of $\mathbb{Z}_p$. Clearly $\mathbb{Z}_p \subseteq \mathbb{G} \subseteq \mathbb{F}$. Then $\pi \in \mathbb{F}$ is transcendental over $\mathbb{Z}_p$, but it is algebraic over $\mathbb{G}$; one can see that its minimum polynomial is $x^p - \tau$. Now, this minimum polynomial is irreducible in $\mathbb{G}[x]$, but it is indeed reducible in $\mathbb{F}[x]$ as

$$x^p - \tau = x^p - \pi^p = (x - \pi)^p.$$

So in its splitting field, the polynomial has only one root. Hence in this case the cardinality of the Galois group $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ is one, even if $[\mathbb{F} : \mathbb{K}] = p$.

From now on, we shall work with fields of characteristic zero. For those we have:

Lemma 625 ("Separability")**.** *Let $\mathbb{K}$ be a field of characteristic zero. Let $f \in \mathbb{K}[x]$. If f is irreducible in $\mathbb{K}[x]$, then in the splitting field of f all roots are distinct.*

Proof. Recall that the derivative of a polynomial $f = a_0 + a_1 x + \ldots + a_n x^n$ is the polynomial $D(f) \stackrel{\text{def}}{=} a_1 + 2a_2 x + \ldots + n a_n x^{n-1}$, and our characteristic-zero assumption guarantees that $n a_n \neq 0$. Since f is irreducible, $\gcd(f, D(f)) = 1$, so by Proposition 419 we conclude. □

Lemma 626 (Primitive element)**.** *Let a, b be algebraic over a field $\mathbb{K}$ of characteristic zero. For some element c algebraic over $\mathbb{K}$ (called "primitive element") one has $\mathbb{K}(a, b) = \mathbb{K}(c)$.*

Proof á-la-Van der Waerden. The idea is to show that setting $c \stackrel{\text{def}}{=} a + tb$, one has always that $\mathbb{K}(c)$ contains b (and hence also $a = c - tb$), except for finitely many values of $t \in \mathbb{K}$. Since $\mathbb{K}$ is infinite, this finite amount of exceptions does not constitute an obstacle. So, let f (respectively, g) be the minimum polynomial of a (respectively, of b) over $\mathbb{K}$. Set $c \stackrel{\text{def}}{=} a + tb$ and define the new polynomial

$$H(x) \stackrel{\text{def}}{=} f(c - tx) \ \in \mathbb{K}(c)[x].$$

By definition, $H(b) = f(c - tb) = f(a) = 0$. So b is a root of H. Let now h be the minimum polynomial of b **over** $\mathbb{K}(\mathbf{c})$. Clearly h divides H. Also, h will divide g, which is the minimum polynomial of b **over** $\mathbb{K}$. So h divides $\gcd(g, H)$ and

$$[\mathbb{K}(b, c) : \mathbb{K}(c)] = \deg h \leq \gcd(g, H).$$

We now claim that "choosing t randomly" (or more precisely, outside a certain finite set) one has $\deg \gcd(g, H) = 1$. This directly implies that $[\mathbb{K}(b, c) : \mathbb{K}(c)] = 1$, i.e. that $b \in \mathbb{K}(c)$, as desired. By contradiction, suppose $\deg \gcd(g, H) \geq 2$. This means that g and H have another root in common, different than b (cf. Lemma 625). Call it b'. Since $f(c - tb') = H(b') = 0$, the element $c - tb'$ is one of the roots of f different than a. Call $a' \stackrel{\text{def}}{=} c - tb'$. But then $a' + tb' = c = a + tb$, which means that

$$t = \frac{a - a'}{b' - b}.$$

Now, consider the subset S of $\mathbb{K}$ formed by the ratios $\frac{a-a'}{b'-b}$, where a' ranges among the $\deg f - 1$ roots of f different from a, and b' ranges among the $\deg g - 1$ roots of g different from b. Clearly, S is finite, because the expression $\frac{a-a'}{b'-b}$ can attain at most $(\deg f - 1)(\deg g - 1)$ different values. Yet $\mathbb{K}$ is infinite. So if we choose t "randomly", it won't be in S; and since $\deg \gcd(g, H) \geq 2$ leads to a contradiction, we conclude that $\deg \gcd(g, H) = 1$. □

Definition 627 (Normal)**.** A finite field extension $\mathbb{F}$ of a field with characteristic zero $\mathbb{K}$ is called *normal over* $\mathbb{K}$ if any irreducible polynomial $p \in \mathbb{K}[x]$ that has a root in $\mathbb{F}$, has all roots in $\mathbb{F}$.

Lemma 628 (Normality of splitting fields)**.** *Let $\mathbb{F}$ be a finite field extension of $\mathbb{K}$, a field of characteristic zero. The following are equivalent:*

(1) *$\mathbb{F}$ is the splitting field of some polynomial $f \in \mathbb{K}[x]$.*

(2) *$\mathbb{F}$ is normal over $\mathbb{K}$.*

Proof (Sketch). By Lemma 626, applied iteratively, $\mathbb{F} = \mathbb{K}(c)$ for some c in $\mathbb{F}$.

(2) ⇒ (1). Let p be the minimum polynomial of c over $\mathbb{K}$. Since it has a root in $\mathbb{F}$ (namely, c), by the assumption p splits into linear factors over $\mathbb{F}$. Also, any field in which p splits into linear factors, must contain c and therefore $\mathbb{K}(c)$. So $\mathbb{K}(c)$ is the splitting field of p.

(1) ⇒ (2). Let p be an irreducible polynomial with a root α in $\mathbb{F}$. Let β be any other root of p; we want to prove that also $\beta \in \mathbb{F}$. Let $\mathbb{F}'$ be the splitting field of the product fp. By definition, $\mathbb{F}' \supset \mathbb{F}$. Now choose an injective homomorphism $\psi : \mathbb{K}(\alpha) \to \mathbb{F}'$ that sends

α to β while fixing $\mathbb{K}$. (This ψ is an element of the Galois group of the polynomial fp.) Now, extend this homomorphism to an injective homomorphism $\jmath : \mathbb{F} \to \mathbb{F}''$, for some finite extension $\mathbb{F}''$ of $\mathbb{F}'$. We claim that

$$\jmath(\mathbb{F}) = \mathbb{F}.$$

From this the conclusion follows immediately, because then $\beta = \jmath(\alpha) \in \jmath(\mathbb{F})$, so $\beta \in \mathbb{F}$. To prove the claim, note first that $\jmath(\mathbb{F})$ is a field. Since $\jmath$ is an extension of ψ and ψ fixes $\mathbb{K}$, then also $\jmath$ fixes $\mathbb{K}$, so $\jmath(\mathbb{F})$ is a field containing $\mathbb{K}$. Also, since $f \in \mathbb{K}[x]$, one has $\jmath(f) = f$. But then inside $\jmath(\mathbb{F})$ the polynomial f splits into linear factors, because if $f = (x-\gamma_1)(x-\gamma_2)\cdots(x-\gamma_n)$ inside $\mathbb{F}$, then

$$f = \jmath(f) = (x-\jmath(\gamma_1))(x-\jmath(\gamma_2))\cdots(x-\jmath(\gamma_n)).$$

And finally, there is no proper subfield of $\jmath(\mathbb{F})$ over which f would split into linear factors, because $\jmath$ is injective. In conclusion, $\jmath(\mathbb{F})$ is the splitting field of f over $\mathbb{K}$. □

Theorem 629. *Let f be polynomial of $\mathbb{K}[x]$, with $\mathbb{K}$ field of characteristic zero. Let $\mathbb{F}$ be the splitting field of f. Then*

$$\#\operatorname{Gal}(\mathbb{F}/\mathbb{K}) \;=\; [\mathbb{F}:\mathbb{K}].$$

Proof. By Lemma 626, $\mathbb{F} = \mathbb{K}(c)$ for some c in $\mathbb{F}$. So if $d \stackrel{\text{def}}{=} [\mathbb{F}:\mathbb{K}] = [\mathbb{K}(c):\mathbb{K}]$, by Theorem 398 we have that the minimum polynomial p of c over $\mathbb{K}$ has degree d, and the powers $1, c, c^2, \ldots, c^{d-1}$ form a basis for $\mathbb{F}$ over $\mathbb{K}$. On the other hand, by Lemma 628 the field $\mathbb{F}$ is normal, so *all* roots of p are in $\mathbb{F}$. By Lemma 625, since p is irreducible, its d roots are all distinct. Let us call them $c = \alpha_1, \alpha_2, \ldots, \alpha_d$. Then we can define d different elements $\phi_1, \ldots, \phi_d$ of $\operatorname{Gal}(\mathbb{F}/\mathbb{K})$ by

$$\begin{array}{rccc} \phi_i : & \mathbb{F} & \longrightarrow & \mathbb{F} \\ & c & \longmapsto & \alpha_i. \end{array}$$

This proves that $\#\operatorname{Gal}(\mathbb{F}/\mathbb{K}) \geq d$. The other inequality is easier: For every ϕ in $\#\operatorname{Gal}(\mathbb{F}/\mathbb{K})$, we have

$$p(\phi(c)) = \phi(p(c)) = \phi(0) = 0.$$

So $\phi(c)$ must be one of the roots of p. Since p has degree d, it follows that there are at most d elements in $\operatorname{Gal}(\mathbb{F}/\mathbb{K})$. □

We conclude exhibiting an example of polynomial with Galois group $\mathcal{S}_5$.

Lemma 630. *The polynomial $f = x^5 - 4x^4 + 2x + 2 \in \mathbb{Q}[x]$ has exactly 3 real roots.*

Proof. The polynomial is irreducible by Eisenstein's criterion (theorem 352) and hence has five distinct roots in $\mathbb{C}$ by Lemma 625. Since

$$f(-1) = -5 < 0, \quad f(0) = 2 > 0, \quad f(2) = -26 < 0, \quad f(4) = 10 > 0,$$

by the intermediate value theorem there are at least three real roots $x_1 \in (-5, 0)$, $x_2 \in (0, 2)$, $x_3 \in (2, 4)$. On the other hand, the first derivative is $f' = 5x^4 - 16x^3 + 2$, the second one is $f'' = 20x^3 - 48x^2 = 4x^2(5x - 12)$. Thus f'' vanishes in 0 and in $\frac{12}{5}$. By Rolle's theorem, between any two zeroes of f' there must be a zero for f''. It follows that f' has at most three zeroes in $\mathbb{R}$. But f' has degree four, so it cannot have *exactly* three zeroes in $\mathbb{R}$ (because if a complex number were a root of f', so would its conjugate). So f'' has at most two zeroes in $\mathbb{R}$, and by Rolle's theorem, f has at most three zeroes in $\mathbb{R}$. Hence, f has exactly three real roots and two complex conjugate ones. □

Proposition 631. *The Galois group of the polynomial* $f = x^5 - 4x^4 + 2x + 2 \in \mathbb{Q}[x]$ *is* $\mathcal{S}_5$.

Proof. We know that the Galois group G is a subgroup of $\mathcal{S}_5$. By Proposition 494, it suffices to show that the Galois group contains a 5-cycle and a flip.

