4H GALOIS THEORY

CHRISTIAN VOIGT (2015) AND ULRICH KRÄHMER (2016)

Abstract. These are the lecture notes for MATHS4105 4H Galois Theory (17 lectures + 5 tutorials, 50 minutes each). Throughout, "ring" means unital associative ring, and ring homomorphisms are understood to preserve units.

Here is a rough time plan that allows for one reserve lecture in case a lecture gets cancelled, one is not fast enough in the beginning, wants to mention additional topics or revise and summarise the whole course.

(1) Moitation: solving polynomial equations by radicals (Lecture 1)
(2) Review: Rings, ideals, fields (Lectures 2,3)
(3) Polynomial rings and irreducibility (Lectures 4,5,6)
(4) Field extensions and their degree (Lecture 7)
(5) Algebraic extensions and algebraic closure (Lectures 8,9)
(6) Ruler and compass constructions (Lecture 10)
(7) Automorphisms, splitting fields, normal extensions (Lectures 11,12)
(8) Separable extensions and perfect fields (Lecture 13)
(9) The main theorem and examples (Lectures 14,15)
(10) the quintic equation (Lecture 16)

Suggestions for future lecturers: give out revision material just before start of term, Topics in Algebra prepares less than expected, especially linear algebra. Corollary 4.20 is used heavily at the end, can be presented more prominently. in the perfect field section we fall back to work inside $\mathbb{C}$ but the general case of characteristic 0 fields is not much harder. have tutorials every week.

Contents

1. Polynomial equations

In this introductory section we shall discuss the historical background of Galois theory, which is the problem of solving polynomial equations. Already the Babylonians were able to solve quadratic equations using the method of completing the square. Progress in the cubic and quartic cases was made during the Renaissance, most notably by the Italian mathematicians Tartaglia, Ferrari and Cardano in the 16th century. However, higher degree equations remained elusive, and it was Niels Henrik Abel who proved in 1824 that a general quintic equation cannot be solved by radicals. This fundamental result was significantly refined by Évariste Galois in 1830, who studied permutations of the roots of polynomial equations, and explained in this way how to decide wether a given quintic equation is solvable or not.

1.1. **Polynomials.**

If not stated otherwise, "*ring*" means unital associative commutative ring in this course, and a *ring homomorphism* is assumed to preserve the identity elements (see the next section for a review of some basic definitions from ring theory).

Definition 1.1. A *polynomial of degree* n with coefficients in a ring R is an expression of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

where $a_0, \ldots, a_n \in R$ and $a_n \neq 0$. One calls a_n the *leading coefficient* of $f(x)$, and $f(x)$ *monic* if $a_n = 1$. An element $c \in R$ is a *root* of $f(x)$ if

$$f(c) = a_n c^n + a_{n-1} c^{n-1} + \cdots + a_1 c + a_0 = 0.$$

If we add a polynomial 0 that we consider to be of degree $-\infty$ and define the notion of addition and multiplication of polynomials suggested by the notation, then the set of all polynomials forms a ring that we denote by $R[x]$.

To find the roots of a given polynomial, that is, to solve a single polynomial equation in one variable, is the topic we will be talking about in this course. If R is a ring with zero divisors, then studying this question will be a bit tedious and not many general statements can be made, so we will usually work under the assumption that R is an *integral domain.* We will also see in the next section that such rings can always be realised as subrings of a field; therefore, it is sufficient to study polynomials with coefficients in a field.

The first question one may ask about the roots of a given polynomial is whether they exist at all. This leads to:

Definition 1.2. A field R is *algebraically closed* if every nonconstant polynomial in $R[x]$ has at least one root.

Next, one can ask how many roots a polynomial has. One easily shows (using the division algorithm for polynomials) that $c \in R$ is a root of $f(x)$ if and only if there exists a polynomial $g(x)$ and some $m \in \mathbb{N}$ with

$$f(x) = (x - c)^m g(x).$$

The maximal m for which there exists such a $g(x)$ is the *multiplicity* of the root c, and as we will prove in Section 4.5, we have:

Theorem 1.3. *A polynomial of degree n with coefficients in a field R has at most n roots, counted with multiplicity, and exactly n if R is algebraically closed.*

Example 1.4. $x^n \in \mathbb{C}[x]$ has the single root $c = 0$ with multiplicity n.

So there is growing evidence that we might want to work over algebraically closed fields, and now the question is how to find these. In this respect, a key result is of course the *fundamental theorem of algebra*:

Theorem 1.5. *The field $\mathbb{C}$ of complex numbers is algebraically closed.*

Despite its name, the theorem is rather a part of analysis than of algebrai and all proofs use some analytic or topological argument - here is a short one based on complex analysis:

Proof. Assume that $f(x) \in \mathbb{C}[x]$ is a nonconstant polynomial. This defines a holomorphic function $f : \mathbb{C} \to \mathbb{C}$. Assume that f has no zeros. Then $g(x) = 1/f(x)$ defines a holomorphic function g on $\mathbb{C}$ as well. Since f is nonconstant we see that

$$\lim_{|z|\to\infty} |f(z)| = \infty.$$

It follows that g is bounded. Therefore g must be a constant function according to Liouville's theorem, which contradicts the fact that f is nonconstant. □

This should not be too surprising: recall that the existence of the natural numbers is postulated as an axiom (we work in ZFC throughout), and then one defines step by step larger sets of numbers in terms of the already constructed one,

$$\mathbb{N} \to \mathbb{Z} \to \mathbb{Q} \to \mathbb{R} \to \mathbb{C},$$

but the construction of $\mathbb{R}$ out of $\mathbb{Q}$ is about analysis and topology, so the complex numbers inherit this nonalgebraic nature from the reals.

The final question one can ask is how to concretely determine the roots of $f(x)$, and this is one of the key applications and the historic motivation of Galois theory: by the end of this course we will learn to decide which polynomial eqations can be solved by *radical expressions* (or radicals in short) in the coeffiicents, that is, formulas obtained by taking iteratively sums, differences, products, quotients, and also k-th roots.

Firstly, in order to motivate the problem of finding solutions to polynomial equations a bit further, we shall give a brief review of its history, and discuss explicitly how solutions by radicals can be written down for complex polynomials of degree $n = 1, 2, 3, 4$. To simplify the presentation, we consider monic polynomials; the general case can be easily reduced to this one by applying the theory to $\frac{1}{a_n} f(x)$.

1.2. **Linear equations.** This is the case $n = 1$, and it is really elementary. The linear equation

$$f(x) = x + a_0 = 0$$

has the unique solution $x = -a_0$.

1.3. **Quadratic equations.** This is the case $n = 2$. The quadratic equation

$$f(x) = x^2 + a_1 x + a_0 = 0$$

can be solved by completing the square. We rewrite the equation as

$$\left(x + \frac{a_1}{2}\right)^2 - \frac{a_1^2}{4} + a_0 = 0$$

and obtain the solutions

$$x = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_0}. \tag{1.1}$$

Although this formula should look very familiar, be aware that

$$\frac{a_1^2}{4} - a_0$$

is in general not a real number since the coefficients a_1, a_0 are complex numbers. Accordingly, it is not entirely obvious how the square root has to be interpreted.

Our notation for the square root $\sqrt{c}$ of a complex number c means that we fix $\sqrt{c} \in \mathbb{C}$ satisfying $\sqrt{c}^2 = c$. Observe that the fundamental theorem of algebra applied to the polynomial $x^2 - c$ ensures that we find such a number $\sqrt{c}$. For $c = 0$ this actually fixes $\sqrt{c} = 0$ uniquely, but for $c \neq 0$ there are exactly two possible choices. We obtain one from the other by multiplication with -1. Note that our choice of $\sqrt{c}$ does not affect the set of solutions since the square root appears with both signs in formula (1.1).

1.4. **Cubic equations.** This is the case $n = 3$. The history of the cubic equation is quite interesting, but we shall not review it in detail here. The formula we present below is usually referred to as Cardano's formula. Gerolamo Cardano (1501-1576) stated it in his book *Ars magna de Regulis Algebraicis* in 1545, but he had actually learnt it from Niccolò Tartaglia (1500-1557).

Let us discuss the method leading to Cardano's formula. The cubic equation

$$f(x) = x^3 + a_2 x^2 + a_1 x + a_0 = 0 \tag{1.2}$$

can first be simplified similarly to the procedure used for completing the square. More precisely, we rewrite

$$x^3 + a_2 x^2 = -a_1 x - a_0$$

as

$$\left(x + \frac{a_2}{3}\right)^3 = \frac{a_2^2}{3}x + \frac{a_2^3}{27} - a_1 x - a_0 = \left(\frac{a_2^2}{3} - a_1\right)x + \frac{a_2^3}{27} - a_0.$$

Setting

$$y = x + \frac{a_2}{3}$$

we obtain

$$y^3 = \left(\frac{a_2^2}{3} - a_1\right)\left(y - \frac{a_2}{3}\right) + \frac{a_2^3}{27} - a_0 = \left(\frac{a_2^2}{3} - a_1\right)y - \frac{2a_2^3}{27} + \frac{a_1 a_2}{3} - a_0.$$

If we write

$$p = a_1 - \frac{a_2^2}{3}, \qquad q = a_0 - \frac{a_1 a_2}{3} + \frac{2a_2^3}{27},$$

we thus arrive at the cubic equation

$$y^3 + py + q = 0 \tag{1.3}$$

which has no quadratic term anymore; solving this reduced cubic equation (1.3) is equivalent to solving our original equation (1.2). The reduction from equation (1.2) to equation (1.3) is sometimes called *Tschirnhaus transformation.*

Let us first assume that p is nonzero. Then

$$-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

is nonzero. Indeed, if this expression were zero then

$$\frac{q^2}{4} = \left(-\frac{q}{2}\right)^2 = \left(\sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\right)^2 = \frac{q^2}{4} + \frac{p^3}{27}$$

would imply $p = 0$. We may therefore fix a cubic root

$$u = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

and according to our previous consideration this number $u \in \mathbb{C}$ is nonzero. By construction, we then have

$$\begin{aligned} u^6 + u^3 q - \frac{p^3}{27} &= \left(-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\right)^2 - \frac{q^2}{2} - q\sqrt{\frac{q^2}{4} + \frac{p^3}{27}} - \frac{p^3}{27} \\ &= \frac{q^2}{4} + q\sqrt{\frac{q^2}{4} + \frac{p^3}{27}} + \frac{q^2}{4} + \frac{p^3}{27} - \frac{q^2}{2} - q\sqrt{\frac{q^2}{4} + \frac{p^3}{27}} - \frac{p^3}{27} \\ &= 0, \end{aligned}$$

that is,

$$u^6 + u^3 q - \frac{p^3}{27} = 0, \tag{1.4}$$

compare subsection 1.3.
Let us consider the expression

$$y = u - \frac{p}{3u},$$

note here that u is nonzero to see that this is well-defined. We compute

$$\begin{aligned} y^3 + py + q &= \left(u - \frac{p}{3u}\right)^3 + p\left(u - \frac{p}{3u}\right) + q \\ &= u^3 - 3u^2\frac{p}{3u} + 3\frac{up^2}{9u^2} - \frac{p^3}{27u^3} + pu - \frac{p^2}{3u} + q \\ &= u^3 - \frac{p^3}{27u^3} + q = 0 \end{aligned}$$

using equation (1.4) in the last step. In other words,

$$y = u - \frac{p}{3u} = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} - \frac{p}{3\sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}}.$$

is a solution of the reduced cubic equation (1.3). We may rewrite this using

$$\begin{aligned} -\frac{p^3}{27\left(-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\right)} &= -\frac{p^3\left(-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}\right)}{27\left(\frac{q^2}{4} - \left(\frac{q^2}{4} + \frac{p^3}{27}\right)\right)} \\ &= -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} \end{aligned}$$

in the form

$$y = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \tag{1.5}$$

which is usually called Cardano's formula. Observe that the choice of the cubic root in the second term of this formula is already fixed by our choice of the cubic root defining u, this is implicit in Cardano's formula.

A nice feature of (1.5) is that it also provides a solution of the reduced cubic equation when $p = 0$. Indeed, if $p = 0$ then one of the two summands in equation (1.5) is zero, and the other one reduces to a cubic root of $-q$ as desired. To summarise, Cardano's formula describes in fact a solution of the reduced cubic equation (1.3) for all possible coefficients p, q.

According to the fundamental theorem of algebra there exist two further complex solutions of this equation. To determine the remaining solutions it suffices to replace u in our above argument by ζu and $\zeta^2 u$ where

$$\zeta = \exp(2\pi i/3) = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$

is a primitive third root of unity. If we fix the choice of cubic roots for our first solution as in Cardano's formula then the remaing solutions are given by

$$\zeta\sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \zeta^2\sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

and

$$\zeta^2\sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \zeta\sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

respectively.

To check that these are indeed the solutions let us assume that p is nonzero and write

$$y_1 = u - \frac{p}{3u}, \qquad y_2 = \zeta u - \frac{\zeta^2 p}{3u}, \qquad y_3 = \zeta^2 u - \frac{\zeta p}{3u}$$

for these numbers. Then we obtain

$$(y - y_1)(y - y_2)(y - y_3) = y^3 - (y_1 + y_2 + y_3)y^2 + (y_1y_2 + y_1y_3 + y_2y_3)y - y_1y_2y_3.$$

Since $1 + \zeta + \zeta^2 = 0$ we see that $y_1 + y_2 + y_3 = 0$. We have

$$\begin{aligned} y_1y_2 + y_1y_3 + y_2y_3 &= \left(u - \frac{p}{3u}\right)\left(\zeta u - \frac{\zeta^2 p}{3u}\right) + \left(u - \frac{p}{3u}\right)\left(\zeta^2 u - \frac{\zeta p}{3u}\right) \\ &\quad + \left(\zeta u - \frac{\zeta^2 p}{3u}\right)\left(\zeta^2 u - \frac{\zeta p}{3u}\right) \\ &= (\zeta + \zeta^2 + 1)u^2 + (\zeta^2 + \zeta + 1)\frac{p^2}{3u^2} \\ &\quad - \frac{p}{3}((\zeta + \zeta^2) + (\zeta^2 + \zeta) + (\zeta + \zeta^2)) \\ &= -p(\zeta + \zeta^2) = p \end{aligned}$$

since $\zeta + \zeta^2 = -1$. Finally, we compute

$$\begin{aligned} y_1y_2y_3 &= \left(u - \frac{p}{3u}\right)\left(\zeta u - \frac{\zeta^2 p}{3u}\right)\left(\zeta^2 u - \frac{\zeta p}{3u}\right) \\ &= u^3 - \frac{p^3}{27u^3} + \frac{p^2}{9u}(\zeta + \zeta^2 + 1) - \frac{pu}{3}(\zeta^2 + \zeta + 1) \\ &= u^3 - \frac{p^3}{27u^3} = -q \end{aligned}$$

so that the above polynomial equals

$$(y - y_1)(y - y_2)(y - y_3) = y^3 + py^2 + qy.$$

This shows that y_1, y_2, y_3 are indeed precisely the required roots. Again, this holds independently of whether p is zero or nonzero.

If the coefficients p and q are real numbers, which happens if a_0, a_1, a_2 are real, there always exists a real solution of the cubic equation (1.2). This follows from the fundamental theorem of algebra and the fact that complex conjugation must preserve the set of solutions. More precisely, since there are precisely three complex solutions we see that one of them must be fixed under complex conjugation. If the discriminant

$$\frac{q^2}{4} + \frac{p^3}{27}$$

is a nonnegative real number, then formula (1.5) produces a real solution if we choose the canonical real cubic root to define u. Note that every real number r, no matter whether it is positive or not, has the real cubic root

$$\operatorname{sign}(r)\sqrt[3]{|r|}$$

where $\operatorname{sign}(r) = \pm 1$ denotes the sign of r, and $\sqrt[3]{|r|}$ is the unique positive cubic root of $|r|$.

1.5. **Quartic equations.** This is the case $n = 4$. The first complete solution for the quartic equation was given by Lodovico Ferrari (1522-1565). It was published by Ferrari's teacher Gerolamo Cardano in the book *Ars magna de Regulis Algebraicis* in 1545.

Let us explain Ferrari's solution. In the same way as in the cubic case, the quartic equation

$$f(x) = x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0 \tag{1.6}$$

can be simplified using a Tschirnhaus transformation. More precisely, using

$$\left(x + \frac{a_3}{4}\right)^4 = x^4 + a_3x^3 + 6\frac{a_3^2}{4^2}x^2 + 4\frac{a_3^3}{4^3}x + \frac{a_3^4}{4^4}$$

we can rewrite this equation in the form

$$\begin{aligned}\left(x + \frac{a_3}{4}\right)^4 &= \frac{3a_3^2}{8}x^2 + \frac{a_3^3}{16}x + \frac{a_3^4}{256} - a_2x^2 - a_1x - a_0 \\ &= \left(\frac{3a_3^2}{8} - a_2\right)x^2 + \left(\frac{a_3^3}{16} - a_1\right)x + \frac{a_3^4}{256} - a_0.\end{aligned}$$

Hence setting

$$y = x + \frac{a_3}{4}$$

yields

$$\begin{aligned}y^4 &= \left(\frac{3a_3^2}{8} - a_2\right)\left(y - \frac{a_3}{4}\right)^2 + \left(\frac{a_3^3}{16} - a_1\right)\left(y - \frac{a_3}{4}\right) + \frac{a_3^4}{256} - a_0 \\ &= \left(\frac{3a_3^2}{8} - a_2\right)y^2 - \left(\frac{3a_3^3}{16} - \frac{a_2a_3}{2}\right)y + \left(\frac{3a_3^4}{128} - \frac{a_2a_3^2}{16}\right) \\ &\quad + \left(\frac{a_3^3}{16} - a_1\right)y - \frac{a_3^4}{64} + \frac{a_1a_3}{4} + \frac{a_3^4}{256} - a_0 \\ &= \left(\frac{3a_3^2}{8} - a_2\right)y^2 - \left(\frac{a_3^3}{8} - \frac{a_2a_3}{2} + a_1\right)y + \frac{3a_3^4}{256} - \frac{a_2a_3^2}{16} + \frac{a_1a_3}{4} - a_0.\end{aligned}$$

We thus obtain

$$y^4 + py^2 + qy + r = 0 \tag{1.7}$$

where

$$p = a_2 - \frac{3a_3^2}{8}, \qquad q = a_1 + \frac{a_3^3}{8} - \frac{a_2a_3}{2}, \qquad r = a_0 - \frac{3a_3^4}{256} + \frac{a_2a_3^2}{16} - \frac{a_1a_3}{4}.$$

We conclude that it suffices to solve equation (1.7).

In order to do this we may assume without loss of generality that both r and q are nonzero. Indeed, if $r = 0$ it is evident that $y = 0$ is a solution of (1.7). We may then divide by y to reduce the problem to a cubic equation, which in turn can be solved by the method described in the previous subsection. In the case $q = 0$ equation (1.7) reduces to a quadratic equation for y^2. Solving this quadratic equation allows us to easily determine the solutions of (1.7) in this case as well.

Let us therefore assume that both r and q are nonzero, and let us choose $u \in \mathbb{C}$ such that

$$q^2 - 4(2u - p)(u^2 - r) = -8u^3 + 4pu^2 + 8ru - 4pr + q^2 = 0.$$

Observe that this expression is a cubic equation for u, therefore we can apply the procedure from the previous subsection to do this.

Moreover, note that both $2u - p$ and $u^2 - r$ are nonzero since we assume that q is nonzero. We choose square roots of these numbers such that

$$q = 2\sqrt{2u - p}\sqrt{u^2 - r}. \tag{1.8}$$

Observe that fixing the square root $\sqrt{u^2 - r}$ determines the root $\sqrt{2u - p}$ uniquely, that is, there are no further sign ambiguities.

Let us now assume in addition that $v \in \mathbb{C}$ is a solution of equation (1.7). Such a solution exists by the fundamental theorem of algebra. Since we assume $r \neq 0$ the number v is nonzero.

Let us consider the complex number $z = v^2 + u$. Since v is a solution of (1.7) we obtain

$$\begin{aligned} z^2 &= v^4 + 2uv^2 + u^2 \\ &= -pv^2 - qv - r + 2uv^2 + u^2 \\ &= (2u - p)v^2 - qv + u^2 - r \\ &= (\sqrt{2u - p}\, v - \sqrt{u^2 - r})^2, \end{aligned}$$

using formula (1.8) in the last step. Hence

$$v^2 = z - u^2 = -u^2 \pm (\sqrt{2u - p}\, v - \sqrt{u^2 - r}).$$

Equivalently, we may write

$$v^2 \pm \sqrt{2u - p}\, v \mp \sqrt{u^2 - r} + u^2 = 0.$$

This is a quadratic equation with solutions

$$v = -\frac{\sqrt{2u - p}}{2} \pm \sqrt{\frac{2u - p}{4} + \sqrt{u^2 - r} - u^2}$$

and

$$v = \frac{\sqrt{2u - p}}{2} \pm \sqrt{\frac{2u - p}{4} - \sqrt{u^2 - r} + u^2},$$

respectively.

In this way we have actually obtained explicit formulas for four solutions of equation (1.7). Since u can be expressed in terms of radicals the same holds for these solutions.

1.6. **Exercises.** This first set of exercises should help you test your knowledge of material from previous semesters and give you an indication what notions and techniques will be used as prerequisites in 4H Galois theory. Some of these you will have forgotten, some you might not have understood back when they were taught, and in this case it is time to refresh your memory and close the gaps, and we will also use the first week of lectures and the first tutorial for this.

Each exercise has a letter grade assigned to it, as a rough guide which level of attainment of the relevant intended learning outcomes I think it can be used well to assess. This refers to the *At the end of this course, students should be able to...* with an emphasis on *end*, so you should revisit this when preparing for the exams in spring and I will do so, too.

Exercise 1.1 (C). State the definition of a *ring* and of a *ring homomorphism*.

Exercise 1.2 (C). Give two examples of rings, one commutative, and one non-commutative.

Exercise 1.3 (C). State the definition of the *complex conjugate* $\overline{z}$ and of the *absolute value* $|z|$ of a complex number z.

Exercise 1.4 (C). State the definition of the *polar form* of a complex number. Determine the polar form of

$$z = -\frac{1}{2} + i\frac{\sqrt{3}}{2}.$$

Draw the location of this number in the complex plane, and compute z^3.

Exercise 1.5 (C). State the definition of a *divisor* of a natural number $n \in \mathbb{N}$ and of the *greatest common divisor* $gcd(m, n)$ of two natural numbers $m, n \in \mathbb{N}$. Determine $gcd(45, 12)$ and $gcd(23, 315)$.

Exercise 1.6 (B). State the definition of a (two-sided) *ideal* in a ring. Verify that $n\mathbb{Z} \subset \mathbb{Z}$ is an ideal for any $n \in \mathbb{Z}$.

Exercise 1.7 (B). Let $f : R \to S$ be a ring homomorphism.

