M3P11: GALOIS THEORY

LECTURES BY PROF. KEVIN BUZZARD; NOTES BY ALEKSANDER HORAWA

These are notes from the course M3P11: Galois Theory taught by Prof. Kevin Buzzard in Fall 2015 at Imperial College London. They were LaTeX'd by Aleksander Horawa.

This version is from May 30, 2016. Please check if a new version is available at my website https://sites.google.com/site/aleksanderhorawa/. If you find a mistake, please let me know at aleksander.horawa13@ic.ac.uk.

Course website: http://wwwf.imperial.ac.uk/~buzzard/maths/teaching/15Aut/M3P11/

Contents

Introduction

Historically, the subject of Galois theory was motivated by the desire to solve polynomial equations:

$$p(x) = 0 \tag{1}$$

with $p(x) = a_0 + a_1 x + \cdots + a_d x^d$, a polynomial. If $a_d \neq 0$, d is the *degree* of $p(x)$. How to solve equation (1)?

If $d = 1$, it is easy: if $a_0 + a_1 x = 0$, then $x = -\frac{a_0}{a_1}$.

If $d = 2$, there is a formula for solutions of a quadratic equation. We get it by completing the square. For example: to solve

$$x^2 + 4x - 19 = 0$$

we note that it is equivalent to

$$(x+2)^2 = x^2 + 4x + 4 = 23$$

and hence $x = -2 \pm \sqrt{23}$.

If $d = 3$, the cubic equation was solved by Cardano (1545). Here is a practical method for solving cubics. Suppose we want to solve

$$Ax^3 + Bx^2 + Cx + D = 0$$

where $A \neq 0$.

Step 1. Divide by A to get an equation of the form

$$x^3 + B'x^2 + C'x + D' = 0$$

Step 2. Complete the cube: substitute $x \mapsto x + B'/3$ to get an equation of the form

$$x^3 + Ex + F = 0.$$

Step 3. Substitute $x = p + q$ with p, q unknowns (picking up an extra degree of freedom). Expand out

$$(p+q)^3 + E(p+q) + f = 0$$

to get

$$p^3 + q^3 + 3pq(p+q) = -E(p+q) - F.$$

Now we lose the extra degree of freedom by demanding that $3pq = -E$. After cancellation, we get two equations:

$$p^3 + q^3 = -F \text{ and } 3pq = -E.$$

Hence we know the sum and the product of the number p^3 and q^3:

$$p^3 + q^3 = -F \text{ and } p^3q^3 = \left(\frac{-E}{3}\right)^3,$$

so p^3 and q^3 are roots of the quadratic equation

$$X^2 + FX + \left(\frac{-E}{3}\right)^3 = 0.$$

Solve the quadratic to get that p^3 is one root (and we get 3 choices for p). For each choice of p, we get $q = \frac{-E}{3p}$ and $x = p + q$.

The quartic ($d = 4$) was solved shortly after by Cardano and his student Ferrari. Just like the quadratic and the cubic, the formula for the solutions of a quartic equation, the formula only involves the operations

$$+ \quad - \quad \times \quad \div \quad \sqrt[n]{}$$

for some n's.

The question of $d = 5$ remained open for hundreds of years. Lagrange (1770) wrote a general treatise on solving polynomial equations. He still could not solve the quintic though. This was finally resolved by Raffini (1799) (incomplete proof) and Abel (1823) (complete proof):

there is **no** formula for roots of a quintic using only $+ \quad - \quad \times \quad \div \quad \sqrt[n]{}$.

Évariste Galois (1831) gave a far more conceptual proof, inventing group theory on the way. He died at 20 in a duel (1831). The paper was only published in 1846.

1. Rings and fields

Recall that a *ring* is

- a set R,
- elements $0, 1 \in R$,
- maps $+\colon R \times R \to R$ and $\times\colon R \times R \to R$,

subject to 3 axioms

(1) $(R, +)$ is an abelian group with identity 0,
(2) $(R, \times)$ is a commutative semigroup, i.e. $a \times (b \times c) = (a \times b) \times c$, $a \times 1 = 1 \times a = a$, and $a \times b = b \times a$ for all $a, b, c \in R$,[1]
(3) *distributivity*: $a \times (b + c) = a \times b + a \times c$ for all $a, b, c \in R$.

Examples of rings: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}/n\mathbb{Z}$ for $n \geq 1$.

We are not really interested in general rings in this course. This course is about fields.

Definition. A *field* is a ring R that satisfies the following extra properties

- $0 \neq 1$,
- every non-zero element of R has a multiplicative inverse: if $r \in R$ and $r \neq 0$, then there exists $s \in R$ such that $rs = 1$; in other words: $R \setminus \{0\}$ is a group under $\times$ with identity 1.

Non-example of a field: $\mathbb{Z}$. Indeed, $3 \in \mathbb{Z}$ and $7 \in \mathbb{Z}$, but there is no integer x such that $3x = 7$, so $3/7 \notin \mathbb{Z}$.

However, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields. ($\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is prime.)

What is the point of fields? It is the correct language for vector spaces: we can talk about a vector space over a field. In this course, we will delve deeper into the internal structure of fields.

Proposition 1.1. *Suppose K is a field and $X \subseteq K$ is a subset of K, with the following properties:*

(1) $0, 1 \in X$,
(2) *if $x, y \in X$, then $x + y, x - y, x \times y \in X$ and if $y \neq 0$ then $x/y \in X$.*

Then X is a field.

Proof. By assumption, X is closed under addition and multiplication. Moreover, X is clearly a ring, because X inherits all the axioms from K. Finally, $0 \neq 1$, and if $0 \neq x \in X$, then $x^{-1} \in X$ by assumption. Therefore, X is a field. □

[1] Some people do not assume that rings are commutative. In this course, all rings are commutative.

Remark. We call X a *subfield* of K.

So let us try and write down some new fields. Think of the smallest subfield of $\mathbb{C}$ containing $\mathbb{Q}$ and i. Remember: if x, y are in this subfield, then $x+y$ and xy are too. If a field contains $\mathbb{Q}$ and i, then it must contain $x+iy$ for all $x, y \in \mathbb{Q}$. So let us consider the set

$$X = \{x + iy : x, y \in \mathbb{Q}\}.$$

Question: Is X a field, i.e. is X a subfield of $\mathbb{C}$? By Proposition 1.1, we know what we have to do. First, clearly $0, 1 \in X$ and $a+b, a-b, a \times b \in X$ for $a, b \in X$. For division, say $a = x + iy$ with $x, y \in \mathbb{Q}$ and $b = s + it$ with $s, t \in \mathbb{Q}$, not both 0. Then

$$\frac{a}{b} = \frac{x+iy}{s+it} = \frac{(x+iy)(s-it)}{(s+it)(s-it)} = \frac{\text{element of } X}{s^2+t^2} \in X,$$

since $s^2 + t^2 \in \mathbb{Q}$ is non-zero. Hence by Proposition 1.1, it is a field.

Now, set $\alpha = \sqrt[3]{2} \in \mathbb{R}$ and try to build a subfield of $\mathbb{R}$ containing $\mathbb{Q}$ and α. What do we **need** in this field? Clearly, need $x + y\alpha$ for all $x, y \in \mathbb{Q}$. Note that $\alpha^3 = 2$, but $\alpha^2 \notin \mathbb{Q}$ and it is not obviously of the form $x + y\alpha$. Challenge question: is α^2 of the form $x + y\alpha$ with $x, y \in \mathbb{Q}$? Let us throw in α^2. Set

$$X = \{x + y\alpha + z\alpha^2 \mid x, y, z \in \mathbb{Q}\}.$$

Easy: check that if $a, b \in X$, then so is $a+b, a-b, a \times b$. Challenge question: if $b \neq 0$, is $a/b \in X$?

Similarly, suppose $\alpha^{2015} = 2$ and let

$$X = \{x_0 + x_1\alpha^1 + \cdots + x_{2014}\alpha^{2014} \mid x_i \in \mathbb{Q}\}.$$

Two tricky questions:

(1) X is clearly a vector space over $\mathbb{Q}$ and $\{1, \alpha, \ldots, \alpha^{2014}\}$ is clearly a spanning set, but is it a basis?
(2) X is clearly a ring, but is it a field?

Here is an answer to question (2).

Proposition 1.2. *Suppose $V \subseteq \mathbb{C}$ is a subring (i.e. $0, 1, \in V$ and it is closed under $+, -, \times$) and $\mathbb{Q} \subseteq V$. Then V can clearly be regarded as a vector space over $\mathbb{Q}$. Assume furthermore that V is a finite-dimensional as a vector space over $\mathbb{Q}$. Then V is a field.*

Remark. The set X above satisfies the hypotheses of the proposition: since $\{1, \alpha, \ldots, \alpha^{2014}\}$ is a finite spanning set, X has a finite basis. Hence X is a field.

Proof of Proposition 1.2. By Proposition 1.1, all we need to check is that if $0 \neq v \in V$, then $1/v \in V$. Consider the map $\varphi_v \colon V \to \mathbb{C}$ given by $\varphi_v(x) = vx$ for $x \in V$. Since V is closed under $\times$, we know that $\mathrm{im}(\varphi_v) \subseteq V$, so $\varphi_v \colon V \to V$. Regarding V as a $\mathbb{Q}$-vector space, φ_v is a linear map. The kernel of φ_v is

$$\{x \in V : \varphi_v(x) = 0\} = \{x \in V : vx = 0\},$$

but $v, x \in V$ and $v \neq 0$ by assumption, which shows that $\ker(\varphi_v) = \{0\}$. In other words, the nullity of φ_v is 0, and hence, by rank-nullity theorem, the rank of φ_v is $\dim_{\mathbb{Q}}(V)$. Therefore,

$\dim \operatorname{im} \varphi_v = \dim_{\mathbb{Q}}(V)$ and $\operatorname{im} \varphi_v \subseteq V$, which shows that $\operatorname{im} \varphi_v = V$. Hence φ_v is surjective. Thus there exists $w \in V$ such that $vw = \varphi_v(w) = 1 \in V$, which shows that $1/v \in V$. □

Remark. In the language of commutative algebra, we have just proved that if R is an integral domain, finite-dimensional over a subfield K, then R is a field.

Tedious interlude about polynomial rings over a field. If R is a ring[2], we will denote by $R[x]$ the ring of polynomials in x with coefficients in R. Formally, an element of $R[x]$ is a finite formal sum

$$r_0 + r_1 x + r_2 x^2 + \cdots + r_d x^d$$

with $d \in \mathbb{Z}_{\geq 0}$. The addition, subtraction, and multiplication is defined in the obvious sense: if $f = \sum_i r_i x^i$, $g = \sum_i s_i x^i$, then

$$f \pm g = \sum_i (r_i \pm s_i) x^i$$

$$fg = \sum_k t_k x^k$$

with $t_k = \sum_{i+j} r_i s_j$, where, by convention, if $f = \sum_{i=0}^{d} r_i x^i$, then we set $r_i = 0$ if $i > d$.

Tedious check: $R[x]$ is indeed a ring.

Remark (technical but important)**.** A polynomial $f(x) \in R[x]$ gives a function: $f(x) = \sum_{i=0}^{d} a_i x^i$ and if $r \in R$, the we can define $f\colon R \to R$ by $f(r) = \sum_{i=0}^{d} a_i r^i \in R$. However, in this generality, different polynomials can give rise to the same function! For example, if $R = \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$, the different polynomials

$$f(x) = x$$

$$g(x) = x^2$$

give rise to the same function $R \to R$, the identity, and in fact

$$x^2 - x = (x - 0)(x - 1) = \prod_{\lambda \in R} (x - \lambda).$$

Polynomial facts and definitions.

The *degree* of a polynomial is the largest power of x with a non-zero coefficient, i.e.

$$\deg \left(\sum_{i=0}^{d} a_i x^i \right) = d \quad \text{if } a_d \neq 0.$$

If $f(x) = \sum_{i=0}^{d} a_i x^i$ of degree d, we say that f is *monic* if $a_d = 1$. The *leading coefficient* of $f(x)$ is the coefficient of x^d for $d = \deg(f)$ and the *constant coefficient* is the coefficient of x^0.

For example, take $R = \mathbb{R}$. The degree of $x^2 + 2x + 3$ is 2, the degree of $f(x) = 7$ is 0. The zero polynomial $f(x) = 0$ has, by convention, degree $-\infty$.

[2]We will mostly be concerned with the case when R is a field, but the following will work over a ring.

We will now look at the internal structure of a ring $K[x]$ where K is a field. Crucial observation: we can do division with remainder.

Lemma 1.3. *If K is a field and $f, g \in K[x]$ with $g \neq 0$, then there exist polynomials $q, r \in K[x]$ with $\deg(r) < \deg(g)$ and $f = q \cdot g + r$. Furthermore, q and r are unique (up to multiplication by units).*

Proof. For existence, use induction on $\deg(f)$. For uniqueness, use the fact that if $s(x)$ is a multiple of $g(x)$ but $\deg(s(x)) < \deg(g(x))$, then $s = 0$. □

Remark. Division with remainder can fail for polynomials over a general ring. For example, in $\mathbb{Z}[x]$ set $f(x) = x^2$ and $g(x) = 2x$.

By Lemma 1.3, we get the Euclid's algorithm holds for polynomial rings $K[x]$ that allows us to find greatest common divisors of polynomials.

Corollary 1.4 (Euclid's algorithm)**.** *Euclid's algorithm holds for polynomials over a field: we have*

$$\begin{aligned} f &= q_1 g + r_1 && \deg(r_1) < \deg(g) \\ g &= q_2 r_1 + r_2 && \deg(r_2) < \deg(r_1) \\ &\vdots \\ r_{n-1} &= q_n r_n + 0 && \deg(r_n) < \deg(r_{n-1}) \end{aligned}$$

and r_n is the greatest common divisor.

In a polynomial ring $K[x]$, the *units* are the non-zero constant polynomials $f(x) = c \neq 0$. Upshot: highest common divisors are only defined up to **units**. We can use this to our advantage—we can always find a greatest common divisor which is *monic*, i.e. it has leading term equal to 1.

We get the following corollary from Euclid's algorithm.

Corollary 1.5. *If $h = \gcd(f, g)$, then there exist polynomials $\lambda(x)$ and $\mu(x)$ such that*

$$h = \lambda f + \mu g.$$

Theorem 1.6. *If K is a field, then any polynomial $0 \neq f(x) \in K[x]$ can be written as:*

$$f(x) = c \times p_1 \times \cdots \times p_n$$

for $c \neq 0$, $c \in K$, and p_i irreducible polynomials. Moreover, the factorization is unique up to reordering and scaling.

What is an irreducible polynomial?

Definition. Fix a field K and say $f(x) \in K[x]$ is a polynomial. Then $f(x)$ is *irreducible* if

(1) $f(x) \neq 0$,
(2) $f(x) \neq$ constant,
(3) if $f = gh$ with $g, h \in K[x]$, then either g or h is constant.

Example. Is x^2+1 irreducible? It depends on the field K. For $K=\mathbb{C}$, it is not; indeed, $x^2+1=(x-i)(x+i)$. For $K=\mathbb{Q}$, it is irreducible. If it factored as $f=gh$ with g,h non-constant, then $\deg(g)=\deg(h)=1$, but the linear factors give rise to roots of x^2+1, and x^2+1 has no roots in $\mathbb{Q}$.

However, this method could fail. For example, the polynomial $f(x)=x^4+5x^2+4\in\mathbb{Q}[x]$ is reducible: $f(x)=(x^2+1)(x^2+4)$, but it has no roots in $\mathbb{Q}$.

We need tricks to factor polynomials and check for irreducibility. Recall that finding a linear factor of f is essentially the same as finding a root of the polynomial. When does a polynomial have a root? That depends very much on the field K.

(1) Let $K=\mathbb{C}$. If $\deg(f)\geq 1$, $f\in\mathbb{C}[x]$, then f always has a root by the fundamental theorem of algebra. As a consequence, $f\in\mathbb{C}[x]$ is irreducible if and only if f is linear.
(2) Let $K=\mathbb{R}$. If $\deg(f)$ is odd, then f has a root.
(3) Let $K=\mathbb{Q}$. Say $f(x)=\sum_{i=0}^{d} c_i x^i$. Does f have a root $a/b\in\mathbb{Q}$? By clearing denominators, we can assume that $f(x)\in\mathbb{Z}[x]$, i.e. $c_i\in\mathbb{Z}$ for all i. Suppose $x=a/b$ is a root with $a,b\in\mathbb{Z}$ coprime. Substituting in $f(x)$, we get

$$\sum_{i=0}^{d} c_i\frac{a^i}{b^i}=0$$

and multiplying by b^d, we get

$$c_0b^d+c_1ab^{d-1}+c_2a^2b^{d-2}+\cdots+c_{d-1}a^{d-1}b+c_da^d=0.$$

Hence

$$c_0b^d=-(ac_1b^{d-1}+c_2ab^{d-2}+\cdots)=\text{ mulitple of } a$$

but a,b are coprime, so a divides c_0. Similarly, b divides c_aa^d and hence b divides c_d. Therefore, if $c_0,c_d\neq 0$, there are only finitely many possibilities for a/b: check them all![3]

Unfortunately, it is not just about finding roots. For example, $x^2-2\in\mathbb{Q}[x]$ has no roots in $\mathbb{Q}$ and is irreducible, but $(x^2-2)^2\in\mathbb{Q}[x]$ has no roots in $\mathbb{Q}$ either, but it is reducible. Here is another technique.

Proposition 1.7 (Gauss's Lemma)**.** *If $f\in\mathbb{Z}[x]$ and f factors as $f=gh$ with $\deg(g)=m\geq 0$, $\deg(h)=n\geq 0$, and $g,h\in\mathbb{Q}[x]$, then it factors as $f=g'h'$ with $\deg(g')=m$, $\deg(h)=n$, and $g',h'\in\mathbb{Z}[x]$.*

Proof. Since $g,h\in\mathbb{Q}[x]$, by clearing denominators, there exists $N\in\mathbb{Z}_{\geq 1}$ such that Nf can be factored as the product of two polynomials in $\mathbb{Z}[x]$ of degrees m and n. Let M be the smallest positive integer such that $Mf=GH$ with $\deg(G)=m$, $\deg(H)=n$, and $G,H\in\mathbb{Z}[x]$. We claim that $M=1$. Assume for a contradiction that $M>1$ and let p be a prime factor of M.