- Let $\alpha_1, \ldots, \alpha_5$ be the five roots of f. Let $\mathbb{G} \stackrel{\text{def}}{=} \mathbb{Q}(\alpha_1, \ldots, \alpha_5)$. Since f is irreducible, it is the minimum polynomial of any of its roots. Hence, $[\mathbb{Q}(\alpha_i) : \mathbb{Q}] = 5$ for all i. Then by theorem 629 and by Lagrange's theorem 394 we have

$$\#G = [\mathbb{G} : \mathbb{Q}] = [\mathbb{G} : \mathbb{Q}(\alpha_1)] \cdot [\mathbb{Q}(\alpha_1) : \mathbb{Q}] = [\mathbb{G} : \mathbb{Q}(\alpha_1)] \cdot 5.$$

  So by Theorem 527, the group G contains an element of order 5. But within $\mathcal{S}_5$, an element of order 5 is just a 5-cycle.
- By Lemma 630, f has exactly two non-real (complex) roots, necessarily conjugate of one another. Then the conjugate homomorphism exchanges these two roots, while fixing the other three roots and the whole of $\mathbb{Q}$. So clearly the conjugate homomorphism is a flip in $\text{Gal}(\mathbb{G}, \mathbb{Q})$. □

Remark 632. The previous proof shows a more general fact: any irreducible degree-5 polynomial of $\mathbb{Q}[x]$ with exactly *three* real roots, has Galois group $\mathcal{S}_5$. But what about irreducible polynomials of $\mathbb{Q}[x]$ with exactly *one* real root?

The situation here is more complicated. First of all, there are now two pairs of complex conjugate roots, so if α_1 is the real root, the conjugate homomorphism in $\text{Gal}(\mathbb{G}, \mathbb{Q}) \subseteq \mathcal{S}_5$ is the product of two disjoint flips $(2,3)(4,5)$. There are several subgroups of $\mathcal{S}_5$ that strictly contain the 5-cycle: $\mathcal{S}_5$, A_5, six groups isomorphic to the dihedral group $D_{10} \stackrel{\text{def}}{=} < (12345), (23)(45) >$, and six further groups isomorphic to the general affine group $GA(1,5) \stackrel{\text{def}}{=} < (12345, 2354) >$. All of these groups can appear:

- $f(x) = x^5 + 20x + 4$ (irreducible[22] by the modulo-3 criterion, Proposition 346) has Galois group $\mathcal{S}_5$.
- $g(x) = x^5 + 20x + 16$ (irreducible by the modulo-3 criterion, in fact $\overline{g} = \overline{f}$) has Galois group A_5.
- $h(x) = x^5 - 2x^4 + x^3 + x^2 - x + 1$ (again irreducible by the modulo-3 criterion) has Galois group D_{10}.
- $i(x) = x^5 - 5$ (irreducible by Eisenstein) has Galois group $GA(1,5)$.

All three polynomials have exactly one real root. This is obvious for f, g and i, because they are increasing functions (their derivative is positive). For h, note that $h'(x) = 0$ has two solution, a local max in $m_1 \approx -0.49$ and a local min in $m_2 \approx 0.46$. Yet $h(m_2) > 0$. So h has exactly one zero in the interval $(-\infty, m_1)$, and no zero elsewhere.

This does not exhaust all possible degree-5 irreducible polynomials. In fact, don't forget Corollary 364: An irreducible polynomial in $\mathbb{Q}[x]$ could also have *five* real roots...

Deeper thoughts 633. Given a field $\mathbb{F}$, there is a quantity associated to every polynomial

$$f = a_0 + a_1 x + \ldots + a_n x^n \in \mathbb{F}[x],$$

called *discriminant*, and defined in terms of the roots $\alpha_1, \ldots, \alpha_n$ by

$$D \stackrel{\text{def}}{=} (a_n)^{2n-2} \cdot \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

[22] The irreducible quadratic polynomials of $\mathbb{Z}_3[x]$ are $x^2 + 1$, $x^2 + x - 1$, $x^2 - x - 1$. So all you need to check irreducibility mod 3, is the lack of roots, plus the fact that the Euclidean division of f by each of the three polynomials above gives a nonzero remainder.

A priori $D \in \overline{\mathbb{F}}$, but one can prove that D belongs always to $\mathbb{F}$. In case $\mathbb{F} = \mathbb{Q}$, furthermore, if the discriminant is a perfect square then the Galois group of the polynomial is contained in A_5; otherwise, it is $\mathcal{S}_5$. (This should suggest you that having Galois group $\mathcal{S}_5$ is rather common). In case of degree-5 polynomials of the form $x^5 + dx + e$, the discriminant is easy to compute: It is $2^8 d^5 + 5^5 e^4$.

Deeper thoughts 634. If you want to know the Galois group of a specific polynomial, say,

$$f = x^5 - 15x^4 + 85x^3 - 225x^2 + 274x - 119,$$

go to the website http://magma.maths.usyd.edu.au/calc/, type in the following four-line program

```
P< x >:=PolynomialAlgebra(Rationals());
f:=x^5-15*x^4+85*x^3-225*x^2+274*x-119;
G:=GaloisGroup(f);
print G;
```

and hit "Submit".

8.5 The Galois Correspondence

In this section we prove the so-called "fundamental theorem of Galois theory", that reveals a correspondence between subgroups of $\mathcal{S}_n$ and field extensions. This correspondence will also reveal why "normal field extensions" have been called "normal": The reason is that they correspond to normal subgroups.

Lemma 635. *Let $\mathbb{K} \subseteq \mathbb{G} \subseteq \mathbb{F}$ be three fields of characteristic zero. If $\mathbb{F}$ is the splitting field of some polynomial f of $\mathbb{K}[x]$, it is also the splitting field of some polynomial of $\mathbb{G}[x]$.*

Proof. Choose the same polynomial f, and view it as an element of $\mathbb{G}[x]$. Clearly f splits in $\mathbb{F}$ and does not split in any proper subfield. □

Theorem 636 (Fundamental theorem of Galois theory)**.** *Let f be a polynomial of $\mathbb{K}[x]$, where $\mathbb{K}$ is any field of characteristic zero. Let $\mathbb{F}$ be the splitting field of f. Then:*

1. *For any 'intermediate field' $\mathbb{K} \subseteq \mathbb{G} \subseteq \mathbb{F}$, $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ is a subgroup of $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$, with cardinality $[\mathbb{F} : \mathbb{G}]$. In particular, if $\mathrm{Gal}(\mathbb{F}/\mathbb{G}) = \mathrm{Gal}(\mathbb{F}/\mathbb{K})$, then $\mathbb{K} = \mathbb{G}$.*
2. *For any subgroup H of $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$, the set $\mathrm{Fix}(H)$ of elements of $\mathbb{F}$ that are fixed by all automorphism in H is an intermediate field $\mathbb{K} \subseteq \mathrm{Fix}(H) \subseteq \mathbb{F}$.*
   In particular, when $H = \mathrm{Gal}(\mathbb{F}/\mathbb{K})$, one has $\mathrm{Fix}(H) = \mathbb{K}$.
3. *The two correspondences $\mathbb{G} \mapsto \mathrm{Gal}(\mathbb{F}/\mathbb{G})$ and $H \mapsto \mathrm{Fix}(H)$ described above are inverses of one another, in the sense that*
$$\mathrm{Fix}(\mathrm{Gal}(\mathbb{F}/\mathbb{G})) = \mathbb{G} \qquad \text{and} \qquad \mathrm{Gal}(\mathbb{F}/\mathrm{Fix}(H)) = H.$$
4. $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ *is a* normal *subgroup of* $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$ *if and only if* $\mathbb{G}$ *is a* normal *extension of* $\mathbb{K}$. *In this case,*
$$\mathrm{Gal}(\mathbb{F}/\mathbb{K}) \Big/ \mathrm{Gal}(\mathbb{F}/\mathbb{G}) \cong \mathrm{Gal}(\mathbb{G}/\mathbb{K}).$$

Proof.

(1) If φ is in $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$, it is an automorphism of $\mathbb{F}$ that fixes all elements of $\mathbb{G}$, so in particular it fixes all elements of $\mathbb{K}$. Hence, φ is in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$. This shows that $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ is a subgroup of $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$. Now, if $\mathbb{F}$ is the splitting field of some polynomial f of $\mathbb{K}[x]$, it is also the splitting field of some polynomial of $\mathbb{G}[x]$, by Lemma 635. So by Theorem 629 $\#\ \mathrm{Gal}(\mathbb{F}/\mathbb{G}) = [\mathbb{F} : \mathbb{G}]$. In particular, if $\mathrm{Gal}(\mathbb{F}/\mathbb{G}) = \mathrm{Gal}(\mathbb{F}/\mathbb{K})$, then $[\mathbb{F} : \mathbb{G}] = [\mathbb{F} : \mathbb{K}]$, so by Lagrange's theorem 394 we conclude that $[\mathbb{G} : \mathbb{K}] = 1$.

(2) Suppose x and y are in $\mathrm{Fix}(H)$. Then for any automorphism φ in H we have $\varphi(x) = x$ and $\varphi(y) = y$. So obviously

$$\begin{array}{lll} \varphi(x+y) & = \varphi(x) + \varphi(y) & = x + y, \\ \varphi(x \cdot y) & = \varphi(x) \cdot \varphi(y) & = x \cdot y, \end{array}$$

which means that $x+y$ and $x \cdot y$ are also in $\mathrm{Fix}(H)$. Moreover, $\mathrm{Fix}(H)$ contains the whole $\mathbb{K}$. (In fact, H is contained in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$, so any φ in H fixes $\mathbb{K}$.) In particular, $\mathrm{Fix}(H)$ contains the elements 0 and 1. This means that for any automorphism φ in H we have $\varphi(0) = 0$ and $\varphi(1) = 1$. But then, for any z in $\mathrm{Fix}(H)$,

$$\begin{array}{llll} \varphi(-z) + z & = \varphi(-z) + \varphi(z) & = \varphi(-z+z) & = \varphi(0) \quad = 0, \\ \varphi(z^{-1}) \cdot z & = \varphi(z^{-1}) \cdot \varphi(z) & = \varphi(z^{-1} \cdot z) & = \varphi(1) \quad = 1, \end{array}$$

which show that $\varphi(-z) = -z$ and $\varphi(z^{-1}) = z^{-1}$: Therefore, $-z$ and z^{-1} are also fixed by φ. But this holds for any φ in H, so by definition $-z$ and z^{-1} are in $\mathrm{Fix}(H)$. Hence $\mathrm{Fix}(H)$ is a field.

Now suppose $H = \mathrm{Gal}(\mathbb{F}/\mathbb{K})$. By contradiction, assume $\mathbb{K} \subsetneq \mathrm{Fix}(H)$. This means that there is an element e outside $\mathbb{K}$ that is fixed by all automorphisms of $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$. On the other hand, by Lemma 626, $\mathbb{F} = \mathbb{K}(c)$ for some c in $\mathbb{F}$. So if $d \stackrel{\text{def}}{=} [\mathbb{F} : \mathbb{K}] = [\mathbb{K}(c) : \mathbb{K}]$, by Theorem 398 we have that the minimum polynomial p of c over $\mathbb{K}$ has degree d, and the powers $1, c, c^2, \ldots, c^{d-1}$ form a basis for $\mathbb{F}$ over $\mathbb{K}$. So we can write

$$e = \sum_{j=0}^{d-1} k_j \cdot c^j \text{ for suitable coefficients } k_j \in \mathbb{K}.$$

By construction, $p(c) = 0$. More generally, let $\alpha_1 = c, \alpha_2, \ldots, \alpha_d$ be the d roots of the minimum polynomial p. The roots are all distinct by Lemma 625 and all in $\mathbb{F}$ by Lemma 628. We have seen in Theorem 629 that the elements of the Galois group $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$ are the maps $\varphi_1, \ldots, \varphi_d$, where φ_i is the map that that sends c to α_i. Now fix i in $\{1, \ldots, d\}$: Since e is fixed by all automorphisms of $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$,

$$e = \varphi_i(e) = \sum_{j=0}^{d-1} \varphi_i(k_j) \cdot (\varphi_i(c))^j = \sum_{j=0}^{d-1} k_j \cdot (\alpha_i)^j.$$

Since this holds for all i, the polynomial of $\mathbb{F}[x]$

$$f(x) \stackrel{\text{def}}{=} -e + \sum_{i=0}^{d-1} k_i \cdot x^i$$

has d zeroes: Namely, $f(\alpha_i) = 0$ for all $i \in \{0, \ldots, d-1\}$. But then f is the zero polynomial, because otherwise it would be a polynomial of degree $< d$, with d roots, in a domain. Hence, all coefficients of f are zero, except for the constant term f_0, which must be equal to e. But then e belongs to $\mathbb{K}$, a contradiction.