(1) State the definition of the *kernel* $\ker(f)$ of f. Show that $\ker(f)$ is an ideal in R.
(2) How is the *image* $\mathrm{im}(f)$ of f defined? Show that $\mathrm{im}(f)$ is a subring of S.

Exercise 1.8 (B). Let R be a ring. Show that the unit element $1 \in R$ is uniquely determined. That is, if $1' \in R$ is any element satisfying $1'a = a = a1'$ for all $a \in R$ then $1' = 1$.

Exercise 1.9 (B). State the definition of a *subring* of a ring. Is $2\mathbb{Z} \subset \mathbb{Z}$ a subring? Justify your answer.

Exercise 1.10 (B). Show that

$$\mathbb{Z}[\sqrt{5}] = \{a + b\sqrt{5} \in \mathbb{R} \mid a, b \in \mathbb{Z}\}$$

is a subring of $\mathbb{R}$.

Exercise 1.11 (B). Using *long division of polynomials*, find a polynomial of highest possible degree dividing $f(x) = x^3 + x^2 + x - 3$ and $g(x) = 2x^4 + x^3 - 2x^2 - 13x - 6$.

Exercise 1.12 (A). Determine all ideals in the ring $\mathbb{Z}$.

Exercise 1.13 (A). Determine all ring homomorphisms $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$.

2. Rings, fields and integral domains

In this section we discuss some basic material on rings, integral domains and fields. For further information and details consult any textbook on general or on commutative algebra.

2.1. **Rings.** Let us begin with the definition of a ring.

Definition 2.1. A *ring* is a set R together with two $+, \cdot : R \times R \to R$ called addition and multiplication and an element $1 \in R$ that satisfy the following conditions:

a) $(R, +)$ is an abelian group.
b) (Associativity) We have $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in R$.
c) (Unitarity) We have $1 \cdot a = a = a \cdot 1$ for all $a \in R$.
d) (Distributivity) We have

$$(a + b) \cdot c = (a \cdot c) + (b \cdot c), \qquad c \cdot (a + b) = (c \cdot a) + (c \cdot b)$$

for all $a, b, c \in R$.

We will usually write ab instead of $a \cdot b$. The unit element of the abelian group $(R, +)$ will be denoted by 0. Unlike some authors, we do not require that $1 \neq 0$. Note that $1 = 0$ holds iff $R = 0$, that is, iff R consists of exactly one element.

Definition 2.2. Let R and S be rings. A *homomorphism* of rings is a homomorphism $f :: R \to S$ of abelian groups such that $f(ab) = f(a)f(b)$ and $f(1) = 1$ for all $a, b \in R$.

A ring *isomorphism* is a bijective ring homomorphism. If $f : R \to S$ is a ring isomorphism, then also the inverse map $f^{-1} : S \to R$ is a ring isomorphism.

The image

$$\operatorname{im}(f) = f(R) = \{f(a) \mid a \in R\} \subset S$$

of a ring homomorphism $f : R \to S$ is a subring of S.

Definition 2.3. A ring R is called *commutative* if $ab = ba$ for all $a, b \in R$.

We will almost exclusively be interested in commutative rings in this course, but generalisations of some results are in fact an active current field of research and a great topic for an MSci project.

2.2. **Ideals and quotient rings.** Ideals in rings play a role analogous to normal subgroups of groups.

Definition 2.4. Let R be a ring. A (two-sided) ideal in R is a subset $I \subset R$ such that

a) I is a subgroup of the abelian group $(R, +)$.
b) For all $a \in R$ and $x \in I$ we have $ax \in I$ and $xa \in I$.

Every ring R has at least two ideals, namely the zero ideal $\{0\}$ and the ideal $I = R$. If R is commutative it is of course enough in *b*) to require $ax \in I$ for all $a \in R$ and $x \in I$.

If $f : R \to S$ is a ring homomorphism, then

$$\ker(f) = \{a \in R \mid f(a) = 0\} \subset R$$

is an ideal in R.

Since R is an abelian group with respect to addition, any abelian subgroup of $(R, +)$ is normal. In particular, if $I \subset R$ is an ideal then the quotient group R/I is naturally defined. Recall that R/I consists of all cosets $a + I$ for $a \in R$.

A proof of the following proposition can be found for example in Fraleigh's book, see Theorem 26.9 and Corollary 26.14 in [2].

Proposition 2.5. *Let R be a ring and let $I \subset R$ be an ideal. Then R/I is a ring with addition and multiplication defined by*

$$(a + I) + (b + I) = (a + b) + I$$

and

$$(a + I)(b + I) = ab + I,$$

and with unit element $1 + I$, respectively.

We call R/I the quotient ring of R by the ideal I. Observe that the canonical map $\pi : R \to R/I$ given by $\pi(a) = a + I$ is a ring homomorphism.

Theorem 2.6 (First isomorphism theorem for rings)**.** *Let $f : R \to S$ be a ring homomorphism. Then there is a ring isomorphism*

$$R/\ker(f) \cong \mathrm{im}(f).$$

More precisely, there exists an injective ring homomorphism $F : R/\ker(f) \to S$ such that the diagram

$$\begin{tikzcd} R \arrow[r, "\pi"] \arrow[dr, "f"'] & R/\ker(f) \arrow[d, "F"] \\ & S \end{tikzcd}$$

is commutative.

For a proof of Theorem 2.6 we refer to Fraleigh's book, see Section 26 and in particular Theorem 26.17 in [2]. The map F is explicitly given by the formula $F(a + I) = f(a)$ for $a + I \in R/I$.

Similarly, the other isomorphism theorems and the correspondence theorem admit straightforward generalisations from groups to rings. We will only need the correspondence theorem:

Theorem 2.7 (Correspondence theorem for rings)**.** *If $I \subset R$ is an ideal, then*

$$J \mapsto \{a + I \mid a \in J\}$$

defines a bijection between the set of ideals $J \subset R$ containing I and the set of ideals in R/I. This preserves the partial orderingis of the two sets given by $\subset$.

If R is a ring and $(I_\lambda)_{\lambda \in \Lambda}$ is a family of ideals in R, then the intersection

$$\bigcap_{\lambda \in \Lambda} I_\lambda$$

is an ideal in R as well.

Definition 2.8. Let R be a ring and let $X \subset R$ be a subset. The ideal (X) in R generated by X is defined as the intersection of all ideals of R containing X. That is,

$$(X) = \bigcap_{\substack{I \subset R \text{ ideal} \\ X \subset I}} I.$$

Note that the set of ideals we intersect over is nonempty; indeed, it always contains the ideal $I = R$.

Proposition 2.9. *Let R be a ring and let $X \subset R$ be a subset. The ideal (X) consists precisely of all finite sums*

$$\sum_{i=1}^{n} a_i x_i b_i$$

where $n \in \mathbb{N}$, $a_i, b_i \in R$ and $x_i \in X$ for all $i = 1, \dots, n$. If R is commutative it suffices to consider all finite sums of the form

$$\sum_{i=1}^{n} a_i x_i$$

where $n \in \mathbb{N}$, $a_i \in R$ and $x_i \in X$ for all $i = 1, \dots, n$.

Proof. Let us write $[X]$ for the set of all finite sums as indicated above. Then $X \subset [X]$ because $x = 1x1 \in [X]$ for all $x \in X$. More precisely, to see that x is of the form above we may take $n = 1$ and $a_1 = b_1 = 1$ as well as $x_1 = x$. By construction, the set $[X]$ is closed under addition, and since

$$-\left(\sum_{i=1}^{n} a_i x_i b_i\right) = \sum_{i=1}^{n} (-a_i) x_i b_i,$$

we conclude that it is a subgroup of the additive group of R. Moreover, from the relations

$$a\left(\sum_{i=1}^{n} a_i x_i b_i\right) = \sum_{i=1}^{n} (a a_i) x_i b_i, \qquad \left(\sum_{i=1}^{n} a_i x_i b_i\right) a = \sum_{i=1}^{n} a_i x_i (b_i a),$$

we see that $[X]$ is closed under left and right multiplication by elements of R. We conclude that $[X]$ is an ideal.

Now let $I \subset R$ be any ideal containing X. Then $axb \in I$ for all $a, b \in R$ and $x \in X$ by the ideal property. Since I is additively closed we see that all finite sums

$$\sum_{i=1}^{n} a_i x_i b_i$$

with $a_i, b_i \in R$ and $x_i \in X$ for all i are contained in I. That is, we have $[X] \subset I$. By definition of (X) we therefore obtain

$$[X] \subset \bigcap_{\substack{I \subset R \text{ ideal} \\ X \subset I}} I = (X).$$

Since $[X]$ is itself an ideal containing X it appears in the set of ideals we intersect over on the right hand side of the previous formula. Hence this inclusion is in fact an equality, that is, we have $[X] = (X)$ as desired.

The remaining assertion concerning commutative rings is obvious, indeed, in this case we can rewrite $a_i x_i b_i = a_i b_i x_i$ for all i. Therefore it is enough to multiply elements of R from the left to elements of the generating set X. □

2.3. **Fields.** For most of the course we will be mainly interested in fields. Fields form a very particular class of commutative rings. In order to define the notion of a field we first introduce the following concept.

Definition 2.10. Let R be a ring. An element $u \in R$ is called a unit if it is invertible, that is, if there exists an element $v \in R$ such that $uv = 1 = vu$.

One can show that the element v in this definition is unique, and we usually write $v = u^{-1}$. We write $R^{\times}$ for the set of invertible elements in R. With the multiplication induced from R, the set $R^{\times}$ is a group with neutral element 1.

Definition 2.11. A field is a commutative ring K such that $1 \neq 0$ and every nonzero element of K is a unit.

Prominent examples of fields are $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$.

Proposition 2.12. *A commutative ring K is a field if and only if the only ideals in K are $\{0\}$ and K.*

Proof. If $I \subset K$ is a nonzero ideal, then there exists a nonzero element $u \in I$. If K is a field, u is invertible, and by the ideal property we find $1 = u^{-1}u \in I$. Hence the ideal $(1) = K$ generated by 1 is contained in I, which means $I = K$. Cinversely, if K contains only two ideals, then $(u) = K$ for every nonzero element $u \in K$, that is, $1 \in (u)$, which means u is invertible. □

Corollary 2.13. *Let K be a field and let $f : K \to R$ be a ring homomorphism to a nonzero ring R. Then f is injective.*

Proof. The kernel $\ker(f)$ is an ideal of K, and since $f(1) = 1 \neq 0$ by our assumption that R is not the zero ring we have $\ker(f) \neq K$. Hence $\ker(f) = 0$ according to Proposition 2.12. □

2.4. **Integral domains.** Even though our main interest will be in fields, we will also have to work with commutative rings which are not fields.

Definition 2.14. Let R be a commutative ring. A zero-divisor in R is a nonzero element $a \in R$ for which there exists another nonzero element $b \in R$ such that $ab = 0$.

If a ring does not contain any zero divisors then many familiar arguments for solving equations are valid.

Definition 2.15. A commutative ring R is called an integral domain if $1 \neq 0$ and R does not contain zero-divisors.

A basic example of an integral domain is the ring $\mathbb{Z}$ of integers.

In an integral domain we may cancel elements in equations in the same way as we would do when computing with, say, real numbers. More precisely, if

$$ra = rb$$

for some elements $r, a, b \in R$ with r nonzero, then we have in fact $a = b$. To see this, consider the difference $d = a - b$. The above relation means $rd = r(a - b) = 0$. Since r is assumed to be nonzero we conclude $d = 0$ because there are no zero-divisors in R. This means of course nothing but $a = b$.

Proposition 2.16. *Every field is an integral domain.*

Proof. Assume that K is a field and let $a, b \in K$ such that $ab = 0$. If a is nonzero then multipliying this equation with a^{-1} yields $b = 1b = a^{-1}ab = 0$. Hence K does not contain zero divisors. □

2.5. **Principal ideal domains.** Principal ideal domains are special integral domains, namely those for which all ideals are principal ideals.

Definition 2.17. Let R be a commutative ring. An ideal $I \subset R$ is called a principal ideal if it is generated by a single element. That is, we have

$$I = \{rp \mid r \in R\}$$

for some element $p \in R$. We write $I = (p)$ in this case.

If R is any commutative ring then the ideal $I = R$ is a principal ideal generated by the unit element 1, that is, we have $(1) = R$. However, in general there may be many other ideals in R which are not generated by just a single element.

Definition 2.18. A commutative ring R is called a principal ideal ring if every ideal in R is principal. If in addition R is an integral domain then R is called a principal ideal domain.

The ring $R = \mathbb{Z}$ is a principal ideal domain. In fact, every ideal of $\mathbb{Z}$ is of the form $(n) = n\mathbb{Z}$ for some $n \in \mathbb{Z}$. This can be shown using the Euclidean algorithm.

2.6. **Maximal ideals and prime ideals.** Many examples of fields and integral domains can be constructed as quotient rings. In order to explain this let us introduce some terminology.

Definition 2.19. Let R be a commutative ring.

a) An ideal $I \subset R$ is *prime* if $I \neq R$ and $pq \in I \Rightarrow p \in I$ or $q \in I$ for all $p, q \in R$.
b) An ideal $I \subset R$ is *maximal* if $I \neq R$ and for every ideal $J \subset R$ such that $I \subset J \subset R$ we have either $J = I$ or $J = R$.

The following theorem explains the link between prime ideals and maximal ideals on the one hand, and integral domains and fields on the other.

Theorem 2.20. *Let R be a commutative ring and let $I \subset R$ be an ideal.*

a) The ideal I is maximal iff the quotient ring R/I is a field.
b) The ideal I is prime iff the quotient ring R/I is an integral domain.

Proof. *a*) follows directly from the correspondence theorem (Theorem 2.7) and Proposition 2.12.

b) We have $(a + I)(b + I) = ab + I = 0 + I$ (the zero element in R/I) iff $ab \in I$, and $a + I = 0 + I$ iff $a \in I$ and $b + I = 0 + I$ iff $b \in I$. □

Using Theorem 2.20 we can exhibit further examples of integral domains and fields.

Proposition 2.21. *For any $p \in \mathbb{N} \setminus \{0\}$, the following statements are equivalent.*

a) p is a prime number.
b) The ideal (p) in $\mathbb{Z}$ is a prime ideal.
c) The ideal (p) in $\mathbb{Z}$ is a maximal ideal.
d) The quotient ring $\mathbb{Z}/(p)$ is an integral domain.
e) The quotient ring $\mathbb{Z}/(p)$ is a field.

Proof. *a*) $\Leftrightarrow$ *b*) If p is prime and $ab \in (p)$, then $ab = pk$ for some $k \in \mathbb{Z}$, which means that either a or b is divisible by p. Hence $a \in (p)$ or $b \in (p)$ which means that (p) is a prime ideal. Conversely, assume that (p) is a prime ideal and that

$ab = p$. Then $ab \in (p)$, so either $a \in (p)$ or $b \in (p)$, say the former. Then $a = pr$ for some $r \in \mathbb{Z}$, hence $ab = prb = p$ implies $rb = 1$ hence $b = \pm 1$ and $a = \pm p$, so p is prime.

a) $\Leftrightarrow$ *c)* Two integers p, q are coprime (have greatest common divisor 1) iff there are $m, n \in \mathbb{Z}$ such that $mq + np = 1$ (if you forgot how this works wait until we recall the analogous statement for polynomials and reconstruct the proof for $\mathbb{Z}$!). Now note that $mq + np = 1$ means precisely that the residue classes $m + (p)$ and $q + (p)$ of m and q are inverses of each other in $\mathbb{Z}_p = \mathbb{Z}/(p)$.

b) $\Leftrightarrow$ *d)* and *c)* $\Leftrightarrow$ *e)* are a consequence of Theorem 2.20. □

We conclude in particular that $\mathbb{Z}/(p)$ is a field for every prime p.

Definition 2.22. We will write $\mathbb{F}_p = \mathbb{Z}/(p)$ in the sequel and refer to this field as the field with p elements.

2.7. **Prime fields and characteristic.** A basic invariant of a field is its characteristic. In this subsection we introduce this concept and the closely related notion of a prime field.

If L is a field then a subfield $K \subset L$ is subring which is itself a field. A basic example is $\mathbb{R}$ as a subfield of $\mathbb{C}$.

Definition 2.23. The *prime field* of a field K is the smallest subfield of K.

Equivalently, K is the intersection of all subfields of K, or the subfield generated by $1 \in K$. Note that arbitrary intersections of subfields are again subfields.

Theorem 2.24. *Let K be a field. The prime field of K is either isomorphic to $\mathbb{Q}$ or to $\mathbb{F}_p$ for a uniquely determined prime p.*

Proof. Let us write $P \subset K$ for the prime field of K. We first define a map $f : \mathbb{Z} \to K$ by setting

$$f(n) = \begin{cases} 1 + 1 + \cdots + 1 \ \ (n \text{ times}) & \text{if } n > 0 \\ 0 & \text{if } n = 0 \\ -1 - 1 - \cdots - 1 \ \ (-n \text{ times}) & \text{if } n < 0. \end{cases}$$

It is easy to verify that this is indeed a ring homomorphism. Moreover, we have $\operatorname{im}(f) \subset P$ as P is the subfield generated by 1. If $\ker(f) = 0$ then $\operatorname{im}(f) \cong \mathbb{Z}$ according to the first isomorphism theorem. Moreover, all nonzero elements of $\operatorname{im}(f)$ are invertible because K is a field. Hence, by the very definition of the rational numbers $\mathbb{Q}$, we can extend f to an injective homomorphism $F : \mathbb{Q} \to K$ by setting

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

This implies $P \cong \mathbb{Q}$.

If $\ker(f) = (p)$ for some $p > 0$, then p must be a prime according to Proposition 2.21. Indeed, any subring of a field is an integral domain, and $\mathbb{Z}/\ker(f)$ is isomorphic to a subring of K according to the first isomorphism theorem. Since $1 \in P$ we see that $\operatorname{im}(f) \cong \mathbb{Z}/\ker(f) = \mathbb{Z}/(p) = \mathbb{F}_p$ is in fact the prime field of K in this case. □

Definition 2.25. Let K be a field. The *characteristic* of K is defined as $\operatorname{char}(K) = 0$ if the prime field of K is $\mathbb{Q}$, and $\operatorname{char}(K) = p$ if the prime field of K is $\mathbb{F}_p$.

2.8. **The field of quotients of an integral domain.** In many examples we obtain integral domains naturally as subrings of fields. A prototypical case is the ring $\mathbb{Z}$ as a subring of $\mathbb{Q}$. The main aim of this subsection is to show that this is actually the most general way in which an integral domain can appear:

Theorem 2.26. *Let R be an integral domain. Then there exists a field* $\mathrm{Quot}(R)$ *together with an injective ring homomorphism $\iota : R \to \mathrm{Quot}(R)$ which has the property that for every injective ring homomorphism $f : R \to K$ into a field K there exists a unique homomorphism of fields $F : \mathrm{Quot}(R) \to K$ such that the diagram*

$$\begin{tikzcd}
R \arrow[r, "\iota"] \arrow[dr, "f"'] & \mathrm{Quot}(R) \arrow[d, "F"] \\
& K
\end{tikzcd}$$

is commutative.

Proof. On the set

$$P(R) = \{(a,b) \mid a, b \in R \text{ and } b \neq 0\} \subset R \times R$$

we define a relation

$$(a,b) \sim (c,d) :\Leftrightarrow ad = bc.$$

It is straightforwardly verified that this is an equivalence relation: reflexivity and symmetry are immediate from the definition, and if $(a,b) \sim (c,d)$ and $(c,d) \sim (e,f)$i, then we have $ad = bc$ and $cf = de$, so that

$$daf = adf = bcf = bde = dbe.$$

Since R is an integral domain we can cancel d from this equation to obtain $af = be$. This means $(a,b) \sim (e,f)$, so that $\sim$ is also transitive.

We now define

$$\mathrm{Quot}(R) := P(R)/\sim$$

and denote the equivalence class of (a,b) by

$$\frac{a}{b}.$$

Think of a formal fraction of two elements of R, generalising the construction of $\mathbb{Q}$ out of $\mathbb{Z}$. Note that by definition of $\sim$, we have

$$\frac{a}{b} = \frac{c}{d} \quad \Leftrightarrow \quad ad = bc.$$

We leave it to the reader to verify that

$$\frac{a}{b} + \frac{c}{d} := \frac{ad+bc}{bd}, \qquad \frac{a}{b}\frac{c}{d} := \frac{ac}{bd}$$

yield a well-defined addition and multiplication on $\mathrm{Quot}(R)$ turning it into a field. The zero and unit element are

$$\frac{0}{1}, \qquad \frac{1}{1}$$

and we will simply denote them by 0 and 1 from now on.

An element $(a,b) \in P(R)$ is equivalent to $(0,1)$ iff $a = 0$. Hence we have

$$\frac{a}{b} \neq 0$$

in $\mathrm{Quot}(R)$ iff $a \neq 0$, so that

$$\frac{b}{a}$$

is well-defined, and we clearly have

$$\frac{a}{b}\frac{b}{a} = 1.$$

Thus $\mathrm{Quot}(R)$ is a field.

Next, we define the map $\iota : R \to \mathrm{Quot}(R)$ by setting

$$\iota(r) = \frac{r}{1}.$$

From the definition of addition and multiplication in $\mathrm{Quot}(R)$ it is clear that ι is a ring homomorphism. Assume that $\iota(r) = 0$. Then $(r, 1) \sim (0, 1)$ which means $r = 0$. We conclude that ι is injective.

Finally, assume that $f : R \to K$ is an injective homomorphism with values in a field K. Then $f(b)$ is nonzero and hence invertible for every nonzero element $b \in R$. We define $F : \mathrm{Quot}(R) \to K$ by

$$F(\frac{a}{b}) = f(a)f(b)^{-1}.$$

This definition does not depend on the choice of representatives for $\frac{a}{b}$. Indeed, if $(a, b) \sim (c, d)$ then $ad = bc$ so that $f(a)f(d) = f(ab) = f(bc) = f(b)f(c)$ and

$$f(a)f(b)^{-1} = f(c)f(d)^{-1}.$$

It is clear from the construction that $F\iota = f$. If $G : \mathrm{Quot}(R) \to K$ is another homomorphism such that $G\iota = f$ then $G(\frac{a}{1}) = f(a)$ for all $a \in R$. In particular, for any nonzero $b \in R$ the element $G(\frac{b}{1}) \in K$ is invertible because K is a field. Moreover

$$G(\frac{1}{b})G(\frac{b}{1}) = G(\frac{1}{1}) = 1$$

shows that $G(\frac{1}{b})$ is the inverse of $G(\frac{b}{1})$. Since inverses are unique we have $G(\frac{1}{b}) = G(\frac{b}{1})^{-1} = f(b)^{-1}$. We conclude

$$G(\frac{a}{b}) = G(\frac{a}{1})G(\frac{1}{b}) = f(a)f(b)^{-1} = F(\frac{a}{b})$$

and hence $G = F$. This shows uniqueness of F and finishes the proof. □

Definition 2.27. We call $\mathrm{Quot}(R)$ the *field of fractions* or the *quotient field* of R.