[3]Note that $c_d\neq 0$ if $\deg(f)=d$ and if $c_0=0$, then x divides f.

We claim that either all the coefficients of G are multiples of p, or all the coefficients of H are. If not, set $G = \sum_i a_i x^i$, $H = \sum_i b_j x^j$ and say $\alpha \leq m$ is the smallest integer such that a_α is not a multiple of p and say $\beta \leq n$ is the smallest integer such that b_β is not a multiple of p. Consider the coefficient of $x^{\alpha+\beta}$ in $Mf = GH$:

$$a_0 b_{\alpha+\beta} + a_1 b_{\alpha+\beta-1} + \cdots + a_\alpha b_\beta + a_{\alpha+1} b_{\beta-1} + \cdots + a_{\alpha+\beta} b_0.$$

But $a_\alpha b_\beta$ is not a multiple of p and a_i is a multiple of p for $i < \alpha$, b_j is a multiple of p for $j < \beta$. Hence the coefficient of $x^{\alpha+\beta}$ in Mf is not a multiple of p, which contradicts $p|M$ and $f \in \mathbb{Z}[x]$. Hence either all the a_i or all the b_j are multiples of p.

Without loss of generality, suppose all the a_i are multiples of p. Then setting $G' = G/p \in \mathbb{Z}[x]$, we get that $\frac{M}{p} f = G'H$, contradicting the minimality of M. □

Corollary 1.8 (Eisenstein's criterion)**.** *Let $f(x) \in \mathbb{Z}[x]$ and $f(x) = \sum_{i=0}^{d} a_i x^i$. If there exists a prime number p such that*

(1) $p \nmid a_d$,
(2) $p \mid a_i$ *for* $0 \leq i < d$,
(3) $p^2 \nmid a_0$,

then $f(x)$ is irreducible in $\mathbb{Q}[x]$.

Proof. Assume f factors in $\mathbb{Q}[x]$ as $f = gh$, $\deg g = m$, $\deg h = n$, $m, n \geq 1$ and seek a contradiction. By Gauss's Lemma 1.7, we can assume that $g, h \in \mathbb{Z}[x]$. Then set

$$g = \sum_{i=0}^{m} b_i x^i,$$

$$h = \sum_{j=0}^{n} c_j x^j.$$

Then the leading coefficient of f is $b_m c_n = a_d$ which is not a multiple of p by (1). Hence b_m and c_n are not multiples of p. Set

$$\beta = \text{smallest } i \text{ such that } p \nmid b_i,$$

$$\gamma = \text{smallest } j \text{ such that } p \nmid c_j.$$

Using the same trick as in Gauss's Lemma 1.7, we obtain that $p \nmid a_{\beta+\gamma}$. Hence $\beta + \gamma = d$ by (2). Hence $\beta = m$ and $\gamma = n$, since $\beta \leq m$, $\gamma \leq n$. In particular, $p|b_0$ and $p|c_0$, since $m, n \geq 1$, and hence $p^2|b_0 c_0 = a_0$, contradicting (3). □

Example. The polynomial $x^{100} - 2$ is irreducible in $\mathbb{Q}[x]$. To see this, use the Eisenstein's criterion 1.8 with $p = 2$.

If p is prime, then $f(x) = 1 + x + \cdots + x^{p-1}$ is irreducible in $\mathbb{Q}[x]$. To see this, apply Eisenstein's criterion 1.8 to $f(x + 1)$.

2. Field extensions

In this chapter, we will analyze the situation where we have two fields $K \subseteq L$. (Given L, construct some K's, and given K, construct some L's.)

Let us start with a fundamental property of subfields.

Proposition 2.1. *If L is a field and K_i (finitely or infinitely many) are subfields of L, then*

$$M = \bigcap_i K_i = \{\lambda \in L : \lambda \in K_i \text{ for all } i\}$$

is also a subfield.

Proof. By Proposition 1.1, we need to check that $0, 1 \in M$ and if $a, b \in M$, then so is $a + b$, $a - b$, ab, and, if $b \neq 0$, a/b. But this is clear. Since K_i are fields for all i, $0, 1 \in K_i$ for all i, and hence $0, 1 \in M$. If $a, b \in M$, then $a, b \in K_i$ for each i, and hence $a + b$, $a - b$, ab, and, if $b \neq 0$, a/b are in K_i for each i, and therefore also in M. □

Consequence: If L is a field and $S \leq L$ is a subset, we can consider every subfield of L containing S (such subfields exist, e.g. L itself). Then by Proposition 2.1, the intersection of all of them is also a subfield of L containing S. The intersection is hence *the subfield of L generated by S*.

Note that this definition is highly non-constructive. For example, let $L = \mathbb{C}$ and $S = \{\pi\}$. What are all subfields of L containing S? Goodness knows.

Another question we can ask: given a field L, what is the intersection of all subfields of L?[4] Let us try and figure this out. We start with a field L. If $K \leq L$, then by definition $0, 1 \in K$ and $0 \neq 1$. We also know that K is a group under $+$ with identity 0. We know that $1+1 \in K$, $1 + 1 + 1 \in K$, $-(1 + 1 + 1) \in K$ etc. More formally, we have a group homomorphism

$$\theta \colon \mathbb{Z} \to K$$

such that $\theta(0_{\mathbb{Z}}) = 0_K$, $\theta(n) = \underbrace{1_K + 1_K + \cdots + 1_K}_{n \text{ times}}$ for $n > 0$, and $\theta(-n) = -\theta(n)$ for $n > 0$. The image of θ will be the cyclic subgroup of L generated by 1. Let us call this cyclic subgroup $C \leq L$ and the argument above shows that any subfield $K \leq L$ contains C.

Examples.

(1) Let $L = \mathbb{Z}/p\mathbb{Z}$ for a prime number p. This is a field and $C = \mathbb{Z}/p\mathbb{Z}$, so the only subfield of L is L itself.
(2) Let $L = \mathbb{C}$. Then $C = \mathbb{Z} \subseteq \mathbb{C}$. Note that C is not a subfield, it is only a subring.

The examples above show that $\theta \colon \mathbb{Z} \to L$ that θ might or might not be injective.

Case 1. θ is not injective.

[4] For a group, the intersection of all its subgroups is the trivial subgroup. However, for a field, we assume that $0 \neq 1$, so this intersection will be non-trivial.

Then $\ker\theta$ is a subgroup of $\mathbb{Z}$ and this subgroup is not $\{0\}$. This means that the subgroup $\ker\theta$ must be $n\mathbb{Z}$, i.e. multiples of n, where n is the smallest positive element of $\ker\theta$. By the first isomorphism theorem:

$$C = \operatorname{im}\theta = \mathbb{Z}/n\mathbb{Z}.$$

What can we say about n? Recall that $C \leq L$ and L is a field. First, $n \neq 1$; otherwise $0 = \theta(1) = 1$. Furthermore, n cannot be composite. Otherwise, $n = ab$ with $1 < a, b < n$, and then $\theta(a) \neq 0$, $\theta(b) \neq 0$, but $\theta(a)\theta(b) = \theta(ab) = \theta(n) = 0$. This cannot happen in a field. Therefore, n is prime and set $n = p$.

Then $C \cong \mathbb{Z}/p\mathbb{Z}$ and C is contained in any subfield of L. But in this case C is a field, and hence the intersection of all subfields of L must be $C \cong \mathbb{Z}/p\mathbb{Z}$.

In this case, we say that L has *characteristic* p. Note that L has characteristic p if and only if $\underbrace{1+1+\cdots+1}_{p \text{ times}} = 0$.

Case 2. θ is injective.

Then $C \cong \mathbb{Z} \leq L$. In this case, C is not a field. But L is a field, so we can do division in L. This means that L must contain a copy of $\mathbb{Q}$. (If $\frac{m}{n} \in \mathbb{Q}$, $n \neq 0$, then L contains $\theta(m)$ and $\theta(n) \neq 0$, and thus also $\theta(m)/\theta(n)$. We call this $\frac{m}{n}$.) Of course, if $K \leq L$ is a subfield, then $C \subseteq K$, so K also contains a copy of $\mathbb{Q}$.

Upshot: $\mathbb{Q} \leq L$ and $\mathbb{Q}$ is the intersection of all subfields of L.

If $\mathbb{Q} \leq L$, we say L has *characteristic* 0.

Examples (characteristic p fields that are not $\mathbb{Z}/p\mathbb{Z}$)**.** Let $K = \mathbb{Z}/3\mathbb{Z}$ and note that $x^2 = -1$ has no solution. Set $L = \mathbb{Z}/3\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}/3\mathbb{Z}\}$ with $i^2 = -1$. Check: L is a field and $L \neq C$. Then $L \neq \mathbb{Z}/9\mathbb{Z}$, but $\#L = 9$.

There are also infinite examples. Let k be the field $\mathbb{Z}/p\mathbb{Z}$ and $k[X]$ be the polynomial ring. Then $k[X]$ is an integral domain and its field of fractions $k(X) = \{f(x)/g(x) : f, g \in k[X]\}$ is an infinite field of characteristic p. Another example would be the algebraic closure of $\mathbb{Z}/p\mathbb{Z}$.

Definition. The *prime subfield* of a field L is the intersection of all subfields of L.

We have shown that the prime subfield of L is isomorphic to $\mathbb{Z}/p\mathbb{Z}$ (characteristic p) or $\mathbb{Q}$ (characteristic 0).

Here is an application of Proposition 2.1 we will see most often:
Suppose L is a field, $K \subseteq L$ is a subfield, and say $S = \{a_1, a_2, \ldots, a_n\}$ (or even $S = \{a_1, a_2, a_3, \ldots\}$) is a subset of L. We are interested in the smallest subfield of L that contains K and S. It exists by Proposition 2.1. We use the notation

$$K(a_1, a_2, \ldots, a_n) = \text{smallest subfield of } L \text{ contatining } K \text{ and } S = \{a_1, a_2, \ldots, a_n\},$$

$$K(a_1, a_2, \ldots) = \text{smallest subfield of } L \text{ contatining } K \text{ and } S = \{a_1, a_2, \ldots\}.$$

Example. Take $K = \mathbb{Q}$ and $L = \mathbb{C}$. What is $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt[100]{2})$, $\mathbb{Q}(\pi)$? What about $\mathbb{Q}(\sqrt{2}, \sqrt{3})$?

Fact (Lindemann)**.** *If $p(X) \in \mathbb{Q}[X]$ is a polynomial and $p(\pi) = 0$, then $p(X) = 0$, i.e. π is* transcendental.

Easy observation:
If $K \subseteq L$, $a \in L$, and $M = K(a) \subseteq L$, then M has the following property: if $p(X) \in K[X]$, then $p(a) \in M$.
(Why? Say $N \leq L$ is any subfield such that $K \subseteq N$ and $a \in N$. Then $p(X) = c_0 + c_1 x + \cdots + c_d x^d$ with $c_i \in K \subseteq N$, and $c_i a^i \in N$, since $a \in N$ (closed under multiplication), so $p(a) \in N$ (closed under addition).)
For example: $\sqrt[100]{2} + 3(\sqrt[100]{2})^2 \in \mathbb{Q}(\sqrt[100]{2})$, $\pi + 9\pi^2 + 4 \in \mathbb{Q}(\pi)$.
(Exercise: If $p(X_1, \ldots, X_n) \in K[X_1, \ldots, X_n]$, then $p(a_1, \ldots, a_n) \in K(a_1, \ldots, a_n)$.)

Example. Let us work out $\mathbb{Q}(\sqrt{2})$. By the easy observation above, we know that

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

We claim that $\mathbb{Q}(\sqrt{2}) = \{a+b\sqrt{2} : a, b \in \mathbb{Q}\}$. Why? Because the right hand side is a subring of $\mathbb{C}$ with dimension at most 2 as a vector space over $\mathbb{Q}$, since $1, \sqrt{2}$ span it. Therefore, by Proposition 1.2, the right hand side is a field. Therefore, equality must hold.

Example. Now say $\alpha^{100} = 2$ and $\alpha \in \mathbb{R}_{>0}$. By the same argument

$$\mathbb{Q}(\alpha) \supseteq \{a_0 + a_1\alpha + \cdots + a_{99}\alpha^{99} : a_i \in \mathbb{Q}\},$$

and the right hand side is a ring and hence a field by Proposition 1.2 and hence equality holds.[5]

Example. Note that $\mathbb{Q}(\pi) \subseteq \mathbb{C}$. Clearly, $\mathbb{Q}(\pi) \supseteq \{f(\pi) : f(X) \in \mathbb{Q}[X]\} = R$. In fact, the map $\mathbb{Q}[X] \to R$ sending $f(X) \mapsto f(\pi)$ is an isomorphism of rings. This is different from the previous cases, because π is transcendental. In this case, $\dim_{\mathbb{Q}}(R) = \infty$, because $\{1, \pi, \pi^2, \ldots\}$ is an infinite subset, linearly independent over $\mathbb{Q}$. Therefore, we cannot use Proposition 1.2, and indeed it is easy to check that $\frac{1}{\pi} \notin R$ even though $\frac{1}{\pi} \in \mathbb{Q}(\pi)$. In fact, one can check that

$$\mathbb{Q}(\pi) = \left\{ \frac{f(\pi)}{g(\pi)} : f(X), g(X) \in \mathbb{Q}[X], g(X) \neq 0 \right\},$$

the field of fractions of R.

Definition. An element $a \in \mathbb{C}$ is *algebraic* over $\mathbb{Q}$ if there exists $0 \neq p(X) \in \mathbb{Q}[X]$ such that $p(a) = 0$. Otherwise, a is *transcendental*.

Moral of this lecture: $\mathbb{Q}(a)$ depends on whether a is algebraic or transcendental.

We can generalize the definition to other fields than $\mathbb{Q}$.

Definition. Say $K \subseteq L$ are fields. If $a \in L$, we say a is *algebraic* over K if there exists $0 \neq p(x) \in K[x]$ such that $p(a) = 0$. We say L is *algebraic* over K, or that the *extension* L/K is *algebraic* if every $a \in L$ is algebraic over K.

Remarks.

(1) "L/K" is pronounced "L over K" and it is not a quotient: it just means $K \subseteq L$.
(2) Whether or not $a \in L$ is *algebraic* depends strongly on K. For example $a = \pi \in L = \mathbb{C}$ is **not** algebraic over $\mathbb{Q}$, but π **is** algebraic over $\mathbb{R}$—it is a root of $x - \pi \in \mathbb{R}[x]$.

[5]Note that we only know that $\dim_{\mathbb{Q}}(\text{RHS}) \leq 100$. We will only be able to check whether equality holds or not later.

Exercise. Show that $\mathbb{C}$ is algebraic over $\mathbb{R}$. Hint: if $z \in \mathbb{C}$, then $(x-z)(x-\overline{z}) \in \mathbb{R}[x]$.

We will now see that if $a \in L$ is algebraic over K, there is a *best* polynomial $p(x) \in K[x]$ such that $p(a) = 0$.

Proposition 2.2 (Existence of the minimal polynomial)**.** *Say $K \subseteq L$ are fields and $a \in L$ is algebraic over K. Then there exists a unique monic irreducible $p(x) \in K[x]$ such that $p(a) = 0$. This $p(x)$ is called the* minimal polynomial *of a over K. Furthermore, if $p(x)$ is the minimal polynomial of a, then for every $f(x) \in K[x]$ such that $f(a) = 0$, we have that $p(x)$ divides $f(x)$.*

Proof. The proof contains one idea. Let d be the smallest degree of all the polynomials $q(x) \in K[x]$, $q(x) \neq 0$, such that $q(a) = 0$. Choose $p(x) \in K[x]$ such that $\deg(p(x)) = d$ and $p(a) = 0$. By scaling ($p(x) \mapsto \lambda p(x)$), we may assume that $p(x)$ is monic.

We claim that $p(x)$ is irreducible. For certainly $p(x)$ is non-constant (as $p \neq 0$ and $p(a) = 0$), and if $p(x) = q(x)r(x)$ with $\deg(q) < d$, $\deg(r) < d$, then $q(a)r(a) = p(a) = 0$ and (because L is a field) $q(a) = 0$ or $r(a) = 0$. This contradicts the minimality of d.

Next, say $f(x) \in K[x]$ and $f(a) = 0$. We write

$$f(x) = q(x)p(x) + r(x)$$

with $\deg(r) < d = \deg(p)$. Then $0 = f(a) = q(a)p(a) + r(a) = r(a)$, so, by definition of d, we must have $r(x) = 0$. Therefore, $f(x) = q(a)p(x)$, a multiple of $p(x)$.

Finally, if $p_1(x)$ is a second monic irreducible polynomial with $p_1(a) = 0$, then $p(x)$ divides $p_1(x)$ by what we just showed, and therefore $p_1(x) = c \cdot p(x)$ for some constant $c \in K$. But they are both monic, so $c = 1$, and hence $p_1(x) = p(x)$. □

Example. Say $\alpha = 2^{1/100} \in \mathbb{R}_{>0}$. Then α is algebraic over $\mathbb{Q}$, since α is a root of $x^{100} - 2 \in \mathbb{Q}[x]$. What is the minimal polynomial of α? Well, $p(x) = x^{100} - 2$ is monic and irreducible over $\mathbb{Q}$ by Eisenstein's criterion 1.8. By Proposition 2.2, $x^{100} - 2$ must be the minimal polynomial of α.

Here then are some important facts about $K(a)$.