(3) If $\mathbb{F}$ is the splitting field of some polynomial f of $\mathbb{K}[x]$, it is also the splitting field of some polynomial of $\mathbb{G}[x]$, by Lemma 635. So applying the second part of item (2), with $\mathbb{G}$ playing the role of $\mathbb{K}$, we immediately get that $\mathrm{Fix}(\mathrm{Gal}(\mathbb{F}/\mathbb{G})) = \mathbb{G}$. The containment $\mathrm{Gal}(\mathbb{F}/\mathrm{Fix}(H)) \supset H$ is obvious, because by definition any homomorphism in H fixes all elements of $\mathrm{Fix}(H)$. To prove equality, let's just show that these subgroups have equal cardinalities. Suppose H consists of h automorphisms $\sigma_1 = id, \sigma_2, \ldots, \sigma_h$. By theorem 629, $\mathrm{Gal}(\mathbb{F}, \mathrm{Fix}(H))$ has exactly $[\mathbb{F} : \mathrm{Fix}(H)]$ elements. So we already know $h \leq [\mathbb{F} : \mathrm{Fix}(H)]$ and we need to show that $h \geq [\mathbb{F} : \mathrm{Fix}(H)]$. By Lemma 626, $\mathbb{F} = \mathrm{Fix}(H)(c)$ for some c in $\mathbb{F}$. So $[\mathbb{F} : \mathrm{Fix}(H)]$ is the degree of the minimum polynomial of c over $\mathrm{Fix}(H)$. Since we want to show that h is larger than or equal to this degree, it suffices to construct a degree-h polynomial with coefficients in $\mathrm{Fix}(H)$ that vanishes in c. For this, set

$$f(x) \stackrel{\text{def}}{=} (x - c) \cdot (x - \sigma_2(c)) \cdot \ \ldots \ \cdot (x - \sigma_h(c)).$$

Clearly $f(c) = 0$. Moreover, for any σ_i in H, by symmetry one has $\sigma_i(f) = f$. So the coefficients of f are fixed by all automorphisms in H. By definition, then, all coefficients of f are in $\mathrm{Fix}(H)$.

(4) Suppose that $\mathbb{G}$ is a normal extension of $\mathbb{K}$. We want to show that $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ is normal in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$. That is, for each σ in $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ and for each τ in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$, we want to show that $\tau^{-1}\sigma\tau$ is in $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$. So pick any element x of the field $\mathbb{G}$. We have to show

$$\tau^{-1}\sigma\tau(x) = x, \quad \text{or equivalently,} \ \ \sigma\tau(x) = \tau(x).$$

Let f be the minimum polynomial of x over $\mathbb{K}$. Since the coefficients of f are in $\mathbb{K}$, and τ fixes $\mathbb{K}$, we have

$$f(\tau(x)) = \tau(f(x)) = f(x) = 0.$$

But then because of normality, since x is in $\mathbb{G}$, also $\tau(x)$ is in $\mathbb{G}$. Since σ fixes $\mathbb{G}$, we have $\sigma\tau(x) = \tau(x)$, as desired.

To sketch the converse inclusion, suppose that $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ is normal in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$. We want to show that $\mathbb{G}$ is normal over $\mathbb{K}$. So, let p be any irreducible polynomial in $\mathbb{K}[x]$ that has a root x in $\mathbb{G}$. Let y be any other root of p. All we need to show is that y is also in $\mathbb{G}$. Choose a map τ in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$ that sends x to y. By the normality assumption, for any σ in $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$ we have that $\tau^{-1}\sigma\tau$ fixes $\mathbb{G}$, so in particular

$$\tau^{-1}\sigma\tau(x) = x \quad \text{or equivalently,} \ \ \sigma\tau(x) = \tau(x). \tag{79}$$

Recall that $y = \tau(x)$, by the way τ was chosen. So Equation 80 tells us that y is fixed by all maps σ in $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$. So y is in $\mathrm{Fix}(\mathrm{Gal}(\mathbb{F}, \mathbb{G})$, which by item (3) equals $\mathbb{G}$.
This shows the "if and only if". As for the final line of the theorem, the proof is a bit technical, but the main idea is just "restricting the homomorphism". Let $\overline{\sigma}$ be a generic element of $X \stackrel{\text{def}}{=} \mathrm{Gal}(\mathbb{F}/\mathbb{K})/\mathrm{Gal}(\mathbb{F}/\mathbb{G})$. By definition of quotient, this is the class of equivalence of some element σ in $\mathrm{Gal}(\mathbb{F}/\mathbb{K})$; and by definition of Galois group, this σ is an automorphism of $\mathbb{F}$ that fixes $\mathbb{K}$. Now, consider the restriction $\sigma_{\mathbb{G}}$ of σ to $\mathbb{G}$: Obviously, this will be an automorphism of $\mathbb{G}$ that fixes $\mathbb{K}$. Now consider the map ψ from X to $\mathrm{Gal}(\mathbb{G}/\mathbb{K})$, that sends $\overline{\sigma} \mapsto \sigma_{\mathbb{G}}$. This map ψ is:

- well-defined and injective: By definition of quotient, $\overline{\sigma} = \overline{\tau}$ if and only if $\sigma^{-1}\tau$ is in $\mathrm{Gal}(\mathbb{F}/\mathbb{G})$, which means that $\sigma^{-1}\tau$ fixes $\mathbb{G}$. This happens if and only if for every g in $\mathbb{G}$, $\sigma^{-1}\tau(g) = g$; if and only if for every g in $\mathbb{G}$, $\tau(g) = \sigma(g)$; if and only if $\sigma_{\mathbb{G}} = \tau_{\mathbb{G}}$; if and only if $\psi(\sigma) = \psi(\tau)$.

- group homomorphism: $\psi(\overline{\sigma}\cdot\overline{\tau}) = \sigma_{\mathbb{G}}\tau_{\mathbb{G}} = (\sigma\tau)_{\mathbb{G}} = \psi(\overline{\sigma})\cdot\psi(\overline{\tau})$.
- surjective: Let ρ be any element of $\mathrm{Gal}(\mathbb{G}/\mathbb{K})$, that is, an automorphism of $\mathbb{G}$ that fixes $\mathbb{K}$. Extend it any way you wish to an automorphism ρ' of $\mathbb{F}$. Choose any automorphism τ of $\mathbb{F}$ that fixes $\mathbb{G}$ (and in particular $\mathbb{K}$). Then the map $\sigma \stackrel{\text{def}}{=} \tau\rho'$ is an automorphism of $\mathbb{F}$ that fixes $\mathbb{K}$. Moreover, for any $g \in \mathbb{G}$,

$$\sigma(g) = \tau\rho'(g) = \tau\rho(g) = \rho(g).$$

  So $\psi(\overline{\sigma}) = \sigma_{\mathbb{G}} = \rho$.

Hence, ψ is an isomorphism. □

Example 637. Going back to Example 619: Between $\mathbb{Q}$ and $\mathbb{F} \stackrel{\text{def}}{=} \mathbb{Q}(\sqrt{2},\sqrt{3})$ there are two different intermediate fields, $\mathbb{G}_1 \stackrel{\text{def}}{=} \mathbb{Q}(\sqrt{2})$ and $\mathbb{G}_2 \stackrel{\text{def}}{=} \mathbb{Q}(\sqrt{3})$. Both $\mathbb{G}_1$ and $\mathbb{G}_2$ are normal over $\mathbb{Q}$, as they are splitting fields of the polynomials x^2-2 and x^2-3, respectively. Note that $\varphi_1(\sqrt{2}) = \sqrt{2}$, so φ_1 fixes $\mathbb{G}_1$; similarly, $\varphi_2(\sqrt{3}) = \sqrt{3}$, so φ_2 fixes $\mathbb{G}_2$. Hence, $\mathrm{Gal}(\mathbb{F}/\mathbb{G}_1)$ is the subgroup of $\mathrm{Gal}(\mathbb{F}/\mathbb{Q})$ consisting of the identity and φ_1. Similarly $\mathrm{Gal}(\mathbb{F}/\mathbb{G}_1)$ is the subgroup of $\mathrm{Gal}(\mathbb{F}/\mathbb{Q})$ consisting of the identity and φ_2. Both are normal subgroups, since $\mathrm{Gal}(\mathbb{F}/\mathbb{Q}) \cong \mathbb{Z}_2\times\mathbb{Z}_2$ is Abelian.

Example 638. Set $r \stackrel{\text{def}}{=} \sqrt[4]{3}$. The polynomial x^4-3 has the four distinct roots $r, ir, -r, -ir$. The splitting field is $\mathbb{Q}(r,i)$, which has degree 8 over $\mathbb{Q}$. So the Galois polynomial has 8 elements. If we define the two maps

$$\begin{array}{rccc} a: & \mathbb{Q}(r,i) & \to \mathbb{Q}(r,i) \\ & r & \mapsto ir \\ & i & \mapsto i \end{array} \qquad \begin{array}{rccc} b: & \mathbb{Q}(\sqrt{2},\sqrt{3}) & \to \mathbb{Q}(\sqrt{2},\sqrt{3}) \\ & r & \mapsto \ \ r \\ & i\sqrt{3} & \mapsto -i \end{array}$$

one can see $a^4 = id = b^2$ and the Galois group consists of the elements

$$\{id,\ a,\ a^2,\ a^3,\ b,\ ba,\ ba^2,\ ba^3\}.$$

This is called *dihedral group* D_4 in the literature. It's the symmetry group of the square, where a is the rotation of 90° anticlockwise, and b is the reflection with respect to the x axis, say. It's not commutative: One has

$$ab = ba^3 \neq ba. \tag{80}$$

The subgroup H_a generated by a is normal, because it has half of the elements: Compare Proposition 480. The subgroup H_b generated by b, instead, is not normal: since by (81)

$$aba^2 = ba^3a^2 = ba^5 = ba,$$

multiplying by a^{-1} on the left we get

$$ba^2 = a^{-1}ba,$$

so $a^{-1}ba$ is neither the identity, nor b. And indeed, $\mathrm{Fix}(H_a) = \mathbb{Q}(i)$ is a normal extension of $\mathbb{Q}$ (it is the splitting field of x^2+1), whereas $\mathrm{Fix}(H_b) = \mathbb{Q}(r) = \mathbb{Q}(\sqrt[4]{3})$ is not a normal extension of $\mathbb{Q}$. The ascending sequence of normal subgroups

$$(id) \subseteq (a^2) \subseteq (a) \subseteq \mathrm{Gal}(\mathbb{Q}(r,i)/\mathbb{Q})$$

corresponds (by looking at $\mathrm{Fix}(*)$) to the descending sequence of normal extensions

$$\mathbb{Q}(r,i) \supset \mathbb{Q}(r^2,i) \supset \mathbb{Q}(i) \supset \mathbb{Q}.$$

Deeper thoughts 639. A deep theorem by Cayley states that every finite group is a subgroup of some $\mathcal{S}_n$. Also, we will prove at the end of this chapter that for any prime $p \geq 5$, $\mathcal{S}_p$ is the Galois group of some degree-p polynomial. Together with these two theorems, Theorem 636 implies that every finite group is the Galois group of a suitable polynomial. The problem of finding a polynomial with given Galois group is called *inverse Galois problem*.