As already indicated above, the construction of the rational numbers $\mathbb{Q}$ is a special instance of Theorem 2.26, namely the case $R = \mathbb{Z}$.

Further examples of fields of quotients arise by starting from polynomial rings over fields. We will come back to this when studying polynomial rings in the next section.

2.9. **Exercises.** The solutions to questions marked with a * are to be handed in on **Tuesday 2nd February** by groups of 1-3 students (your choice). One question per student on the submission will be marked and the homework will be returned and discussed in the tutorial in that week.

*Exercise** 2.1 (C). State the definition of a principal ideal and of a principal ideal domain.

*Exercise** 2.2 (B). Let R be an arbitrary commutative ring. Show that every maximal ideal in R is a prime ideal.

*Exercise** 2.3 (B). Mark each of the following true or false.

(1) $\mathbb{Q}$ is isomorphic to the field of fractions of $\mathbb{Z}$.
(2) $\mathbb{R}$ is isomorphic to the field of fractions of $\mathbb{Z}$.
(3) $\mathbb{C}$ is isomorphic to the field of fractions of $\mathbb{R}$.
(4) Every element of an integral domain R is a unit in the field of quotients $\mathrm{Quot}(R)$.
(5) If R and S are isomorphic integral domains then their field of quotients $\mathrm{Quot}(R)$ and $\mathrm{Quot}(S)$ are isomorphic.

Exercise 2.4 (B). Let R be a ring and let $u \in R$ be a unit. Show that the element $v \in R$ satisfying $vu = 1 = uv$ is uniquely determined.

Exercise 2.5 (B). Determine all units in the rings $R = \mathbb{Z}, R = \mathbb{Z}\times\mathbb{Z}, R = \mathbb{Z}/(7), R = \mathbb{Z}/(4)$.

Exercise 2.6 (B). Determine all maximal ideals in $R = K \times K$ where K is a field. What are the corresponding quotient rings isomorphic to?

*Exercise** 2.7 (A). Let R and S be rings and let $f : R \to S$ be a ring homomorphism.

(1) Show that if $u \in R$ is a unit then $f(u) \in S$ is a unit.
(2) Assume in addition that R is a field and that S is a nonzero ring. Show that in this case $f(u)$ is nonzero for every nonzero element $u \in R$. Conclude that f is injective.
(3) Is the previous statement still true if S is the zero ring?

Exercise 2.8 (A). This is an excercise on ring homomorphisms.

(1) Determine all ring homomorphisms $\mathbb{Q} \to \mathbb{Z}/(4)$.
(2) Determine all ring homomorphisms $\mathbb{Z}/(4) \to \mathbb{Q}$.
(3) Determine all ring homomorphisms $\mathbb{Q} \to \mathbb{C}$.

Exercise 2.9 (A). Let R and S be rings and let $f : R \to S$ be a ring homomorphism.

(1) Show that if $J \subset S$ is an ideal, then
$$f^{-1}(J) \subset R = \{r \in R | f(r) \in J\}$$
is an ideal in R.
(2) Show that if $I \subset R$ is an ideal then $f(I) = \{f(r) \mid r \in R\}$ is an ideal in the subring $f(R) = \mathrm{im}(f)$ of S.
(3) Is $f(I)$ necessarily an ideal in S? Give a proof or a counterexample.
(4) Show that if f is surjective then $f(f^{-1}(J)) = J$ for all ideals $J \subset S$.

Exercise 2.10 (A). This question deals with quotients of principal ideal rings.

(1) Show that if R is a principal ideal ring and $f : R \to S$ is a surjective ring homomorphism onto a nonzero ring S, then S is a principal ideal ring as well. (*Hint: If $J \subset S$ is an ideal you may consider $f^{-1}(J) \subset R$ and use the results of the previous excercise*).

(2) Show that $\mathbb{Z}/(n)$ is a principal ring for all $n > 1$.
(3) Determine all ideals in the ring $\mathbb{Z}/(12)$.

Exercise 2.11 (A). Let R be an integral domain. Verify in detail that addition and multiplication in the field of quotients $\mathrm{Quot}(R)$ are well-defined.

Exercise 2.12 (A). Let K be a field. Show that the canonical homomorphism $\iota : K \to \mathrm{Quot}(K)$ is an isomorphism.

3. Rings of polynomials

In this section we take a closer look at rings of polynomials. We shall study irreducible polynomials and methods for determining irreducibility.

3.1. **Basic properties of $R[x]$.**

We begin by an important property of $R[x]$: it is the *free commutative R-algebra in one generator* - compare this to the notion of a free group or a free abelian group:

Proposition 3.1 (Homomorphism extension property)**.** *If $f : R \to S$ is a homomorphism of rings and $y \in S$ is any element, then there exists a unique ring homomorphism $F : R[x] \to S$ such that $F(a) = f(a)$ for all $a \in R$ and $F(x) = y$.*

Proof. We define F be setting

$$F\Big(\sum_{j=0}^{\infty} a_j x^j\Big) = \sum_{j=0}^{\infty} f(a_j) y^j$$

It is straightforward to check from the definition of addition and multiplication in $R[x]$ that this yields indeed a ring homomorphism $F : R[x] \to S$ with the desired properties. The uniqueness of F follows from the fact that the above requirements determine F on all constant polynomials and on $x \in R[x]$, and these elements generate $R[x]$ as a ring. □

For example let R be a ring and let $r \in R$. Then we obtain an evaluation homomorphism $\mathrm{ev}_r : R[x] \to R$ by sending $f(x)$ to its value $\mathrm{ev}_r(f(x)) = f(r) \in R$. In the above notation, this corresponds to the case $S = R$ and $y = r$. Later we will be interested in particular in determining those r for which $f(r) = 0$.

We mentioned in the beginning that we will usually work with integral domains, so the following seems also worth mentioning:

Proposition 3.2. *Let R be a commutative ring. Then R is an integral domain iff the polynomial ring $R[x]$ is an integral domain.*

Proof. It is clear that any subring of an integral domain is again an integral domain. Therefore, if $R[x]$ is an integral domain then the same holds for R.

Conversely, assume that R is an integral domain. Let $f(x), g(x) \in R[x]$ be nonzero polynomials and assume that $f(x)g(x) = 0$. If $\deg(f) = m$ and $\deg(g) = n$ we may write

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$$
$$g(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$$

and obtain

$$f(x)g(x) = a_m b_n x^{m+n} + \cdots + a_0 b_0.$$

In particular, $f(x)g(x) = 0$ implies $a_m b_n = 0$. Since R is an integral domain this can only happen if $a_m = 0$ or $b_n = 0$, thus contradicting our assumption that

$\deg(f) = m$ and $\deg(g) = n$. That is, $f(x)g(x) = 0$ implies that one of $f(x), g(x)$ must be the zero polynomial. Therefore $R[x]$ is an integral domain. □

In particular, we conclude that the polynomial ring $K[x]$ over any field K is an integral domain. Hence we can apply the construction of the field of quotients to such rings.

Definition 3.3. The field of quotients $\operatorname{Quot}(K[x])$ is called the *field of rational functions* over K.

Recall that its elements can be written as fractions

$$\frac{f(x)}{g(x)}$$

for $f(x), g(x) \in K[x]$ with $g(x) \neq 0$. Such a quotient defines a function

$$K \setminus P \to K, \qquad c \mapsto \frac{f(c)}{g(c)}$$

where $P := \{c \in K \mid g(c) = 0\}$ is the set of *poles* of the rational function. Note also that for $K = \mathbb{C}$ all rational functions are meromorphic.

As we mentioned already in Section 1, the usual procedure of long division can be applied to polynomials with coefficients in arbitrary fields:

Theorem 3.4 (Long division). *Let K be a field and let $f(x), g(x) \in K[x]$ with $g(x) \neq 0$. Then there exist unique $q(x), r(x) \in K[x]$ with $\deg(r(x)) < \deg(g(x))$ such that*

$$f(x) = g(x)q(x) + r(x).$$

Proof. We follow the usual Euclidean algorithm for real polynomials familiar from first year algebra. In fact, the calculations involved in this algorithm only use the fact that $\mathbb{R}$ is a field, and therefore work over any field K. □

Using this, we now show:

Theorem 3.5. *If K is a field, then $K[x]$ is a principal ideal domain.*

Proof. According to Proposition 3.2 $K[x]$ is an integral domain. Now assume that $I \subset K[x]$ is a nonzero ideal, and let $g(x) \in I$ be a nonzero polynomial of smallest degree. We shall prove that $I = (g(x))$. To this end assume that $f(x) \in I$ is arbitrary. Since $g(x) \neq 0$, Theorem 3.4 implies that we find elements $q(x), r(x) \in K[x]$ such that $f(x) = g(x)q(x) + r(x)$ with $\deg(r(x)) < \deg(g(x))$. Since $g(x) \in I$ we have $g(x)q(x) \in I$, and combining this with $f(x) \in I$ shows that $r(x) = f(x) - g(x)q(x) \in I$. However, $g(x)$ was chosen to be a nonzero element of I of minimal degree, so we must have $r(x) = 0$. This means $f(x) = g(x)q(x)$ as had to be shown. □

3.2. **Irreducible polynomials.**

Irreducible polynomials in polynomial rings are analogues of the prime numbers in $\mathbb{Z}$. They play a crucial role in Galois theory. We start with the following definition.

Definition 3.6. Let K be a field and let $f(x), g(x) \in K[x]$. We say that $g(x)$ divides $f(x)$ if there exists $q(x) \in K[x]$ such that

$$f(x) = g(x)q(x)$$

We write $g(x)|f(x)$ in this case.

We observe that if $g(x)$ is nonzero then $g(x)|f(x)$ means precisely that the remainder polynomial in the long division of $f(x)$ by $g(x)$ is zero.

Example 3.7. Consider the polynomials $f(x) = x^2 - 4$ and $g(x) = x - 2$ in $\mathbb{Q}[x]$. Then $g(x)|f(x)$ because $f(x) = x^2 - 4 = (x - 2)(x + 2)$.

Definition 3.8. Let K be a field. A *greatest common divisor* or *highest common factor* of $f(x), g(x) \in K[x]$ is a polynomial $d(x) \in K[x]$ such that

a) $d(x)|f(x)$ and $d(x)|g(x)$
b) If $e(x)|f(x)$ and $e(x)|g(x)$ then $e(x)|d(x)$.

Example 3.9. Retaining our example from above, a highest common factor of $f(x) = x^2 - 4$ and $g(x) = x - 2$ is $g(x)$ itself.

Let us prove a classic that was already covered in the first year algebra course:

Proposition 3.10. *Let K be a field and let $f(x), g(x) \in K[x]$ be nonzero. Then there exist $a(x), b(x) \in K[x]$ such that $a(x)f(x) + b(x)g(x)$ is a highest common factor of $f(x)$ and $g(x)$.*

Proof. Assume without loss of generality that $\deg(g(x)) \leq \deg(f(x))$, set $r_0(x) := f(x), r_1(x) := g(x)$, and let $I := \{a(x)f(x) + b(x)g(x) \mid a(x), b(x) \in K[x]\}$ be the ideal generated by $f(x)$ and $g(x)$.

Now use the division algorithm to construct inductively a sequience of polynomials $r_i(x), q_i(x)$ such that

(3.1)
$$r_i(x) = r_{i+1}(x)q_i(x) + r_{i+2}(x), \qquad \deg(r_{i+2}(x)) < \deg(r_{i+1}(x)) \text{ or } r_{i+2}(x) = 0$$

with must terminate for degree reasons after fnitely many steps in 0. Note that if $r_i(x)$ and $r_{i+1}(x)$ are both in I then so is $r_{i+2}(x)$ as a direct consequence of (3.1). So since this is true for $i = 0$, we see by induction that $r_i(x) \in I$ for all i.

For the last nonzero $r_d(x)$ we then have

$$r_{d-1}(x) = r_d q_{d-1}(x),$$

so $r_d(x)$ divides $r_{d-1}(x)$. However, it follows again directly from (3.1) that if a polynomial divides both $r_{i+2}(x)$ and $r_{i+1}(x)$, it also divides $r_i(x)$. Thus we obtain by an induction starting at $i = d$ and going down that $r_d(x)$ divides in fact all $r_i(x)$ and in particular $r_0(x) = f(x)$ and $r_1(x) = g(x)$. Thus we have found a polynomial $r_d(x)$ of the form $a(x)f(x) + b(x)g(x)$ which is a divisor of both $f(x)$ and $g(x)$.

To finish the proof note that if a polynomial divides $f(x)$ and $g(x)$, then it divides all polynomials of the form $a(x)f(x) + b(x)g(x)$, so our polynomial $r_d(x)$ not just a divisor but a greatest common divisor. □

Definition 3.11. Let R be an integral domain and let $f(x) \in R[x]$ be a nonconstant polynomial. Then $f(x)$ is called *reducible* over R if $\deg(f(x) < 1$ of if it can be expressed as a product

$$f(x) = g(x)h(x)$$

of two polynomials $g(x), h(x) \in R[x]$ which are both of strictly lower degree then $f(x)$. Otherwise we say that $f(x)$ is *irreducible* over R.

We note that it is important to keep reference to the base ring here, as we shall see from our examples below. By iterative decomposition we can write every polynomial as a product of irreducible factors. We may view irreducible polynomials in $R[x]$ as analogous of the prime numbers in $\mathbb{Z}$.

We remark that in the same way as the decomposition of a natural number as a product of primes is unique, the decomposition of a polynomial into irreducible factors is unique. A proof can be found in Theorem 45.29 of [2].

Theorem 3.12 (Uniqueness of factorisation)**.** *Let K be a field. The factorisation of $f(x) \in K[x]$ into irreducibles is unique apart from multiplication with constant factors and reordering.*

More generally, this works in every principal ideal domain, a PID is a UFD (unique factorisation domain).

Taking the analogy between prime numbers and irreducible polynomials further we shall now prove the following theorem.

Theorem 3.13. *Let K be a field and let $f(x) \in K[x]$ be a nonconstant polynomial. Then the following conditions are equivalent.*

a) $f(x)$ is irreducible.
b) $K[x]/(f(x))$ is an integral domain.
c) $K[x]/(f(x))$ is a field.

Proof. *a)* ⇒ *c)* Let $g(x) \in K[x]$ be a polynomial such that the coset $g(x)+(f(x))$ in $K[x]/(f(x))$ is nonzero. This means $g(x) \notin (f(x))$, and since $f(x)$ is irreducible the highest common factor of $f(x)$ and $g(x)$ must be a constant polynomial. According to Proposition 3.10 this means we find $a(x), b(x) \in K[x]$ such that

$$a(x)f(x) + b(x)g(x) = 1.$$

This implies $(b(x) + (f(x)))(g(x) + (f(x))) = (1 + (f(x)))$ in the quotient ring $K[x]/(f(x))$. That is, $g(x) + (f(x))$ is invertible, which means that $K[x]/(f(x))$ is a field.

c) ⇒ *b)* is clear.

b) ⇒ *a)* Assume that $f(x) = g(x)h(x)$ is reducible, where $g(x), h(x)$ are polynomials of degree strictly less than $\deg(f(x))$. Then $(g(x)+(f(x)))(h(x)+(f(x))) = 0$ in $K[x]/(f(x))$. Since $K[x]/(f(x))$ is assumed to be an integral domain this implies $(g(x) + (f(x))) = 0$ or $(h(x) + (f(x))) = 0$, which in turn means $g(x) \in (f(x))$ or $h(x) \in (f(x))$. This contradicts the fact that all nonzero polynomials in $(f(x))$ have degree at least $\deg(f(x))$. □

The theorem is in fact true for arbitrary principal ideal domains and will usually be stated as saying that in a principal ideal domain all irreducible elements are prime and that a principal ideal domain whic is not a field has Krull dimension 1, which means that every nonzero prime ideal is maximal.

We shall be interested in particular in the case of roots. If α is a zero of order m of $f(x)$ then we can write

$$f(x) = (x - \alpha)^m q(x)$$

for some $q(x) \in K[x]$. The polynomial $q(x)$ in turn is not divisible by $(x - \alpha)$.

Proposition 3.14. *Let K be a field and let $f(x) \in K[x]$ be a a nonconstant polynomial. If $\alpha_1, \ldots, \alpha_k$ are the roots of $f(x)$ in K, with multiplicities $m_1, \ldots, m_k$, then*

$$f(x) = (x - \alpha_1)^{m_1} \cdots (x - \alpha_k)^{m_k} q(x)$$

for a polynomial $q(x) \in K[x]$ which has no roots in K. In particular, every polynomial $f(x)$ of degree $n > 0$ has at most n zeros in K, counted with multiplicities.

Proof. We apply iteratively long division of polynomials to $f(x)$ and the linear factors $(x-\alpha_j)$. More precisely, if $f(\alpha_j)=0$ we can write $f(x)=(x-\alpha_j)q(x)+r(x)$ and conclude $r(\alpha_j)=0$. Since $\deg(r(x))<1$ we know that $r(x)$ is constant, and we conclude that $r(x)=0$. Repeating this process with $q(x)$ yields the desired expression after finitely many steps. □

Example 3.15. Note that "irreducible" does *not* mean "has no root"! For example, $x-2\in\mathbb{Q}[x]$ is irreducible but has a root, and $(x^2+1)(x^2+2)=x^4+3x^2+2$ is reducible but has no root. all we can say is that a polynomial of degree greater than 1 that has a root is reducible.

3.3. **Criteria for testing irreducibility.** In this subsection we discuss some criteria which help us to do decide whether a given polynomial is irreducible.

Firstly, let us briefly consider the situation for polynomials of low degrees. More precisely, let K be a field and let $f(x)\in K[x]$ be a nonzero polynomial. If $\deg(f(x))=1$ then $f(x)=\alpha_1x+\alpha_0$ for some $\alpha_1,\alpha_0\in K$ such that $\alpha_1\neq 0$. It follows immediately from the definition of irreducibility that all such polynomials are irreducible. If $\deg(f(x))=2$ then $f(x)$ is reducible iff we find polynomials $g(x),h(x)$ of degree 1 such that $f(x)=g(x)h(x)$. Notice that factorising $f(x)$ in this way is equivalent to saying that $f(x)$ has two roots in K. In other words, a polynomial $f(x)$ of degree 2 is irreducible over K iff it has no root in K.

Example 3.16. Consider the polynomial $f(x)=x^2+1\in\mathbb{Q}[x]$. This polynomial is irreducible over $\mathbb{Q}$, and even over $\mathbb{R}$, because otherwise we could write $f(x)$ as the product of two linear polynomials with real coefficents. This is impossible since $f(x)$ has no real zeros. Of course, $f(x)=x^2+1$ is reducible over $\mathbb{C}$, because $x^2+1=(x+i)(x-i)$ in $\mathbb{C}[x]$.

A similar reasoning works if $\deg(f(x))=3$. Indeed, in this case a factorisation $f(x)=g(x)h(x)$ into polynomials of smaller degree implies that one of $g(x)$ or $h(x)$ must be of degree one, and the other of degree 2. Having a factorisation with a linear polynomial as factor is equivalent to saying that $f(x)$ has a zero. We can rephrase this by saying that a polynomial $f(x)\in K[x]$ of degree 3 is irreducible over K iff it has no zeros in K.

Despite these facts, we should point out that it is typically difficult to decide whether a given polynomial of degree greater than 3 is irreducible. We shall now discuss some results which can help to prove irreducibility.

Lemma 3.17 (Gauss' Lemma). *Let $f(x)\in\mathbb{Z}[x]$ be irreducible over $\mathbb{Z}$. Then $f(x)$ is also irreducible over $\mathbb{Q}$.*

Proof. Assume that $f(x)=g(x)h(x)$ for some polynomials $g(x),h(x)\in\mathbb{Q}[x]$ of smaller degree. Multiplying both sides by the product of all denominators of the coefficients of $g(x)$ and $h(x)$ we can write $nf(x)=g'(x)h'(x)$ where now $g'(x),h'(x)\in\mathbb{Z}[x]$.

We now inductively cancel out prime factors of n: let p be a prime factor of n. We claim that if we write

$$g'(x)=g_0+g_1x+\cdots g_rx^r,\qquad h'(x)=h_0+h_1x+\cdots h_sx^s$$

then p divides all coefficients g_i or all coefficients h_j. To prove this assume that the assertion is wrong. Then there exist smallest values i and j such that p does neither divide g_i nor h_j. However, since p divides all coefficients of $nf(x)=g'(x)h'(x)$ we know that p divides the coefficient of x^{i+j} in $g'(x)h'(x)$, which is given by

$$g_0h_{i+j} + g_1h_{i+j-1} + \cdots + g_ih_j + \cdots g_{i+j}h_0.$$

By our choice of i and j, the prime p divides every term in this expression except g_ih_j. This is a contradiction to the fact that the entire sum is divisible by p.

We may therefore assume without loss of generality that p divides all coefficients of $g'(x)$. Hence we can write $g'(x) = pg''(x)$ where $g''(x)$ is again contained in $\mathbb{Z}[x]$. We may now divide the equation $nf(x) = pg''(x)h'(x)$ by p, and still remain within $\mathbb{Z}[x]$. Proceeding in this way we see that we can factorise $f(x)$ over $\mathbb{Z}[x]$. □

The point of Lemma 3.17 is that there are a priori much more possibilities to factorise a polynomial over $\mathbb{Q}$ than over $\mathbb{Z}$. We shall it in the proof of the following irreducibility criterion.

Proposition 3.18 (Eisenstein's criterion)**.** *Let*

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{Z}[x]$$

be a polynomial with integer coefficients. Assume that there exists a prime $p \in \mathbb{N}$ such that

a) $a_0, a_1, \dots a_{n-1}$ are divisible by p.
b) a_n is not divisible by p.
c) a_0 is not divisible by p^2.

Then $f(x)$ is irreducible over $\mathbb{Q}$.

Proof. Due to Lemma 3.17 it suffices to show that $f(x)$ is irreducible over $\mathbb{Z}$. To prove this, assume that $f(x) = g(x)h(x)$ where

$$g(x) = g_0 + g_1x + \cdots g_rx^r, \qquad h(x) = h_0 + h_1x + \cdots h_sx^s$$

are polynomials in $\mathbb{Z}[x]$ of degree smaller than $\deg(f(x))$. Then clearly $r, s \geq 1$ and $r + s = n$. Now $g_0h_0 = a_0$ and iusing assumptions *a*) and *c*) we see that p divides precisely one of g_0 and h_0. Without loss of generality let us assume that p divides g_0 but not h_0. If all coeffients g_i were divisible by p then a_n would be divisible by p, which contradicts assumption *b*). Hence there exists a smallest index $j < n$ such that g_j is not divisible by p. Observe that

$$a_j = g_0h_j + g_1h_{j-1} + \cdots + g_jh_0 \Rightarrow g_jh_0 = g_0h_j + g_1h_{j-1} + \cdots + g_{j-1}h_1 - a_j$$

is divisible by p due to *a*), so since g_j is not divisble by p, h_0 is divisible by p, which contradicts our previous observation that only one of g_0, h_0 is divisible by p. □

Example 3.19. Consider for instance the polynomial

$$f(x) = 3x^5 + 15x^2 + 5x + 10$$

in $\mathbb{Z}[x]$. We may apply Eisenstein's criterion for $p = 5$ to conclude that $f(x)$ is irreducible.