Proposition 2.3. *Say $K \subseteq L$ and $a \in L$.*

(a) *If a is algebraic over K, then $K(a)$ is finite-dimensional as a K-vector space, and $\dim_K(K(a)) = d$ is the degree of the minimal polynomial of a, and a K-basis for $K(a)$ is $\{1, a, a^2, \ldots, a^{d-1}\}$.*
(b) *If a is not algebraic over K, then $K(a)$ has infinite dimension as a K-vector space.*[6]

Proof. Part (b) is easy: $K(a) \ni 1, a, a^2, a^3, \ldots$ and a is not algebraic over K, so this is an infinite linearly independent[7] set.

For part (a), let $p(x)$ be the minimal polynomial of a over K. Say $p(x) = x^d + \cdots$ has degree $d \geq 1$. Consider the K-vector subspace V of L spanned by $1, a, a^2, \ldots, a^{d-1}$. Then V is clearly an abelian group under addition and $0, 1 \in V$. Say $v, w \in V$. Write $v = f(a)$,

[6] In fact, $K(a) \cong \mathrm{Frac}(K[x])$.
[7] No **finite** non-trivial linear combination is 0.

$w = g(a)$ for $f(x), g(x) \in K[x]$ with $\deg(f), \deg(g) \leq d-1$. Set $h(x) = f(x)g(x)$ so that $vw = h(a)$. But $h(x) = q(x)p(x) + r(x)$ with $\deg(r) < d$, and therefore $h(a) = 0 + r(a)$, so $vw = r(a) \in V$. Therefore, V is closed under multiplication. Moreover, $\dim_K(V) \leq d$ and we can use Proposition 1.2 to conclude that V is a subfield of L.

Now, $\{1, a, \ldots, a^{d-1}\}$ is a spanning set for $K(a)$ as a vector space over K. For linear independence, note that if $\lambda_i \in K$, not all zero, and

$$\sum_{i=0}^{d-1} \lambda_i a^i = 0,$$

then set $f(x) = \sum_{i=0}^{d-1} \lambda_i x^i$. Then $f(a) = 0$ and, by Proposition 2.2, $f(x)$ is a multiple of $p(x)$, but $f(x) < d$, so $f(x) = 0$, and $\lambda_i = 0$ for all i. □

Remark. There is an *evaluation map* $K[x] \to K(a)$ given by $f(x) \mapsto f(a)$. If a is transcendental, this map is injective and not surjective (for example $1/a$ is not in the image). On the other hand, if a is algebraic, the map is surjective (as $1, a, \ldots, a^{d-1}$ is a basis for $K(a)$) and not injective (for example, if $p(x)$ is the minimal polynomial of a, then $p(x)$ goes to $p(a) = 0$). Note that $K[x]$ and $K(a)$ are both K-vector spaces, and the evaluation map is K-linear. In fact, they are also rings and the evaluation map is a ring homomorphism.

What is the kernel of the map for a algebraic? By definition, it is

$$\{f(x) \in K[x] : f(a) = 0\}.$$

Let $p(x)$ be the minimal polynomial of a. Obviously, if $f(x)$ is a multiple of $p(x)$, say $f(x) = p(x)q(x)$, then $f(a) = p(a)q(a) = 0 \times q(a) = 0$, so $f(x)$ is in the kernel. Conversely, if $f(x)$ is in the kernel, then $f(x)$ is a multiple of $p(x)$ by Proposition 2.2.

Upshot: the kernel of the evaluation map $f(x) \mapsto f(a)$ is the multiples of $p(x)$ in $K[x]$, i.e. the principal ideal generated by $p(x)$.

What just happened? Given the data L, $K \subseteq L$, $a \in L$ algebraic over K, we built the minimal polynomial $p(x) \in K[x]$ and a field $K(a)$ containing K and a.

Question. Say K is a field and $p(x) \in K[x]$ is an irreducible polynomial. Can we build a in some bigger field L, with minimal polynomial $p(x)$, and then build $K(a)$?

Example ($K = \mathbb{Q}$)**.** Let $p(x) = x^{10} + 5x + 5$, which is irreducible by Eisenstein's criterion 1.8. Idea: take $L = \mathbb{C}$. Since $\mathbb{C}$ is algebraically closed, we can find a root $a \in \mathbb{C}$ of $p(x)$. Then a is algebraic over $\mathbb{Q}$ and by Proposition 2.2 there exists a unique monic irreducible polynomial $q(x)$ such that $q(a) = 0$, the minimal polynomial. Hence, it must be $p(x)$. Therefore, $K(a) = \mathbb{Q}(a) \subseteq \mathbb{C}$ and by Propositon 2.3, it has a $\mathbb{Q}$-basis $1, a, \ldots, a^9$.

Example ($K = \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$)**.** Let $p(x) = x^3 + x + 1$. Then $p(0) = 1$, $p(1) = 1$, so $p(x)$ is irreducible (since it has degree 3).

There is no ring homomorphism $\mathbb{Z}/2\mathbb{Z} \to \mathbb{C}$, since $0 \mapsto 0$, $1 \mapsto 1$, so $1 + 1 = 0 \mapsto 1 + 1 = 2$, a contradiction. Therefore, we cannot use the same argument as in the example for $K = \mathbb{Q}$.

Here is what we can do instead. By the remark earlier, if L exists, then there is a map $K[x] \to K(a)$ with kernel $I = \{\text{multiples of } p(x)\}$. Then the first isomorphism theorem

shows that

$$K(a) = K[x]/I.$$

For the right hand side, we only need K and $p(x)$.

Theorem 2.4. *Let K be a field and let $p(x) \in K[x]$ be an irreducible polynomial. Let $I \subseteq K[x]$ be the ideal (or at least a subgroup under addition, if you do not know what an ideal is) consisting of multiples of $p(x)$. Then the quotient $K[x]/I$ is naturally a field containing K. Let us call it L. Let $a \in L$ be the image of the polynomial x (i.e. $a = x + I \in K[x]/I$). Then a is a root of $p(x)$ in L and so, in particular, $p(x)$ is the minimal polynomial of a over K. The K-dimension of L is $d = \deg(p(x))$.*

Proof. Note that $K[x]/I$ is a quotient group, so it is certainly a group under $+$.

Multiplication. Given $f + I$ and $g + I$ in $K[x]/I$, define the product to be $fg + I$. Note: if we change f to $\tilde{f} = f + i$ with $i \in I$, then $f + I = \tilde{f} + I$, but $\tilde{f}g = (f + i)g = fg + ig$. However, ig is a multiple of $p(x)$, so it is in I, and hence

$$\tilde{f}g + I = fg + I.$$

Similarly for a different choice $\tilde{g}$ of g.

It is easy to check that $K[x]/I$ is now a ring (axioms are inherited from $K[x]$).

Next, we claim that $1, a, a^2, \ldots, a^{d-1}$ is a K-basis for $K[x]/I$, where $d = \deg(p(x))$.

Spaning. If $f + I \in K[x]/I$, then $f = qp + r$ with $\deg(r) < d$, and $f + I = r + I$, since $qp \in I$. If $r(x) = \sum_{i=0}^{d-1} \lambda_i x^i$, then $f + I = r + I = \sum_{i=0}^{d-1} \lambda_i a^i$, since $a = x + I$.

Linear independence. If $\sum_{i=0}^{d-1} \mu_i a^i = 0$, then this is the same as saying $\sum_{i=0}^{d-1} \mu_i x^i \in I$, which is a multiple of $p(x)$, so $\mu_i = 0$ for all i.

Finally, $p(a) = p(x) + I = I$, the zero of $K[x]/I$.

We have to show L is a field, i.e. non-zero elements have inverses. We will mimic the proof of Proposition 1.2. Take $0 \leq \lambda \in L$ and define $\varphi_\lambda \colon L \to L$ by $\varphi_\lambda(t) = \lambda t$, a K-linear map. We need to show φ_λ is injective (so it will be surjective by rank-nullity theorem). So suppose that $\varphi_\lambda(t) = 0$ and write $\lambda = f + I$, $t = g + I$, $f, g \in K[x]$. Then $\lambda \neq 0$ if and only if $f + I \neq 0$ if and only if $f \notin I$, i.e. f is not a multiple of p. But $\lambda t = \varphi_\lambda(t) = 0$ means that $fg + I = I$ so fg is a multiple of p. Hence g must be a multiple of p (i.e. $g \in I$), since f is not and p is irreducible. Therefore, $t = 0$ in L, as requested. □

Notation: if K, L, M are fields and $K \subseteq L$, $K \subseteq M$, we write

$$\mathrm{Hom}_K(L, M) = \{f \colon L \to M \text{ a homomorphism of fields such that } f_{|K} = \text{ identity}\}.$$

Good exercise:

$$\mathrm{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C}) = \{\mathrm{id}, \text{complex conjugation}\}.$$

Lemma 2.5. *Let K be a field, $p(x) \in K[x]$ an irreducible polynomial. Then let $K \subseteq L := K[x]/I$ and $a \in L$ is the image of x, i.e. $x + I$. If M is* **any** *field containing a copy of K*

and $b \in M$ is a root of $p(x)$ in M, then there is a unique field map $\alpha\colon L \to M$ such that $\alpha(k) = k$ for all $k \in K$, and $\alpha(a) = b$. This map is an injection. The image of α is the subfield $K(b)$ of M. In particular, if $M = K(b)$, then α is an isomorphism. If $M \supseteq K$ is **any** *field, then there is a natural bijection*

$$\mathrm{Hom}_K(L, M) \longleftrightarrow \{\textit{roots of } p(x) \textit{ in } M\}$$

$$(\beta\colon L \to M) \longmapsto \beta(a) \in M.$$

Proof. Existence. Given $M \ni b$, how to build $\alpha\colon L \to M$? Consider the evaluation map $K[x] \to M$ given by $f(x) \mapsto f(b)$. What is the kernel? If b is a root of $p(x)$, then $p(x)$ is in the kernel, so $I \subseteq \ker(\alpha)$, and therefore we get an induced map $L = K[x]/I \to M$.

Uniqueness. Suppose α' is a K-map $L \to M$ such that $\alpha'(a) = b$. Since α' is a field map, $\alpha'(a^i) = b^i$ for all $0 \leq i < d$, and $\alpha'(k) = k$ for all $j \in K$. Therefore, α' is a K-linear map that agrees with α on a K-basis; hence $\alpha' = \alpha$.

Injectivity follows from the general fact: if $\varphi\colon L \to M$ is a map of fields, then φ is injective. Indeed, if $0 \neq \lambda$, $\varphi(\lambda) = 0$, then $\varphi(1) = \varphi(\lambda)\varphi(1/\lambda) = 0 \cdot \lambda(1/\lambda) = 0$. But $\varphi(1) = 1 \neq 0$, a contradiction.

Image of α. Think of L and M as K-vector spaces. Then L is spanned by $1, a, a^2, \ldots, a^{d-1}$, so $\operatorname{im}\alpha$ is spanned by $1, b, b^2, \ldots, b^{d-1}$, a K-basis for $K(b) \subseteq M$ by Proposition 2.3. In particular, if $M = K(b)$, then α is injective and surjective.

Finally, let us write down maps

$$\mathrm{Hom}_K(L, M) \longleftrightarrow \{\text{roots of } p(x) \text{ in } M\}$$

both ways and check they are inverses to each other. We have already seen 'left to right': send $\beta\colon L \to M$ to $\beta(a)$. This works because $p(a) = 0$ in L, so $\beta(p(a)) = p(\beta(a)) = 0$ in M. (Note that $\beta \circ p = p \circ \beta$ because β is a field map and p is a polynomial.) 'Right to left': we have just done this! Send b to α. Easy check: these are inverse bijections. □

Say $K \subseteq L$ are fields. Then L becomes a vector space over K (check the axioms). Now say $K \subseteq L \subseteq M$ are all fields. What is the relationship between $[M : K]$, $[M : L]$, and $[L : K]$?

Theorem 2.6 (Tower law)**.** *If $K \subseteq L \subseteq M$, then $[M : K] = [M : L][L : K]$.*

Proof. This is true even when some dimensions are infinite, but we will just show that if $[M : L] = m < \infty$ and $[L : K] = n < \infty$, then $[M : K] = mn < \infty$. The infinite case can be found on Problem Sheet 2, Exercise 7.

In the finite case, let $\{e_1, e_2, \ldots, e_m\}$ be an L-basis for M and $\{f_1, f_2, \ldots, f_n\}$ be a K-basis for L. The idea of the proof is to set $g_{ij} = e_i f_j$ for $1 \leq i \leq m$, $1 \leq j \leq n$. We claim that g_{ij} are a basis for M as a K-vector space.

Spanning. Take any $v \in M$. By definition of $\{e_i\}$, we can write

$$v = \sum_{i=1}^{m} \lambda_i e_i \qquad \text{for } \lambda_i \in L.$$

By definition of $\{f_j\}$, each $\lambda_i \in L$ can be written

$$\lambda_i = \sum_{j=1}^{n} \mu_{ij} f_j \qquad \text{for } \mu_{ij} \in K.$$

Putting this together, we obtain

$$v = \sum_{i,j} \mu_{ij} e_i f_j = \sum_{i,j} \mu_{ij} g_{ij} \qquad \text{for } \mu_{ij} \in K.$$

Linear independence. Say we have $\mu_{ij} \in K$ such that

$$\sum_{i,j} \mu_{ij} g_{ij} = 0.$$

Expand out to get:

$$\sum_{i=1}^{m} \underbrace{\left(\sum_{j=1}^{n} \mu_{ij} f_j \right)}_{\lambda_i \in L} e_i = 0.$$

By linear independence of $\{e_i\}$, we obtain

$$\sum_{j=1}^{m} \mu_{ij} f_j = \lambda_i = 0,$$

so by linear independence of $\{f_j\}$, we obtain $\mu_{ij} = 0$ for all i and j. □

Example. We have that $[\mathbb{Q}(\sqrt{2} : \mathbb{Q}] = 2$, since the minimal polynomial of $\sqrt{2}$ over $\mathbb{Q}$ is $X^2 - 2$. Also, $i \notin \mathbb{Q}(\sqrt{2})$ as $i \notin \mathbb{R}$, $i = \sqrt{-1}$, so its minimal polynomial over $\mathbb{Q}(\sqrt{2})$ is $X^2 + 1$. Therefore,

$$\mathbb{Q}(\sqrt{2}, i) = \mathbb{Q}(\sqrt{2})(i)$$

has dimension 2 over $\mathbb{Q}(\sqrt{2})$, and the Tower law 2.6 tells us

$$[\mathbb{Q}(\sqrt{2}, i) : \mathbb{Q}] = 4.$$

In fact, the proof of the Tower law tells us that $\{1, \sqrt{2}, i, i\sqrt{2}\}$ is a $\mathbb{Q}$-basis for $\mathbb{Q}(\sqrt{2}, i)$.

Reminder: If $K \subseteq L$, then we say the extension L/K is *algebraic* if every $\lambda \in L$ is algebraic over K.

A priori, checking that L/K is algebraic seems hard. For instance, to check that $\mathbb{Q}(\sqrt{2}, i)$ is algebraic over $\mathbb{Q}$, we would have to check that $3i - \sqrt{2} + 7\sqrt{2}i$ is a root of a polynomial with coefficients in $\mathbb{Q}$. This is far from immediate, but fortunately, we can give a (non-constructive) proof.

Corollary 2.7.

(1) *If $K \subseteq L$ are fields and $[L : K] < \infty$, then L/K is algebraic.*
(2) *If $K \subseteq L$ are fields and $a_1, a_2, \ldots, a_n \in L$ are all algebraic over K, then*

$$[K(a_1, a_2, \ldots, a_n) : K]$$

is finite.

(3) *If $K \subseteq L$ are fields and $a, b \in L$ are algebraic over K, then so are $a+b$, $a-b$, ab, and, if $b \neq 0$, a/b.*

Proof. For (1), say $\lambda \in L$. Then $K \subseteq K(\lambda) \subseteq L$. By assumption $\dim_K L < \infty$, hence $\dim_K(K(\lambda)) \leq [L : K] < \infty$ is also finite. Hence λ is algebraic by Proposition 2.2.

For (2), recall that by Proposition 2.2 we know that $[K(a_1) : K] < \infty$. Note that a_2 is algebraic over K by definition, so a_2 is algebraic over the larger field $K(a_1)$. Again, by Proposition 2.2 we know that $[K(a_1)(a_2) : K(a_1)] < \infty$, so by the Tower law 2.6, we have

$$[K(a_1, a_2) : K] < \infty.$$

We continue by induction on n to prove (2).

For (3), we know that a is algebraic over K, so $[K(a) : K] < \infty$, and b is algebraic over K, so it is algebraic over $K(a)$, and hence $[K(a, b) : K(a)] < \infty$. Now, by the Tower law 2.6, we get that

$$[K(a, b) : K] < \infty,$$

and by (1), $K(a, b)$ is algebraic over K. But $a + b$, $a - b$, ab, and, if $b \neq 0$, a/b are all in $K(a, b)$, so they are algebraic over K. □

Example. Note that $\sqrt[5]{7}$ and $\sqrt[5]{11}$ both have degree 5 over $\mathbb{Q}$ (to prove this, use Eisenstein's criterion 1.8). Hence $\sqrt[5]{7}$ and $\sqrt[5]{11}$ is algebraic over $\mathbb{Q}$, and one can check that the degree of $\sqrt[5]{7}+\sqrt[5]{11}$ over $\mathbb{Q}$ is at most 25. Can you find its minimal polynomial? Probably not without a computer.

Corollary 2.8. *If K is a field and $p(x) \in K[x]$ is a polynomial, then there exists a finite extension L/K (i.e. $K \subseteq L$, L a field, and $[L : K]$ finite) such that $p(x)$ factors into linear factors over L.*

Proof. Use the abstract root adjoining method (Lemma 2.5) to throw in one root of one irreducible factor of $p(x)$. This extension is finite over K by Corollary 2.7, so we can finish the proof by induction on the degree of $p(x)$ and using the Tower law 2.6. □

Proposition 2.9. *If $K \subseteq L \subseteq M$ and both L/K and M/L are algebraic, then M/K is algebraic.*

Remark. In Problem Sheet 2, Exercise 10, we see that if $\overline{\mathbb{Q}} := \{\lambda \in \mathbb{C} : \lambda \text{ is algebraic over } \mathbb{Q}\}$, then $\overline{\mathbb{Q}}$ is a subfield of $\mathbb{C}$, and $[\overline{\mathbb{Q}} : \mathbb{Q}]$ is infinite. In particular, you can be algebraic and infinite, and Tower law 2.6 does not imply Proposition 2.9.