8.6 Equations that cannot be solved by radicals

Finally, we are ready to reflect more carefully on what it means to have a "formula" for solving an equation. Recall that by "n-th root" of an element t of a field $\mathbb{F}$, we mean another element w in a finite extension field of $\mathbb{F}$, such that $w^n = t$. Intuitively, the polynomial equation $f(x) = 0$ is *solvable by radicals* if all its roots can be computed from the coefficients of f in a finite number of steps, by means of sums, products, or root extractions. For example, the roots of $ax^2+bx+c = 0$ can be computed as $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$.

Now, if you are in a field $\mathbb{F}$, taking sums and products of elements of $\mathbb{F}$ keeps you within that field. Root extractions are the only possible steps that enable you to "move on" to a bigger field, namely, a finite field extension. Note also that these extension are somewhat of a "special kind", in the sense that the minimum polynomial of the element you are adding, **is always of the type $\mathbf{x^n - t}$**, for some t in your field.

Definition 640. Let $\mathbb{F}$ be a field of characteristic zero, and let $f \in \mathbb{F}[x]$. The polynomial equation $f(x) = 0$ is *solvable by radicals with exponents* $(n_0, \ldots, n_s)$ if all roots of f lie in a field $\mathbb{F}_s$, where

$$\mathbb{F} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \ldots \subseteq \mathbb{F}_s, \tag{81}$$

such that for all i in $\{0, \ldots, s-1\}$ one has

$$\mathbb{F}_{i+1} = \mathbb{F}_i(w_{i+1}) \qquad \text{and} \qquad (w_{i+1})^{n_i} = f_i \in \mathbb{F}_i.$$

The expression 82 is called a *root tower for f over $\mathbb{F}$*. The positive integers $(n_0, \ldots, n_s)$ are called the *exponents* of the root tower. Without loss of generality, they are all greater than 1.

We didn't request these extensions $\mathbb{F}_i$ to be normal over $\mathbb{F}$; but as we shall now see, there is no loss of generality in assuming it. This is perhaps the most crucial step in Galois' theory. We begin with two Lemmas.

Lemma 641. *Let $\mathbb{F} \subseteq \mathbb{E}$ be fields of characteristic zero. Let $f \in \mathbb{F}$. If $x^n - f$ splits into linear factors in $\mathbb{E}$, then $\mathbb{E}$ contains all n-th roots of unity.*

Sketch of proof. Let r be a root of $x^n - f$ inside $\mathbb{E}$. Let u be an n-th root of unity. Let k be any integer in $\{0, \ldots, n-1\}$. Then

$$(u^k r)^n = u^{kn} r^n = (u^n)^k f = 1^k f = f.$$

So each $u^k r$ is a root of $x^n - f$. Since there are n choices for k and they give rise to different roots, we have that $r, ur, u^2r, \ldots, u^{n-1}r$ are the n roots of $x^n - f$. It follows that the splitting field of $x^n - f$ is $\mathbb{F}(r, ur, u^2r, \ldots, u^{n-1}r)$. Now, there is an isomorphic copy of this field inside $\mathbb{E}$; for simplicity, let us identify $\mathbb{F}(r, ur, u^2r, \ldots, u^{n-1}r)$ with its isomorphic image and just write $\mathbb{F}(r, ur, u^2r, \ldots, u^{n-1}r) \subseteq \mathbb{E}$. Since $\mathbb{E}$ contains ur and u, it contains also their ratio u. □

Lemma 642. *Let $\mathbb{F} \subseteq \mathbb{E}$ be fields of characteristic zero. Suppose $\mathbb{F}$ already contains all n-th roots of unity. Let $f \in \mathbb{F}$. If $\mathbb{E}$ is a splitting field for $x^n - f$ over $\mathbb{F}$, then* $\mathrm{Gal}(\mathbb{E}/\mathbb{F})$ *is cyclic.*

Proof. Let r be a root of $x^n - f$ in $\mathbb{E}$. Let u be the first n-th root of unity. Then as above, the roots of $x^n - f$ can be described as

$$r, ur, u^2r, \ldots, u^{n-1}r.$$

So $\mathbb{E}$, the splitting field of $x^n - f$, is (isomorphic to) $\mathbb{F}(r, ur, u^2r, \ldots, u^{n-1}r)$, which is equal to $\mathbb{F}(r)$ because $u \in \mathbb{F}$. Now write

$$\mathrm{Gal}(\mathbb{E}/\mathbb{F}) = \{\sigma_1, \ldots, \sigma_m\}.$$

Fix i in $\{1, \ldots, m\}$ and consider the automorphism $\sigma_i \in \mathrm{Gal}(\mathbb{E}/\mathbb{F})$. Since it maps roots to roots, there is some integer k_i such that

$$\sigma_i(r) = u^{k_i} r.$$

But u^{k_i} belongs to $\mathbb{F}$, since $\mathbb{F}$ contains all roots of unity. And by definition of Galois group, σ_i fixes the elements of $\mathbb{F}$. So $\sigma_i(u^{k_i}) = u^{k_i}$ and

$$\sigma_i(\sigma_j r) = \sigma_i(u^{k_j} r) = \sigma_i(u^{k_j})\sigma_i(r) = u^{k_j}\sigma_i(r) = u^{k_j}u^{k_i}r = u^{k_i+k_j}r.$$

So we can define a group isomorphism $\psi : \mathrm{Gal}(\mathbb{E}/\mathbb{F}) \to \mathbb{Z}_m$ by mapping σ_i to k_i. □

Theorem 643 (Galois)**.** *Let $f(x) = 0$ be a polynomial in $\mathbb{F}[x]$, where $\mathbb{F}$ is any field of characteristic zero. Assume f is solvable by radicals with exponents $(n_0, \ldots, n_s)$.*

(1) *There exists a root tower $\mathbb{F} = \mathbb{K}_0 \subseteq \ldots \subseteq \mathbb{K}_s$ for f over $\mathbb{F}$, again with exponents $(n_0, \ldots, n_s)$, in which all the fields $\mathbb{K}_i$ are normal over $\mathbb{F}$.*

(2) *Moreover, if $n \stackrel{\text{def}}{=} \mathrm{lcm}(n_0, \ldots, n_s)$, then all n-th roots of unity are contained in $\mathbb{K}_s$.*

(3) *Moreover, there exists an n-th root of unity u such that the group*

$$\mathrm{Gal}(\mathbb{K}_s/\mathbb{F}(u))$$

is solvable.

(4) *Finally, $\mathrm{Gal}(\mathbb{K}_s/\mathbb{F})$ is solvable.*

Proof.

(1) By definition, $\mathbb{F} = \mathbb{F}_0 \subseteq \mathbb{F}_1 \subseteq \ldots \subseteq \mathbb{F}_s$, with $\mathbb{F}_{i+1} = \mathbb{F}_i(w_{i+1})$ and $(w_{i+1})^{n_i} = f_i \in \mathbb{F}_i$. We inductively construct a new root tower $\mathbb{F} = \mathbb{K}_0 \subseteq \ldots \subseteq \mathbb{K}_s$ for f over $\mathbb{F}$, as follows.

- Set $\mathbb{K}_0 \stackrel{\text{def}}{=} \mathbb{F}$ (obviously).
- If $\mathbb{F}_1 = \mathbb{F}(w_1)$, with $w_1^{n_0} = f_0 \in \mathbb{F}_0 = \mathbb{F}$, define $\mathbb{K}_1$ to be the splitting field of the polynomial
$$g_0 \stackrel{\text{def}}{=} x^{n_0} - f_0.$$
Clearly $\mathbb{K}_1$ is normal over $\mathbb{F}$ by Lemma 628. Note that $\mathbb{K}_1 \supset \mathbb{F}_1$.
- Induction step: Suppose that we have constructed $\mathbb{K}_i$ normal over $\mathbb{F}$ and containing $\mathbb{F}_i$. Let $\sigma_1, \sigma_2, \ldots, \sigma_{m_i}$ be the elements of $\mathrm{Gal}(\mathbb{K}_i, \mathbb{F})$. Suppose also that $\mathbb{F}_{i+1} = \mathbb{F}_i(w_{i+1})$ with $(w_{i+1})^{n_i} = f_i \in \mathbb{F}_i$. Consider the polynomial
$$g_i \stackrel{\text{def}}{=} (x^{n_i} - \sigma_1(f_i)) \cdot (x^{n_i} - \sigma_2(f_i)) \cdot \ldots \cdot (x^{n_i} - \sigma_{m_i}(f_i)).$$
Let $\mathbb{G}_{i,1}$ be the splitting field for $(x^{n_i} - \sigma_1(f_i))$ over $\mathbb{K}_i$. Inductively, let $\mathbb{G}_{i,j}$ be the splitting field for $(x^{n_i} - \sigma_j(f_i))$ over $\mathbb{G}_{i,j-1}$. Eventually, set $\mathbb{K}_{i+1} \stackrel{\text{def}}{=} \mathbb{G}_{i,m_i}$. (Recall that m_i was the number of elements of $\mathrm{Gal}(\mathbb{K}_i, \mathbb{F})$.)
By the way g_i was defined, each of the σ's in $\mathrm{Gal}(\mathbb{K}_i, \mathbb{F})$ sends g_i to itself. But by Theorem 636, part (2), what is fixed by all σ's in $\mathrm{Gal}(\mathbb{K}_i, \mathbb{F})$ must be in $\mathbb{F}$. Hence, all the coefficients of $g_i(x)$ are in $\mathbb{F}$. So if we consider the polynomial
$$g_0 \cdot g_1 \cdot \ldots \cdot g_i$$
its coefficients lie in $\mathbb{F}$, and its splitting field is $\mathbb{K}_{i+1}$. By Lemma 628, $\mathbb{K}_{i+1}$ is normal.

(2) Fix i in $\{1, \ldots, s-1\}$. Since $\mathrm{Gal}(\mathbb{K}_i, \mathbb{F})$ is a group, one of its elements $\sigma_1, \sigma_2, \ldots, \sigma_{m_i}$ is the identity. Up to reordering them, suppose it is σ_1. Then $\mathbb{G}_{i,1}$ is by definition the splitting field for $(x^{n_i} - \sigma_1(f_i)) = (x^{n_i} - f_i)$ over $\mathbb{K}_i$. Since this holds for all i, it follows that inside $\mathbb{K}_s$ all polynomials $x^{n_i} - f_i$ split into linear factors. Then $\mathbb{K}_s$, for each i, contains the first n_i-th root of unity, u_i. But then it is easy to see that the product of the u_i's is a primitive n-th root of unity. So $\mathbb{K}_s$ contains all of its powers, that is, all n-th roots of unity.