If $f(x) \in K[x]$ is a polynomial and $a \in K$ is a constant, then $f(x + a)$ is the polynomial in $K[x]$ obtained by expanding powers $(x + a)^k$ using the binomial formula. Equivalently, using Theorem 3.1 we may say that $f(x + a)$ is the image of $f(x)$ under the homomorphism $\phi : K[x] \to K[x]$ determined by sending x to $x + a$ and the identity map on K.

Sometimes it is useful to apply the following observation in order to prove irreduciblity.

Lemma 3.20. *Let R be an integral domain, $f(x) \in R[x]$ be a polynomial, and $a \in R$. Then $f(x)$ is reducible over R iff $f(x+a)$ is reducible over R.*

Proof. Assume we can write $f(x) = g(x)h(x)$ in $R[x]$ such that both $g(x)$ and $h(x)$ have strictly lower degree than $f(x)$. Then we also get $f(x+a) = g(x+a)h(x+a)$, and we observe that $\deg(g(x+a)) = \deg(g(x))$ and $\deg(h(x+a)) = \deg(h(x))$. It follows that $f(x+a)$ is reducible. Hence reducibility of $f(x)$ implies reduciblity of $f(x+a)$, and applying the same argument to $f(x+a)$ and $-a$ shows that the converse holds as well. □

Example 3.21. Consider $f(x) = x^5 + 10x^4 + 40x^3 + 80x^2 + 83x + 41 \in \mathbb{Q}[x]$. Notice that we cannot apply Eisenstein's criterion to $f(x)$ directly. However, observing that

$$f(x) = (x+2)^5 + 3x + 9$$

we obtain

$$f(x-2) = x^5 + 3x + 3.$$

The latter is irreducible by Eisenstein's criterion for $p = 3$, so Lemma 3.20 shows that $f(x)$ is irreducible over $\mathbb{Q}$ as well.

Finally, we discuss reduction modulo primes to test irreducibility.

Proposition 3.22 (Mod p test)**.** *Let $f(x) = a_n x^n + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$. Assume that p is a prime not dividing a_n. Let us write $\overline{a}$ for the coset of a in $\mathbb{Z}/(p) = \mathbb{F}_p$, and let*

$$\overline{f}(x) = \overline{a_n} x^n + \cdots + \overline{a_1} x + \overline{a_0}$$

be the corresponding polynomial in $\mathbb{F}_p[x]$. If $\overline{f}(x)$ is irreducible over $\mathbb{F}_p$, then $f(x)$ is irreducible over $\mathbb{Q}$.

Proof. Assume that $\overline{f}(x)$ is irreducible over $\mathbb{F}_p$ but that $f(x)$ is reducible over $\mathbb{Q}$. Then $f(x)$ is reducible over $\mathbb{Z}$ by Lemma 3.17, hence we may write $f(x) = g(x)h(x)$ for polynomials $g(x), h(x) \in \mathbb{Z}[x]$ of strictly positive degree such that $\deg(g(x)) + \deg(h(x)) = \deg(f(x))$. Therefore we have $\overline{f}(x) = \overline{g}(x)\overline{h}(x)$ for the polynomials obtained using the canonical quotient homomorphism $\mathbb{Z}[x] \to \mathbb{F}_p[x]$. Since a_n is not divisible by p, $\overline{a_n}$ is nonzero, and we conclude that the highest coefficients of $\overline{g}(x)$ and $\overline{h}(x)$ must be nonzero as well. This would imply that $\overline{f}(x)$ is reducible, which contradicts our assumption. Hence $f(x)$ is irreducible over $\mathbb{Q}$. □

Example 3.23. Consider the polynomial

$$f(x) = x^4 + 15x^3 + 7$$

in $\mathbb{Z}[x]$. Over $\mathbb{F}_5$ this becomes

$$\overline{f}(x) = x^4 + 2,$$

and we argue that $\overline{f}(x)$ is irreducible over $\mathbb{F}_5$. If we could write $\overline{f}(x)$ as a product of two factors $\overline{f}(x) = g(x)h(x)$ for some $g(x), h(x) \in \mathbb{F}_5[x]$ then either both $g(x), h(x)$ are of degree 2, or one of them has degree 1. In the latter case $\overline{f}(x)$ would have a zero in $\mathbb{F}_5$, which is easily checked not to be the case. Indeed, one simply has to insert the five different elements of $\mathbb{F}_5$ into $\overline{f}(x)$ to see this. If both $g(x)$ and $h(x)$ are of degree 2 we can write

$$x^4 + 2 = (x^2 + ax + b)(x^2 + cx + d)$$

and obtain

$$a + c = 0, \qquad ac + b + d = 0, \qquad bd = 2$$

or equivalently

$$c = -a, \qquad b + d = a^2, \qquad bd = 2$$

However, in $\mathbb{Z}_5$ we have

$$0^2 = 0, \quad 1^2 = 1, \quad 2^2 = 4, \quad 3^2 = 4, \quad 4^2 = 1$$

so $b + d = a^2$ implies

$$d = -b, \quad d = 1 - b, \text{ or } d = 4 - b.$$

Inserting this into the last question yields

$$-b^2 = 2, \quad b(1 - b) = 2, \text{ or} b(4 - b) = 2$$

and inserting all possible five values of b shows that none of these equations has a solution. We conclude that $\overline{f}(x)$ is irreducible over $\mathbb{F}_5$, and according to Proposition 3.22 we deduce that $f(x)$ is irreducible over $\mathbb{Z}$.

3.4. **Exercises.**

*Exercise** 3.1 (C). State the definition of an irreducible polynomial.

Exercise 3.2 (C). Let $K = \mathbb{F}_3$. Find quotient and remainder when performing long division of $f(x) = 2x^3 + 2x^2 + x + 1$ by $g(x) = 2x^2 + 2$ in $K[x]$.

*Exercise** 3.3 (B). Show that the following polynomials are irreducible.

(1) $x^7 + 48x + 24 \in \mathbb{Q}[x]$.
(2) $x^n - p \in \mathbb{Q}[x]$ for a prime p and any $n > 0$.
(3) $x^3 + 7x + 38 \in \mathbb{Q}[x]$.

Exercise 3.4 (B). Mark each of the following true or false.

(1) Every polynomial which is irreducible over $\mathbb{Q}$ is irreducible over $\mathbb{R}$.
(2) Polynomials of prime degree are irreducible.
(3) A polynomial which is irreducible over $\mathbb{Z}$ is irreducible over $\mathbb{Q}$.
(4) Every polynomial is a constant multiple of an irreducible polynomial.
(5) A polynomial of degree n over a field K has at most n zeros in K.

Exercise 3.5 (B). Let $f(x) = x^3 + 2x^2 + 2x + 1 \in \mathbb{F}_7[x]$. Find a factorisation of this polynomial into linear factors in $\mathbb{F}_7[x]$.

Exercise 3.6 (B). Consider the polynomials $f(x), g(x) \in \mathbb{F}_5[x]$ given by

$$f(x) = x^4 - 3x^3 + 2x^2 + 4x - 1, \qquad g(x) = x^2 - 2x + 3.$$

(1) Use long division to determine $q(x), r(x) \in \mathbb{F}_5[x]$ such that $f(x) = g(x)q(x) + r(x)$.
(2) Determine a highest common factor $d(x) \in \mathbb{F}_5[x]$ of $f(x)$ and $g(x)$.
(3) Determine $a(x), b(x) \in \mathbb{F}_5[x]$ such that $a(x)f(x) + b(x)g(x) = d(x)$.

Exercise 3.7 (B). Determine wether the following polynomials in $\mathbb{Z}[x]$ satisfy the Eisenstein criterion for irreducibility for some prime p.

(1) $f(x) = x^2 - 12$
(2) $f(x) = 4x^{10} - 9x^3 + 24x - 18$
(3) $f(x) = 8x^3 + 6x^2 - 9x + 24$

Exercise 3.8 (B). Mark each of the following true or false.

(1) $f(x) = x - 2$ is is irreducible over $\mathbb{Q}$.
(2) $f(x) = 3x - 6$ is irreducible over $\mathbb{Q}$.
(3) $f(x) = x^2 - 3$ is irreducible over $\mathbb{Q}$.
(4) If K is a field then the units of $K[x]$ are precisely the nonzero elements of K, viewed as constant polynomials.
(5) If K is a field then every polynomial $f(x) \in K[x]$ of degree 1 has at least one zero in K.

Exercise 3.9 (A). In this excercise you shall show that the converse of the Mod p-test does not hold. More precisely, give an example of a polynomial $f(x) \in \mathbb{Z}[x]$ and a prime p such the following three conditions hold.

(1) p does not divide the highest coefficient of $f(x)$.
(2) $f(x)$ is irreducible over $\mathbb{Q}$.
(3) $\overline{f}(x) \in \mathbb{F}_p[x]$ is reducible over $\mathbb{F}_p$.

Exercise 3.10 (A). Let p be a prime. In this excercise we show that the so-called p-th *cyclotomic polynomial*

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1 \in \mathbb{Q}[x]$$

is irreducible.

a) Verify that $\Phi_p(x)(x-1) = x^p - 1$.
b) Show that a polynomial $f(x) \in \mathbb{Q}[x]$ is irreducible iff $f(x+1) \in \mathbb{Q}[x]$ is irreducible.
c) Use the Eisenstein criterion to prove that $\Phi_p(x+1)$ is irreducible.

Exercise 3.11 (A). Consider the polynomial ring $\mathbb{Z}[x]$.

(1) Show that
$$I = \left\{ \sum_{j=0}^{\infty} a_j x^j \in R \mid a_0 \text{ even} \right\}$$
is an ideal in $\mathbb{Z}[x]$.
(2) Show that I is not of the form $(f(x))$ for some $f(x) \in \mathbb{Z}[x]$.
(3) Conclude that $\mathbb{Z}[x]$ is not a principal ideal domain.

4. Field extensions

A central concept in Galois theory is the notion of a field extension. In this section we collect important definitions and results related to field extensions, and discuss some examples. We will also study geometric applications concerning ruler and compass constructions.

4.1. Field extensions and subfields.

Let us begin with the following definition.

Definition 4.1. Let K and F be fields and suppose that $K \subset F$ is a subring. Then we say that K is a *subfield* of F, and that F is a *field extension* of K. We write $F|K$ in this situation. If $K \subset E \subset F$ are fields, we call E an *intermediate field* of the field extension $F|K$.

Basic examples of field extensions are $\mathbb{C}|\mathbb{R}$ or $\mathbb{R}|\mathbb{Q}$.

Definition 4.2. Let $F|K$ be a field extension and let $\Lambda \subset F$ be a subset. We write $K(\Lambda) \subset F$ for the intermediate field of $F|K$ generated by Λ. Formally,

$$K(\Lambda) = \bigcap_{\substack{K \subset E \subset F \text{ field} \\ \Lambda \subset E}} E.$$

If $\Lambda = \{\lambda\}$ consists of a single element we simply write $K(\lambda)$ instead of $K(\Lambda)$.

The field $K(\Lambda)$ is the collection of everything that can be obtained from K and Λ using finitely many additions, subtractions, multiplications, and divisions. We will also call it the field obtained by adjoining all elements from Λ to K.

Example 4.3. Consider for instance the field extension $\mathbb{R}|\mathbb{Q}$ and the field $\mathbb{Q}(\sqrt{2})$ obtained by adjoining the single element $\sqrt{2}$ to $\mathbb{Q}$. Then we have

$$\mathbb{Q}(\sqrt{2}) = \{a + \sqrt{2}b \mid a, b \in \mathbb{Q}\}.$$

Indeed, the set on the right hand side is clearly contained in the subfield $\mathbb{Q}(\sqrt{2})$ generated by $\mathbb{Q}$ and $\sqrt{2}$. Moreover one checks that this set is itself already a subfield of $\mathbb{R}$, which shows that it equals $\mathbb{Q}(\sqrt{2})$.

4.2. The degree of a field extension.

Given a field extension $F|K$ we view F as a vector space over the smaller field K - the addition is the addition in F and the multiplication of a vector $v \in F$ by a scalar $\alpha \in K$ is the product of the two in the field F.

Recall that the *dimension* of a K-vector space V is $\dim_K(V) = n$ if there exist n linearly independent vectors $v_1, \dots, v_n \in V$ spanning V. We set $\dim_K(V) = \infty$ otherwise.

Definition 4.4. The *degree* $[F : K]$ of a field extension $F|K$ is the dimension of the K-vector space F,

$$[F : K] = \dim_K(F).$$

We say that $F|K$ is a finite extension if $[F : K]$ is finite, and an infinite extension otherwise.

Example 4.5. The extension $\mathbb{C}|\mathbb{R}$ is finite with $[\mathbb{C} : \mathbb{R}] = 2$. Indeed, the elements $1, i$ form a basis of $\mathbb{C}$ as a real vector space.

Example 4.6. $\mathbb{R}|\mathbb{Q}$ is an infinite extension: $\mathbb{R}$ is an uncountable set, and every finite-dimensional $\mathbb{Q}$-vector space is countable (or 0).

Proposition 4.7. *[Tower law] Let $K \subset E \subset F$ be field extensions. If $F|E$ and $E|K$ are finite then $F|K$ is finite and*

$$[F : K] = [F : E][E : K].$$

Moreover $[F : K]$ is infinite iff $[F : E]$ or $[E : K]$ is infinite.

Proof. Assume first that $F|E$ and $E|K$ are finite. Let $\alpha_1, \dots, \alpha_m$ be a K-basis of E and let $\beta_1, \dots, \beta_n$ be an E-basis of F. Then the elements $\alpha_i\beta_j$ for $1 \leq i \leq m$ and $1 \leq j \leq n$ form a K-basis of F. Indeed, every element γ of F can be written as a linear combination

$$\gamma = \sum_{j=1}^{n} \lambda_j \beta_j$$

where $\lambda_j \in E$ for every j. Therefore we can write

$$\lambda_j = \sum_{i=1}^{m} \mu_{ij}\alpha_i$$

for uniquely determined elements $\mu_{ij} \in K$, and we obtain

$$\gamma = \sum_{j=1}^{n}\sum_{i=1}^{m} \mu_{ij}\alpha_i\beta_j.$$

This shows that the vectors $\alpha_i\beta_j$ form a generating set for F as a K-vector space.

Now assume that

$$\sum_{j=1}^{n}\sum_{i=1}^{m} \mu_{ij}\alpha_i\beta_j = 0$$

for some coefficients $\mu_{ij} \in K$. Since the β_j form an E-basis of F we conclude

$$\sum_{i=1}^{m} \mu_{ij}\alpha_i = 0$$

for all $j = 1, \dots, n$. Since the α_i form a K-basis of E it follows that $\mu_{ij} = 0$ for all i, j. This means that the vectors $\alpha_i\beta_j$ are linearly independent.

The same arguments work with minor modifications if $F|E$ or $E|K$ are infinite. In particular, we obtain a K-basis of infinite length for F if one of $[E : F]$ or $[E : K]$ are infinite. □

4.3. **Algebraic extensions.** We will be mainly interested in algebraic field extensions.

Definition 4.8. Let $F|K$ be a field extension. An element $\alpha \in F$ is called *algebraic* over K if there exists a nonzero polynomial $f(x) \in K[x]$ such that $f(\alpha) = 0$. Otherwise α is called *transcendental* over K. The extension $F|K$ is called algebraic if every element of F is algebraic over K.

Example 4.9. Any $\alpha \in K$ is algebraic over K, as it is the zero of $x - \alpha \in K[x]$. The element $\sqrt{2} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$ because it is a zero of $f(x) = x^2 - 2 \in \mathbb{Q}[x]$. It can be shown that the Euler number $e \in \mathbb{R}$ is transcendental over $\mathbb{Q}$. Similarly, π is transcendental over $\mathbb{Q}$, a fact which we will use to show that *squaring the circle* by ruler and compass is impossible. See the final section of this chapter for more information.

Proposition 4.10. *Every finite extension $F|K$ is algebraic.*

Proof. We have to show that any element $\alpha \in F$ is algebraic over K. Assume $[F : K] = n$ and consider the elements $1, \alpha, \alpha^2, \dots \alpha^n$. Since these are $n+1$ elements they cannot be linearly independent. Therefore there are $a_0, a_1, \dots, a_n \in K$ such that

$$a_0 1 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n = 0,$$

and not all a_j are zero. This means precisely that α is a zero of

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in K[x],$$

the latter being a nonzero polynomial. Therefore α is algebraic over K. □

Using Proposition 4.10 we obtain a large supply of algebraic extensions. For instance, $\mathbb{C}|\mathbb{R}$ is algebraic, and similarly $\mathbb{Q}(\sqrt{2})|\mathbb{Q}$ since these extensions are finite. We note that algebraic extensions need not be finite however.

4.4. **Simple extensions.** Here is a different approach to the definition of $\alpha \in F$ being algebraic over a subfield K: α defines an *evaluation map*

$$\mathrm{ev}_\alpha : K[x] \to F, \qquad \sum_{i=0}^n a_i x^i \mapsto \sum_{i=0}^n a_i \alpha^i,$$

where $\sum_{i=0}^n a_i\alpha^i$ is computed in F. This map is both a ring homomorphism and a linear map between K-vector spaces, and by very definition of what it means to be algebraic, we have:

Lemma 4.11. *An element $\alpha \in F$ is algebraic over K if and only if $\ker \mathrm{ev}_\alpha \neq 0$, that is, if ev_α is not injective.*

The benefit of this viewpoint is that we now can use all our knowledge about polynomial rings and their ideals:

Lemma 4.12. *If $\alpha \in F$ is algebraic over K, then there exists a unique monic polynomial $\mathrm{m}_{\alpha,K}(x) \in K[x]$ of smallest degree such that $\mathrm{m}_{\alpha,K}(\alpha) = 0$, and $\mathrm{m}_{\alpha,K}(x)$ divides any polynomial $g(x) \in K[x]$ with $g(\alpha) = 0$.*

Proof. $\ker \mathrm{ev}_\alpha \lhd K[x]$ is an ideal (kernel of a ring map) and hence principal ($K[x]$ is a PID), and each nonzero ideal has a unique monic generator. □

Definition 4.13. We call $\mathrm{m}_{\alpha,K}(x) \in K[x]$ the *minimal polynomial* of α over K.

Example 4.14. $f(x) = x^2 - 2$ is the minimal polynomial for $\sqrt{2}$ over $\mathbb{Q}$.

Next step, again not difficult but a new observation:

Lemma 4.15. $\mathrm{m}_{\alpha,K}(x)$ *is irreducible.*

Proof. If $\mathrm{m}_{\alpha,K}(x) = p(x)q(x)$ with $\deg(p(x)), \deg(q(x)) < \deg(\mathrm{m}_{\alpha,K}(x))$, then $p(\alpha)q(\alpha) = 0$, so $p(\alpha) = 0$ or $q(\alpha) = 0$, contradicting the minimality of $\mathrm{m}_{\alpha,K}(x)$. □

Recall that $K(\alpha) \subset F$ is the smallest subfield of F containing K and α. Combining the above with the first isomorphism theorem yields now a complete description of this field:

Lemma 4.16. *There is an isomorphism of rings and of K-vector spaces*

$$K(\alpha) \cong K[x]/(\mathrm{m}_{\alpha,K}(x)) \cong \operatorname{im} \mathrm{ev}_\alpha \subset F.$$

Proof. Recall that $\mathrm{m}_{\alpha,K}(x)$ being irreducible means that the ideal it generates is maximal in $K[x]$, but this ideal is $\ker \mathrm{ev}_\alpha$. So $K[x]/\ker \mathrm{ev}_\alpha$ is a field. The first isomorphism theorem (for rings respectively K-vector spaces) yields an isomorphism with $\operatorname{im} \mathrm{ev}_\alpha \subset F$. So $\operatorname{im} \mathrm{ev}_\alpha$ is a field that clearly contains K as the image of the constant polynomials and α as the image of x, hence $K(\alpha) \subset \operatorname{im} \mathrm{ev}_\alpha$. Conversely, any field containing K and α contains all polynomial expressions in α with coefficients in K, hence $K(\alpha) \subset \operatorname{im} \mathrm{ev}_\alpha$. □

So we see that if $F|K$ is a field extension and $\alpha \in F$, then there are two possible situations: either α is transcendental, and then ev_α is an embedding of $K[x]$ into F, or α is algebraic, in which case $\operatorname{im} \mathrm{ev}_\alpha$, that is, the ring of polynomial expressions in α with coefficients in K, is already a field (quite nontriovial I think!).

Finally, we observe the following consequence of the division algorithm:

Lemma 4.17. *If $f(x) \in K[x]$, then $\dim_K K[x]/(f(x)) = \deg(f(x))$.*

Proof. By the division algorithm, every polynomial $g(x) \in K[x]$ can be uniquely wirtten as $g(x) = q(x)f(x) + r(x)$ with $\deg(r(x)) < \deg(f(x)) =: d$. A direct restatement of this fact is that the K-vector space $K[x]$ is the direct sum of the subspace $(f(x))$ and the subspace $K[x]_{<d}$ of all polynomials of degree less than d,

$$K[x] = (f(x)) \oplus K[x]_{<d},$$

which means that $K[x]/(f(x)) \cong K[x]_{<d}$, and this vector space has dimension d (the elements $1, x, x^2, \dots, x^{d-1}$ form a basis). □

The following summarises all these insights (for the final comment recall that a finite field extension is algebraic):

Theorem 4.18. *Let $F|K$ be a field extension and let $\alpha \in F$ be algebraic over K with minimal polynomial $\mathrm{m}_{\alpha,K}(x) \in K[x]$. Then there is an isomorphism*

$$K[x]/(\mathrm{m}_{\alpha,K}(x)) \to K(\alpha), \quad f(x) + (\mathrm{m}_{\alpha,K}(x)) \mapsto f(\alpha)$$

of fields and K-vector spaces. In particular,

$$[K(\alpha) : K] = \deg(\mathrm{m}_{\alpha,K}(x))$$

so that $K(\alpha)|K$ is an algebraic extension.

The field extensions which are obtained in this way by adding a single algebraic element to K deserve a special name:

Definition 4.19. A field extension $F|K$ is called simple iff $F = K(\alpha)$ for some element $\alpha \in F$.

Corollary 4.20. *Let $K(\alpha)|K$ and $K(\beta)|K$ be simple algebraic extensions. Then the following conditions are equivalent.*

a) There exists a field isomorphism $\theta : K(\alpha) \to K(\beta)$ fixing all elements of K such that $\theta(\alpha) = \beta$.
b) α and β have the same minimal polynomial over K.