Proof of Proposition 2.9. Suppose $m \in M$. Since M/L is algebraic there exists $0 \neq p(x) \in L[x]$ such that $p(m) = 0$. Write $p(x) = \sum_{i=0}^{d} \lambda_i x^i$, $\lambda_i \in L$. The trick is to note that

$$[K(\lambda_0, \lambda_1, \ldots, \lambda_d) : K] < \infty$$

by Corollary 2.7 (b). Moreover, m is algebraic over $N = K(\lambda_0, \lambda_1, \ldots, \lambda_d)$, since $p(x) \in N[x]$, so

$$[K(\lambda_0, \ldots, \lambda_d, m) : N] < \infty$$

and $[N : K] < \infty$, so by the Tower law 2.6

$$[K(\lambda_0, \ldots, \lambda_m, m) : K] < \infty$$

and hence m is algebraic over K. □

3. Straightedge and compass construction

The ancients were interested in numbers as lengths, and operations involving numbers as constructions.

Tools:

(1) long straight object (could use it to draw a line between two points),
(2) pair of compasses (given points A and B in a plane, they could draw a circle centered at A going through B).

Can make new points from old by looking at where lines and circles intersect.

Example: They could drop a perpendicular, take square roots, divide a length by any natural number!

To get a feeling for what you can do, go to `euclidthegame.com`.

Things the Greeks could not do:

(1) Trisect an angle (e.g. construct $\cos(20^\circ)$).
(2) Duplicate the cube (i.e. construct $\sqrt[3]{2}$).
(3) Square the circle (i.e. construct π).

In fact, none of these things are possible with ruler and compasses!

Rules: If $S \subseteq \mathbb{R}^2$ is a finite set of points, we say $p \in \mathbb{R}^2$ is *constructible in one step from* S if we can draw a finite set of circles, with center $s_0 \in S$ and radius equal to $\mathrm{dist}(s_1, s_2)$, $s_1, s_2 \in S$, and lines going through two distinct points of S, such that p is the intersection of two of these.

We say $p \in \mathbb{R}^2$ is *constructible from* S, if there exist $p_1, p_2, \ldots, p_n = p$ all in $\mathbb{R}^2$ such that p_{i+1} is constructible in one step from $S \cup \{p_1, p_2, \ldots, p_i\}$.

Typically, we start with

$$S = \{(0, 0), (1, 0)\}.$$

We can certainly then construct $(q, 0)$ for all $q \in \mathbb{Q}$ and $(\cos 60^\circ, \sin 60^\circ)$ etc.

The following lemma is the key to this whole chapter.

Lemma 3.1. *Suppose* $S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ *and* $p = (\alpha, \beta)$ *is constructible in one step from* S. *Let* $K = \mathbb{Q}(x_1, y_1, \ldots, x_n, y_n)$. *Then* $[K(\alpha) : K] \leq 2$ *and* $[K(\beta) : K] \leq 2$.

Proof. Recall that p is either the intersection of two lines through points in S, or the intersection of a line and a circle, or the intersection of two circles.

Case 1. Two lines. Lines will be defined by equations of the form $\lambda x + \mu y = 1$ and lines go through two points in S. Then $\lambda, \mu \in K$. Hence the intersection of two such lines is (α, β) with $\alpha, \beta \in K$.

Case 2. A line and a circle. Substitute the equation for a line into the equation for a circle and end up proving that α is a root of a quadratic equation with coefficients in K. If it factors, $\alpha \in K$. If it does not, $[K(\alpha) : K] = 2$ by Proposition 2.3.

Case 3. Two circles. They are of the form

(1) $x^2 + y^2 + ax + by + c = 0$, $a, b, c \in K$,
(2) $x^2 + y^2 + dx + ey + f = 0$, $d, e, f \in K$.

The miracle is that (1)−(2) is linear, so solving (1) and (1)−(2) at the same time is the previous case.

In all cases, $[K(\alpha) : K]$ is 1 or 2. Same for β. □

Corollary 3.2. *If $p = (\alpha, \beta)$ is constructible from $S = \{(x_1, y_1), \ldots, (x_n, y_n)\}$ and $K = \mathbb{Q}(x_1, y_1, \ldots, x_n, y_n)$, then $[K(\alpha, \beta) : K]$ is a power of 2.*

Proof. Use Lemma 3.1, the Tower law 2.6, and observe that if $[K(\beta) : K] = 2$ then

$$[K(\alpha, \beta) : K(\alpha)] \leq 2$$

to obtain the result. □

Corollary 3.3. *If $S_0 = \{(0, 0), (1, 0)\}$, then we cannot construct any of*

$$(\cos(20^\circ), 0), (\sqrt[3]{2}, 0), (\pi, 0).$$

Proof. We know that $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, so $\mathbb{Q}(\sqrt[3]{2})$ cannot be contained in any field L such that $[L : \mathbb{Q}]$ is a power of 2, or else it contradicts the Tower law 2.6.

Since π is transcendental, $[\mathbb{Q}(\pi), \mathbb{Q}] = \infty$.

Finally, for $a = \cos 20^\circ$, a is algebraic of degree 3 (to see this, note that $\cos(3\theta)$ is a cubic in $\cos(\theta)$ and $\cos(60^\circ) = \frac{1}{2}$), so it cannot be constructed. □

4. Splitting fields

Let E be a field, $p(x) \in E[x]$ an irreducible polynomial of degree $d > 1$. Then $p(x)$ has no roots in E. In Chapter 2, we saw two methods for adding a root.

Method 1. Suppose $E \subseteq L$ and $p(x)$ has a root $a \in L$. Set $M = E(a)$ which contains E and a. Then p has a root in M.

Problem: For $E \subseteq L'$ and p has a root $a' \in L'$, set $M' = E(a') \supseteq L'$. Is $M \cong M'$? This is not clear at the beginning of Chapter 2, but is resolved by Method 2.

Method 2. Let $I = (p(x))$, the ideal of multiples of $p(x)$. Then

$$E[x]/I$$

is a new field with a root of p. Then Theorem 2.4 and Lemma 2.5 imply that

$$M \cong E[x]/I \cong M'.$$

Example. Let $E = \mathbb{Q}$, $L = L' = \mathbb{C}$, and $p(x) = x^3 - 2$. We can set $a = \sqrt[3]{2} \in \mathbb{R} \subseteq \mathbb{C} = L$, and $a' = e^{2\pi i/3}a \in \mathbb{C} \setminus \mathbb{R}$. They are both roots of $p(x)$, whence

$$M = \mathbb{Q}(a) \subseteq \mathbb{R};\ M' = \mathbb{Q}(a') \not\subseteq \mathbb{R}.$$

Note that $M \neq M'$ but $M \cong M'$ via $1 \mapsto 1$, $a \mapsto a'$, $a^2 \mapsto a'^2$, since they are both isomorphic to $\mathbb{Q}[x]/(p(x))$.

In this chapter, we take a field E and any non-zero polynomial $p(x) \in E(x)$ of degree d, and we want to construct a bigger field $K \supseteq E$ such that *all* roots of $p(x)$ are in K. If roots of $p(x)$ in K are $a_1, a_2, \ldots, a_d$, we will be interested in the field

$$E(a_1, \ldots, a_d) = \text{smallest subfield of } K \text{ containing } E \text{ and all the roots of } p(x).$$

Technical issue: If $p(x) \in E[x]$ and $E \subseteq L$, $E \subseteq L'$, and p has roots $a_1, \ldots, a_d \in L$, $a'_1, \ldots, a'_d \in L'$, then we would hope that

$$L \supseteq E(a_1, \ldots, a_d) \cong E(a'_1, \ldots, a'_d) \subseteq L'.$$

The general idea of the proof is induction on the number of roots we have added so far. The problem is that the naive approach does not work.

Chapter 2 is all about **irreducible** polynomials. We proved that

$$E(a) \cong E(a') \text{ if } a, a' \text{ are roots of an irreducible polynomial } p.$$

If p is reducible, the technique fails. For example, take $E = \mathbb{Q}$, $p(x) = (x^3 - 2)(x^3 - 3)$, a reducible polynomial, and $L = L' = \mathbb{C}$. Then $a = \sqrt{2}$, $a' = \sqrt{3}$ are both roots of p. However

$$\sqrt{3} \notin \mathbb{Q}(a) \ncong \mathbb{Q}(a') \ni \sqrt{3}.$$

The problem is that the quotient $E[x]/(p(x))$ is not a field if p is reducible!

Terrifying thing: Let $p(x) \in E[x]$ be an irreducible polynomial and $E \subseteq L$ with $a \in L$ a root of $p(x)$. Then $K = E(a) \supseteq E$. Now, consider $p(x) \in K[x]$. It is clearly reducible, because $a \in K$ is a root of p, so $p(x) = (x - a)q(x)$ with $\deg(q(x)) = d - 1$ in $K[x]$.
Question: Can we prove that $q(x)$ is irreducible?
Answer: No, because it might not be!
Example where it is true: $E = \mathbb{Q}$, $p(x) = x^3 - 2$, $a = \sqrt[3]{2}$, $K = E(a)$. Then $p(x) = (x-a)q(x) \in K[x]$, where $\deg(q(x)) = 2$. Since $q(x)$ has degree 2, it is irreducible if and only if $q(x)$ has no roots. Does $q(x)$ have its roots in $K = \mathbb{Q}(a)$? No, because the roots ($e^{2\pi i/3}a$ and $e^{4\pi i/3}a$) of $q(x)$ are not in $\mathbb{R}$, and $K \subseteq \mathbb{R}$.
Example where it is not true: $E = \mathbb{Q}$, $p(x) = x^4 + x^3 + x^2 + x + 1 = \frac{x^5-1}{x-1}$, which is irreducible over $\mathbb{Q}$ by Eisenstein's criterion 1.8 (using the trick). Set $\zeta = e^{2\pi i/5} \neq 1$, $\zeta^5 = 1$, a root of $p(x)$. Set $K = \mathbb{Q}(\zeta)$. In $K[x]$, $p(x) = (x - \zeta)q(x)$ with $\deg(q(x)) = 3$. Note that $\overline{\zeta} \in K = \mathbb{Q}(\zeta)$, since $\overline{\zeta} = \zeta^4$. In fact, the roots of $q(x)$ are $\zeta^2, \zeta^3, \zeta^4 \in K = \mathbb{Q}(\zeta)$. Therefore, $q(x)$ factors into three linear factors in K.

Definition. Suppose E is a field and $f(x) \in E[x]$ is a non-zero polynomial. We say $f(x)$ *splits completely* in E, or (for simplicity) $f(x)$ *splits* in E if $f(x)$ factors into linear factors in $E[x]$, i.e.

$$f(x) = c(x - a_1)(x - a_2)\dots(x - a_d) \text{ with } a_i \in E.$$

We say $L \supseteq E$ is a *splitting field* for $f(x)$ over E if $f(x)$ splits in L,

$$f(x) = c(x - a_1)(x - a_2)\dots(x - a_d) \text{ with } a_i \in L,$$

and furthermore $L = E(a_1, \dots, a_d)$.

For example, $f(x) = x^2 + 1 \in \mathbb{Q}$. Then $f(x)$ splits in $\mathbb{C}$ but $\mathbb{C}$ is not a splitting field for $f(x)$ over $\mathbb{Q}$. It is too big; the splitting field is $\mathbb{Q}(i)$.

To prove uniqueness of splitting fields up to isomorphism, we introduce a temporary definition.[8]

Temporary definition. Now, say E is a field, $0 \neq f(x) \in E[x]$, and $E \subseteq L$, a field. We say L has *property* $(*)$ for the pair $(E, f(x))$ if it has the following property: If $E \subseteq K$, a field, then $f(x)$ splits completely in K if and only if $\mathrm{Hom}_E(L, K)$ is non-empty.

Recall that $\mathrm{Hom}_E(L, K)$ are field homomorphisms $L \to K$ which are identity on E.

Remark. Note that if L has property $(*)$, then $\mathrm{Hom}_E(L, L)$ contains the identity map, so $f(x)$ splits completely in L.

Lemma 4.1. *If E is a field and $0 \neq f(x) \in E[x]$, then there exists $L \supseteq E$ such that L has property $(*)$ for $(E, f(x))$, and $[L : E] < \infty$.*

Proof. Induction on $\deg(f)$. If $\deg(f)$ is 0 or 1, then f splits completely in E, and thus $L = E$ works.

Inductive step. Let $p(x)$ be an irreducible factor of $f(x)$ in $E[x]$. By Chapter 2, we can add one root of $p(x)$ to E and get $F = E[x]/(p(x))$. Now, F contains a root of p, so also a root of f, and we can write

$$f(x) = (x - \lambda)g(x), \text{ where } g(x) \in F[x] \text{ and } \deg(g) < \deg(f).$$

The inductive hypothesis implies that there exists a field $L \supseteq F$ such that L has property $(*)$ for (F, g), and $[L : F] < \infty$.

We claim that L has property $(*)$ for (E, f) and $[L : E] < \infty$. Note first that the Tower law 2.6 implies that $[L : E] = [L : F][F : E] = [L : F]\deg(p) < \infty$. We still have to show that L has property $(*)$ for (E, f). First, note that because L has property $(*)$ for (F, g), by an earlier remark, $g(x)$ splits in L. Note also $\lambda \in F \subseteq L$, and therefore $f(x)$ splits in L.

Now, check $(*)$. If $K \supseteq E$ and $\mathrm{Hom}_E(L, K) \neq 0$, then take $\varphi\colon L \to K$, and if $a_1, \dots, a_d$ are roots of f in L, then $\varphi(a_1), \dots, \varphi(a_d)$ are the roots of $\varphi(f)$ in K. But $\varphi_{|E} = \mathrm{id}$, so $\varphi(f) = f$, and hence f splits in K. Conversely, say f splits in K. We need a field map $L \to K$, which is the identity on E. Since f splits in K, p splits in K, so p has a root $\alpha \in K$. Therefore, there exists a field map $F \to E(\alpha) \subseteq K$, which is identity on E, by Lemma 2.5. By $(*)$ for L

[8] One can prove uniqueness of splitting fields up to isomorphism without this temporary definition. See, for example, *Abstract Algebra* by D. Dummit and R. Foote.

for the pair (F, g) this map $F \to K$ extends to a map $L \to K$ which is the identity on F, so it is the identity on E. □

Lemma 4.2. *Suppose E is a field and $0 \neq f \in E[x]$. If $L \supseteq E$ satisfies $(*)$ for (E, f), and $L' \supseteq E$ satisfies $(*)$ for (E, f), and if $[L : E]$, $[L' : E] < \infty$, then $L \cong L'$ via an isomorphism which is the identity on E.*

Proof. Since L satisfies $(*)$, f splits in L, and since L' satisfies $(*)$, there exists $\varphi\colon L' \to L$ which is the identity on E. All field maps are injections, so $\dim_E L \geq \dim_E L'$. By symmetry, $\dim_E L' \geq \dim_E L$. Therefore, $[L' : E] = [L : E]$. Now, φ is an injection between two vector spaces of the same dimension, and hence it must be an isomorphism. □

Theorem 4.3. *Suppose E is a field and $0 \neq f \in E[x]$. Then the following are equivalent for an extension $L \supseteq E$:*

(1) *L is a splitting field for f over E,*
(2) *L satisfies property $(*)$ for (E, f).*

Proof. Suppose L satisfies (2). Then $f(x) = c\prod_{i=1}^{d}(x - a_i)$ with $a_i, c \in L$ by the remark. Therefore, $L \supseteq K := E(a_1, \ldots, a_d)$, where $d = \deg(f)$. Therefore, $[L : E] \geq [K : E]$. Since f splits in K, property $(*)$ implies that there exists $\varphi\colon L \to K$, injective, so $[K : E] \geq [L : E]$. Hence $[K : E] = [L : E]$, so $L = E(a_1, \ldots, a_d)$, and hence L is the splitting field for f over E.

Conversely, suppose (1): L is a splitting field for f over E, i.e. $L = E(a_1, \ldots, a_d)$, where a_i are the roots of f in L. By Lemma 4.1, there exists a field M satisfying property $(*)$ for (E, f) and $[M : F] < \infty$. Therefore, there exists a map $\varphi\colon M \to L$ which is the identity on E. Clearly, φ is injective (all field maps are injective). But f splits in M, so φ maps the roots of f in M to the roots of f in L. Therefore, $\operatorname{im}\varphi$ contains E and the roots $a_1, \ldots, a_d$. The image of φ is a field containing $E(a_1, \ldots, a_d) = L$, so it is equal to L, and φ is surjective. Since M has property $(*)$, $L \cong M$ also has property $(*)$. □

Corollary 4.4. *Splitting fields are unique up to isomorphism.*

Proof. Use Lemma 4.2 and Theorem 4.3. □

Why are we talking about splitting fields? Galois theory is about permuting the roots of polynomials, so we need all the roots.

Algebraic closures. We could go on now to build algebraic closure, but we need too much algebra (existence of maximal ideals, Zorn's lemma, etc.) The theorem we would get to would be.

Theorem. *If E is a field, then there exists an extension $E \hookrightarrow \overline{E}$ such that*

(1) *$\overline{E}$ is algebraic over E,*
(2) *$\overline{E}$ is algebraically closed (i.e. any polynomial of degree greater or equal to 1 with coefficients in $\overline{E}$ factors into linear factors in $\overline{E}$).*

More intuitively: *The field $\overline{E}$ is the smallest algebraically closed field containing E.*

The proof is a long exercise in algebra (which one could do if one was taking the Algebra III course).