(3) Let u be the first of the n-th roots of unity (corresponding to taking $k = 1$ in Definition 609.) We claim that $\mathrm{Gal}(\mathbb{K}_s/\mathbb{F}(u))$ is solvable. To see this, consider the chain of fields

$$\mathbb{F} = \mathbb{K}_0 \subseteq \mathbb{K}_1 \subseteq \mathbb{G}_{1,1} \subseteq \mathbb{G}_{1,2} \subseteq \ldots \subseteq \mathbb{G}_{1,m_1} = \mathbb{K}_2 \subseteq \mathbb{G}_{2,1} \subseteq \ldots\ldots \subseteq \mathbb{G}_{s,m_s} = \mathbb{K}_s$$

produced in part (1). To clean up the notation, let us reindex these fields in a natural order. By setting $\mathbb{L}_0 \stackrel{\text{def}}{=} \mathbb{F}$, $\mathbb{L}_1 \stackrel{\text{def}}{=} \mathbb{K}_1$, $\mathbb{L}_2 \stackrel{\text{def}}{=} \mathbb{G}_{1,1}$, and so on, the chain above becomes

$$\mathbb{F} = \mathbb{L}_0 \subseteq \mathbb{L}_1 \subseteq \ldots \subseteq \mathbb{L}_t = \mathbb{K}_s.$$

For each $j \in \mathbb{N}$, let v_{j+1} be an element of $\mathbb{K}_s$ such that

$$\mathbb{L}_{j+1} = \mathbb{L}_j(v_{j+1}).$$

Then define recursively $\mathbb{E}_0 \stackrel{\text{def}}{=} \mathbb{F}(u)$ and

$$\mathbb{E}_{j+1} \stackrel{\text{def}}{=} \mathbb{E}_j(v_{j+1}).$$

By definition, $\mathbb{E}_j \supset \mathbb{L}_j$ for each j; in particular, $\mathbb{E}_t \supset \mathbb{L}_t = \mathbb{K}_s$. However, we showed in item (2) that $u \in \mathbb{K}_s$; so $\mathbb{E}_t \subseteq \mathbb{K}_s$, or in other words, $\mathbb{E}_t = \mathbb{K}_s$. But then $\mathbb{E}_t$ is normal over $\mathbb{F}$; so by Lemma 635, $\mathbb{E}_t$ is normal over any intermediate field $\mathbb{E}_j$. Set

$$G_j \stackrel{\text{def}}{=} \mathrm{Gal}(\mathbb{E}_t/\mathbb{E}_j).$$

Translated into this notation, what we want to show is that G_0 is solvable. By Theorem 636, we have a chain

$$G_0 \supseteq G_1 \supseteq G_2 \supseteq \ldots \supseteq G_t = (0),$$

and we know that each G_{j+1} is normal in G_j. It remains to show that G_j/G_{j+1} is Abelian. In fact, these quotients are even cyclic, because of Lemma 642.

(4) Since $\mathbb{E}_0$ is normal over $\mathbb{F}$, by Theorem 636, item(4), we get that

$$\mathrm{Gal}(\mathbb{E}_t/\mathbb{F})/\mathrm{Gal}(\mathbb{E}_t/\mathbb{E}_0) \cong \mathrm{Gal}(\mathbb{E}_0/\mathbb{F}).$$

We have already shown in part (3) that $G_0 = \mathrm{Gal}(\mathbb{E}_t/\mathbb{E}_0)$ is solvable. So if we show that $\mathrm{Gal}(\mathbb{E}_0/\mathbb{F})$ is Abelian, we are done: a chain proving solvability would then be

$$\mathrm{Gal}(\mathbb{E}_t/\mathbb{F}) \supseteq G_0 \supseteq G_1 \supseteq G_2 \supseteq \ldots \supseteq G_t = (0).$$

Is $\mathrm{Gal}(\mathbb{E}_0/\mathbb{F})$ Abelian? Recall that $\mathbb{E}_0 = \mathbb{F}(u)$, where u was a primitive n-th root of unity. So the automorphisms σ_i of $\mathbb{F}(u)$ that fix $\mathbb{F}$ are determined by their value on u, which again must be an n-th root of unity, and thus of the type

$$\sigma_i(u) = u^{k_i}.$$

From this it follows that any two such automorphisms σ_i and σ_j must commute. □

Corollary 644. *If the polynomial equation $f(x) = 0$ is solvable by radicals, and the coefficients of f belong to a field $\mathbb{F}$ of characteristic zero, then the Galois group of f is solvable.*

Proof. Let us use the notation of the previous Theorem. Since f splits over $\mathbb{K}_s$, the splitting field $\mathbb{L}$ of f will be contained in $\mathbb{K}_s$. Being a splitting field, $\mathbb{L}$ is normal over $\mathbb{F}$ by Lemma 628. $\mathbb{K}_s$ is also normal over $\mathbb{F}$. So we can apply Theorem 636, item(4):

$$\mathrm{Gal}(\mathbb{K}_s/\mathbb{F})\Big/_{\mathrm{Gal}(\mathbb{K}_s/\mathbb{L})} \cong \mathrm{Gal}(\mathbb{L}/\mathbb{F}).$$

Our goal is to show that $\mathrm{Gal}(\mathbb{L}/\mathbb{F})$ is solvable. But this group is (isomorphic to) a quotient of the solvable group $\mathrm{Gal}(\mathbb{K}_s/\mathbb{F})$: And quotient of solvable is solvable, by Proposition 594. □

Corollary 645. *Some polynomial equations of degree 5 are not solvable by radicals.*

Proof. There are polynomials like $x^5 + 20x + 16$ and $x^5 + 20x + 4$ with Galois group A_5 and $\mathcal{S}_5$, respectively (cf. Remark 632). These groups are not solvable (cf. Examples 595, 596). □

Corollary 646. *For every $d \geq 5$, some polynomial equations of degree d are not solvable by radicals.*

Proof. If f is a polynomial unsolvable by radicals, so is $x^{d-5}f$. □

One may wonder if one can construct polynomials with Galois group $\mathcal{S}_d$, for some $d > 5$. The answer is yes. The proof is complicated. Here we present a simpler version that works **when $d > 5$ is prime**.

Lemma 647. *If g in $\mathbb{R}[x]$ has distinct real roots, for $M \in \mathbb{N}$ suitably large the polynomial $g(x) + \frac{2}{M}$ has the same number of real roots as g.*

Proof. Let $z_1, \ldots, z_t$ be the zeroes of the derivative g'. None of them is a zero of g as well: In fact, were $g(z_i) = 0 = g'(z_i)$ for some i, then $(x - z_i)^2$ would divide g (by the derivative trick: compare the proof of Lemma 625), against the assumption that g has distinct roots. So, set

$$\alpha \stackrel{\text{def}}{=} \min\{|g(z_i)| \ : \ i \in \{1, \ldots, t\}\}.$$

By what we just said, α is positive. By definition, every local minimum or maximum for f has vertical distance at least α from the x-axis. If we shift the graph of f upward or downward by less than α, then, the number of roots won't change. If we choose $M > \frac{2}{\alpha}$, we have $\frac{2}{M} < \alpha$, and the conclusion follows. □

Theorem 648. *Let d be any prime larger than 5. There exists a degree-d polynomial with Galois group $\mathcal{S}_d$.*

Proof. Choose[23] any sequence of $d - 1$ positive integers $b_1, \ldots, b_{d-1}$. Set

$$g \stackrel{\text{def}}{=} (x^2 + 2b_1)(x - 2b_2)(x - 2b_3) \cdots (x - 2b_{d-1}).$$

Clearly, g has 2 roots that are complex conjugate and $d - 2$ real roots. Moreover, the leading coefficient of g is 1, but all other coefficients are even. The constant term is even a multiple of 2^{d-1}: call it $2^{d-1}k$, with $k \in \mathbb{N}$. Now let M be a large odd integer and consider

$$f(x) \stackrel{\text{def}}{=} g(x) + \frac{2}{M}.$$

[23] Proof from Math Stack Exchange, https://math.stackexchange.com/questions/1194928/irreducible-polynomial-in-mathbb-qx-of-degree-n-having-exactly-n-2-real, edited April 2017.

By Lemma 647, we may assume that f has 2 complex conjugate roots and $d-2$ real roots. So the Galois group must contain a single flip. By Remark 495, to conclude that the Galois group is $\mathcal{S}_d$ it suffices to show that the Galois group contains a d-cycle.

Let us prove it! The leading coefficient of f is 1; all other coefficients of f are of the form e/M, for some even number e. The constant term is $\frac{2^{d-1}k+2}{M}$, so it is of the form a_0/M, with a_0 congruent to 2 modulo 4. Thus the polynomial Mf is of the form

$$Mx^d + 2z_{d-1}x^{d-1} + \ldots + 2z_1x + 2O,$$

with $z_1, \ldots, z_{d-1}$ integers and M, O odd integers. By Eisenstein's theorem, Mf is irreducible over $\mathbb{Q}[x]$; but since M is invertible in $\mathbb{Q}$, this is the same as saying that f is irreducible in $\mathbb{Q}[x]$. Thus by the derivative trick, f has d distinct roots; let us call them $\alpha_1, \ldots, \alpha_d$, with $\alpha_{d-1} = \overline{\alpha_d} \notin \mathbb{R}$. Since f is irreducible and monic, it is the minimum polynomial of any of its roots. So $[\mathbb{Q}(\alpha_i) : \mathbb{Q}] = d$ for all i. As in Lemma 630, set $\mathbb{G} \stackrel{\text{def}}{=} \mathbb{Q}(\alpha_1, \ldots, \alpha_d)$. By theorem 629 and by Lagrange's theorem 394,

$$\#\operatorname{Gal}(\mathbb{G} : \mathbb{Q}) = [\mathbb{G} : \mathbb{Q}] = [\mathbb{G} : \mathbb{Q}(\alpha_1)] \cdot [\mathbb{Q}(\alpha_1) : \mathbb{Q}] = [\mathbb{G} : \mathbb{Q}(\alpha_1)] \cdot d.$$

So by Theorem 527 (here we use that d is a prime number!), the group G contains an element of order d. But within $\mathcal{S}_d$, an element of order d is just a d-cycle. □

8.7 Exercises

1. How many automorphism does a cyclic group of prime cardinality have?
2. Describe each automorphism of the Galois group of $x^4 - 5$ as permutation of the roots.
3. Let f be a polynomial in $\mathbb{Q}[x]$. Let f' be its derivative. Let $g = \gcd(f, f')$. Show that $\frac{f}{g}$ is a polynomial with the same roots as f, but no multiple root.
4. Prove that $\mathcal{S}_3$ is solvable and find a monic polynomial of $\mathbb{Q}[x]$ that has $\mathcal{S}_3$ as Galois group.
5. Find a polynomial of degree 9 that has $\mathcal{S}_7$ as Galois group.

9 Modules

We are now going to define *modules*, which are a generalization of vector spaces in which the scalars are taken from a C-ring instead of a field. This section is based on the beautiful book by David Eisenbud, *Commutative Algebra: With a View Towards Algebraic Geometry*, Springer.

9.1 Modules, submodules, quotients and direct sums

Definition 649. Let A be a C-ring with 1. An *A-module* is an Abelian group $(M,+)$, with additive notation, endowed with an operation $(a,m) \mapsto am$ from $A \times M$ to M (called *scalar multiplication*) that satisfies the following axioms, for all a, b in A and for all m, m' in M:

(M1) $a(bm) = (ab)m$.

(M2) $a(m+m') = am + am'$.

(M3) $(a+b)m = am + bm$.

(M4) $1m = m$.

Example 650. When A is a field: "A-modules" are the same as "A-vector spaces". (The definitions are literally the same.)