Proof. *a)* $\Rightarrow$ *b)* Assume that $\theta : K(\alpha) \to K(\beta)$ is an isomorphism such that $\theta(\alpha) = \beta$ fixing K. If $f(x) \in K[x]$ is the minimal polynomial of α then $f(\beta) = \theta(f(\alpha)) = 0$. Since $f(x)$ is irreducible this means that it is the minimal polynomial of β as well.

b) ⇒ *a)* Assume that α and β have the same minimal polynomial $f(x)$. Then we obtain field isomorphisms

$$K(\alpha) \cong K[x]/(f(x)) \cong K(\beta)$$

according to Theorem 4.18. Under the resulting isomorphism $\theta : K(\alpha) \to K(\beta)$, the element α is identified with β. Moreover we have $\theta(\lambda) = \lambda$ for all $\lambda \in K$, that is, all elements of K remain fixed under θ. □

Finally we remark that we can turn the arguments round to show that every polynomial $f(x) \in K[x]$ with coefficients in any field K admits a zero in some extension field of K:

Theorem 4.21 (Kronecker's theorem)**.** *Let K be a field and let $f(x) \in K[x]$ be a polynomial. Then there exists a field extension $F|K$ and $\alpha \in F$ such that $f(\alpha) = 0$.*

Proof. We may assume without loss of generality that $f(x)$ is irreducible. Let us define $F = K[x]/(f(x))$. According to Theorem 3.13 we know that F is a field. Since $K \subset K[x]$ we obtain in this way a field extension $F|K$. Indeed, the canonical maps $K \to K[x] \to K[x]/(f(x))$ induces a nonzero ring homomorphism $\iota : K \to F$, and since K is a field the ideal $\ker(\iota)$ must be zero. That is, the map ι is injective, and we may and will identify K as a subring of F in this way.

If we write $\alpha = x+(f(x))$ for the coset of x in the quotient, then we have $f(\alpha) = 0$ by construction. More precisely, we get $f(\alpha) = f(x) + (f(x)) = 0 + (f(x))$, which is the zero element in $K[x]/(f(x))$. □

Observe that the above construction yields $F = K(\alpha)$ since the polynomial ring is generated by K and x. That is, the proof of Theorem 4.21 provides us in fact with a simple extension of K in which $f(x)$ has a zero.

4.5. **Algebraic closures.**

In this section we discuss briefly the notion of an algebraically closed field and of algebraic closures. For our purposes in the sequel we will mainly need the fact that the field $\mathbb{C}$ of complex numbers is algebraically closed.

Proposition 4.22. *Let K be a field. Then the following properties are equivalent.*

- *a)* *Every nonconstant polynomial $f(x) \in K[x]$ of degree n has n roots in K, counted with multiplicity.*
- *b)* *K is algebraically closed.*
- *c)* *Every nonconstant polynomial $f(x) \in K[x]$ splits into a product of linear polynomials in $K[x]$.*

Proof. *a)* ⇒ *b)* is clear.
b) ⇒ *c)* If $\alpha \in K$ is a root of $f(x)$ then we can apply long division to write $f(x) = (x - \alpha)g(x)$ with some lower degree polynomial $g(x) \in K[x]$. Continuing with a root of $g(x)$ we obtain the claim inductively.

c) ⇒ *a)* If $f(x)$ can be written as a product

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$$

of linear polynomials in $K[x]$ then it has clearly exactly $n = \deg(f(x))$ roots in K, counted with multiplicity. □

We shall now define the algebraic closure of an arbitrary field K. Loosely speaking, this is the smallest algebraically closed field extension of K.

Definition 4.23. Let K be a field. An algebraic closure of K is an algebraic field extension $\overline{K}|K$ such that $\overline{K}$ is algebraically closed.

An algebraic closure $\overline{K}$ of K is always uniquely determined up to isomorphism. We will usually simply fix such a field $\overline{K}$.

Theorem 4.24. *Every field K has an algebraic closure.*

For a proof we refer to Section 31 in [2].

For our purposes it is sufficient to know that the field $\mathbb{C}$ is algebraically closed, as we already mentioned in Theorem 1.5 in Section 1.

For a subfield K of $\mathbb{C}$ we may use Theorem 1.5 to identify the algebraic closure $\overline{K}$ with a subfield of $\mathbb{C}$. In fact, one can take

$$\overline{K} = \{\alpha \in \mathbb{C} \mid \alpha \text{ algebraic over } K\}$$

in this case. Indeed, using degree considerations, one can show that $\overline{K}$ is indeed a field. It follows then by construction that the extension $\overline{K}|K$ is algebraic. Moreover, every polynomial $f(x) \in \overline{K}[x]$ has a zero in $\mathbb{C}$ because $\mathbb{C}$ is algebraically closed. Using further degree considerations, we see that such a zero is in fact already contained in $\overline{K}$. That is, $\overline{K}$ is algebraically closed.

4.6. **Constructions with Ruler and Compass.** We now use field extensions to solve some classical problems involving *constructions by ruler and compass.* For more information and historical background we refer to Chapter 7 of [5].

The aim of the game is to start with any two points in the plane $\mathbb{R}^2$ and to construct step by step new points by adding points that can be *contructed with ruler and compass* from the already existing points:

Definition 4.25. Let $P \subset \mathbb{R}^2$ be a set and each of $U, V \subset \mathbb{R}^2$ be either a line

$$\{(x, y) \in \mathbb{R}^2 \mid (x - a)(d - b) = (y - b)(c - a)\}$$

through two points $(a, b), (c, d) \in P$ or a circle

$$\{(x, y) \in \mathbb{R}^2 \mid (x - a)^2 + (y - b)^2 = r^2 = (c - e)^2 + (d - f)^2\}$$

centered at a point $(a, b) \in P$ and whose radius equals the distance of two points $(c, d), (e, f) \in P$. Then we say that the points in the intersection $U \cap V$ are *constructible from P in one step.*

The key idea that leads to an understanding which points are constructible in a finite number of steps using ruler and compass is to translate the problem into the language of field extensions by considering the smallest subfield of $\mathbb{R}$ which contains all the coordinates of already constructed points. If we start with say $(0, 0)$ and $(1, 0)$, then the first field to consider is $E_0 = \mathbb{Q}$.

Lemma 4.26. *Let $E \subset \mathbb{R}$ be a subfield. If $(x, y) \in \mathbb{R}^2$ is constructible in one step from E^2 and $F = E(x, y)$, then we have $[F : E] \in \{1, 2\}$.*

Proof. Consder first two nonparallel lines through the points $(a_i, b_i), (c_i, d_i)$, $i = 1, 2$. They intersect in a unique point (x, y) obtained by solving the equations

$$(x - a_1)(d_1 - b_1) = (y - b_1)(c_1 - a_1), \quad (x - a_2)(d_2 - b_2) = (y - b_2)(c_2 - a_2).$$

and by doing so one immediately finds that $x, y \in E$, so that $[F : E] = 1$.

Similarly, the case line meets circle leads to solving the equations

$$(x-a_1)(d_1-b_1)=(y-b_1)(c_1-a_1), \quad (x-a_2)^2+(y-b_2)^2=r^2.$$

If $c_1 \neq a_1$, then the first equation can be solved to give a relation $y=\alpha+x\beta$ with $\alpha, \beta \in E$. Inserting this into the second equation yields a quadratic equation for x with coefficients in E. Therefore the minimal polynomials of x over E has degree 1 or 2, and accordingly the field extension $E(x)|E$ has degree 1 or 2. In view of the equation $y=\alpha x+\beta$, we have $y \in E(x)$ so $E(x,y)=E(x)$. If instead $c_1=a_1$, then we must have $b_1 \neq d_1$ and we can instead write $x=\gamma y+\delta$ for some $\gamma, \delta \in E$, and obtain similarly that $x \in E(y)=E(x,y)$ and $[E(y):E] \in \{1,2\}$.

Finally, the case circle meets circle corresponds to the equations

$$r_j^2=(x-p_j)^2+(y-q_j)^2=x^2-2p_jx+p_j^2+y^2-2q_jy+q_j^2, \quad j=1,2$$

where (p_j,q_j) are the centres of the two circles and r_j are their radii. Subtracting the equations from one another yields

$$2(p_1-p_2)x+2(q_1-q_2)y=r_2^2-r_1^2.$$

which is equivalent to

$$(x-a)(d-b)=(y-b)(c-a)$$

where

$$d=p_1-p_2+b, \quad c=q_2-q_1+a$$

and (a,b) is any point satisfying the equation

$$a(p_1-p_2)+b(q_1-q_2)=\frac{r_2^2-r_1^2}{2}.$$

In other words, we can also construct the point (x,y) from E^2 by intersecting a line and a circle, and this case has already been covered above. □

Using Lemma 4.26 we can formulate the following central result on ruler and compass constructions.

Theorem 4.27. *Let $E \subset \mathbb{R}$ be a field. If $(x,y) \in \mathbb{R}^2$ is constructible in finitely many steps from E^2, then $[E(x,y):E]$ is a power of 2.*

Proof. For a constructible point (x,y) we find a tower of fields

$$E=E_0 \subset E_1 \subset \cdots \subset E_n$$

such that $(x,y) \in E_n^2$ and each extension $E_{j+1}|E_j$ is obtained by adjoining the coordinates of a point constructible from E_j^2. According to Lemma 4.26 the degree $[E_{j+1}:E_j]$ is a power of 2 for all $j=0,\ldots,n-1$, hence the tower law implies that $[E_n:E]$ is a power of 2 as well. Since $E(x,y) \subset E_n$ is a subfield, we conclude that $[E(x,y):E]$ is a power of 2 a s well, again using the tower law. □

We now use Theorem 4.27 to discuss two geometric applications. More precisely, we show that the problem of *trisecting arbitrary angles* and the problem of *squaring the circle* cannot be solved using only ruler and compass.

4.6.1. *Trisecting an angle.* An ancient problem was to trisect any given angle with a ruler and compass construction. The following theorem due to Pierre Wantzel, a student of Gauss, shows this can not be done for the angle $\pi/3$:

Theorem 4.28 (Wantzel (1837))**.** *There is no point $(x,y) \in \mathbb{R}^2$ which is constructible form $\mathbb{Q}^2$ and for which the line through (x,y) and $(0,0)$ has an angle of $\pi/9$ with the x-axis.*

Proof. It suffices to prove that $(\cos(\pi/9), \sin(\pi/9)) \in \mathbb{R}^2$ is not constructible from $\mathbb{Q}^2$. Assume otherwise. It is not difficult to show that then $(\cos(\pi/9), 0)$ is constructible as well. Moreover, whenever we can construct $(r, 0)$ we can also construct $(2r, 0)$, simply draw a circle of radius $(r, 0)$ at $(r, 0)$ and consider its intersection with the line passing through $(0, 0)$ and $(1, 0)$. In particular, $(\beta, 0)$ where $\beta = 2\cos(\pi/9)$ is constructible as well under our assumption. According to Theorem 4.27, this implies that β is contained in a finite extension of $K = \mathbb{Q}$ of even degree.

We will derive a contradiction by computing the minimal polynomial of β. Using the trigonometric identities

$$\begin{aligned}\cos(a+b) &= \cos(a)\cos(b) - \sin(a)\sin(b)\\ \sin(a+b) &= \cos(a)\sin(b) + \sin(a)\cos(b)\end{aligned}$$

we obtain

$$\begin{aligned}\cos(3\theta) &= \cos(\theta + 2\theta)\\ &= \cos(\theta)\cos(2\theta) - \sin(\theta)\sin(2\theta)\\ &= \cos(\theta)(\cos(\theta)^2 - \sin(\theta)^2) - \sin(\theta)(2\cos(\theta)\sin(\theta))\\ &= \cos(\theta)(2\cos(\theta)^2 - 1) - 2\cos(\theta)(1 - \cos(\theta)^2)\\ &= 4\cos(\theta)^3 - 3\cos(\theta)\end{aligned}$$

for any angle θ. Setting $\theta = \pi/9$ and observing $\cos(\pi/3) = 1/2$ yields

$$4\cos(\pi/9)^3 - 3\cos(\pi/9) - \frac{1}{2} = 0,$$

or equivalently,

$$\beta^3 - 3\beta - 1 = 0.$$

Now the polynomial $f(x) = x^3 - 3x - 1$ is irreducible over $\mathbb{Q}$. Indeed, we have

$$f(x+1) = x^3 + 3x^2 - 3,$$

and this is irreducible by Eisenstein's criterion. We conclude $[\mathbb{Q}(\beta) : \mathbb{Q}] = 3$, which contradicts the statement that the extension $\mathbb{Q}(\beta)|\mathbb{Q}$ should have even degree. □

Let us note that the angle π can easily be trisected using ruler and compass, so it is not true that no angle can be trisected.

4.6.2. *Squaring the circle.* The task is now to construct a square whose circumference agrees with the circumference of the unit circle

$$S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Since the length of the unit circle is $L = 2\pi$, the side length of our desired square should be $L/4 = \pi/2$. In other words, we are supposed to construct a square with sides of length $\pi/2$.

Theorem 4.29. *It is impossible to square the circle using ruler and compass.*

Proof. Here the claim is that $(\pi/2, 0)$ is not constructible form $\mathbb{Q}^2$ which follows easily from the fact that π is transcendental, $[\mathbb{Q}(\pi) : \mathbb{Q}] = \infty$ (for a proof of this fact we refer to Chapter 24 of [5]). □

4.7. **Exercises.**

*Exercise** 4.1 (C). Show that $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subset \mathbb{R}$ is a subfield of $\mathbb{R}$.

Exercise 4.2 (B). Consider the sets

$$\mathbb{Z}[i] = \{u + vi \mid u, v \in \mathbb{Z}\}, \qquad \mathbb{Q}(i) = \{u + vi \mid u, v \in \mathbb{Q}\}.$$

(1) Show that $\mathbb{Z}[i]$ is a subring of $\mathbb{C}$. Conclude that $\mathbb{Z}[i]$ is an integral domain.
(2) Show that $\mathbb{Q}(i)$ is a subfield of $\mathbb{C}$.
(3) Show that the field of fractions $\mathrm{Quot}(\mathbb{Z}[i])$ of $\mathbb{Z}[i]$ is isomorphic to $\mathbb{Q}(i)$.

*Exercise** 4.3 (A). Let $f(x) = x^2 - 2 \in \mathbb{Q}[x]$. In this excercise we study the field $\mathbb{Q}[x]/(f(x))$. We will show carefully that this field is isomorphic to the field

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subset \mathbb{R}.$$

In particular, this will show that the abstractly defined field $\mathbb{Q}[x]/(f(x))$ can be identified with a concrete set of real numbers.

(1) Verify that $f(x)$ is irreducible over $\mathbb{Q}$. Conclude that $\mathbb{Q}[x]/(f(x))$ is a field.
(2) Explain that the assignment $\mathbb{Q}[x] \ni h(x) \mapsto h(\sqrt{2}) \in \mathbb{R}$ defines a ring homomorphism $\pi : \mathbb{Q}[x] \to \mathbb{R}$.
(3) Show that the ideal $(f(x)) \subset \mathbb{Q}[x]$ generated by $f(x)$ is contained in $\ker(\pi)$.
(4) Show that π induces a homomorphism $\Pi : \mathbb{Q}[x]/(f(x)) \to \mathbb{R}$ such that

$$\Pi(h(x) + (f(x))) = \pi(h(x))$$

for all $h(x) \in \mathbb{Q}[x]$.
(5) Use the fact that $\mathbb{Q}[x]/(f(x))$ is a field to conclude that Π is injective.
(6) Show that the image of Π is given by $\mathrm{im}(\Pi) = \mathbb{Q}(\sqrt{2})$, and conclude that Π induces an isomorphism $\mathbb{Q}[x]/(f(x)) \to \mathbb{Q}(\sqrt{2})$.

5. Automorphisms, normality and separability

5.1. **K-algebras and their automorphisms.** The theme in the previous lectures was to classify field extensions up to an isomorphism of fields and of K-vector spaces. Here is a notion that generalises the notion of a field extension and allows us to discuss this in a more compact form:

Definition 5.1. An *algebra over a ring* K is a ring homomorphism $\eta : K \to F$. A *homomorphism* between a K-algebra $\eta : K \to F$ and a K-algebra $\eta' : K \to F'$ is a ring homomorphism $\varphi : F \to F'$ such that $\varphi \circ \eta = \eta'$.

Remark 5.2. Recall that we only consider commutative rings; more generally a noncommutative algebra F over a commutative ring K is a ring homomorphism $\eta : K \to Z(F) := \{a \in F \mid ab = ba \, \forall \, b \in F\}$. A standard example is $F := M_n(K)$ with $\eta(\lambda)$ equal to the diagonal matrix with diagonal entries equal to λ.

Example 5.3. $F := K[x]$ is a K-algebra, η being the inclusion of K as the constant polynomials. By composition with the canonical projection, any quotient ring $K[x]/(f(x))$ becomes a K-algebra.

Example 5.4. Any field extension $F|K$ defines a K-algebra with η being the inclusion map.

Example 5.5. The isomorphism $K[x]/(\mathrm{m}_{\alpha,K}(x)) \to K(\alpha)$ that we studied in the previous section is an isomorphism of K-algebras - this combines the statement that it is an isomorphism of rings and an isomorphism of K-vector spaces.

Remark 5.6. Often one is sloppy and just calls F the K-algebra, especially when η is just the inclusion of a subring. We will exclusively deal with the case that K is a field so that η is injective. One could therefore identify K with the subring $\eta(K) \subset F$ and suppress η, just writing λ for $\eta(\lambda)$. However, there might be several different such inclusions and hence it is helpful to treat K as a separate object.

Definition 5.7. Let $F|K$ be a field extension. We denote by $\mathrm{Aut}(F|K)$ the group of all K-algebra isomorphisms $F \to F$ (with composition as group multiplication). Its elements will be simply called the K-automorphisms of F.

So explicitly, a K-automorphism is a field automorphism $\sigma : F \to F$ which fixes all elements of $K \subset F$, that is, for which $\sigma(\alpha) = \alpha$ for all $\alpha \in K$. Equivalently, it is a field and K-vector space isomorphism.

The key insight of Galois was that the solvability of a polynomial $f(x) \in K[x]$ for $K \subset \mathbb{C}$ is determined by the structure of the group $\mathrm{Aut}(F|K)$, where $F \subset \mathbb{C}$ is the subfield generated by the zeros $\alpha_1, \ldots, \alpha_n$ of $f(x)$. That is, F is the smallest subfield containing K such that $f(x)$ splits into linear factors in $F[x]$.

An important first observation in this context is that any automorphism $\sigma \in \mathrm{Aut}(F|K)$ permutes the zeros of $f(x)$. More precisely, if we have

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

then $f(\alpha) = 0$ implies

$$\begin{aligned} f(\sigma(\alpha)) &= a_n \sigma(\alpha)^n + a_{n-1} \sigma(\alpha)^{n-1} + \cdots + a_1 \sigma(\alpha) + a_0 \\ &= \sigma(a_n \alpha^n + a_{n-1} \alpha^{n-1} + \cdots + a_1 \alpha + a_0) \\ &= \sigma(f(\alpha)) = 0, \end{aligned}$$

so that $\sigma(\alpha)$ is a zero of $f(x)$ as well.

In this section we discuss several results about field extensions needed to formulate the main theorem of Galois theory. We shall begin with formulating how one obtains, starting from an arbitrary field K, a field extension $F|K$ in which a given polynomial $f(x) \in K[x]$ splits into linear factors.

5.2. **Splitting fields.** Let K be a field. Recall that a monic polynomial $f(x) \in K[x]$ is said to *split into linear factors* if we can write

$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$$

for some elements $\alpha_1, \dots, \alpha_n \in K$. We have seen already that we usually have to enlarge our field to obtain a factorisation of $f(x)$ into linear factors. For instance,

$$f(x) = x^2 - 2$$

is irreducible over $\mathbb{Q}$, but factorises as $f(x) = (x - \sqrt{2})(x + \sqrt{2})$ over $\mathbb{Q}(\sqrt{2})$.

Definition 5.8. Let K be a field and let $f(x) \in K[x]$ be a polynomial. An extension field F of K is called a *splitting field* for $f(x)$ if $f(x)$ factorises into linear factors over F, and if there is no intermediate field $K \subset E \subsetneq F$ with this property.

So F is supposed to be minimal amongst the fields over which $f(x)$ splits, we do not want to add more elements to K than necessary.

We now show existence and uniqueness - see this as a generalisation of the story in the previous chapter, there we adjoined one root of a polynomial to a field K, now we adjoin all roots of that polynomial:

Theorem 5.9. *Let K be a field. For every polynomial $f(x) \in K[x]$ there exists a splitting field $F|K$. We can choose F as a subfield of an algebraic closure $\overline{K}$ of K.*

Proof. Let $g(x)$ be any irreducible factor in $f(x)$ and define $F_1 = K[x]/(g(x))$ which is an extension of K as discussed before, and the polynomial $g(x)$ and hence also $f(x)$ has a root $\alpha_1 := x + (g(x)) \in F_1$. Hence we may write

$$f(x) = (x - \alpha_1) f_1(x)$$

in $F_1[x]$. Now apply the same procedure to $f_1(x) \in F_1[x]$ to define a field F_2 with at least two zeros of $f(x)$ and continue inductively. After at most $\deg(f(x))$ steps we obtain a splitting field F in this way.

Alternatively, given an algebraic closure $\overline{K}|K$ we can factorise

$$f(x) = (x - \alpha_1) \cdots (x - \alpha_n)$$

with elements $\alpha_1, \dots, \alpha_n \in \overline{K}$ and define $F = K(\alpha_1, \dots, \alpha_n) \subset \overline{K}$. □

The following theorem shows that splitting fields are essentially unique.

Theorem 5.10. *Let $\phi : K_1 \to K_2$ be a field isomorphism. Moreover let $F_1|K_1$ be a splitting field for $f(x) \in K_1[x]$ and let $F_2|K_2$ be a splitting field for $\phi(f(x)) \in K_2[x]$. Then there exists an isomorphism $\Phi : F_1 \to F_2$ extending ϕ.*

Proof. Let us clarify first that if

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

then we write $\phi(f(x)) \in K_2[x]$ for the polynomial

$$\phi(f(x)) = \phi(a_n) x^n + \phi(a_{n-1}) x^{n-1} + \cdots + \phi(a_1) x + \phi(a_0)$$

obtained by applying ϕ componentwise.

Let $\alpha_1, \dots, \alpha_n$ be the zeros of $f(x)$ in F_1. Then $F_1 = K(\alpha_1, \dots, \alpha_n)$, and we shall construct the desired extension Φ inductively. Firstly, let $f_1(x) \in K_1[x]$ be the minimal polynomial of α_1. Then $\phi(f_1(x)) \in K_2[x]$ is irreducible, and we find a zero $\beta_1 \in F_2$ of $\phi(f_1(x))$. Now using Corollary 4.20 we obtain an isomorphism $\Phi_1 : K_1(\alpha_1) \to K_2(\beta_1)$ extending ϕ. Similarly, we let $f_2(x) \in K_1(\alpha_1)[x]$ be the minimal polynomial of α_2 over $K_1(\alpha_1)$ (not over K_1!). Then $\Phi_1(f_2(x))$ is an irreducible polynomial, and it has a zero $\beta_2 \in F_2$ (think about this claim!). Using again Corollary 4.20 we obtain an isomorphism $\Phi_2 : K_1(\alpha_1, \alpha_2) \to K_2(\beta_1, \beta_2)$ extending Φ_1. After n steps we arrive at a field isomorphism $\Phi = \Phi_n : F_1 = K_1(\alpha_1, \dots, \alpha_n) \to K_2(\beta_1, \dots, \beta_n)$. Since by construction $\phi(f(x))$ splits over the subfield $K_2(\beta_1, \dots, \beta_n) \subset F_2$, we see that in fact $K_2(\beta_1, \dots, \beta_n) = F_2$ by definition of a splitting field. □

In particular, we conclude that the splitting field of a polynomial $f(x) \in K[x]$ over a field K is uniquely determined up to isomorphism.