Special easier case: if $E \subseteq \mathbb{C}$, we can set $\overline{E} = \{z \in \mathbb{C} : z \text{ is algebraic over } E\}$. Exercise: such a definition works.

Remark. The algebraic closure $\overline{E}$ is unique up to isomorphism, i.e. if L_1 and L_2 are algebraic closures of E, then $L_1 \cong L_2$ via an isomorphism which is the identity on E.

Definition. Suppose E is a field. We say that an algebraic extension $F \supseteq E$ is *normal* over E (alternatively, F/E is *normal*) if it has the following property:
If $p(x) \in E[x]$ is irreducible over E and has a root in F, then it splits completely in F.

Non-example: Let $E = \mathbb{Q}$, $F = \mathbb{Q}(\sqrt[3]{2})$. Then F/E is not normal, because $p(x) = x^3 - 2$ is irreducible over E by Eisenstein and has a root in F, but $F \hookrightarrow \mathbb{R}$ and the other two roots of $p(x)$ are not real, so $p(x)$ does not split completely.

Rubbish example: E a field, $\overline{E}$ its algebraic closure, then $\overline{E}/E$ is normal. For example, $\mathbb{C}/\mathbb{R}$ is normal.

What about $E = \mathbb{Q}$? Clearly, $F = \mathbb{Q}$ is normal over E. Can we think of any other example with $[F : \mathbb{Q}]$ finite? (Note that $[\overline{\mathbb{Q}} : \mathbb{Q}] = \infty$.)

Say $p(x) \in \mathbb{Q}[x]$ is irreducible and F is its splitting field. How can we check that every irreducible $q(x) \in \mathbb{Q}[x]$ with a root in F splits completely in F?

Theorem 4.5. *If F/E is a finite extension, then the following are equivalent*

(1) *F/E is normal,*
(2) *there exists $0 \neq f(x) \in E[x]$ such that F is a splitting field for f over E.*

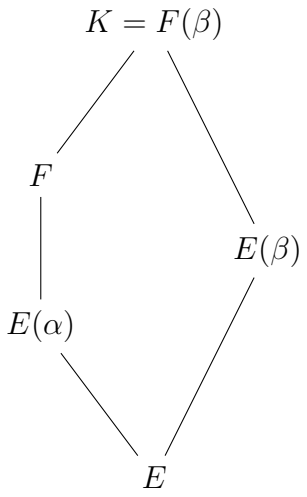
Proof. First, we prove that (1) implies (2). Since F/E is finite, choose an E-basis $a_1, a_2, \ldots, a_d$ as an E-vector space. Clearly, $F = E(a_1, \ldots, a_d)$. Let p_i be the minimal polynomial of a_i over E, which exists because F/E is finite and hence algebraic (Corollary 2.7). Then $p_i \in E[x]$ is irreducible over E, $a_i \in F$, so since F/E is normal, p_i splits completely in F. Set $f(x) = \prod_{i=1}^{d} p_i(x)$. We have seen that $f(x)$ splits completely in F. Moreover:

$$F \supseteq E(\text{all roots of } f) \supseteq E(a_1, \ldots, a_d) = F,$$

and thus F is the splitting field of f over E.

Now, we prove that (2) implies (1). Let F be a splitting field for $f(x)$ over E. Say $p(x) \in E[x]$ is irreducible and has a root $\alpha \in F$. We know that $p(x)$ factors into irreducibles in $F[x]$, so let $q(x)$ be one of these irreducibles in $F[x]$. We want to show that $\deg q = 1$. By Chapter 2, there exists an extension $K = F(\beta)$ over F, where $q(x)$ has a root β.

We have the following diagram of field extensions.



Easy checks:

(1) $E(\alpha) \cong E(\beta)$ by Chapter 2, since they are both isomorphic to $E[x]/(p(x))$
(2) F is the splitting field of f over $E(\alpha)$
(3) $F(\beta)$ is the splitting field of f over $E(\beta)$

By Corollary 4.4, $[F(\beta) : E(\beta)] = [F : E(\alpha)]$. Also, $[E(\beta) : E] = \deg p = [E(\alpha) : E]$. Tower law 2.6 implies that $[F(\beta) : E] = [F : E]$. But $F \subseteq F(\beta)$, and hence $F = F(\beta)$ (subspace has same dimension as the whole space, so the subspace is the whole space). Thus $\beta \in F$, so $\deg(q) = 1$. □

Lemma 4.6. *Say $K \supseteq F \supseteq E$ and $[K : E]$, $[K : F]$, $[F : E]$ all finite. If K/E is normal, then K/F is normal.*

Proof. Suppose $p(x) \in F[x]$ is irreducible and has a root $\alpha \in K$. Set $q(x)$ to be the minimal polynomial for α over E. We know that K/E is normal, $q(x)$ is irreducible over E, and $q(x)$ has a root $\alpha \in K$, so q splits completely in K. But α is a root of q, so p divides q in $F[x]$, because it is the minimal polynomial. And hence $p(x)$ splits completely. □

What about the other possibilities. Here, we have counter-examples!

(1) $\mathbb{Q} \subset \mathbb{Q}(\sqrt[3]{2}) \subset$ (splitting field of $x^3 - 2$ over $\mathbb{Q}$), and $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ is not normal.
(2) $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\sqrt[4]{2})$ and $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$, $\mathbb{Q}(\sqrt[4]{2}/\mathbb{Q}(\sqrt{2})$ are normal, but $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}$ is not normal.

5. Separable extensions

Galois theory is about how automorphisms of fields can permute roots of polynomials. For it to work, we need

(1) polynomials have all their roots in extension field (*normal* extension),
(2) irreducible polynomials do not have repeated roots (*separable* extension).

Definition. Let E be a field.

(1) An irreducible polynomial $p(x) \in E[x]$ is called *separable* if the roots of $p(x)$ in a splitting field $K \supseteq E$ of $p(x)$ are distinct.
(2) We say $0 \neq f(x) \in E[x]$ is *separable* if all its irreducible factors are separable.
(3) If F/E is an extension of fields and $\alpha \in F$ is algebraic over E, then α is *separable* over E if the minimal polynomial of α over E is separable.
(4) If F/E is an algebraic extension of fields, then F is *separable* over E if for all $\alpha \in F$, α is separable over E.

Remark. Ad (1). A polynomial $p(x)$ is separable if and only if $p(x)$ has distinct roots in **any** extension $K \supseteq E$ where $p(x)$ splits completely (replace K by subfield generated over E by roots of $p(x)$).

Example. Let $E = \mathbb{Q}$ and $p(x) = x^3 - 2$, which is irreducible over $\mathbb{Q}$, and it is separable, because the roots in $\mathbb{C}$ are $\alpha = \sqrt[3]{2}$, $\omega\alpha$, $\omega^2\alpha$, where $\omega = e^{2\pi i/3}$, all distinct.

Let us try to think of a non-example. Note that the polynomial x^2 is also separable over $E[x]$, because both factors are (every degree 1 polynomial is separable).

Let us assume that my field E is a subfield of $\mathbb{C}$. We can use calculus to prove that every $0 \neq f(x) \in E[x]$ is separable over E. Suppose for a contradiction that there exists $0 \neq p(x) \in E[x]$ that is **not** separable. Without loss of generality, $p(x)$ is irreducible (because an irreducible factor of it is not separable, so we can take that factor instead). Since $p(x)$ has a repeated root $\alpha \in \mathbb{C}$, write $p(x) = (x - \alpha)^2 q(x)$. We can differentiate $p(x)$ to get $p'(x) \in E[x]$ that still has the root α. Set $h(x) = \gcd(p(x), p'(x))$, which can be computed using Euclid's algorithm, and clearly $h(x) \in E[x]$. By definition, $h(x) \neq 0$, and α is a root of $h(x)$. Therefore, $h(x)$ is non-constant, and $h(x)$ divides $p(x)$, which is irreducible. Therefore, $h(x) = c \cdot p(x)$ for a non-zero constant $c \in E$. But $h(x)$ divides $p'(x)$, which contradicts $0 < \deg(p'(x)) < \deg(p(x)) = \deg(h(x))$.

As a consequence: If $E \subseteq F \subseteq \mathbb{C}$ and F/E is algebraic, then F/E is separable.

How much calculus did we actually use here? We will come back to this soon.

For now, say $E \subseteq F \subseteq K$ and some of these are separable. Are others?

Lemma 5.1. *If K/E is separable, then so are K/F and F/E.*

Proof. The separability of F/E is obvious. For separability of K/F, suppose $\alpha \in K$ and let $p(x)$ be the minimal polynomial of α over E, and let $q(x)$ be the minimal polynomial of α over F. Let L be the splitting field of $p(x)$ over K. Since K/E is separable, $p(x)$ has distinct roots in L. Now, $q(x)$ is the minimal polynomial of α over F, and $p(x) \in E[x] \subseteq F[x]$ with $p(\alpha) = 0$. Therefore, $q(x)$ divides $p(x)$ in $F[x]$ by the definition of a minimal polynomial. Hence $q(x)$ has distinct roots in L, which shows K/F is separable. □

The converse is true, but harder to prove: if K/F and F/E is separable, then so is K/E. The proof will be in the next chapter.

Let us go back to calculus. Let us develop enough calculus over an **arbitrary** field to try and understand separability better. We cannot hope to *differentiate* a general function $E \to E$ (even *continuity* makes no sense). However, we **can** differentiate polynomials.

Definition. Let E be any field. Define the *derivative map* $D\colon E[x] \to E[x]$ on the basis by $D(x^n) = nx^{n-1}$ and extend it E-linearly to $E[x]$. (Note that $n \in \mathbb{Z}$, but there is a homomorphism $\theta\colon \mathbb{Z} \to E$, so by n we really mean $\theta(n)$.) Explicitly,

$$D\left(\sum_{i=0}^{d} a_i x^i\right) = \sum_{i=1}^{d} i a_i x^{i-1}.$$

Remark. If $E = \mathbb{Z}/p\mathbb{Z}$ or, more generally, any field of characteristic p, then

$$D(x^p) = px^{p-1} = 0.$$

Lemma 5.2 (Product rule)**.** *For $f, g \in E[x]$, we have that $D(fg) = fD(g) + gD(f)$.*

Proof. Let us first check it in the special case $f(x) = x^m$, $g(x) = x^n$. We have

$$\text{LHS} = D(x^{m+n}) = (m+n)x^{m+n-1},$$

$$\text{RHS} = x^m D(x^n) + x^n D(x^m) = nx^{m+n-1} + mx^{m+n-1} = \text{LHS}.$$

Next, let us consider the case when $f \in E[x]$ is arbitrary and $g(x) = x^n$. We can think of f as a variable. Both sides are linear maps $E[x] \to E[x]$ and they agree on a basis, so they are the same.

Finally, when f, g are arbitrary, think of f as fixed, g as a variable. Both sides are linear maps and agree on a basis. Therefore, they are equal. □

(If you do not like the above argument, check the general case $f(x) = \sum a_i x^i$, $g(x) = \sum b_j x^j$.)

Corollary 5.3. *We have that $D\left((x-a)^n\right) = n(x-a)^{n-1}$ if $n \in \mathbb{Z}_{\geq 1}$, $a \in E$.*

Proof. We either use induction on n and Lemma 5.2, or work it out using the binomial theorem. □

What happens if we mimic the proof that if $E \subseteq \mathbb{C}$, then every $0 \neq f(x) \in E[x]$ is separable?

Proposition 5.4. *Let E be a field and $0 \neq f(x) \in E[x]$. Suppose $E \subseteq L$ and f splits completely in L. Then the following are equivalent:*

(1) *f has a repeated root in L,*
(2) *there exists $\alpha \in L$ such that $f(\alpha) = 0$ and $(Df)(\alpha) = 0$,*
(3) *$\gcd(f, Df)$ has positive degree.*

Proof. We first show (1) implies (2). Supposing (1), we get that f factors over L as

$$f(x) = (x-\alpha)^2 g(x), \text{ where } g(x) \in L[x].$$

Hence:

$$\begin{aligned} Df &= (x-\alpha)^2 Dg + gD\left((x-a)^2\right) && \text{by Lemma 5.2} \\ &= (x-\alpha)^2 Dg + 2g(x-\alpha) && \text{by Corollary 5.3} \end{aligned}$$

so $(Df)(\alpha) = 0 + 0 = 0$.

For (2) implies (3), say α is a zero of both f and Df. Since $f(\alpha) = 0$, α is algebraic over E, so let $p(x)$ be the minimal polynomial of α over E. But $f \in E[x]$ and $f(\alpha) = 0$, so $p | f$,

and similarly $Df \in E[x]$ and $Df(\alpha) = 0$, so $p|Df$. Therefore, $p|\gcd(f, Df)$ and $\deg(p) \geq 1$. Thus (3) holds.

Finally, we show that (3) implies (1). Suppose $h(x) = \gcd(f, Df)$ has positive degree. Then $h(x)|f(x)$ and $f(x)$ splits completely in L, so $h(x)$ splits completely in L. Since $\deg(h) > 0$, h has a root $\alpha \in L$. Then α is a root of f. Say $f(x) = (x - \alpha)q(x)$ with $q(x) \in L[x]$. Therefore, by Lemma 5.2

$$Df = q(x) + (x - \alpha)Dq(x).$$

We know that h divides Df, so $Df(\alpha) = 0$. Substitute $x = \alpha$ into the formula for Df to get $0 = q(\alpha)$. Therefore, $f(x) = (x - \alpha)q(x)$ has α as a repeated root. □

Corollary 5.5. *If $p(x) \in E[x]$ is irreducible and not separable, then $(Dp)(x) = 0$.*

Proof. Let L be a splitting field for $p(x)$. Then $p(x)$ has a repeated root, so by Proposition 5.4, we obtain that $h(x) = \gcd(p, Dp)$ and $\deg(h) > 0$. Hence $h = \text{constant} \cdot p$, since p is irreducible, so $\deg(h) = \deg(p)$. But h divides Dp and $\deg(Dp) < \deg(p) = \deg(h)$. Therefore, Dp must be identically 0. □

Corollary 5.6. *If $f(x) \in E[x]$ is irreducible and not separable then the characteristic of E is a prime number $p > 0$, and $f(x) = a_0 + a_1x^p + a_2x^{2p} + \cdots + a_dx^{dp}$.*

Proof. Such polynomials are the only ones of positive degree such that $Df = 0$. □

It is still hard to come up with an inseparable polynomial though. For example, let us fix a prime p and set $E = \mathbb{Z}/p\mathbb{Z}$. Let us try $f(x) = x^p - 2$. Then indeed $Df = 0$, but the problem is that $f(x)$ is not irreducible. In fact, $f(x) = (x - 2)^p$.

Frobenius map. Suppose E is a field (or even a commutative ring), and $\text{char}(E) = p$, a prime number (i.e. $\underbrace{1 + 1 + \cdots + 1}_{p \text{ times}} = 0$.) Consider $\varphi\colon E \to E$ given by $\varphi(x) = x^p$.

We claim that this is a field (ring) homomorphism. Clearly, $\varphi(0) = 0$, $\varphi(1) = 1$, $\varphi(xy) = (xy)^p = x^py^p = \varphi(x)\varphi(y)$. Moreover:

$$\varphi(x + y) = (x + y)^p = \sum_{i=0}^{p} \binom{p}{i} x^i y^{p-i} = x^p + y^p = \varphi(x) + \varphi(y)$$

because if $1 \leq i \leq p - 1$, then $\binom{p}{i} = \frac{p!}{i!(p-i)!}$ and p does not cancel.

We call φ the *Frobenius homomorphism.*

Example. Let $E = \mathbb{Z}/p\mathbb{Z}$. Then $x^p \equiv x \bmod p$ by Fermat's Little Theorem, so φ is the identity element.

On the other hand, if $\mathbb{Z}/p\mathbb{Z} \subset E$, $\mathbb{Z}/p\mathbb{Z} \neq E$, then φ will be non-trivial. For example, say $E = \mathbb{F}_3(i)$, the splitting field of $x^2 + 1$ over $\mathbb{F}_3$. Then $\varphi(a) = a$ if $a = 0, 1, 2$, but $\varphi(i) = i^3 = -i$. Therefore, φ acts like *complex conjugation.*

Corollary 5.7. *If E is a field of characteristic p and $\varphi\colon E \to E$ is a bijection, then* **every** *polynomial $0 \neq f(x) \in E[x]$ is separable.*

Proof. Suppose f is not separable. Without loss of generality, f is irreducible. By Corollary 5.6, we have that

$$f(x) = a_0 + a_1 x^p + \cdots + a_d x^{pd}.$$

But φ is a bijection, so for all $0 \leq i \leq d$, there exists $b_i \in E$ such that $(b_i)^p = a_i$. Set

$$g(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_d x^d,$$

so that $g(x)^p = f(x)$, and f is not irreducible. □

Definition. A field E is *perfect* if every finite extension L/E is separable; equivalently, if every polynomial in $E[x]$ is separable.

Corollary 5.8.