Example 651. When $A = \mathbb{Z}$: "$\mathbb{Z}$-modules" are the same as "Abelian groups". In fact, all modules are Abelian groups by definition. Conversely, if g is an element of an Abelian group $(G,+)$, it makes sense to define $2g = g+g$, and recursively, $ng \stackrel{\text{def}}{=} g+(n-1)g$; all these elements, plus their inverses $-g, -2g$, etc., live inside G. Condition (M2) is met because G is Abelian:

$$2(g+g') = g+g'+g+g' \stackrel{!}{=} g+g+g'+g' = 2g+2g'.$$

Example 652. For any ideal $I \subseteq A$, both I and A/I are A-modules. In particular, (0) and A are A-modules.

Notation. For convenience, from now on we will simply write 0 instead of (0).

Definition 653. Let M be an A-module. A *submodule* $S \subseteq M$ is an Abelian subgroup of $(M,+)$ such that for all a in A and for all $s \in S$, one has $as \in S$. The restriction to $A \times S$ of the scalar multiplication of M makes S an A-module.

Example 654. When A is a field: "submodules" are the same as "subspaces".

Example 655. When $A = \mathbb{Z}$: "submodules" are the same as "subgroups".

Example 656. When you view A itself as A-module, the submodules are the ideals of A.

Example 657. Let I be an arbitrary set, not necessarily finite. Let $(M_i)_{i \in I}$ be a family of A-modules. The *direct sum*

$$\bigoplus_i M_i \stackrel{\text{def}}{=} \{(m_i)_{i\in I} \text{ such that } m_i \in M_i \text{ for all } i \in I, \text{ and } m_i = 0 \text{ for all but finitely many } i\}$$

with the operation $a\,(m_i)_{i\in I} \stackrel{\text{def}}{=} (am_i)_{i\in I}$, is a submodule of the *Cartesian product*

$$\prod_i M_i \stackrel{\text{def}}{=} \{(m_i)_{i\in I} \text{ such that } m_i \in M_i \text{ for all } i \in I\}.$$

Remark 658. Direct sums obviously coincide with Cartesian products if the set I is finite. In general, the direct sum is a submodule of the Cartesian product. It turns out that direct sums are much easier to understand than Cartesian products. For this reason, we tend to prefer the notation $M_1 \oplus M_2$ to $M_1 \times M_2$, even if they obviously are the same, since the set $I = \{1, 2\}$ is finite. So when you see $M_1 \oplus M_2$, remember that it is just the classical Cartesian product.

Definition 659. Let S be any submodule of an A-module M. Then the quotient group ${}^{M}/_{S}$ is well-defined and it also an A-module, with respect to the operation $a\overline{m} \stackrel{\text{def}}{=} \overline{am}$. We will refer to it as the *quotient (module).*

9.2 Homomorphisms, short exact sequences, and projective modules

Definition 660. A function $f : M \longrightarrow N$ between two A-modules is called a *homomorphism (of modules)* if for all $a, b \in A$ and for all $m \in M$,

- $f(a + b) = f(a) + f(b)$ (that is, f is a homomorphism of Abelian groups);
- $f(am) = af(m)$.

A bijective homomorphism is called *isomorphism*; it is an easy exercise to see that the inverse of a bijective homomorphism is still a homomorphism. Two modules are *isomorphic* if there exist an isomorphism between them.

Example 661. Let A be any C-ring with 1. Let M be any A- module. Let $m \in M$. Consider the "multiplication by m" , $\mu : A \longrightarrow M$, $\mu(a) \stackrel{\text{def}}{=} am$. By Axioms (M3) and (M1)

$$\mu(a + b) = (a + b)m = am + bm = \mu(a) + \mu(b) \text{ and } \mu(ab) = (ab)m = a(bm) = a\mu(b),$$

and thus μ is a homomorphism of modules.

Remark 662. If you fix an element m of A, the multiplication by m is a map $\alpha : A \longrightarrow A$ that is a module homomorphism, but *not* a C-ring homomorphism (unless m is idempotent).

Definition 663. A homomorphism $f : M_1 \longrightarrow M_2$ is called a *mono* if for any two homomorphisms $M_0 \underset{v}{\overset{u}{\rightrightarrows}} M_1 \xrightarrow{f} M_2$ if $fu = fv$ then $u = v$.

Definition 664. A homomorphism $f : M_1 \longrightarrow M_2$ is called an *epi* if for any two homomorphisms $M_1 \xrightarrow{f} M_2 \underset{v}{\overset{u}{\rightrightarrows}} M_3$ if $uf = vf$ then $u = v$.

Lemma 665. *Monos are exactly the injective homomorphisms, and epis are the and surjective homomorphisms.*

Proof. Left as exercise. *Hint*: To prove that monos are injective, take f mono; then choose as M_0 the kernel of f, as u the inclusion, as v the zero map. To prove that epis are surjective: Take f epi; then choose $M_3 = {}^{M_2}/_{\operatorname{Im} f}$, as u the projection, and as v the zero map. □

Definition 666. Let $(M_i)_{i \in \mathbb{Z}}$ be A-modules. A sequence of homomorphisms

$$\ldots \xrightarrow{f_{-3}} M_{-2} \xrightarrow{f_{-2}} M_{-1} \xrightarrow{f_{-1}} M_0 \xrightarrow{f_0} M_1 \xrightarrow{f_1} M_2 \xrightarrow{f_2} \ldots$$

is called *exact* if for each i, the image of f_i equals the kernel of f_{i+1}.

Definition 667. A *short exact sequence* is an exact sequence of the type

$$0 \longrightarrow M_1 \xrightarrow{\alpha_1} M_2 \xrightarrow{\alpha_2} M_3 \longrightarrow 0.$$

Note that the only homomorphism starting from 0, is the zero map, which has image $\{0\}$; and the only homomorphism mapping to 0 is the zero map, whose kernel is the whole of M_2. So essentially a short exact sequence consists of an injective map $\alpha : M_1 \longrightarrow M_2$ and a surjective map $\alpha_2 : M_2 \longrightarrow M_3$ such that the image of α_1 is equal to the kernel of α_2.

Example 668. Any homomorphism $\eta : M \to N$ gives rise to a short exact sequence

$$0 \longrightarrow \ker \eta \xrightarrow{\iota} M \xrightarrow{\eta} \operatorname{Im} \eta \longrightarrow 0.$$

Example 669. Let $n > 0$ be an integer. If $n\cdot$ is the map that multiplies any integer by n, then

$$0 \longrightarrow \mathbb{Z} \xrightarrow{n\cdot} \mathbb{Z} \xrightarrow{\pi} \mathbb{Z}_n \longrightarrow 0$$

is a short exact sequence of $\mathbb{Z}$-modules.

Example 670. Let I be any ideal of a C-ring A with 1. Then

$$0 \longrightarrow I \xrightarrow{\iota} A \xrightarrow{\pi} {}^{A}\!/_{I} \longrightarrow 0$$

is a short exact sequence of A-modules.

Example 671. Let S be any submodule of an A-module M. Then the following is a short exact sequence of A-modules:

$$0 \longrightarrow S \xrightarrow{\iota} M \xrightarrow{\pi} {}^{M}\!/_{S} \longrightarrow 0.$$

Moreover, every short exact sequence is of this form, because if

$$0 \longrightarrow M_1 \xrightarrow{\alpha_1} M_2 \xrightarrow{\alpha_2} M_3 \longrightarrow 0,$$

then $M_1 \cong \operatorname{Im} \alpha_1 = \ker \alpha_2$, which is a submodule S of M_2; and using the First Isomorphism Theorem for groups, it is easy to see that $M_3 \cong {}^{M_2}\!/_{\ker \alpha_2}$.

We are going to encode this information in the following "**commutative diagram**". Commutative means, map compositions starting from the same set and ending with the same set always give the same result.

$$\begin{array}{ccccccccc}
0 & \longrightarrow & M_1 & \xrightarrow{\alpha_1} & M_2 & \xrightarrow{\alpha_2} & M_3 & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \cong} & & \downarrow{\scriptstyle =} & & \downarrow{\scriptstyle \cong} & & \\
0 & \longrightarrow & \ker \alpha_2 & \xrightarrow{\iota} & M_2 & \xrightarrow{\pi} & {}^{M_2}\!/_{\ker \alpha_2} & \longrightarrow & 0
\end{array}$$

Lemma 672. *Given any two submodules S, T of M, let*

$$S + T \stackrel{\text{def}}{=} \{s + t \text{ such that } s \in S, t \in T\}.$$

Then there is a short exact sequence

$$0 \longrightarrow S \cap T \xrightarrow{\alpha} S \oplus T \xrightarrow{\beta} S + T \longrightarrow 0,$$

where $\alpha = (\iota_S, \iota_T)$ is the map formed by the two inclusions, and β is the difference $(s, t) \mapsto s - t$. In particular, when $S \cap T = 0$, then $S \oplus T \cong S + T$.

Proof. Left as exercise. □

Definition 673. A short exact sequence $0 \longrightarrow M_1 \xrightarrow{\alpha_1} M_2 \xrightarrow{\alpha_2} M_3 \longrightarrow 0$ *splits* if there is an isomorphism $\varphi : M_2 \xrightarrow{\cong} M_1 \oplus M_3$ such that the following diagram commutes:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M_1 & \xrightarrow{\alpha_1} & M_2 & \xrightarrow{\alpha_2} & M_3 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle =} & & \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle =} & & \\ 0 & \longrightarrow & M_1 & \xrightarrow{\iota} & M_1 \oplus M_3 & \xrightarrow{\pi} & M_3 & \longrightarrow & 0 \end{array}$$

Non-Example 674. Not all short exact sequences split. For example, within $\mathbb{Z}$-modules (aka Abelian groups), consider for any prime p the short exact sequence

$$0 \longrightarrow \mathbb{Z}_p \xrightarrow{p\cdot} \mathbb{Z}_{p^2} \xrightarrow{\pi} \mathbb{Z}_p \longrightarrow 0$$

where the map $p\cdot$ is $1 \mapsto p$ and the map π is the projection mod p. This does not split, because $\mathbb{Z}_{p^2}$ is not isomorphic as Abelian group to $\mathbb{Z}_p \times \mathbb{Z}_p$. Analogously, the short exact sequence of Example 669 does not split ($\mathbb{Z}$ is not isomorphic to $\mathbb{Z} \oplus \mathbb{Z}_n$ because they have different torsion, cf. Lemma 588).