Corollary 5.11. *Let K be a field and let $f(x) \in K[x]$. If F_1 and F_2 are splitting fields for $f(x)$ then there exists an isomorphism $F_1 \cong F_2$ iof K-algebras.*

Proof. The assertion follows by considering the special case $K_1 = K_2$ and $\phi = \mathrm{id}$ of Theorem 5.10. □

5.3. **Normality.** In this subsection we introduce the concept of a normal extension. The idea is this I'd say: we want to find all solutions to a polynomial equation $f(x) = 0$, $f(x) \in K[x]$, typically $K = \mathbb{Q}$; this is our overall topic since Lecture 1. In other words, we will have a splitting field F of $f(x)$ that we want to describe, and then we would like to have explicit formulas for the roots of $f(x)$ now viewed in $F[x]$. However, in this process we might have to solve auxiliary polynomial equations with coefficients in the new field, and the problem arises that they could have no solutions. This is what the following notion takes care of:

Definition 5.12. A field extension $F|K$ is called *normal* if every irreducible polynomial $f(x) \in K[x]$ with a zero in F splits over F.

And the folliwng tells us not to worry, this is no problem as long as we work in a splitting field, that is, do not make our field bigger than we need:

Theorem 5.13. *Let $E|K$ be a finite extension. Then the following conditions are equivalent.*

a) $E|K$ is normal.
b) $E|K$ is the splitting field of some polynomial $f(x) \in K[x]$.

Proof. a) $\Rightarrow$ *b)* Since $E|K$ is finite we can write $E = K(\alpha_1, \dots, \alpha_n)$ for some elements $\alpha_1, \dots, \alpha_n \in E$. If $f_j(x)$ is the minimal polynomial of α_j then $f_j(x)$ splits over E into linear factors by normality. We conclude that $E|K$ is the splitting field of $f(x) = f_1(x)f_2(x)\cdots f_n(x)$.

b) $\Rightarrow$ *a)* Assume that E is the splitting field of $f(x) \in K[x]$. Let $g(x) \in K[x]$ be any irreducible polynomial with a zero in E. We have to show that $g(x)$ splits in $E[x]$. To this end let F be a splitting field of $f(x)g(x)$ such that $E \subset F$ (e.g. view $g(x)$ as an element of $E[x]$ and adjoin the zeros of $g(x)$ in an algebraic closure $\bar{E}$ to E). Moreover let $\beta_1, \beta_2 \in F$ be zeros of $g(x)$. We claim that

$$[E(\beta_1) : E] = [E(\beta_2) : E]. \tag{5.1}$$

This is proved as follows. Consider the towers of fields

$$\begin{aligned} K \subset K(\beta_1) \subset E(\beta_1) \subset F \\ K \subset K(\beta_2) \subset E(\beta_2) \subset F. \end{aligned}$$

For $j = 1, 2$ we have

$$[E(\beta_j) : E][E : K] = [E(\beta_j) : K] = [E(\beta_j) : K(\beta_j)][K(\beta_j) : K]. \tag{5.2}$$

Since $g(x) \in K[x]$ is irreducible we have a K-isomorphism $K(\beta_1) \cong K(\beta_2)$ according to Corollary 4.20, in particular

$$[K(\beta_1) : K] = [K(\beta_2) : K]. \tag{5.3}$$

Now $E(\beta_j)$ is the splitting field of $f(x)$ over $K(\beta_j)$, and by Theorem 5.10 we conclude that $E(\beta_1) \cong E(\beta_2)$ and

$$[E(\beta_1) : K(\beta_1)] = [E(\beta_2) : K(\beta_2)]. \tag{5.4}$$

Combining equations (5.4), (5.2) and (5.3) we obtain equation (5.1) as desired.

Now if $\beta_1 \in E$ then $E(\beta_1) = E$ and therefore $[E(\beta_1) : E] = 1$. By our above considerations we deduce $[E(\beta_2) : E] = 1$, which in turn means $\beta_2 \in E$. That is, if $g(x)$ has a zero in E then every other zero of $g(x)$ will be contained in E as well. This means that $E|K$ is normal. □

Example 5.14. Consider the field $\mathbb{Q}(\sqrt[3]{2})$ obtained by adjoining the cubic root $\alpha = \sqrt[3]{2} \in \mathbb{R}$ to $\mathbb{Q}$. The minimal polynomial of α over $\mathbb{Q}$ is $f(x) = x^3 - 2$. Over $\mathbb{C}$ we have

$$f(x) = (x - \sqrt[3]{2})(x - \sqrt[3]{2}\zeta)(x - \sqrt[3]{2}\zeta^2)$$

where $\zeta = e^{2\pi i/3}$. In particular, we conclude that $\mathbb{Q}(\sqrt[3]{2})|\mathbb{Q}$ is not normal because $f(x)$ has a zero in $\mathbb{Q}(\sqrt[3]{2})$, but it does not split over $\mathbb{Q}(\sqrt[3]{2})$. To obtain a normal extension we have to adjoin ζ as well, indeed, according to Theorem 5.13 the extension $\mathbb{Q}(\sqrt[3]{2}, \zeta)|\mathbb{Q}$ is normal because $\mathbb{Q}(\sqrt[3]{2}, \zeta)$ is the splitting field of $f(x)$.

As a consequence of Theorem 5.13 we obtain the following extension theorem.

Theorem 5.15. *Let $F|K$ be a normal extension. If $K \subset E \subset F$ is an intermediate field, then any K-algebra homomorphism $\sigma : E \to F$ extends to a K-algebra homomorphism $F \to F$.*

Proof. According to Theorem 5.13 we know that F is the splitting field of a polynomial $f(x) \in K[x]$. We may thus write $F = E(\alpha_1, \ldots, \alpha_n)$ where $\alpha_j \in F$ are zeros of $f(x)$ not contained in E. Let $f_1(x) \in E[x]$ be the minimal polynomial of α_1 over E. Then the polynomial $\sigma(f_1(x))$ obtained by applying σ to the coefficients of $f_1(x)$ is irreducible over $\sigma(E)$, and since $\sigma(f_1(x))$ divides $\sigma(f(x)) = f(x)$ in $F[x]$ we see that it has a zero β_1 in F. We may use σ to identify the fields E and $\sigma(E)$, and under this identification the minimal polynomial $\sigma(f_1(x))$ of β_1 corresponds to $f_1(x)$. Using Corollary 4.20 we obtain an extension of σ to a homomorphism $\sigma_1 : E(\alpha_1) \to \sigma(E)(\beta_1)$ such that $\sigma_1(\gamma) = \sigma(\gamma)$ for $\gamma \in E$ and $\sigma_1(\alpha_1) = \beta_1$. We may now repeat the argument with E replaced by $E(\alpha_1)$ and obtain inductively the desired extension. □

Remark 5.16. We did *not* claim this extension is unique! Even starting the inclusion σ could extend in different ways to F and in this way we will get some subgroup of $\mathrm{Aut}(F|K)$. We'll develop that thought further very soon.

The following is a bit weird I admit. It somehow says that being normal can be tested within suitable bigger fields in which E is somewhat rigidly embedded once one embeds K:

Theorem 5.17. *Let $E|K$ be a finite extension. Then the following conditions are equivalent.*

a) *$E|K$ is normal.*
b) *There exists a finite normal extension $F|K$ such that $K \subset E \subset F$ and for every K-homomorphism $\sigma : E \to F$ one has $\operatorname{im}(\sigma) = \sigma(E) = E$.*

Proof. *a)* $\Rightarrow$ *b)* We may simply take $F = E$.

b) $\Rightarrow$ *a)* Let $f(x) \in K[x]$ be a polynomial having a zero $\alpha \in E$. Without loss of generality we may assume that $f(x)$ is irreducible. Then $f(x)$ splits in F by normality of $F|K$.

Now let $\beta \in F$ be any other zero of $f(x)$. Using Corollary 4.20 we obtain a K-homomorphism $\sigma : K(\alpha) \to K(\beta) \subset F$ such that $\sigma(\alpha) = \beta$. According to Theorem 5.15 we can extend σ to a homomorphism $F \to F$, which we shall again denote by σ. Our assumption implies that $\operatorname{im}(\sigma) = E$, in particular we obtain $\beta = \sigma(\alpha) \in E$. We conclude that any zero of $f(x)$ is contained in E, and this means precisely that $E|K$ is normal. □

To conclude this section let us discuss a specific property of normal extensions which will be important later on:

Proposition 5.18. *If $K(\alpha)|K$ is normal, then $|\operatorname{Aut}(K(\alpha)|K)| \leq [K(\alpha) : K]$.*

Proof. If $f(x) \in K[x]$ is the minimal polynomial of α, then $[K(\alpha) : K] = \deg(f(x))$ according to Proposition 4.18. Due to normality, the polynomial $f(x)$ splits over $K(\alpha)$, say with m pairwise different zeros $\alpha_1, \alpha_2, \dots, \alpha_m \in K(\alpha)$. Some of these zeros may have multiplicity greater than 1, therefore we have $m \leq \deg(f(x))$. According to Corollary 4.20, we obtain a unique K-homomorphism $\sigma_j : K(\alpha) \to K(\alpha)$ such that $\sigma_j(\alpha) = \alpha_j$. Conversely, every K-homomorphism $K(\alpha) \to K(\alpha)$ maps zeros of $f(x)$ to zeros of $f(x)$, and therefore equals σ_j for some j. That is, we have $|\operatorname{Aut}(E|K)| = m$. □

We will be interested in extensions where this becomes an equality. This is closely related to the concept of separability, to be discussed next.

5.4. **Separability.** As indicated at the end of the previous subsection, we are interested in the following class of polynomials.

Definition 5.19. Let K be a field. An irreducible polynomial $f(x) \in K[x]$ is called *separable* if every root of $f(x)$ in a splitting field F of $f(x)$ is simple. An algebraic element in an extension field of K is called separable if its minimial polynomial is separable. An algebraic extension $F|K$ is called separable if every element of F is separable.

That is, if $f(x) \in K[x]$ is a polynomial of degree n then it is separable if we have
$$f(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$$
over the splitting field F for pairwise distinct roots $\alpha_j \in F$.

Example 5.20. The *cyclotomic polynomial*
$$\Phi_5(x) = x^4 + x^3 + x^2 + x + 1$$
is separable over $\mathbb{Q}$ because its zeros are the distinct roots of unity $\zeta_k = e^{2\pi ik/5} \in \mathbb{C}$ for $0 < k < 5$. To see that these are indeed the roots of $\Phi_5(x)$ observe that $\Phi_5(x)(x - 1) = x^5 - 1$, that is, $\Phi_5(x)$ is obtained by splitting off the obvious zero $x = 1$ of the polynomial $x^5 - 1$.

Using the observations at the end of the previous subsection we immediately obtain the following result.

Proposition 5.21. *Let $E|K$ be a finite normal and separable simple extension. Then*

$$[E:K] = |\mathrm{Aut}(E|K)|.$$

The following theorem shows that the restriction to simple extensions in the previous proposition is actually no loss of generality.

Theorem 5.22 (Primitive element theorem)**.** *Let $E|K$ be a finite separable extension. Then there exists $\alpha \in E$ such that $E = K(\alpha)$.*

Proof. Let us first assume that K has infinitely many elements. Using induction, it suffices to show that $E = K(\beta, \gamma)$ is generated by a single element. Let $f(x) \in K[x]$ and $g(x) \in K[x]$ be the minimal polynomials of β and γ, respectively. Moreover let F be a field containing E such that $f(x)$ and $g(x)$ split over F. For instance, we can take F to be the splitting field of the polynomial $f(x)g(x)$, viewed as an element of $E[x]$. We let $\beta = \beta_1$ and $\beta_2, \dots, \beta_m$ be the zeros of $f(x)$ in F, and $\gamma = \gamma_1$ as well as $\gamma_2, \dots, \gamma_n$ be the zeros of $g(x)$ in F. Since $E|K$ is separable, the elements $\beta_1, \dots, \beta_m$ are pairwise distinct, and similarly $\gamma_1, \dots, \gamma_n$ are pairwise distinct. According to our assumption that K is infinite, we find a nonzero element $a \in K$ such that

$$a \neq \frac{\beta_i - \beta}{\gamma - \gamma_j}$$

for all $i = 1, \dots, m$ and all $j = 2, \dots, n$. If we set $\alpha = \beta + a\gamma$ we have

$$\alpha = \beta + a\gamma \neq \beta_i + a\gamma_j,$$

and therefore

$$\alpha - a\gamma_j \neq \beta_i$$

for all i and all $j \neq 1$. Recall that $f(x) \in K[x]$ is the minimal polynomial of β, and consider the polynomial

$$h(x) = f(\alpha - ax) \in K(\alpha)[x].$$

By construction we have $h(\gamma) = f(\beta) = 0$ and $h(\gamma_j) = f(\alpha - a\gamma_j) \neq 0$ for $j \neq 1$, since the elements β_i are the zeros of $f(x)$ in F. It follows that $h(x)$ and the minimal polynomial $g(x) \in K[x]$ of γ have a common factor in $K(\alpha)[x]$, namely the minimal polynomial of γ over $K(\alpha)$. Moreover, this minimal polynomial must be linear since γ is the only common zero of $g(x)$ and $h(x)$ in F. We conclude that $\gamma \in K(\alpha)$, and therefore also $\beta = \alpha - a\gamma \in K(\alpha)$. This means $K(\alpha) = K(\beta, \gamma)$ and finishes the proof.

If K is a finite field then E is also finite because $E|K$ is assumed to be a finite extension. In this case the group $E^\times = E \setminus \{0\}$ of invertible elements in E is cyclic: by the classification of finitee abelian groups it is isomorphic to $\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_d}$; now let m be the least common multiple of the n_i so that $x^m = 1$ for all $x \in E$. As an m-th order equation this has at most m solutions in the field E, so $m \geq |E^\times|$. However, the order of every element divides $|E^\times|$, so $m \leq |E^\times|$ and we get equality and that the group is cyclic. So we can simply take α to be a generator. □

An element α such that $E = K(\alpha)$ as in Theorem 5.22 is called a primitive element.

Applying Theorem 5.22 to Proposition 5.21 we obtain

Corollary 5.23. *Let $E|K$ be any finite normal and separable extension. Then*

$$[E : K] = |\mathrm{Aut}(E|K)|.$$

To conclude this subsection we show that separability behaves well with respect to intermediate fields.

Proposition 5.24. *Let $E|K$ and $F|E$ be finite extensions. If $F|K$ is separable then $F|E$ and $E|K$ are separable.*

Proof. Since $F|K$ is separable every element of F has a separable minimal polynomial over K. In particular, this applies to all elements of E, and therefore $E|K$ is separable as well. Now let $\alpha \in F$ and let $f(x) \in E[x]$ be the minimal polynomial of α over E. Moreover let $g(x)$ be the minimal polynomial of α over K. Then we have $f(x)|g(x)$ in $E[x]$. In particular, if $g(x)$ has only simple roots in a splitting field, then the same is true for $f(x)$. □

One can show that the converse of Proposition 5.24 holds as well. That is, if $F|E$ and $E|K$ are separable then $F|K$ is separable. We refer to section 51 in [2] for more information.

5.5. **Perfect fields.**

For applications of Galois theory to the solvability of polynomial equations over the complex numbers, separability does not create any complications. More precisely, in this subsection we show that all irreducible polynomials over subfields of $\mathbb{C}$ are separable.

Let us introduce the following terminology.

Definition 5.25. A field K is said to be perfect if every finite extension of K is separable.

Therefore, we may say equivalently that the aim of this subsection is to show that every subfield of $\mathbb{C}$ is perfect.

Lemma 5.26. *Let K be a subfield of $\mathbb{C}$ and let $f(x) \in K[x]$ be a nonzero polynomial. Then $f(x)$ has multiple roots in $\mathbb{C}$ iff $f(x)$ and the derivative $f'(x)$ have in $K[x]$ a common factor of positive degree.*

Proof. Throughout the argument we may assume without loss of generality that $f(x)$ is monic. If $f(x)$ has the repeated zero α then

$$f(x) = (x - \alpha)^2 g(x)$$

over $\mathbb{C}$ for some $g(x) \in \mathbb{C}[x]$. Hence

$$f'(x) = (x - \alpha)(2g(x) + (x - \alpha)g'(x))$$

and $f(x), f'(x)$ have the common factor $(x - \alpha)$ over $\mathbb{C}$. It follows that the minimal polynomial of α over K divides both $f(x)$ and $f'(x)$.

Conversely, assume that $f(x) = (x - \alpha_1) \cdots (x - \alpha_n)$ over $\mathbb{C}$ where all $\alpha_1, \ldots, \alpha_n$ are the pairwise distinct zeros of $f(x)$. Then

$$f'(\alpha_j) = (\alpha_j - \alpha_1) \cdots (\alpha_j - \alpha_{j-1})(\alpha_j - \alpha_{j+1}) \cdots (\alpha_j - \alpha_n) \neq 0$$

for all j, o $f'(x), f(x)$ share no zero and hence no common factor. Let us assume that $d(x)$ is a common factor of degree > 0 of $f(x)$ and $f'(x)$. Then the previous consideration shows $d(\alpha_j) \neq 0$ for all j. Since $\deg(f'(x)) < \deg(f(x))$ we have $\deg(d(x)) < n$. If $q(x)$ is such that

$$f(x) = d(x)q(x)$$

we conclude $q(\alpha_j) = 0$ for all $j = 1, \dots, n$, which contradicts the fact that $0 < \deg(q(x)) < n$. This shows that $f(x)$ and $f'(x)$ do not have a common factor of positive degree. □

Proposition 5.27. *Let K be a subfield of $\mathbb{C}$. Then every irreducible polynomial $f(x) \in K[x]$ is separable. Hence K is a perfect field.*

Proof. Let us assume that $f(x)$ has a multiple root in $\mathbb{C}$. Then according to Lemma 5.26 it has a nontrivial common factor with $f'(x)$. Since $\deg(f'(x)) < \deg(f(x))$ and $f(x)$ is irreducible this implies $f'(x) = 0$. Therefore $f(x)$ is constant. Since $f(x)$ has a multiple root we conclude that $f(x) = 0$. This in turn contradicts the assumption that $f(x)$ is irreducible, so in particular nonzero. Hence $f(x)$ has only simple roots in $\mathbb{C}$. Now if $F|K$ is any finite extension then the minimal polynomial of any $\alpha \in F$ is separable. That is, $F|K$ is a separable extension. This means precisely that K is perfect. □

Minor modifications of the above arguments show that in fact every field of characteristic zero is perfect. Finite fields and obviously algebraically closed fields are also perfect.

5.6. **Exercises.**

*Exercise** 5.1 (C). Let K be a field. Give an example (including a justification) of

(1) a finite dimensional K-vector space.
(2) an infinite dimensional K-vector space.

Exercise 5.2 (C). Consider the field extension $\mathbb{C}|\mathbb{R}$. Describe the subfield $\mathbb{R}(\Lambda)$ where $\Lambda = \{1, 2i+5\}$.

Exercise 5.3 (B). In this exercise we show that $\mathbb{R}|\mathbb{Q}$ is an infinite field extension. Throughout, we shall use the following terminology: A set X is called countably infinite if there exists a bijective map $f : \mathbb{N} \to X$.

a) Explain how to obtain a bijective map $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$. Deduce that the direct product of two countably infinite sets is again countably infinite.
b) Explain that the rational numbers $\mathbb{Q}$ form a countably infinite set. You may use without proof that an infinite subset of a countably infinite set is again countably infinite.
c) Show that every finite dimensional $\mathbb{Q}$-vector space V is countable.
d) Show that $\mathbb{R}|\mathbb{Q}$ is an infinite extension. You may use without proof that $\mathbb{R}$ is uncountable, that is, not countably infinite.

*Exercise** 5.4 (B). Find the splitting fields $F \subset \mathbb{C}$ and determine $[F : \mathbb{Q}]$ for the following polynomials in $\mathbb{Q}[x]$.

(1) $f(x) = x^2 - 3$
(2) $f(x) = x^4 - 1$
(3) $f(x) = (x^2 - 2)(x^2 - 3)$

*Exercise** 5.5 (B). Consider the field extension $\mathbb{R}|\mathbb{Q}$. In this excercise we study the field $\mathbb{Q}(\sqrt{2}, \sqrt{3}) \subset \mathbb{R}$, that is, the field $\mathbb{Q}(\Lambda)$ generated by the set $\Lambda = \{\sqrt{2}, \sqrt{3}\}$.

(1) Show that $\sqrt{3}$ is not contained in $\mathbb{Q}(\sqrt{2})$.
(2) Show that $f(x) = x^2 - 3$ is the minimal polynomial of $\sqrt{3}$ over $\mathbb{Q}(\sqrt{2})$.
(3) Show $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] = 2$ and $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$.
(4) Write down a $\mathbb{Q}$-basis of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$.

Exercise 5.6 (B). In this excercise we show that $\mathrm{Aut}(\mathbb{R}|\mathbb{Q})$ is the trivial group. In order to do this we consider an arbitrary element $\sigma \in \mathrm{Aut}(\mathbb{R}|\mathbb{Q})$ and show in several steps that it equals the identity map $\mathrm{id} : \mathbb{R} \to \mathbb{R}$.

(1) Using that every positive real number r can be written as a product $r = b^2$ for some $b \in \mathbb{R}$ show that $\sigma(r) > 0$ if $r > 0$.
(2) Show that $a < b$ implies $\sigma(a) < \sigma(b)$.
(3) Now let $a \in \mathbb{R}$ and $\epsilon > 0$. Choose a rational number q such that $0 < q \leq \epsilon$. Taking $\delta = q$ show that for any $x \in \mathbb{R}$ such that $|x - a| < \delta$ we have
$$|\sigma(x) - \sigma(a)| < \epsilon.$$
Conclude that $\sigma : \mathbb{R} \to \mathbb{R}$ is continuous.
(4) Show that a continuous map $\mathbb{R} \to \mathbb{R}$ is uniquely determined by its restriction to $\mathbb{Q}$, and conclude that $\sigma = \mathrm{id}$.

Exercise 5.7 (B). Show that the following complex numbers are algebraic over $\mathbb{Q}$ by finding $f(x) \in \mathbb{Q}[x]$ such that $f(\alpha) = 0$.

(1) $\alpha = 1 - \sqrt{2}$.
(2) $\alpha = \sqrt{2} + \sqrt{5}$.