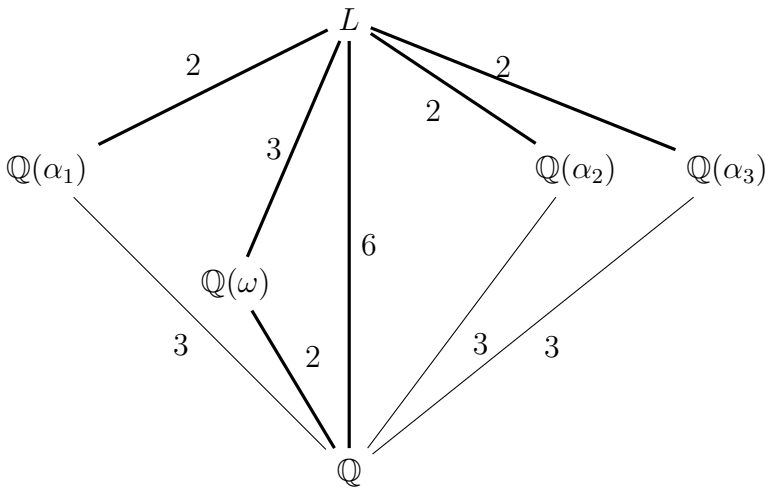
(1) *If* $\mathrm{char}(E) = 0$*, then* E *is perfect.*
(2) *If* E *is finite, then* E *is perfect.*

Proof. We already proved (1) in Corollary 5.6. For (2), note that φ is a field map, so φ is injective. But E is finite, so φ is bijection, and the statement follows from Corollary 5.7. □

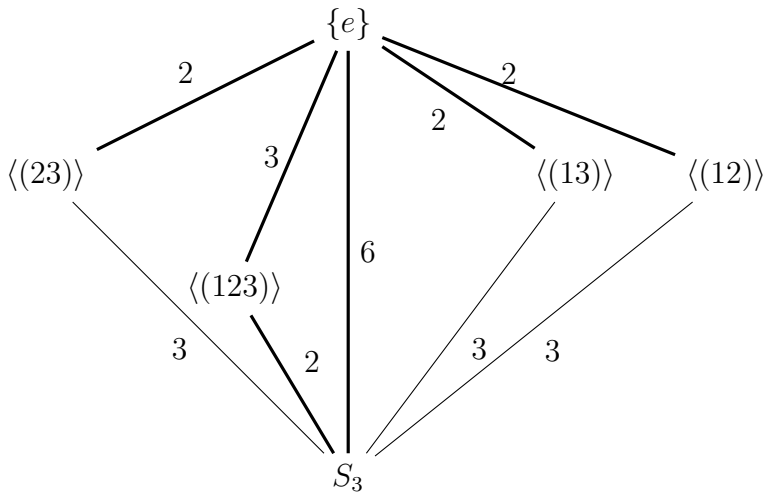
6. The fundamental theorem of Galois theory

Example (The fundamental theorem of Galois theory)**.** Let $K = \mathbb{Q}$ and L be the splitting field of $x^3 - 2$. Write $x^3 - 2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$ in $\mathbb{C}[x]$, where $\alpha_1 = \alpha = \sqrt[3]{2}$, and if $\omega = e^{2\pi i/3}$, then $\alpha_2 = \omega\alpha$, $\alpha_3 = \omega^2\alpha$. Then $L = \mathbb{Q}(\alpha, \omega)$.

The fundamental theorem of Galois theory says that the subfields of L are all those in the picture. The thin lines mean subfields, the thick lines mean normal extensions, and the numbers represent degrees.



Picture of S_3 and its subgroups. (Groups get bigger as we go down, and thick lines represent normal subgroups, and numbers represent indices.)



Definition. An algebraic extension L/K of fields is *Galois* if it is normal and separable.

We only consider finite extensions.

Definition. An extension L/K of fields is *finite Galois* if it is finite, normal, and separable.

Examples.

(1) The extension $\mathbb{C}/\mathbb{R}$ is finite (degree 2) (and hence algebraic), normal (splitting field of x^2+1), and separable (because characteristic is 0).
(2) For $K=\mathbb{Q}$ (or any characteristic 0 field), $f(x)=x^3-2$ (or any non-zero polynomial), let L be the splitting field of $f(x)$. If f has roots $\alpha_1,\dots,\alpha_d$, then $L=K(\alpha_1,\dots,\alpha_d)$, L/K is finite by the tower law. Finally, L/K is normal (Theorem 4.5) and separable (characteristic 0).

Idea: If L/K is Galois, we can associate to it a Galois group.

If L/K is finite and Galois, this group will be finite. A part of the fundamental theorem of Galois theory will be a correspondence

$$\{\text{subfields } K\subseteq M\subseteq L\}\longleftrightarrow\{\text{subgroups of Galois group}\}.$$

Here is how to get the group. Suppose $K\subseteq L$ are two fields (not necessarily Galois). Define

$$\mathrm{Aut}_K(L)=\{\text{field isomorphisms } \varphi\colon L\to L \text{ such that } \varphi(k)=k \text{ for all } k\in K\}.$$

Great example: if $L=\mathbb{C}$, $K=\mathbb{R}$, then $\mathrm{Aut}_{\mathbb{R}}(\mathbb{C})$ contains the identity $z\mapsto z$ on $\mathbb{C}$ and complex conjugation.

Exercise. Check that $\mathrm{Aut}_{\mathbb{R}}(\mathbb{C})=\{\text{identity, complex conjugation}\}$. Hint: if $\varphi\colon\mathbb{C}\to\mathbb{C}$ is a field map, then $\varphi(i)^2=\varphi(i^2)$.

Remark. Note that $[\mathbb{C}:\mathbb{R}]=2$ and $|\mathrm{Aut}_{\mathbb{R}}(\mathbb{C})|=2$.

Lemma 6.1. *Suppose $K\subseteq L$ are fields. Then $\mathrm{Aut}_K(L)$ is a group, where the group law is composition of functions.*

Proof. If φ and ψ are isomorphisms of fields $L \to L$ then so is $\varphi \circ \psi$ (easy check). If φ and ψ do not move elements of K, then neither does $\varphi \circ \psi$. Therefore, the group law is well-defined.

Associativity: always true for composition of functions.

Identity: identity map $L \to L$ is in $\mathrm{Aut}_K(L)$.

Inverses: inverse of an isomorphism is an isomorphism. □

Notation: If L/K is finite and Galois, then define $\mathrm{Gal}(L/K) = \mathrm{Aut}_K(L)$. For example, $\mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \{\text{identity, complex conjugation}\}$.

Theorem (Fundamental theorem of Galois theory)**.** *Say L/K is finite and Galois, and set $G = \mathrm{Gal}(L/K)$. Then*

(a) *G is finite and $|G| = [L : K]$.*
(b) *There is an order-reversing bijection*

$$\Phi\colon \{\textit{subgroups of } G\} \to \{\textit{subfields of } L \textit{ containing } K\},$$

$$\Phi(H) = \{\lambda \in L : h(\lambda) = \lambda \textit{ for all } h \in H\}.$$

(c) *If $H \subseteq G$ corresponds to $K \subseteq M \subseteq L$ via bijection Φ, then L/M is Galois, and $\mathrm{Gal}(L/M) = H$, and in particular $[L : M] = |H|$, so $[M : K] = |G/H|$.*
(d) *If H is a normal subgroup of G, then M/K is normal, and $\mathrm{Gal}(M/K) = G/H$.*

Suppose L/K is finite. We will provide a criterion for normality.

Here is an example of a normal extension: $K = \mathbb{Q}$, $L = \mathbb{Q}(\sqrt{2})$, the splitting field of $x^2 - 2$. Suppose $i\colon L \to \mathbb{C}$ is a field map and $i_{|K}$ is the identity (automatic). Then

$$i(\sqrt{2})^2 = i(\sqrt{2}^2) = i(2) = 2,$$

so $i(\sqrt{2}) = \pm\sqrt{2}$. In particular, $i(\sqrt{2}) \in L \subseteq \mathbb{C}$, and $i(L) \subseteq L$, so $i(L) = L$ for dimension reasons.

In fact, more generally, if L/K is finite and normal, and $L \subseteq M$ and $i\colon L \to M$ such that $i_{|K}$ is the identity, then $i(L) = L$. (Example Sheet 4, Question 8.)

If L/K is finite and not normal, this may fail. For example, if $L = \mathbb{Q}(\sqrt[3]{2})$ and

$$\beta = e^{2\pi i/3}\sqrt[3]{2} \in \mathbb{C}.$$

Then there is a field map $L \to \mathbb{C}$ that maps $\sqrt[3]{2}$ to β by Lemma 2.5. The image of i contains β so it is not contained in L.

Lemma 6.2. *A finite extension L/K is normal if and only if for any field extension $L \subseteq M$ and for any field map $i\colon L \to M$ such that $i_{|K}$ is the identity, we have $i(L) = L$.*

This lemma actually holds for any algebraic extension, but we only prove it in this generality.

Proof. The 'only if' implication is Question 8 from Exercise Sheet 4. For the 'if' implication, say $i(L) = L$ for all i. We will prove that L/K is normal. Suppose we have L/K finite. Let $L \subseteq M$, where M is a field large enough to make the proof work. For example, we can

take M to be the algebraic closure of L (but we did not prove these exist). It is, in fact, enough to let M be the normal closure of L/K (see Question 6 on Exercise Sheet 4).

Say $p(x) \in K[x]$ is irreducible and has a root $\alpha \in L$. We want to show $p(x)$ splits completely in L. We know M/K is normal, so $p(x)$ splits completely in M. Choose any root β of $p(x)$ in M. We want to show $\beta \in L$. By Lemma 2.5, the fields $K(\alpha)$ and $K(\beta)$ are isomorphic. So choose $i\colon K(\alpha) \to K(\beta)$ an isomorphism which is identity on K. Write $L = K(\alpha, \gamma_1, \gamma_2, \ldots, \gamma_r)$, since L/K is finite. We extend $i\colon K(\alpha) \to K(\beta)$ to a map $i\colon K(\alpha, \gamma_1, \ldots, \gamma_r) \to M$, recursively on r. (At each stage, let $p_j(x)$ be the minimal polynomial of γ_j over $K(\alpha, \gamma_1, \ldots, \gamma_{j-1})$. This works because M is normal over $K(\alpha, \gamma_1, \ldots, \gamma_j)$.) We end up with $i\colon L \to M$ and by hypothesis $i(L) = L$. But $i(L) \ni i(\alpha) = \beta$, and therefore $\beta \in L$. □

Now, say L/K is a finite field extension and $L \subseteq M$ for M big enough, i.e. M is an algebraic closure of L, or a normal closure of L/K.

Definition. The *separable degree* $[L : K]_s$ is the number of field maps $L \to M$ which are identity on K.

Example. Let $K = \mathbb{Q}$, $p(x) \in \mathbb{Q}[x]$ be irreducible, and α be a root of $p(x)$ in $\mathbb{C}$. Moreover, let $L = K(\alpha)$, and $M = \overline{\mathbb{Q}} \subseteq \mathbb{C}$. To give a map $L \to \overline{\mathbb{Q}}$ which is identity on $K = \mathbb{Q}$, we just have to decide where α goes. But α can go to any root of $p(x)$ in $\overline{\mathbb{Q}}$ by Lemma 2.5. If $d = \deg p(x)$, there will be d roots. Hence

$$[\mathbb{Q}(\alpha) : \mathbb{Q}]_s = d = [\mathbb{Q}(\alpha) : \mathbb{Q}].$$

Lemma 6.3. *If $L = K(\alpha)$, then*

$$[L : K]_s \leq [L : K]$$

with equality if and only if α is separable over K.

Proof. Let $p(x)$ be the minimal polynomial of α, of degree d. Then we know $[L : K] = d$ by Proposition 2.3. If M is the algebraic closure of L, then the number of field maps $i\colon L \to M$ which are identity on K is the number of roots of $p(x)$ in M by Lemma 2.5. Since $p(x)$ splits completely in M, the number of roots of $p(x)$ is at most d, and equal to d if and only if $p(x)$ is separable, i.e. α is separable. □

Lemma 6.4 (Tower law for separable degree)**.** *If $M \supseteq L \supseteq K$ and all the extensions are finite, then*

$$[M : K]_s = [M : L]_s [L : K]_s.$$

Proof. To give a field map $M \to \overline{M}$ which is identity on K it to do the following two things:

(1) give a map $i\colon L \to \overline{M} = \overline{L}$ which is identity on K ($[L : K]_s$ ways to do this),
(2) (regarding L as living in $\overline{M}$ via $i\colon L \to \overline{M}$) giving $M \to \overline{M}$ which is identity on $i(L)$ ($[M : L]_s$ ways to do this).

(Note that $\overline{M}$, $\overline{L}$, $\overline{i(L)}$, and $\overline{K}$ can all be thought of as the same field.) Therefore, $[M : K]_s = [L : K]_s [M : L]_s$. □

Corollary 6.5. *If L/K is a finite extension then*

(1) $[L : K]_s \leq [L : K]$
(2) *equality if and only if L/K is separable.*

Proof. Since L/K is finite, there exist $\gamma_1, \ldots, \gamma_n \in L$ such that $L = K(\gamma_1, \ldots, \gamma_n)$. Then we set $K_0 = K$ and $K_{i+1} = K_i(\gamma_i)$ to get

$$K = K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_n = K(\gamma_1, \ldots, \gamma_n) = L.$$

Recall that $[K_{i+1} : K_i]_s \leq [K_{i+1} : K_i]$ by Lemma 6.3, and so (1) follows from the two tower laws (the usual one 2.6 and Lemma 6.4).

For (2), note that if L/K is separable, then by Lemma 5.1 each K_{i+1}/K_i is separable, and hence equality holds by Lemma 6.3 and the tower laws. Conversely, if equality holds, choose $\alpha \in L$ arbitrary. We want to show α is separable over K. The trick is to note that $L \supseteq K(\alpha) \supseteq K$. Therefore:

$$\begin{array}{lll} [L : K]_s & = [L : K(\alpha)]_s[K(\alpha) : K]_s & \text{by Lemma 6.4} \\ & \leq [L : K(\alpha)][K(\alpha) : K] & \text{by part (1)} \\ & = [L : K] & \text{by the tower law} \\ & = [L : K]_s & \text{by assumption} \end{array}$$

Hence equalities must hold everywhere; in particular, $[K(\alpha) : K]_s = [K(\alpha) : K]$, showing that α is separable over K by Lemma 6.3. □

Theorem 6.6 (Fundamental theorem of Galois theory (a))**.** *If F/E is a finite extension of fields, then $|\mathrm{Aut}_E(F)| \leq [F : E]$, and equality holds if and only if F/E is Galois.*

Proof. Note that

$$\mathrm{Aut}_E(F) = \{\varphi\colon F \to F : \varphi_E = \mathrm{id}_E \text{ and } \varphi \text{ isomorphism}\} \subseteq \{\varphi\colon F \to \overline{F} : \varphi_E = \mathrm{id}_E\},$$

and the size of the latter set is $[F : E]_s$. Hence

$$|\mathrm{Aut}_E(F)| \leq [F : E]_s \leq [F : E],$$

as required. Finally, note that $\subseteq$ above is an equality if and only if F/E is normal (by Lemma 6.2), and the $\leq$ above is an equality if and only if F/E is separable (by Corollary 6.5). □

Now let us prove the fundamental theorem of Galois theory (b): If L/K is finite and Galois, and $G = \mathrm{Gal}(L/K)$, then there is a natural bijection

$$\{\text{subfields of } L \text{ containing } K\} \longleftrightarrow \{\text{subgroups of } G\}.$$

Here are the maps between these sets:

- Given M such that $K \subseteq M \subseteq L$, define

  $$\Gamma(M) = \{h \in G : h(m) = m \text{ for all } m \in M\}.$$

  Easily checked to be a subgroup of G.

- If $H \subseteq G$ is a subgroup, then set
$$\Phi(H) = \{m \in L : h(m) = m \text{ for all } h \in H\}.$$
This is a subfield because for all $h \in H$, $h(k) = k$ for any $k \in K$ by definition of G. Therefore, $K \subseteq \Phi(H)$, and in particular $0, 1 \in \Phi(H)$. Moreover, $\Phi(H)$ is clearly closed under the field operations, because h is a field map.

Setting
$$\mathcal{G} = \{\text{subgroups of } G\},$$
$$\mathcal{F} = \{\text{subfields } K \subseteq M \subseteq L\},$$
we have just defined
$$\Gamma \colon \mathcal{F} \to \mathcal{G},$$
$$\Phi \colon \mathcal{G} \to \mathcal{F}.$$
We want to show that $\Gamma \circ \Phi = \mathrm{id}$, $\Phi \circ \Gamma = \mathrm{id}$.

Lemma 6.7 (Easy observations about Γ and Φ)**.**

(1) *The maps Γ and Φ are order reversing (i.e. If $H_1 \subseteq H_2$ then $\Phi(H_1) \supseteq \Phi(H_2)$, and if $K \subseteq M_1 \subseteq M_2 \subseteq L$ then $\Gamma(M_1) \supseteq \Gamma(M_2)$).*
(2) *For $M \in \mathcal{F}$, we have $M \subseteq \Phi(\Gamma(M))$, and for $H \in \mathcal{G}$, we have $H \subseteq \Gamma(\Phi(H))$.*
(3) *We have $\Gamma \circ \Phi \circ \Gamma = \Gamma$ and $\Phi \circ \Gamma \circ \Phi = \Phi$.*

Proof. In each part, we only prove one statement, and the other is left as an exercise.

In (1), say $M_1 \subseteq M_2$. If $h \in \Gamma(M_2)$ then $h(m_2) = m_2$ for all $m_2 \in M_2$ by definition, and hence $h(m_1) = m_1$ for all $m_1 \in M_1 \subseteq M_2$; therefore $h \in \Gamma(M_1)$.

In (2), say $K \subseteq M \subseteq L$. Then
$$H := \Gamma(M) = \{h \in G : h(m) = m \text{ for all } m \in M\}.$$
Hence
$$\Phi(\Gamma(M)) = \Phi(H) = \{l \in L : h(l) \text{ for all } h \in \Gamma(M)\},$$
which clearly contains M, because if $l \in M$ then $h(l) = l$ by definition of h. Hence $M \subseteq \Phi(\Gamma(M))$.

For (3), we have
$$\begin{aligned}(\Gamma \circ \Phi \circ \Gamma)(M) &= \Gamma(\Phi(\Gamma(M))) \\ &\subseteq \Gamma(M) \qquad \text{by (1) and (2)}\end{aligned}$$
and
$$\begin{aligned}(\Gamma \circ \Phi \circ \Gamma)(M) &= (\Gamma \circ \Phi)(\Gamma(M)) \\ &\supseteq \Gamma(M) \qquad \text{by (1) and (2)}\end{aligned}$$
which completes the proof. □

Example. Let $K = \mathbb{Q}$, $L = \mathbb{Q}(\sqrt{2}, i)$. Then L/K is finite and Galois because

- $\mathrm{char}(\mathbb{Q}) = 0$, so L/K is separable,
- L is the splitting field of $f(x) = (x^2 - 2)(x^2 + 1)$, so L/K is finite and normal.