Example 675. Let $a, b \geq 2$ be two integers with $\gcd(a, b) = 1$. The following short exact sequence splits

$$0 \longrightarrow \mathbb{Z}_a \xrightarrow{b\cdot} \mathbb{Z}_{ab} \xrightarrow{\pi} \mathbb{Z}_b \longrightarrow 0$$

Proposition 676. *Let A be a C-ring with 1 and let $0 \longrightarrow M_1 \xrightarrow{\alpha_1} M_2 \xrightarrow{\alpha_2} M_3 \longrightarrow 0$ be a short exact sequence of A-modules. The following are equivalent:*

(1) *The sequence splits.*

(2) *There is a homomorphism $\gamma : M_3 \longrightarrow M_2$ such that $\alpha_2\gamma$ is the identity on M_3.*

(3) *There is a homomorphism $\beta : M_2 \longrightarrow M_1$ such that $\beta\alpha_1$ is the identity on M_1.*

Proof.
(1) $\Rightarrow$ (2). By the assumption, we have

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M_1 & \xrightarrow{\alpha_1} & M_2 & \xrightarrow{\alpha_2} & M_3 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle =} & & \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle =} & & \\ 0 & \longrightarrow & M_1 & \xrightarrow{\iota} & M_1 \oplus M_3 & \xrightarrow{\pi} & M_3 & \longrightarrow & 0 \end{array}$$

Using the inclusion $j : M_3 \to M_1 \oplus M_3$, let us define the map $\gamma \stackrel{\text{def}}{=} \varphi^{-1} j$. Indeed, for any x in M_3, using the commutativity of the diagram, we have

$$\alpha_2\gamma(x) = \pi\varphi\gamma(x) = \pi\varphi\varphi^{-1}j(x) = \pi j(x) = x.$$

(2) $\Rightarrow$ (1). Since α_1 is injective, we can invert it on its image. For any $m \in M_2$, we define

$$\varphi : M_2 \longrightarrow M_1 \oplus M_3, \qquad \varphi \stackrel{\text{def}}{=} (\alpha_1^{-1}(id_{M_2} - \gamma\alpha_2),\ \alpha_2).$$

Let us first prove that φ is well-defined. Since α_1 is injective but (a priori) not surjective, we need to check whether for all m the element $m - \gamma\alpha_2(m)$ belongs to $\operatorname{Im}\alpha_1 = \ker\alpha_2$. Indeed, by the assumption,

$$\alpha_2(m - \gamma\alpha_2(m)) = \alpha_2(m) - \alpha_2\gamma(\alpha_2(m)) = \alpha_2(m) - \alpha_2(m) = 0.$$

Now let us check that φ makes the diagram commute. By construction, if $m \in M_1$,

$$\varphi(\alpha_1(m_1)) = \alpha_1^{-1}(\alpha_1(m_1) - \gamma\alpha_2\alpha_1(m_1)) = \alpha_1^{-1}(\alpha_1(m_1)) = m_1,$$

because $\alpha_2\alpha_1$ is the zero map. So the first square commutes. The commutativity of the second square is obvious. So it remains to see that φ is a bijection. We leave it to you to verify that its inverse is

$$\psi : M_1 \oplus M_3 \longrightarrow M_1, \qquad \psi(x, y) = \alpha_2(x) + \gamma(y).$$

(1) $\Rightarrow$ (3). Left as exercise.
(3) $\Rightarrow$ (1). Left as exercise. Hint: Define $\varphi \stackrel{\text{def}}{=} (\beta,\ \alpha_2)$. □

Definition 677. A *free A-module* is a module that is isomorphic to the direct sum of (possibly infinitely many) copies of A.

Definition 678. An A-module P is called *projective* if there exists an A-module Q such that $P \oplus Q$ is free.

Remark 679. All free modules are obviously projective (choose $Q = A$). The converse is true for $\mathbb{Z}$-modules (cf. Corollary 697), but not in general: For example, $\mathbb{Z}_2$ and $\mathbb{Z}_3$ are $\mathbb{Z}_6$-modules. Since $\mathbb{Z}_2 \oplus \mathbb{Z}_3 = \mathbb{Z}_6$, both $\mathbb{Z}_2$ and $\mathbb{Z}_3$ are projective, but not free. Cf. Deeper Thoughts 698.

Proposition 680. *Let A be a C-ring with 1. For any A-module P the following are equivalent:*

(1) *P is projective.*

(2) *For any homomorphism $f : P \longrightarrow N$ and any epi $g : M \to N$, there is a homomorphism $h : P \to M$ such that $f = gh$.*

(3) *Every short exact sequence $0 \longrightarrow M_1 \longrightarrow M_2 \longrightarrow P \longrightarrow 0$ splits.*

Proof.
(1) $\Rightarrow$ (2). Let $(b_i)_i \in I$ be a basis of $P \oplus Q$. We do not know if g is injective, but it is surjective; so for every i, choose (any way you want) some m_i in M such that $g(m_i) = f(b_i)$. Since we can define a map $H : P \oplus Q \to M$ by declaring where we want to send the basis elements, let us impose $H(b_i) \stackrel{\text{def}}{=} m_i$ for all i. The restriction of H to P is the desired map h.
(2) $\Rightarrow$ (3). Since $M_2 \longrightarrow P$ is epi, call it g, and apply the assumption (2) to $f = id_P$. This gives you a map $h : P \longrightarrow M_2$ such that gh is the identity on P. By Proposition 676 (applied to $h = \gamma$), this implies that the sequence splits.
(3) $\Rightarrow$ (1). Let F be any free module from which some epi map $\eta : F \longrightarrow P$ exists. (To construct one such F: choose a set of generators $\{g_i\}_{i \in I}$ for P, set $F \stackrel{\text{def}}{=} \bigoplus_{i \in I} A$, and define η by mapping e_i to g_i.) The map η can be expanded to a short exact sequence

$$0 \longrightarrow \ker \eta \longrightarrow F \stackrel{\eta}{\longrightarrow} P \longrightarrow 0.$$

By assumption this sequence splits, so $(P \oplus \ker \eta) \cong F$ is free and P is projective. □

Corollary 681. *When A is a field, all short exact sequences of A-modules split.*

Proof. All vector spaces have a basis, hence they are all free and in particular they are all projective. Apply Proposition 680. □

9.3 Homology and free resolutions

Let us start with a generalization of the notion of "exact sequences".

Definition 682. Let $(M_i)_{i\in\mathbb{Z}}$ be A-modules. A sequence of homomorphisms

$$\left(M_i \xrightarrow{f_i} M_{i+1}\right)_{i\in\mathbb{Z}} \stackrel{\text{def}}{=} \ldots \xrightarrow{f_{-1}} M_0 \xrightarrow{f_0} M_1 \xrightarrow{f_1} M_2 \xrightarrow{f_2} \ldots$$

is called a *complex* if for each i, the image of f_{i-1} is *contained* in the kernel of f_i. Or in other words, if the composition of any two consecutive maps is the zero map.
The *i-th homology* of this complex is the A-module

$$H_i \stackrel{\text{def}}{=} \frac{\ker f_i}{\operatorname{Im} f_{i-1}}.$$

Remark 683. Exact sequences are precisely the complexes in which the i-th homology is 0, for all i. For this reason, sometimes in the literature you may find the expression "exact at i" instead of "the i-th homology (module) is 0".

Definition 684. Let $\mathbf{M} \stackrel{\text{def}}{=} \left(M_i \xrightarrow{f_i} M_{i+1}\right)_{i\in\mathbb{Z}}$ and $\mathbf{N} \stackrel{\text{def}}{=} \left(N_i \xrightarrow{g_i} N_{i+1}\right)_{i\in\mathbb{Z}}$ be two complexes of A-modules. A *homomorphism of complexes* is a sequence $\boldsymbol{\alpha}$ of homomorphisms $\alpha_i : M_i \to N_i$ such that all squares in the diagram below commute, i.e. $\alpha_{i+1} f_i = g_i \alpha_i$ for all i.

$$\begin{array}{ccccccccc}
\ldots & \longrightarrow & M_{i-1} & \xrightarrow{f_{i-1}} & M_i & \xrightarrow{f_i} & M_{i+1} & \xrightarrow{f_{i+1}} & \ldots \\
 & & \downarrow{\scriptstyle \alpha_{i-1}} & & \downarrow{\scriptstyle \alpha_i} & & \downarrow{\scriptstyle \alpha_{i+1}} & & \\
\ldots & \longrightarrow & N_{i-1} & \xrightarrow{g_{i-1}} & N_i & \xrightarrow{g_i} & N_{i+1} & \xrightarrow{g_{i+1}} & \ldots
\end{array}$$

A homomorphism of complexes is called *mono* (or *epi*, or *isomorphism*), if all the homomorphisms of modules that compose it are monos (resp. epis, resp. isomorphisms).

Definition 685. A *short exact sequence* of complexes

$$\mathbf{0} \longrightarrow \mathbf{L} \xrightarrow{\boldsymbol{\alpha}} \mathbf{M} \xrightarrow{\boldsymbol{\beta}} \mathbf{N} \longrightarrow \mathbf{0}$$

is given by a sequence $\boldsymbol{\alpha}$ of monos $\alpha_i : L_i \to M_i$ and a sequence $\boldsymbol{\beta}$ of epis $\beta_i : M_i \to N_i$ such that $\operatorname{Im}\alpha_i = \ker\beta_i$ for all i.

Theorem 686. *Every short exact sequence of complexes*

$$\mathbf{0} \longrightarrow \mathbf{L} \xrightarrow{\boldsymbol{\alpha}} \mathbf{M} \xrightarrow{\boldsymbol{\beta}} \mathbf{N} \longrightarrow \mathbf{0}$$

induces a long exact sequence in homology

$$\ldots \longrightarrow H_{n-1}(\mathbf{N}) \longrightarrow H_n(\mathbf{L}) \longrightarrow H_n(\mathbf{M}) \longrightarrow H_n(\mathbf{N}) \longrightarrow H_{n+1}(\mathbf{L}) \longrightarrow \ldots$$

Sketch of proof. The details are left as exercise; we sketch the main idea. Let $e_i : L_i \to L_{i+1}$ be the maps of the complex $\mathbf{L}$. As above, we call f_i and g_i the maps of the complexes $\mathbf{M}$ and $\mathbf{N}$. For each n, the map $H_n(\mathbf{L}) \longrightarrow H_n(\mathbf{M})$ is the one induced by α; you should check that it's well-defined. Similarly, for each n, the map $H_n(\mathbf{M}) \longrightarrow H_n(\mathbf{N})$ is the one induced by β (and you should check that it's well-defined). The hard part of the proof is to come up, for each n, with a map $H_n(\mathbf{N}) \longrightarrow H_{n+1}(\mathbf{L})$. This is done in the following way. Pick $\overline{c} \in H_n(\mathbf{N}) \stackrel{\text{def}}{=} \frac{\ker g_n}{\operatorname{Im} g_{n-1}}$.

This is the class of equivalence of an element $c \in \ker g_n$, so $g_n(c) = 0$. Because β_n is surjective for all n, then there is a $b \in L_n$ such that $c = \beta_n(b)$. Applying g_n to this equality, and using that all squares commute, we get

$$0 = g_n(c) = g_n\beta_n(b) = \beta_{n+1}f_n(b).$$

So $f_n(b)$ belongs to $\ker \beta_{n+1}$, which is equal to $\operatorname{Im} \alpha_{n+1}$. Since α_{n+1} is injective, there is a unique a in L_{n+1} such that $f_n(b) = \alpha_{n+1}(a)$. We claim that this a is in $\ker e_{n+2}$. To see this it suffices to show that $\alpha_{n+2}e_{n+2}(a) = 0$, because α_{n+2} is injective; and indeed,

$$\alpha_{n+2}e_{n+2}(a) = f_{n+1}\alpha_{n+1}(a) = f_{n+1}f_n(b) = 0,$$

because the composition of two consecutives f_i's is zero. So we can define the map

$$\begin{array}{rccc} \Delta_n : & H_n(\mathbf{N}) & \longrightarrow & H_{n+1}(\mathbf{L}) \\ & \overline{c} & \longmapsto & \overline{a}. \end{array}$$

It remains to verify (exercise!) that:

- Δ_n is well-defined, i.e. $\overline{c} = \overline{c'}$ (or equivalently, if $c - c' \in \operatorname{Im} g_{n+1}$) then $\Delta_n(\overline{c}) = \Delta_n(\overline{c'})$;
- the sequence formed by the maps we constructed is exact. □

Here is a particularly interesting exact sequence:

Definition 687. Let A be a C-ring with 1. A *free resolution* of a module M is a long exact sequence of free A-modules

$$\ldots \longrightarrow F_n \xrightarrow{\varphi_n} F_{n-1} \xrightarrow{\varphi_{n-1}} \ldots \longrightarrow F_1 \xrightarrow{\varphi_1} F_0 \xrightarrow{\varphi_0} M \longrightarrow 0$$

$\operatorname{Im} \varphi_i$ is called *the i-th syzygy module* of M (with respect to the given resolution). The last (surjective) map $\varphi_0 : F_0 \longrightarrow M$ of a resolution is called a *presentation* of M.