(3) $\alpha = (1+2i)^3$.
(4) $\alpha = \sqrt{\sqrt[5]{2}+1}$.

*Exercise** 5.8 (A). Mark each of the following true or false.

(1) Every finite extension of a field is algebraic.
(2) An extension $F|K$ is finite if it can be obtained by adjoining a finite number of elements to K.
(3) Every algebraic extension is finite.
(4) $\mathbb{Q}$ is algebraically closed.
(5) $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$.

*Exercise** 5.9 (A). Let $F|K$ be a field extension. Show that the following conditions are equivalent.

(1) $F|K$ is finite.
(2) There are elements $\alpha_1, \dots, \alpha_n \in F$ which are algebraic over K such that $F = K(\alpha_1, \dots, \alpha_n)$.

Exercise 5.10 (A). Let $F|K$ be a field extension such that $[F : K] = 3$ and let $\alpha, \beta \in F \setminus K$. Show that there exist elements $a, b, c, d \in K$ such that

$$\beta = \frac{a + b\alpha}{c + d\alpha}.$$

6. Galois theory

This section contains the core results of Galois theory. They provide a strong link between subfields of certain field extensions $F|K$ and subgroups of the corresponding automorphism group $\mathrm{Aut}(F|K)$.

Definition 6.1. A finite field extension $F|K$ is called *Galois* if it is both normal and separable.

If we consider subfields of $\mathbb{C}$ then separability holds automatically due to Proposition 5.27. Therefore, given fields $K \subset F \subset \mathbb{C}$ the extension $F|K$ is Galois iff it is finite and normal iff F is the splitting field of a polynomial $f(x) \in K[x]$.

Example 6.2. The extension $\mathbb{C}|\mathbb{R}$ is a Galois extension because $\mathbb{C}$ is the splitting field of $f(x) = x^2 + 1 \in \mathbb{R}[x]$.

In honour of the work of Galois, the automorphism group $\mathrm{Aut}(F|K)$ of a Galois extension is also called the Galois group of $F|K$.

6.1. The main theorem of Galois theory.

Let $F|K$ be a field extension. If $U \subset \mathrm{Aut}(F|K)$ is a subgroup we write

$$\mathrm{Fix}(F, U) = \{\alpha \in F \mid \sigma(\alpha) = \alpha \text{ for all } \sigma \in U\}$$

for the fixed points of the group U. It is easy to check that $\mathrm{Fix}(F, U)$ is a subfield of F containing K.

Conversely, let $K \subset E \subset F$ be an intermediate field. Then

$$\mathrm{Aut}(F|E) = \{\sigma \in \mathrm{Aut}(F|K) \mid \sigma(u) = u \in E \text{ for all } u \in E\}$$

is a subgroup of $\mathrm{Aut}(F|K)$.

Proposition 6.3. *Let $F|K$ be a Galois extension and let $K \subset E \subset F$ be an intermediate field. Then $F|E$ is a Galois extension and* $\mathrm{Fix}(F, \mathrm{Aut}(F|E)) = E$.

Proof. Since $F|K$ is separable the same holds for $F|E$ by Proposition 5.24. Since $F|K$ is normal it is the splitting field of some $f(x) \in K[x]$ according to Theorem 5.13. Clearly $F|E$ is the splitting field of $f(x)$ viewed as an element of $E[x]$ as well, which due to Theorem 5.13 shows that $F|E$ is normal. Hence $F|E$ is a Galois extension. By definition, we have

$$E \subset \mathrm{Fix}(F, \mathrm{Aut}(F|E)). \tag{6.1}$$

Hence assume that $\alpha \in F$ is not contained in E. Let $f(x) \in E[x]$ be the minimal polynomial of α, and let $\beta \in F$ be another zero of $f(x)$. Note that there exists another zero because $\deg(f(x)) > 1$, and we have $\beta \neq \alpha$ since $f(x)$ is separable. According to Corollary 4.20 we obtain a field homomorphism $\sigma : E(\alpha) \to E(\beta) \subset F$ such that $\sigma(\alpha) = \beta$. Since $F|E$ is normal we conclude from Theorem 5.15 that this extends to an element of $\mathrm{Aut}(F|E)$ which we denote again by σ. Since $\beta = \sigma(\alpha) \neq \alpha$ we see that α is not contained in $\mathrm{Fix}(F, \mathrm{Aut}(F|E))$. This shows $E = \mathrm{Fix}(F, \mathrm{Aut}(F|E))$. □

Proposition 6.4. *Let $F|K$ be a Galois extension with Galois group $G = \mathrm{Aut}(F|K)$ and let $U \subset G$ be a subgroup. Then $F|\,\mathrm{Fix}(F, U)$ is a Galois extension with Galois group U.*

Proof. Since $\mathrm{Fix}(F,U)$ is an intermediate field we know from Proposition 6.3 that $F|\,\mathrm{Fix}(F,U)$ is a Galois extension. This implies

$$|\mathrm{Aut}(F|\,\mathrm{Fix}(F,U))| = [F : \mathrm{Fix}(F,U)] \tag{6.2}$$

according to Corollary 5.23.

Since $F|\,\mathrm{Fix}(F,U)$ is separable there exists $\alpha \in F$ such that $F = \mathrm{Fix}(F,U)(\alpha)$ by the primitive element Theorem 5.22. Consider the polynomial

$$f(x) = \prod_{\sigma\in U}(x-\sigma(\alpha)) \in F[x]$$

Then $f(\alpha) = 0$, and we actually have $f(x) \in \mathrm{Fix}(F,U)[x]$ since the elements of U only permute the linear factors $(x-\sigma(\alpha))$. According to Theorem 4.18 it follows that $[F : \mathrm{Fix}(F,U)] \le |U|$ since the minimal polynomial of α over $\mathrm{Fix}(F,U)$ divides $f(x)$ and $\deg(f(x)) = |U|$. Combining this with equation (6.2) we obtain

$$|\mathrm{Aut}(F|\,\mathrm{Fix}(F,U))| \le |U|,$$

and using $U \subset \mathrm{Aut}(F|\,\mathrm{Fix}(F,U))$ we conclude $U = \mathrm{Aut}(F|\,\mathrm{Fix}(F,U))$. □

We shall now formulate and prove the main theorem of Galois theory.

Theorem 6.5. *[Main theorem of Galois Theory] Let $F|K$ be a Galois extension with Galois group $G = \mathrm{Aut}(F|K)$. Moreover let*

$$\mathcal{F} = \{E \mid K \subset E \subset F \textit{ intermediate field}\}$$

be the set of all intermediate fields and

$$\mathcal{G} = \{U \mid U \subset G \textit{ subgroup}\}$$

the set of all subgroups of the Galois group. Then the following assertions hold.

a) (Galois correspondence) The maps $\alpha : \mathcal{F} \to \mathcal{G}$ defined by

$$\alpha(E) = \mathrm{Aut}(F|E)$$

and $\phi : \mathcal{G} \to \mathcal{F}$ defined by

$$\phi(U) = \mathrm{Fix}(F,U)$$

are mutually inverse bijections.

b) The maps α and ϕ are order reversing. That is,

$$E_1 \subset E_2 \Longleftrightarrow \mathrm{Aut}(F|E_2) \subset \mathrm{Aut}(F|E_1)$$

for $E_1, E_2 \in \mathcal{F}$ and

$$U_1 \subset U_2 \Longleftrightarrow \mathrm{Fix}(F,U_2) \subset \mathrm{Fix}(F,U_1)$$

for $U_1, U_2 \in \mathcal{G}$.

c) If $K \subset E \subset F$ is an intermediate field then $F|E$ is Galois and

$$[F : E] = |\mathrm{Aut}(F|E)|, \qquad [E : K] = \frac{|G|}{|\mathrm{Aut}(F|E)|}.$$

d) A subgroup $U \subset G$ is normal iff the corresponding field extension $\mathrm{Fix}(F,U)|K$ is Galois. In this case we have

$$\mathrm{Aut}(\mathrm{Fix}(F,U)|K) \cong \mathrm{Aut}(F|K)/\mathrm{Aut}(F|\,\mathrm{Fix}(F,U)) = G/U.$$

Proof. *a)* According to Proposition 6.3 we have

$$E = \mathrm{Fix}(F, \mathrm{Aut}(F|E)) = \phi(\alpha(E))$$

for $E \in \mathcal{F}$, and due to Proposition 6.4 we have

$$U = \mathrm{Aut}(F|\,\mathrm{Fix}(F,U)) = \alpha(\phi(U))$$

for $U \in \mathcal{G}$. This shows that α and ϕ are mutually inverse bijections.

b) It follows immediately from the definitions that

$$E_1 \subset E_2 \Longrightarrow \operatorname{Aut}(F|E_2) \subset \operatorname{Aut}(F|E_1)$$

for $E_1, E_2 \in \mathcal{F}$ and

$$U_1 \subset U_2 \Longrightarrow \operatorname{Fix}(F, U_2) \subset \operatorname{Fix}(F, U_1)$$

for $U_1, U_2 \in \mathcal{G}$. Since α and ϕ are bijections this yields the claim.

c) According to Proposition 6.3 the extension $F|E$ is Galois, and hence we have $[F : E] = |\operatorname{Aut}(F|E)|$ according to Corollary 5.23. From the tower law $[F : K] = [F : E][E : K]$ obtained in Proposition 4.7 and $[F : K] = |G|$ we therefore obtain

$$[E : K] = \frac{[F : K]}{[F : E]} = \frac{|G|}{|\operatorname{Aut}(F|E)|}$$

as desired.

d) Let us abbreviate $E = \operatorname{Fix}(F, U)$. Assume first that $U \subset F$ is a normal subgroup. That is, for all $\sigma \in G$ and $\tau \in U$ we have $\sigma^{-1}\tau\sigma \in U$. Since $F|K$ is normal it follows from Theorem 5.15 that every K-homomorphism $E \to F$ extends to an element of $\operatorname{Aut}(F|K) = G$. Now if $\alpha \in E = \operatorname{Fix}(F, U)$ then for any $\sigma \in G$ and $\tau \in U$ we have

$$\tau\sigma(\alpha) = \sigma(\sigma^{-1}\tau\sigma)(\alpha) = \sigma(\alpha)$$

and hence $\sigma(\alpha) \in \operatorname{Fix}(F, U)$ as well. It follows that every homomorphism $E \to F$ fixing K must preserve E, so that $E|K$ is normal according to Theorem 5.17. Since $E|K$ is separable by Proposition 5.24 we conclude that $E|K$ is Galois.

Conversely, assume that $E|K$ is Galois. Then $E|K$ is in particular normal, so that E is the splitting field of a polynomial $g(x) \in K[x]$. Now let $\sigma \in G$ be arbitrary. Then the automorphism σ permutes the zeros of $g(x)$, and since all these zeros are contained in E it follows that $\sigma(E) \subset E$, where $\sigma(E)$ denotes the image of E under σ. In fact, we obtain $\sigma(E) = E$ using that $\sigma : F \to F$ is bijective. For $\tau \in U$ and $\alpha \in E = \operatorname{Fix}(F, U)$ we therefore have

$$\sigma\tau\sigma^{-1}(\alpha) = \sigma\sigma^{-1}(\alpha) = \alpha$$

because $\sigma^{-1}(\alpha) \in E$ is fixed by all elements of U. In other words, we have $\sigma\tau\sigma^{-1} \in U$, which means that $U \subset G$ is a normal subgroup.

It remains to verify the claim on the Galois group of $E|K$. From the above observations we obtain a well-defined group homomorphism $\pi : G \to \operatorname{Aut}(E|K)$ by letting $\pi(\sigma) \in \operatorname{Aut}(E|K)$ be the restriction of $\sigma \in \operatorname{Aut}(F|K)$ to E. The map π is surjective because any field homomorphism $\sigma : E \to E \subset F$ fixing K can be extended to a K-automorphism of F according to Theorem 5.15. If $\sigma \in \ker(\pi)$ then $\sigma(\alpha) = \alpha$ for all $\alpha \in E$, which means $\sigma \in \operatorname{Aut}(F|E) = \operatorname{Aut}(F|\operatorname{Fix}(F, U)) = U$. Hence π induces an isomorphism $G/U \cong \operatorname{Aut}(E|K)$ as claimed. □

Part *d*) of Theorem 6.5 is the origin of the notion of a normal subgroup of a group. In fact, the history of group theory as a precise mathematical subject started with the work of Galois.

6.2. **An example.** In this section we study a simple example of a Galois extension and work out the Galois correspondence explicitly. A more interesting example is treated in detail in Chapter 13 of [5].

We shall consider the field $F = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ as an extension of $K = \mathbb{Q}$. Firstly, the field F is the splitting field of the polynomial

$$f(x) = x^4 - 5x^2 + 6 = (x^2 - 2)(x^2 - 3) \in \mathbb{Q}[x]$$

since

$$f(x) = (x - \sqrt{2})(x + \sqrt{2})(x - \sqrt{3})(x + \sqrt{3})$$

over $\mathbb{C}$. Hence according to Theorem 5.13 we see that $F|K$ is normal. Since $K \subset \mathbb{C}$ we conclude from Proposition 5.27 that $F|K$ is separable. Therefore $F|K$ is a Galois extension.

Let us compute the degree $[F : K]$. We have already seen that $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$, and it is not hard to check that $\sqrt{3}$ is not contained in $\mathbb{Q}(\sqrt{2})$. This implies that $f(x) = x^2 - 3$ is the minimal polynomial of $\sqrt{3}$ over $\mathbb{Q}(\sqrt{2})$, and we obtain $[\mathbb{Q}(\sqrt{2}, \sqrt{3}), \mathbb{Q}(\sqrt{2})] = 2$. Hence the tower law shows

$$[F : K] = [\mathbb{Q}(\sqrt{2}, \sqrt{3}), \mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}), \mathbb{Q}] = 2 \cdot 2 = 4.$$

Following the proof of Proposition 4.7 we see that a K-basis of F is given by the elements $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$. That is, every element $\alpha \in F$ can be written uniquely in the form

$$\alpha = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$$

with $a, b, c, d \in \mathbb{Q}$.

The main theorem of Galois theory implies that the Galois group $G = \mathrm{Aut}(F|K)$ has order 4. In order to determine the structure of G observe that an element $\sigma \in G$ is uniquely determined by its action on the four roots $\pm\sqrt{2}, \pm\sqrt{3}$ of $f(x)$. Moreover since σ preserves the roots of the irreducible factors $(x^2 - 2)$ and $(x^2 - 3)$ of $f(x)$ separately, we see that G must be a subgroup of $\mathbb{Z}_2 \times \mathbb{Z}_2$. Here we identify elements of $\mathbb{Z}_2$ with the permutations of $\pm\sqrt{2}$ and $\pm\sqrt{3}$, respectively. Since we already know $|G| = 4$ we conclude $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Let us identify explicitly the four elements of G. The neutral element corresponds to the identity automorphism $\mathrm{id} : F \to F$. Using Corollary 4.20 for $\mathbb{Q}(\sqrt{2})$ and $\alpha = \sqrt{3}, \beta = -\sqrt{3}$ we obtain an automorphism $\sigma \in \mathrm{Aut}(F|K)$ such that

$$\sigma(\sqrt{2}) = \sqrt{2}, \qquad \sigma(\sqrt{3}) = -\sqrt{3}$$

Similarly, we obtain $\tau \in \mathrm{Aut}(F|K)$ such that

$$\tau(\sqrt{2}) = -\sqrt{2}, \qquad \tau(\sqrt{3}) = \sqrt{3}$$

Note that under the isomorphism $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ indicated above, the element σ corresponds to $(0, 1)$ and τ corresponds to $(1, 0)$. The composition $\tau\sigma = \sigma\tau$ determines the remaining element of $\mathrm{Aut}(F|K)$, corresponding to the element $(1, 0) + (0, 1) = (1, 1) \in \mathbb{Z}_2 \times \mathbb{Z}_2$. Explicitly, we have

$$\tau\sigma(\sqrt{2}) = -\sqrt{2}, \qquad \tau\sigma(\sqrt{3}) = -\sqrt{3}$$

on the generators.

The only subgroups of G apart from the trivial group and G itself are of order 2, and given by the subgroups $\langle\sigma\rangle, \langle\tau\rangle, \langle\sigma\tau\rangle$, generated by σ, τ and $\tau\sigma$, respectively. Using this information, the main theorem allows us to determine all intermediate fields of $F|K$. Apart from the obvious intermediate fields K and F we find

$$\mathrm{Fix}(F, \langle\sigma\rangle) = \mathbb{Q}(\sqrt{2}), \qquad \mathrm{Fix}(F, \langle\tau\rangle) = \mathbb{Q}(\sqrt{3})$$

and

$$\mathrm{Fix}(F, \langle\sigma\tau\rangle) = \mathbb{Q}(\sqrt{6})$$

corresponding to the subgroups of order 2.

6.3. **Exercises.**

*Exercise** 6.1 (B). Let $F|K$ be a field extension.

(1) Let $U \subset \mathrm{Aut}(F|K)$ be a subgroup. Show that $\mathrm{Fix}(F, U)$ is a subfield of F containing K.
(2) Let $K \subset E \subset F$ be an intermediate field. Show that $\mathrm{Aut}(F|E)$ is a subgroup of $\mathrm{Aut}(F|K)$.

*Exercise** 6.2 (A). Let $F|K$ be a normal extension. According to theorem 5.6. the field F is the splitting field of some polynomial $f(x) \in K[x]$.
In this excercise we verify that the automorphism group $\mathrm{Aut}(F|K)$ of $F|K$ can be identified with a group of permutations of the zeros of $f(x)$. To this end let

$$Z_f = \{\alpha_1, \dots, \alpha_n\}$$

be the set of zeros of $f(x)$ in F, and denote by S_{Z_f} the group of all permutations of Z_f.

(1) Show that
$$\iota(\sigma)(\alpha) = \sigma(\alpha)$$
for $\sigma \in \mathrm{Aut}(F|K)$ and $\alpha \in Z_f$ defines a i group homomorphism
$$\iota : \mathrm{Aut}(F|K) \to S_{Z_f}.$$
(2) Show that ι is injective.
(3) Give an example of a normal extension $F|K$ as above such that ι is not surjective.

*Exercise** 6.3 (A). Consider the field $F = \mathbb{Q}(\sqrt[3]{2}, \zeta)$ inside $\mathbb{C}$ where $\zeta = e^{2\pi i/3}$.

(1) Show that $F|\mathbb{Q}$ is a Galois extension and compute the degree $[F : \mathbb{Q}]$.
(2) Prove that $\mathrm{Aut}(F|\mathbb{Q})$ is isomorphic to the symmetric group S_3 on 3 elements.

*Exercise** 6.4 (A). Mark each of the following true or false.

(1) Every polynomial splits over some field.
(2) $\mathbb{Q}(\sqrt{5})|\mathbb{Q}$ is a finite normal extension.
(3) If $F|K$ is a normal extension then every polynomial $f(x) \in K[x]$ with a zero in F splits over F.
(4) Every finite extension of a subfield of $\mathbb{C}$ is separable.
(5) $\mathbb{C}|\mathbb{R}$ is a Galois extension.

Exercise 6.5 (B). Consider the subset

$$\overline{\mathbb{Q}} = \{\alpha \in \mathbb{C} \mid \alpha \text{ algebraic over } \mathbb{Q}\}$$

of $\mathbb{C}$.

(1) Show that $\overline{\mathbb{Q}}$ is a field.
(2) Show that $\overline{\mathbb{Q}}|\mathbb{Q}$ is an infinite algebraic extension.
(3) Show that $\overline{\mathbb{Q}}$ is algebraically closed.

Exercise 6.6. A Galois extension $F|K$ is called abelian if $\mathrm{Aut}(F|K)$ is an abelian group.

(1) Show that if $K \subset E \subset F$ is an intermediate field then $E|K$ and $F|E$ are abelian Galois extensions as well.
(2) Give an example of an abelian Galois extension.

Exercise 6.7. Let $F \subset \mathbb{C}$ be the splitting field of $f(x) = x^4 + 1 \in \mathbb{Q}[x]$.

(1) Verify that the complex zeros of $f(x)$ are
$$\alpha = \frac{1+i}{\sqrt{2}}, \quad \beta = \frac{1-i}{\sqrt{2}}, \quad \gamma = -\frac{1+i}{\sqrt{2}}, \quad \delta = -\frac{1-i}{\sqrt{2}}$$
and show $F = \mathbb{Q}(\alpha)$.
(2) Show that $f(x)$ is irreducible over $\mathbb{Q}$.
(3) Show that $F|\mathbb{Q}$ is a Galois extension and determine $\mathrm{Aut}(F|\mathbb{Q})$.
(4) Determine all intermediate fields for $F|\mathbb{Q}$.

Exercise 6.8. Consider the field $F = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$ inside $\mathbb{C}$.

(1) Show that $F|\mathbb{Q}$ is a Galois extension.
(2) Compute the degree $[F : \mathbb{Q}]$.
(3) Show that $\mathrm{Aut}(F|\mathbb{Q}) \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$.
(4) Determine all subgroups of $\mathrm{Aut}(F|\mathbb{Q})$ and the number of intermediate fields for the extension $F|\mathbb{Q}$.

Exercise 6.9. Compute the Galois group of $f(x) = x^4 - 2 \in \mathbb{Q}[x]$ as follows.

(1) Show that the splitting field of $f(x)$ is $F = \mathbb{Q}(\sqrt[4]{2}, i)$ and compute $[F : \mathbb{Q}] = 8$.
(2) Show that there exist elements $\sigma, \tau \in \mathrm{Aut}(F|\mathbb{Q})$ such that
$$\sigma(\sqrt[4]{2}) = i\sqrt[4]{2}, \quad \sigma(i) = i, \quad \tau(\sqrt[4]{2}) = \sqrt[4]{2}, \quad \tau(i) = -i,$$
respectively.
(3) Use σ and τ to define 8 distinct elements of $\mathrm{Aut}(F|\mathbb{Q})$.
(4) Identify $\mathrm{Aut}(F|\mathbb{Q})$ as an abstract group.

Exercise 6.10. Let $\mathbb{F}_p$ be the field with p elements for a prime p, and let $\overline{\mathbb{F}}_p$ be an algebraic closure of $\mathbb{F}_p$. Moreover let $n \in \mathbb{N}$ and consider the polynomial $f(x) = x^q - x \in \mathbb{F}_p[x]$ where $q = p^n$.

(1) Show that $\overline{\mathbb{F}}_p|\mathbb{F}_p$ is an infinite extension by proving that any algebraically closed field has infinitely many elements.
(2) Let $\phi : \overline{\mathbb{F}}_p \to \overline{\mathbb{F}}_p$ be the Frobenius homomorphism of $\overline{\mathbb{F}}_p$, given by $\phi(a) = a^p$. Show that ϕ is surjective.
(3) Show that $\phi \in \mathrm{Aut}(\overline{\mathbb{F}}_p|\mathbb{F}_p)$.
(4) Let $F_q \subset F$ be the set of all zeros of $f(x)$. Show that F_q is a subfield of F containing q elements.