First question: What is $\mathrm{Gal}(L/K)$? By the fundamental theorem of Galois theory (a), we know that

$$|\mathrm{Gal}(L/K)| = [L:K] = \Big[\mathbb{Q}(\sqrt{2},i):\mathbb{Q}\Big] = \Big[\mathbb{Q}(\sqrt{2},i):\mathbb{Q}(\sqrt{2})\Big]\Big[\mathbb{Q}(\sqrt{2}):\mathbb{Q}\Big] = 2\cdot 2 = 4.$$

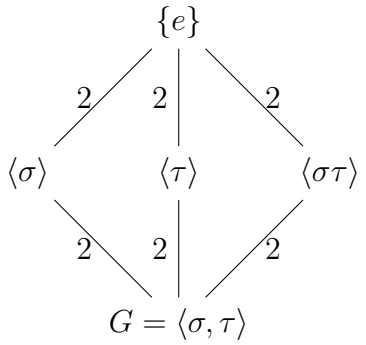
A map $\varphi\colon L\to L$ which is identity on $K=\mathbb{Q}$ is determined by where it sends $\sqrt{2}$, i, because these elements generate L over $\mathbb{Q}$. The usual trick: $\varphi(\sqrt{2})^2 = \varphi(\sqrt{2}^2) = \varphi(2) = 2$, and hence $\varphi(\sqrt{2}) = \pm\sqrt{2}$. Similarly, $\varphi(i) = \pm i$. Thus we have at most four choices for φ. But $|\mathrm{Gal}(L/K)| = 4$, and hence all four choices must work.

For each choice of pair $(a,b)\in\{\pm1\}\times\{\pm1\}$, there exists a field isomorphism $\varphi_{a,b}\colon L\to L$ which is identity on $\mathbb{Q}$ such that $\varphi_{a,b}(\sqrt{2}) = a\sqrt{2}$, $\varphi_{a,b}(i) = bi$.

Define $\sigma\colon L\to L$ such that $\sigma(\sqrt{2}) = \sqrt{2}$, $\sigma(i) = -i$ and $\tau\colon L\to L$ such that $\tau(\sqrt{2}) = -\sqrt{2}$, $\tau(i) = i$. Hence

$$G = \mathrm{Gal}(L/K) = \{\mathrm{id}, \sigma, \tau, \sigma\tau\}.$$

Subgroups of G?



The fundamental theorem of Galois theory (b) says there exists a corresponding picture for subfields of L containing K. Subfields of $L = \mathbb{Q}(\sqrt{2},i)$ containing K:

$$\Phi(H) = \{\lambda\in L : h(\lambda) = \lambda \text{ for all } h\in H\}.$$

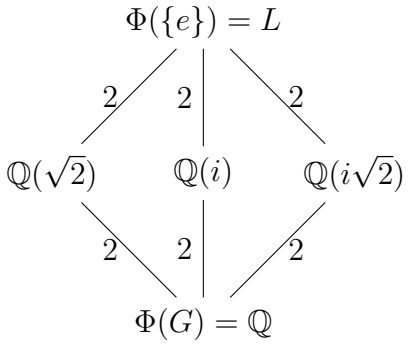
Clearly, $\Phi(\{e\}) = L = \mathbb{Q}(i,\sqrt{2})$. For the rest, we first note that $[L:K] = 4$ and a $\mathbb{Q}$-basis for L is $\{1, i, \sqrt{2}, i\sqrt{2}\}$. For $\lambda = \alpha+\beta i+\gamma\sqrt{2}+\delta i\sqrt{2}$, $\alpha,\beta,\gamma,\delta\in\mathbb{Q}$, a general element of L, we have

$$\sigma(\lambda) = \alpha-\beta i+\gamma\sqrt{2}-\delta i\sqrt{2},$$

$$\tau(\lambda) = \alpha+\beta i-\gamma\sqrt{2}-\delta i\sqrt{2},$$

$$(\sigma\tau)(\lambda) = \alpha-\beta i-\gamma\sqrt{2}+\delta i\sqrt{2}.$$

Therefore, by checking which elements of the field remain fixed, we obtain the following intermediate field diagram (by applying Φ to the group diagram above):



Now, the fundamental theorem of Galois theory (b) says that these are all the subfields.

Proposition 6.8. *Let L/K be a finite Galois extension. Then $\Phi \circ \Gamma$ is the identity map $\mathcal{F} \to \mathcal{F}$.*

Proof. First, here is an observation. If L/K is finite Galois and $L \supseteq M \supseteq K$, M a subfield, then L/M is finite Galois: finiteness follows from the Tower law 2.6, normality follows from Lemma 4.6, separability follows from Lemma 5.1. Furthermore,

$$\mathrm{Gal}(L/M) = \{\varphi \colon L \to L : \text{isomorphism such that } \varphi_{|M} = \mathrm{id}_M\} \subseteq \mathrm{Gal}(L/K).$$

Moreover,

$$\mathrm{Gal}(L/M) = \{h \in \mathrm{Gal}(L/K) : h(m) = m \text{ for all } m \in M\} = \Gamma(M).$$

Now, say $M \in \mathcal{F}$ and set $N = \Phi\Gamma(M)$. We want to show $N = M$. By Lemma 6.7 (2), $N \supseteq M$. By Lemma 6.7 (3),

$$\Gamma(N) = \Gamma\Phi\Gamma(M) = \Gamma(M).$$

By the fundamental theorem of Galois theory (a) 6.6, we know that $|\mathrm{Gal}(L/M)| = |\mathrm{Gal}(L/N)|$ implies that

$$[L : M] = [L : N].$$

Therefore, $[L : M] = [L : N][N : M]$, and hence $[N : M] = 1$, so $N = M$. □

Corollary 6.9. *The map Γ is injective.*

In particular, because $\mathcal{G}$ is obviously finite, we can deduce that $\mathcal{F}$ is finite.

Corollary 6.10. *If F/E is a finite separable extension, then there are only finitely many intermediate fields M with $E \subseteq M \subseteq F$.*

This is not obvious (and it is, in fact, false if we drop the assumption of separability—see Exercise Sheet 5, Question 5.)

Proof. Let L be the normal closure of F over E, i.e. L/E is the smallest normal extension of E containing F. Since L/E is the normal closure of a separable extension, it is separable (by Exercise Sheet 5, Question 3d). Therefore, there are only finitely many subfields of L containing E, and hence there are only finitely many subfields of F containing E. □

Corollary 6.11 (Theorem of the primitive element)**.** *If L/K is finite and separable then there exists $\alpha \in L$ such that $L = K(\alpha)$.*

Again, this statement is false when we drop the assumption of separability—see Exercise Sheet 5, Question 5.

Proof. There are two cases.

Case 1. K is finite (as a set). Then L is finite, so $L^\times$ is a finite subgroup of the non-zero elements of a field, so $L^\times$ is cyclic (by Exercise Sheet 3, last question). Therefore, there exists α such that $L^\times = \langle\alpha\rangle$. Then $L = K(\alpha)$.

Case 2. K is infinite. Then L is a finite-dimensional vector space over an infinite field. But, by Exercise Sheet 6, Question 5, this implies that L is not the union of finitely many proper subspaces. Therefore,

$$L \neq \bigcup_{K \subseteq M \subset L} M.$$

Choose $\alpha \in L$ such that $\alpha \notin \bigcup_{K \subseteq M \subset L} M$. But $K(\alpha)$ is a subfield of L containing α, so it cannot be proper, and hence $K(\alpha) = L$. □

Here is a proposition motivated by the definition of Φ.

Proposition 6.12. *Say L is a field and G is a finite subgroup of* $\mathrm{Aut}(L)$*. Set*

$$K = \{\lambda \in L : g(\lambda) = \lambda \text{ for all } g \in G\}.$$

Then $[L : K]$ is finite, L/K is finite and Galois, and $\mathrm{Gal}(L/K) = G$*. In particular,*

$$[L : K] = |G|.$$

Proof. First, let us prove that $[L : K]$ is finite and is at most $n = |G|$. We will do this by showing that if $\alpha_1, \alpha_2, \ldots, \alpha_{n+1} \in L$ are arbitrary then they are K-linearly-dependent. Let V be the vector space L^n. An element of V is just n elements of L. Say $G = \{g_1, \ldots, g_n\}$. Given $(\lambda_1, \ldots, \lambda_n) \in V$, we can consider the function $G \to L$ such that $g_i \mapsto \lambda_i$. This sets up a natural bijection:

$$V = \{\text{functions } G \to L\}.$$

From now on, think of V as this set of functions. Here are $n + 1$ elements of V. For $1 \leq i \leq n + 1$, define $\varphi_i \colon G \to L$ by $\varphi_i(g) = g(\alpha_i)$. These are $n + 1$ elements in an n-dimensional vector space over L, so they are linearly dependent. Therefore, there exist $\mu_i \in L$ not all equal to 0 such that

$$\sum_{i=1}^{n+1} \mu_i \varphi_i = 0.$$

Choose a linear relation of smallest *length*, i.e. $\mu_i = 0$ for $i > r$, $\mu_r \neq 0$ and $\sum_{i=1}^{r} \mu_i \varphi_i = 0$ and r is as small as possible with this property. (I.e. the first r of the φ_i are linearly dependent but the first $r - 1$ are not). We know that $\mu_r \neq 0$, so replacing μ_i with μ_i/μ_r, we can assure

that $\mu_r = 1$ (and $\mu_i = 0$ for $i > r$). We will now show that all $\mu_i \in K$. We know that for all $g \in G$:

$$\sum_{i=1}^{r} \mu_i g(\alpha_i) = \sum_{i=1}^{r} \mu_i \varphi_i(g) = 0. \tag{1}$$

Applying $h\colon L \to L$, any element of G, to the equation, we obtain

$$\sum_{i=1}^{r} h(\mu_i) hg(\alpha_i) = h(0) = 0$$

and since $hg \in G$, we can substitute it to equation (1) to obtain

$$\sum_{i=1}^{r} \mu_i hg(\alpha_i) = 0.$$

Subtracting, we have for all $g \in G$

$$\sum_{i=1}^{r} (h(\mu_i) - \mu_i) hg(\alpha_i) = 0$$

so writing $k = hg$, we have for all $k \in K$

$$\sum_{i=1}^{r} (h(\mu_i) - \mu_i)\varphi_i(k) = \sum_{i=1}^{r} (h(\mu_i) - \mu_i) k(\alpha_i) = 0.$$

But $h(\mu_r) - \mu_r = 1 - 1 = 0$, and the new linear relation has *smaller length*. Therefore, $h(\mu_i) = \mu_i$ for all $i \leq r$, and hence $h(\mu_i) = \mu_i$ for $1 \leq i \leq n+1$. But $h \in G$ was arbitrary. Since μ_i are fixed by all G, $\mu_i \in K$. Then equation (1) with $g = \mathrm{id}$ yields

$$\sum_{i=1}^{r} \mu_i \alpha_i = 0$$

and $\mu_r = 1$, so not all μ_i are 0. Therefore, α_i are K-linearly-dependent.

We have shown that any $n+1$ elements of L are K-linearly-dependent, and hence $[L : K] \leq n = |G|$. In particular, L/K is finite. We will prove it is Galois.[9]

Separability. Choose $\alpha \in L$. Set $S = \{g(\alpha) : g \in G\}$ and note that $|S| \leq n = |G|$. Set

$$p(x) = \prod_{s \in S} (x - s) \in L[x].$$

We claim that $p(x) \in K[x]$. Choose $h \in G$. Then

$$h(S) = \{h(s) : s \in S\} = \{hg(\alpha) : g \in G\} = S.$$

Therefore,

$$h(p(x)) = \prod_{s \in h(S)} (x - s) = p(x),$$

[9]There is a quicker way to finish the proof. By Theorem 6.6, $|\mathrm{Aut}_K(L)| \leq [L : K]$ with equality if and only if L/K is Galois, and clearly $G \subseteq \mathrm{Aut}_K(L)$, so

$$|G| \leq |\mathrm{Aut}_K(L)| \leq [L : K] \leq |G|.$$

Hence equalities hold, and L/K is Galois with Galois group G.

so if $p(x) = \sum a_i x^i$, then $h(a_i) = a_i$ for all i, $h \in G$. Therefore, $p(x) \in K[x]$. We now claim that $p(x)$ is the minimal polynomial of α over K. Since $\alpha \in S$, $p(\alpha) = 0$. Moreover, if $q(x) \in K[x]$ and $q(\alpha) = 0$, then for any $\beta \in S$, there exists $g \in G$ such that $g(\alpha) = \beta$. Since $q(\alpha) = 0$, we have that $g(q(\alpha)) = 0$, and so $g(q)(g(\alpha)) = 0$. But $g(q) = q$ and $g\alpha = \beta$, and so $q(\beta) = 0$. Hence all $\beta \in S$ are roots of $q(x)$, so $p(x)$ must divide $q(x)$. Hence $p(x)$ is the minimal polynomial of α, and its roots are distinct, so α is separable.

Normality. If $f(x) \in K[x]$ is irreducible and has a root $\alpha \in L$, then build $p(x)$ for α as above. Then $p(x)$ is the minimal polynomial, and hence $p(x)$ divides $f(x)$, so $p = f$ (up to scaling). But all roots of p are in L by construction, and thus f splits completely.

Therefore, L/K is Galois, so by definition $G \hookrightarrow \mathrm{Gal}(L/K)$. Hence

$$|G| \leq |\mathrm{Gal}(L/K)| = [L : K] \leq |G|,$$

so the injection is an isomorphism. □

This is the last thing we needed to prove the fundamental theorem of Galois theory.

Theorem 6.13 (Fundamental theorem of Galois theory)**.** *Let L/K be a finite Galois extension and $G = \mathrm{Gal}(L/K)$; write $\mathcal{F} = \{K \subseteq M \subseteq L$ subfields$\}$, $\mathcal{G} = \{H \subseteq G$ subgroups$\}$. Define two maps*

$$\Phi \colon \mathcal{G} \to \mathcal{F}$$
$$\Phi(H) = \{\lambda \in L : h(\lambda) = \lambda \text{ for all } h \in H\}$$
$$\Gamma \colon \mathcal{F} \to \mathcal{G}$$
$$\Gamma(M) = \{g \in G : g(m) = m \text{ for all } m \in M\}$$

Then

(a) $[L : K] = |G|$
(b) Γ *and* Φ *are order-reversing bijections*
(c) *If $M \in \mathcal{F}$ corresponds to $H \in \mathcal{G}$ then $H = \mathrm{Gal}(L/M)$.*
(d) *The subgroup H of G is normal if and only if M/K is normal, and in this case*

$$\mathrm{Gal}(M/K) = G/H$$

(i.e. they are isomorphic and there is a natural *isomorphism between them.)*

Proof. We proved (a) in Theorem 6.6.

For (b), we already proved that Γ and Φ are order-reversing in Lemma 6.7, and that $\Phi \circ \Gamma = \mathrm{id}$ in Proposition 6.8. We only need to prove that $\Gamma \circ \Phi = \mathrm{id}$. So say $H \subseteq G$ is a subgroup and set $M = \Phi(H)$. We need to check that

$$\Gamma(M) = \{g \in G : g(m) = m \text{ for all } m \in M\}$$

is equal to H. Since

$$M = \Phi(H) = \{\lambda \in L : h(\lambda) = \lambda \text{ for all } h \in H\},$$

Proposition 6.12 says that L/M is finite and Galois, and by definition the elements of G that fix M element-wise are precisely $\mathrm{Gal}(L/M)$. Therefore, $\mathrm{Gal}(L/M) = \Gamma(M)$. Then Proposition 6.12 says that $\mathrm{Gal}(L/M) = H$, and hence $H = \Gamma(M)$.

Now, (c) is obvious: if $H \longleftrightarrow M$, then $H = \Gamma(M) = \mathrm{Gal}(L/M)$.

For (d), let us first suppose that M/K is normal for $M \in \mathcal{F}$. Then M/K is separable (by Lemma 5.1, since L/K is separable) and finite (since $M \subseteq L$), and hence M/K is finite and Galois. Set $Q = \mathrm{Gal}(M/K)$ and say $g \in G$. Then $g\colon L \to L$, and hence g restricts to a map $g_{|M}\colon M \to L$, fixing K pointwise. By Lemma 6.2, $g(M) = M$, since M is normal. Therefore, $g_{|M}\colon M \to M$ is the identity on K, and $g_{|M} \in Q$. In particular, there is a natural map $G \to Q$ that maps g to $g_{|M}$. It is easy to check it is a group homomorphism. The kernel of this map is

$$\{g \in G : g_{|M} = \mathrm{id}_M\} = \Gamma(M).$$

Set $H = \Gamma(M)$. Then H is a kernel, and therefore H is a normal subgroup of G. We have shown that M normal implies that H is normal.

In particular, the first isomorphism theorem implies that $G/H \cong \mathrm{im}(G) \subseteq Q$. By counting,

$$\begin{aligned} |G/H| &= \tfrac{|\mathrm{Gal}(L/K)|}{|\mathrm{Gal}(L/M)|} && \text{by part (c)} \\ &= \tfrac{[L:K]}{[L:M]} && \text{by part (a)} \\ &= [M:K] && \text{by the Tower law 2.6} \end{aligned}$$

On the other hand, $Q = \mathrm{Gal}(M/K)$, and hence $|Q| = [M : K]$ by part (a). Therefore, $|\mathrm{im}(G)| = |Q|$, and the map $G \to Q$ is hence surjective. In particular, by the first isomorphism theorem, $G/H \cong Q$.