Every module has a (possibly infinite) free resolution: Start by taking some generators for M, and map a free module onto M by sending the generators of the free module to the generators of M, as we have done in the proof of Proposition 680. Then call M_1 the kernel of this map, and iterate (i.e. find a free resolution of M_1).

An important class of modules with a *finite* free resolution, is the ideals in a polynomial ring S with finitely many variables:

Example 688. Let $S \stackrel{\text{def}}{=} \mathbb{C}[x, y, z]$. The ideal $I = (xy, xz) \subseteq S$ has a free resolution

$$0 \longrightarrow S \xrightarrow{\varphi_1} S^2 \xrightarrow{\varphi_0} I \longrightarrow 0$$

where φ_1 is the injection $f \mapsto (zf, -yf)$, and φ_0 is the 'presentation' $(f, g) \mapsto (xyf + xzg)$.

Definition 689. Let $\mathbb{K}$ be a field. Let $\mathfrak{M} \stackrel{\text{def}}{=} (x_1, \ldots, x_n)$, the so-called "irrelevant ideal" of $\mathbb{K}[x_1, \ldots, x_n]$. A free resolution of an ideal I of $\mathbb{K}[x_1, \ldots, x_n]$ is called *minimal* if the image of each f_i is contained in $\mathfrak{M} \cdot F_i$.

Non-Example 690. The resolution of Example 688 is minimal. The same ideal I has also non-minimal resolutions, for example

$$0 \longrightarrow S \xrightarrow{\varphi_2} S^2 \xrightarrow{\varphi_1} S^2 \xrightarrow{\varphi_0} I \longrightarrow 0$$

where φ_2 is the injection $u \mapsto (-yu, u)$, φ_1 is the map $(a, b) \mapsto (az - by, ayz - by^2)$ and φ_0 is the presentation $(f, g) \mapsto xyf + xzg$. This resolution is not minimal because of φ_2.

Theorem 691 (Hilbert's syzygy theorem)**.** *Every ideal I of $\mathbb{K}[x_1, \ldots, x_n]$ has a finite minimal free resolution, which is unique up to isomorphism of complexes. The number of other nonzero modules involved in such resolution of I, called* projective dimension of I*, is at most n.*

9.4 *Smith normal forms and finitely generated modules

In this section, we extend the result of Section 7.7 from finitely generated Abelian groups to finitely generated P-modules, where P is any Principal Ideal Domain.

Definition 692 (Smith-Normal)**.** Let P be a PID. We say that a nonzero $m \times n$ matrix D with entries in P is *Smith-normal* if there exists an integer $t \in \{1, \dots, \min(m, n)\}$ such that:

- all entries of D are zero except for $d_{11}, \dots, d_{tt}$;
- for all $i < j$ in $\{1, \dots, t\}$, d_{ii} divides d_{jj}.

Theorem 693 (Smith Normal Form)**.** *Let A be any nonzero matrix with entries in a PID. Then we can find invertible square matrices U, V, such that the matrix $D \stackrel{\text{def}}{=} UAV$ is Smith-normal. Moreover, D is unique up to multiplying some of the d_{ii} by invertible elements. (For this reason, we shall call D "the Smith normal form of A".)*

Proof. The proof is very similar to that of Theorem 574. The differences with respect to Theorem 574 are that

(i) we do not have a concept of "smallest" in terms of absolute value. So we'll have to replace it with a notion of "smallest" in the terms of "with the fewest prime factors": this makes sense because every PID is a UFD.

(ii) when we want to clean up rows/columns, we cannot use Euclidean division. We will need to use the fact (cf. Remark 323) that in a PID, any ideal (a, b) generated by two elements is also generated by their greatest common divisor.

(iii) we need to prove something slightly stronger, namely, that each d_{ii} divides d_{jj} for $j > i$.

Let us try to set up a reduction algorithm. As in Theorem 574, the following moves are allowed: (1) We can exchange rows or columns in A (because it corresponds to multiplying A to the left or to the right by elementary matrices of the first kind); (2) for any $a \in A$, we can replace a row/column of A by that row/column plus a times another row/column (because it corresponds to multiplying A by elementary matrices of the second kind); (3) we change all signs in a single row/column of A (and that's an elementary matrix of the third kind). The crucial point in the proof is to introduce a fourth move available. We'll explain it only for rows (for columns it is analogous). Namely, (4): For any a, b in P with $\gcd(a, b) = 1$, and for any distinct indices i, j in $\{1, \dots, m\}$, we can pass from A to a new matrix A' where, if we denote by R_i and R'_i the i-th row of A and A', respectively, then

✓ $R'_i = aR_i + bR_j$,

✓ $R'j = cR_i + dR_j$, for any c, d in P such that $ad - bc = 1$,

✓ all other rows of A and A' are the same.

It is easy to see that $A' = FA$ where F is a matrix that "looks like the identity but contains $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ as submatrix", and for this reason, the determinant of F is precisely $ad - bc$, which is 1 because of how c and d were chosen. So F is also invertible, and move (4) is perfectly valid.

Here is the right strategy:

- Find the element p with the smallest number of (not necessarily distinct) prime factors. It exists, because P is a PID, and thus a UFD. With operations of row exchange or column exchange, let's bring p to the top-left position: Row 1, column 1.
- Now we want to "zero out" all other entries in the first column. Let's start with the element in row 2, column 1. Call it q. If $q = 0$ we do nothing. If $q \neq 0$, then q does not divide p (otherwise it would have fewer prime factors, a contradiction with how p was chosen). So let us set $g \stackrel{\text{def}}{=} \gcd(p, q)$. Since by Remark 323 $(p, q) = (g)$, then there are elements a, b in P such that
$$ap + bq = g.$$

So dividing the expression above by g, if we set $c \stackrel{\text{def}}{=} -\frac{q}{g} \in P$ and $d \stackrel{\text{def}}{=} \frac{p}{g} \in P$, we obtain

$$ad - bc = 1.$$

In particular, $\gcd(a, b) = 1$ (otherwise if they had a prime factor in common, such factor would also divide $ad - bc$). So let us apply our new move (4) to the indices $i = 1, j = 2$, which are the rows where p and q are located. From A we get to pass to a new matrix A' where

- ✓ $a'_{11} = a \cdot a_{11} + b \cdot a_{21} = ap + bq = g$, which has fewer prime factors than p;
- ✓ $a'_{21} = c \cdot a_{11} + d \cdot a_{21} = cp + dq$, which is a multiple of g;
- ✓ all other rows of A and A' are the same.

But since a'_{21} is now a multiple of a'_{11}, if we now replace row 2 by "row 2 minus g times row 1", we zero out the entry in row 2, column 1.

- Similarly, we zero out the element in row i, column 1, for fixed $i = 3, 4, 5...$, in this order.
- Now the first column has a nonzero element $d_{1,1}$ at the top, and all zeroes below. Next, we "zero out" all other entries in the first row. We proceed in analogous way as before, but this time working on the columns. Note: It might seem that by zeroing out the first row, you "spoil" the work have done before, i.e. in the first column you might get nonzero elements. If this is the case, you'll have to zero out the first column again. And then after that, maybe you'll have to fix the first row again. In any case do not worry: The process certainly ends, because the "pivot" p gets "smaller and smaller" (i.e. it has fewer and fewer factors).
- At this point the first row and the first column have all zeroes, except for the element at the top left corner. Now we induct: "ignore" the first row and the first column, and proceed to find, in the resulting submatrix, an element p with the fewest prime factors. Place it with swaps in row 2, column 2; then zero out the second row and the second column.

This way, by induction, we get to a diagonal matrix. But suppose that for some $i < j$, d_{ii} does not divide d_{jj}. What can we do? For this we need an additional trick, which is not present in Theorem 574. The trick is simply to change A by replacing column i by "column i plus column j". (This corresponds to multiplying A on the right by an elementary matrix of the second kind.) Now below d_{ii}, instead of all zeroes, there is exactly one nonzero entry, which is equal to d_{jj}. So we can reapply the strategy to "zero out column i" as we did above: This will replace the entry d_{ii} by a new entry $d_{ii}{}' = \gcd(d_{ii}, d_{jj})$, that by construction divides d_{jj}. Since this can be done for all $i < j$, but it modifies only d_{ii} (and not d_{jj}) let us start doing it for $i = t - 1$ and $j = t$; and then let us proceed upwards, considering next the case $i = t - 2$ and $j = t - 1$. And so on. □

Example 694. For any positive integers a, b, the matrix $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ is diagonal-positive (cf. Definition 572), but unless a divides b, it is not Smith normal. Its normal form is

$$\begin{pmatrix} \gcd(a, b) & 0 \\ 0 & \operatorname{lcm}(a, b) \end{pmatrix}.$$

Remark 695. If U is any square invertible matrix, its determinant must also be invertible, because $1 = \det(UU^{-1}) = \det U \det U^{-1}$. Thus if $D = UAV$ is the Smith normal form of a *square* matrix A, with entries in a PID C-ring P, then U and V are square and

$$\det D = \det U \det A \det V = \det A(\det U \det V).$$

So the determinant of a square matrix and of its Smith normal form are the same, up to a factor that is invertible in P. More generally, it is an easy linear algebra exercise to show that the Smith normal form of *any* matrix A (square or not) has the same *rank* of A.

Theorem 696. *Let P be any PID C-ring. Every submodule of any free P-module is itself free. Moreover, for every finitely generated P-module M, there is a decreasing sequence of proper ideals $(d_1) \supset (d_2) \supset \ldots \supset (d_n)$ such that*

$$M \cong P/(d_1) \oplus P/(d_2) \oplus \ldots \oplus P/(d_n).$$

Sketch of proof. The idea is the same of Theorems 585 and 586: One takes a finite presentation of M, writes it down as matrix, and then puts it in Smith normal form. One then gets that

$$M \cong P/(d_{11}) \oplus P/(d_{22}) \oplus \ldots \oplus P/(d_{tt}).$$

Since d_{ii} divides d_{jj} for $j > i$, we have $(d_i) \supset (d_j)$. Now, if some d_{ii}'s are invertible, they result in zero summands in the expression above. So let us relabel by $d_1, \ldots, d_n$ the "non-invertible elements" of the list $d_{11}, \ldots, d_{tt}$, in the same order. Then each (d_i) is proper and

$$M \cong P/(d_1) \oplus P/(d_2) \oplus \ldots \oplus P/(d_n).$$ □

Corollary 697. *If A is a PID, "projective A-modules" and "free A-modules" are the same.*

Proof. Free modules are projective. Projective modules are by definition submodules of free modules; so if A is a PID, they are free by Theorem 696. □

Deeper thoughts 698. One may wonder if the converse holds, i.e. if every C-ring A such that all projective A-modules are free, must be a PID. The answer is negative. One beautiful way to see this is via the Quillen–Suslin theorem, which solved a famous conjecture by Serre: *for any field $\mathbb{K}$ and any positive integer n, if $S \stackrel{\text{def}}{=} \mathbb{K}[x_1, \ldots, x_n]$, then every projective S-module is free.*