Exercise 6.11. Let $K = \mathbb{F}_p$ for a prime $p \in \mathbb{N}$, and let $E = K(t) = \mathrm{Quot}(K[t])$ be the field of fractions of $K[t]$. In this excercise we show that E admits inseparable finite extensions, or in other words, that E is an imperfect field. To this end we consider the polynomial
$$f(x) = x^p - t$$
in $E[x]$.

(1) Show that $f(x)$ has no zeros in E, and conclude that the Frobenius homomorphism $\phi : E \to E$ given by $\phi(a) = a^p$ fails to be surjective.
(2) Assume that $E \subset F$ is an extension field such that there exists $\alpha \in F$ with $f(\alpha) = 0$. Show that $f(x) = (x - \alpha)^p$ in $F[x]$.
(3) Show that $f(x)$ is irreducible over E.
(4) Now let E be an arbitrary field of characteristic p, and assume that the Frobenius homomorphism $\phi : E \to E$ of E is not surjective. Show that E is imperfect.

Exercise 6.12. Let p be a prime and consider the cyclotomic polynomial

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$$

in $\mathbb{Q}[x]$.

(1) Show that $\zeta = e^{2\pi i/p} \in \mathbb{C}$ is a zero of $\Phi_p(x)$.
(2) Show that $\mathbb{Q}(\zeta)$ is the splitting field of $\Phi_p(x)$ and determine $[\mathbb{Q}(\zeta) : \mathbb{Q}]$.
(3) Show that $\mathrm{Aut}(\mathbb{Q}(\zeta)|\mathbb{Q}) \cong \mathbb{Z}_{p-1}$.

Exercise 6.13. Let R be a commutative ring and let $R[x_1, \ldots, x_n]$ be the polynomial ring over R in n variables, obtained by iterating the construction studied in the lectures. Elements of $R[x_1, \ldots, x_n]$ are expressions of the form

$$f(x_1, \ldots, x_n) = \sum_{j_1, \ldots, j_n = 0}^{\infty} a_{j_1, \ldots, j_n} x_1^{j_1} \cdots x_n^{j_n}$$

such that only finitely many of the coefficients $a_{j_1, \ldots, j_n} \in R$ are nonzero.

(1) Show that $R[x_1, \ldots, x_n]$ has the following homomorphism extension property. If S is another commutative ring, $f : R \to S$ a ring homomorphism and $y_1, \ldots, y_n \in S$ are arbitary elements, then there exists a unique ring homomorphism $F : R[x_1, \ldots, x_n] \to S$ such that F agrees with f on constant polynomials and $f(x_j) = y_j$ for all $j = 1, \ldots, n$.
(2) Let $n > 1$. Show that $R[x_1, \ldots, x_n]$ is isomorphic to $R[x_1, \ldots, x_{n-1}][x_n]$, that is, to the polynomial ring in one variable over $R[x_1, \ldots, x_{n-1}]$.
(3) Show that $R[x_1, \ldots, x_n]$ is an integral domain iff R is an integral domain.

Exercise 6.14. Let K be a field and let $K^\times = K \setminus \{0\}$. Consider

$$G_K = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid a \in K^\times, b \in K \right\} \subset GL_2(K)$$

where $GL_2(K)$ is the group of all invertible 2×2-matrices with entries in K.

(1) Show that G_K is a subgroup of $GL_2(K)$.
(2) Show that G_K is solvable.
(3) Show that $G_{\mathbb{F}_3} \cong S_3$ and conclude that S_3 is solvable.

The group G_K is often called the $ax + b$-group of K.

Exercise 6.15. Is every polynomial $f(x) \in \mathbb{Q}[x]$ of the form

$$f(x) = ax^8 + bx^6 + cx^4 + dx^2 + e$$

solvable by radicals? Justify your answer.

Here are some further questions if you want an extra challenge.

Exercise 6.16. Let $\mathbb{C}[x_1, \ldots, x_n]$ be the polynomial ring in n variables over $\mathbb{C}$. This ring is an integral domain, and the corresponding field of quotients $\mathrm{Quot}(\mathbb{C}[x_1, \ldots, x_n])$ is denoted by $\mathbb{C}(x_1, \ldots, x_n)$. It is also called the *field of rational functions in n variables over* $\mathbb{C}$.

(1) Explain how a typical element of $\mathbb{C}(x_1, \ldots, x_n)$ looks like.
(2) Let $\sigma \in S_n$ be a permutation. Construct an element $\phi_\sigma \in \mathrm{Aut}(\mathbb{C}(x_1, \ldots, x_n)|\mathbb{C})$ such that $\phi_\sigma(x_j) = x_{\sigma(j)}$ for all $j = 1, \ldots, n$. Prove that we obtain in this way an injective group homomorphism $\iota : S_n \to \mathrm{Aut}(\mathbb{C}(x_1, \ldots, x_n)|\mathbb{C})$ such that $\iota(\sigma) = \phi_\sigma$ for all $\sigma \in S_n$.

(3) Let $U = \operatorname{im}(\iota) \cong S_n$ be the subgroup of $\operatorname{Aut}(\mathbb{C}(x_1,\dots,x_n)|\mathbb{C})$ obtained in (b) and set $K = \operatorname{Fix}(U, \mathbb{C}(x_1,\dots,x_n))$. Show that $\mathbb{C}(x_1,\dots,x_n)|K$ is a Galois extension with Galois group S_n. You may use without proof that every finite extension of a field of characteristic zero is separable.

Exercise 6.17. Let G be a finite group.

(1) Find a Galois extension $F|K$ such that G is isomorphic to the Galois group $\operatorname{Aut}(F|K)$.
(2) Which groups G do you know for which there exists a Galois extension $F|\mathbb{Q}$ such that G is isomorphic to $\operatorname{Aut}(F|\mathbb{Q})$?

The *inverse problem of Galois theory* asks wether *every* finite group G is isomorphic to $\operatorname{Aut}(F|\mathbb{Q})$ for some Galois extension $F|\mathbb{Q}$. It is an open problem - solve it and you will become famous!

7. The quintic equation

We now finally return to our original question and show that a quintic polynomial equation with rational coefficients is in general not solvable by radicals. This result was one of the original motivations for the development of Galois theory. Its proof relies on the Galois correspondence between subfields of a Galois extensions and subgroups of the Galois group. In fact, the solvability of polynomial equations is closely related to solvability of the corresponding Galois groups.

7.1. **Solvable groups.**

In this subsection we review some background from group theory, compare section 35 in [2].

Definition 7.1. A group G is called solvable if there is a series

$$\{e\} = G_0 \lhd G_1 \lhd G_2 \lhd \cdots \lhd G_n = G$$

of subgroups of G such that

a) For each j the group G_j is a normal subgroup of G_{j+1}.
b) Each quotient group G_{j+1}/G_j is abelian.

A series of groups as in definition 7.1 is called a subnormal series for G. We remark that the group G_j in a subnormal series is typically not a normal subgroup of the entire group G, indeed it is only required to be normal in G_{j+1}.
As an example of a solvable group consider the group

$$G = G_K = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid a \in K^\times, b \in K \right\} \subset GL_2(K)$$

where K is a field and $K^\times = K \setminus \{0\}$ denotes the multiplicative group of K. One may check directly that G is indeed a subgroup of $GL_2(K)$. It is often called the $ax+b$-group, because it can be viewed as the group of affine transformations of K. If we let $G_1 \subset G$ be the subgroup of all matrices with $a = 1$ then we obtain a subnormal series

$$\{e\} = G_0 \lhd G_1 \lhd G_2 = G$$

of length 2. In particular, G is solvable.
The class of solvable groups is closed under taking quotients and subgroups in the following sense.

Lemma 7.2. *Let G be a solvable group.*

a) If $N \subset G$ is a normal subgroup then G/N is solvable.
b) If $H \subset G$ is a subgroup then H is solvable as well.

Proof. a) Let

$$\{e\} = G_0 \lhd G_1 \lhd G_2 \lhd \cdots \lhd G_n = G$$

be a subnormal series for G. Then G_jN is a normal subgroup of $G_{j+1}N$ for all j, and G/N has a series

$$\{e\} = N/N = G_0N/N \lhd G_1N/N \lhd G_2N/N \lhd \cdots \lhd G_nN/N = G/N.$$

Using the second isomorphism theorem for groups we have

$$(G_{j+1}N/N)/(G_jN/N) \cong (G_{j+1}N)/(G_jN),$$

and moreover we obtain

$$(G_{j+1}N)/(G_jN) = G_{j+1}(G_jN)/(G_jN) \cong G_{j+1}/(G_{j+1} \cap G_jN).$$

The latter group is a quotient of the abelian group G_{j+1}/G_j, therefore itself abelian. This means that the above series is a subnormal series for G/N, and therefore G/N is solvable.
b) Choose the groups G_j as in *a*) and consider the groups $H_j = H \cap G_j \subset H$ for all j. If we consider the distinct groups among this family of subgroups of H we obtain a subnormal series for H. For this we use that

$$H_{j+1}/H_j = (H \cap G_{j+1})/(H \cap G_j) \cong G_j(H \cap G_{j+1})/G_j$$

by the second isomorphism theorem for groups, see theorem 34.5 in [2]. Since $G_j(H \cap G_{j+1})$ is a subgroup of G_{j+1} we see that the quotient group on the right hand side is abelian. □
For our study of polynomical equations we will be rather interested in groups that fail to be solvable.

Definition 7.3. A nontrivial group G is called simple if if does not have any nontrivial normal subgroups.

Here we recall that a subgroup of a group G is called nontrivial iff it is not equal to $\{e\}$ or G itself.
If G is a nonabelian simple group then G is not solvable. Indeed, in this case we cannot write down any subnormal series of G.
Recall that the symmetric group S_n is the group of all permutations of n elements. Inside S_n we have the subgroup A_n consisting of all even permutations, that is, all permutations that can be written as an even number of transpositions. We refer to section 9 in [2] for further information.

Theorem 7.4. *For $n \geq 5$ the alternating groups A_n are simple.*

This theorem is stated as theorem 15.15 in [2], we shall not prove it here. Combining theorem 7.4 and lemma 7.2 yields the following result.

Corollary 7.5. *For $n \geq 5$ the symmetric group S_n is not solvable.*

In our discussion below we will use the following fact.

Proposition 7.6. *Let G be a subgroup of S_5 containing a transposition and a 5-cycle. Then $G = S_5$.*

Proof. We may label the elements of the five-element set $\{1, 2, 3, 4, 5\}$ in such a way that the 2-cycle contained in G is (12) and that G contains the 5-cycle $c = (12345)$. This is justified by the fact that there exists a power of c that maps 1 to 2 since c has prime order.
We have to show that $t = (12)$ and $c = (12345)$ generate the group S_5. First we observe that

$$c^{-1}(12)c = (23), \qquad c^{-1}(23)c = (34), \qquad c^{-1}(34)c = (45).$$

Using

$$(12)(23)(12) = (13), \qquad (13)(34)(13) = (14), \qquad (14)(45)(14) = (15),$$

we conclude that G contains all transpositions $(1m)$ for $m \leq 5$. Therefore G contains all products

$$(1m)(1n)(1m) = (mn)$$

for $1 \leq m, n \leq 5$. Since every element of S_5 is a product of transpositions this finishes the proof. □

7.2. **Solution by radicals.** In this section we investigate the structure of radical extensions, that is, extensions obtained by iteratively adjoining n-th roots to a given field. We show that the resulting Galois groups are solvable. The basic idea is that the field obtained by adjoining an n-th root to a given field has an abelian Galois group. In this way we obtain a subnormal series for the Galois group of a radical extension, which in turn shows that this group is solvable.

Definition 7.7. Let $F|K$ be a field extension. An element $\alpha \in F$ is called a *radical* over K if there exists $n \in \mathbb{N}$ such that $\alpha^n \in K$.

In other words, α is a radical if it is a zero of a polynomial of the form $f(x) = x^n - a$ for some $a \in K$.

Definition 7.8. A radical extension is a field extension $F|K$ of the form

$$F = K(\alpha_1, \dots, \alpha_m)$$

such that α_1 is a radical over K and for all $j = 2, \dots, m$ the elements α_j are radical over $K(\alpha_1, \dots, \alpha_{j-1})$. The elements α_j are said to form a radical sequence in this case.

The concept of a radical extension formalises what we mean to solve an equation by radicals. More precisely, we shall adopt the following definition.

Definition 7.9. Let $K \subset \mathbb{C}$ be a field and $f(x) \in K[x]$. If $K \subset E \subset \mathbb{C}$ is a splitting field of $f(x)$ then we say that $f(x)$ is solvable by radicals if there exists $K \subset E \subset F \subset \mathbb{C}$ such that $F|K$ is a radical extension.

We point out here that we do not require the splitting field E itself to be a radical extension of K.
The next lemmas show that certain Galois groups related to radicals are abelian.

Lemma 7.10. *Let $K \subset \mathbb{C}$ and $E \subset \mathbb{C}$ be the splitting field of $x^n - 1$ over K where $n \in \mathbb{N}$. Then* $\mathrm{Aut}(E|K)$ *is abelian.*

Proof. The n-th roots of unity in $\mathbb{C}$, that is, the roots of $x^n - 1$ in $\mathbb{C}$ are all simple. If we write $\zeta = \exp(2\pi i/n)$ then we have $E = K(\zeta)$, so that any K-automorphism of E is determined by its value on ζ. Every K-automorphism permutes the zeros of $x^n - 1$, so that any $\alpha \in \mathrm{Aut}(E|K)$ is of the form $\alpha = \alpha_j$ for some $0 \le j < n$ where $\alpha_j(\zeta) = \zeta^j$. Since

$$\alpha_i \alpha_j(\zeta) = \zeta^{ij} = \alpha_j \alpha_i(\zeta)$$

we see that $\alpha_i \alpha_j = \alpha_j \alpha_i$, and we conclude that $\mathrm{Aut}(E|K)$ is abelian. □
Note that if K already contains all roots of unity then we have $E = K$ and $\mathrm{Aut}(E|K)$ is the trivial group.

Lemma 7.11. *Let $n \in \mathbb{N}$ and let $E \subset \mathbb{C}$ be a subfield in which $x^n - 1$ splits. Morever let $a \in E$ and $F \subset \mathbb{C}$ be the splitting field of $x^n - a$. Then* $\mathrm{Aut}(F|E)$ *is abelian.*

Proof. Let $\alpha \in F$ be any zero of $x^n - a$. Then the zeros of $x^n - a$ are given by $\zeta^j \alpha$ where $\zeta = \exp(2\pi i/n)$ and $1 \le j < n$. In particular, using $\zeta \in E$ we conclude $F = E(\alpha)$. Since the elements of $\mathrm{Aut}(F|E)$ permute the zeros of $x^n - a$ we see in the same way as in the proof of lemma 7.10 that $\mathrm{Aut}(F|E)$ is abelian. □
Combining these two lemmas we obtain the following result.

Proposition 7.12. *Let $K \subset \mathbb{C}$ and $a \in K$. If F is the splitting field of $f(x) = x^n - a$ then* $\mathrm{Aut}(F|K)$ *is solvable.*

Proof. Since F is the splitting field of a polynomial we know that $F|K$ is Galois. If $\alpha_1, \dots, \alpha_n \in F$ are the roots of $f(x)$ then $\zeta_j = \alpha_j \alpha_1^{-1} \in F$ are the roots of $x^n - 1$. We consider the intermediate field $E \subset F$ generated by $\zeta_1, \dots, \zeta_n$. Note that it may well happen that $E = K$, but this does not affect our argument. Since E is a splitting field the extension $E|K$ is Galois, and due to lemma 7.10 we see that $\mathrm{Aut}(E|K)$ is abelian. Now we can write $F = E(\alpha_1)$, and this is the splitting field of the minimal polynomial of α_1 over E. By the main theorem of Galois theory 6.5 we obtain a subnormal series

$$\{1\} \lhd \mathrm{Aut}(F|E) \lhd \mathrm{Aut}(F|K)$$

since $\mathrm{Aut}(F|K)/\mathrm{Aut}(F|E) \cong \mathrm{Aut}(E|K)$ is abelian. This means that $\mathrm{Aut}(F|K)$ is solvable. □

We are now ready to formulate the main theorem on solvability of polynomials.

Theorem 7.13. *Let $K \subset E \subset F \subset \mathbb{C}$ be fields such that $E|K$ is normal and $F|K$ is a radical extension. Then the group* $\mathrm{Aut}(E|K)$ *is solvable.*

Proof. Since $F|K$ is a radical extension we can write

$$F = K(\alpha_1, \dots, \alpha_m)$$

where $\alpha_1^{n_1} \in K$ and $\alpha_j^{n_j} \in K(\alpha_1, \dots, \alpha_{j-1})$ for all j.

Our first aim is to show inductively that there is a normal radical extension $L|K$ such that $E \subset F \subset L$ and $\mathrm{Aut}(L|K)$ is solvable. The field L will be obtained as the last step in a tower

$$K = L_0 \subset L_1 \subset \cdots \subset L_{m-1} \subset L_m = L$$

of radical Galois extensions of K such that $K(\alpha_1, \dots, \alpha_j) \subset L_j$ and $\mathrm{Aut}(L_j|K)$ is solvable for all j. Firstly, we define L_1 to be the splitting field of $f_1(x) = x^{n_1} - \alpha_1^{n_1} \in K[x]$. It is clear that $L_1|K$ is a radical extension. Due to theorem 5.13 we know that $L_1|K$ is normal and thus Galois, and according to proposition 7.12 we see that $\mathrm{Aut}(L_1|K)$ is solvable.

Assume now that the Galois extension $L_j|K$ in this tower is already construced, and consider the polynomial

$$f_{j+1}(x) = \prod_{\sigma \in \mathrm{Aut}(L_j|K)} (x^{n_{j+1}} - \sigma(\alpha_{j+1})^{n_{j+1}})$$

in $L_j[x]$. Since this polynomial is invariant under the action of any $\sigma \in \mathrm{Aut}(L_j|K)$ we see from proposition 6.3 that $f_{j+1}(x)$ is in fact contained in $K[x]$. We define L_{j+1} as the splitting field of $f_{j+1}(x)$ over L_j. Then L_{j+1} is a splitting field over K as well. Therefore $L_{j+1}|K$ is normal and hence Galois, and it is clear that it is a radical extension. Applying proposition 7.12 iteratively to the polynomials $x^{n_{j+1}} - \sigma(\alpha_{j+1})^{n_{j+1}} \in L_j(x)$ for $\sigma \in \mathrm{Aut}(L_j|K)$ we see that $\mathrm{Aut}(L_{j+1}|L_j)$ is solvable.

To conclude the proof we note that $\mathrm{Aut}(E|K) \cong \mathrm{Aut}(L|K)/\mathrm{Aut}(L|E)$ by the main theorem of Galois theory 6.5, using that the extension $E|K$ is normal. Therefore $\mathrm{Aut}(E|K)$ is solvable due to part $a)$ of lemma 7.2. □

7.3. **Nonsolvability of the quintic equation.** In this section we show that quintic polynomial equations are not solvable by radicals in general. More precisely, we show explicitly that this happens for the equation

$$f(x) = x^5 - 6x + 3 = 0.$$

In the sequel we shall say that $\mathrm{Aut}(F|K)$ is the Galois group of $f(x) \in K[x]$ over K if $F|K$ is the splitting field of a polynomial $f(x)$.

Lemma 7.14. *Let $f(x) \in \mathbb{Q}[x]$ be an irreducible polynomial of degree 5. If $f(x)$ has precisely two nonreal zeros in $\mathbb{C}$ then the Galois group of $f(x)$ over $\mathbb{Q}$ is isomorphic to the symmetric group S_5.*

Proof. According to the fundamental theorem of algebra, the splitting field F of $f(x)$ can be viewed as a subfield of $\mathbb{C}$. Let G be the Galois group of $F|\mathbb{Q}$. Then G is naturally a subgroup of the permutation group of the set of zeros of $f(x)$ in $\mathbb{C}$. Since $f(x)$ is assumed to be irreducible we see from proposition 5.27 that its zeros are all distinct. Therefore we can identify G with a subgroup of the permutation group S_5.

In the construction of F we first adjoin an algebraic element of degree 5, and therefore $[F : \mathbb{Q}]$ is divisible by 5. According to theorem 6.5 c) we see that 5 divides $|G|$. Hence due to Cauchy's theorem there exists an element of order 5 in G, see theorem 36.3 in [2]. Since 5 is prime, the only such elements in S_5 are 5-cycles. That is, the group G contains a 5-cycle σ.

Complex conjugation is an automorphism $\kappa : \mathbb{C} \to \mathbb{C}$ that fixes $\mathbb{Q}$. By our assumption on the structure of the zeros of $f(x)$ it follows that κ restricts to an automorphism of F. Indeed, f has precisely two complex zeros which are necessarily complex conjugate to each other. The remaining zeros are assumed to be real, so that they are fixed by complex conjugation. Since F is generated by the zeros of $f(x)$ this shows that κ induces an automorphism of F which clearly defines a 2-cycle inside G.

Using proposition 7.6 we conclude that $G = S_5$ as desired. □

We are now in a position to exhibit concrete quintic polynomials that are not solvable by radicals.

Theorem 7.15. *The polynomial $f(x) = x^5 - 6x + 3 \in \mathbb{Q}[x]$ is not solvable by radicals.*

Proof. By Eisenstein's criterion 3.18 for $p = 3$ the polynomial $f(x)$ is irreducible over $\mathbb{Q}$. In order to apply lemma 7.14 we will show that $f(x)$ has exactly 3 real zeros, each with multiplicity 1. We calculate

$$f(-2) = -17, \qquad f(-1) = 8, \qquad f(0) = 3, \qquad f(1) = -2, \qquad f(2) = 23$$

and conclude from the intermediate value theorem that $f(x)$ has at least 3 distinct real zeros. Since $f(x)$ defines a differentiable function $f : \mathbb{R} \to \mathbb{R}$, Rolle's theorem implies that the real zeros of f are separated by the zeros of the derivative f'. We compute

$$f'(x) = 5x^4 - 6$$

which has precisely

$$\pm\sqrt[4]{\frac{6}{5}}$$

as real zeros. Hence $f(x)$ has exactly three real roots.

According to lemma 7.14 we see that the Galois group of $f(x)$ is equal to S_5. The group S_5 is not solvable, and hence according to theorem 7.13 the polynomial $f(x)$ is not solvable by radicals. □

References

[1] A. Baker, An introduction to Galois theory, lecture notes
[2] J.B. Fraleigh, A first course in abstract algebra, 7th edition, Pearson
[3] J. Howie, Fields and Galois Thoery, Springer
[4] T.W. Hungerford, Algebra, Springer
[5] I. Stewart, Galois Theory, 3rd edition, Chapman&Hall/CRC Mathematics

[6] S. Weintraub, Galois Thoery, Springer

School of Maths and Stats, University of Glasgow, University Gardens, Glasgow G12 8QW, UK

E-mail address: ulrich.kraehmer@glasgow.ac.uk