Finally, say $H \in \mathcal{G}$ and $H \trianglelefteq G$ is a normal subgroup. Set $M = \Phi(H)$. We want to show that M is normal. We know that

$$M = \Phi(H) = \{\lambda \in L : h(\lambda) = \lambda \text{ for all } h \in H\}.$$

If $g \in G$, what is $\Phi(g^{-1}Hg)$? Well, for $h \in H$, $h(\lambda) = \lambda$, and hence

$$(g^{-1}hg)(g^{-1}\lambda) = g^{-1}h\lambda = g^{-1}\lambda.$$

Therefore, $\Phi(g^{-1}Hg) = g^{-1}M$. So if H is normal, $g^{-1}Hg = H$ for all $g \in G$, and hence $g^{-1}M = M$ for all $g \in G$, so $gM = M$ for all $g \in G$. Hence $g \in G$ induces a field map $M \to M$. In particular, we get a natural map $G \to \mathrm{Aut}_K(M)$ with the kernel $\Gamma(M) = H$. Hence, we get an injection

$$G/H \hookrightarrow \mathrm{Aut}_K(M).$$

Now, $H = \mathrm{Gal}(L/M)$, so

$$|G/H| = \frac{[L:K]}{[L:M]} = [M:K],$$

which shows that $[M : K] \leq |\mathrm{Aut}_K(M)|$. But Corollary 6.5 shows that $|\mathrm{Aut}_K(M)| \leq [M : K]$ with equality if and only if M/K is Galois. Therefore, M/K is normal. □

Example. Let $N \in \mathbb{Z}_{\geq 1}$. Say $f(x) = x^N - 1 \in \mathbb{Q}[x]$ and let L be the splitting field of $f(x)$. Then $L/\mathbb{Q}$ is finite and Galois. What is the Galois group?

What is L? Suppose $\zeta_N = e^{2\pi i/N} \in \mathbb{C}$. The roots of $f(x)$ are $1, \zeta_N, \zeta_N^2, \ldots, \zeta_N^{N-1}$, and hence

$$L = \mathbb{Q}(1, \zeta_N, \zeta_N^2, \ldots, \zeta_N^{N-1}) = \mathbb{Q}(\zeta_N).$$

We call L the N^{th} *cyclotomic field*. We know that $[L : \mathbb{Q}]$ is the degree of the minimal polynomial of ζ_N over $\mathbb{Q}$. In general, we have not computed this. The answer is that it is (by definition) the N^{th} *cyclotomic polynomial*. Tricky exercise: its degree is $\varphi(N)$, where φ

is the Euler totient function. We proved this in the special case where N is a prime number. In this case, ζ_N is a root of the irreducible polynomial

$$\frac{x^N-1}{x-1} = 1 + x + \cdots + x^{N-1}.$$

Here is a map $\mathrm{Gal}(L/\mathbb{Q}) \to (\mathbb{Z}/N\mathbb{Z})^\times$. If $g \in \mathrm{Gal}(L/\mathbb{Q})$, $g(\zeta_N)$ is a root of $x^N - 1$, which is $(\zeta_N)^a$ for some $0 \leq a \leq N-1$. The map is clearly injective. It is slightly tricky to show that it is an isomorphism. If N is prime, we can prove this, since the LHS has size $[L : \mathbb{Q}] = N - 1$ and the RHS has size $N - 1$, because for N prime

$$(\mathbb{Z}/N\mathbb{Z})^\times = \{1, 2, \ldots, N-1\}.$$

In fact, if N is prime, $\mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) = (\mathbb{Z}/N\mathbb{Z})^\times$, the non-zero elements of $\mathbb{Z}/N\mathbb{Z}$.

We give a trick for computing Galois groups. Say $f(x) \in K[x]$ is a non-zero polynomial, and assume (for simplicity) $\mathrm{char}(K) = 0$ (or that f is separable). Suppose $\deg(f) = n$ and say $\alpha_1, \ldots, \alpha_n \in \overline{K}$ are the roots of f. Set $L = K(\alpha_1, \ldots, \alpha_n)$, the splitting field of f. Then L/K is finite, separable, and normal (as it is a splitting field), and hence L/K is finite and Galois. Set $G = \mathrm{Gal}(L/K)$. Then the fundamental theorem of Galois theory 6.13 (a) says that $|G| = [L : K]$.

The trick is that G is a subgroup of S_n. Here is why: If $g \in G$, then $g\colon L \to L$, so $g(\alpha_i)$ is a root of $g(f(x)) = f(x)$, as $g_{|K} = \mathrm{id}_K$. Thus g permutes the roots of $f(x)$. So given $g \in G$, we get a permutation of $S = \{\alpha_1, \ldots, \alpha_n\}$, so we get a homomorphism $G \to \mathrm{Sym}(S) \cong S_n$. Furthermore, the map is injective: if $g \in G$ and $g(\alpha_i) = \alpha_i$ for all i, then g fixes all α_i and g fixes all $\lambda \in K$, and therefore g fixes $K(\alpha_1, \ldots, \alpha_n) = L$, and hence $g = \mathrm{id}_L$.

Example. Let $K = \mathbb{Q}$, $f(x) = x^3 - 2$. Then $L = \mathbb{Q}(\alpha, \omega\alpha, \omega^2\alpha) = \mathbb{Q}(\alpha, \omega)$ for $\alpha = \sqrt[3]{2}$ and $\omega = e^{2\pi i/3}$. We know that $[L : \mathbb{Q}] = 6$, so $|\mathrm{Gal}(L/\mathbb{Q})| = 6$, and the injection $\mathrm{Gal}(L/\mathbb{Q}) \hookrightarrow S_3$ has to be an isomorphism. An element $g \in \mathrm{Gal}(L/\mathbb{Q})$ is determined by the corresponding permutation of $\{\alpha, \omega\alpha, \omega^2\alpha\}$.

7. Insolvability of the quintic (by radicals)

Idea: Solutions of $ax^2 + bx + c = 0$ (for $a \neq 0$) are $x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$. There are analogous formulae for the cubic and quartic, involving only $+ - \times \div \sqrt[n]{\ }$ for $n = 2, 3, 4, \ldots$

Abel–Ruffini and later Galois showed that there was no such formula for the quintic.

The operations $+ - \times \div$ we can do in any field. However, $\sqrt[n]{\ }$ we need to be able to *insert into the system.*

Assume throughout this chapter that all fields in it have characteristic 0.

Definition. A finite extension L/K of fields (of characteristic 0) is an *extension by radicals* if there exist fields $K = L_0 \subseteq L_1 \subseteq L_2 \subseteq L_3 \subseteq \cdots \subseteq L_n = L$ such that for each i, $1 \leq i \leq n$, there exists $\alpha_i \in L_i$ such that $L_i = L_{i-1}(\alpha_i)$ and there exists $n_i \in \mathbb{Z}_{\geq 1}$ such that $\alpha_i^{n_i} \in L_{i-1}$, i.e. "$L_i$ is a field generated by L_{i-1} and one n_ith root".

Example. If $\beta = \sqrt[3]{1+\sqrt{2}}$ then there exists a field $L \subseteq \mathbb{C}$ such that $L \ni \beta$ and L is an extension of $\mathbb{Q}$ by radicals. Explicitly:

$$\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\sqrt{2})\left(\sqrt[3]{1+\sqrt{2}}\right) \ni \beta.$$

Exercise. Prove that if $\gamma = \sqrt[50]{\frac{1+\sqrt[3]{3}}{1+\sqrt[4]{5}}}$, then there exists an extension by radicals $L/\mathbb{Q}$ with $\gamma \in L$.

Remark. If L/K is an extension by radicals then the normal closure M/K of L/K is too (for the proof, see Exercise Sheet 6, Question 7.). Hence we can assume that our extensions are actually Galois (by making them bigger, if necessary).

Idea. Suppose there exists a formula for the roots of a quintic involving only $+ - \times \div \sqrt[n]{}$. Apply it to the quintic $x^5 - 6x + 3$. Let L be the splitting field of this quintic over $\mathbb{Q}$. Then $L/\mathbb{Q}$ is finite and Galois, and an extension by radicals. The idea is that that being an extension by radicals tells us something about $\mathrm{Gal}(L/\mathbb{Q})$. We will show that $\mathrm{Gal}(L/\mathbb{Q})$ does **not** have this property if L is the splitting field of $x^5 - 6x + 3$.

(Spoilers: the Galois group is S_5 and S_5 is not a solvable group.)

We now prove a key lemma.

Lemma 7.1. *Suppose E is a field of characteristic 0, p is a prime number, and assume that $x^p - 1$ splits completely in E. Then for all $\beta \in E$, the polynomial $x^p - \beta$ either splits into linear factors or is irreducible in $E[x]$.*

In the irreducible case, if $\alpha \in \overline{E}$ is a root of $x^p - \beta$ then $F = E(\alpha)$ is the splitting field of $x^p - \beta$ and $[F : E] = p$.

Proof. Let F be the splitting field of $x^p - \beta$. Then, if pth roots of 1 in E are $\{1, \omega, \omega^2, \ldots, \omega^{p-1}\}$, we know $F = E(\alpha, \omega\alpha, \omega^2\alpha, \ldots, \omega^{p-1}\alpha) = E(\alpha)$, α a pth root of β.

Case 1: $\alpha \in E$. Then $x^p - \beta$ splits completely.

Case 2: $\alpha \notin E$. Say $x^p - \beta$ were reducible in $E[x]$, so it factors into 2 polynomials g, h of degrees m, n with $1 \leq m, n < p$, $m + n = p$. By scaling, we can assume that g, h are monic. Consider $F[x]$, where polynomials factor uniquely. Then

$$gh = x^p - \beta = (x - \alpha)(x - \omega\alpha)\ldots(x - \omega^{p-1}\alpha),$$

so the irreducible factors of $g(x)$ in $F[x]$ are all of the form $(x - \omega^i\alpha)$ for some i. Therefore:

$$g(x) = x^m + \cdots \pm \omega^j\alpha^m$$

for some j. But $g(x) \in E[x]$ and therefore $\pm\omega^j\alpha^m \in E$, and hence $\alpha^m \in E$. Also, $\alpha^p = \beta \in E$. But $1 \leq m \leq p - 1$, and thus $\gcd(m, p) = 1$, and hence there exist $\alpha, \mu \in \mathbb{Z}$ such that $\lambda m + \mu p = 1$, which shows that

$$\alpha = \alpha^{\lambda m + \mu p} = (\alpha^m)^\lambda \cdot (\alpha^p)^\mu \in E.$$

Hence $\alpha \in E$, which gives a contradiction. □

Exercise. What is $\mathrm{Gal}(F/E)$ in case 2?

Funny examples of extensions by radicals where the degree of the of extensions is not what you expect.

(1) Let $\zeta = e^{2\pi i/p}$, then $\zeta \notin \mathbb{Q}$, $\zeta^p \in \mathbb{Q}$, p prime, and $[\mathbb{Q}(\zeta) : \mathbb{Q}] = p - 1 \neq p$.
(2) If $\alpha = \sqrt[3]{2} \in \mathbb{R}$, $\omega = e^{2\pi i/3}$, $\beta = \omega\alpha$, and $E = \mathbb{Q}(\alpha)$, $F = E(\beta)$, then $[F : E] = 2$.

Idle question. If $1 \leq m \leq n$, can you find a field of characteristic 0, E, and an extension $E \subseteq F$, such that $F = E(\alpha)$, $\alpha^n \in E$ and $[F : E] = m$?

If G is a finite group, then a *finite filtration* on G is a sequence

$$G = H_0 \supseteq H_1 \supseteq H_2 \supseteq \cdots \supseteq H_n = \{1\}$$

of subgroups H_i of G.

Fact from the Group Theory course: The following are equivalent for a finite group G:

(1) G has a finite filtration by subgroups H_i such that $H_{i+1} \trianglelefteq H_i$ and the quotient group H_i/H_{i+1} is abelian,
(2) same, but H_i/H_{i+1} is cyclic,
(3) same, but H_i/H_{i+1} is cyclic of prime order.

A group is *solvable* or *soluble* if it satisfies any (and hence all) of these equivalent conditions.

Basic facts: If G is solvable then any subgroup and any quotient G/N are solvable.

Conversely, if G is a finite group, $N \trianglelefteq G$, and N and G/N are both solvable, then G is solvable.

Examples. Abelian groups are solvable. The symmetric group S_4 is solvable:

$$S_4 \rhd A_4 \rhd V_4 = \{e, (12)(34), (13)(24), (14)(23)\} \rhd \{1\}.$$

However, S_5 is not solvable, (and therefore, S_n is not solvable for all $n \geq 5$).

Cheap proof that S_n and A_n, $n \geq 5$, are not solvable. Say $G = S_n$ or A_n for $n \geq 5$. Set $s_1 = (12)(34) \in G$, $s_2 = (135) \in G$, and $s_3 = s_1 s_2 = (14352) \in G$. Suppose for a contradiction that G is solvable. Write

$$G = H_0 \supseteq H_1 \supseteq H_2 \supseteq \cdots \supseteq H_d = \{1\}$$

with $H_{i+1} \lhd H_i$, H_i/H_{i+1} cyclic of prime order.

We claim that $s_1, s_2, s_3 \in H_i$ for all i. This is a contradiction, since they are not in H_d. We prove the claim by induction on i. Assume $s_1, s_2, s_3 \in H_i$. Consider the images $t_1, t_2, t_2 \in H_i/H_{i+1} \cong C_p$, the order of t_i divides the order of s_i, order of t_i is 1 or p, since $p = |H_i/H_{i+1}|$. Hence p is at most one of 2, 3, 5, and so at least 2 of t_i are identity. But $t_1 t_2 = t_3$ and hence t_i is the identity, so $s_i \in H_{i+1}$. □

Theorem 7.2. *If F/E is finite Galois for E of characteristic 0, and F/E is an extension by radicals then* $\mathrm{Gal}(F/E)$ *is solvable.*

Proof. Since F/E is solvable by radicals, suppose

$$E = F_0 \subseteq F_1 \subseteq F_2 \subseteq \cdots \subseteq F_d = F$$

with $F_{i+1} = F_i(\alpha_i)$, $\alpha_i^{n_i} \in F_i$. Without loss of generality, we can assume that all n_i are prime (for example, if $n = pq$, replace α by α^p, α to get $E \subseteq E(\alpha^p) \subseteq E(\alpha)$ with $(\alpha^p)^q \in E$.)

Set N to be the product of all the n_i. Throw in an Nth root of unity to apply Lemma 7.1. Set $K = E(\zeta_N) =$ splitting field of $x^N - 1$ over E, where $(\zeta_N)^N = 1$. Let $L_0 = K$ and $L_i = F_i(\zeta_N)$, the splitting field of $x^N - 1$ over F_i. Then $L_{i+1} = F_{i+1}(\zeta_N) = F_i(\zeta_N, \alpha_i) = L_i(\alpha_i)$, so

$$E \subseteq K = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_d \subseteq F(\zeta_N)$$

is still an extension by radicals.

By applying the fundamental theorem of Galois theory 6.13 to the tower of extensions above, we get that if $G = \mathrm{Gal}(F(\zeta_N)/E)$ then

$$G = G_d \supseteq G_{d-1} \supseteq \cdots \supseteq G_0 \supseteq \{1\}.$$

The fundamental theorem of Galois theory 6.13 applied to L_{i+1}/L_i shows that $G_i \lhd G_{i+1}$ and the quotient is cyclic of prime order. Finally, we need to check $\mathrm{Gal}(K/E)$ is abelian. But $K = E(\zeta_N)$. If $g, h \in \mathrm{Gal}(K/E)$, then $g(\zeta_N) = \zeta_N^A$ for some A and $h(\zeta_N) = \zeta_N^B$ for some B, and so $gh(\zeta_N) = \zeta_N^{BA} = \zeta_N^{AB} = hg(\zeta_N)$. Thus gh and hg agree on E and ζ_N, and hence they agree on K. Therefore, $gh = hg$ and we have shown that $\mathrm{Gal}(K/E)$ is abelian.

Thus $\mathrm{Gal}(F(\zeta_N)/E)$ is solvable, hence so is $\mathrm{Gal}(F/E)$, since it is its quotient by the fundamental theorem of Galois theory 6.13. □

Corollary 7.3. *If $f \in \mathbb{Q}[x]$ is solvable by radicals (i.e. its roots are in an extension by radicals) and L is the splitting field of f over $\mathbb{Q}$, then $\mathrm{Gal}(L/\mathbb{Q})$ is solvable.*

Proposition 7.4. *If $f(x) = x^5 - 6x + 3$ and L is the splitting field of f over $\mathbb{Q}$, then $\mathrm{Gal}(L/\mathbb{Q}) \cong S_5$.*

Proof. Note that f is irreducible by Eisenstein's criterion with $p = 3$. If $\alpha \in L$ is a root then $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 5$, so $[L : \mathbb{Q}] = [L : \mathbb{Q}(\alpha)] \times 5$, a multiple of 5. Also, $\mathrm{Gal}(L/\mathbb{Q}) = G$ is a subgroup of S_5, and hence $|G|$ is a multiple of 5. By Cauchy's theorem, G contains an element of order 5, a 5-cycle. Next, note $G \ni$ complex conjugation. Now, $f(-1) = 8$, $f(1) = -2$, and thus f has at least 3 real roots. But $f'(x) = 5x^4 - 6$ has 2 zeroes, and therefore f has exactly 3 real roots. Thus $G \subseteq S_5$, and it contains a 5-cycles and complex conjugation. Without loss of generality, $(12) = x$, and replacing the 5-cycle by a power, we can assume that the 5-cycle is $(12345) = y$. Then

$$\begin{gathered} yxy^{-1} = (23) \\ y(23)y^{-1} = (34) \\ y(24)y^{-1} = (45) \\ (12)(23)(12) = (13) \\ (13)(34)(13) = (14) \\ (14)(45)(14) = (15) \end{gathered}$$

so G contains $(1i)$ for all i, and hence it contains $(ij) = (1i)(1j)(1i)$ for all i, j, all the transpositions. Therefore, $G = S_5$. □

Corollary 7.5. *There is no formula for the roots of a general quintic involving the coefficients and* $+ - \times \div \sqrt[n]{}$.